Free Energy in a Markovian Model
of a Lattice Spin System

RICHARD HOLLEY[*]

Department of Mathematics Princeton University, Princeton, New Jersey, USA

Abstract. A Markov process which may be thought of as a classical lattice spin system is considered. States of the system are probability measures on the configuration space, and we study the evolution of the free energy of these states with time. It is proved that for all initial states the free energy is nonincreasing and that it strictly decreases from any initial state which is shift invariant but not an equilibrium state. Finally we show that the state of the system converges weakly to the set of Gibbsian Distributions for the given interaction, and that all shift invariant equilibrium states are Gibbsian Distributions.

1. Introduction

In this paper we study a model of a classical lattice spin system which is also a Markov process. We let Z be the integers and consider the lattice Z^ν. At each point of the lattice we have a particle which is spinning either up of down. Thus the configuration of spins can be represented by functions, ξ, from Z^ν into $\{-1, 1\}$ with the interpretation that the spin at the site x is up (down) if $\xi(x) = 1(-1)$. For each subset R of Z^ν we have a number J_R, and throughout this paper it will be assumed that

$$J_R = J_{R+x} \quad \text{for all} \quad x \in Z^\nu, \tag{1.1a}$$

and

$$\text{there is a positive integer } L \text{ such that if } 0 \in R \text{ and } R \text{ is not contained in } [-L, L]^\nu \text{ then } J_R = 0. \tag{1.1b}$$

Expressions like $[-L, L]^\nu$ will always mean $[-L, L]^\nu$ restricted to Z^ν.

Let $\sigma_R(\xi) = \prod_{x \in R} \xi(x)$ and let $\beta > 0$. Then we define

$$c(x, \xi) = \exp\{\beta \sum_{R \ni x} J_R \sigma_R(\xi)\}.$$

The Markov process, $\xi(t)$, can then be described intuitively as follows: if at time t the configuration is $\xi(t)$, then the particle at x reverses

---

its spin some time during $[t, t + \Delta t]$ with probability

$$c(x, \xi(t)) \Delta t + 0((\Delta t)^2).$$

All of the particles on the lattice are doing this independently of each other; except of course the frequency at which a particle flips over, $c(x, \xi(t))$, is affected by the spins of the other particles.

It is clear that what we are describing is the infinitesimal generator of the process. Thus if $E = \{-1, 1\}^{Z^\nu}$ is given the product topology and f is a continuous real valued function on E which depends on only finitely many coordinates, then the infinitesimal generator, $\mathfrak{A}$, of the Markov process is given by

$$\mathfrak{A}f(\xi) = \sum_{x \in Z^\nu} c(x, \xi) \left[f(\xi_x) - f(\xi) \right], \tag{1.2}$$

where ξ_x is the function which is equal to ξ except at x where it is equal to $-\xi$.

The equation in (1.2), or more precisely its consequencs Lemma (2.3) below, is closely related to what is often called the Master equation.

This model seems to have been first proposed by R. Glauber [3]. However, for the case of an infinite lattice, it is not at all obvious that there exists a Markov process whose infinitesimal generator is given by (1.2) or that (1.2) is enough to determine the process uniquely. These problems have been solved recently by R. L. Dobrushin, I. I. Pyatetski-Shapiro, and N. B. Vasilev in [2] and by T. Liggett in [8].

If A is a Borel subset of E let $P^\xi(\xi(t) \in A)$ denote the probability that the configuration is an element of A at time t given that at time zero it was equal to ξ. Then if μ_0 is a probability measure on the Borel sets of E we define

$$\mu_t(A) = \int_E P^\xi(\xi(t) \in A) \, \mu_0(d\xi). \tag{1.3}$$

If μ_0 is the state of the system at time zero, then μ_t is the state at time t.

We call states μ_0 which have the property that $\mu_0 = \mu_t$ for all $t \geq 0$ equilibrium states.

As in [2] or [5] it can be shown that the Gibbsian Distributions in the sense of Dobrushin [1] are equilibrium states.

In section 2 we define the free energy per site of a state μ (denoted by $A(\mu)$) in the usual way and prove that $A(\mu_t)$ is a nonincreasing function of t.

In section 3 we strengthen this result and prove that if μ_0 is shift invariant but not a Gibbsian Distribution, then $A(\mu_t)$ is strictly less then $A(\mu_0)$ for all $t > 0$. We also prove that if μ_0 is shift invariant, then μ_t converges weakly to the set of Gibbsian Destributions. In the case

where the Gibbsian Distribution is unique this gives us an ergodic theorem. These results show that *the only shift invariant equilibrium states are Gibbsian Distributions* and *that a variational equilibrium state in the sense of D. Ruelle* [9] *page* 187 *is a Gibbsian Distribution.* This last result is already known (see O. Lanford and D. Ruelle [7]).

The techniques used in this paper are the same as those in [6]; however, the results here are more complete.

2. Free Energy

If Λ is a finite subset of Z^ν we will let $S(\Lambda)$ be the set of functions from Λ into $\{-1, 1\}$. For $\eta \in S(\Lambda)$ we let $B(\Lambda, \eta)$ be the Borel set in E defined by

$$B(\Lambda, \eta) = \{\xi : \xi(x) = \eta(x) \quad \text{for all} \quad x \in \Lambda\} \,.$$

If $\Lambda = [-N, N]^\nu$ we write $S(N)$ instead of $S([-N, N]^\nu)$ and $B(N, \eta)$ instead of $B([-N, N]^\nu, \eta)$. If $\Lambda = [-N - L, N + L]^\nu \setminus [-N, N]^\nu$ we write $S(N, L)$ for $S(\Lambda)$.

The free energy per site of an arbitrary state μ at temperature $1/\beta$ is then defined to be

$$A(\mu) = \lim_{N \to \infty} \sup (2N + 1)^{-\nu} \Bigg[\sum_{\eta \in S(N)} \mu(B(N, \eta)) \, U(N, \eta)$$
$$+ \beta^{-1} \sum_{\eta \in S(N)} \mu(B(N, \eta)) \, ln \, \mu(B(N, \eta)) \Bigg] . \tag{2.1}$$

Here $U(N, \eta) = \Sigma \, J_R \sigma_R(\eta)$ and the summation is taken over all $R \subset [-N, N]^\nu$.

Remark. If we set $P(N, \eta) = \exp\{-\beta U(N, \eta)\}$ then we may rewrite (2.1) to get

$$A(\mu) = \lim_{N \to \infty} \sup \beta^{-1}(2N + 1)^{-\nu} \sum_{\eta \in S(N)} \mu(B(N, \eta)) \, ln \, \frac{\mu(B(N, \eta))}{P(N, \eta)} \,.$$

(2.2) Theorem. *Let μ_0 be an arbitrary state and μ_t be as in* (1.3). *Then $A(\mu_t)$ is a nonincreasing function of t.*

If we were dealing with a system of finitely many spins, instead of infinitely many, this would be a standard result (see for example M. Kac [4] page 98). Lemma (2.8) below is essentially Theorem (2.2) for the case of finitely many spins.

The proof of Theorem (2.2) is accomplished by a series of lemmas.

If $\eta \in S(N)$ and $\psi \in S(N, L)$ we let $\psi\eta$ be the element of $S(N+L)$ which is equal to η on $[-N, N]^\nu$ and is equal to ψ on

$$[-N-L, N+L]^\nu \backslash [-N, N]^\nu .$$

(2.3) Lemma.

$$\frac{d}{dt} \mu_t(B(N, \eta)) = \sum_\psi \sum_x c(x, \psi\eta_x) \mu_t(B(N+L, \psi\eta_x))$$
$$- \sum_\psi \sum_x c(x, \psi\eta) \mu_t(B(N+L, \psi\eta)) .$$

Here the summation over ψ extends over all elements of $S(N, L)$ and the summation over x is over all elements of $[-N, N]^\nu$.

Because of (1.1 b), if $x \in [-N, N]^\nu$, then $c(x, \psi\eta_x)$ and $c(x, \psi\eta)$ can be defined as in the introduction.

Lemma (2.3) follows immediately from (1.1 b) and 1.2). The proof is left to the reader.

(2.4) Lemma.

$$\frac{d}{dt} \sum_\eta \mu_t(B(N, \eta)) \, ln \, \frac{\mu_t(B(N, \eta))}{P(N, \eta)}$$

$$= \sum_\eta \left\{ \frac{d}{dt} \mu_t(B(N, \eta)) \right\} ln \, \frac{\mu_t(B(N, \eta))}{P(N, \eta)}$$

$$= \sum_\psi \sum_\eta \sum_x c(x, \psi\eta_x) \mu_t(B(N+L, \psi\eta_x)) \, ln \, \frac{\mu_t(B(N, \eta))}{P(N, \eta)}$$

$$- \sum_\psi \sum_\eta \sum_x c(x, \psi\eta) \mu_t(B(N+L, \psi\eta)) \, ln \, \frac{\mu_t(B(N, \eta))}{P(N, \eta)} .$$

Proof. The first equality follows by simply exchanging the order of summation and differentiation and then using the product rule. The only difficulty occurs if $\mu_t(B(N, \eta)) = 0$ for some η. In that case if $\frac{d}{dt} \mu_t(B(N, \eta)) > 0$, then both sides are minus infinity; while if $\frac{d}{dt} \mu_t(B(N, \eta)) = 0$, then by first using (1.2) to show that $\mu_t(B(N, \eta))$ has two continuous derivatives we see that

$$\frac{d}{dt} \{ \mu_t(B(N, \eta)) \, ln \, \mu_t(B(N, \eta)) \} = 0 .$$

Thus if we interpret $0 \, ln \, 0$ to be 0, which we will do throughout, the equality still holds.

The second equality follows by using the results of Lemma (2.3). From now on we will drop the subscript t from the notation.

(2.5) Lemma.

$$\sum_{\psi} \sum_{\eta} \sum_{x}{}' c(x, \psi \eta_x^{\cdot}) \, \mu(B(N+L, \psi \eta_x)) \ln \frac{\mu(B(N, \eta))}{P(N, \eta)}$$

$$- \sum_{\psi} \sum_{\eta} \sum_{x}{}' c(x, \psi \eta) \, \mu(B(N+L, \psi \eta)) \ln \frac{\mu(B(N, \eta))}{P(N, \eta)}$$

$$\leq 2 e^K (e^{-1} + 2K)(2N+1)^{\nu-1} L \nu,$$

where Σ' is the sum over all $x \in [-N, N]^{\nu} \backslash [-N+L, N-L]^{\nu}$, and $K = \beta \sum_{R \ni 0} |J_R|$.

Proof. Noting that $(\eta_x)_x = \eta$ and that the map $\eta \rightarrow \eta_x$ is a bijection of $S(N)$ with itself we see that the expression on the left hand side of (2.5) is equal to

$$\sum_{\psi} \sum_{\eta} \sum_{x}{}' c(x, \psi \eta) \, \mu(B(N+L, \psi \eta)) \left[\ln \frac{\mu(B(N, \eta_x))}{P(N, \eta_x)} - \ln \frac{\mu(B(N, \eta))}{P(N, \eta)} \right]. \quad (2.6)$$

Note that $\mu(B(N+L, \psi \eta)) \leq \mu(B(N, \eta))$. Therefore if a factor of the form $\ln x - \ln 0$ appears in (2.6), it is multiplied by zero, and that whole term is equal to zero. If $\mu(B(N, \eta_x)) = 0$ *and* $\mu(B(N, \eta)) > 0$, *then* $\mu(B(N+L, \psi \eta)) > 0$ for some ψ and (2.6) is equal to minus infinity. Thus, since we are trying to bound (2.6) above, we may assume that if $\mu(B(N, \eta)) > 0$, then $\mu(B(N, \eta_x))$ is also greater than zero for all x.

Let $\ln^+ x = \max(\ln x, 0)$. Then (2.6) can be bounded above by

$$e^K \sum_{\psi} \sum_{\eta} \sum_{x}{}' \mu(B(N+L, \psi \eta)) \ln^+ \frac{\mu(B(N, \eta_x))}{\mu(B(N, \eta))}$$

$$+ 2K e^K \sum_{\psi} \sum_{\eta} \sum_{x}{}' \mu(B(N+L, \psi \eta))$$

$$= e^K \sum_{\eta} \sum_{x}{}' \frac{\mu(B(N, \eta))}{\mu(B(N, \eta_x))} \ln^+ \frac{\mu(B(N, \eta_x))}{\mu(B(N, \eta))} \mu(B(N, \eta_x))$$

$$+ 2K e^K \sum_{\eta} \sum_{x}{}' \mu(B(N, \eta)).$$

Since $x \ln^+(x^{-1}) \leq e^{-1}$, our last expression is less than or equal to

$$e^K e^{-1} \sum_{x}{}' \sum_{\eta} \mu(B(N, \eta_x)) + 2K e^K \sum_{x}{}' \sum_{\eta} \mu(B(N, \eta)).$$

The number of terms in the summation on x is less than $2(2N+1)^{\nu-1} L\nu$, and for fixed x the summation over η is equal to one. Thus substituting $2(2N+1)^{\nu-1} L\nu$ for each of the double summations completes the proof.

The difference between the expression appearing in the statement of Lemma (2.4) and the left hand side of the inequality in Lemma (2.5) is

$$\sum_{\psi}\sum_{\eta}\sum_x{}'' c(x,\psi\eta_x)\,\mu(B(N+L,\psi\eta_x))\,ln\frac{\mu(B(N,\eta))}{P(N,\eta)}$$
$$-\sum_{\psi}\sum_{\eta}\sum_x{}'' c(x,\psi\eta)\,\mu(B(N+L,\psi\eta))\,ln\frac{\mu(B(N,\eta))}{P(N,\eta)}, \tag{2.7}$$

where $\sum''$ is the summation over x in $[-N+L, N-L]^\nu$.

From (1.1b) we see that if $x \in [-N+L, N-L]^\nu$, then $c(x,\psi\eta)$ does not depend on ψ. Thus we may perform the summation on ψ and obtain

$$\sum_\eta\left[\sum_x{}'' c(x,\eta_x)\,\mu(B(N,\eta_x)) - \sum_x{}'' c(x,\eta)\,\mu(B(N,\eta))\right] ln\frac{\mu(B(N,\eta))}{P(N,\eta)}.$$

Let

$$F_0(x) = \begin{cases} x - x\,ln\,x - 1 & \text{if} \quad x > 0 \\ -1 & \text{if} \quad x = 0 \end{cases}$$

and

$$F(\mu, N, \eta, \xi) = \begin{cases} F_0\left(\dfrac{P(N,\eta)}{\mu(B(N,\eta))}\dfrac{\mu(B(N,\xi))}{P(N,\xi)}\right)\mu(B(N,\eta)) \\ \quad\text{if}\quad \mu(B(N,\eta)) > 0 \\ -\infty \quad \text{if}\quad \mu(B(N,\eta)) = 0 \quad \text{and} \quad \mu(B(N,\xi)) > 0 \\ 0 \quad \text{if}\quad \mu(B(N,\eta)) = \mu(B(N,\xi)) = 0. \end{cases}$$

(2.8) Lemma.

$$\sum_\eta\left[\sum_x{}'' c(x,\eta_x)\,\mu(B(N,\eta_x))\right.$$
$$\left. - \sum_x{}'' c(x,\eta)\,\mu(B(N,\eta))\right] ln\frac{\mu(B(N,\eta))}{P(N,\eta)}$$
$$= \sum_\eta\sum_x{}'' F(\mu, N, \eta, \eta_x)\,c(x,\eta_x)\,P(N,\eta_x)/P(N,\eta).$$

Proof. We write $\eta \in S(N)$ as $\psi\xi$, where $\psi \in S(N-L,L)$ and $\xi \in S(N-L)$. For fixed ψ consider the kernel $\mathfrak{A}_\psi(\cdot,\cdot)$ on $S(N-L) \times S(N-L)$ given by

$$\mathfrak{A}_\psi(\delta,\gamma) = \begin{cases} c(x,\psi\delta) & \text{if} \quad \gamma = \delta_x \\ -\sum_x{}'' c(x,\psi\delta) & \text{if} \quad \gamma = \delta \\ 0 \quad \text{otherwise}. \end{cases}$$

Then the left hand side in the statement of the lemma can be written

$$\sum_{\psi} \left[\sum_{\delta} \sum_{\gamma} \mathfrak{A}_{\psi}(\delta, \gamma)\, \mu(B(N, \psi\delta))\, ln\, \frac{\mu(B(N, \psi\gamma))}{P(N, \psi\gamma)} \right]. \qquad (2.9)$$

$\mathfrak{A}_{\psi}$ has the following properties:

$$\sum_{\gamma} \mathfrak{A}_{\psi}(\delta, \gamma) = 0 \quad \text{for all } \delta, \qquad (2.10a)$$

$$\sum_{\delta} P(N, \psi\delta)\, \mathfrak{A}_{\psi}(\delta, \gamma) = 0 \quad \text{for all } \gamma. \qquad (2.10b)$$

The first of these properties is obvious from the definition of $\mathfrak{A}_{\psi}$ and the second follows from an easy computation using the definitions of $P(\cdot, \cdot)$ and $c(\cdot, \cdot)$.

Using (2.10a) and (2.10b) it is easily seen that (2.9) is equal to

$$\sum_{\psi} \left[\sum_{\delta} \sum_{\gamma} F(\mu, N, \psi\gamma, \psi\delta)\, \mathfrak{A}_{\psi}(\delta, \gamma)\, P(N, \psi\delta)/P(N, \psi\gamma) \right]$$
$$= \sum_{\psi} \sum_{\gamma} \sum_{\delta \neq \gamma} F(\mu, N, \psi\gamma, \psi\delta)\, \mathfrak{A}_{\psi}(\delta, \gamma)\, P(N, \psi\delta)/P(N, \psi\gamma).$$

We may delete the terms where $\delta = \gamma$ because

$$F(\mu, N, \eta, \eta) = 0 \quad \text{for all } \eta.$$

The proof is completed by substituting the definition of $\mathfrak{A}_{\psi}$ into our last expression.

Returning to the proof of Theorem (2.2) we remark that it will be sufficient to show that for some constant C independent of N

$$\sum_{\eta \in S(N)} \mu_t(B(N, \eta))\, ln\, \frac{\mu_t(B(N, \eta))}{P(N, \eta)} \qquad (2.11)$$

$$\leqq \sum_{\eta \in S(N)} \mu_0(B(N, \eta))\, ln\, \frac{\mu_0(B(N, \eta))}{P(N, \eta)} + CN^{\nu-1}t.$$

Lemma (2.4) gives us an expression for the derivative of the left hand side of (2.11) which we split into the two terms in Lemmas (2.5) and (2.8). Lemma (2.5) gives us a bound on the first term of the form $CN^{\nu-1}$, and Lemma (2.8) shows that the second term is nonpositive. This last statement follows since the function F is nonpositive. Thus the derivative is bounded above by $CN^{\nu-1}$ and (2.11) follows.

3. Shift Invariant States

We begin this section with a description of the Gibbsian Distributions on E. The reader is referred to [1] where this notion is first defined and studied. Our notation is the same as in Section 2.

For $\psi \in S(N, L)$ let $\lambda_{\psi, N}$ be the measure on $S(N)$ given by

$$
\lambda_{\psi, N}(\xi) = \begin{cases} Z^{-1}(\psi, N) \exp\{-\beta P(N, \xi)\} & \text{if} \quad \xi = \psi \eta \\ \qquad \text{for some} \quad \eta \in S(N - L) \\ 0 \quad \text{otherwise}, \end{cases}
$$

where $Z^{-1}(\psi, N)$ is the normalizing constant which makes $\lambda_{\psi, N}$ a probability measure. Let

$$
\overline{S}(N) = \{\xi \in E : \xi(x) = 1 \text{ if } x \notin [-N, N]^\nu\}.
$$

There is an obvious identification of $\overline{S}(N)$ with $S(N)$ and by means of this identification we may think of $\lambda_{\psi, N}$ as a measure on E. Let λ_N be any convex combination of the $\lambda_{\psi, N}$, where the averaging is done over ψ. If we do this for each N we get a sequence $\{\lambda_N\}$ which, since E is compact, has a weakly convergent subsequence, $\lambda_{N_k} \to \lambda$. We denote the set of all λ obtained in this manner by $\mathscr{A}$. $\mathscr{A}$ is the set of Gibbsian Distributions.

A state μ is called shift invariant if for all finite Λ, all $\eta \in S(\Lambda)$, and all $a \in Z^\nu$

$$
\mu(B(\Lambda, \eta)) = \mu(B(\Lambda + a, \eta^a)),
$$

where η^a is the element in $S(\Lambda + a)$ given by

$$
\eta^a(x) = \eta(x - a). \tag{3.1}
$$

Let $\mathscr{M}$ be the set of all shift invariant states with the weak topology. We need the following facts about the elements of $\mathscr{M}$.

If $\mu_0 \in \mathscr{M}$, then $\mu_t \in \mathscr{M}$ for all $t > 0$. $\qquad$ (3.2)

The map $(\mu_0, t) \to \mu_t$ is continuous in the product topology on $\mathscr{M} \times [0, \infty)$. $\qquad$ (3.3)

In the definition of $A(\mu)$ the limit supremum is actually a limit. That is, if $\mu \in \mathscr{M}$, then

$$
\text{(3.4)}
$$

$$
A(\mu) = \lim_{N \to \infty} \beta^{-1}(2N + 1)^{-\nu} \sum_{\eta \in S(N)} \mu(B(N, \eta)) \ln \frac{\mu(B(N, \eta))}{P(N, \eta)}.
$$

An easy proof of (3.2) and (3.3) can be given by making use of the proofs of existence of the Markov processes. These proofs are however quite long and irrelevant for our present purposes; therefore, we refer the reader to [2] or [8].

A proof of (3.4) can be found in [9] page 180.

If m is large enough so that $2^m - 1 \geq L$, let

$$
H_m(\mu) = \sum_\eta \sum_x F(\mu, 2^m - 1, \eta, \eta_x) c(x, \eta_x) P(2^m - 1, \eta_x)/P(2^m - 1, \eta), \tag{3.5}
$$

where the summation on η is over all $\eta \in S(2^m - 1)$ and the summation on x is over all $x \in [-2^m + 1 + L, 2^m - 1 - L]^\nu$.

Comparing the definition of H_m with the tight hand side of the expression in Lemma (2.8), we see that H_m represents the rate that the free energy in the cube $[-2^m + 1, 2^m - 1]^\nu$ is decreasing due to the flips of the particles inside the cube $[-2^m + 1 + L, 2^m - 1 - L]^\nu$. Observe that if μ is shift invariant, then the two cubes used in the definition of $H_m(\mu)$ could be centered at any point of Z^ν and yeild the same value for $H_m(\mu)$. We will make use of this observation in the proof of the next lemma. One other important property of H_m is that it is upper semi-continuous on $\mathcal{M}$. This follows easily from the definition of F given in section 2.

The next lemma says that for shift invariant states the rate of decrease of the free energy in a cube grows at least as fast as the volume of the cube.

(3.6) Lemma. *If* $\mu \in \mathcal{M}$, *then* $H_m(\mu) \leqq 2^\nu H_{m-1}(\mu)$.

Proof. Since F is nonpositive, $H_m(\mu)$ is less than or equal to the expression obtained if the summation on x is only taken over the 2^ν disjoint cubes of side $2^m - 2L - 2$ which are in the corners of the cube $[-2^m + 1 + L, 2^m - 1 - L]^\nu$. Consider the cube

$$D = [L + 1, 2^m - 1 - L]^\nu .$$

Writing each $\eta \in S(2^m - 1)$ as $\psi\xi$, where $\xi \in S([1, 2^m - 1]^\nu)$ and $\psi \in S([-2^m + 1, 2^m - 1]^\nu \backslash [1, 2^m - 1]^\nu)$, we have

$$\sum_\eta \sum_{x \in D} F(\mu, 2^m - 1, \eta, \eta_x) c(x, \eta_x) P(2^m - 1, \eta_x)/P(2^m - 1, \eta) \tag{3.7}$$

$$= \sum_\xi \sum_{x \in D} \sum_\psi F(\mu, 2^m - 1, \psi\xi, \psi\xi_x) c(x, \psi\xi_x) P(2^m - 1, \psi\xi_x)/P(2^m - 1, \psi\xi) .$$

The idea is to show that for fixed ξ and x

$$\sum_\psi F(\mu, 2^m - 1, \psi\xi, \psi\xi_x) c(x, \psi\xi_x) P(2^m - 1, \psi\xi_x)/P(2^m - 1, \psi\xi)$$
$$\leqq F(\mu, 2^{m-1} - 1, \xi^a, \xi^a_{x+a}) c(x + a, \xi^a_{x+a}) \tag{3.8}$$
$$\times P(2^{m-1} - 1, \xi^a_{x+a})/P(2^{m-1} - 1, \xi^a) ,$$

where $a = (-2^{m-1}, \ldots, -2^{m-1}) \in Z^\nu$.

If (3.8) is true, then since $\xi \to \xi^a$ maps $S([1, 2^m - 1]^\nu)$ one to one onto $S(2^{m-1} - 1)$ and $x \to x + a$ maps D one to one onto $[-2^{m-1} + 1 + L, 2^{m-1} - 1 - L]^\nu$, we see by substituting (3.8) into (3.7) that (3.7) is bounded above by $H_{m-1}(\mu)$. The same argument holds for each of the 2^ν corners of the cube. Thus the lemma will be proved as soon as we establish (3.8).

From (1.1a) and (1.1b) we see that for $x \in D$

$$c(x, \psi\xi_x) = c(x, \xi_x) = c(x + a, \xi^a_{x+a})$$

and also that

$$P(2^m - 1, \psi\xi_x)/P(2^m - 1, \psi\xi) = P(2^{m-1} - 1, \xi^a_{x+a})/P(2^{m-1} - 1, \xi^a). \quad (3.9)$$

Thus we only need to show that

$$\sum_\psi F(\mu, 2^m - 1, \psi\xi, \psi\xi_x) \leq F(\mu, 2^{m-1} - 1, \xi^a, \xi^a_{x+a}). \quad (3.10)$$

If $\mu(B(2^m - 1, \psi\xi)) = 0$ and $\mu(B(2^m - 1, \psi\xi_x)) > 0$ for some ψ, then (3.10) follows immediately from the definition of F. If $\mu(B(2^m - 1, \psi\xi)) = \mu(B(2^m - 1, \psi\xi_x)) = 0$, then $F(\mu, 2^m - 1, \psi\xi, \psi\xi_x) = 0$, and we may delete such ψ from the summation. Thus we may assume that the summation is only over those ψ for which $\mu(B(2^m - 1, \psi\xi)) > 0$. (3.10) then follows by substituting the definition of F and the using (3.9), Jensen's inequality (F_0 is concave), and the shift invariance of μ.

(3.11) **Lemma.** $H(\mu) = \beta^{-1} \lim_{m \to \infty} (2^{m+1} - 1)^{-\nu} H_m(\mu)$ *exists for all* μ *in* $\mathcal{M}$ *(it is possibly minus infinity) and is upper semi-continuous. Moreover, if* $\mu_0 \in \mathcal{M}$, *then*

$$A(\mu_t) - A(\mu_0) \leq \int_0^t H(\mu_s) \, ds \, .$$

Proof. Let $G(m) = \prod_{j=m}^{\infty} (2^{j+2} - 2)^\nu/(2^{j+2} - 1)^\nu$.

Then by Lemma (3.6)

$$G(m+1)(2^{m+2} - 1)^{-\nu} H_{m+1}(\mu) \leq G(m)(2^{m+1} - 1)^{-\nu} H_m(\mu) \, .$$

Therefore, $G(m)(2^{m+1} - 1)^{-\nu} H_m(\mu)$ is a decreasing sequence of upper semi-continuous functions on $\mathcal{M}$. Hence the limit exists and is upper semi-continuous. Since $G(m)$ approaches one as m approaches infinity, this limit when divided by β is equal to $H(\mu)$.

From the proof of (2.11) we know that

$$\sum_{\eta \in S(N)} \mu_t(B(N, \eta)) \ln \frac{\mu_t(B(N, \eta))}{P(N, \eta)} - CN^{\nu-1}t$$

is a nonincreasing function of t. Therefore an application of Lebesgue's Theorem and Fatou's Lemma yields

$$\sum_\eta \mu_t(B(N, \eta)) \ln \frac{\mu_t(B(N, \eta))}{P(N, \eta)} - CN^{\nu-1}t$$

$$- \sum_\eta \mu_0(B(N, \eta)) \ln \frac{\mu_0(B(N, \eta))}{P(N, \eta)}$$

$$\leq \int_0^t \left[\frac{d}{ds} \left(\sum_\eta \mu_s(B(N, \eta)) \ln \frac{\mu_s(B(N, \eta))}{P(N, \eta)} \right) - CN^{\nu-1} \right] ds \, .$$

Using (3.2) and (3.4) together with this inequality we get

$A(\mu_t) - A(\mu_0)$

$$\leqq \beta^{-1} \limsup_{m \to \infty} (2^{m+1} - 1)^{-\nu} \int_0^t \frac{d}{ds} \left[\sum_\eta \mu_s(B(2^m - 1, \eta)) \ln \frac{\mu_s(B(2^m - 1, \eta))}{P(2^m - 1, \eta)} \right] ds$$

$$\leqq \limsup_{m \to \infty} \beta^{-1} (2^{m+1} - 1)^{-\nu} \int_0^t [H_m(\mu_s) + C(2^m - 1)^{\nu-1}] \, ds.$$

This last inequality follows exactly as in the proof of Theorem (2.2).

Since $0 < G(m) < 1$ and $H_m(\mu) \leqq 0$, it follows that $H_m(\mu) \leqq G(m) H_m(\mu)$. Therefore by monotone convergence we have

$A(\mu_t) - A(\mu_0)$

$$\leqq \limsup_{m \to \infty} \int_0^t \beta^{-1} (2^{m+1} - 1)^{-\nu} G(m) H_m(\mu_s) \, ds = \int_0^t H(\mu_s) \, ds,$$

and the proof is complete.

(3.12) Lemma. *Let $\mu \in \mathcal{M}$. If $\mu \notin \mathcal{A}$ then $H(\mu) < 0$.*

Proof. For each N the measure $\mu(B(N, \cdot))$ is defined on $S(N)$, and just as in the description of $\mathcal{A}$ we may think of it as a measure on E. It is clear that as N goes to infinity $\mu(B(N, \cdot))$ converges weakly to μ. Thus to complete the proof of the lemma it will suffice to show that if $H(\mu) = 0$, then $\mu(B(2^m - 1, \cdot))$ is equal to one of the λ_{2^m-1} used in the description of $\mathcal{A}$. Since

$$H(\mu) \leqq \beta^{-1} G(m) (2^{m+1} - 1)^{-\nu} H_m(\mu) \leqq 0,$$

if $H(\mu)$ is zero then $H_m(\mu)$ is zero for all m. But from (3.5) we see that if $H_m(\mu) = 0$, then for all $\eta \in S(2^m - 1)$ and all $x \in [-2^m + 1 + L, 2^m - 1 - L]^\nu$ we have

$$P(2^m - 1, \eta)/P(2^m - 1, \eta_x) = \mu(B(2^m - 1, \eta))/\mu(B(2^m - 1, \eta_x))$$

or else

$$\mu(B(2^m - 1, \eta)) = \mu(B(2^m - 1, \eta_x)) = 0. \tag{3.13}$$

Thus if $\psi \in S(2^m - 1 - L, L)$ and ξ' and ξ'' are in $S(2^m - 1 - L)$, then

$$P(2^m - 1, \psi\xi')/P(2^m - 1, \psi\xi'') = \mu(B(2^m - 1, \psi\xi'))/\mu(B(2^m - 1, \psi\xi'')),$$

or else

$$\mu(B(2^m - 1, \psi\xi')) = \mu(B(2^m - 1, \psi\xi'')) = 0. \tag{3.14}$$

Indeed, there is a sequence $x_1, x_2, \ldots, x_k$ such that $\xi'' = (\cdots((\xi'_{x_1})_{x_2} \cdots)_{x_k}$, and (3.14) follows by applying (3.13) k times. (3.14) is the same as the

statement that $\mu(B(2^m - 1, \cdot))$ is equal to one of the λ_{2^m-1} used in the description of $\mathscr{A}$.

(3.15) Theorem. *Let $\mu \in \mathscr{M}$ and $\mu \notin \mathscr{A}$. Then there exists a weakly open set G_μ containing μ and $\varepsilon, \delta > 0$ such that if $v_0 \in G_\mu$ and $0 \leq s \leq \varepsilon$, then*

$$A(v_s) - A(v_0) \leq -\delta s.$$

Proof. By Lemma (3.11) $H(\mu) < 0$. Therefore, since H is upper semi-continuous, there is a $\delta > 0$ such that if

$$G_1 = \{v \in \mathscr{M} : H(v) < -\delta\},$$

then G_1 is open and contains μ. From (3.3) it follows that there is an open set G_2 in $\mathscr{M} \times [0, \infty)$ containing $(\mu, 0)$ and such that if $(v_0, s) \in G_2$, then $v_s \in G_1$. Since $(\mu, 0) \in G_2$, there is an $\varepsilon > 0$ and an open set $G_\mu \subset \mathscr{M}$ such that $G_\mu \times [0, \varepsilon) \subset G_2$. Thus if $v_0 \in G_\mu$ and $0 \leq s < \varepsilon$, then $H(v_s) < -\delta$. The proof is completed by an application of Lemma (3.11).

(3.16) Corollary. *Let $\mu_0 \in \mathscr{M}$ and suppose that $t_n \to \infty$ and that μ_{t_n} converges weakly to μ. Then $\mu \in \mathscr{A}$.*

Proof. This follows immediately from Theorems (3.15) and (2.2) and the easily verified fact that there are two real numbers m and M such that for all states μ we have $m \leq A(\mu) \leq M$.

(3.17) Corollary. *Let G be a weakly open subset of $\mathscr{M}$ which contains $\mathscr{A} \cap \mathscr{M}$, and let $\mu_0 \in \mathscr{M}$. Then for all t sufficiently large $\mu_t \in G$. In particular if the parameters are such that there is a unique Gibbsian Distribution, then the state at time t converges weakly to that Gibbsian Distribution as t goes to infinity.*

Proof. $\mathscr{M}$ is weakly compact, therefore Corollary (3.17) follows from Corollary (3.16).

(3.18) Corollary. *All shift invariant equilibrium states are Gibbsian Distributions.*

A shift invariant state μ is said to be an equilibrium state in the variational sense if it satisfies

$$A(\mu) = \inf_{v \in \mathscr{M}} A(v)$$

(see Ruelle [7] pages 187–190).

(3.19) Corollary. *An equilibrium state in the variational sense is a Gibbsian Distribution.*

Proof. This is an immediate consequence of Theorem (3.14).

References

1. Dobrushin, R. L.: Gibbsian random fields for lattice systems with pairwise interactions, Functional Analysis and its Applications, **2**, 292—302 (1968).
2. — Patetski-Shapiro, I. I., Vasilev, N. B.: Markov processes in an infinite product of discrete spaces. Soviet-Japanese Symposium in Probability Theory. Khavarovsk, 1969.
3. Glauber, R. J.: Time dependent statistics of the Ising model. Journal of Mathematical Physics, **4**, 294—307 (1963).
4. Kac, M.: Probability and related topics in physical sciences. New York: Interscience 1959.
5. Holley, R.: A class of interactions in an infinite particle system. Advances in Mathematics **5**, 291—309 (1970).
6. — Pressure and Helmholtz free energy in a dynamic model of a lattice gas. Proceedings of the Sixth Berkeley Symposium on Probability and Mathematical Statistics 1971.
7. Lanford, O. E., Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. Commun. math. Phys. **13**, 194—215 (1969).
8. Liggett, T. M.: Existence theorems for infinite particle systems, to appear.
9. Ruelle, D.: Statistical mechanics. Amsterdam: Benjamin 1969.

R. Holley
Dept. of Mathematics
Fine Hall
Princeton University
Princeton, New Jersey 08540, USA