

Decay of Correlations in Classical Lattice Models at High Temperature*

Leonard Gross

Department of Mathematics, Cornell University, Ithaca, NY 14853, USA

Abstract. In classical statistical mechanical lattice models with many body potentials of finite or infinite range and arbitrary spin it is shown that the truncated pair correlation function decays in the same weighted summability sense as the potential, at high temperature.

1. Introduction

We consider a lattice model on Z^m with single spin space X and interaction φ . Here φ denotes a real valued function $\cup_A X^A$, the union running over all nonempty finite subsets, A , of the lattice Z^m . Under suitable conditions on φ , and for sufficiently small β , the interaction $\beta\varphi$ has a unique Gibbs state σ . We shall study the decay of the truncated two point function for σ . More precisely, and somewhat more generally, we study the behavior of $\psi(a) \equiv \sigma(fg_a) - \sigma(f)\sigma(g_a)$ where f and g are bounded real functions on $X^{(Z^m)}$ and g_a is the translate of g by an element a in Z^m .

For what potentials φ does $\psi(a)$ decay in some specified manner as $a \rightarrow \infty$ (e.g., exponentially)? This question has been addressed in numerous works, in which, under a variety of different conditions on the form of φ , an assumed decay rate for φ at large distances is shown to imply some related kind of decay rate for $\psi(a)$ as $a \rightarrow \infty$. The restrictions on the form of φ usually involve a restriction on either the range of φ (e.g., finite range), the many bodiedness of φ (e.g., pair interactions), or on the cardinality of X (e.g., cardinality two), or some combination of these, and sometimes further form restrictions designed to allow use of transfer matrix methods or correlation inequalities. For a survey of the extensive literature up to 1974 we refer the reader to Duneau, Souillard and Iagolnitzer [4], and to Ruelle [16]. We also refer the reader to Ruelle [17] for background structure. We mention here only the recent papers of Israel [11], Sylvester [18], and Holley and Stroock [8], which obtain exponential decay for some interactions of what may be called locally finite range. For other recent work see also [20–22].

* Research partially supported by the National Science Foundation under Grant MCS 78-00680

Our results are roughly of the following form. Let X be an arbitrary measure space with a single spin probability measure. Assume φ is translation invariant and put

$$\zeta(a) = \sup_{s \in X^{Z^m}} \left| \sum_{A \supset \{0, a\}} \varphi(s|A) \right|.$$

Let $d(\cdot, \cdot)$ be a translation invariant metric on Z^m . If

$$\sum_{a \in Z^m} \zeta(a) e^{d(0, a)} < \infty \quad (1.1)$$

then for sufficiently small β and suitable f and g , we prove that ψ satisfies

$$\sum_{a \in Z^m} \psi(a) e^{d(0, a)} < \infty. \quad (1.2)$$

We require no form restrictions on φ but only size restriction and (1.1). The precise statement is in Theorem 3. Translation invariance of φ is unnecessary when suitably formulated. The key results are Theorems 1 and 2 which are formulated in terms of conditional probabilities of σ and allow application also to the case of arbitrary temperature and small activity (Corollary 4.4). In view of the identical appearance of (1.1) and (1.2) our main result can be described succinctly by saying that the truncated two point function decays at high temperature (or low activity) at the same rate as the potential, if by “rate of decay” we understand the weighted summability condition (1.1), which is determined by the metric d . If $d(a, b) = k|a - b|$ for some constant $k > 0$ then (1.1) is satisfied if φ is of finite range or if φ (more precisely, ζ) is exponentially decaying. In either case ψ also decays exponentially. If $d(a, b) = k \log(1 + |a - b|)$ then (1.1) allows polynomial decay and (1.2) forces at least polynomial decay. The technique of proof, which is based on modifications of Lanford’s and Vasershtein’s proofs [12, 19] of Dobrushin’s uniqueness theorem [1, Theorem 2], seems to be new. It seems likely that the technique can be extended to study higher order correlations.

2. Statement of Main Theorems

Let L be a countable set and X a compact metric space. We write $\Omega = X^L$ for the space of all functions s from L into X , and we denote by s_a the value of s at a point a in L . Ω is the configuration space of a lattice model with individual spin space X . We use the fact that X is a compact metric space just for ease of formulation of results. Actually any measurable space, X , will suffice for all our purposes except existence of Gibbs states. One need only replace the space of continuous functions on Ω by the space of bounded measurable functions on Ω which are uniformly approximable by cylinder functions. But the cases of principle current interest are $X =$ finite set, S^m , compact group, or a finite product of these. We expect that in field theoretic models, where $X =$ real line, the more quantitative hypotheses are unlikely to hold.

For each point a in L and each point s in Ω let $\mu_a(\cdot | s)$ be a probability measure on the Borel sets of X such that

- i. $\mu_a(\cdot | s)$ does not depend on s_a and
- ii. $s \rightarrow \int_X f(x) \mu_a(dx | s)$ is continuous on Ω for each continuous function f on X .

$C(\Omega)$ will denote the space of continuous real valued functions on Ω with supremum norm $|f|_\infty$. For each point a in L define $\tau_a : C(\Omega) \rightarrow C(\Omega)$ by

$$(\tau_a f)(s) = \int_X f(x \vee \hat{s}) \mu_a(dx|s) \quad (2.1)$$

where \hat{s} is the restriction of s to $L - \{a\}$ and $x \vee \hat{s}$ is the point of Ω taking the value x at a and which is s elsewhere on L .

Definition. A probability measure σ on the Borel sets of Ω will be said to have *conditional probabilities* $\{\mu_a\}_{a \in L}$ if

$$\sigma(\tau_a f) = \sigma(f) \quad (2.2)$$

for all a in L and all f in $C(\Omega)$. If the measures μ_a arise from a potential, then, as is well known [1, 13, 17], the Eqs. (2.2) are special cases of the D.L.R. equations for the Gibbs states, σ , of the potential. This is the case of main interest to us. It will be discussed in Sect. 4.

For any two distinct points a and b in L put

$$\varrho_{a,b} = \frac{1}{2} \sup \{ \|\mu_b(\cdot|s) - \mu_b(\cdot|t)\|_{\text{var}} : s = t \text{ off } \{a\} \} \quad (2.3)$$

Let $d(\cdot, \cdot)$ be a semimetric on L . That is, d is a real valued nonnegative symmetric function on $L \times L$ satisfying the triangle inequality. The example $d=0$ is not without some interest. Given the system $\{\mu_a\}$ and a semimetric d we put

$$\beta_{a,b} = \varrho_{a,b} e^{d(a,b)} \quad a \neq b \quad \text{and} \quad \beta_{a,a} = 0 \quad \text{for all } a. \quad (2.4)$$

Furthermore we write

$$\alpha = \sup_{b \in L} \sum_{a \in L} \beta_{a,b} \quad (2.5)$$

and

$$\gamma = \sup_{a \in L} \sum_{b \in L} \beta_{a,b}. \quad (2.6)$$

If f is in $C(\Omega)$ and a is in L put

$$\delta_a(f) = \sup \{ |f(s) - f(t)| : s = t \text{ off } \{a\} \} \quad (2.7)$$

and

$$\|f\|_a = \sum_{b \in L} e^{d(a,b)} \delta_b(f). \quad (2.8)$$

The first sentence of Theorem 1 is Dobrushin's uniqueness theorem [1] when $d=0$. See also [12, 19].

Theorem 1. *If $\alpha < 1$ then there is at most one probability measure σ on Ω with conditional probabilities $\{\mu_a\}_{a \in L}$. Moreover for any functions f and g in $C(\Omega)$*

$$|\sigma(fg) - \sigma(f)\sigma(g)| \leq e^{-d(a,b)} \|f\|_a \|g\|_b (1 - \alpha)^{-2} (1 - \alpha^2)^{-1} \quad (2.9)$$

Theorem 2. *Assume $L = Z^m$ and assume that the semimetric d is translation invariant. If $\alpha < 1$ and $\gamma(1 + \alpha) < 1$, and if σ is a (unique by Theorem 1) probability measure*

with conditional probabilities $\{\mu_a\}_{a \in L}$ then for any functions f and g in $C(\Omega)$ and any point c in L

$$\sum_{a \in L} |\sigma(fg_a) - \sigma(f)\sigma(g_a)| e^{d(a,c)} \leq \|f\|_c \|g\|_0 (1-\alpha)^{-1} \cdot (1-\gamma-\alpha\gamma)^{-1} \quad (2.10)$$

where g_a refers to the translate of g by a and the subscript on $\|g\|_0$ refers to the origin of Z^m .

Explicitly, $g_a(s) = g(s^{(a)})$ where $s_b^{(a)} = s_{b+a}$.

Example. Let $X = \{0, 1\}$, and let L be an arbitrary countable set. Take $f(s) = s_a$ and $g(s) = s_b$. Then $\|f\|_a = 1$ and $\|g\|_b = 1$. So Theorem 1 asserts that if $\alpha < 1$ then

$$|\sigma(s_a s_b) - \sigma(s_a)\sigma(s_b)| \leq e^{-d(a,b)} (1-\alpha)^{-2} (1-\alpha^2)^{-1}. \quad (2.11)$$

Thus the conclusion of Theorem 1 reduces to a decay estimate on the truncated two point function. Somewhat more generally, if X is arbitrary, and u and v are in $C(X)$ and $f(s) = u(s_a)$ and $g(s) = u(s_b)$ then the right side of (2.9) is finite, resulting in an inequality similar to (2.11).

If $L = Z^m$ and $d(a,b) = r|a-b|$ where $|a|$ is the Euclidean norm on Z^m and $r > 0$ then (2.11) asserts that the decay is exponential with correlation length r^{-1} . If, on the other hand, $d(a,b) = r \log(1+|a-b|)$ then this is also a metric and $e^{-d(a,b)} = (1+|a-b|)^{-r}$, so that (2.11) implies power law decay of the correlation function. We shall see in Sect. 4 that if a potential satisfies condition (1.1) then at high enough temperature we have $\alpha < 1$, and consequently the two point function decays as in (2.11). But note that the decay in (2.11) is weaker than the decay required for (1.1). Thus there is a loss of qualitative information in passing from (1.1) to (2.11). Theorem 2 remedies this at the price of going to a slightly higher temperature (which has the effect of reducing α and γ). If in the preceding example we take $f(s) = s_0$ and $g(s) = s_0$ and $c = 0$ then the conclusion of Theorem 2 reads

$$\sum_{a \in Z^m} |\sigma(s_0 s_a) - \sigma(s_0)\sigma(s_a)| e^{d(a,0)} \leq (1-\alpha)^{-1} (1-\gamma-\alpha\gamma)^{-1}.$$

Thus the truncated two point function decays in the same weighted summability sense as the potential at high temperature. See [4] for further discussion of background.

3. Proofs of the Main Theorems

Choose an enumeration a_1, a_2, \dots of the points of L . We shall write $\delta_n(f)$ instead of $\delta_{a_n}(f)$ and τ_n and μ_n instead of τ_{a_n} and μ_{a_n} . Similarly for $\varrho_{j,k}$ and $\beta_{j,k}$. The operators τ_n are bounded operators on $C(\Omega)$ of norm one as follows from (2.1). Put

$$T_{n,p} = \tau_n \tau_{n+1} \dots \tau_p \quad 1 \leq n \leq p. \quad (3.1)$$

If f is a function in $C(\Omega)$ which depends only on the first m coordinates s_{a_1}, \dots, s_{a_m} of its argument s then $\tau_k f = f$ if $k > m$. Consequently $\lim_{p \rightarrow \infty} T_{n,p} f$ exists. Since functions depending on only finitely many coordinates are dense in $C(\Omega)$ in supremum norm and since $\|T_{n,p}\| = 1$ for all n, p , it follows that the uniform limit of $T_{n,p} f$ exists for all f in $C(\Omega)$ as $p \rightarrow \infty$. We put

$$T_n f = \lim_{p \rightarrow \infty} T_{n,p} f \quad f \in C(\Omega) \quad (3.2)$$

and we write also $T = T_1$. Clearly $\|T_n\| = 1$ for all n .

Now if μ_1 and μ_2 are probability measures on some space on which f is a bounded real valued measurable function then, putting $c = (\sup f + \inf f)/2$ we have

$$\begin{aligned} |\mu_1(f) - \mu_2(f)| &= |(\mu_1 - \mu_2)(f - c)| \\ &\leq \|\mu_1 - \mu_2\|_{\text{var}} \sup |f - c| \\ &= \frac{1}{2} \|\mu_1 - \mu_2\|_{\text{var}} |\sup f - \inf f|. \end{aligned} \quad (3.3)$$

Thus if f is in $C(\Omega)$, and s, t are two points of Ω which coincide except at the point a in L then for any point $b \neq a$.

$$\begin{aligned} |(\tau_b f)(s) - (\tau_b f)(t)| &= \left| \int f(x \vee \hat{s}) \mu_b(dx|s) - \int f(x \vee \hat{t}) \mu_b(dx|t) \right| \\ &\leq \left| \int (f(x \vee \hat{s}) - f(x \vee \hat{t})) \mu_b(dx|s) \right| + \left| \int f(x \vee \hat{t}) (\mu_b(dx|s) - \mu_b(dx|t)) \right| \\ &\leq \delta_a(f) + \varrho_{a,b} \delta_b(f) \end{aligned} \quad (3.4)$$

by (3.3). Thus $\delta_a(\tau_b f) \leq \delta_a(f) + \varrho_{a,b} \delta_b(f)$ if $a \neq b$. On the other hand if $a = b$ then since $(\tau_b f)(s)$ is independent of s_b the left side of (3.4) is zero. Thus

$$\delta_j(\tau_k f) \leq \begin{cases} 0 & \text{if } j = k. \\ \delta_j(f) + \varrho_{j,k} \delta_k(f) & \text{if } j \neq k. \end{cases} \quad (3.5)$$

We have taken the derivation of the basic inequality (3.5) from Lanford [12]. The proofs of Theorems 1 and 2 are based on repeated application of a slightly extended form of (3.5). By the triangle inequality for d we have

$$\begin{aligned} e^{d(a,a_j)} \varrho_{j,k} \delta_k(f) &\leq e^{d(a_j,a_k)} \varrho_{j,k} e^{d(a,a_j)} \delta_k(f) \\ &= \beta_{j,k} e^{d(a,a_k)} \delta_k(f). \end{aligned}$$

Hence for any point a in L ,

$$e^{d(a,a_j)} \delta_j(\tau_k f) \leq \begin{cases} 0 & \text{if } j = k \\ e^{d(a,a_j)} \delta_j(f) + \beta_{j,k} [e^{d(a,a_k)} \delta_k(f)] & \text{if } j \neq k. \end{cases} \quad (3.6)$$

Following Vasershtein [19] we shall manipulate these inequalities in matrix form. Define matrices $A^{(n)}$ by

$$(A^{(n)})_{i,k} = \begin{cases} \delta_{ik} & \text{if } k \neq n \\ \beta_{ik} & \text{if } k = n \end{cases} \quad (3.7)$$

for $i, k, n = 1, 2, 3, \dots$ and put

$$B^{n,p} = A^{(n)} A^{(n+1)} \dots A^{(p)} \quad \text{for } n \leq p. \quad (3.8)$$

Since each row of $A^{(r)}$ has at most two non zero entries and since

$$(B^{n,p})_{ik} = \sum_{j_1, \dots, j_r} A_{ij_1}^{(n)} A_{j_1 j_2}^{(n+1)} \dots A_{j_r k}^{(p)}, \quad r = p - n$$

it is clear that each entry of $B^{n,p}$ is a finite sum. Moreover it follows from (3.7) that

$$(B^{n,p})_{i,k} = \begin{cases} \delta_{ik} & \text{if } k \notin [n, p] \\ (B^{n,k})_{i,k} & \text{if } k \in [n, p]. \end{cases} \quad (3.9)$$

Thus the limit $B^{(n)} = \lim_{p \rightarrow \infty} B^{n,p}$ exists entrywise. In fact

$$B_{i,k}^{(n)} = \begin{cases} \delta_{ik} & \text{if } k < n \\ (B^{n,k})_{i,k} & \text{if } k \geq n. \end{cases} \quad (3.10)$$

In particular we put $A = B^{(1)}$.

The first of the following four inequalities is due to Vasershtein [19].

Lemma 3.1. *If $\sup_k \sum_{i=1}^{\infty} \beta_{i,k} \equiv \alpha < 1$ then*

$$\sum_{i=1}^{\infty} A_{i,k} \leq \alpha \quad \text{for all } k \quad (3.11)$$

and

$$\sum_{n=1}^{\infty} B_{n,k}^{(n+1)} \leq (1-\alpha)^{-1} \quad \text{for all } k. \quad (3.12)$$

If instead, $\sup_i \sum_{k=1}^{\infty} \beta_{i,k} \equiv \gamma < 1$, then

$$\sum_{k=1}^{\infty} B_{i,k}^{(n)} \leq (1-\gamma)^{-1} \quad \text{for all } i \text{ and all } n, \quad (3.13)$$

and

$$\sum_{k=1}^{\infty} A_{i,k} \leq \gamma(1-\gamma)^{-1} \quad \text{for all } i. \quad (3.14)$$

Proof. From the definition (3.7) we see that $\sum_{i=1}^{\infty} A_{i,k}^{(n)} = 1$ if $k \neq n$ and is $\leq \alpha \leq 1$ if $k = n$. Hence

$$\begin{aligned} \sum_{i=1}^{\infty} A_{i,k} &= \sum_{i=1}^{\infty} (B^{1,k})_{i,k} \quad \text{by (3.9)} \\ &= \sum_{j_{k-1}} \sum_{j_{k-2}} \dots \left(\sum_{j_1} \left(\sum_i A_{i,j_1}^{(1)} \right) A_{j_1,j_2}^{(2)} \right) \dots A_{j_{k-1},k}^{(k)}. \end{aligned}$$

The innermost sum is at most one. Hence so is the sum on j_1 , and so on, until we reach the sum on j_{k-1} which is at most α . This proves (3.11).

To prove (3.12) we use induction on k . If $k=1$ then by (3.10)

$$\sum_{n=1}^{\infty} B_{n,k}^{(n+1)} = B_{1,1}^{(2)} + B_{2,1}^{(3)} + \dots = 1 + 0 + \dots \leq (1-\alpha)^{-1}.$$

Thus the induction hypothesis is satisfied for $k=1$. Suppose $k > 1$. Now if $n+1 \leq k-1$ then by (3.10) and (3.8)

$$\begin{aligned} B_{n,k}^{(n+1)} &= (B^{n+1,k})_{n,k} \\ &= \sum_{j=1}^{\infty} (B^{n+1,k-1})_{n,j} A_{j,k}^{(k)} \\ &= \sum_{j=1}^{k-1} (B^{n+1,k-1})_{n,j} \beta_{j,k} \end{aligned}$$

by (3.9) because $\delta_{nj}=0$ if $j>k-1$. Sum this equality from $n=1$ to $k-2$ and separate the $j=k-1$ term to get

$$\begin{aligned} \sum_{n=1}^{k-2} B_{n,k}^{(n+1)} &= \sum_{j=1}^{k-2} \sum_{n=1}^{k-2} (B^{n+1,k-1})_{n,j} \beta_{j,k} + \sum_{n=1}^{k-2} (B^{n+1,k-1})_{n,k-1} \beta_{k-1,k} \\ &= \sum_{j=1}^{k-2} \left[\sum_{n=1}^{k-2} B_{n,j}^{(n+1)} \right] \beta_{jk} + \left\{ \sum_{n=1}^{k-2} B_{n,k-1}^{(n+1)} \right\} \beta_{k-1,k}. \end{aligned}$$

By the induction hypothesis the term in square brackets is at most $(1-\alpha)^{-1}$ since each j that occurs is less than k . Moreover since $B_{n,k-1}^{(n+1)}=1$ by (3.10) when $n=k-1$, the term in braces $\{ \}$ is $\sum_{n=1}^{k-1} B_{n,k-1}^{(n+1)} - 1$, which by the induction hypothesis is at most $(1-\alpha)^{-1} - 1 = \alpha/(1-\alpha)$. Thus

$$\sum_{n=1}^{k-2} B_{n,k}^{(n+1)} \leq (1-\alpha)^{-1} \sum_{j=1}^{k-2} \beta_{jk} + \alpha(1-\alpha)^{-1} \beta_{k-1,k}.$$

But

$$\sum_{n=k-1}^{\infty} B_{n,k}^{(n+1)} = B_{k-1,k}^{(k)} + B_{k,k}^{(k+1)} + B_{k+1,k}^{(k+2)} + \dots = A_{k-1,k}^{(k)} + 1 + 0 + \dots$$

by (3.10) and (3.8). Hence

$$\begin{aligned} \sum_{n=1}^{\infty} B_{n,k}^{(n+1)} &\leq (1-\alpha)^{-1} \sum_{j=1}^{k-2} \beta_{jk} + \alpha(1-\alpha)^{-1} \beta_{k-1,k} + \beta_{k-1,k} + 1 \\ &= (1-\alpha)^{-1} \sum_{j=1}^{k-1} \beta_{j,k} + 1 \\ &\leq (1-\alpha)^{-1} \alpha + 1 \\ &= (1-\alpha)^{-1}. \end{aligned}$$

This concludes the proof of (3.12).

To prove (3.13) and (3.14) note that for $n+1 \leq k \leq p$ and any i

$$\begin{aligned} (B^{n,p})_{i,k} &= (B^{n,k})_{i,k} \quad \text{by (3.9)} \\ &= \sum_{j=1}^{\infty} (B^{n,k-1})_{i,j} \beta_{j,k} \\ &= \sum_{j=n}^{k-1} (B^{n,k-1})_{i,j} \beta_{j,k} + \sum_{j \notin \{n, k-1\}} \delta_{ij} \beta_{jk} \quad \text{by (3.9)} \\ &= \sum_{j=n}^{k-1} (B^{n,p})_{i,j} \beta_{j,k} + \sum_{j \notin \{n, k-1\}} \delta_{ij} \beta_{jk} \quad \text{by (3.9)}. \end{aligned}$$

For $k=n$ we have $(B^{n,p})_{i,k}=(B^{n,n})_{i,n}=\beta_{in}$. Hence

$$\begin{aligned} \sum_{k=n}^p (B^{n,p})_{i,k} &= \sum_{k=n+1}^p (B^{n,p})_{i,k} + \beta_{i,n} \\ &= \sum_{k=n+1}^p \sum_{j=n}^{k-1} (B^{n,p})_{i,j} \beta_{jk} + \sum_{k=n+1}^p \sum_{j \notin [n, k-1]} \delta_{ij} \beta_{jk} + \beta_{i,n} \\ &\leq \sum_{k=n+1}^p \sum_{j=n}^p (B^{n,p})_{i,j} \beta_{j,k} + \sum_{k=n+1}^p \beta_{i,k} + \beta_{i,n} \\ &\leq \sum_{j=n}^p (B^{n,p})_{i,j} \gamma + \gamma. \end{aligned}$$

Thus $(1-\gamma) \sum_{k=n}^p (B^{n,p})_{i,k} \leq \gamma$, and consequently

$$\sum_{k=n}^p (B^{n,p})_{i,k} \leq \gamma / (1-\gamma). \quad (3.15)$$

If $k \in [n, p]$ then $B_{i,k}^{(n)} = (B^{n,p})_{i,k}$ by (3.9) and (3.10) so that (3.15) may be written $\sum_{k=n}^p B_{i,k}^{(n)} \leq \gamma / (1-\gamma)$ and therefore

$$\sum_{k=n}^{\infty} B_{i,k}^{(n)} \leq \gamma / (1-\gamma). \quad (3.16)$$

If $n=1$ this reduces to (3.14). If $n>1$ then by (3.10) $\sum_{k=1}^{n-1} B_{i,k}^{(n)} = \sum_{k=1}^{n-1} \delta_{i,k}$ which is at most one. Combining this with (3.16) yields (3.13). This concludes the proof.

Corollary 3.2. *If $\alpha < 1$ then for all f in $C(\Omega)$*

$$\|Tf\|_a \leq \alpha \|f\|_a \quad \text{for all } a \in L. \quad (3.17)$$

Proof. Fix a in L . For any function f in $C(\Omega)$ let Df be the column vector whose j th component is the nonnegative number $D_j f \equiv e^{d(a, a_j)} \delta_j(f)$. Then the basic inequality (3.6) may be written

$$D(\tau_k f) \leq A^{(k)} Df \quad (3.18)$$

where the inequality refers to the components. Repeated application gives $D(\tau_1 \dots \tau_p f) \leq A^{(1)} \dots A^{(p)} D(f) = B^{1,p} Df$. If $p \geq i$ then by (3.9) and (3.10) $(B^{1,p})_{i,k} = 0$ if $k > p$ and equals $A_{i,k}$ if $k \leq p$. Thus if $p \geq i$ then

$$D_i(\tau_1 \dots \tau_p f) \leq \sum_{k=1}^{\infty} (B^{1,p})_{i,k} D_k(f) = \sum_{k=1}^p A_{ik} D_k(f).$$

But $\tau_1 \dots \tau_p f$ converges to Tf uniformly as $p \rightarrow \infty$. Hence $D_i(\tau_1 \dots \tau_p f)$ converges to $D_i(Tf)$. It follows that

$$D_i(Tf) \leq \sum_{k=1}^{\infty} A_{ik} D_k(f). \quad (3.19)$$

Taking the sum on i yields, by (3.11),

$$\|Tf\|_a = \sum_{i=1}^{\infty} D_i(Tf) \leq \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} A_{ik} D_k(f) \leq \alpha \sum_{k=1}^{\infty} D_k(f) = \alpha \|f\|_a.$$

Corollary 3.3. *If $\alpha < 1$ then there is at most one probability measure σ on Ω with conditional probabilities $\{\mu_a\}_{a \in L}$. Moreover for any function f in $C(\Omega)$ $T^n f$ converges uniformly and*

$$\lim_{n \rightarrow \infty} T^n f = \sigma(f). \quad (3.20)$$

Proof. The semimetric d plays no role in this corollary. If $\alpha < 1$ when computed for a semimetric d it is also less than one when computed with $d=0$. So we assume $d=0$ in this proof. In this case the semi-norms $\|f\|_a$ do not depend on a . We drop the subscript a and write $\|f\| = \sum_{j=1}^{\infty} \delta_j(f)$.

By Corollary 3.2 we have $\|T^n f\| \leq \alpha^n \|f\|$, $n = 1, 2, \dots$ for any function f in $C(\Omega)$. Now by changing one coordinate at a time one sees that $\sup f - \inf f \leq \|f\|$. Suppose that $\|f\| < \infty$. Put $u_n = \inf T^n f$ and $v_n = \sup T^n f$. Then $v_n - u_n \leq \|T^n f\| \leq \alpha^n \|f\|$ which goes to zero as $n \rightarrow \infty$. However since T is positivity preserving it is also order preserving. Since T takes constant functions to constant functions we may apply T^p to the inequality $u_n \leq T^n f \leq v_n$ to get $u_n \leq T^{n+p} f \leq v_n$. It follows that $u_n \leq u_{n+p} \leq v_{n+p} \leq v_n$. Hence v_n and u_n converge to a common value $\varrho(f)$ which is also the uniform limit of $T^n f$. Clearly ϱ is a positive linear functional on the set $C^1(\Omega)$ of functions f with $\|f\| < \infty$. $\varrho(1) = 1$ since $T1 = 1$. Moreover $|\varrho(f)| \leq \|f\|_{\infty}$ because T has norm 1. Since $C^1(\Omega)$ is dense in $C(\Omega)$ ϱ extends uniquely to a positive continuous linear functional on $C(\Omega)$ given by a probability measure. Now if σ is any probability measure with conditional probabilities $\{\mu_a\}_{a \in L}$ then $\sigma(\tau_k f) = \sigma(f)$ for all k , and therefore $\sigma(Tf) = \sigma(f)$. Hence if f is in $C^1(\Omega)$, $\sigma(f) = \sigma(T^n f)$ which converges to $\sigma(\varrho(f)) = \varrho(f)$.

Lemma 3.4. *For any functions f and g in $C(\Omega)$*

$$|T(fg) - (Tf)(Tg)|_{\infty} \leq \sum_{n=1}^{\infty} \delta_n(T_{n+1}f) \delta_n(T_{n+1}g). \quad (3.21)$$

Proof. If μ is a probability measure on some space and f and g are bounded real measurable functions then

$$\begin{aligned} |\mu(fg) - \mu(f)\mu(g)| &= |\mu[(f - \mu(f))(g - \mu(g))]| \\ &\leq \sup |f - \mu(f)| \sup |g - \mu(g)| \\ &\leq (\sup f - \inf f)(\sup g - \inf g) \end{aligned}$$

because $\inf f \leq \mu(f) \leq \sup f$. Hence in particular

$$|\tau_n(fg) - (\tau_n f)(\tau_n g)|_{\infty} \leq \delta_n(f) \delta_n(g) \quad (3.22)$$

if f and g are in $C(\Omega)$.

Now $T_n = \tau_n T_{n+1}$ so that

$$\begin{aligned} |T_n(fg) - (T_n f)(T_n g)|_\infty &\leq |\tau_n T_{n+1}(fg) - \tau_n [(T_{n+1} f)(T_{n+1} g)]|_\infty \\ &\quad + |\tau_n [(T_{n+1} f)(T_{n+1} g)] - (\tau_n T_{n+1} f)(\tau_n T_{n+1} g)|_\infty \\ &\leq |T_{n+1}(fg) - (T_{n+1} f)(T_{n+1} g)|_\infty + \delta_n(T_{n+1} f) \delta_n(T_{n+1} g). \end{aligned}$$

Summing this collapsing sequence of inequalities from $n=1$ to p we get

$$\begin{aligned} |T(fg) - (Tf)(Tg)|_\infty &\leq \sum_{n=1}^p \delta_n(T_{n+1} f) \delta_n(T_{n+1} g) \\ &\quad + |T_{p+1}(fg) - (T_{p+1} f)(T_{p+1} g)|_\infty. \end{aligned} \quad (3.23)$$

But if f and g depend only on the coordinates s_1, \dots, s_m then the last term on the right is zero as soon as $p \geq m$ because $T_{p+1} f = f$, etc. A standard approximation argument for bilinear expressions then shows that

$$\lim_{p \rightarrow \infty} |T_{p+1}(fg) - (T_{p+1} f)(T_{p+1} g)|_\infty = 0$$

for all f and g in $C(\Omega)$. The lemma follows then by letting $p \rightarrow \infty$ in (3.23).

Proof of Theorem 1. We first estimate the right side of (3.21). By the triangle inequality for d we have

$$\begin{aligned} &\sum_{n=1}^{\infty} \delta_n(T_{n+1} f) \delta_n(T_{n+1} g) \\ &\leq \sum_{n=1}^{\infty} e^{-d(a,b)} [e^{d(a,a_n)} \delta_n(T_{n+1} f)] [e^{d(b,a_n)} \delta_n(T_{n+1} g)] \\ &\leq e^{-d(a,b)} \left[\sum_{n=1}^{\infty} e^{d(a,a_n)} \delta_n(T_{n+1} f) \right] \left[\sum_{n=1}^{\infty} e^{d(b,a_n)} \delta_n(T_{n+1} g) \right]. \end{aligned} \quad (3.24)$$

Referring to the proof of Corollary 3.2 for notation, we see that the same argument that establishes (3.19) also shows that

$$D_i(T_n f) \leq \sum_{k=1}^{\infty} B_{i,k}^{(n)} D_k(f). \quad (3.25)$$

Hence

$$\begin{aligned} \sum_{n=1}^{\infty} e^{d(a,a_n)} \delta_n(T_{n+1} f) &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} B_{n,k}^{(n+1)} D_k(f) \\ &\leq (1-\alpha)^{-1} \sum_{k=1}^{\infty} D_k(f) \quad \text{by (3.12)} \\ &\leq (1-\alpha)^{-1} \|f\|_a. \end{aligned} \quad (3.26)$$

Similarly $\sum_{n=1}^{\infty} e^{d(b,a_n)} \delta_n(T_{n+1} g) \leq (1-\alpha)^{-1} \|g\|_b$. Combining these inequalities with (3.24) and (3.21) we get

$$|T(fg) - (Tf)(Tg)|_\infty \leq e^{-d(a,b)} \|f\|_a \|g\|_b (1-\alpha)^{-2}. \quad (3.27)$$

Now we repeat part of the technique of Lemma 3.4 thus:

$$\begin{aligned}
|T^n(fg) - (T^n f)(T^n g)|_\infty &\leq |T[T^{n-1}(fg) - (T^{n-1} f)(T^{n-1} g)]|_\infty \\
&\quad + |T[(T^{n-1} f)(T^{n-1} g)] - (T^n f)(T^n g)|_\infty \\
&\leq |T^{n-1}(fg) - (T^{n-1} f)(T^{n-1} g)|_\infty + e^{-d(a,b)} \|T^{n-1} f\|_a \|T^{n-1} g\|_b (1-\alpha)^{-2} \\
&\leq |T^{n-1}(fg) - (T^{n-1} f)(T^{n-1} g)|_\infty + e^{-d(a,b)} \|f\|_a \|g\|_b \alpha^{2(n-1)} (1-\alpha)^{-2}.
\end{aligned}$$

In the last line we have used Corollary 3.2 and in the line before that we used (3.27). By induction we get

$$|T^m(fg) - (T^m f)(T^m g)|_\infty \leq e^{-d(a,b)} \|f\|_a \|g\|_b (1-\alpha)^{-2} \sum_{n=1}^m \alpha^{2(n-1)}. \quad (3.28)$$

By Corollary 3.3 the left side converges to $|\sigma(fg) - \sigma(f)\sigma(g)|$ while the sum on the right converges to $(1-\alpha^2)^{-1}$. This proves (2.9). The uniqueness of σ is proved in Corollary 3.3.

Proof of Theorem 2. Let f and g be in $C(\Omega)$ and apply Lemma 3.4 to the functions $T^r f$ and $T^r g_a$. Multiply by $e^{d(a,c)}$ and use the triangle inequality for d to get

$$\begin{aligned}
|T[(T^r f)(T^r g_a)] - (T^{r+1} f)(T^{r+1} g_a)|_\infty &e^{d(a,c)} \\
&\leq \sum_{n=1}^{\infty} [e^{d(c,a_n)} \delta_n(T_{n+1} T^r f)] [e^{d(a,a_n)} \delta_n(T_{n+1} T^r g_a)].
\end{aligned} \quad (3.29)$$

Apply (3.19) and (3.25) successively to get

$$\begin{aligned}
e^{d(a,a_n)} \delta_n(T_{n+1} T^r g_a) &\equiv D_n(T_{n+1} T^r g_a) \\
&\leq (B^{(n+1)} A^r D g_a)_n \\
&= \sum_{k=1}^{\infty} (B^{(n+1)} A^r)_{n,k} D_k(g_a).
\end{aligned} \quad (3.30)$$

Now $\delta_b(g_a) = \delta_{b-a}(g)$ as follows from the definitions of δ_b and g_a . Moreover by the definition of D_j in Corollary 3.2 we have (the element a of Corollary 3.2 is the same a as in g_a)

$$D_k g_a = e^{d(a,a_k)} \delta_k(g_a) = e^{d(0,a_k-a)} \delta_{a_k-a}(g)$$

by the translation invariance of d .

Summing over a we get

$$\begin{aligned}
\sum_{a \in L} D_k g_a &= \sum_{a \in L} e^{d(0,a_k-a)} \delta_{a_k-a}(g) \\
&= \|g\|_0.
\end{aligned}$$

Therefore if we sum (3.30) over a we get

$$\sum_{a \in L} e^{d(a,a_n)} \delta_n(T_{n+1} T^r g_a) \leq \|g\|_0 \sum_{k=1}^{\infty} (B^{(n+1)} A^r)_{n,k} \quad (3.31)$$

But the inequalities (3.13) and (3.14) applied repeatedly yield

$$\sum_{k=1}^{\infty} (B^{(n+1)}A^r)_{n,k} \leq (1-\gamma)^{-1}(\gamma/(1-\gamma))^r.$$

Inserting this in (3.31) we may now sum (3.29) over a and get

$$\begin{aligned} & \sum_{a \in L} |T[(T^r f)(T^r g_a)] - (T^{r+1} f)(T^{r+1} g_a)|_{\infty} e^{d(a,c)} \\ & \leq \sum_{n=1}^{\infty} [e^{d(c,a_n)} \delta_n(T_{n+1} T^r f)] \|g\|_0 (1-\gamma)^{-1} (\gamma/(1-\gamma))^r \\ & \leq (1-\alpha)^{-1} \|T^r f\|_c \|g\|_0 (1-\gamma)^{-1} (\gamma/(1-\gamma))^r \\ & \leq (1-\alpha)^{-1} \alpha^r \|f\|_c \|g\|_0 (1-\gamma)^{-1} (\gamma/(1-\gamma))^r \end{aligned} \quad (3.32)$$

where we have used (3.26) in the next to the last line and Corollary 3.2 in the last line.

We now use a part of the technique of Lemma 3.4 for the third time.

$$\begin{aligned} & \sum_a |T^{r+1}(fg_a) - (T^{r+1} f)(T^{r+1} g_a)|_{\infty} e^{d(a,c)} \\ & \leq \sum_a |T[T^r(fg_a) - (T^r f)(T^r g_a)]|_{\infty} e^{d(a,c)} \\ & \quad + \sum_a |T[(T^r f)(T^r g_a)] - (T^{r+1} f)(T^{r+1} g_a)|_{\infty} e^{d(a,c)} \\ & \leq \sum_a |T^r(fg_a) - (T^r f)(T^r g_a)|_{\infty} e^{d(a,c)} \\ & \quad + \|f\|_c \|g\|_0 (1-\alpha)^{-1} (1-\gamma)^{-1} (\alpha\gamma/(1-\gamma))^r \end{aligned}$$

by (3.32). Summing this collapsing sequence of inequalities from $r=0$ to $n-1$ we get

$$\begin{aligned} & \sum_a |T^n(fg_a) - (T^n f)(T^n g_a)|_{\infty} e^{d(a,c)} \\ & \leq \|f\|_c \|g\|_0 (1-\alpha)^{-1} (1-\gamma)^{-1} (1-\alpha\gamma(1-\gamma)^{-1})^{-1} \end{aligned}$$

(which is clearly valid also even if some terms of the preceding inequalities were infinite). Using Corollary 3.3 and Fatou's Lemma we may take the limit as $n \rightarrow \infty$ to conclude the proof of Theorem 2.

4. Applications to Classical Lattice Spin Systems

Continuing the notation of §2 we let X be a compact metric space, L a countable set, and $\Omega = X^L$. We let ν be a finite measure on X – the a priori single spin measure. An interaction is given by a real valued function φ on the union of all X^A where A runs over all non empty finite subsets of L . It will be convenient to denote the restriction of φ to X^A by φ_A , and, moreover, we shall regard $\varphi_A(s)$ as a function on Ω which depends only on $s|_A$, i.e., only on the configuration s inside A . We assume further that each φ_A is continuous. The potentials φ_A determine the total internal

energy of that portion of the system which lies in a finite set A by means of $U_A(s) = \sum_{A \subset A} \varphi_A(s)$. $U_A(s)$ depends of course only on the spin configuration inside A .

Write $|A|$ for the cardinality of A and put

$$\|\varphi\|_k = \sup_{a \in L} \sum_{A \ni a} |A|^{k-1} \sup_{s \in \Omega} |\varphi_A(s)| \quad k=0, 1, 2, \dots \quad (4.1)$$

Let \mathcal{P}_k be the Banach space of potentials φ for which $\|\varphi\|_k < \infty$.

Although the pressure is definable in the translation invariant subspace, $\hat{\mathcal{P}}_0$, of \mathcal{P}_0 [16], an example of Fisher [5] shows that the pressure does not depend continuously on the density at some points of this space. An abstract analog (i.e., lack of strict convexity of the pressure) of this physically unrealistic circumstance has been shown to occur on a dense set in $\hat{\mathcal{P}}_0$ by Israel [10, Theorem 3]. We shall not consider \mathcal{P}_0 further. In contrast, strict convexity of the pressure (taking into account equivalence of potentials) is known in the translation invariant subspace, $\hat{\mathcal{P}}_1$, of \mathcal{P}_1 [7] and hence in $\hat{\mathcal{P}}_k$, $k \geq 1$. A proof of a strong form of strict convexity in $\hat{\mathcal{P}}_2$ has been given in [3]. We shall not be concerned with the pressure nor require translation invariance of the potentials in this article.

Let φ be in \mathcal{P}_1 . For any finite set $A \subset L$ and s in Ω put

$$W_A(s) = \sum_{A \cap A \neq \emptyset} \varphi_A(s). \quad (4.2)$$

The series converges uniformly, since each term is dominated by some corresponding term of the convergent series $\sum_{a \in A} \sum_{A \ni a} \sup_s |\varphi_A(s)|$ whose sum is no greater than $|A| \|\varphi\|_1$. We write, for any finite set $A \subset L$,

$$v^A = \prod_{a \in A} v$$

which is a finite measure on X^A . For each t in X^{A^c} put

$$Z_A(t) = \int_{X^A} e^{-W_A(y \vee t)} v^A(dy) \quad (4.3)$$

and

$$\mu_A(dy|t) = Z_A(t)^{-1} e^{-W_A(y \vee t)} v^A(dy) \quad (4.4)$$

where $y \in X^A$ and $y \vee t$ is the product element in Ω . Then $\mu_A(\cdot|t)$ is a probability measure on X^A for each t . As in the previous section we may regard $\mu_A(\cdot|s)$ as a measure valued function of $s \in \Omega$ which depends only on $s|A^c$. If A consists of a single point a we write $\mu_a(\cdot|s)$, $W_a(s)$ and $Z_a(s)$.

For each finite set $A \subset L$ and each continuous real valued function f on Ω define

$$(\tau_A f)(s) = \int_{X^A} f(y \vee (s|A^c)) \mu_A(dy|s). \quad (4.5)$$

It is easy to verify that $\tau_A f$ is again in $C(\Omega)$.

A *Gibbs state* for φ is a probability measure σ on Ω satisfying the following equations (the D.L.R. equations [1], [13]).

$$\sigma(\tau_A f) = \sigma(f) \quad \text{for all } f \text{ in } C(\Omega) \text{ and all finite } A \subset L. \quad (4.6)$$

Now let

$$\zeta(a, b) = \sup_s \left| \sum_{A \supset \{a, b\}} \varphi_A(s) \right| \quad a, b \in L, \quad a \neq b. \quad (4.7)$$

$\zeta(a, b)$ is a measure of the strength of the total interaction between particles at points a and b .

Theorem 3. *Let φ be in \mathcal{P}_1 and let d be a semimetric on L . Put*

$$\alpha_0 = 2e^{4\|\varphi\|_1} \sup_{a \in L} \sum_{b \in L, b \neq a} \zeta(a, b) e^{d(a, b)}. \quad (4.8)$$

i) *If $\alpha_0 < 1$ then φ has a unique Gibbs state σ and for any functions f and g in $C(\Omega)$ and points a, b in L*

$$|\sigma(fg) - \sigma(f)\sigma(g)| \leq e^{-d(a, b)} \|f\|_a \|g\|_b (1 - \alpha_0)^{-2} (1 - \alpha_0^2)^{-1}. \quad (4.9)$$

ii) *If $L = \mathbb{Z}^m$, d is translation invariant and $\alpha_0 < (\sqrt{5} - 1)/2 \approx 0.62$ then for any functions f and g in $C(\Omega)$ and any point c in L*

$$\sum_{a \in \mathbb{Z}^m} |\sigma(fg_a) - \sigma(f)\sigma(g_a)| e^{d(a, c)} \leq \|f\|_c \|g\|_0 (1 - \alpha_0)^{-1} (1 - \alpha_0 - \alpha_0^2)^{-1}. \quad (4.10)$$

Lemma 4.1. *If $a \neq b$ and $s = t$ except at b then*

$$\|\mu_a(\cdot | t) - \mu_a(\cdot | s)\|_{\text{var}} \leq e^{4\zeta(a, b)} - 1. \quad (4.11)$$

Moreover

$$\varrho_{b, a} \leq 2e^{4\|\varphi\|_1} \zeta(a, b) \quad (4.12)$$

where $\varrho_{a, b}$ is given by (2.3).

Proof. It will be convenient to regard s and t as elements of $X^{(a)^c}$. Then

$$\begin{aligned} \|\mu_a(\cdot | t) - \mu_a(\cdot | s)\|_{\text{var}} &= \int_X \left| \frac{e^{-W_a(x \vee t)}}{Z_a(t)} - \frac{e^{-W_a(x \vee s)}}{Z_a(s)} \right| \nu(dx) \\ &= \int \left| \frac{Z_a(s)}{Z_a(t)} e^{W_a(x \vee s) - W_a(x \vee t)} - 1 \right| \frac{e^{-W_a(x \vee s)}}{Z_a(s)} \nu(dx) \\ &\leq \sup_{x \in X} \left| \frac{Z_a(s)}{Z_a(t)} e^{W_a(x \vee s) - W_a(x \vee t)} - 1 \right|. \end{aligned} \quad (4.13)$$

Now

$$\begin{aligned} |W_a(x \vee s) - W_a(x \vee t)| &= \left| \sum_{A \ni a} (\varphi_A(x \vee s) - \varphi_A(x \vee t)) \right| \\ &= \left| \sum_{A \supset \{a, b\}} (\varphi_A(x \vee s) - \varphi_A(x \vee t)) \right| \\ &\leq 2\zeta(a, b). \end{aligned} \quad (4.14)$$

But

$$\frac{Z_a(s)}{Z_a(t)} = \int e^{-(W_a(x \vee s) - W_a(x \vee t))} \left(\frac{e^{-W_a(x \vee t)}}{Z_a(t)} \nu(dx) \right).$$

Thus the ratio on the left is dominated above and below by the maximum and minimum, respectively, of the first factor under the integral, since the other factor is a probability measure. Thus in view of (4.14) we have $e^{-2\zeta(a,b)} \leq Z_a(s)/Z_a(t) \leq e^{2\zeta(a,b)}$. Hence

$$e^{-4\zeta(a,b)} \leq Z_a(s)Z_a(t)^{-1} e^{W_a(x \vee s) - W_a(x \vee t)} \leq e^{4\zeta(a,b)}. \quad (4.15)$$

Since $1 - e^{-u} \leq e^u - 1$ when $u \geq 0$ (4.11) now follows from (4.13) and (4.15). But

$$\zeta(a, b) \leq \|\varphi\|_1 \quad \text{and} \quad e^x - 1 \leq xe^x \quad \text{when} \quad x \geq 0.$$

Hence (4.12) also follows.

Proof of Theorem 3. Since $\zeta(a, b) = \zeta(b, a)$ we see, referring to Eqs. (2.4), (2.5) and (4.12), that

$$\begin{aligned} \sum_{b \in L} \beta_{a,b} &= \sum_{b \in L} \varrho_{a,b} e^{d(a,b)} \\ &\leq 2e^{4\|\varphi\|_1} \sum_{b \in L, b \neq a} \zeta(a,b) e^{d(a,b)}. \end{aligned}$$

Thus $\alpha \leq \alpha_0$, and similarly $\gamma \leq \alpha_0$. Hence $\alpha < 1$ and Theorem 1 applies to yield i). The existence of the Gibbs state σ is well known (see e.g. [1] or [17]). If $\alpha_0 < (\sqrt{5}-1)/2$ then $\gamma(1+\alpha) \leq \alpha_0 + \alpha_0^2 < 1$ and Theorem 2 may be applied to yield ii).

Here is another special case of Dobrushin's uniqueness theorem.

Corollary 4.2. *If φ is in \mathcal{P}_2 and*

$$2\|\varphi\|_2 e^{4\|\varphi\|_1} < 1 \quad (4.16)$$

then φ has only one Gibbs state.

Proof.

$$\begin{aligned} \sum_{b: b \neq a} \zeta(a, b) &= \sum_{b: b \neq a} \sup_s \left| \sum_{A \supset \{a,b\}} \varphi_A(s) \right| \\ &\leq \sum_{b: b \neq a} \sum_{A \supset \{a,b\}} \sup_s |\varphi_A(s)| \\ &= \sum_{A \ni a} (|A| - 1) \sup_s |\varphi_A(s)| \\ &\leq \|\varphi\|_2. \end{aligned} \quad (4.17)$$

Thus $\alpha_0 \leq 2e^{4\|\varphi\|_1} \|\varphi\|_2$ when $d \equiv 0$. Theorem 3i) is applicable.

To make the temperature dependence explicit one need only replace φ by $\beta\varphi$ in Theorem 3 where $\beta = 1/kT$. Thus

Corollary 4.3. *If φ is in \mathcal{P}_1 on Z^m and d is a translation invariant semimetric on Z^m such that*

$$\sup_{a \in Z^m} \sum_{b: b \neq a} \zeta(a, b) e^{d(a,b)} < \infty$$

then for all sufficiently small real β the strong decay inequality (4.10) holds for the Gibbs state belonging to $\beta\varphi$.

The results of this section have so far not made use of any particular properties of the a priori single spin measure ν . The extent to which ν is “concentrated at a point” on the one hand or “smeared out” on the other will clearly affect the conditional probabilities μ_A , (4.4), and their variations, $\|\mu_a(\cdot|t) - \mu_a(\cdot|s)\|_{\text{var}}$. It is sometimes convenient to regard the external field component of an interaction φ (i.e., φ_A with $|A|=1$) as influencing the single spin measure ν . We illustrate in the simplest case how the theorems of the preceding section yield results on decay of correlations at any temperature but at strong external field (or equivalently, low activity).

We take $X = \{0, 1\}$. We consider only potentials φ (so-called vacuum potentials [3]) such that $\varphi_A(s) = 0$ unless $s \equiv 1$ on A . (Such a potential is determined by the set function $\psi(A) \equiv \varphi_A(1)$.) We write φ' for the potential obtained from φ by replacing all one particle potentials $\varphi_{\{a\}}$ by zero. For simplicity of statement and ease of comparison with other results in the literature we assume $\varphi_{\{a\}}(1) = h$ is independent of a .

Corollary 4.4. *With the notation of the preceding paragraph, with $\varphi \in \mathcal{P}_1$ and $\nu(\{0\}) = \nu(\{1\}) = 1/2$, put*

$$\alpha' = e^{-h} e^{\|\varphi'\|_1} \sup_{a \in L} \sum_{b: b \neq a} (1 - e^{-\zeta(a,b)}) e^{d(a,b)} \quad (4.18)$$

i) *If $\alpha' < 1$ then φ has a unique Gibbs state σ and*

$$|\sigma(s_a s_b) - \sigma(s_a)\sigma(s_b)| \leq e^{-d(a,b)} (1 - \alpha')^{-2} (1 - \alpha'^2)^{-1}. \quad (4.19)$$

ii) *If $\alpha' < (\sqrt{5} - 1)/2$, $L = \mathbb{Z}^m$, and the semimetric d is translation invariant then*

$$\sum_{a \in \mathbb{Z}^m} |\sigma(s_c s_a) - \sigma(s_c)\sigma(s_a)| e^{d(a,c)} \leq (1 - \alpha')^{-1} (1 - \alpha' - \alpha'^2)^{-1} \quad (4.20)$$

for any point c .

Remark. The conclusions i) and ii) hold also in the stronger forms given in Theorems 1 and 2. Many of the details of the proof of the corollary are extant in the literature, but we repeat them here for the reader’s convenience.

Proof. $W_a(s) \equiv \sum_{A \ni a} \varphi_A(s)$ is zero if $s_a = 0$ while

$$W_a(s) = \left[\sum_{A \ni a} \varphi'_A(s) \right] + h \quad \text{if } s_a = 1. \quad (4.21)$$

Thus if $t \in X^{\{a\}^c}$ then $Z_a(t) = (1 + e^{-W_a(1 \vee t)})/2$ where $1 \vee t$ is 1 at a and t elsewhere. For two probability measures on X with weights p_0, p_1 and q_0, q_1 respectively the variation of their difference is $|p_0 - q_0| + |p_1 - q_1| = 2|p_0 - q_0|$. But

$$\mu_a(\{0\}|t) = e^{-W_a(0 \vee t)} (2Z_a(t))^{-1} = (2Z_a(t))^{-1}.$$

Hence if $s \in X^{\{a\}^c}$ and differs from t only at b then

$$\|\mu_a(\cdot|t) - \mu_a(\cdot|s)\|_{\text{var}} = 2|(2Z_a(t))^{-1} - (2Z_a(s))^{-1}| = |Z_a(s) - Z_a(t)|(Z_a(t)Z_a(s))^{-1}.$$

Now we may assume without loss of generality that $W_a(1 \vee t) - W_a(1 \vee s) \leq 0$. Then

$$\begin{aligned} \|\mu_a(\cdot|t) - \mu_a(\cdot|s)\|_{\text{var}} &= 2 \left| \frac{e^{W_a(1 \vee t) - W_a(1 \vee s)} - 1}{1 + e^{-W_a(1 \vee s)}} \right| \frac{e^{-W_a(1 \vee t)}}{1 + e^{-W_a(1 \vee t)}} \\ &\leq 2(1 - e^{W_a(1 \vee t) - W_a(1 \vee s)})e^{-W_a(1 \vee t)}. \end{aligned}$$

From (4.21) $W_a(1 \vee t) \geq h - \|\varphi'\|_1$ for all t while

$$\begin{aligned} W_a(1 \vee t) - W_a(1 \vee s) &\approx \sum_{A \ni a} (\varphi_A(1 \vee t) - \varphi_A(1 \vee s)) \\ &\approx \sum_{A \supset \{a, b\}} (\varphi_A(1 \vee t) - \varphi_A(1 \vee s)). \end{aligned}$$

If $t(b)=0$ the first summand $\varphi_A(1 \vee t)=0$ and if $s(b)=0$ the second summand is zero. Thus in either case we have $|W_a(1 \vee t) - W_a(1 \vee s)| \leq \zeta(a, b)$. Hence

$$(1/2)\|\mu_a(\cdot|t) - \mu_a(\cdot|s)\|_{\text{var}} \leq (1 - e^{-\zeta(a, b)})e^{-h}e^{\|\varphi'\|_1}.$$

Thus the right side of this inequality gives an upper bound for $q_{b, a}$. Hence $\alpha \leq \alpha'$ where α is given by (2.5). Since ζ is symmetric $\gamma \leq \alpha'$ also. Therefore Theorems 1 and 2 apply. We have formulated the conclusion of the corollary in the weak form described in the Example of §2 to make it more familiar looking to some readers.

Example 4.5. If φ involves at most two body interactions then $\zeta(a, b) = |\varphi_{\{a, b\}}(1)|$ in the context of the last corollary. In particular if the interaction as well as d is translation invariant then the hypothesis of Corollary 4.4 becomes

$$\alpha' \equiv e^{-h} e^{a \sum_0^{v(a)}} \sum_{b \neq 0} (1 - e^{-|v(b)|}) e^{d(0, b)} < 1$$

or $\alpha' < (\sqrt{5} - 1)/2$ for the parts i) and ii) respectively, where $v(a - b) = \varphi_{\{a, b\}}(1)$. Moreover if the potential v is nonnegative then one may estimate $e^{-W_a(1 \vee t)}$ by e^{-h} in the proof of Corollary 4.4, and then the conclusions i) and ii) hold with α' replaced by

$$\alpha' = e^{-h} \sum_{b \neq 0} (1 - e^{-v(b)}) e^{d(0, b)}.$$

Remark 4.6. In case $d=0$ in Corollary 4.4 we recover some well known uniqueness theorems (c.f. [2, 6, 12, 16]) for Gibbs states. By (4.17) $\sum_{b: b \neq a} \zeta(a, b) \leq \|\varphi'\|_2$. Hence it follows from Corollary 4.4 that φ has a unique Gibbs state if $\|\varphi'\|_2 < \infty$ and h is sufficiently large. Explicitly, uniqueness holds if $e^{-h} e^{\|\varphi'\|_1} \|\varphi'\|_2 < 1$ while the summability conclusion (4.20) holds with $d=0$ if $e^{-h} e^{\|\varphi'\|_1} \|\varphi'\|_2 < (\sqrt{5} - 1)/2$. We note however that stronger results have been proven for the case $d=0, X = \{0, 1\}$ by Gallavotti and Miracle-Sole [6]. They only require $\|\varphi'\|_1 < \infty$ but allow $\|\varphi'\|_2 = \infty$.

Remark 4.7. From the point of view of the infinite system a mathematically more natural object than the potential that describes the interaction is the relative Hamiltonian, $H(s, t)$, defined by

$$H(s, t) = \sum_{A \subset L} (\varphi_A(s) - \varphi_A(t)) \quad (4.22)$$

for two spin configurations s, t which differ at only finitely many sites. The relative Hamiltonian is the basic object in several treatments of lattice models and is more intimately connected with its Gibbs states than are the potentials [9, 14, 15]. Theorem 3 can also be formulated directly in terms of the relative Hamiltonian. Define

$$\|H\|_1 = \sup_{a \in L} \sup \{ |H(s, t)| : s = t \text{ off } \{a\} \}. \quad (4.23)$$

If \mathcal{D} is the set of ordered pairs (s, t) of spin configurations which differ at only finitely many sites, then the space \mathcal{H}_1 of anti-symmetric $(H(s, t) = -H(t, s))$ continuous functions on \mathcal{D} which satisfy $H(s, t) = H(s, u) + H(u, t)$ whenever $s = t = u$ except on a finite set, is a Banach space in the norm $\|\cdot\|_1$. The mapping $\varphi \rightarrow H$ given by (4.22) clearly satisfies $\|H\|_1 \leq 2\|\varphi\|_1$, and hence is continuous from \mathcal{P}_1 into \mathcal{H}_1 . Moreover, if $a_1 \neq a_2$ put $\lambda(a_1, a_2) = \sup \{ |H(s, t) - H(u_1, t) - H(u_2, t)| : s = t \text{ except at } a_1 \text{ and } a_2 \}$ where $u_j = s$ at a_j and equals t otherwise. Then one sees easily that $\lambda(a, b) \leq 4\zeta(a, b)$ for $a \neq b$. If the interaction has range R (i.e., $\varphi_A(s) = 0$ if $\text{diam } A > R$) then $\lambda(a, b) = 0$ and $\zeta(a, b) = 0$ if $\text{dist}(a, b) > R$. Both $\lambda(a, b)$ and $\zeta(a, b)$ measure the deviation from finite range. Put $\|H\|_2 = \|H\|_1 + \sup_{a, b} \lambda(a, b)$ where we take $\lambda(a, a) \equiv 0$. Then $\|H\|_2 \leq 2\|\varphi\|_1 + 4\|\varphi\|_2$ and in the translation invariant subspace of \mathcal{P}_2 $\|H\|_2 \leq 4\|\varphi\|_2$. The techniques of Lemma 4.1 and Corollary 4.2 also show that $q_{b, a} \leq (e^{2\lambda(a, b)} - 1)/2 \leq e^{4\|H\|_1} \lambda(a, b)$. Hence uniqueness of Gibbs states holds if $\|H\|_2 e^{2\|H\|_1} \leq 1$. Similarly decay of correlations can be studied in terms of $\lambda(a, b)$ instead of $\zeta(a, b)$. Thus all the hypotheses of this section can be formulated in terms of the relative Hamiltonian H . Higher order difference norms on H are readily defined corresponding to the norms $\|\varphi\|_k$ with corresponding Banach spaces \mathcal{H}_k . We conjecture that the free energy is k times continuously differentiable in a neighborhood of the origin in the translation invariant subspace of \mathcal{P}_k (or \mathcal{H}_k) for any single spin space X . But compare [6, 11].

Note. After completion of the manuscript for this article I learned from Robert B. Israel that he has independently obtained correlation decay inequalities for a general class of classical lattice models. His techniques and norms are quite different from those used in the present article.

Acknowledgements. The author wishes to acknowledge many illuminating discussions of statistical mechanics with H. Epstein, M. E. Fisher, R. B. Griffiths, H. Kesten, J. L. Lebowitz, D. Ruelle, Ya. G. Sinai, K. Wilson and especially with G. Gallavotti, O. McBryan, F. Spitzer, and B. Widom. I would also like to thank Professor N. H. Kuiper for extending the hospitality of the I.H.E.S.

References

1. Dobrushin, R.L.: The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory Probab. Its Appl.* **13**, 197–224 (1968)
2. Dobrushin, R.L.: The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. *Funct. Anal. Its Appl.* **2**, 302–312 (1968)
3. Dobrushin, R.L., Nakhapetyan, B.S.: Strong convexity of the pressure for lattice systems of classical statistical physics. *Theor. Math. Phys.* **20**, 782–790 (1974)
4. Duneau, M., Souillard, B., Iagolnitzer, D.: Decay of correlations for infinite-range interactions. *J. Math. Phys.* **16**, 1662–1666 (1975)
5. Fisher, M.E.: On discontinuity of the pressure. *Commun. Math. Phys.* **26**, 6–14 (1972)

6. Gallavotti, G., Miracle-Sole, S.: Correlation functions of a lattice system. *Commun. Math. Phys.* **7**, 274–288 (1968)
7. Griffiths, R., Ruelle, D.: Strict convexity (continuity) of the pressure in lattice systems. *Commun. Math. Phys.* **23**, 169–175 (1971)
8. Holley, R.A., Stroock, D.W.: Applications of the stochastic Ising model to the Gibbs states. *Commun. Math. Phys.* **48**, 249–265 (1976)
9. Holsztynski, W., Slawny, J.: Peierls condition and the number of ground states. *Commun. Math. Phys.* **61**, 177–190 (1978)
10. Israel, R.B.: Existence of phase transitions for long-range interactions. *Commun. Math. Phys.* **43**, 59–68 (1975)
11. Israel, R.B.: High temperature analyticity in classical lattice systems. *Commun. Math. Phys.* **50**, 245–257 (1976)
12. Lanford, O.E. III: Entropy and equilibrium states in classical statistical mechanics. In: *Lecture notes in physics*, Vol. 20. *Statistical mechanics and mathematical problems*, Lenard, A. (ed.). Berlin, Heidelberg, New York: Springer 1973
13. Lanford, O.E. III, Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. *Commun. Math. Phys.* **13**, 194–215 (1969)
14. Pirogov, S.A., Sinai, Ya.G.: Phase diagrams of classical lattice systems. *Theor. Math. Phys.* **25**, 358–369 (1975)
15. Pirogov, S.A., Sinai, Ya.G.: Ground states in two-dimensional Boson quantum field theory. *Ann. Phys.* **109**, 393–400 (1977)
16. Ruelle, D.: *Statistical mechanics*, New York: Benjamin 1969
17. Ruelle, D.: *Thermodynamic Formalism*. Reading, Mass.: Addison-Wesley 1978
18. Sylvester, G.S.: Weakly coupled Gibbs measures. Preprint, Rockefeller University (1977)
19. Vasershtein, L.N.: Markov processes over denumerable products of spaces, describing large systems of automata. *Probl. Trans. Inf.* **5**, 64–72 (1969) (English Translation)
20. Duneau, M., Souillard, B.: Cluster properties of lattice and continuous systems. *Commun. Math. Phys.* **47**, 155–166 (1976)
21. Iagolnitzer, D., Souillard, B.: Decay of correlations for slowly decreasing potentials. *Phys. Rev. A* **16**, 1700–1704 (1977)
22. Iagolnitzer, D., Souillard, B.: On the analyticity in the potential in classical statistical mechanics. *Commun. Math. Phys.* **60**, 131–152 (1978)

Communicated by E. Lieb

Received January 15, 1979

