Notes on correlation inequalities and the existence of Gibbs measures for the Ising model

Benjamin C. Wallace

April 2015

1 Introduction

We follow [1] and [2].

1.1 Definitions and notation

Throughtout, let $\Lambda \subseteq \mathbb{Z}^d$ be finite and let $\partial\Lambda$ denote the external boundary of Λ. Define $\Omega = \{\pm 1\}^{\mathbb{Z}^d}$ and $\Omega_\Lambda = \{\pm 1\}^\Lambda$. Let $\omega_\Lambda \in \Omega_\Lambda$ denote the restriction of $\omega \in \Omega$ to Λ. We may sometimes abuse notation by letting ω_Λ denote the cylinder set

$$\{\tilde{\omega} \in \Omega : \tilde{\omega}_\Lambda = \omega_\Lambda\}. \tag{1}$$

In particular, if $\delta^x \in \Omega_x$ is the configuration defined by $\delta^x_y = \delta_{xy}$, then the event δ^x occurs if the spin at x is positive.

Given disjoint Λ_1 and Λ_2 we denote the concatenation of configurations $\omega_{\Lambda_i} \in \Omega_{\Lambda_i}$ for $i = 1, 2$ by $\omega_{\Lambda_1}\omega_{\Lambda_2} \in \Omega_{\Lambda_1 \cup \Lambda_2}$.

We also define the projection maps $\sigma_x : \omega \mapsto \omega_x$. These will become random variables once Ω is equipped with a measure.

Definition 1. The *Hamiltonian* $H^{\eta,\Lambda}_{\beta,h} : \Omega_\Lambda \to \mathbb{R}$ *for the nearest-neighbour Ising model* at inverse temperature $\beta > 0$, with external field $h \in \mathbb{R}$, and with boundary condition $\eta \in \Omega$ off of Λ is defined by

$$H^{\eta,\Lambda}_{\beta,h}(\omega) = -\beta \left(\sum_{\substack{x\sim y \\ x,y\in\Lambda}} \omega_x\omega_y + h\sum_{x\in\Lambda}\omega_x + \sum_{\substack{x\sim y \\ x\in\Lambda, y\in\partial\Lambda}} \omega_x\eta_y \right). \tag{2}$$

We define a Gibbs specification by

$$\mu^{\eta,\Lambda}_{\beta,h}(\omega) = \frac{1}{Z^{\eta,\Lambda}_{\beta,h}} e^{-H^{\eta,\Lambda}_{\beta,h}(\omega)}, \quad Z^{\eta,\Lambda}_{\beta,h} = \sum_{\omega\in\Omega_\Lambda} e^{-H^{\eta,\Lambda}_{\beta,h}(\omega)}. \tag{3}$$

We will often drop the subscripts β and h.

Recall that the finite-volume measures satisfy the following spatial Markov property.

Proposition 2 (Markov property). *If $\Lambda_1 \subseteq \Lambda_2$, then*

$$\mu^{\xi,\Lambda_1}(\omega_{\Lambda_1}) = \mu^{\xi,\Lambda_2}(\omega_{\Lambda_1} \mid \xi_{\Lambda_2 \setminus \Lambda_1}). \tag{4}$$

We are ultimately interested in the structure of the set of *(infinite-volume) Gibbs measures* or *Gibbs states* of the Ising model, i.e. measures μ on Ω satisfying the Dobrushin-Lanford-Ruelle (DLR) equations

$$\mu(\omega_\Lambda \mid \eta_{\partial\Lambda}) = \mu^{\eta,\Lambda}_{\beta,h}(\omega), \tag{5}$$

for all ω, η, and Λ.

1.2 Main results

The first question that comes up is whether any Gibbs measure exists. This has already been established in class. In fact, we showed the following stronger result.

Theorem 3. *There exists a shift-invariant Gibbs measure for the Ising model on $\mathbb{Z}^d$.*

This measure was constructed (roughly) by averaging shifts of finite-volume measures over boxes and taking the limit. In this talk, we will discuss the construction of an infinite-volume measure without any kind of averaging.

Definition 4. Define the partial order on configurations in Ω component-wise. With respect to this ordering, we have a notion of *increasing* function on Ω. A measure μ on Ω has *positive correlations* if $\mu(fg) \geq \mu(f)\mu(g)$ for all bounded increasing functions f and g on Ω.

In particular, if μ has positive correlations and $\sigma \sim \mu$ (σ is a random variable with distribution μ), then $\mathrm{Cov}(\sigma_x, \sigma_y) \geq 0$ for all x, y. This means the spins have a tendency to align with one another.

Theorem 5.

1. *The (weak) limit*

$$\mu^+_{\beta,h} := \lim_{\Lambda \uparrow \mathbb{Z}^d} \mu^{+,\Lambda}_{\beta,h} \tag{6}$$

   exists and is a Gibbs measure for the Ising model.

2. *The Gibbs measure $\mu^+_{\beta,h}$ is invariant under (graph) automorphisms of $\mathbb{Z}^d$.*

3. *The Gibbs measure $\mu^+_{\beta,h}$ has positive correlations.*

4. *Analogues of the above statements hold for $\mu^-_{\beta,h}$.*

Note.

1. The lattice $\mathbb{Z}^d$ can be replaced by a locally finite graph.

2. We will focus on the case of + boundary conditions.

3. The limit as $\Lambda \uparrow \mathbb{Z}^d$ refers to the existence of a limit along any sequence $\Lambda_n \uparrow \mathbb{Z}^d$ as $n \to \infty$. It is easy to see that this limit, if it exists for all such sequences, is independent of the sequence Λ_n.

   For instance, fix sequences Λ_n and $\tilde{\Lambda}_n$ let $\mu^+ = \lim_{n\to\infty} \mu^{+,\Lambda_n}$ and $\tilde{\mu}^+ = \lim_{n\to\infty} \mu^{+,\tilde{\Lambda}_n}$. Consider the sequence

$$\hat{\Lambda}_n = \begin{cases} \Lambda_n & n \text{ even} \\ \tilde{\Lambda}_n & n \text{ odd} \end{cases} \tag{7}$$

   and let $\hat{\mu}^+ = \lim_{n\to\infty} \mu^{+,\hat{\Lambda}_n}$. Then μ^{+,Λ_n} and $\mu^{+,\tilde{\Lambda}_n}$ are both subsequences of $\mu^{+,\hat{\Lambda}_n}$. Since all three sequences converge, they must have the same limit.

2 Generalities

Recall the metric on Ω: except in trivial cases, $d(\omega, \omega') = 2^{-k(\omega,\omega')}$, where $k(\omega, \omega') = \max(j : \omega_{B_j} = \omega'_{B_j})$. Hence, $\omega^n \to \omega$ if ω^n can be made to agree with ω on an arbitrarily large central block by taking n large.

In particular, if $\omega^n \to \omega$, then $\mathbb{1}_{\tilde{\omega}_\Lambda}(\omega^n) \to \mathbb{1}_{\tilde{\omega}_\Lambda}(\omega)$. Thus, indicator functions on cylinder sets are continuous. Hence, if μ_n, μ are probability measures on Ω and $\mu_n \Rightarrow \mu$, then $\mu_n(\omega_\Lambda) \to \mu(\omega_\Lambda)$ (i.e. cylinder sets are μ-continuity sets). Of course, the converse holds as well since any measure is defined by the values it takes on cylinder sets.

Lemma 6. *If $\mu^{+,\Lambda}(f)$ converges for all bounded continuous f, then μ^+ exists and is an infinite-volume Gibbs measure.*

Proof. The existence of μ^+ is a consequence of the Riesz representation theorem.

That μ^+ is a Gibbs measure essentially follows from the fact that satisfaction of the DLR equations is a "microscopic condition". That is, we can write

$$\mu^+(\omega_\Lambda \mid \xi_{\partial\Lambda}) = \frac{\mu^+(\omega_\Lambda \xi_{\partial\Lambda})}{\mu^+(\xi_{\partial\Lambda})} = \lim_{\Delta\uparrow\mathbb{Z}^d} \frac{\mu^{+,\Delta}(\omega_\Lambda \xi_{\partial\Lambda})}{\mu^{+,\Delta}(\xi_{\partial\Lambda})}. \tag{8}$$

Taking Δ much larger than Λ, the Markov property allows us to replace the Δ inside the limit by Λ (for Δ large). Doing so, the inside becomes independent of Δ and we can get rid of the limit altogether; we are left with

$$\mu^+(\omega_\Lambda \mid \xi_{\partial\Lambda}) = \mu^{+,\Lambda}(\omega_\Lambda \mid \xi_{\partial\Lambda}), \tag{9}$$

which shows that μ^+ satisfies the DLR equations. $\square$

In fact, we can restriction our attention to a much smaller class of functions.

Definition 7. A function on Ω is *local* if it only depends on the values of spins in a finite box.

Lemma 8. *Any continuous function f on Ω can be uniformly approximated by local functions.*

For a proof, see [1, Lemma 6.35].

Thus, we need only check convergence on local functions. Indeed, suppose $\mu^{+,\Lambda}(g)$ converges for all local g. For f continuous, let f_n be a sequence of local functions converging to f uniformly. Write

$$|\mu_n(f) - \mu_m(f)| \leq \mu_n(|f - f_k|) + |\mu_n(f_k) - \mu_m(f_k)| + \mu_m(|f_k - f|). \tag{10}$$

Setting k large so that $\sup|f - f_k| < \epsilon/2$ and then taking the limit as $m, n \to \infty$, we see that $\mu_n(f)$ is a Cauchy sequence.

The task of checking convergence on local functions is greatly simplified by the existence of a good basis.

For $A \subseteq \Lambda$, let $\sigma_A : \Omega_\Lambda \to \mathbb{R}$ be defined by $\sigma_A(\omega) = \prod_{x\in\Lambda} \omega_x$. Define an inner-product on $\mathbb{R}^{\Omega_\Lambda}$ by $\langle f, g\rangle = 2^{-|\Lambda|} \sum_{\tilde{\omega}\in\Omega_\Lambda} f(\tilde{\omega})g(\tilde{\omega})$.

Lemma 9. *The collection $\{\sigma_A : A \subseteq \Lambda\}$ is an orthonormal basis of $\mathbb{R}^{\Omega_\Lambda}$ under this inner-product.*

Proof. Since there are the same number ($2^{|\Lambda|}$) of subsets of Λ and configurations on Λ, it suffices to show that the σ_A form an orthonormal set. It is easy to see that $\langle \sigma_A, \sigma_A\rangle = 1$. For $A \neq B$, we have

$$\langle \sigma_A, \sigma_B\rangle = \sum_{\tilde{\omega}} \prod_{x\in A} \tilde{\omega}_x \prod_{y\in B} \tilde{\omega}_y \tag{11}$$

$$= \sum_{\tilde{\omega}} \prod_{x\in A\setminus B} \tilde{\omega}_x \prod_{y\in B\setminus A} \tilde{\omega}_y \tag{12}$$

$$= \sum_{\tilde{\omega}} \prod_{x\in A\triangle B} \tilde{\omega}_x. \tag{13}$$

The result follows from the following claim (with $S = A\triangle B$): for any finite Λ and $S \subseteq \Lambda$,

$$\sum_{\tilde{\omega}\in\Omega_\Lambda} \prod_{x\in S} \tilde{\omega}_x = 0. \tag{14}$$

The claim is easily proved by induction (the correct inductive hypothesis is to assume that the claim holds for *all* finite Λ and all $S \subseteq \Lambda$ with $|S| < n$). □

Lemma 10 (Binomial theorem)**.** *For any set X and numbers $a_x, b_x \in \mathbb{R}$ for each $x \in X$,*

$$\prod_{x\in X}(a_x + b_x) = \sum_{Y\subseteq X} \prod_{x\in X\setminus Y} a_x \prod_{y\in Y} b_y. \tag{15}$$

Example 11. Let $a_x = a, b_x = b$ and suppose $|X| = n$. Then

$$(a+b)^n = \prod_{x\in X}(a+b) = \sum_{Y\subseteq X} a^{n-k}b^k = \sum_{k=0}^{n} \sum_{\substack{Y\subseteq X\\ |Y|=k}} a^{n-k}b^k = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}b^k. \tag{16}$$

Let $\mathbb{1}_{\delta^A}$ be the indicator function of the event $\delta^A = \{\omega_x = 1, x \in A\}$.

Corollary 12. *The collection* $\{\mathbb{1}_{\delta^A} : A \subseteq \Lambda\}$ *spans* $\mathbb{R}^{\Omega_\Lambda}$.

Proof. By the lemma, $f \in \mathbb{R}^{\Omega_\Lambda}$ can be written

$$f = \sum_{A \subseteq \Lambda} f_A \sigma_A, \tag{17}$$

where $f_A = \langle f, \sigma_A \rangle$. But observe that $\sigma_x = 2\mathbb{1}_{\delta^x} - 1$ and so by the binomial theorem,

$$\sigma_A = \sum_{Y \subseteq A} (-1)^{|A \setminus Y|} 2^{|Y|} \mathbb{1}_{\delta^Y}. \tag{18}$$

Thus, by interchanging the sums,

$$f = \sum_{Y \subseteq \Lambda} \tilde{f}_Y \mathbb{1}_{\delta^Y}, \tag{19}$$

where

$$\tilde{f}_Y = \sum_{A \supseteq Y} f_A (-1)^{|A \setminus Y|} 2^{|Y|}. \tag{20}$$

□

As we will see, the $\mathbb{1}_{\delta^A}$ are more convenient than the σ_A due to the fact that they are increasing functions.

We conclude that Theorem 5(1) is a consequence of the the following theorem.

Theorem 13. *For any* $A \subseteq \mathbb{Z}^d$ *finite,* $\mu^{+,\Lambda}(\delta^A)$ *converges.*

Since $\mu^{+,\Lambda}(\delta^A)$ is bounded, its convergence will follow if we can establish monotonicity in Λ.

Intuitively, if we increase Λ, then A will be further from $\partial\Lambda$, so the boundary condition will have less of an effect. Since the effect of the + boundary condition is to encourage positive spins, we would expect $\mu^{+,\Lambda}(\delta^A)$ to decrease. Most of our efforts will be devoted to proving the following lemma.

Lemma 14. $\mu^{+,\Lambda}(\delta^A)$ *decreases as* Λ *increases.*

3 Coupling and stochastic dominance

The key to proving Theorem 5 is a notion of monotonicity for measures.

Definition 15 (Stochastic dominance)**.** We write $\mu_1 \leq \mu_2$ if $\mu_1(f) \leq \mu_2(f)$ for all continuous bounded increasing functions f.

Note. The property of stochastic dominance is preserved under weak limits.

Example 16.

1. Suppose $|\Lambda| = 1$, so μ_1 and μ_2 are measures on $\mathbb{R}$. Then $\mu_1 \leq \mu_2$ implies that $F_1(x) \geq F_2(x)$ for all x, where F_i is the cdf of μ_i.

   In fact, the converse holds: just approximate a continuous f by simple functions and note that the indicator function of an interval $[a, b)$ is a difference of *increasing* indicator functions, which have the form $\mathbb{1}_{[r,\infty)}$.

2. Suppose

$$\mu^{+,\Lambda_2} \leq \mu^{+,\Lambda_1} \text{ whenever } \Lambda_1 \subseteq \Lambda_2. \tag{21}$$

   Then Lemma 14 follows immediately.

How does this simplify our problem? It turns out that there is a particularly nice way to establish stochastic dominance of measures.

Definition 17. A *coupling* of random variables $X_1 \sim \mu_1$ and $X_2 \sim \mu_2$ is a joint distribution μ for (X_1, X_2), i.e. a measure μ with marginals μ_1 and μ_2.

Example 18. Suppose X_1 and X_2 are Bernoulli(1/2) random variables with values in $\{\pm 1\}$. Thus, $\Pr(X_i = \pm 1) = \mu_i(\pm 1) = 1/2$ for $i = 1, 2$. The *independent coupling* of X_1 and X_2 is just the product measure $\mu_1 \otimes \mu_2$ with $\mu_1 \otimes \mu_2(\pm 1, \pm 1) = 1/4$, under which X_1 and X_2 are independent.

We can also construct a dependent coupling μ. For instance, consider μ as defined by the following table:

	-1	1	μ_1
-1	1/3	1/6	1/2
1	1/6	1/3	1/2
μ_2	1/2	1/2	

.

Theorem 19 (Strassen)**.** *The following are equivalent:*

1. $\mu_1 \leq \mu_2$
2. *There exists a coupling μ of μ_1 and μ_2 such that if $(X_1, X_2) \sim \mu$, then* $\Pr(X_1 \leq X_2) = 1$.

Definition 20. A coupling of X_1 and X_2 under which $\Pr(X_1 \leq X_2) = 1$ is known as a *monotone coupling.*

Note. It is easy to see that the existence of a monotone coupling implies stochastic dominance. The converse is the true content of Strassen's theorem. However, due to the length of its proof, we will not discuss it here.

The idea for measures μ_1 and μ_2 with cdf's F_1 and F_2 is as follows (from [3]). By Example 16(1), $\mu_1 \leq \mu_2$ implies $F_1 \geq F_2$. Let $F_i^*(u) = \inf(x : F_i(x) \geq u)$ be the generalized inverse of F_i. Then $F_1^* \leq F_2^*$. Now if U is uniformly distributed on $[0, 1]$, then

$$\Pr(F_i^*(U) \leq x) = \Pr(U \leq F_i(x)) = F_i(x) \tag{22}$$

for $x \in [0, 1]$. Thus, $X_i = F_i^*(U) \sim F_i$ and X_1, X_2 are jointly distributed with $X_1 \leq X_2$.

4 Correlation inequalities

We will now construct the desired coupling. In fact, we will deal with a somewhat more general situation.

Theorem 21 (Holley). *Let μ_1 and μ_2 be probability measures on Ω_Λ. Assume that μ_2 assigns positive probability to all elements of Ω_Λ. If*

$$\mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \leq \mu_2(\delta^x \mid \eta_{\Lambda\setminus x}), \tag{23}$$

for all $x \in \Lambda$ and boundary conditions $\xi \leq \eta$ with positive probability under μ_1 and μ_2, respectively, then

$$\mu_1 \leq \mu_2. \tag{24}$$

Note. This actually holds (with minor modifications) for spins taking values in any finite set $S \subseteq \mathbb{R}$.

The proof makes use of the *Gibbs sampler*, which is a Markov chain Monte Carlo algorithm for sampling from a distribution μ on Ω_Λ. The Gibbs sampler constructs a Markov chain σ_k with stationary distribution μ. If $\sigma_i = \omega$, then σ_{i+1} is sampled as follows:

1. Pick a site $x \in \Lambda$ uniformly at random;
2. Re-sample ω_x according to the conditional measure $\mu(\cdot \mid \omega_{\Lambda\setminus x})$.

The output is a configuration ω', which we set as the state of σ_{i+1}.

Since

$$\sum_{\omega_{\Lambda\setminus x}} \mu(\omega'_x \mid \omega_{\Lambda\setminus x})\mu(\omega_{\Lambda\setminus x}) = \mu(\omega'_x), \tag{25}$$

the measure μ is indeed stationary for this Markov chain.

Proof of Theorem 21. We must find a monotone coupling of μ_1 and μ_2. We do this by coupling their respective Markov chains σ and τ (generated by the Gibbs sampler) in such a way that, once $\sigma_i \leq \tau_i$, we will have $\sigma_j \leq \tau_j$ for all $j \geq i$. We note that the (random) time i at which this occurs must be finite as τ must eventually reach the all 1's state (since it is an irreducible Markov chain on a finite state space).

The coupled Markov chain (σ, τ) is defined as follows. Let U_k be an iid sequence of uniformly distributed random variables on $[0, 1]$ and suppose $(\sigma^i, \tau^i) = (\xi, \eta)$. As before, choose $x \in \Lambda$ at random, but now set

$$\sigma_x^{i+1} = \begin{cases} 1 & \text{if } \mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \geq U_i \\ -1 & \text{otherwise} \end{cases}, \tag{26}$$

$$\tau_x^{i+1} = \begin{cases} 1 & \text{if } \mu_2(\delta^x \mid \eta_{\Lambda\setminus x}) \geq U_i \\ -1 & \text{otherwise} \end{cases}. \tag{27}$$

Then

$$\Pr(\sigma_x^{i+1} = 1) = \Pr(\mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \geq U_i) = \mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \tag{28}$$

and similarly for τ. In other words, the individual Markov chains are being updated according to the conditional measures as before so this does define a coupling of σ and τ.

Moreover, this coupling achieves its goal. Indeed, suppose that $(\sigma^i, \tau^i) = (\xi, \eta)$ with $\xi \leq \eta$. Let x be the site at which we are re-sampling the spin in order to produce $(\sigma^{i+1}, \tau^{i+1})$. Then $\sigma_x^{i+1} = 1$ if $\mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \geq U_i$. But then by (23), $\mu_2(\delta^x \mid \eta_{\Lambda\setminus x}) \geq U_i$, which means we will have $\tau_x^{i+1} = 1$, which preserves $\sigma^i \leq \tau^i$. A similar argument shows that $\sigma_x^{i+1} = -1$ if $\tau_x^{i+1} = -1$.

Now (σ, τ) is an irreduible, aperiodic Markov chain on the finite state space $\Omega_\Lambda \times \Omega_\Lambda$, so it has a limiting distribution μ. Since μ_1 and μ_2 are stationary for σ and τ, respectively, μ is a coupling of μ_1 and μ_2. Moreover, since $\sigma^i \leq \tau^i$ for almost all i (almost surely), μ is a monotone coupling. □

Another important fact we will need is the FKG inequality, which follows from Holley's inequality.

Definition 22. A measure μ on Ω_Λ is *monotone* if (23) holds with $\mu_1 = \mu_2 = \mu$.

Example 23. The measure $\mu^{+,\Lambda}$ is monotone. To see this, begin by writing

$$\mu^{+,\Lambda}(\delta^x \mid \xi_{\Lambda\setminus x}) = \frac{\mu^{+,\Lambda}(\delta^x \xi_{\Lambda\setminus x})}{\mu^{+,\Lambda}(\xi_{\Lambda\setminus x})} = \left(1 + \frac{\mu^{+,\Lambda}((-\delta^x)\xi_{\Lambda\setminus x})}{\mu^{+,\Lambda}(\delta^x \xi_{\Lambda\setminus x})}\right)^{-1}. \tag{29}$$

We must check that the ratio on the right-hand side is a decreasing function of ξ.

This ratio is easily computed. For instance, for x not on the internal boundary of Λ, the Hamiltonians appearing in the numerator and denominator differ only by

$$\left(-\beta \sum_{y:y\sim x} \xi_y\right) - \left(-\beta \sum_{y:y\sim x} -\xi_y\right) = -2\beta \left(\sum_{y:y\sim x} \xi_y\right). \tag{30}$$

This is a decreasing function of ξ, as required.

A similar computation can be done for x on the internal boundary.

Theorem 24 (FKG (Fortuin-Kasteleyn-Ginibre) inequality)**.** *Suppose μ is a probability measure on Ω_Λ that assigns positive probability to all elements of Ω_Λ. If μ is monotone, then it has positive correlations.*

Note. The FKG inequality holds more generally for the Ising model with arbitrary *positive* couplings. That is, let $J = (J_{xy})_{x,y\in\mathbb{Z}^d}$ and define a modified Ising Hamiltonian by

$$H_{\beta,h,J}^{\eta,\Lambda}(\omega) = \beta \left(\sum_{\substack{x\sim y \\ x,y\in\Lambda}} J_{xy}\omega_x\omega_y + h\sum_{x\in\Lambda} \omega_x + \sum_{\substack{x\sim y \\ x\in\Lambda, y\in\partial\Lambda}} J_{xy}\omega_x\eta_y \right). \tag{31}$$

The Ising model with this Hamiltonian is said to be *ferromagnetic* if $J_{xy} \geq 0$ for all $x \sim y$. The FKG inequality can be proven for the general ferromagnetic Ising model.

Proof of Theorem 24. Let $\mu_1 = \mu$. For bounded increasing functions f, g, we need to show that $\mu_1(fg) \geq \mu_1(f)\mu_1(g)$. By decomposing g into its positive and negative parts, it suffices to consider the case of $g \geq 0$.

Thus, $\mathrm{d}\mu_2 = \frac{1}{Z} g \mathrm{d}\mu_1$ defines a probability measure if $Z = \mu_1(g)$. Since g is increasing,

$$\frac{\mu_2(\delta^x \mid \xi_{\Lambda\setminus x})}{\mu_2(-\delta^x \mid \xi_{\Lambda\setminus x})} = \frac{\mu_1(\delta^x \xi_{\Lambda\setminus x}) g(\delta^x \xi_{\Lambda\setminus x})}{\mu_1((-\delta^x)\xi_{\Lambda\setminus x}) g((-\delta^x)\xi_{\Lambda\setminus x})} \tag{32}$$

$$\geq \frac{\mu_1(\delta^x \xi_{\Lambda\setminus x})}{\mu_1((-\delta^x)\xi_{\Lambda\setminus x})} \tag{33}$$

$$= \frac{\mu_1(\delta^x \mid \xi_{\Lambda\setminus x})}{\mu_1(-\delta^x \mid \xi_{\Lambda\setminus x})}, \tag{34}$$

which implies that

$$\mu_2(\delta^x \mid \xi_{\Lambda\setminus x}) \geq \mu_1(\delta^x \mid \xi_{\Lambda\setminus x}). \tag{35}$$

By monotonicty, we get

$$\mu_1(\delta^x \mid \xi_{\Lambda\setminus x}) \leq \mu_2(\delta^x \mid \xi_{\Lambda\setminus x}) \leq \mu_2(\delta^x \mid \eta_{\Lambda\setminus x}). \tag{36}$$

This verifies (23) and Holley's inequality implies that $\mu_1 \leq \mu_2$.

Therefore,

$$\mu_1(fg) = Z\mu_2(f) \geq Z\mu_1(f) = \mu_1(f)\mu_1(g). \tag{37}$$

□

Proof of Lemma 14. As discussed, it suffices to prove (21). But by the Markov property and the FKG inequality (whose hypothesis was verified above), for any bounded increasing f,

$$\mu^{+,\Lambda_1}(f) = \mu^{+,\Lambda_2}(f \mid \omega_{\Lambda_2\setminus\Lambda_1} = 1) \tag{38}$$

$$= \frac{\mu^{+,\Lambda_2}(f \mathbb{1}_{\omega_{\Lambda_2\setminus\Lambda_1}=1})}{\mu^{+,\Lambda_2}(\mathbb{1}_{\omega_{\Lambda_2\setminus\Lambda_1}=1})} \tag{39}$$

$$\geq \mu^{+,\Lambda_2}(f). \tag{40}$$

□

5 Properties of $\mu^{\pm}$

First, let us complete the proof of Theorem 5.

The proof requires the following lemma.

Lemma 25.

1. *The measure $\mu^{\xi,\Lambda}$ has positive correlations for any ξ.*

2. *If $\xi \leq \eta$, then*

$$\mu^{\xi,\Lambda} \leq \mu^{\eta,\Lambda}. \tag{41}$$

3. *Suppose* $\mu^\omega := \lim_{\Lambda \uparrow \mathbb{Z}^d} \mu^{\omega,\Lambda}$ *exist for* $\omega = \xi, \eta$. *If* $\xi \leq \eta$, *then*

$$\mu^\xi \leq \mu^\eta. \tag{42}$$

Proof of theorem 5. We have already established (1).

For (2), we use the fact that the limit defining μ^+ is independent of the sequence of boxes Λ along which it is taken. Thus, if T is an automorphism of $\mathbb{Z}^d$, then (letting T also denote the induced action on Ω)

$$\mu^+(T^{-1}C) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \mu^{+,\Lambda}(T^{-1}C) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \mu^{+,T\Lambda}(C) = \mu^+(C) \tag{43}$$

for any cylinder set $C \subseteq \Omega$.

Lastly, the positive correlations property (3) is an immediate consequence of Lemma 25(1). $\square$

By (42), any Gibbs measure μ satisfies

$$\mu^- \leq \mu \leq \mu^+. \tag{44}$$

This identity has a few more implications.

Proposition 26. *The following are equivalent:*

1. *The infinite-volume Gibbs measure is unique.*
2. $\mu^-_{h,\beta} = \mu^+_{h,\beta}$.
3. $\mu^-_{h,\beta}(\delta^x) = \mu^+_{h,\beta}(\delta^x)$ *for some* $x \in \mathbb{Z}^d$.

Proof of Theorem 26. Clearly, (1) $\Rightarrow$ (2) $\Rightarrow$ (3).

Suppose (3) holds. By automorphism invariance this statement holds for *any* $x \in \mathbb{Z}^d$; it also holds with 1 replaced by -1. By (44) and Strassen's theorem, there exists a monotone coupling μ of μ^+ and μ^-. Let $(\sigma^1, \sigma^2) \sim \mu$. Then $\sigma^1_x = 1$ forces $\sigma^2_x = 1$ and, likewise, $\sigma^2_x = -1$ forces $\sigma^1_x = -1$ (both statements almost surely), so

$$\begin{aligned} \Pr(\sigma^1_x = \sigma^2_x) &= \Pr(\sigma^1_x = 1) + \Pr(\sigma^2_x = -1) && (45)\\ &= \Pr(\sigma^2_x = 1) + \Pr(\sigma^2_x = -1) && (46)\\ &= 1. && (47) \end{aligned}$$

Since this holds for any x, we get $\Pr(\sigma^1 = \sigma^2) = 1$, so (2) holds.

Lastly we note that (2) implies (1) by (44). $\square$

Finally, we observe that (44) also implies extremality of $\mu^\pm$.

Proposition 27. *Let* μ_1 *and* μ_2 *be Gibbs measures such that*

$$\mu^\pm = \lambda\mu_1 + (1-\lambda)\mu_2 \tag{48}$$

for $\lambda \in (0,1)$. *Then* $\mu_1 = \mu_2 = \mu^\pm$.

Proof. Suppose $\mu_1 \neq \mu^\pm$. Then by (44), $\mu_1 < \mu^\pm$ and $\mu_2 \leq \mu^\pm$. Thus,

$$\mu^\pm = \lambda\mu_1 + (1-\lambda)\mu_2 < \mu^\pm, \tag{49}$$

a contradiction. □

Note. The Aizenman-Higuchi theorem states that $\mu^\pm$ are the *only* extremal Gibbs measures for $d = 2$. However, this is false for $d > 2$. So while all Gibbs measures are "squeezed" between μ^- and μ^+ in the sense of stochastic dominance, this is not true in the sense of convexity.

We end with the proof of Lemma 25.

Proof of Lemma 25. Note that

$$H^{\xi,\Lambda}(\omega) - H^{\eta,\Lambda}(\omega) = \sum_{\substack{x\sim y \\ x\in\Lambda, y\notin\Lambda}} \omega_x(\eta_y - \xi_y), \tag{50}$$

so

$$\exp(-H^{\eta,\Lambda}(\omega)) = \exp(-H^{\xi,\Lambda}(\omega)) I^\Lambda_{\xi,\eta}(\omega) \tag{51}$$

with

$$I^\Lambda_{\xi,\eta}(\omega) = \exp\left(\sum_{\substack{x\sim y \\ x\in\Lambda, y\notin\Lambda}} \omega_x(\eta_y - \xi_y)\right). \tag{52}$$

Now observe that, since $\xi \leq \eta$, the function $I^\Lambda_{\xi,\eta}$ is increasing. Therefore, the FKG inequality implies that

$$\mu^{\eta,\Lambda}(f) = \frac{\sum_{\omega\in\Omega_\Lambda} f(\omega) e^{-H^{\xi,\Lambda}(\omega)} I^\Lambda_{\xi,\eta}(\omega)}{\sum_{\omega\in\Omega_\Lambda} e^{-H^{\xi,\Lambda}(\omega)} I^\Lambda_{\xi,\eta}(\omega)} = \frac{\mu^{\xi,\Lambda}(f I^\Lambda_{\xi,\eta})}{\mu^{\xi,\Lambda}(I^\Lambda_{\xi,\eta})} \geq \mu^{\xi,\Lambda}(f) \tag{53}$$

for any bounded increasing local f. □

References

[1] Sacha Friedli and Yvan Velenik. Equilibrium statistical mechanics of classical lattice systems: a concrete introduction.

[2] H. O. Georgii, O. Häggström, and C. Maes. The random geometry of equilibrium phases, 1999.

[3] Torgny Lindvall. *Lectures on the coupling method.* Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.