FERROMAGNETIC PHASE TRANSITIONS IN A STACKED LATTICE ISING MODEL

JOHN HOPPER

Abstract. In this paper we discuss some of the mathematics behind the Ising model, a physics model of magnet formation. After a brief introduction to the physics, this paper will focus on the existence of a ferromagnetic state given certain model parameters on the infinite lattice. This paper investigates an extension of this model where we consider a finite number of stacked lattices. We then observe how the stacked lattice model compares to the standard model on a lattice of one higher dimension. In particular, this paper is interested in how and where some of the techniques shown in [1] for the standard lattice extend to a stacked lattice.

Contents

1. Some Physics Background to the Ising Model and Paramagent-Ferromagnet Phase Transision

This problem was first introduced within the physics community and conventions are reminiscent of this origin. To prevent losing mathematicians to the somewhat cumbersome equations, subscripts, and names this section will provide a brief background to the problem thereby making sense of some of the conventions.

As a good chemist may inform you, the magnetic properties of certain metals and alloys are due to the existence of "free electrons" in the spin states of the atoms composing these metals. These electrons, located more or less equally spaced throughout a piece of metal, have one of two spin states: up or down (colloquially north and south). Basic magnetism tells us opposites attract, so we would imagine

Date: December 20, 2021.

that the most stable (lowest energy) state of a magnet would be that with alternating spins. Unfortunately, classical physics misses a key aspect of this story and, in reality, the state with the least potential energy occurs when adjacent electrons have the same spin[1].

In short, free electrons cause certain metals to be magnets. This begs the question: if magnets are made out of iron why are other iron products (like needles, hammers, and other tools) often not magnetized? It is precisely that: they are not magnetized. Any iron product is capable of being a magnet but it must be magnetized first (you can do this yourself: take a steel sewing needle and place a large magnet next to it, in a day or so you will have a small magnetic sewing needle). You may notice that there are metals with free electrons which are attracted to magnets, yet, no matter how long you put a magnet next to them, they remain unmagnetized (an excellent example is a refrigerator door, often made out of aluminum, spends all day and night with magnets attached but never attains a permanent magnetic field which is why you can put a fridge magnet on backwards). These are called paramagnets, and they are metals whose free electron interaction is too weak to overcome entropy and become a permanent magnet. In other words, the energy associated with a more random or chaotic system overcame the energy associated with magnetic interactions. This causes the paramagnet, in the absence of an external magnetic field, to lose its magnetization and revert to a more stable chaotic state. Paramagnets can only gain a stable magnetization in the presence of a magnetic field, thus paramagnets can be attracted to a nearby magnet, but paramagnets lose this magnetization as soon as the external magnet is far enough away.

One can imagine that the spins of free electrons are under a sort of 'tug-of-war' between the forces of entropy and magnetization. Since entropy increases with temperature, we can affect this 'tug-of-war' by changing the ambient temperature. An increase in temperature effectively dampens electron interactions by augmenting the entropy of the system. Thus, when heated to very high temperatures, metals that are typically ferromagnets (like iron) become paramagnets. Conversely, typical paramagnets (like aluminum) can be cooled down to low temperatures (and so low entropies) to produce ferromagnets. This temperature-dependent transition between paramagnets and ferromagnets has interested many physicists. The Ising model was developed to get a better understanding of this phase transition. Since electrons in a metal are typically regularly spaced, we may model their locations with an integer lattice. At every vertex of a subset of the integer lattice sits a free electron that is either spin-up or spin-down ($+1$ or -1). We then suppose that nearby electrons have interactions based on their relative spins. Furthermore, we suppose that electrons in non-adjacent sites interact very weakly and may be ignored because the quantum level interactions at play are extremely short range. With these rules in place, we may apply standard statistical physics formulas to factor in entropy and find the probabilities of given sets of configurations.

[1]Note, we are only dealing with energy from electricity and magnetism; entropy will make some of these states less stable. For those interested in why adjacent particles of the same spin have lower energy, I suggest reading physics text on this model. A brief explanation is that electrons must be anti-symmetric under exchange, thus they must be either anti-symmetric in their spin or spatial functions. If they are anti-symmetric in their spatial functions (and thus the same spin) the electrons will, on average, be farther apart (since anti-symmetric functions vanish at the origin) reducing the Coulomb interaction which is much stronger than the magnetic interaction.

The formulas are straightforward when we consider finite subsets of the lattice, but they require more sophisticated mathematics when we consider the infinite lattice. We will also see that the nature of the problem changes greatly with the number of dimensions we consider. Ising, whom the model is named after, rejected the model as trivial after observing a lack of ferromagnets in the 1 dimensional infinite lattice. He failed to realize that there is a ferromagnetic phase transition in higher dimensions, and the problem for the three dimensional lattice is hardly trivial since an exact mathematical solution is still unknown.

Since the dimension has such a strong effect on the outcome of the model, this paper will explore what will be referred to as "stacked lattices" or a finite number of infinite integer lattices stacked on top of each other where electrons interact with those adjacent in their own lattice as well as those 'just above' and 'just below'. The motivation for the stacked lattice is two-fold. First, in real life, not every compound produces a lattice that is infinitely long in every direction. A common example is the non-magnet graphite whose lattice structure is that of a small stack of layers which extend greatly in two directions. This type of structure is better modeled by a stack of 2 dimensional lattices than a 3 dimensional lattice. Second, the stacked lattice can be thought of as an intermediate step between two dimensions. Since the nature of this model changes greatly with the number of dimensions it may be helpful to investigate such an intermediate step. With this intuition, we begin our investigation of the magnetic phase transition of stacked lattices.

2. Notation and Definitions

We consider the space $\Omega = \mathcal{S}^{\mathbb{Z}^d}$ where $\mathcal{S}$ is the possible states at a point on the lattice, classically $\mathcal{S} = \{\pm 1\}$ indicating the spin of a free electron, but when we have a stacked lattice $\mathcal{S}$ can become more complicated. For an element $\sigma \in \Omega$, we represent the spin at a particular site of the lattice $I = (i_1, \ldots, i_d)$ with a subscript: $\sigma_I \in \mathcal{S}$. Finite subsets of the lattice are written as $\Lambda \subset \mathbb{Z}^d$, and we let $\mathscr{F}_\Lambda$ denote the σ-algebra generated by cylinder sets of the form $A = \{\sigma \in \Omega; I \in \Lambda, \sigma_I = s_I\}$ where $s_I \in \mathcal{S}$ for all $I \in \Lambda$. We then let $\mathscr{F}$ denote the Borel σ-algebra with respect to the product topology on Ω. We write energy functions based on the interactions of a particular subset Λ as $\Psi_\Lambda(\sigma)$ where Ψ_Λ is a $\mathscr{F}_\Lambda$ measurable function; if $\sigma, \tau \in \Omega$ agree on Λ, then $\Psi_\Lambda(\sigma) = \Psi_\Lambda(\tau)$. We define the shift map for θ_i^j, where if $\omega \in \Omega$, then $(\theta_i^j(\omega))_{(k_1,\ldots,k_j,\ldots,k_d))} = \omega_{(k_1,\ldots,k_j+i,\ldots,k_d))}$, we can think of this as moving the origin by i in the jth coordinate.

Definition 2.1 (Absolutely summable shift-invariant potential)**.** $\Psi = \{\Psi_\Lambda : \Lambda \subset \mathbb{Z}^d$ is finite, $\Psi_\Lambda : \Omega \to \mathbb{R}\}$ is an absolutely summable shift-invariant potential if for all $i \in \mathbb{Z}^d$ and finite $\Lambda \subset \mathbb{Z}^d$ the following hold:

(1) Ψ_Λ is $\mathscr{F}_\Lambda$-measurable
(2) $\Psi_{\Lambda+i} = \Psi_\Lambda \circ \theta_i$
(3) $||\Psi|| = \sum_{\Lambda, 0\in\Lambda, |\Lambda|<\infty} ||\Psi_\Lambda||_\infty < \infty$, with $||\Psi_\Lambda||_\infty = \sup_{\sigma\in\Omega} |\Psi_\Lambda(\sigma)|$

Recall that the physics behind this model suggests that only adjacent electrons should interact, thus only the relative position of electrons should affect the energy. Shift-invariant potentials can be thought of as this type of energy function; they only depend on the relative position of spins. In our case, we will want Ψ_Λ to be non-zero when Λ is the set of two adjacent particles (and so define their quantum interaction) or when Λ is a single particle (and so define the internal energy, or

interaction of a state with itself and with the ambient magnetic field h). Shift invariant potentials will be the building-blocks of many of our $\mathscr{F}_\Lambda$ measurable functions. In our work, we will want to move between different σ-algebras and the measurable functions on them, one method of doing so is with something called a Markov Kernel.

Definition 2.2 (Markov (or Stochastic) Kernel)**.** Given two measurable spaces $(\mathcal{X}, \mathscr{X})$ and $(\mathcal{Y}, \mathscr{Y})$, a Markov kernel from $(\mathcal{X}, \mathscr{X})$ to $(\mathcal{Y}, \mathscr{Y})$ is a map $\pi : \mathcal{X} \times \mathscr{Y} \to [0, 1]$ where

(1) For each $x \in \mathcal{X}$, $\pi(x, \cdot) : \mathscr{Y} \to [0, 1]$ is a probability measure.
(2) Given a set $Y \in \mathscr{Y}$, $\pi(\cdot, Y) : \mathcal{X} \to [0, 1]$ is an $\mathscr{X}$-measurable function.

It is not uncommon to write $\pi(x, \cdot)$ as $\pi_x(\cdot)$. We can view these as a generalization of transition probabilities applied to measure spaces. Like transition probabilities these kernels are particularly useful for changing the domains of measures and functions. In particular, we will use these kernels to relate potentials on one subset of the lattice to larger subsets. There are many Markov kernels between subsets of the lattice, but we want some sort of definition for a collection of them which work together nicely (we will end up calling these nice sets specifications). One such nice property will relate to composition, so we first need to define how to compose Markov kernels. Let f be a bounded real-valued $\mathscr{Y}$-measurable function, μ be a finite measure on $\mathscr{X}$, and π be a stochastic kernel from $(\mathcal{X}, \mathscr{X})$ to $(\mathcal{Y}, \mathscr{Y})$, then we can define the following compositions,

$$\pi f(x) = \int_{\mathcal{Y}} f(y)\pi(x, dy)$$
$$\mu\pi(A) = \int_{\mathcal{X}} \pi(x, A)\mu(dx).$$

If we think of a Markov kernel as both a bounded measurable function and a measure, then we can combine these two definitions to compose Markov kernels. Let π be a Markov kernel from $(\mathcal{X}, \mathscr{X})$ to $(\mathcal{Y}, \mathscr{Y})$ and ρ be a Markov kernel from $(\mathcal{Y}, \mathscr{Y})$ to $(\mathcal{Z}, \mathscr{Z})$, then we define their composition as a Markov kernel from $(\mathcal{X}, \mathscr{X})$ to $(\mathcal{Z}, \mathscr{Z})$ where,

$$\pi\rho(x, Z) = \int_{\mathcal{Y}} \pi(x, dy)\rho(y, Z).$$

With composition of Markov Kernels defined, we can define set Markov kernels which compose in a convenient way, which we will call specifications. One of the major conveniences that we desire is an analogue of the tower property of conditional expectation (the third property in our definition). In fact, most of the specifications we will work with can be viewed as conditional expectations because conditional expectations have these three properties 'baked in'.

Definition 2.3 (Specification)**.** A specification is a family of Markov Kernels $\Pi = \{\pi_\Lambda | \Lambda \subset \mathbb{Z}^d \text{ finite}\}$ where

(1) π_Λ is a Markov kernel from $(\Omega, \mathscr{F}_{\Lambda^c})$ to $(\Omega, \mathscr{F})$.
(2) π_Λ is $\mathscr{F}_{\Lambda^c}$-proper: if $B \in \mathscr{F}_{\Lambda^c}$, then $\pi_\Lambda(\sigma, B) = \mathbb{1}_B(\sigma)$ for all $\sigma \in \Omega$.
(3) If $\Delta \subset \Lambda \subset \mathbb{Z}^d$ are finite, then $\pi_\Lambda \pi_\Delta = \pi_\Lambda$

Remark 2.4. In our particular problem, we will want specifications which come from absolutely summable shift invariant potentials. These specifications are sometimes

called Gibbs specifications in other areas of literature. Since we are working toward a similarly named and more important definition, Gibbs measure, we will avoid calling these Gibbs specifications so as not to confuse the reader.

There is a fairly standard method in physics to create a specification from a potential using a Hamiltonian. Hamiltonian functions output the total energy of a state $\sigma \subset \mathcal{S}^\Lambda$ for a given boundary condition. By a boundary condition we mean a fixed $\omega \in \mathcal{S}^{\mathbb{Z}^d}$. Given a boundary condition we can extend a finite state of σ to all of $\mathbb{Z}^d$: $\sigma^\omega \in \Omega$ where for $a \in \Lambda$ and $b \in \Lambda^c$ we have that $(\sigma^\omega)_a = \sigma_a$ and $(\sigma^\omega)_b = \omega_b$. We can then write the Hamiltonian as a function $\mathcal{H}^\omega_\Lambda : \mathcal{S}^\Lambda \to \mathbb{R}$ where

$$\mathcal{H}^\omega_\Lambda(\sigma) = \sum_{\tau \subset \Lambda} \Psi_\tau(\sigma) + \sum_{\Lambda \cap \tau \neq \emptyset, \Lambda^c \cap \tau \neq \emptyset} \Psi_\tau(\sigma^\omega).$$

The first summation is the energy of the state and the second summation is the interaction energy between the state and the boundary. We can see that this is really the total contribution to the potential energy by the electrons in the finite state σ, so the concept of the Hamiltonian as an energy function is well founded. From this construction we can define a specification Π_β in terms of Ψ and the Hamiltonian it induces:

$$\pi_\Lambda(\omega, A) = \frac{\sum_{\sigma \in A} e^{-\beta \mathcal{H}^\omega_\Lambda(\sigma)}}{\sum_{\tau \in \mathcal{S}^\Lambda} e^{-\beta \mathcal{H}^\omega_\Lambda(\tau)}}.$$

In the realm of physics, β is the inverse temperature, for our purposes it acts as a variable that determines how uniform the individual measures are.

It is a simple application of the definitions to check that this is indeed a specification. At this point, we have mostly introduced notations and equations from physics. All these equations can be directly computed and have no phase transitions, but more interesting behavior happens in the thermodynamic limit. Thus, we look at the limit as the size of our finite lattice approaches infinity. It is this move from finite to infinite lattices where the more mathematical notation comes in.

As we transition to infinite lattices, the definitions based on ratios of energies will be dropped since the sums in the numerator and denominator will both diverge. Instead, we will consider measures and their limits. We want some sort of infinite measure that complies with physics on finite subsets. In effect, we want this measure to 'project down' to the regular physics equations (these physics equations are compiled in our specification). In probability, conditional expectations are considered projections, and so this is how we want the infinite measure to comply with our specification. We can see that the definition of Gibbs measure formalizes this understanding.

Definition 2.5 (Set of Gibbs Measures)**.** Given a specification Π, we define the set of Gibbs measures of this specification, $\mathscr{G}^\Pi$, to be the set of probability measures on $(\Omega, \mathscr{F})$, where for all bounded $\mathscr{F}$ measurable functions f and finite subsets $\Lambda \subset \mathbb{Z}^d$ we have that $\mathbb{E}^\mu\{f|\mathscr{F}_{\Lambda^c}\} = \pi_\Lambda f$ for $\mu \in \mathscr{G}^\Pi$.

We will focus our attention to Gibbs measures of specifications which are generated by our absolutely summable shift invariant potential Ψ which is derived from magnetic interactions. As such, from now on, we will assume that a Gibbs measure is with reference to this specification.

The set of Gibbs measures has physical significance because its size is directly linked to parramagnet to ferromagnet phase transitions. Consider that a ferromagnet may be given a permanent magnetic field, but also might not be magnetized (for example not every piece of iron you find is a magnet, but some are). We can imagine that if a piece of iron has zero average magnetic charge, then - on average - free electrons are going to be spin up exactly one half of the time. On the other hand, if a piece of iron is magnetized to have some positive average charge and if we blindly choose a free electron, then it is more likely to be spin up than spin down. In this light, we can see that the same piece of iron can have two distinct probability measures on the spin states. We can imagine that each Gibbs measure is such a probability measure. Thus, ferromagnets should have a set of Gibbs measures with more than one element[2]. On the other hand, we can see that this is not the case with parramagnets. Since every parramagnet cannot maintain a magnetic field, we can understand that any isolated parramagnet has zero overall magnetic charge. Therefore, the probability of any given free electron to be spin up is always one half. Thus, there is always only one probability measure on the states of the free electrons of a parramagnet, as such we can realize that the set of Gibbs measures of a parramagnet is a singleton set.

With this understanding, we now have a goal which is to show when certain stacked lattices and shift invariant potentials create a singleton set of Gibbs measures or create a larger set of Gibbs measures. You may protest, asking what if the set of Gibbs measures is empty, as it turns out the sets of shift invariant potentials we work with always admit at least one Gibbs measure as is shown in [1]. We will begin with the one dimensional case of stacked lattices which relies heavily on a lemma from linear algebra.

3. The One-Dimensional Regular and stacked Lattices

We will be able to reduce the one-dimensional Ising model to a problem in linear algebra. Luckily, the types of matrices that arise are in a fairly restrictive class: positive, real, symmetric matrices. This allows us to obtain a result with some fairly basic linear algebra facts.

3.1. The Eigenvalue Gap of Positive, Real, Symmetric Matrices.

Proposition 3.1. *Let M be a real symmetric $n \times n$ matrix, where $n \geq 2$, with entries $M_{i,j}$ such that, for all $1 \leq i, j \leq n$, $M_{i,j} > 0$. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of M. Then, for all $n \geq j > 1$, $\lambda_1 > |\lambda_j|$.*[3]

Let $e_1, \ldots, e_n$ denote the standard basis of $\mathbb{R}^n$, hence $Me_j = \sum_{i=1}^{n} M_{i,j} e_i$. By the spectral theorem, there exists an orthonormal eigenbasis $u_1, \ldots, u_n$ for M where $\lambda_i u_i = M u_i$. Let $u_1 = \sum_{i=1}^{n} x_i e_i$. Our method of proof uses that u_1 has all positive (or negative) entries, but we will first show something more general about eigenvectors of M with positive and negative entries. For this proof we will use the

[2]It is easy to check that the set of Gibbs measures is convex, so the existence of two distinct Gibbs measures is equivalent to the existence of uncountably many distinct Gibbs measures.

[3]Upon giving a presentation on my work, it was pointed out to me that this is the Peron-Forbenius theorem with an added supposition that our matrix M is symmetric. Symmetry appears naturally in our problem making the facts equally useful in this paper and I am biased toward my own proof believing it to be simpler and more intuitive than those of the full theorem.

following notation: if $v = \sum_{i=1}^{n} y_i e_i$, then $\overline{v} := \sum_{i=1}^{n} |y_i| e_i$. Note that this preserves the norm of v: i.e.,

$$||\overline{v}||^2 = \sum_{i=1}^{n} |y_i|^2 = \sum_{i=1}^{n} y_i^2 = ||v||^2.$$

Lemma 3.2. *Let M be as in Proposition 3.1. If u is a unit eigenvector of M with both positive and negative entries and λ is the corresponding eigenvalue, then $\overline{u}^T M \overline{u} > |\lambda|$.*

Proof. Write $u = \sum_{i=1}^{n} y_i e_i$. We know that if u has both positive and negative entries then for all $a_1, \dots a_n > 0$ that $|\sum_{i=1}^{n} a_i y_i| < \sum_{i=1}^{n} a_i |y_i|$.

Now consider the action of M on the vector $\overline{u}$: $M\overline{u} = \sum_{i=1}^{n} (\sum_{j=1}^{n} M_{i,j}|y_j|) e_i$, and since $M_{i,j} > 0$, we know that $\sum_{j=1}^{n} M_{i,j}|y_j| > |\sum_{j=1}^{n} M_{i,j} y_j|$. Pick $1 \leq k \leq n$ such that $|y_k| > 0$ (such a k exists since u is not the zero vector). Since $\sum_{j=1}^{n} M_{i,j}|y_j| > |\sum_{j=1}^{n} M_{i,j} y_j|$, we know that

$$\begin{aligned}
\overline{u}^T M \overline{u} &= \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} M_{i,j}|y_j| \Big) |y_i| \\
&= \sum_{i \neq k} \Big(\sum_{j=1}^{n} M_{i,j}|y_j| \Big) |y_i| + \Big(\sum_{j=1}^{n} M_{i,j}|y_j| \Big) |y_k| \\
&> \sum_{i \neq k} \Big(\sum_{j=1}^{n} M_{i,j}|y_j| \Big) |y_i| + \left| \Big(\sum_{j=1}^{n} M_{i,j} y_j \Big) y_k \right| \\
&\geq \sum_{i=1}^{n} \left| \Big(\sum_{j=1}^{n} M_{i,j} y_j \Big) y_i \right| \geq \left| \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} M_{i,j} y_j \Big) y_i \right| = |u^T M u|.
\end{aligned}$$

If we put these inequalities together we get that

$$\overline{u}^T M \overline{u} > |u^T M u| = |\lambda|.$$

□

Proof of Proposition 3.1. Recall Rayleigh's Principal:

$$\lambda_1 = \max_{||w||=1, w \in \mathbb{R}^n} w^T M w.$$

The entries of u_1 are either all non-negative or all non-positive, since otherwise Lemma 3.2 implies $\lambda_1 < \overline{u_1}^T M \overline{u_1} \leq \max_{||w||=1, w \in \mathbb{R}} w^T M w = \lambda_1$.

Without loss of generality, let the entries of u_1 be non-negative. We claim that they are all positive. Observe that the first eigenvalue is positive since, by Rayleigh's Principal, $\lambda_1 \geq e_1^T M e_1 = M_{1,1} > 0$. Then, letting $u_1 = \sum_{i=1}^{n} x_k e_k$, there exists some $1 \leq k \leq n$ where $x_k > 0$. Hence, for every $1 \leq i \leq n$,

$$x_i = \frac{\sum_{j=1}^{n} M_{i,j} x_j}{\lambda} = \frac{\sum_{j=1, j \neq k}^{n} M_{i,j} x_j}{\lambda} + \frac{M_{i,k} x_k}{\lambda} \geq \frac{M_{i,k} x_k}{\lambda} > 0.$$

Thus, u_1 has all positive entries. Now, let $i > 1$. Eigenvectors are orthonormal, $u_1 \cdot u_i = 0$, so u_i must have both positive and negative entries. Applying Rayleigh's Principal and Lemma 3.2, we conclude that

$$\lambda_1 \geq \overline{u_i}^T M \overline{u_i} > |\lambda_i|.$$

□

3.2. **Proof that the One-Dimensional Stacked Lattice Represents a Paramagnet.** Consider instead of a state space $\mathcal{S} = \{\pm 1\}$, we viewed an n-tuple of plus and minus ones which represent multiple lattices stacked on top of each other: $\mathcal{S} = \{\pm 1\}^n = \{(x_1, \ldots, x_n); x_i \in \{\pm 1\}\}$. We will still preserve the nearest neighbor interaction, as such the ith element of each n-tuple will interact with the $i+1$ and $i-1$ elements of the same tuple as well as with the ith elements of the adjacent tuples. Though this may initially seem complicated the only information about this interaction that will be important for our investigation is that only adjacent tuples will interact.

For two adjacent points of the lattice $a \sim b \in \mathbb{Z}^d$, with spins $\sigma_a = (\sigma_a^{(1)}, \ldots, \sigma_a^{(d)})$ and $\sigma_b = (\sigma_b^{(1)}, \ldots, \sigma_b^{(d)})$, the standard shift invariant potential of their interaction is $\Psi^{J,h}_{\{a,b\}}(\sigma) = \sum_{i=1}^{n} -J\sigma_a^{(i)}\sigma_b^{(i)}$, the internal potentials are given by $\Psi^{J,h}_{\{a\}}(\sigma) = \sum_{i=1}^{n}(-h\sigma_a^{(i)}) + \sum_{i=1}^{n-1}(-J\sigma_a^{(i)}\sigma_a^{(i+1)})$, and $\Psi^{J,h}_{\Lambda} = 0$ if $\Lambda = \emptyset$, $|\Lambda| > 2$, or $\lambda = \{a, b\}$ where $a \not\sim b$. Given a fixed boundary condition, $\omega \in \Omega$ we can then write the Hamiltonian as a function $\mathcal{H}^{\omega}_{\Lambda} : \mathcal{S}^{\Lambda} \to \mathbb{R}$ where

$$\mathcal{H}^{\omega}_{\Lambda}(\sigma) = \sum_{\tau \subset \Lambda} \Psi_{\tau}(\sigma) + \sum_{a \in \Lambda, b \in \partial\Lambda} \Psi_{\{a,b\}}(\sigma^{\omega}).$$

This can be an intimidating expression when written generically, for our investigation this is simply the nearest neighbor interactions within and across the boundary. Sometimes it is convenient to alter this definition somewhat and define the Hamiltonian as a function on Ω. First, consider the projection $\varphi_{\Lambda} : \Omega \to \mathcal{S}^{\Lambda}$, where for all $a \in \Lambda$ we have that $\varphi_{\Lambda}(\sigma)_a = \sigma_a$. We can then define $\overline{\mathcal{H}}^{\omega}_{\Lambda} : \Omega \to \mathbb{R}$ so that $\overline{\mathcal{H}}^{\omega}_{\Lambda}(\sigma) = \mathcal{H}^{\omega}_{\Lambda}(\varphi_{\Lambda}(\sigma))$. One reason this is nice, is that $\overline{\mathcal{H}}^{\omega}_{\Lambda}$ is a $\mathscr{F}_{\Lambda}$ measurable function.

When working with Gibbs measures on all of $\mathbb{Z}$, it is important to note what are our measurable sets. When working with the Borel σ-algebra, we often consider a basis of the topology which for the product topology is the collection of cylinder sets. Cylinder sets contain all the elements with certain spin states on a finite subset of the lattice. Formally, $C \subset 2^{\Omega}$ of the form $\{\sigma \in \Omega; \sigma_a = \omega_a \forall x \in \Lambda\}$ for a finite Λ and fixed $\omega \in \mathcal{S}^{\Lambda}$ (note, 2^{Ω} is the power set of Ω). These sets generate $\mathscr{F}$ and so a Gibbs measure μ is determined by the measure of cylinder sets.

By Tychonoff's theorem we know that the product of compact spaces is compact. $\mathcal{S}$ has only finitely many states it is clearly compact. Thus, we know that $\Omega = \mathcal{S}^{\mathbb{Z}}$ is compact. Compactness is used in [1] to show that set of Gibbs measures is non-empty. Our goal is to prove $\mathscr{G}$ is unique for one-dimensional stacked lattices. When it is unique, we say that the system represents a *paramagnet*. If there are multiple Gibbs measures, we say that the system represents a *ferromagnet*. We can understand the physics behind these definitions by noting that a paramagnet can hold no permanent magnetic field, so all stable (low energy) systems have no magnetic field and are the same under the "zoomed out" view of the infinite lattice. Whereas, we know that ferromagnets can hold permanent magnetic fields of various strengths which relate to the different measures on the lattice.

Theorem 3.3. *For all $\beta, J > 0$, $h \in \mathbb{R}$, and $n \in \mathbb{N}$, we have there is a unique Gibbs measure on $\Omega = (\{\pm 1\}^n)^{\mathbb{Z}}$.*

Proof. Let $\rho \in \mathscr{G}$ be a Gibbs measure, we wish to show that for all Borel measurable sets A that $\rho(A)$ has a unique value. We can note that cylinder sets generate the

Borel σ-algebra, so it suffices to show this for cylinder sets A. In fact, without loss of generality, we can further restrict A to be the cylinder set where $[-m, m]$ is given. Fix one such value of m and the set $A = \{\sigma \in \Omega; \sigma_i = \alpha_i \text{ for } |i| \leq m\}$, for some $\alpha \in \mathcal{S}^{[-m,m]}$. Our goal is to show that $\rho(A)$ is unique.

Consider the following collection of measurable functions $Y_k = \mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}_{[k,-k]^c}\}$, for $k < 0$. We can note that these form a reverse (or backwards) martingale with respect to the filtration $\mathscr{F}^k = \mathscr{F}_{[k,-k]^c}$ (for $k \leq -1$). To show this is a reverse martingale we need $\mathbb{E}^\rho\{Y_{k+1}|\mathscr{F}^k\} = Y_k$. Unpacking the definition of Y_k, we note that $\mathbb{E}^\rho\{Y_{k+1}|\mathscr{F}^k\} = \mathbb{E}^\rho\{\mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^{k+1}\}|\mathscr{F}^k\}$. Since $\mathscr{F}^k \subset \mathscr{F}^{k+1}$, we may use the tower property of conditional expectation to show that

$$\mathbb{E}^\rho\{\mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^{k+1}\}|\mathscr{F}^k\} = \mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^k\} = Y_k.$$

So, indeed, Y_k is a reverse martingale. By the reverse martingale convergence theorem, we know that the Y_k converges almost surely and is in L^1. What we will show next is that this limit is a constant independent of the boundary conditions.

Notice that Y_k is a random variable and the only elements of $[k, -k]^c$ that interact with $[k, -k]$ are those on the boundary. As such, we can view Y_k as a function on the boundary of $[k, -k]$ which is simply an element of $\mathcal{S}^2$ where the first entry is the state on the left boundary and the second entry is the sate on the right boundary. So, consider $\omega \in \mathcal{S}^2$ we want can think of this as the leftmost and rightmost edges of the boundary of $[k, -k]$. Formally, we can allow a transformation $T_k : \mathcal{S}^2 \to \mathcal{S}^{[k,-k]^c}$

$$T_k(\omega)_i = \begin{cases} \omega_1 & i > 0 \\ \omega_2 & i < 0 \end{cases}$$

to provide a formal way of recognizing the boundary term as only an element of $\mathcal{S}^2$. We can then consider $\overline{Y}_k(\omega) = Y_k(T_k(\omega))$. We will want to show that for every sequence of boundaries $\{{}_k\sigma \in \mathcal{S}^{[k,-k]^c}\}$, the limit $\lim_{k\to-\infty} Y_k(\sigma_k)$ is the same value. Equivalently we can show that regardless of the sequence $\{{}_k\omega \in \mathcal{S}^2\}$ the limit $\lim_{k\to-\infty} \overline{Y}_k({}_k\omega)$ remains the same value. We will do this by writing out the expression for $\overline{Y}_k$. Note that this expression is trivial when $|k| \leq m$ and does not affect the limit, so without loss of generality, we will consider when $|k| > m$. We will be summing on both sides of the interval $[-m, m]$, to simplify notation if $\pi \in \mathcal{S}^{[k,-m)}$, $\alpha \in \mathcal{S}^{[-m,m]}$ and $\rho \in \mathcal{S}^{(m,-k]}$, then we will let $\pi \times \alpha \times \rho \in \mathcal{S}^{[k,-k]}$ where

$$(\pi \times \alpha \times \rho)_i = \begin{cases} \pi_i & i \in [k, -m) \\ \alpha_i & i \in [-m, m] \\ \rho_i & i \in (m, -k] \end{cases}.$$

so we will In this case we can write

$$\overline{Y}_k(\omega) = \frac{\sum_{\pi\in\mathcal{S}^{[k,-m]}} \sum_{\rho\in\mathcal{S}^{[m,-k]}} e^{-\beta\mathcal{H}^{T_k(\omega)}_{[k,-k]}(\pi\times\alpha\times\rho)}}{\sum_{\tau\in\mathcal{S}^{[k,-k]}} e^{-\beta\mathcal{H}^{T_k(\omega)}_{k,-k]}(\tau)}}.$$

This formula, as is, does not appear very illuminating, but there is a trick to deal with sums of these exponentials using matrices. Consider the component of the quantum interaction between two adjacent points on the lattice, $Q : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ where

$$Q((a_i)_{i=1}^n, (b_i)_{i=1}^n) = \exp\Big(-\beta(\sum_{i=1}^n -Ja_ib_i + \sum_{i=1}^{n-1} -J/2(a_ia_{i+1}+b_ib_{i+1}) + \sum_{i=1}^n -h/2(a_i+b_i)\Big).$$

With this notation we can write $e^{-\beta \mathcal{H}^{\omega}_{[-k,k]}(\sigma)}$ as a product,

$$e^{-\beta \mathcal{H}^{\omega}_{[k,-k]}(\sigma)} = e^{\frac{\beta}{2}(\Phi_{\{k-1\}}(\omega)+\Phi_{\{1-k\}}(\omega))} \prod_{i=k-1}^{-k} Q(\sigma^{\omega}_i, \sigma^{\omega}_{i+1}).$$

The exponential term out front simply cancels the extra internal energies that are included in the product. Since this is just a constant factor and is only dependent on the boundaries it will float around in our computation for a bit before canceling out. Until then, for the sake of more concise notation, let $Z^{\omega} = e^{\beta/2(\Phi_{\{k-1\}}(\omega)+\Phi_{\{1-k\}}(\omega))}$.

Note that Q is symmetric in its indices, $Q(x,y) = Q(y,x)$. Furthermore, that $Q(x,y)$ outputs the exponential of a real number, so the image of Q is the (strictly) positive reals. We can use this notation to create a $2^n \times 2^n$ matrix of the possible outputs of $Q(a,b)$. For ease of notation, create a bijection $f[1,2^n] \to \mathcal{S}$ (you may think of this a binary notation, but the exact bijection is irrelevant), then we can define a matrix $\mathbf{Q} = \Big(Q(f(i),f(j))\Big)_{i,j}$. The beauty of this notation, is that for $i < j \in \mathbb{Z}$

$$\sum_{\sigma \in \mathcal{S}^{[i,j]}} e^{-\beta \mathcal{H}^{\omega}_{[i,j]}(\sigma)} = Z^{\omega} \sum_{\sigma \in \mathcal{S}^{[i,j]}} \prod_{l=i-1}^{j} Q(\sigma^{\omega}_l, \sigma^{\omega}_{l+1}) = Z^{\omega} (\mathbf{Q}^{j-i+2})_{f^{-1}(\omega_{i-1}), f^{-1}(\omega_{j+1})}.$$

We have now pushed through the more cumbersome notation of sums of Hamiltonians to get the power of a matrix $\mathbf{Q}$. Observe the nature of this matrix, $\mathbf{Q}$. The function Q is symmetric in its indices, thus we know that the matrix $\mathbf{Q}$ is symmetric. Next, since the function Q has its image on the positive reals, $\mathbf{Q}$ has positive real entries. Proposition 3.1 proves that $\mathbf{Q}$ has an eigenvalue gap. Order the eigenvalues of $\mathbf{Q}$, as $\lambda = \lambda_1 \geq \lambda_2 \cdots \geq \lambda_{2^n}$. We will want to consider $\mathbf{Q}/\lambda$ and $(\mathbf{Q}/\lambda)^i$, the latter has eigenvalues of 1 and $(\frac{\lambda_j}{\lambda})^i$ the second type is shirking geometrically with respect to i.

With this analysis done, we can start computing the limits and finding $\lim_{k\to-\infty} \overline{Y}_k$. We will start with returning to the earlier sum and expressing it with matrices:

$$\overline{Y}_k(\omega) = \frac{\sum_{\pi \in \mathcal{S}^{[k,-m]}} \sum_{\rho \in \mathcal{S}^{[m,-k]}} e^{-\beta \mathcal{H}^{\omega}_{[k,-k]}(\pi,\alpha,\rho)}}{\sum_{\tau \in \mathcal{S}^{[k,-k]}} e^{-\beta \mathcal{H}(\tau)}}$$

$$= \frac{Z^{\omega} \mathbf{Q}^{-k-m+2}_{f^{-1}(\omega_{k-1}), f^{-1}(\alpha_{-m})} \prod_{i=-m}^{m-1} \mathbf{Q}(\alpha_i, \alpha_{i+1}) \mathbf{Q}^{-k-m+2}_{f^{-1}(\alpha_m), f^{-1}(\omega_{1-k})}}{Z^{\omega} \mathbf{Q}^{2-2k}_{\omega_{k-1}, \omega_{1-k}}}$$

Notice that the factor of Z^{ω} cancels on top and bottom meaning the only boundary terms left are in the entry of the matrix $\mathbf{Q}$. We will now show this is small by replacing $\mathbf{Q}$ with $\mathbf{Q}/\lambda$. To keep the notation concise, let $C = \prod_{i=-m}^{m-1} \mathbf{Q}(\alpha_i, \alpha_{i+1})$ since we will be treating it as a constant from here on. Also, let $\{u_i\}_{i=1}^{2^n}$ be the unit eigenvector basis of $\mathbf{Q}$ with corresponding eigenvalue λ for u_1. Let $\{e_i\}$ be the standard basis of $\mathbf{Q}$, so $e_j^T \mathbf{Q} e_i = \mathbf{Q}_{i,j}$. With this done, the previous fraction can be rewritten as

$$= C\lambda^{-2m} \frac{(e^T_{f^{-1}(\omega_{k-1})} \mathbf{Q}^{2-k-m} e_{f^{-1}(\alpha_{-m})})(e^T_{f^{-1}(\alpha_m)} \mathbf{Q}^{2-k-m} e_{f^{-1}(\omega_{1-k})})}{e^T_{f^{-1}(\omega_{k-1})} \mathbf{Q}^{2-2k} e_{f^{-1}(\omega_{1-k})}}$$

$$= C\lambda^{-2m} \frac{(e_{f^{-1}(\omega_{k-1})} \cdot u_1)(e_{f^{-1}(\omega_{1-k})} \cdot u_1)(e_{f^{-1}(\alpha_{-m})} \cdot u_1)(e_{f^{-1}(\alpha_m)} \cdot u_1) + o(1)}{(e_{f^{-1}(\omega_{k-1})} \cdot u_1)(e_{f^{-1}(\omega_{1-k})} \cdot u_1) + o(1)}$$

Where in the numerator the

$$o(1) = \sum_{i=2}^{2^n} (\lambda_i/\lambda)^{2-k-m} (e_{f^{-1}(\omega_{k-1})} \cdot u_i)(e_{f^{-1}(\omega_{1-k})} \cdot u_i)(e_{f^{-1}(\alpha_{-m})} \cdot u_i)(e_{f^{-1}(\alpha_m)} \cdot u_i)$$

and in the denominator

$$o(1) = \sum_{i=2}^{2^n} (\lambda_i/\lambda)^{2-2k} (e_{f^{-1}(\omega_{k-1})} \cdot u_i)(e_{f^{-1}(\omega_{1-k})} \cdot u_i)$$

both of which are bounded by a constant times $\left(\frac{\max(|\lambda_2|,|\lambda_{2^n}|)}{\lambda}\right)^{-k}$ which goes to zero as k approaches negative infinity. Thus, we can see that

$$\lim_{k\to-\infty} \overline{Y}_n({}_k\omega) = C\lambda^{-2m}(e_{f^{-1}(\alpha_{-m})} \cdot u_1)(e_{f^{-1}(\alpha_m)} \cdot u_1) := C_{\beta,h}(A)$$

which is a constant that only depends on h, β and the set A, importantly it is independent of the sequence of boundaries $\{{}_k\omega\}$.

We have now shown that the limit of the inverse martingale Y_k is depends only on A, h, and β, but we next need to show that this shows that the $\rho(A)$ is likewise only dependent on A, h and β. Using dominated convergence, we can use the following string of equivalences,

$$\begin{aligned}\rho(A) = \mathbb{E}^\rho\{\mathbb{1}_A\} = \mathbb{E}^\rho\{\mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^k\}\} = \lim_{k\to-\infty} \mathbb{E}^\rho\{\mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^k\}\} \\ = \mathbb{E}^\rho\{\lim_{k\to-\infty} \mathbb{E}^\rho\{\mathbb{1}_A|\mathscr{F}^k\}\} = \mathbb{E}\{C_{\omega,h}(A)\} = C_{\omega,h}(A).\end{aligned}$$

We have shown that $\rho(A)$ is dependent on β, h and A. Thus, for a fixed β, h, and A there is a unique value of $\rho(A)$, and so we have shown that a fixed value of β and h that ρ is unique, so $\mathscr{G}$ is a singleton set. □

This proof is an extension of that described in [1]. The move to stacked lattices has little effect on the method of proof and requires only a more general linear algebra argument. Unfortunately, the proof method for the 2 dimensional lattice will not extend as nicely to the stacked lattice and will require some non-trivial reframing of the problem. This is not too surprising when we realize that the point of phase transition is still open for the three dimensional lattice. With the current method of stacking we can, in the limit, approach the lattice of one higher dimension but only with more limited boundary constraints. A cylinder set in a $2m+1$ stacked lattices where $([-m,m]^n)^c$ is given is similar to the cylinder set of the higher dimension where $([-m,m]^{n+1})^c$ is given, but it is neutral for the boundary terms in the $n+1$ direction. If there was a phase transition for all stacked lattices of dimension n for a β and h, then there would also be a phase transition in the lattice of one higher dimension. This is because reducing boundary constraints can only increase the number of possible Gibbs measures (i.e., a positive boundary in all $n+1$ directions should have a greater affinity toward positive spins than a boundary where n dimensions are positive but the $n+1$st is neutral). In this way, if a stacked 2 dimensional lattice represents a ferromagnet for a β and h then so should the 3 dimension lattice for the same β and h. Thus, the stacked lattice can find a lower bound on the critical β for the higher dimensional lattice[4]. It is

[4]We can imagine that this lower bound improves as we increase the number of stacks, possibly to the true value

therefore not surprising that as we move to stacked two dimensional lattices, the β of a phase transition inherits some of the difficulty of the 3 dimensional phase transition problem.

4. Two-Dimensional Phase Transition in a Standard Lattice

We will start our analysis of the two dimensional stacked lattice with the standard two dimensional lattice. We will see that the standard techniques do not generalize as well as they did in the one dimensional model, and we will reconcile the techniques by reworking the problem.

We will use the idea that there is a symmetry between positive and negative spins. The equations are identical after flipping every spin of a lattice (including the boundary terms). This is inherit in the shift invariant potential which, when $h = 0$, is only dependent on relative spins $\sigma_i\sigma_j$. We will use this symmetry to analyze our space.

We will need this new technique since the trick used in the one dimensional case proved the uniqueness of a Gibbs measure. Whereas we are now proving the existence of multiple Gibbs measures. Our method will be to produce two Gibbs measures and to show that they are distinct for large enough β by showing that they given two different probabilities to the same event.

Theorem 4.1. *For all $\beta \geq \frac{\ln(3)+1}{2J}$ the single lattice 2 dimensional Ising model with no external magnetic field, $h = 0$, represents a Ferromagnet.*

Remark 4.2. Notice that the β for which the model represents a ferromagnet depends on J, this comes from physics where the value of J is fixed by the materials of the magnet and the value of β can be changed[5].

We begin our proof by constructing 'positive' and 'negative' Gibbs measures, ρ and ν. Our goal is to show they are distinct for large enough β. Consider that we can define a Gibbs measures as the weak limit of the measures on finite subsets of the lattice. Let $P \in \Omega$ where $P_i = +1$ for all i, and $Q \in \Omega$ where $Q_i = -1$ for all i. We can then define for any cylinder set $A \subset \mathscr{F}$ that $\rho(A) = \lim_{n\to\infty} \pi_{B_n}(P, A)$ and $\nu = \lim_{n\to\infty} \pi_{B_n}(Q, A)$ where $B_n = \{(i, j) \in \mathbb{Z}^2; |i| + |j| \leq n\}$. We can think of these as the limit of positive boundary conditions and negative boundary conditions. We can imagine that these are the two extreme Gibbs measures and are as positive preferring and negative preferring as possible.

Our goal is to show that $\rho \neq \nu$. Consider the event $A \in \mathscr{F}$ where $A = \{\sigma \in \Omega; \sigma_{(0,0)} = 1\}$. We will show that $\rho(A) \neq \nu(A)$. We can imagine that ρ is positive preferring and want to show that $\rho(A) > \frac{1}{2}$ by bounding $\rho(A^c)$. We will follow [1] which uses a trick with contours. Contours are objects in the dual lattice. The dual lattice is simply the dual graph of the standard lattice; the edges of the original lattice are the vertices of the dual, we connect two vertices of the dual if they represent edges that are incident on the same vertex. Readers looking for a more precise work can look into any book on graph theory for the concepts of a dual graph. With this definition we can define a contour.

[5]Some mathematics texts (such as [1]) bundle these two values together caring more about the different behaviors than the physical interpretation.

Definition 4.3 (contour)**.** For a finite subset of the lattice $\Lambda \subset \mathbb{Z}^2$ with boundary ω, a closed simple walk, γ, on the dual lattice is a contour in $\sigma \in \{\pm 1\}^\Lambda$ if every vertex γ_i represents an edge between two points a_i, b_i where $\sigma^\omega_{a_i}\sigma^\omega_{b_i} = -1$.

Colloquially we can view this as a path that separates spin up locations and spin down locations. Notice that the contour can include edges between a point in Λ and a point on the boundary. Since this is a closed path in 2 dimensions there is a good notion of inside and outside the contour (the finite subset of the plane and the infinite subset). We can make this notion rigorous by saying a vertex is *enclosed* by the contour if every infinite simple walk starting at a vertex contains two points which are connected by an edge which is an element of the contour.

Now note that for $n \geq 1$ that for $\sigma \in \{\pm 1\}^{B_n}$ where $\sigma_{(0,0)} = -1$ then there exists a contour γ on σ^P enclosing $(0,0)$ (here it is crucial that we allow elements of the contour to represent edges between a subset and its boundary). The benefit of this is that we can count the number of contours instead of the number of states (this will over-count somewhat since a single state can have multiple contours containing the origin). We want to have an upper-bound on the number of contours so as to have an upper-bound on the number of states where the origin has the opposite spin as the boundary. For simplicity of notation let Γ_k denote the set of closed simple walks of length k on the dual two-dimensional lattice which enclose the origin and $\Gamma = \cup_{k\in\mathbb{N}}\Gamma_k$ be the set of closed simple walks enclosing the origin. The following lemma will put a bound on the size of these sets.

Lemma 4.4. *For all $k \geq 1$, we have that $|\Gamma_k| \leq 4k \cdot 3^{k-2}$*

Proof. Since we know that for every $\gamma \in \Gamma_k$ that one of the edges $\{(0,i),(0,i+1)\}$ for $i \geq 0$ must be in γ since γ encloses the origin. Furthermore we know that since γ must also contain one of the edges $\{(0,i),(0,i-1)\}$ for $i \leq 0$ and must be a connected path, it cannot contain a point more than a distance k from these vertices, thus one of $\{(0,i),(0,i+1)\}$ for $i \in [0,k-1]$ must be in γ. We can say that there are then k choices for one of the vertices of $\gamma \in \Gamma_k$.

Next, if we can note that there are only 4 adjacent vertices to this initial point, and since the path is simple after the second point there are at most 3 adjacent vertices which can be a valid next step in the simple walk. We can put these numbers together to see that assuming $k > 2$ that $|\Gamma_k| \leq 4k \cdot 3^{k-2}$. We can also note that there is no closed walk for $k \leq 2$, so $|\Gamma_k| = 0 < 4k \cdot 3^{k-2}$ for $k = 1, 2$. Thus, the bound holds for all $k \geq 1$.

□

Our proof will revolve around the nature of spin flips. Given a subset of our usual finite subset of the lattice $\Delta \subset \Lambda \subseteq \mathbb{Z}^2$, we can define a spin flip transformation $F^\Lambda_\Delta : \{\pm 1\}^\Lambda \to \{\pm 1\}^\Lambda$ which flips the spin in Δ;

$$F^\Lambda_\Delta(\sigma)_i = \begin{cases} -\sigma_i & i \in \Delta \\ \sigma_i & i \notin \Delta. \end{cases}$$

This is clearly a one-to-one transformation on $\mathcal{S}^\Lambda$. For a contour γ will consider subsets $\Delta_\gamma = \{i \in \Lambda; i \text{ is enclosed by } \gamma\}$ and the transformations $F^\Lambda_{\Delta_\gamma}$ to compare the likely-hood of having a negative spin at the origin to having a positive spin and so show that $\rho(A) > 1/2$. Since our only definition of $\rho(A)$ must deal with projections down to finite subsets let ${}^nA := \{\sigma \in \mathcal{S}^{B_n}; \sigma_{(0,0)} = 1\}$ which is the

projection of A onto $\mathcal{S}^{B_n}$.With these definitions we are ready to preform the major computation for Theorem 4.1.

Lemma 4.5. *For $\beta \geq \frac{\ln(3)+1}{2J}$, we have that $\rho(A) > 1/2$.*

Proof. We will work in reverse to show that $\rho(A^c) < 1/2$ since ρ is a probability measure this will imply desired statement.

We start by simply unpacking our notation, remember for this problem $h = 0$ and we will be using the positive boundary condition P.

$$\rho(A^c) = \lim_{n\to\infty} \pi_{B_n}(P, A^c) = \lim_{n\to\infty} \frac{\sum_{({}^nA)^c} e^{-\beta \mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}}$$

We want to more closely look at this summation in the numerator. Consider that for every $\sigma \in ({}^nA)^c$ we know that there is a contour, γ, enclosing the origin. Fix a closed walk γ on the dual lattice consider the subset of $({}^nA)^c$ where γ is a contour for the positive boundary condition: $({}^nA)^c_\gamma := \{\sigma \in ({}^nA)^c; \gamma \text{ is a contour in } \sigma^P\}$. Since every $\sigma \in ({}^nA)^c$ has a contour, we can say that $({}^nA)^c = \cup_{\gamma\in\Gamma}({}^nA)^c_\gamma$, note that these sets are not necessarily disjoint (their overlap is small enough to allow the bound we desire). With this notation we can bound the sums in the limit,

$$\rho(A^c) = \lim_{n\to\infty} \frac{\sum_{\sigma\in({}^nA)^c} e^{-\beta \mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}} \leq \lim_{n\to\infty} \frac{\sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} \sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta \mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}}.$$

Now we are ready to apply the spin flip, fix a contour of length k, γ, and $\sigma \in ({}^nA)^c_\gamma$. We want to compare $\mathcal{H}^P_{B_n}(\sigma)$ with $\mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}(\sigma))$. We can unpack the Hamiltonian into internal energies, interactions not along an edge of γ, and interactions along an edge in γ. With $h = 0$ the internal energies are zero so the spin flip does not change the value. The interactions are always based on $\sigma_i\sigma_j$, so if $i, j \in \Delta_\gamma$ then $F^{B_n}_{\Delta_\gamma}(\sigma)_i F^{B_n}_{\Delta_\gamma}(\sigma)_j = (-\sigma_i)(-\sigma_j) = \sigma_i\sigma_j$, if $i, j \notin \Delta_\gamma$ then it is simply $F^{B_n}_{\Delta_\gamma}(\sigma)_i F^{B_n}_{\Delta_\gamma}(\sigma)_j = \sigma_i\sigma_j$. Now, the third case were the edge connecting i and j is an element of γ so $i \in \Delta_\gamma$ and $j \notin \Delta_\gamma$, then $F^{B_n}_{\Delta_\gamma}(\sigma)_i F^{B_n}_{\Delta_\gamma}(\sigma)_j = -(\sigma_i\sigma_j) = -(-1) = 1 = \sigma_i\sigma_j + 2$. Now, we know that if a contour is of length k then there are exactly k such interactions which change and the rest are unaffected by the spin flip: $\mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}(\sigma)) = \mathcal{H}^P_{B_n}(\sigma) + 2Jk$. We can use this to bound $\rho(A^c)$,

$$\begin{aligned}
\rho(A^c) &\leq \lim_{n\to\infty} \frac{\sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} \sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta \mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}} \\
&= \lim_{n\to\infty} \frac{\sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} \sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta \mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}(\sigma)) - 2\beta kJ}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}} \\
&= \lim_{n\to\infty} \sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} e^{-2\beta kJ} \frac{\sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta \mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}(\sigma))}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta \mathcal{H}^P_{B_n}(\tau)}}.
\end{aligned}$$

Recall that $F^{B_n}_{\Delta_\gamma}$ is a bijection from $\{\pm 1\}^{B_n}$ to itself, so we know that the multi-set $F^{B_n}_{\Delta_\gamma}(({}^nA)^c_\gamma)$ is a regular set $\{F^{B_n}_{\Delta_\gamma}(\sigma); \sigma \in ({}^nA)^c_\gamma\} \subset \mathcal{S}^{B_n}$[6]. This implies that

[6]It is important to track where our over-counting is. We can note that sets defined by separate contours may have overlap, but the set defined by a single contour has each state at most once.

the fraction $\frac{\sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta\mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}(\sigma))}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta\mathcal{H}^P_{B_n}(\tau)}}$ is less than or equal to 1. Using this, we can continue our computation that

$$\rho(A^c) \leq \lim_{n\to\infty} \sum_{k=1}^{\infty} \sum_{\gamma\in\Gamma_k} e^{-2\beta kJ}(1)$$

and by Lemma 4.4 this is less than or equal to

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} 4k3^{k-2}e^{-2\beta kJ} = \sum_{k=1}^{\infty} 4k3^{k-2}e^{-2\beta kJ}.$$

Finally, we may apply the supposition that $\beta \geq \frac{\ln(3)+1}{2J}$,

$$\rho(A^c) \leq \sum_{k=1}^{\infty} 4k3^{k-2}e^{-k(\ln(3)+1)} = \frac{4}{9}\sum_{k=1}^{\infty} ke^{k(\ln(3)-\ln(3)-1)} = \frac{4}{9}\sum_{k=1}^{\infty} ke^{-k} = \frac{4}{9}\frac{e}{(1-e)^2} < \frac{1}{2}$$

□

Note that this bound on the critical β is by no means tight, the real critical temperature is $\beta_c = \frac{\ln(1+\sqrt{2})}{2J}$ [1]. Yet, in this paper, we are more interested in showing that there is interesting behavior over precisely determining the conditions for this behavior. Speaking of which, we still need to finish off the theorem.

Proof. We know that $\rho(A) > \frac{1}{2}$ by Lemma 4.5. Furthermore, we know that if $h = 0$ there is a global spin symmetry: the Hamiltonian only depends on relative spins. As such, we know that for any finite subset, $\Lambda \subset \mathbb{Z}^2$, and spin state on said subset, $\sigma \in \mathcal{S}^\Lambda$ we know that $\mathcal{H}^P_\Lambda(\sigma) = \mathcal{H}^Q_\Lambda(F^\Lambda_\Lambda\sigma)$. Furthermore, it is true that $F^{B_n}_{B_n}({}^nA) = ({}^nA)^c$ and $F^{B_n}_{B_n}(\mathcal{S}^{B_n}) = \mathcal{S}^{B_n}$. Thus, we know that

$$\nu(A) = \lim_{n\to\infty} \frac{\sum_{\sigma\in{}^nA}\mathcal{H}^Q_{B_n}(\sigma)}{\sum_{\tau\in\mathcal{S}^{B_n}}\mathcal{H}^Q_{B_n}(\tau)} = \lim_{n\to\infty} \frac{\sum_{\sigma\in({}^nA)^c}\mathcal{H}^P_{B_n}(\sigma)}{\sum_{\tau\in\mathcal{S}^{B_n}}\mathcal{H}^P_{B_n}(\tau)}$$

$$= \rho(A^c) < \frac{1}{2} < \rho(A)$$

Thus, we know that $\nu(A) < \rho(A)$, so these are two distinct Gibbs measures and so $\mathscr{G}$ is not a singleton set, thus by definition the Ising model represents a Ferromagnet. □

5. Extension to the Two-Dimensional Stacked Lattice

Unfortunately, the method shown in section 4 does not translate well to the stacked Lattice. We can see this by understanding that the concept behind the spin flip no longer holds. The first hurtle comes with defining a contour; the same trick would require a higher dimensional analogue of a contour. One option would be to invent a new object to fill this role in higher dimensions. If, instead, we wish to continue to use one-dimensional contours we can only look at them on one layer of the lattice. This brings its own issues since it is not clear which spins to flip given a contour. There are two a priori choices: flip only the spins within the layer

The over-counting happens in the summation outside of the fraction, not the summation within the fraction.

of the lattice which has the contour or flip these spins as well as flip the spins of the layers above and below them.

The first option of only flipping the spins in a single layer of the lattice has the issue that there is no consistent reduction in energy as there was before. For the standard lattice, the energy after the flip dropped by $e^{-2\beta Jk}$, here there could also be increases in energy from the flip coming from the vertical interactions. Notice that the number of vertical interactions grows with the area within the contour which can grow at a rate much faster than the length of the contour. Regardless of the relative strength between lattices, so long as it is non-zero, the vertical interaction will dominate in the limit and make for inconsistent changes in energy after a spin flip.

The second option is to flip inside the contour as well as all the sites above and below the contour. This eliminates the issues caused by vertical interactions, but creates a new issue. The interactions on other layers of the lattice could cause increases in energy after the spin flip. Luckily, the size of this increase is bounded by the length of the contour. If the lattice with the contour has stronger interactions than the lattices above or below it, then this method can work. Physically, we can imagine this as two layers of different materials with different J constants such as iron and aluminum.

Suppose that now the spin states $\mathcal{S} = \{\pm 1\}^2$, were the interaction strength on the 'top' layer (or the first term) is J and the interaction strength between sites on the bottom layer is J'. Finally, the vertical interaction strength will be V, but in our more limited technique, the value of V is irrelevant. Formally, we can define the absolutely summable shift invariant potential $\Phi^{J,J',V,h}_{\{a,b\}}(\sigma) = -J\sigma_a^{(1)}\sigma_b^{(1)} - J\sigma_a^{(2)}\sigma_b^{(2)}$ for $a \sim b$, $\Phi^{J,J',V,h}_{\{a\}}(\sigma) = -V\sigma_a^{(1)}\sigma_a^{(2)} - h(\sigma_a^{(1)} + \sigma_a^{(2)})$, and for $|\Lambda| \neq 1, 2$ or $\Lambda = \{a, b\}$ for $a \not\sim b$ $\Phi^{J,J',V,h}_{\Lambda}(\sigma) = 0$. Notice that for $h = 0$ that the internal energy at any location a only depends on the relative spins $\sigma_a^{(1)}\sigma_a^{(2)}$. We can also extend are definition of spin flip to the stacked lattice, for $\Delta \subseteq \Lambda \subset \mathbb{Z}^2$ define $F^\Lambda_\Delta : \mathcal{S}^\Lambda \to \mathcal{S}^\Lambda$ such that for $j \in \{1, 2\}$,

$$\left(F^\Lambda_\Delta(\sigma)\right)^{(j)}_i = \begin{cases} -\sigma_i^{(j)} & i \in \Delta \\ \sigma_i^{(j)} & i \notin \Delta \end{cases}.$$

As with the single lattice, notice that F^Λ_Δ is a bijective map from $\mathcal{S}^\Lambda$ to itself. We can define contours on the lattice with the expanded state-space by simply ignoring the 'bottom layer', in other words a closed simple walk, γ, on the dual lattice of $\mathbb{Z}^2$ is a contour of $\sigma \in \mathcal{S}^{\mathbb{Z}^2}$ if for every $e \in \gamma$ represents and edge $\{a, b\}$ where $\sigma_a^{(1)}\sigma_b^{(1)} = -1$. We can then see that if γ is a contour of σ with $\Delta_\gamma \subseteq \Lambda$ and a boundary condition $\omega \in \Omega$ then for $h = 0$, $\mathcal{H}^\omega_\Lambda(F^\Lambda_{\Delta_\gamma}(\sigma)) \geq 2k(J - J') + \mathcal{H}^\omega_\Lambda(\sigma)$. Thus, by following the old proof we can find a sufficient β for ferromagnetic properties regardless of the vertical interaction strength.

Proposition 5.1. *For all $V \in \mathbb{R}$, $J > J' \geq 0$, if $\beta \geq \frac{\ln(3)+1}{2(J-J')}$ and $h = 0$ then the 2 dimensional stacked lattice Ising model represents a ferromagnet.*

Proof. This will be a more condensed version of that seen in the proof of Theorem 4.1. Consider that we can still use the bound of $|\Gamma_k|$ given in Lemma 4.4. As such, we can define two boundary conditions $P, N \in \Omega$ where $P_{(i,j)} = \{1, 1\}$ and $N_{(i,j)} = \{-1, -1\}$for all $i, j \in \mathbb{Z}$. Furthermore, let $\rho(\cdot)$ be the weak limit of $\pi_{B_n}(P, \cdot)$

as n goes to infinity and $\nu(\cdot)$ be the weak limit of $\pi_{B_n}(N,\cdot)$ as n goes to infinity. As before, define the Borel measurable set $A = \{\sigma \in \Omega; \sigma^{(1)}_{(0,0)} = 1\}$. There is a symmetry of spins since $h = 0$ so $\rho(A) = \nu(A^c) = 1 - \nu(A)$. It suffices to show that $\rho(A) > 1/2$.

This is done much the same way as before, define the projection of A onto $\mathcal{S}^{B_n}$. We define ${}^nA = \{\sigma \in \mathcal{S}^{B_n}; \sigma^{(1)}_{(0,0)} = 1\} \subseteq 2^{\mathcal{S}^{B_n}}$. Given a closed simple walk on the dual lattice, γ we can define $({}^nA)^c_\gamma = \{\sigma \in ({}^nA)^c; \gamma \text{ is a contour in } \sigma^P\}$. From here we can repeat the computation done earlier:

$$\begin{aligned}
\rho(A^c) = \lim_{n\to\infty} \pi_{B_n}(P, A^c) &= \lim_{n\to\infty} \frac{\sum_{\sigma\in({}^nA)^c} e^{-\beta\mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta\mathcal{H}^P_{B_n}(\tau)}} \\
&\le \lim_{n\to\infty} \frac{\sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} \sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta\mathcal{H}^P_{B_n}(\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta\mathcal{H}^P_{B_n}(\tau)}} \\
\le \lim_{n\to\infty} \frac{\sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} e^{-\beta 2k(J-J')} \sum_{\sigma\in({}^nA)^c_\gamma} e^{-\beta\mathcal{H}^P_{B_n}(F^{B_n}_{\Delta_\gamma}\sigma)}}{\sum_{\tau\in\mathcal{S}^{B_n}} e^{-\beta\mathcal{H}^P_{B_n}(\tau)}} \\
\le \lim_{n\to\infty} \sum_{k=1}^\infty \sum_{\gamma\in\Gamma_k} e^{-\beta 2k(J-J')} \le \lim_{n\to\infty} \sum_{k=1}^\infty 4k\cdot 3^{k-2} e^{-\beta 2k(J-J')} \\
\le \lim_{n\to\infty} \sum_{k=1}^\infty 4k\cdot 3^{k-2} e^{-k(\ln 3+1)} < \frac{1}{2}.
\end{aligned}$$

Thus, we have shown that $\rho(A) > \frac{1}{2} > \nu(A)$, thus $\rho \neq \nu$ and there are two distinct Gibbs measures for the stacked lattice, thus the model represents a ferromagnet. □

This proof is a rather brute force application of the previous techniques to the stacked lattice. Our method required some non-trivial reworking of the problem to consider the two different interaction potentials. Nonetheless, what is interesting is this proof works completely independently from the strength of the vertical interactions, V. These interactions should have some effect because we can imagine that as V grows the model becomes more like a single lattice with interactions $J + J'$ and as V drops and the model becomes more like two separate lattices which can represent ferromagnets separately[7]. This intuition implies that V should have an effect, but for suitable J and J' the model is always a ferromagnet, even for negative V when the two lattices act like anti-ferromagnets with each other. In fact, this also applies to what we did in the one-dimensional stacked lattice, if you look back at the one-dimensional proof the precise values of analogies to J, J', and V do not matter and if we allowed different - even negative - interaction constants (analogues of J, J' and V) the proof would still hold. This implies that while the vertical interaction may have some effect it can be dominated by the other factors such as the dimension or the relative strength of different layers of the stacked lattice.

[7]This is based on the idea that if V is infinite then the two layers are identical, and if V is zero the two layers are independent. We can imagine that on a finite subset of the lattice we may pass through the limit, but it is not trivial to pass through the limit on the entire space.

6. Conclusion

It should be noted that this investigation was done by following two particular approaches, one using eigenvalue gaps of an interaction matrix for the one dimensional stacked lattice and another using contours and spin flips for the two dimensional stacked lattice. Both of these methods are introduced in [1]. These are not the only methods, for the single lattice there are other techniques which provide more exact solutions to the critical temperature (such as the more famous result of Onsager). An investigation of all these disparate techniques and attempts to extend them was outside the scope of this paper. These other techniques may fair better or worse than those shown here and could provide more or less satisfactory descriptions of the phase transition for stacked lattices and could be interesting topics for further investigation.

One of the goals of this paper was to explore how much stacking lattices mimicked adding to the dimension. Stacking one dimensional lattices changed very little about the problem and did not exhibit major aspects of higher dimensional Ising models since there was no ferromagnetic behavior. As for the two dimensional stacked lattice there is a lot less to say since our techniques did not extend to the analogue of the three dimensional model, but instead our methods worked for a special case when the two layers have different interaction strengths. In some sense the model did exhibit some of the three dimensional behavior in that it was more difficult, just as the exact formula for the three dimensional Ising model is an open problem and not solvable by extensions of well-known tricks. We did find an interesting set of conditions where the existence of ferromagnetic behavior was independent of the vertical interactions. In sum, more work is needed to understand the effects of stacking lattices particularly in the case of two dimensions where we can hope to approximate the more difficult three dimensional Ising model.

7. Acknowledgements

I have much to owe to my REU mentor, Jin Woo Sung, who guided me through general probability theory, introduced me to the Ising model, helped edit this paper, and put up with my confusion at various points in this process. Peter May also deserves much thanks for putting on the REU program and also helping edit this paper. I would also like to thank Greg Lawler who headed the probability section of the REU and Wendy Zheng who talked with me about some of the differences between the math and physics perspectives of this problem.

References

[1] Rassoul-Agha, Firas and Seppäläinen, Timo. *A Course on Large Deviations with an Introduction to Gibbs Measures*, American Mathematical Society 2015.