Junior Seminar in Mathematical Physics: Proofs from the Book

January 12, 2026

Organizational matters

- Students:
  1. Matias Andia (mati@princeton.edu): Mermin-Wagner Dobrushin-Shlosman or Pfister,
  2. Rees Barnes (rb5132@princeton.edu), LRO for $d \geq 3$, $N \geq 2$.
  3. Kevin Guan (kg8941@princeton.edu) , Lee-Yang a la Newman.
  4. Jerry Han (jerryhan@princeton.edu) Exp Decay at high temp a la Aizenman-Simon
  5. Olivia Kwon (sk9017@princeton.edu) Infinite Volume Gibbs,
  6. Joshua Lin (joshua.lin@princeton.edu) Peierls,
  7. Sonny Lowe (sonny@princeton.edu) Dobrushin Uniqueness,
  8. Selina Marvit (marvit@princeton.edu) McBryan-Spencer,
  9. Ken Matsuzaki (km8416@princeton.edu): Brydges-Fröhlich-Spencer
  10. Neill Saggi (ns9252@princeton.edu): Lee-Yang a la Lieb-Sokal
  11. Kashti Satish Umare (kashti@princeton.edu): Transfer matrices in 1D,
  12. (honorary) Zachary Ducorsky (zd0134@princeton.edu): Corr. inequalities.
- Meetings:
  1. Sep 18: Orientation, organization and intro to SM.
  2. Sep 25: (Olivia) Existence of infinite volume limit, (Kashti) Transfer matrix solution of $d = 1$.
  3. Oct 9: (Zach) Basic inequalities by Zach
  4. Oct 23: (Jerry) Aizenman-Simon, (Selina) Mcbryan-Spencer
  5. Nov 6: (Joshua) Peierls, (Sonny) Dobrushin Uniquness
  6. Nov 13: (Matias) Mermin-Wagner, (Ken) Brydges-Frohlich-Spencer graphical expansion
  7. Nov 20: (Rees) Frohlich-Simon-Spencer's LRO, (Neill) Lieb-Sokal's Lee-Yang

Guideline to students editing this document: Please try as much as you can to:

1. Follow the notation laid out in the beginning and integrate with the notation chosen before you.
2. Use

   ```
   \eq{ \hbar }
   ```

   and

   ```
   \eql{ \label{eq:my equation} \hbar }
   ```

   for *all* displayed equations. This is a shorthand for an align environment. Only use 'eql' if you plan to refer back to this equation or you anticipate others will refer back to this equation.

3. Cross-referencing throughout the document is to be done via the

```
\Cref{my_label}
```

command, rather than 'ref'.

4. Make sure you

```
\cite{my_citation}
```

any source you follow. This means you need to add the appropriate bibtex entry in our bib file beforehand.

5. If there are technical, algebraic or analytic tools that you use (for example properties of the discrete Laplacian) make them into separate lemmas and put those lemmas in an appendix down below.

6. Try your best not to create any custom commands but if you do create them at the very end of the preamble right before the begin document command. You should add a comment with your name, and make sure not to make any conflicts.

7. Try your best not to make any edits outside of your section but also try your best to cross-reference into other people's sections (e.g. if you use a correlation inequality, cross ref its statement).

8. *The document must compile with no errors after you finish editing it.*

9. Conventiosn for custom commands:

   (a) Any blackboard letter appears as double that letter, i.e., for $\mathbb{Z}$ we issue the command

   ```
   \ZZ.
   ```

   If the macro for the letter you need does not exist create it.

   (b) Always use $\mathbb{E}$ for expectation (rather than angular brackets) and $\mathbb{P}$ for the Gibbs probability measure (with or without qualifying super / sub scripts depending on whether confusion arises or not; generally speaking WITHIN calculations and proofs avoid superscripts and subscripts, in statement of theorems try to incorporate them).

   (c) The sphere is always backboarded as $\mathbb{S}^{N-1}$.

Contents

Part I

Statistical Mechanics: the classical $O(N)$ model

The main source of material for mathematical classical statistical mechanics is the Friedli and Velenik book [7]. Another useful source is the Peled Spinka lecture notes [11]. One should also look up lecture notes by Roland Bauerschmidt and Hugo Duminil-Copin.

1 The basic model

Let $d, N \in \mathbb{N}_{\geq 1}$ be given (the space dimension and *spin* dimension). Let $\Lambda \equiv \Lambda_L := [-L, L]^d \cap \mathbb{Z}^d$ be a finite box within $\mathbb{Z}^d$. This box can also be considered as a finite *graph* [1] $G = (V, E)$ i.e. a set of vertices $V = \Lambda \subseteq \mathbb{Z}^d$ and a set of edges E which indicate who is neighbor to whom. We have $|V| = (2L+1)^d$ and $|E| = d \times 2L \times (2L+1)^{d-1}$.

Pick some $\beta \in (0, \infty)$. We define the partition function of the d-dimensional classical $O(N)$ model, at inverse temperature β, initially in finite volume L as:

$$Z_{\beta,L}^{d,O(N)} := \int_{\psi:\Lambda\to\mathbb{S}^{N-1}} \exp\left(-\frac{1}{2}\beta\left\langle\psi, -\Delta\psi\right\rangle\right) \mathrm{d}\mu(\psi).$$

Here

$$\mu = \prod_{x\in\Lambda} \mu_0$$

i.e., the $|\Lambda|$-fold product measure all of the same copy of the measure μ_0, *the a-priori measure*. Naturally we choose the (normalized) volume measure on $\mathbb{S}^{N-1}$: in the case of $N = 1$ this is the (normalized) counting measure on $\{\pm 1\}$ but for $N \geq 2$ it is the natural measure which measures (normalized) area on the unit sphere.

Moreover, the symbol $\langle\psi, -\Delta\psi\rangle$ is somewhat abbreviated notation for

$$\langle\psi, -\Delta\psi\rangle \equiv \sum_{x,y\in\Lambda} (-\Delta)_{xy} \langle\psi_x, \psi_y\rangle_{\mathbb{R}^N}$$

and $-\Delta$ is an $|\Lambda| \times |\Lambda|$ matrix to be specified shortly.

The finite volume is a handy technical tool to avoid talking about probability measures of infinite stochastic processes. Ultimately our aim is to derive any result (read: estimate) uniformly in L so that conclusions are made about the $L \to \infty$ limit (that limit exists, but let us avoid this question for a minute).

The number of points in our box is $k := |\Lambda| = (2L+1)^d$. Hence $\mathbb{S}^{N-1}$ is the $(N-1)$-dimensional sphere within $\mathbb{R}^N$, and we should view ψ as a map from Λ into $\mathbb{R}^N$. I.e., for any $x \in \Lambda$, $\psi_x \in \mathbb{R}^N$ and $\|\psi_x\|_2^2 = 1$. Thus, with some abuse of notation, if $A \in \mathrm{Mat}_{k\times k}(\mathbb{R})$ then

$$\langle\psi, A\psi\rangle \equiv \sum_{x,y\in\Lambda} \sum_{i=1}^{N} (\psi_x)_i A_{xy} (\psi_y)_i .$$

(Truly we should have written $A \otimes \mathbb{1}_N$ instead of A....).

The symbol $-\Delta$ is the discrete Laplacian. For every choice of L, it is a $k \times k$ matrix, with $k = (2L+1)^d$, given as follows:

$$(-\Delta v)_x := \sum_{y\sim x} v_x - v_y \qquad (v : \Lambda \to \mathbb{R}, x \in \Lambda)$$

where $y \in \Lambda$ obeys $y \sim x$ if and only if y is "adjacent" to x in Λ, that is, a nearest neighbor. In terms of matrices,

$$-\Delta = D - A$$

where A is the *adjacency* matrix of the graph Λ (i.e. it equals 1 if there is an edge between two vertices and 0 otherwise) and D is the diagonal *degree* matrix of the graph (specifying the number of edges connected to a given vertex). Here is the point where the discussion of boundary conditions enters: we may decide that for those vertices of Λ at the boundary, they have less neighbors than those in the bulk (*free boundary conditions*), or we may decide to wrap Λ around itself, i.e., to make a torus, to form *periodic boundary conditions*. Simultaneously, we may also consider other custom options, e.g., that the boundary is pinned to a certain range of values. By the way, the values of the boundary need not necessarily be on the sphere. In principle these choices need to be specified when $-\Delta$ is discussed.

[1] There is a whole direction of research to ask how does exotic properties of different graphs affect the phase transitions we are studying. We shall not pursue this direction here and for the most part stick with finite sub-graphs of $\mathbb{Z}^d$.

For example let us illustrate this with the choice $d=1$ and then, say, $L=4$, (so $k=9$). We get

$$-\Delta_{\text{free}}=\begin{bmatrix}1&-1&&&&&&&\\-1&2&-1&&&&&&\\&-1&2&-1&&&&&\\&&-1&2&-1&&&&\\&&&-1&2&-1&&&\\&&&&-1&2&-1&&\\&&&&&-1&2&-1&\\&&&&&&-1&2&-1\\&&&&&&&-1&1\end{bmatrix}$$

and

$$-\Delta_{\text{periodic}}=\begin{bmatrix}2&-1&&&&&&&-1\\-1&2&-1&&&&&&\\&-1&2&-1&&&&&\\&&-1&2&-1&&&&\\&&&-1&2&-1&&&\\&&&&-1&2&-1&&\\&&&&&-1&2&-1&\\&&&&&&-1&2&-1\\-1&&&&&&&-1&2\end{bmatrix}.$$

It should be noted that we could also rewrite the bilinear form as follows. Assume $-\Delta_{\text{periodic}}$ is used for the moment (then $2d|\Lambda|=2|E|$ where E is the set of edges). Then

$$\begin{aligned}\langle\psi,-\Delta\psi\rangle &= \sum_{x\in\Lambda}\psi_x\cdot\sum_{y\sim x}\psi_x-\psi_y\\ &= \sum_{x\in\Lambda}\underbrace{\psi_x\cdot\psi_x}_{=1}\underbrace{\left(\sum_{y\sim x}\right)}_{\deg(x)}-\sum_{x\in\Lambda,y\sim x}\psi_x\cdot\psi_y\\ &= 2|E|-2\sum_{\{x,y\}\in\Lambda:x\sim y}\psi_x\cdot\psi_y\,.\end{aligned}$$

However,

$$\|\psi_x-\psi_y\|^2_{\mathbb{R}^N}=\|\psi_x\|^2_{\mathbb{R}^N}+\|\psi_y\|^2_{\mathbb{R}^N}-2\psi_x\cdot\psi_y=2(1-\psi_x\cdot\psi_y)$$

so

$$\langle\psi,-\Delta\psi\rangle=\sum_{\{x,y\}\in\Lambda:x\sim y}\|\psi_x-\psi_y\|^2_{\mathbb{R}^N}=2|E|-2\sum_{\{x,y\}\in\Lambda:x\sim y}\psi_x\cdot\psi_y\,.$$

We emphasize that constant (ψ-independent) terms in the bilinear form are irrelevant since we are only interested in ratios. Hence we understand

$$\langle\psi,-\Delta\psi\rangle$$

as measuring the total amount of (squared) disagreement throughout the grid Λ: Moreover, since these are unit vectors, $\psi_x\cdot\psi_y$ gives the cosine of the angle between the two vectors as measured using the geodesic length on the sphere. Full agreement is when $\psi_x\cdot\psi_y=1$, so maximal agreement is the *minimum* value of $\langle\psi,-\Delta\psi\rangle$, which we call *the energy* usually denoted by H and also called *the Hamiltonian* or *the interaction*. We can also consider more general energy functions

$$H:\left(\mathbb{S}^{N-1}\right)^\Lambda\to\mathbb{R}\,.$$

Hence generally

$$\exp\left(-\frac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)=\exp\left(-\beta H(\psi)\right)\,.$$

Since $\beta>0$, those configurations $\psi:\Lambda\to\mathbb{S}^{N-1}$ which minimize the energy functional H are those which maximize agreement throughout. For this reason these models are called *ferromagnetic*. Anti-ferromagnetic models maximize disagreement and may be obtained by $H\mapsto -H$.

Note that the matrix $-\Delta$ may be diagonalized easily (it is symmetric). E.g. for $-\Delta_{\text{periodic}}$ one may use the Fourier series (or its discrete version on Λ). The k eigenvalues lie within the interval

$$[0, 4d] \ .$$

Zero is always an eigenvalue and it corresponds to the eigenvector (assuming we *avoid* pinning the field, described right below) which is a constant configuration throughout: that is the energy minimizing configuration.

Pinned boundary conditions are implemented as follows. One takes $-\Delta_{\text{free}}$ or $-\Delta_{\text{periodic}}$, but also picks some fixed $B \subseteq \Lambda$ (the boundary set, though in principle it can be any subset of Λ, also just one vertex in the middle) and a "boundary values field" $\varphi : B \to \mathbb{S}^{N-1}$ (actually the co-domain is allowed to even be a general $\mathbb{R}^N$ vector) and then instead of take the integral over all configurations $\Lambda \to \mathbb{S}^{N-1}$, restrict to the integral over the set

$$\Omega_\varphi := \left\{ \psi : \Lambda \to \mathbb{R}^N \ \middle| \ \psi|_B = \varphi \wedge \|\psi_x\| = 1 \forall x \in \Lambda \setminus B \right\} \ .$$

I.e., really, it is actually an integral over

$$|\Lambda| - |B|$$

spheres.

Finally, we also want to allow for an external magnetic field $h : \Lambda \to \mathbb{R}^N$. It enters into the Hamiltonian as

$$H(\psi) = \frac{1}{2} \langle \psi, -\Delta \psi \rangle - \langle h, \psi \rangle \ .$$

If we take h to be non-zero only along $\partial \Lambda$ (those vertices with less than $2d$ degree) then achieve a similar effect to setting the values of ψ on the boundary of a slightly large graph $\overline{\Lambda}$ to h.

The distinction from $-\Delta_{\text{periodic}}$ to $-\Delta_{\text{free}}$ is not terribly important for us now so going forward, unless otherwise noted, we shall use $-\Delta_{\text{free}}$ together with some given field $\varphi : B \to \mathbb{S}^{N-1}$ (the object φ carries the specification of its domain automatically).

In principle the measure depends on the boundary condition φ also, so we should really denote

$$Z^{d,O(N)}_{\beta,L,\varphi,h} := \int_{\psi \in \Omega_\varphi} \exp\left(-\frac{1}{2} \beta \langle \psi, -\Delta \psi \rangle + \beta \langle h, \psi \rangle \right) \prod_{x \in \Lambda} \mathrm{d}\mu_0(\psi_x) \ .$$

1.1 Terminology

We list some terminology from statistical mechanics:

1. The quantity $Z^{d,O(N)}_{\beta,L,\varphi,h}$ is called *the partition function*. The summand within it is called *the Gibbs factor* and the associated probability measure $\mathbb{P}^{d,O(N)}_{\beta,L,\varphi,h}$ on Ω_φ is called *the Gibbs measure*.

2. If $N = 1$ we have the *Ising model*. If $N = 2$ we have the XY or $O(2)$ model. If $N = 3$ we have the *classical (isotropic) Heisenberg* or $O(3)$ model. The $N \to \infty$ is sometimes referred to as the *spherical* limit.

3. The $L \to \infty$ limit (you cannot take that limit at the level of $Z^{d,O(N)}_{\beta,L,\varphi,h}$, you must only take it for *ratios* such as $\mathbb{P}^{d,O(N)}_{\beta,L,\varphi,h}$ or $\mathbb{E}^{d,O(N)}_{\beta,L,\varphi,h}$) is referred to as *the thermodynamic* or *infinite volume* limit. Hence for now let us take for granted that there is some measure $\mathbb{P}^{d,O(N)}_{\beta,\varphi,h}$ which is to be understood as a measure on the space of functions $\mathbb{Z}^d \to \mathbb{S}^{N-1}$ and is obtained as the limit of sequence of measures $\left\{ \mathbb{P}^{d,O(N)}_{\beta,L,\varphi_L,h_L} \right\}_L$. We will study the existence and nature of this limit very soon. Note that some care has to be taken with the specification of the boundary conditions here because we let $\varphi : B \to \mathbb{S}^{N-1}$ and $B \subseteq \Lambda_L$, so if L varies so does B and hence φ, in principle. Then it remains to be seen if the infinite volume object $\mathbb{P}^{d,O(N)}_{\beta,\varphi,h}$ has any "memory" of φ or not and what sense does it make to keep carrying φ in the notation. We will study this question too below.

4. *Uniqueness of the Gibbs measure* is the general statement that $\mathbb{P}^{d,O(N)}_{\beta,\varphi,h}$ does *not* depend on φ (for a particular class of sequences $\{ \varphi_L \}_L$), i.e., there is a-posteriori only *one* infinite-volume Gibbs measure.

5. *The two-point function* is the map

$$\mathbb{Z}^d \times \mathbb{Z}^d \ni (x, y) \mapsto \mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} \right] \in [-1, 1] \ .$$

The *truncated-two-point-function* is

$$\mathbb{Z}^d \times \mathbb{Z}^d \ni (x,y) \mapsto \left(\mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} \right] - \left\langle \mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\psi_x \right], \mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\psi_y \right] \right\rangle_{\mathbb{R}^N} \right) \in [-1,1] \, .$$

Measures how far away spins are correlated with each other.

6. Magnetization at a site x is $\mathbb{E}^{d,O(N)}_{\beta,\varphi,h} [\psi_x]$ and total magnetization is

$$m^{d,O(N)}_{\beta,\varphi,h} := \lim_{L\to\infty} \mathbb{E}_{d,N,\beta,\varphi,\Lambda} \left[\frac{1}{|\Lambda|} \sum_{x\in\Lambda} \psi_x \right] \in \mathbb{R}^N \, .$$

7. The quantity $F^{d,O(N)}_{\beta,L,\varphi,h} := -\frac{1}{\beta} \log \left(Z^{d,O(N)}_{\beta,L,\varphi,h} \right)$ is called *the free energy.* Its volume-density is

$$f^{d,O(N)}_{\beta,\varphi,h} := \lim_{L\to\infty} \frac{1}{|\Lambda_L|} F^{d,O(N)}_{\beta,L,\varphi,h}$$

should this limit exist.

8. There are two main phenomena we are interested in when studying this model:

   (a) *Long range order* (low temperatures, high β): the system has intrinsic global, collective magnetization, $m^{d,O(N)}_{\beta,\varphi,h} \neq 0$, $\mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} \right]$ does not decay as $\|x-y\| \to \infty$.

   (b) *Disordered phase* (high temperatures, low β): the system is does not show preference to any particular direction, $m^{d,O(N)}_{\beta,\varphi,h} = 0$, $\mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} \right]$ decays as $\|x-y\| \to \infty$ (however slowly). *Correlation length* is the rate of exponential decay.

   (c) *Phase transition*: the shift of the system from one type of the above behavior to another type of the above behavior as a continuous parameter (usually the inverse temperature) is varied.

   (d) *Criticality* or *critical point*: The set of parameters of the system on the boundary between two phases, i.e. the point (or line, or manifold) of phase transition, $\mathbb{E}^{d,O(N)}_{\beta,\varphi,h} \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} \right]$ decays but *polynomially.*

9. *Universality* refers to a type of behavior of certain quantities, usually asymptotically, usually near the critical point.

10. *Gaussianity*, *free-field*, or *spin-wave* behavior is the phenomenon that the random field $\psi : \Lambda \to \mathbb{R}^N$ behaves as if it had a Gaussian measure (it does not due to the a-priori measure μ_0). Gaussian upper bounds are upper bounds (e.g., on the two point function) in terms of the two-point function of the *Gaussian field* or the *Gaussian free field.*

11. By *symmetry* we refer to the operation of rotating a vector in $\mathbb{R}^N$ from one direction to another, and observing that *something* remains the same. For example, the inner product

$$\langle \psi_x, \psi_y \rangle$$

is invariant if we apply an orthogonal (rotation) matrix on both vectors simultaneously. The group of $N \times N$ orthogonal matrices, $O(N)$, is the main group of symmetries of our model, since $Z_{d,N,\beta,L,\varphi=0}$ possesses a global $O(N)$ symmetry, in the sense that the probability density (or the push forward of the probability measure if you wish) remains the same if we apply a global $M \in O(N)$ matrix to the magentization vector on all vertices of Λ. There is a slight issue here with boundary conditions which would spoil that, so in principle if $\varphi \neq 0$ then we have to rotate the boundary conditions too.

   (a) *Continuous symmetry* means that the group of symmetries is a smooth manifold as opposed to a discrete finite group. Compare the discontinuous case $N=1$ (whence $O(1) = \{ \pm 1 \} \cong \mathbb{Z}_2$) with the continuous case $N \geq 2$ ($O(2) \cong \mathbb{S}^1$). It turns out that discrete versus continuous symmetry plays a role.

   (b) *Symmetry breaking or spontaneous symmetry breaking* is the situation where the finite volume Gibbs *factor* is invariant under some symmetry (for any given finite volume) yet *the infinite volume measure is not.* This phenomenon is of utmost interest to us and will have parallels in quantum mechanics as well. Long-range order from above is an example of such symmetry breaking whereas the disordered phase is the absence of symmetry breaking.

(c) *Mermin–Wagner* (sometimes Mermin–Wagner–Hohenberg) is the general result that there is no continuous symmetry breaking if $d \leq 2$.

12. *Kosterlitz–Thouless* or *Berezinskii-Kosterlitz–Thouless* is a phenomenon of a whole critical interval of temperatures, usually $[\beta_c, \infty)$.

13. *Mass gap* or *massive* field is a field whose two-point function decays exponentially in $\|x - y\|$. This is called in this way because in a Gaussian free field, if we replace the Laplacian

$$-\Delta$$

with a massive Laplacian

$$-\Delta + m^2 \mathbb{1}$$

then we indeed get exponentially decay of the two point function with rate m.

1.2 The phases of the classical $O(N)$ model

> "All happy families are alike, each unhappy family is unhappy in its own fashion." This observation from the opening lines of Leo Tolstoy Anna Karenin can well serve as an epigraph to the family of papers that includes the present one. The main goal of these paper is to demonstrate that, contrary to the richness of the behavior exhibited by Gibbs fields at low temperatures, their properties outside the phase transition region are quite uniform.
>
> – Dobrushin and Shlosman in [4].

We will see in Section 7 that all systems we consider are "boring", i.e., disordered if the temperature is sufficiently high (dependent on d and N of course). For the most part this motivates focusing on either low temperatures or studying what happens near the (if there is a single) critical temperature.

<table>
<tr><th>N \ d</th><th>1</th><th>2</th><th>3</th><th>4</th><th>≥ 5</th></tr>
<tr><td>1</td><td colspan="5" rowspan="5">disordered</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td></tr>
<tr><td>≥ 5</td></tr>
</table>

Table 1: Low β (high temperature): all entries are disordered for generic reasons.

<table>
<tr><th>N \ d</th><th>1</th><th>2</th><th>≥ 3</th></tr>
<tr><td>1</td><td rowspan="3">disordered
(transfer matrix)</td><td colspan="2">LRO (Peierls)</td></tr>
<tr><td>2</td><td>BKT (vortex binding)</td><td rowspan="2">LRO (SSB; infrared bounds)</td></tr>
<tr><td>≥ 3</td><td>disordered? (Polyakov conjecture)</td></tr>
</table>

Table 2: High β (low temperature): phases by (N, d) and indicative mechanisms. LRO = long-range order; BKT = Berezinskii–Kosterlitz–Thouless; SSB = spontaneous symmetry breaking.

2 The Gaussian free field (GFF)

The *Gaussian free field* is defined in almost identical way to the $O(N)$ model, with the exception of replacing the a-priori measure μ_0 from the volume measure on $\mathbb{S}^{N-1}$ by the Lebesgue measure on $\mathbb{R}^N$. This can be risky because now we run the risk of integrals not converging. This danger can be mitigated in various different ways, either through boundary conditions or the addition of a mass to the Laplacian.

Let us study the model whose partition function is

$$Z_{\beta,L,m}^{d,\mathrm{GFF}_N} := \int_{\psi:\Lambda\to\mathbb{R}^N} \exp\left(-\frac{1}{2}\beta\left\langle\psi,\left(-\Delta+m^2\mathbb{1}\right)\psi\right\rangle\right)\mathrm{d}\psi$$

where by $\mathrm{d}\psi$ we mean the Lebesgue measure on $\left(\mathbb{R}^N\right)^\Lambda$. Because $m\neq 0$ then all Gaussian integrals converge regardless of the spectrum of $-\Delta$ (whether it has eigenvalue zero or not; with Dirichlet boundary conditions for example $-\Delta$ anyway has no zero mode and then the integral converges even with $m=0$). We then have

$$Z_{\beta,L,m}^{d,\mathrm{GFF}_N} = \int_{\psi:\Lambda\to\mathbb{R}^N} \exp\left(-\frac{1}{2}\beta\left\langle\psi,\left(-\Delta+m^2\mathbb{1}\right)\psi\right\rangle\right)\mathrm{d}\psi = \frac{(2\pi)^{N\frac{|\Lambda|}{2}}\beta^{-\frac{N|\Lambda|}{2}}}{\sqrt{\det_\Lambda\left(-\Delta+m^2\mathbb{1}\right)^N}}$$

and

$$\int_{\psi:\Lambda\to\mathbb{R}^N} \exp\left(-\frac{1}{2}\beta\left\langle\psi,\left(-\Delta+m^2\mathbb{1}\right)\psi\right\rangle+\langle J,\psi\rangle\right)\mathrm{d}\psi = Z_{\beta,L,m}^{d,\mathrm{GFF}_N}\exp\left(\frac{1}{2\beta}\left\langle J,\left(-\Delta+m^2\mathbb{1}\right)^{-1}J\right\rangle\right)$$

so that

$$\mathbb{E}\left[\psi_x\cdot\psi_y\right] = \frac{N}{\beta}\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{xy} \qquad (x,y\in\Lambda)\ .$$

First note that beyond the overall constant outside, $\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{xy}$ is independent of β. In particular the fate of exponential decay or not of the two-point function is independent of β and hence the GFF has no phase transitions. This is also clear by looking at the integral and making a change of variable $\psi_x\mapsto\sqrt{\beta}\psi_x$.

Note that for $\|x\|\gg\frac{1}{m}$, we have

$$\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{0,x} \sim c_d(m)\|x\|^{-\frac{d-1}{2}}\exp\left(-\frac{1}{m}\|x\|\right)\ .$$

However, if we're really interested in the $m=0$ case, then we have the following behavior ($m\to 0^+$ asymptotics at fixed large $\|x\|$):

$$\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{0,x} = \begin{cases}\frac{1}{2m}-\frac{\|x\|}{2}+O(m) & d=1\\ \frac{1}{2\pi}\log\left(\frac{1}{m}\right)-\frac{2}{\pi}\log\left(\|x\|\right)+C+o(1) & d=2\\ \frac{\Gamma\left(\frac{d}{2}-1\right)}{4\pi^{\frac{d}{2}}}\|x\|^{2-d} & d\geq 3\end{cases}.$$

We see that in $d\leq 2$ this limit $m\to 0^+$ does not exist. One way to cure this is to always consider differences:

$$\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{0,x}-\left[\left(-\Delta+m^2\mathbb{1}\right)^{-1}\right]_{0,1} = \begin{cases}-\frac{\|x\|}{2}+\frac{1}{2} & d=1\\ -\frac{2}{\pi}\log\left(\|x\|\right)+o(1) & d=2\\ \frac{\Gamma\left(\frac{d}{2}-1\right)}{4\pi^{\frac{d}{2}}}\left(\|x\|^{2-d}-1\right) & d\geq 3\end{cases}.$$

This suggests that objects like

$$\mathbb{E}\left[\|\psi_x-\psi_0\|^2\right]$$

are more appropriate than

$$\mathbb{E}\left[\psi_x\cdot\psi_0\right]$$

where studying $d\leq 2$ massless Gaussian fields, i.e., it is the *difference* field $x\mapsto\psi_x-\psi_0$ or $(x,y)\mapsto\psi_x-\psi_y$ which is finite no matter what.

3 Olivia: The existence of the infinite volume Gibbs measure and the DLR conditions

In this section, we assume basic measure theory. If necessary, one can refer to, for example, [13, 6, 15]. In addition, we assume basic point set topology (refer to, for instance, [10]). Moreover, we assume that readers are familiar with conditional probability. The preliminaries about conditional probability can be found in Appendix A. Moreover, we assume that the readers are familiar with concepts of probability kernel, specifications, and how they relate to conditional probability. One can read about this in Appendix B.

The goal of this section is to convey to readers that there exists an infinite-volume Gibbs measure for the classical $O(N)$ model. This section mostly follows ideas [7, Chapter 6]. However, we will prove the existence of an infinite-volume Gibbs measure for the general $O(N)$ model, contrary to [7], which proves only for the Ising ($N = 1$) model. To do this, we appeal to functional analysis (especially the Banach-Alaoglu Theorem!). While we will not give proof of the Banach-Alaoglu Theorem, we will state it when we use it, to make this section more self-contained.

3.1 Notations and Introduction to Infinite-Volume Gibbs Measure

We stay consistent with notations from Section 1. To summarize, we denote by $\Lambda_L \equiv [-L, L]^d \subset \mathbb{Z}^d$; $\psi : \Lambda_L \to \mathbb{S}^{N-1}$ is the random spin field; $\varphi : B \subset \Lambda_L \to \mathbb{S}^{N-1}$ is the deterministic field specifying the boundary conditions; and $\Omega_{\Lambda_L,\varphi}$ to be set of all spins configurations ψ on Λ_L with boundary condition φ. Lastly, we denote the space of all possible spin configuration in infinite volume (with no pinning) by $\Omega \equiv \{\psi : \mathbb{Z}^d \to \mathbb{S}^{N-1}\} = \bigtimes_{i\in\mathbb{Z}^d} \mathbb{S}^{N-1}$. To be clear, we will not allow pinning in the infinite volume measures.

Remark 3.1. *We defined Λ_L to be a finite box of side length $2L + 1$. However, we really just needed Λ_L to be a finite subset of $\mathbb{Z}^d$. One can check that there will be nothing dependent on the fact that Λ_L is a box, except for ease of notation. From what follows, we will denote by $\Lambda \Subset \mathbb{Z}^d$ a finite subset of $\mathbb{Z}^d$.*

We moreover introduce convenient notation. Given $\psi \in \Omega$, we denote by $\psi_\Lambda \in \Omega_\Lambda$ the restriction of ψ to the coordinates of Λ.

Now, let us start discussing the Gibbs measure. Let us start with the finite case. Here, we denote by $\beta > 0$ the inverse temperature; $h : \Lambda_L \to \mathbb{R}^N$ is the external magnetic field; $H_\Lambda : \Omega_{\Lambda,\varphi} \to \mathbb{R}_{\geq 0}$ is the Hamiltonian defined by $H(\psi) = \frac{1}{2}\langle\psi, -\Delta\psi\rangle - \langle h, \psi\rangle$ where the discrete Laplacian $-\Delta$ is the finite $|\Lambda| \times |\Lambda|$ matrix which acts on sites in or connected to Λ; $Z^{d,N}_{\beta,L,\varphi}$ is the corresponding partition function with boundary condition φ. All these are consistent with notations from Section 1. In the finite case, the definition of the Gibbs measure is clear.

Definition 3.2. *Given $\Lambda_L \subset \mathbb{Z}^d$, let us define $\mathscr{F}_\Lambda$ as the Borel σ-algebra of Ω_Λ w.r.t. the product topology. Given boundary conditions φ, we define the finite Gibbs measure, a probability measure on $(\Omega, \mathscr{F}_\Lambda)$, by*

$$\mathbb{P}^{d,N}_{\beta,h,L,\varphi}(A) := \int_{\psi\in A} \frac{\exp(-\beta H_\Lambda(\psi))}{Z^{d,N}_{\beta,h,L,\varphi}} \mathrm{d}\left(\prod_{i\in\Lambda_L} \sigma\right)\psi \qquad (A \in \mathscr{F}_\Lambda)\,. \tag{1}$$

Here $\mu = \left(\prod_{i\in\Lambda_L} \mu_0\right)$ where μ_0 is the volume measure on $\mathbb{S}^{N-1}$. In the case when $N = 1$, μ_0 would just be the counting measure.

Now, let us start talking about the infinite-volume Gibbs measure. First, what do we mean by an infinite-volume Gibbs measure? It does not make sense to simply replace Λ with $\mathbb{Z}^d$ in Definition 3.2, since then: while $-\Delta$ becomes an operator instead of a matrix, $H_{\mathbb{Z}^d}$ does *not make sense*: it will not converge.

So without direct reference to $H_{\mathbb{Z}^d}$, we want to somehow take a limit and define infinite-volume Gibbs measure as

$$\mu_{\beta,h} \text{ “ = ” } \lim_{L\to\infty} \mu_{\beta,h,L,\varphi}.$$

However, we will have to address some questions such as:

1. Which σ-algebra $\mathscr{F}$ on Ω are we considering?
2. As $L \to \infty$, Hamiltonian and Partition Function are not well defined. Also, each spin configuration ψ has probability 0 of happening as $L \to \infty$. How should we get away with this?
3. Is the limit dependent on the boundary condition φ we pick, or the sequence of boxes Λ_L which we take?

Out of the above three questions, we can only answer the first immediately: we take for $\mathscr{F}$ *the Borel σ-algebra on* Ω. By definition of the product topology, $\mathscr{F}$ is the σ-algebra generated by sets which are inverse images of finite-subset projections, i.e.

$$\mathscr{F} = \sigma\left(\rho_\Lambda^{-1}(A) : \Lambda \Subset \mathbb{Z}^d, A \in \mathscr{P}\left(\prod_{i\in\Lambda\Subset\mathbb{Z}^d} \mathbb{S}^{N-1}\right)\right)$$

where $\rho_\Lambda : \Omega \to \prod_{i\in\Lambda\Subset\mathbb{Z}^d} \mathbb{S}^{N-1}$ is defined as the projection onto the coordinates within Λ. We want our infinite-volume Gibbs measure, when we construct it, to be a measure on $(\Omega, \mathscr{F})$. Moreover, we will use $M_1(\Omega)$ to denote set of all probability measures on $(\Omega, \mathscr{F})$.

In addition, for every finite subset $\Lambda \Subset \mathbb{Z}^d$, we will denote by the σ-algebra

$$\mathscr{F}_\Lambda = \sigma\left(\rho_\Lambda^{-1}(A) : A \in \mathscr{P}\left(\prod_{i\in\Lambda\Subset\mathbb{Z}^d} \mathbb{S}^{N-1}\right)\right).$$

Notice that $\mathscr{F}_\Lambda$ is a sub-σ algebra of $\mathscr{F}$.

As we describe the σ-algebra on Ω, let us also introduce an essential vocabulary that will be crucial in constructing an infinite-Gibbs measure.

Definition 3.3 (Local Function)**.** *Given $f : \Omega \to \mathbb{R}$ bounded, we say that f is* local on $\Lambda \Subset \mathbb{Z}^d$ *iff $f(\psi) = f(\widetilde{\psi})$ for every $\psi, \widetilde{\psi} \in \Omega$ such that $\psi_\Lambda = \widetilde{\psi}_\Lambda$. In other words, f is local if it only depends on the Ω_Λ components.*

We remark that all bounded local functions are measurable in $\mathscr{F}$; what's more, they are also $\mathscr{F}_\Lambda$-measurable by definition.

We now go back to the questions proposed above. Unfortunately, the next two questions are hard to answer; therefore, we avoid answering this question and rather try to define an infinite-volume Gibbs measure with no reference to a limit. As we will see, *the DLR Condition*, which ensures compatibility with the finite volume measures, is the characterization of infinite-volume Gibbs measure that we will use to define infinite-volume Gibbs measure without reference to a limit.

3.2 The DLR Condition

In this section, we introduce DLR condition in the theorem below, which holds for every finite-volume Gibbs measure.

Theorem 3.4 (Dobrushin, Lanford, and Ruelle)**.** *For every $f : \Omega \to \mathbb{R}$ measurable and bounded and for all $B \subset \Lambda \Subset \mathbb{Z}^d$,*

$$\mathbb{E}_{\Lambda,\varphi}[f] = \mathbb{E}_{\Lambda,\varphi}[\mathbb{E}_{B,\cdot}[f])]. \tag{2}$$

Proof. We start at the right-hand side.

$$\mathbb{E}_{\Lambda,\varphi}[\mathbb{E}_{\Delta,\cdot}(f))] = \int_{\psi\in\Omega_\Lambda} \mathbb{E}_{\Delta,\psi_\Lambda\varphi_{\Lambda^c}}[f] \frac{\exp(-\beta H_\Lambda(\psi_\Lambda\varphi_{\Lambda^c}))}{Z_{\Lambda,\psi_\Lambda\varphi_{\Lambda^c}}} \left(\prod_{i\in\Lambda} \sigma\right)(d\psi). \tag{3}$$

Similarly,

$$\mathbb{E}_{\Delta,\psi_\Lambda\varphi_{\Lambda^c}}[f] = \int_{\eta\in\Omega_\Delta} f(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) \frac{\exp(-\beta H_\Delta(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}))}{Z_{\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}}} \left(\prod_{i\in\Delta} \sigma\right)(d\eta). \tag{4}$$

Observe further that because H_Λ only depends on Ω_Λ components of the spins and the boundary, we have that

$$\begin{aligned} H_\Lambda(\psi_\Lambda\varphi_{\Lambda^c}) - H_\Delta(\psi_\Lambda\varphi_{\Lambda^c}) &= H_\Lambda(\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) - H_\Delta(\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) \\ &= H_\Lambda(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) - H_\Delta(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}). \end{aligned} \tag{5}$$

Therefore, we substitute (4) into (3) and use (5). For the sake of brevity, we write $d\psi_\Lambda$ instead of $\left(\prod_{i\in\Lambda}\sigma\right)(d\psi_\Lambda)$

because it is clear we are taking an integral with respect to the Haar measure. Then, we get that

$$\begin{aligned}
&\mathbb{E}_{\Lambda,\varphi}[\mathbb{E}_{\Delta,\cdot}(f))] \\
&= \int_{\psi_{\Lambda-\Delta}\in\Omega_{\Lambda-\Delta}} \int_{\psi_\Delta\in\Omega_\Delta} \int_{\eta_\Delta\in\Omega_\Delta} f(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) \frac{\exp(-\beta(H_\Delta(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) + H_\Lambda(\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c})))}{Z_{\Delta,\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}} Z_{\Lambda,\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}}} d\eta_\Delta d\psi_\Delta d\psi_{\Lambda-\Delta} \\
&= \int_{\psi_{\Lambda-\Delta}\in\Omega_{\Lambda-\Delta}} \int_{\eta_\Delta\in\Omega_\Delta} \int_{\psi_\Delta\in\Omega_\Delta} f(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) \frac{\exp(-\beta(H_\Delta(\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) + H_\Lambda(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c})))}{Z_{\Delta,\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}} Z_{\Lambda,\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}}} d\psi_\Delta d\eta_\Delta d\psi_{\Lambda-\Delta} \\
&= \int_{\psi_{\Lambda-\Delta}\in\Omega_{\Lambda-\Delta}} \int_{\eta_\Delta\in\Omega_\Delta} f(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}) \frac{\exp(-\beta H_\Delta(\eta_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}))}{Z_{\Lambda,\eta_\Delta\psi_{\Lambda-\Delta_{\varphi_{\Lambda^c}}}}} \underbrace{\frac{\int_{\psi_\Delta\in\Omega_\Delta}\exp(-\beta H_\Lambda(\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}))d\psi_\Delta}{Z_{\Delta,\psi_\Delta\psi_{\Lambda-\Delta}\varphi_{\Lambda^c}}}}_{=1} d\eta_\Delta d\psi_{\Lambda-\Delta} \\
&= \int_{\psi_\Lambda\in\Omega_\Lambda} f(\psi_\Lambda\varphi_{\Lambda^c}) \frac{\exp(-\beta H_\Lambda(\psi_\Lambda\varphi_{\Lambda^c}))}{Z_{\Lambda,\psi_\Lambda\varphi_{\Lambda^c}}} d\psi_\Lambda \qquad (\because \text{Let } \psi_\Lambda = \eta_\Delta\psi_{\Lambda-\Delta}) \\
&= \mathbb{E}_{\Lambda,\varphi}[f].
\end{aligned}$$

□

Of course, if we were to take the 'limit' as $\Lambda \nearrow \mathbb{Z}^d$ after taking expectation, by DLR Condition, we will have that the resulting infinite-volume measure μ satisfies

$$\mathbb{E}_\mu[f] = \mathbb{E}_\mu[E_{\Lambda,\cdot}(f)] \tag{6}$$

for every bounded and measurable function f. However, it turns out that DLR Condition also gives us a way to define infinite-volume Gibbs measure without reference to limits.

To do this, we first notice that for every fixed f continuous, the function

$$\psi \mapsto \mathbb{E}_{\Lambda,\psi}[f]$$

is a continuous, bounded, local function on Λ as a function of ψ and hence measurable. Now, we are ready to define the infinite-volume Gibbs measure indeed without reference to limits. In the definition below, we make strong use of the Kakutani-Markov-Riesz Theorem, whose statement we recall first.

Theorem 3.5 (Kakutani-Markov-Riesz)**.** *Let Ω be a locally compact Hausdorff space and T a positive linear functional on $C_c(\Omega \to \mathbb{C})$. Then, there exists a unique positive measure μ such that for every $f \in C_c(\Omega \to \mathbb{C})$*

$$T(f) = \int_\Omega f(\psi)\mu(d\psi).$$

For the proof of it, please refer to [15, 6, 13].

Using this theorem, we may now make the

Definition 3.6 (Infinite-Volume Gibbs Measure)**.** *Suppose we have a positive linear functional $T : C_c(\Omega) \to \mathbb{R}$ such that*

$$T(f) = T(\mathbb{E}_{\Lambda,\cdot}[f])$$

for every $\Lambda \Subset \mathbb{Z}^d$. Consider μ measure corresponing to T given via Riesz–Markov–Kakutani Theorem. We call this μ an infinite-volume Gibbs measure. Note that this μ will satisfy the DLR condition.

Remark 3.7. *Notice that this definition does not make reference to limits because the function T does not need to come from a limit.*

While readers familiar with conditional probability will smell that the DLR condition has something to do with conditional probability, it is a little tricky to write it in the language of conditional probability just yet. To do this formally, we appeal to the language of probability kernels and specifications. We kindly refer readers unfamiliar with these concepts to Appendix B below'.

3.3 Gibbsian Specification

Consider the collection of probability kernel $\{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ where each $\pi_\Lambda : \mathcal{F} \times \Omega \to [0,1]$ is defined by:

$$\pi_\Lambda(A|\psi) \equiv \int_{\psi \in A} \frac{\exp\left(-\beta H_\Lambda(\psi)\right)}{Z_{\Lambda,\varphi}} d\left(\prod_{i\in\Lambda} \sigma\right)(\psi) = \mu_{\Lambda,\varphi}(A). \tag{7}$$

It is visible from the definition that it is a probability kernel.

We claim that this family is a specification.

Theorem 3.8. $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ *is a specification.*

Proof. We need to check two things: that each π_Λ is a proper probability kernel and that for every $\Delta \subset \Lambda \Subset \mathbb{Z}^d$, $\pi_\Lambda \pi_\Delta = \pi_\Lambda$.

To first show that each π_Λ is proper, fix a finite subset $\Lambda \Subset \mathbb{Z}^d$ and show that for every $B \in \mathcal{F}_{\Lambda^c}$, $\pi_\Lambda(B|\psi) = \mathbb{1}_B(\psi)$. This is very easy to see because

$$\begin{aligned}
\pi_\Lambda(B|\psi) &= \mu^{d,N}_{\beta,h,L,\varphi}(B) \\
&= \int_{\eta_\Lambda \in \Omega_\Lambda} \mathbb{1}_B(\eta_\Lambda \psi_{\Lambda^c}) \frac{\exp\left(-\beta H_\Lambda(\eta)\right)}{Z^{d,N}_{\beta,h,L,\varphi}} \left(\prod_{i\in\Lambda} \sigma\right)(d\eta) \\
&= \mathbb{1}_B(\psi) \int_{\eta_\Lambda \in \Omega_\Lambda} \frac{\exp\left(-\beta H_\Lambda(\eta)\right)}{Z^{d,N}_{\beta,h,L,\varphi}} \left(\prod_{i\in\Lambda} \sigma\right)(d\eta) \\
&= \mathbb{1}_B(\psi)
\end{aligned}$$

where the second equality follows from the fact that $B \in \mathcal{F}_{\Lambda^c}$.

Next, we show that for every $\Delta \subset \Lambda \Subset \mathbb{Z}^d$, $\pi_\Lambda \pi_\Delta = \pi_\Lambda$. However, this is just the same as DLR condition applied to indicator functions. $\square$

This theorem above, along with Theorem B.8, allows us to relate the finite Gibbs measure to conditional probability. Moreover, this above theorem allows us to define infinite-volume Gibbs measure in even more abstract language.

Definition 3.9 (Infinite-Volume Gibbs Measure)**.** *We call probability measure $\mu \in M_1(\Omega)$ an infinite-volume Gibbs measure if it is compatible with specification given above $\{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$, i.e. for every $\Lambda \Subset \mathbb{Z}^d$, $\mu\pi_\Lambda = \mu$.*

Remark 3.10. *It is easy to check that definition above corresponds to Definition 3.6.*

Using notation from Definition B.7, Let us denote set of all infinite-volume Gibbs measures via $\mathcal{G}(\pi)$. Then, proving existence of infinite-volumes Gibbs measure can be rephrased to showing that $\mathcal{G}(\pi) \neq \emptyset$ and showing the uniqueness can be rephrased to showing that $|\mathcal{G}(\pi)| = 1$.

Since our goal is to prove the existence of an infinite-volume Gibbs measure, we show that $\mathcal{G}(\pi) \neq \emptyset$. However, to do this, we first need to deal with a bit of topology.

3.4 Topology of Ω and $M_1(\Omega)$.

From the very beginning, we gave product topology on Ω. Because each S^{N-1} is compact, by Tychonoff's theorem, we know that Ω is a compact space.

Lemma 3.11. *Ω is metrizable with distance function given by*

$$d(x,y) = \sum_{i\in\mathbb{N}} \frac{\|x_i - y_i\|}{2^{i+1}} \tag{8}$$

where we enumerate $\mathbb{Z}^d$. One can trivially check that the d indeed is a distance function.

It turns out that thinking of continuous functions in terms of $\epsilon - \delta$ definition using this metric is much simpler. We will use this metric on Ω to prove Lemma 3.13 and Theorem 3.21.

Remark 3.12. *Consider $N = 1$ case. This is a fun mental exercise that I really enjoyed doing (and is not relevant to the rest of the section). Notice that Ω in this case is homeomorphic to the Cantor set in $[0, 1]$. Therefore, the pullback of the Gibbs measure we give on Ω to the Cantor set and trivially embedding it to $[0, 1]$ will generate a singular continuous measure with respect to the Lebesgue measure on $[0, 1]$.*

Now, before we define topology on $M_1(\Omega)$, the set of all measures on $(\Omega, \mathscr{F})$, we prove a lemma. Before that, however, let us denote set of all continuous functions on Ω by $C(\Omega)$. Because Ω is compact, we know that $C(\Omega) = C_c(\Omega) = C_0(\Omega)$, i.e. all continuous functions on Ω are bounded and have compact support.

Lemma 3.13. *Local functions are dense in $C(\Omega)$ equipped with the sup norm.*

Proof. Given $\epsilon > 0$ and $g \in C(\Omega)$, one can find $\Lambda \Subset \mathbb{Z}^d$ such that for every $\psi, \varphi \in \Omega$ with $\psi_\Lambda = \varphi_\Lambda$, $|g(\psi) - g(\varphi)| < \epsilon$. We can find such a Λ because g is continuous and because the metric on Ω is defined such that for big enough index (so outside of some finite set) whether or not two configurations differ on that coordinate is negligible.

Now pick your favorite $\tilde{\psi} \in \Omega$. Define $f(\psi) = g(\psi_\Lambda \tilde{\psi}_{\Lambda^c})$. By construction, f is local and for every $\psi \in \Omega$,

$$|f(\psi) - g(\psi)| < \epsilon.$$

Therefore, we have that

$$||f - g||_\infty < \epsilon,$$

proving the lemma. □

We also prove a lemma about measures in $M_1(\Omega)$.

Lemma 3.14. *Given $\mu, \nu \in M_1(\Omega)$, the following are equivalent.*

- $\mu = \nu$,
- $\mathbb{E}_\mu(f) = \mathbb{E}_\nu(f)$ *for every f local,*
- *and* $\mathbb{E}_\mu(g) = \mathbb{E}_\nu(g)$ *for every $g \in C(\Omega)$.*

The proof is immediate from the Lemma 3.13.

Now we endow a topology on $M_1(\Omega)$. We define the topology by specifying the mode of convergence.

Definition 3.15. *We say $\mu_n \to \mu$ in $M_1(\Omega)$ if for every $A \in \mathscr{F}$, $\mu_n(A) \to \mu(A)$.*

This topology is exactly the weak topology, i.e., integrals of continuous functions converge; Below we characterize equivalent ways to define convergence on $M_1(\Omega)$.

Lemma 3.16. *Suppose $\{\mu_n\} \subset M_1(\Omega)$ and $\mu \in M_1(\Omega)$. Then, the following are equivalent.*

- $\mu_n \to \mu$,
- $\mathbb{E}_{\mu_n}(f) \to \mathbb{E}_\mu(f)$ *for every f local function,*
- $\mathbb{E}_{\mu_n}(g) \to \mathbb{E}_\mu(g)$ *for every $g \in C(\Omega)$.*

The proof again is immediate from the Lemma 3.13.

Now, we show that $M_1(\Omega)$ is compact in this topology, and hence sequentially compact thanks to the space being metrizable. To do this, we appeal to the Banach-Alaoglu theorem, a classical theorem from functional analysis.

Theorem 3.17 (Banach-Alaoglu)**.** *Suppose X is a Banach space. Then the closed unit ball on X^* under weak-$*$ topology is compact.*

For the proof of it, we refer the readers to [12] or [14]. Now, we prove that $M_1(\Omega)$ is compact.

Theorem 3.18. *$M_1(\Omega)$ is compact under the topology specified via Definition 3.15.*

Proof. Identify every $\mu \in M_1(\Omega)$ with functional $T_\mu \in C(\Omega)^*$ such that $T_\mu(f) = \int_\Omega f(\psi)d\mu(\psi)$ using KMR, i.e. Theorem 3.5. Observe that $C(\Omega)$ is a Banach space and that $\mu_n \to \mu$ iff $T_{\mu_n} \to T_\mu$ in weak-$*$ topology, thanks to Lemma 3.16. This shows that identification that KMR is in fact a homeomorphism.

Moreover, notice that because every $\mu \in M_1(\Omega)$ is a probability measure, $M_1(\Omega) \subset \bar{B}_1^*$, i.e. $M_1(\Omega)$ can be identified with a subset of the closed unit ball in $C(\Omega)^*$. Since the limit of a linear functionals (hence in particular those of the form T_{μ_n}) is a functional, and since by KMR we can identify the limit with an element in $M_1(\Omega)$, we have that $M_1(\Omega)$ identifies with a closed subset of $\bar{B}_1^*$. Now, using Banach-Alaoglu and the fact that a closed subset of every compact set is compact, we have that $M_1(\Omega)$ is compact, as desired. □

Now, we go on to prove the existence of an infinite-volume Gibbs measure.

3.5 Existence of Infinite-Volume Gibbs Measure

First, a definition.

Definition 3.19 (Quasilocal Specification)**.** *We say that a specification $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is quasilocal if for every $\Lambda \Subset \mathbb{Z}^d$ and $A \in \mathcal{F}$, the map $\psi \mapsto \pi_\Lambda(A|\psi)$ is continuous.*

The reason why we call these specifications quasilocal is because they are continuous and as seen in Lemma 3.13, local functions are dense in the set of continuous functions, thus 'quasilocal'.

We now prove a small fact about quasilocal specifications.

Lemma 3.20 (Feller Property)**.** *Let $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ be a quasi-local specifications. Then, for every $\Lambda \Subset \mathbb{Z}^d$ and for every $f \in C(\Omega)$, $\pi_\Lambda f \in C(\Omega)$ where $\pi_\Lambda f$ is defined in (78).*

Proof. It is visible that for every $A \in \mathcal{F}$ and $\alpha \in \mathbb{R}$, $\pi_\Lambda(\alpha \mathbb{1}_A)(\psi) = \int_\Omega \alpha \mathbb{1}_A(\eta)\pi_\Lambda(d\eta|\psi)$ is continuous. Therefore, a linear combination of indicator functions applied to π_Λ is continuous. Now given $f \in C(\Omega)$, approximate f with linear combinations of indicator functions, say f_n. Then,

$$\begin{aligned}
||\pi_\Lambda(f_n) - \pi_\Lambda(f)||_\infty &= \sup_{\psi\in\Omega} |\int_\Omega (f_n(\eta) - f(\eta))\pi_\Lambda(d\eta|\psi)| \\
&\le ||f_n - f||_\infty \sup_{\psi\in\Omega} |\int_\Omega \pi_\Lambda(d\eta|\psi)| \\
&= ||f_n - f||_\infty.
\end{aligned}$$

Thus $\pi_\Lambda f_n \to \pi_\Lambda f$ uniformly. Since uniform limit of continuous functions is continuous, we have that $\pi_\Lambda f \in C(\Omega)$. □

We now show that the specification associated with each finite-volume Gibbs measure is quasilocal.

Theorem 3.21. *Specification $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ defined via $\pi_\Lambda(A|\psi) = \mu^{d,N}_{\beta,h,L,\varphi}(A)$ is a quasilocal specification.*

Proof. Given $\Lambda \Subset \mathbb{Z}^d$, let us enumerate $\mathbb{Z}^d$ and find M for every $i \in \partial\Lambda$, $i < M$ in the enumeration of $\mathbb{Z}^d$. Now given $\epsilon > 0$, let $\delta = \epsilon/2^{2M}$. Then, for every $\psi, \varphi \in \Omega$ such that $d(\psi, \varphi) < \delta$, we know that the first M coordinates of ψ and φ must be the same by definition of distance on Ω as in (8). Therefore, $\psi_{\partial\Lambda} = \varphi_{\partial\Lambda}$ by choice of M and this means $|\pi_\Lambda(A|\psi) - \pi_\Lambda(A|\varphi)| = 0 < \epsilon$. Therefore, we have proved that $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is quasi-local. This proof works for all $O(N)$ model. □

Theorem 3.22. *If $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is quasi-local, then $\mathcal{G}(\pi) \neq \emptyset$. This implies that infinite-volume Gibbs measure exists.*

Proof. Fix your favorite $\varphi \in \Omega$. Define measures $\mu_n(\cdot) = \pi_{\Lambda_n}(\cdot|\varphi)$ where $\Lambda_n \equiv [-n, n]^d$. Given $\Lambda \Subset \mathbb{Z}^d$ since Λ is finite, there exists some n_0 such that for every $n \geq n_0$, $\Lambda \subset \Lambda_n$. Therefore, we have thanks to the consistency condition (part of the definition of specification)

$$\mu_n \pi_\Lambda(\cdot) = \pi_{\Lambda_n} \pi_\Lambda(\cdot|\varphi) = \pi_{\Lambda_n}(\cdot|\varphi) = \mu_n. \tag{9}$$

By Theorem 3.18, we have that $M_1(\Omega)$ is sequentially compact and hence we can find subsequence $(\mu_{n_k})_{k \in \mathbb{N}}$ such that $\mu_{n_k} \to \mu$ for some $\mu \in M_1(\Omega)$. We claim that $\mu \in \pi$ and hence is an infinite Gibbs measure.

First, since π is quasilocal, we know that for every $\Lambda \Subset \mathbb{Z}^d$ and $f \in C(\Omega)$, by Lemma 3.20, $\pi_\Lambda f \in C(\Omega)$. Then, we get that

$$\mathbb{E}_{\mu\pi_\Lambda}[f] = \mathbb{E}_\mu[\pi_\Lambda f] = \lim_{k\to\infty} \mathbb{E}_{\mu_{n_k}}[\pi_\Lambda f] = \lim_{k\to\infty} \mathbb{E}_{\mu_{n_k}\pi_\Lambda}[f] = \lim_{k\to\infty} \mathbb{E}_{\mu_{n_k}}[f] = \mathbb{E}_\mu[f].$$

Above, the first and the third equality follow from Lemma B.6; the second and the last equality follow from Lemma 3.16 and that $\mu_{m_k} \to \mu$; while the fourth equality follows from (9). We remark that in the second equality is where we needed $\pi_\Lambda f \in C(\Omega)$. With this, we have shown that $\mu\pi_\Lambda = \mu$ for every $\Lambda \Subset \mathbb{Z}^d$ thanks to Lemma 3.14. This means $\mu \in \mathcal{G}(\pi)$ and hence μ is an infinite-volume Gibbs measure. □

3.6 Why not Kolmogorov's Extension Theorem?

Now that we have shown the existence of an infinite-volume Gibbs measure, let us meditate on whether all of our hard work was actually necessary. One way to reflect on our hard work is by testing whether 'easy' extensions succeed or not in giving what we wanted. One of the easy and well-known extension techniques is Kolmogorov's Extension Theorem, also known as Kolmogorov existence theorem, the Kolmogorov consistency theorem, or the Daniell-Kolmogorov theorem. The theorem goes as follows in this context.

Theorem 3.23 (Kolmogorov's Extension Theorem)**.** *Let $\{\mu_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ such that each μ_Λ is a probility measure on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$. Moreover, suppose they satisfy the consistency condition in the sense that for every $\Lambda \Subset \mathbb{Z}^d$ and for every $\Delta \subset \Lambda$,*

$$\mu_\Delta = \mu_\Lambda \cdot (\rho_\Delta)^{-1}$$

where $\rho_\Delta : \Omega_\Lambda \to \Omega_\Delta$ is the natural projection map. Then, there exists a unique $\mu \in M_1(\Omega)$ such that $\mu \mid \Lambda = \mu_\Lambda$ for every $\Lambda \Subset \mathbb{Z}^d$.

The reason why we cannot use this extension theorem simply is that the finite-volume Gibbs measures do not satisfy the assumption that Kolmogorov's Extension Theorem requires. For a concrete example, we refer the readers to Example 7.72 of [15].

3.7 Remark on Uniqueness

Let us wrap up this section with a remark on the uniqueness of the infinite-volume Gibbs measure. It is important to observe that a priori, we *do not know* whether an infinite-volume Gibbs measure is unique. In particular, the part in the construction that introduces potential non-uniqueness is when we fix 'our favorite boundary condition φ' in the proof of Theorem 3.22.

Over the years, people have come up with a uniqueness criterion that tells us when we have the uniqueness of an infinite-volume Gibbs measure. One of them is Dobrushin's Uniqueness Theorem, which can be found in Section 6.5 of

[7]. The proof includes looking at total variation distance on $M_1(\Omega)$, or equivalently, thanks to Theorem 3.5, operator norm on $C(\Omega)^*$.

With this, we end our discussion on the existence of an infinite-volume Gibbs measure in $O(N)$ model setting.

4 Zach: Correlation inequalities: Ginibre, FKG, GHS

4.1 Griffiths inequality

4.2 FKG inequality

5 Kashti: Full solution in $d = 1$ via transfer matrices

5.1 The Ising Model, A Primer (FV 1.4.2) [7]

The general theory of the Ising model was developed by Wilhelm Lenz in 1920 in response to he question of weather phase transitions could be described within statistical mechanics. It ended up being the first system of locally interacting units where it was possible to prove the existence of phase transitions.

We can model the crystalline structure corresponding to atoms in a magnet using a finite non-oriented **graph** $G(\Lambda)$, whose of set of vertices Λ is a subset of $\mathbb{Z}^n$. An example is the box of radius 2, where a **box** is defined $\mathrm{B}(n) = \{-n, ..., n\}^d$, shown in Figure 1.1. We will consider the edge between two vertices i, j with $||j - i||_1 = 1$ where we define the norm as $||i||_1 = \sum_{d=1}^{k} |i_k|$. We use $i \backsim j$ to denote two edges as nearest neighbors. Thus, the set of edges in a box are denoted by $\{\{i, j\} \subset B(n) : i \backsim j\}$.

A key point of the **Ising model** is that it assumes a spin is located at each vertex of the graph G, where the spins are restricted to one direction $+1$ for up and -1 for down at each vertex i. A **configuration** is denoted by $\omega \in \Omega_\Lambda$ where

$$\Omega_\Lambda = \{-1, +1\}^\Lambda.$$

5.1.1 Finite Volume Gibbs Distributions (FV 3.1) [7]

For a finite volume $\Lambda \subset \mathbb{Z}^d$ the configurations of the Ising model are given by elements of the set

$$\Omega_\Lambda = \{S^{N-1}\}^\Lambda.$$

Definition 5.1. *At a vertex $i \in \mathbb{Z}^d$, spin is the basic random variable associated to the model. More generally it is the random variable $\sigma_i : \Omega_\Lambda \to S^N$ where we define $\sigma_i(\psi) = \psi_i$.*
Concretely think of it as taking a configuration of the spins for a given box and outputting the spin for a given vertex i.

We identify a finite set Λ with a graph containing all edges by the nearest neighbor pairs of the vertices of Λ. Here we denote the set of edges by

$$\mathcal{E}_\Lambda = \{\{i, j\} \subset \Lambda : i \backsim j\}$$

For each configuration $\psi \in \Omega_\Lambda$ we can associate an energy.

Definition 5.2.

$$H(\psi) = \frac{1}{2} \langle \psi, -\Delta\psi \rangle - \langle h, \psi \rangle$$

where $h \in \mathbb{R}$ is the magnetic field.

Moving forward, we will assume there is no magnetic field meaning that the h term is 0. Thus:

$$H(\psi) = \frac{1}{2} \langle \psi, -\Delta\psi \rangle$$

Now we define having periodic boundary conditions.

Definition 5.3. *The Gibbs distribution of the Ising model in a torus with n vertices denoted V_n with periodic boundary condition at parameter inverse temperature $\beta \in \mathbb{R}^+$ on $\{-1,1\}^{V_n}$ defined by*

$$\mu(\psi) = \frac{1}{\boldsymbol{Z}} exp(-\beta H(\psi))$$

The normalization constant

$$\boldsymbol{Z} = \sum_{\psi \in \Omega_\Lambda} exp(-\beta H)$$

is called the partition function in Λ with periodic boundary condition.

More generally, we define the partition function as

$$Z = \int_{\psi \to S^{N-1}} exp(-\beta H) d\nu(\psi) = \int_{\Omega_\Lambda} exp(-\beta H) d\nu(\psi)$$

We note that our measure ν is usually just the uniform product measure on S^{N-1} (ie uniform measure on the circle, sphere,...), but we then take the product over all vertices $|\Lambda|$, and that is our new measure (the product measure over all the uniform measures). Thus

$$d\nu(\psi) = \otimes_{i \in |\Lambda|} d\nu_{S^{N-1}} \psi_i.$$

To see this agrees with the Ising measure version, note that in that case we simply that ν to be the counting measure.

The partition function "adds up" the Boltzmann weights $\exp(-\beta H)$ *for every possible spin configuration across all vertices of the lattice.*

5.1.2 Deriving The Two Point Function and Transfer Matrix

At each site $j \in \{1, ..., n\}$ configuration $\omega = (s_1, ..., s_n) \in \Omega_\Lambda = (S^{N-1})^n$. We note the Gibbs weight of our configuration is

$$\exp(-\beta H(\omega)) = \prod_{j=1}^{n} e^{\beta s_j \cdot s_{j+1}}$$

with our defined product measure. Then the partition function can be written as

$$Z_n(\beta) = \int_{\Omega_\Lambda} e^{-\beta H(\omega)} d\nu(\omega) = \int_{(S^{N-1})^n} \prod_{j=1}^{n} e^{\beta s_j \cdot s_{j+1}} \prod_{j=1}^{n} d\nu_{S^{N-1}}(s_j)$$

Now we define the bond kernel $K : S^{N-1} \times S^{N-1} \to \mathbb{R}$, where explicitly

$$K(s, s') := e^{\beta s \cdot s'}$$

Then we have that the nearest-neighbor Boltzmann factors as

$$\prod_{j=1}^{n} e^{\beta s_j \cdot s_{j+1}} = \prod_{j=1}^{n} K(s_j, s_{j+1})$$

This is the key point: the Gibbs weight is a product of identical two-site kernels which dependent ONLY on adjacent spins. In a single dimension, this factorization is exactly what allows up to define the full transfer operator (in higher dimension there is no single linear ordering which gives us this kind of single chain factorization of the full Gibbs weight).

Now we define the transfer operator T on $L^2(S^{N-1}, d\nu_{S^{N+1}})$ using the integral kernel:

$$(Tf)(s) = \int_{S^{N-1}} K(s, s') f(s') d\nu_{S^{N-1}}(s')$$

Equivalently, T is the integral operator with kernel $K(s, s')$. In the discrete Ising case ($N = 1$), T simply reduces to the finite matrix with entries $T_{ss'} = K(s, s')$.

Now we compute T^n kernel iteratively. We use $K^m(x, y)$ to be the kernel of T^m. Then :

$$K^{(m)}(x,y) = \int_{(S^{N-1})^{m-1}} K(x,x_1)K(x_1,x_2)...K(x_{m-1},y) \prod_{i=1}^{m-1} d\nu_{S^{N-1}}(x_i)$$

This gives us

$$Z_n(\beta) = \int_{(S^{N-1})^n} \prod_{j=1}^{n} K(s_j,s_{j+1}) \prod_{j=1}^{n} d\nu_{S^{N-1}}(s_j)$$

$$= \int_{S^{N-1}} K^n(x,x)d\nu_{S^{N-1}}(x) = \mathrm{Tr}(T^n)$$

Note that for finite S^{N-1} this is the matrix identity $\mathrm{Tr}(T^n) = \sum_i (T^n)_{ii}$.

Now we fix two sites 0 and r where r is the distance between them. The two point function is just then simply

$$\mathbb{E}[s_0 \cdot s_r] = \frac{1}{Z_n} \int_{(S^{N-1})^n} (s_0 \cdot s_r) \prod_{j=1}^{n} K(s_j,s_{j+1}) \prod_{j=1}^{n} d\nu_{S^{\nu-1}}(s_j)$$

Now we partition the product of kernels into three blocks (two below and one which is the $(s_0 \cdot s_r)$):

$$(K(s_0,s_1)...K(s_{r-1},s_r))(K(s_r,s_{r+1})...K(s_n,s_0))$$

and integrate over the intermediate spins. This gives the multiplication operator A on $L^2(S^{N-1})$ given by

$$(Af)(s) = \Psi(s)f(s)$$

where $\Psi(s)$ for $\Psi : S^{N-1} \to \mathbb{R}$ is a scalar function representing the component of spin)(the component of the spin at site s).

For the dot product observable, we write

$$s_0 \cdot s_r = \sum_{a=1}^{N} \Psi_a(s_0)\Psi_a(s_r)$$

Now we compute the numerator of $\mathbb{E}[s_0 \cdot s_r]$. The argument below works for a fixed component Ψ and extends by linearity to the sum.

$$\mathcal{N} = \int_{S^{N-1}} \int_{S^{N-1}} \Psi(s_0)\Psi(s_r)K^{(r)}(s_0,s_r)K^{(n-r)}(s_r,s_0)d\nu(s_0)d\nu(s_r)$$

Now consider the product operator T^rA which has kernel $L^{(r)}(x,y)$ that

$$(T^rAf)(x) = \int_{S^{N-1}} K^{(r)}(x,y)(Af)(y)d\nu(y) = \int_{S^{N-1}} K^{(r)}(x,y)\Psi(y)f(y)d\nu(y)$$

Thus the kernel is of the form

$$L^{(r)}(x,y) = K^{(r)}(x,y)\Psi(y)$$

Now we take the next form $T^{n-r}A = T^{n-r} \circ A$. The kernel of $T^{n-r}A$ by the same reasoning is.

$$\tilde{L}^{(n-r)}(x,y) = K^{(n-r)}\Psi(y)$$

Now we compute the kernel of the full product $(T^{n-r}A) \circ (T^rA)$ is given by the usual kernel composition formula

$$(T^{n-r}AT^rA)(x,y) = \int_{S^{N-1}} \tilde{L}^{(n-r)}(x,z)L^{(r)}(z,y)d\nu(z)$$

$$= \int_{S^{N-1}} (K^{(n-1)}(x,z)\Psi(z))(K^{(r)}(z,y)\Psi(y))d\nu(z)$$

$$= \Psi(y) \int_{S^{N-1}} K^{(n-r)}(x,z)K^{(r)}(z,y)\Psi(z)d\nu(z)$$

Now finally, we take the trace

$$\mathrm{Tr}(T^{n-r}AT^rA) = \int_{S^{N-1}} \Psi(x)(\int_{S^{N-1}} K^{(n-r)}(x,z)K^{(r)}(z,x)\Psi(z)d\nu(z))d\nu(x)$$

Then we swap the order of integration by Fubini and we rename our dummy variables ($x \mapsto s_0, z \mapsto s_r$)

$$\mathrm{Tr}(T^{n-r}AT^rA) = \iint_{S^{N-1}\times S^{N-1}} \Psi(s_0)\Psi(s_r)K^{(n-r)}(s_0,s_r)K^r(s_r,s_0)d\nu(s_r)d\nu(s_0)$$

Now we note that scalar product commutes so we have

$$\mathrm{Tr}(T^{n-r}AT^rA) = \iint \Psi(s_0)\Psi(s_r)K^{(r)}(s_0,s_r)K^{(n-r)}(s_r,s_0)d\nu(s_0)d\nu(s_r) = \mathcal{N}$$

Finally, we divide by $Z = \mathrm{Tr}(T^n)$.

Therefore the two-point function equals the trace-with-insertions divided by the partition function:

$$\boxed{\mathbb{E}[s_0 \cdot s_r] \;=\; \frac{\mathrm{Tr}\big(T^{N-r} A\, T^{r} A\big)}{\mathrm{Tr}\big(T^{N}\big)}}$$

5.1.3 Essential Background

If we suppose $\lambda_0 > \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ...$ are the eigenvalues of T, by the positivity of the kernel, λ_0 is strictly larger than the next eigenvalue in magnitude for $\beta > 0$. Note such λ_0 must exist as the transfer operator is compact, self-adjoint, and positivity improving as its kernel $e^{\beta\psi\cdot\psi'}$ is strictly positive. Thus by Perron–Frobenius, such a unique largest eigenvalue must exist.

Then we diagonalize and insert resolutions of the identity into the two point function

$$\mathbb{E}[s_0 \cdot s_r] = \frac{\mathrm{Tr}(ST^rST^{N-r})}{\mathrm{Tr}(T^n)}$$

In the limit as $N \to \infty$, this sum becomes proportional to $(\frac{\lambda_1}{\lambda_0})^r$. The leading nontrivial contribution for a seperation r large comes from the first sub-leading eigenvalue

$$\mathbb{E}[s_0 \cdot s_1] \sim C(\frac{\lambda_1}{\lambda_2})^r$$

as $r \to \infty$. Thus the correlations decay exponentially giving us the decay rate (which we call the inverse correlation length):

$$\zeta^{-1} = \log\frac{\lambda_0}{\lambda_1}$$

5.1.4 N=1 (Ising Spins ± 1)

For the 1-D periodic chain having L vertices $i = 1, .., L$ and edges $\mathcal{E} = \{|[i, i+1\} : i = 1, ..., L\}$ where $i+1$ is taken mod L. Our spins are given by $\psi_i \in \{\pm 1\}$. Now we use the discrete Laplacian Convention

$$(-\Delta\psi)_i = \sum_{j:|\{i,j\}\in\mathcal{E}} (\psi_i - \psi_j)$$

Then we compute the quadratic form

$$\langle\psi, -\Delta\psi\rangle = \sum_i \psi_i \sum_{j\sim i}(\psi_i - \psi_j) = \sum_{\{i,j\}\in\mathcal{E}} (\psi_i(\psi_i - \psi_j) + \psi_j(\psi_j - \psi_i))$$

we simplify the inside terms

$$\psi_i(\psi_i - \psi_j) + \psi_j(\psi_j - \psi_i) = \psi_i^2 + \psi_j^2 - 2\psi_i\psi_j = 2 - 2\psi_i\psi_j$$

Thus we obtain

$$\frac{1}{2}\langle\psi, -\Delta\psi\rangle = \sum_{\{i,j\}\in\mathcal{E}} (1 - \psi_i\psi_j) = |\mathcal{E}| - \sum_{\{i,j\}\in\mathcal{E}} \psi_i\psi_j$$

. This means that in a uniform magnetic field h, our Hamiltonian is of the form

$$H(\psi) = \frac{1}{2}\langle \psi, -\Delta\psi\rangle = L - \sum_{i=1}^{L} \psi_i \psi_{i+1}$$

and the partition function at inverse temperature β is

$$Z_L(\beta) = \sum_{\psi \in \{\pm 1\}^L} e^{-\beta H(\psi)} = e^{-\beta L} \sum_{\psi} \exp(\beta \sum_{i=1}^{L} \psi_i \psi_{i+1})$$

Now using the indices $s, s' \in \{-1, +1\}$ we defined the transfer matrix

$$T_{s,s'} = \exp(\beta s s')$$

Then the sum over spins factorizes into a product of T entries and we get the matrix

$$\begin{bmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{bmatrix}$$

Note that going forward we will keep the multiplicative $e^{-\beta L}$ constant in mind but drop it in calculations as it does not effect the ratio $\frac{\lambda_0}{\lambda_1}$.

Now we compute the eigenvalues of T. Let us denote $e^{\beta} = a$ and $e^{-\beta} = b$.

Then

$$\det(T - \lambda I) = (a - \lambda)(a - \lambda) - b^2 = 0$$

so

$$\lambda^2 - (2a)\lambda + (a^2 - b^2) = 0$$

. We then compute

$$2a = e^{\beta} + e^{\beta} = 2e^{\beta}$$

and

$$a - a = 0$$

additionally with $b^2 = e^{-2\beta}$ and $a^2 = e^{2\beta}$. Now we introduce a term

$$\Delta = \sqrt{(\frac{a-a}{2})^2 + b^2} = \sqrt{e^{-2\beta}} = e^{-\beta}$$

giving us our eigenvalues

$$\lambda_{\pm} = e^{\beta} \pm \Delta = e^{\beta} \pm e^{-\beta}$$

and finally, we order them by $\lambda_0 = \lambda_+$ and $\lambda_1 = \lambda_-$ for $\beta > 0, \lambda_0 > \lambda_1 > 0$

Now let $S = \text{diag}(1, -1)$. The two point function is given by

$$\mathbb{E}[s_0 \cdot s_r]_L = \frac{\text{Tr}(ST^r S T^{L-r})}{\text{Tr}(T^L)}$$

Now we diagonalize the sum $T = \sum_{k=0,1} \lambda_k \Pi_k$ (Π_k being the rank 1 projectors). Now expand the numerator and denominator giving a finite sum of terms of the form $\lambda_i^r \lambda_j^{L-r}$. Thus concretely we get

$$\mathbb{E}[s_0 \cdot s_r]_L = \frac{\sum_{i,j \in \{0,1\}} c_{ij} \lambda_i^r \lambda_j^{L-r}}{\lambda_0^L + \lambda_1^L}$$

for some constants c_{ji} independent of L, r. Now we take the limit $L \to \infty$. First divide the numerator and denominator by λ_0^L Since $|\lambda_0/\lambda_1| > 1$ for $\beta > 0$, all terms with a factor $(\lambda_1/\lambda_0)^L$ vanish in the limit. The surviving leading $L \to \infty$ contribution for a fixed r is

$$\mathbb{E}[s_0 \cdot s_r] = c(\frac{\lambda_1}{\lambda_0})^r$$

where c is a nonzero amplitude coming for eigenvector overlaps. In particular, we have

$$\mathbb{E}[s_0 \cdot s_r] \propto (\frac{\lambda_1}{\lambda_0})^r.$$

Thus the exponential decay is proven and the decay factor per step is the eigenvalue ratio λ_1/λ_0.

Now, the decay ratio explicity can be given by

$$\frac{\lambda_1}{\lambda_0} = \frac{e^{\beta} - e^{-\beta}}{e^{\beta} + e^{-\beta}} = \frac{2\sinh\beta}{2\cosh\beta} = \tanh\beta$$

.

Now we take $\beta \to \infty$ to get

$$\frac{\lambda_1}{\lambda_0} \to 1$$

From below, which means for any given temperature, this decays exponentially.

5.2 N=2 (XY)

In the XY model, spins are unit vectors in the plane, where we will denote as $\psi_j = e^{i\theta_j}$. Here the discrete Laplacian acts as

$$(\Delta\psi)_j = \psi_{j+1} + \psi_{j-1} - 2\psi_j$$

Thus we have

$$\frac{1}{2}\langle\psi, -\Delta\psi\rangle = \frac{1}{2}\sum_j \overline{\psi_j}(2\psi_j - \psi_{j+1} - \psi_{j-1})$$

$$= \sum_j |\psi_j|^2 - \frac{1}{2}\sum_j(\overline{\psi_j}\psi_{j+1} + \overline{\psi_{j+1}}\psi_j)$$

$$= \sum_j |\psi_j|^2 - \sum_j \mathrm{Re}(\overline{\psi_j}\psi_{j+1]})$$

If $|\psi_j| = 1$, our first term $\sum_j |\psi_j|^2$ will simply be a constant independent of the confirmation; drop constants in the Gibbs weight. Thus up to a constant, we have

$$H(\{\theta_j\}) = -\sum_j \mathrm{Re}(\overline{\psi_j}\psi_{j+1}) = -\sum_j \cos(\theta_{j+1} - \theta_j)$$

Note that when $h = 0$ the magnetization vanishes because of symmetry and so the full two point function equals the connected piece. Now we provide more detail:

Now our Hamiltonian is of the form

$$H(\psi) = \frac{1}{2}\langle\psi, -\Delta\psi\rangle$$

As we have seen for spins $\psi_j = e^{i\theta_j} \in S^1$, that

$$H(\{\theta_j\}) + \text{constant} = -\sum_{j=1}^{n} \cos(\theta_{j+1} - \theta_j)$$

But we drop the constant so we get with periodic boundary

$$H(\{\theta_j\}) = -\sum_{j=1}^{n} \cos(\theta_{j+1} - \theta_j)$$

Now the Gibbs measure at inverse temperature β is

$$\mu(\{\theta_j\}) \propto \exp(\beta\sum_{j=1}^{n} \cos(\theta_{j+1} - \theta_j))$$

thus the partition function is

$$Z_n = \int_{[0,2\pi]^n} \prod_{j=1}^{n} d\theta_j \prod_{j=1}^{n} e^{\beta\cos(\theta_{j+1}-\theta_j)}$$

Now the transfer operator acting on functions $f \in L^2([0, 2\pi])$ we define as

$$(Tf)(\theta) = \int_0^{2\pi} K(\theta,\theta')f(\theta')d\theta$$

where the kernel is $K(\theta, \theta') = e^{\beta\cos(\theta-\theta')}$ and we note that K is only dependent on $\theta - \theta'$. Thus the partition function can simply be written as the trace

$$Z_n = \mathrm{Tr}(T^n)$$

Now we commute the eigen-functions and eigenvalues. Since $K(\theta, \theta')$ is only dependent on $\theta - \theta'$, T is simply the convolution operator on the circle meaning it can be diagonalized by the Fourier basis.

Take $e_m(\theta) = e^{im\theta}$ for $m \in \mathbb{Z}$. We compute

$$(Te_m)(\theta) = \int_0^{2\pi} e^{\beta\cos(\theta-\theta')}e^{im\theta'}d\theta'$$

Now we do a change of variables and let $\gamma = \theta - \theta'$ so we have $\theta' = \theta - \gamma$ so $d\theta' = -d\gamma$. Thus the integral becomes

$$(Te_m)(\theta) = \int_0^{2\pi} e^{\beta\cos\gamma}e^{im(\theta-\gamma)}d\gamma$$

and we factor out $e^{im\theta}$ getting

$$(Te_m)(\theta) = e^{im\theta}\int_0^{2\pi} e^{b\cos\gamma}e^{-im\gamma}d\gamma$$

So we have e_m is an eigenfunction with eigenvalue

$$\lambda_m = \int_0^{2\pi} e^{b\cos\gamma}e^{-in\gamma}d\gamma$$

This is exactly the modified Bessel function

$$\frac{1}{2\pi}\int_0^{2\pi} e^{b\cos\gamma}e^{-in\gamma}d\gamma$$

so we have

$$\gamma_m = 2\pi I_m(\beta)$$

meaning we have an eigenbasis $\{e_m\}$ with eigen values $\{2\pi I_m(\beta)\}$, with the largest value being the expected λ_0.

Now using transform operator formalism we have

$$\mathbb{E}[e^{i(\theta_0-\theta_r)}] = \frac{1}{Z_n}\mathrm{Tr}(T^{n-r}A^\dagger T^r A)$$

where A is the multiplicative operator $(Af)(\theta) = e^{i\theta}f(\theta)$. Thus

$$Ae_m(\theta) = e^{i\theta}e^{im\theta} = e^{i(m+1)\theta} = e_{m+1}(\theta)$$

and A raises the Fourier index by 1 while $A^\dagger$ lowers it.

Thus A maps $g_m \mapsto g_{m+1}$ for $g_m = e^{im\theta}$ Now use the spectral resolution of T,

$$Tg_m = \lambda_m\, g_m.$$

Then

$$\mathrm{Tr}\big(T^{n-r}A^\dagger T^r A\big) = \sum_m \mathbb{E}[g_m,\, T^{n-r}A^\dagger T^r A g_m].$$

Since $Ag_m = g_{m+1}$ and $Tg_{m+1} = \lambda_{m+1} g_{m+1}$, we obtain

$$\mathbb{E}[g_m,\, T^{n-r} A^\dagger T^r A g_m] = \lambda_m^{n-r}\, \lambda_{m+1}^r.$$

Therefore,

$$\mathrm{Tr}\big(T^{n-r} A^\dagger T^r A\big) = \sum_m \lambda_m^{n-r}\, \lambda_{m+1}^r.$$

Thus we have

$$\mathrm{Tr}(T^{n-r} A^\dagger T^r A) = \sum_{m\in\mathbb{Z}} \lambda_m^{n-r} \lambda_{m+1}^r$$

and then we divide by partition function

$$Z_n = \sum_{m\in\mathbb{Z}} \lambda_m^n$$

Finally in the limit $n \to \infty$. the sum is dominated by the largest eigenvalue λ_0 os

$$\mathbb{E}[e^{i(\theta_0-\theta_r)}] \sim \frac{\lambda_0^{n-r}\lambda_1^r}{\lambda_0^n} = (\frac{\lambda_1}{\lambda_0})^r$$

Note the cosine term is simply the real part so we obtain

$$\mathbb{E}[\cos(\theta_0-\theta_r)] = (\frac{\lambda_1}{\lambda_0})^r = (\frac{I_1(\beta)}{I_0(\beta)})^r = e^{-r/\zeta}$$

where $\zeta = \frac{1}{-\log(\frac{I_1(\beta)}{I_0(\beta)})}$ as needed.

Now we study $\beta \to \infty$. Make the e:

$$\frac{\lambda_1(\beta)}{\lambda_0(\beta)} \approx \frac{I_1(\beta)}{I_0(\beta)} = \frac{\frac{e^\beta}{\sqrt{2\pi\beta}}(1-\frac{3}{8\beta})}{\frac{e^\beta}{\sqrt{2\pi\beta}}(1+\frac{3}{8\beta})}$$

Now we take the limit as $\beta \to \infty$:

$$\frac{\frac{e^\beta}{\sqrt{2\pi\beta}}(1-\frac{3}{8\beta})}{\frac{e^\beta}{\sqrt{2\pi\beta}}(1+\frac{3}{8\beta})} \to 1$$

from below, so we have exponential decay.

5.3 N=3 [8]

We do this with the assumption that the magnetic field is 0. To do this take our Hamiltonain

$$H(\psi) = \frac{1}{2}\langle\psi, -\Delta\psi\rangle$$

On a 1-D with nearest neighbor discrete Laplacian $(\Delta\psi)_j = \psi_{j+1} + \psi_{j-1} - 2\psi_j$, so neglecting the magnetic filed term,

$$\langle\psi, -\Delta\psi\rangle = \sum_j ||\psi_j|^2 - \sum_j \mathrm{Re}(\overline{\psi_j}\psi_{j+1})$$

In this case the spin is constrained to a unit vector $|\psi_j| = 1$. The first term is constant and can be dropped from the Gibbs weight. Thus our Hamilatonain is of the form

$$H(\{\psi_j\}) = -\sum_j \psi_j \cdot \psi_{j+1}$$

For a unit vectors, $s \in S^2$ parametrize relative angle by $\cos\theta = s \cdot s'$. The Boltzmann weight for a bond is given by

$$\exp(\beta J s \cdot s') = \exp(E\cos\theta)$$

where $E = \beta J$

Now we define the transfer operator as an integral operator on $L^2(S^2)$ with kernel

$$(Tf)(s') = \int_{S^2} e^{Es'\cdot s} f(s) d\omega(s)$$

where we have $d\Omega$ the usual solid angle measure on the sphere. As the kernel depends only on $s' \cdot s = \cos\theta$, it is rotationally invariant and symmetric meaning T is self adjoint, compact (Hilbert Schmidt) and positive. The partition function is the operator trace

$$Z_n = \mathrm{Tr}(T^n)$$

As the kernel of the function is $s \cdot s'$ the eigenfunctions are simply the spherical harmonics $Y_{\ell m}(s)$ where the eigenvalue depends on ℓ (with degeneracy $2\ell+1$). Now we expand the kernel in Legendre polynomials. This gives us

$$e^{E\cos\theta} = \sqrt{\frac{\pi}{2E}} \sum_{\ell=0}^{\infty} (2\ell+1) I_{\ell+1/2}(E) P_\ell(\cos\theta)$$

Since the $\{P_\ell\}$ are orthogonal on $[-1,1]$ with

$$\int_{-1}^{1} P_\ell(t) p_m(t) dt = \frac{2}{2\ell+1}\delta_{\ell m}$$

the coefficient a_ℓ in the expansion $e^{Et} = \sum a_\ell P_\ell(t)$ is

$$a\ell = \frac{2\ell+1}{2}\int_{-1}^{1} e^{Et} P_\ell(t) dt$$

This integral is evaluated by the substitution $cos\phi = t$ and reorganizing the standard integral version of the modified Bessel function with half integer order. this gives us $a_\ell = (2\ell+1)\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E)$

According to the addition theorem for spherical harmonics

$$P_\ell(s\cdot s') = \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} Y_{\ell m}(s)\overline{Y_{\ell m}(s')}$$

Now we insert this into our $e^{E\cos\theta}$ term to get

$$e^{Es\cdot s'} = \sum_{\ell=0}^{\infty}(2\ell+1)\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E) P_\ell(s\cdot s')$$

$$= \sum_{\ell=0}^{\infty}(2\ell+1)\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E)\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} Y_{\ell m}(s)\overline{Y_{\ell m}(s')}$$

$$= 4\pi\sum_{\ell=0}^{\infty}\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E)\sum_{m=-\ell}^{\ell} Y_{\ell m}(s)\overline{Y_{\ell m}(s')}$$

Thus the kernel is an explicit sum over the spherical harmonics.

Take standard orthonormal spherical harmonics $Y_{\ell m}$ with

$$\int_{S^2}\overline{Y_{\ell m}} Y_{\ell' m'} d\Omega = \delta_{\ell\ell'}\delta_{mm'}$$

Now we act T on $Y_{\ell' m'}$ to get

$$(TY_{\ell' m'})(s') = \int_{S^2} e^{Es'\cdot s} Y_{\ell' m'} d\Omega(s)$$

$$= 4\pi\sum_{\ell=0}^{\infty}\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E)\sum_{m=-\ell}^{\ell} Y_{\ell m}(s')\int_{S^2}\overline{Y_{\ell m}(s)} Y_{\ell' m'}(s) d\Omega(s)$$

$$= 4\pi\sum_{\ell-0}^{\infty}\sqrt{\frac{\pi}{2E}} I_{\ell+1/2}(E)\sum_{m=-\ell}^{\ell} Y_{\ell m}(S')\delta_{\ell\ell'}\delta_{mm'}$$

$$= 4\pi\sqrt{\frac{\pi}{2E}}I_{\ell+1/2}(E)Y_{\ell' m'}(s')$$

Thus each subspace $\{Y_{\ell m} : m = -\ell, ..., \ell\}$ is an eigenspace with corresponding eigenvalue

$$\lambda_\ell = 4\pi\sqrt{\frac{\pi}{2E}}I_{\ell+1/2}(E)$$

for $\ell = 0, 1, 2, ..$

As T is real, symmetric, and Hilbert Schmidt, the eigenvalues are real, positive and the spectrum is discrete. The degeneracy of λ_ℓ is $2\ell + 1$

From the spectral decomposition

$$\text{Tr}(T^n) = \sum_{\ell=0}^{\infty}(2\ell+1)\gamma_\ell^n = \sum_{\ell=0}^{\infty}(2\ell+1)(4\pi\sqrt{\frac{\pi}{2E}}I_{\ell+1/2}(E))^n$$

In the thermodynamic limit $n \to \infty$ the $\ell = 0$ term dominates as $\lambda_0 > \lambda_1 \geq ...$

Now we want toe two point function

$$C(r) = \langle s_0 \cdot s_r \rangle = \frac{1}{Z_n}\int (s_0 \cdot s_r)e^{E\sum_j s_j \cdot s_{j+1}}\prod_j d\Omega(s_j)$$

Evaluating this we see

$$C_n(r) = \frac{\text{Tr}(T^{n-r}(s)T^r(s))}{\text{Tr}(T^m)}$$

To evaluate this, we write the product

$$s_0 \cdot s - r = P_1(s_0 \cdot s_r) = \frac{4\pi}{3}\sum_{m=-1}^{1} Y_{1m}(S_0)\overline{Y_{1m}(s_r)}$$

note this just the $\ell = 1$ term.

The kernel of T^r

$$K^{(r)}(s_r, s_0) = (T^r)(s_r, s_0) = \sum_{\ell=0}^{\infty}\lambda_\ell^r \sum_{m=-\ell}^{\ell} Y_{\ell m}(s_r)\overline{Y_{\ell m}(s_0)}$$

note that here the projection kernel on the ℓ subspace is

$$\sum_m Y_{\ell m}\overline{Y_{\ell m}(t)}$$

Now the numerator of $C_N(r)$ can be written

$$\mathcal{N} = \int_{S^2}\int_{S^2}(P_1(s_0 \cdot s_r))K^{(r)}(s_r, s_0)d\Omega(s_0)d\Omega(s_r)$$

As integrating out the other spins and using the transfer factorization will leave exactly the double integral (ie the left T^{n-r} will contribute a factor this in the large N limit will be dominated by $\ell = 0$ thus the computation below isolates the ℓ-dependence of the r-step propagation)

Now we insert the two expansions to get

$$\mathcal{N} = \int\int \frac{4\pi}{3}\sum_{m'=-1}^{1} Y_{lm'}(s_0)\overline{Y_{1m'}(s_r)}\sum_{\ell=0}^{\infty}\lambda_\ell^r\sum_{m=-\ell}^{\ell} Y_{\ell m}(s_r)\overline{Y_{\ell m}(s_0)}d\Omega(s_0)d\Omega(s_r)$$

$$= \frac{4\pi}{3}\sum_{\ell=0}^{\infty}\lambda_\ell^m \sum_{m'=-1}^{1}\sum_{m=-\ell}^{\ell}(\int_{S^2} Y_{1m'}(s_0)\overline{Y_{\ell m}(s_0)}d\Omega(s_0))(\int_{S^2} \overline{Y_{1m'}(s_r)}Y_{\ell m}(s_r)d\Omega(s_r))$$

We use orthonormality $\int_{S^2} Y_{LM}\overline{Y_{L'M'}} = \delta_{LL'}\delta_{MM'}$. Each of the integrals above will force $\ell = 0$ amd $m = m'$. Thus for all terms $\ell \neq 1$. Thus only the $\ell = 1$ term survives meaning

$$\mathcal{N} = \frac{4\pi}{3}\lambda_1^r \sum_{m'=-1}^{1}(\int Y_{1m'}(s_0)\overline{Y_{1m'}(s_0)}d\Omega(s_0))(\int \overline{Y_{1m'}(s_r)}Y_{1m'}(s_r)d\Omega(s_r))$$

Each inner integral equals 1 by normalization and the sum over m' has three terms meaning the factor $\frac{4\pi}{3}\cdot 3$ simplifies back to 4π. Thus we have

$$\mathcal{N} = 4\pi\lambda_1^r$$

Now consider the denominator in the large n limit.

$$\text{Tr}(T^n) = \sum_\ell (2\ell+1)\lambda_\ell^N$$

. as $n \to \infty$, this is dominated by $\ell = 0$ meaning

$$\text{Tr}(T^n) = \lambda_0^n(1 + o(1))$$

We now put this together and restore the left T^{n-r} factor which contributes the same dominant λ_0^{n-r} -piece that cancels with the denominator. This gives us in $n \to \infty$ at fixed r

$$\mathbb{E}[s_0 \cdot s_r] = (\frac{\lambda_1}{\lambda_0})^r = (\frac{\sqrt{\frac{\pi}{2E}}I_{3/2}(E)}{\sqrt{\frac{\pi}{2E}}I_{1/2}(E)})^r$$

Using the half integer Bessel closed form

$$I_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sinh x$$

$$I_{3/2}(x) = \sqrt{\frac{2}{\pi x}}(\cosh x - \frac{\sinh x}{x})$$

Thus we have

$$\frac{I_{3/2}(E)}{I_{1/2}(E)} =$$

$$\frac{\cosh(E) - \frac{\sinh E}{E}}{\sinh E} = \coth E - 1/E$$

.

Thus we have

$$\mathbb{E}[s_0 \cdot s_r] = (\coth E - 1/E)^r$$

where $E = \beta$. This is exponential decay in r. Now we define the correlation length ζ by

$$\mathbb{E}[s_0 \cdot s_r] \sim e^{-r/\zeta}$$

meaning we have

$$\zeta^{-1} = -\log(\coth(\beta) - 1/\beta)$$

Now we will study the limit as $\beta \to \infty$, First note that our ratio of

$$\frac{\lambda_1}{\lambda_0} \simeq \coth\beta - 1/\beta \ = \ 1 \ + \ 2e^{-2\beta} \ + \ 2e^{-4\beta} \ + \ O\big(e^{-6\beta}\big) - 1/\beta = 1 - 1/\beta \to 1.$$

From below, again giving us exponential decay for any given β.

6 Sonny: Dobrushin Uniqueness Criterion, Extremal Decomposition, Clustering and Correlation Decay.

6.1 Motivation, Setup, and What a Gibbs Measure Is

Basic spaces and lattice setup

- **Lattice.** We work on the d–dimensional integer lattice $\mathbb{Z}^d$ $(d \geq 1)$. Its elements are called *sites*.
- **Spin space.** Fix a compact Polish space S of *spins* (labels/states available at each site). Typical examples:
  - Ising model: $S = \{\pm 1\}$ (two labels).
  - $O(N)$ model: $S = \mathbb{S}^{N-1} \subset \mathbb{R}^N$ (unit vectors).
- **Configuration space.** A (global) configuration is a function $\omega : \mathbb{Z}^d \to S$. The set of all configurations is

$$\Omega := S^{\mathbb{Z}^d}.$$

- **Locality and finite regions.** For a finite set $\Lambda \Subset \mathbb{Z}^d$, write ω_Λ for the restriction of ω to Λ and $\mathcal{F}_\Lambda$ for the σ–algebra generated by $\{\pi_i : i \in \Lambda\}$. A *local function* depends only on finitely many coordinates, i.e. it is $\mathcal{F}_\Lambda$–measurable for some finite Λ.
- **Outside configuration ("boundary").** For a finite Λ, ω_{Λ^c} denotes the configuration on the complement. We use the notational shorthand $\omega = \omega_\Lambda \omega_{\Lambda^c}$ to emphasize "inside + outside".
- **Neighborhood graph.** When needed, we view $\mathbb{Z}^d$ as a graph (e.g. nearest-neighbor edges), which determines which sites directly interact. None of the definitions below require a specific choice at this point.

What is a Gibbs measure, in plain language?

Think of Ω as all possible labelings of the infinite grid by symbols from S. A *probability measure* on $(\Omega, \mathcal{F})$ tells us how likely each global labeling is.

A **Gibbs measure** is a special kind of probability measure that is *locally consistent*: if you look at any finite window Λ and you are told everything outside the window (ω_{Λ^c}), then the way the measure assigns probabilities to the patterns *inside* the window is determined by a fixed, prescribed local rule (the *specification*). Informally:

> *Whatever the rest of the grid looks like, the conditional law inside any finite window follows the same local recipe that was specified in advance.*

This makes a Gibbs measure the right notion of "equilibrium" for an infinite system: it does not assert a global formula on the entire infinite grid; instead, it enforces a consistent family of conditional rules on all finite windows.

In other words, the Gibbs measure is the mathematical compromise between order (low energy) and disorder, ENTROPY (many possibilities). At low temperature (β large) energy dominates, and the system becomes more ordered; at high temperature (β small) entropy dominates, and configurations become more random.

Why are Gibbs measures important?

They isolate the essential idea of *local coherence*: if you know the outside of a window, there is a standard, agreed-upon way to randomize the inside. This viewpoint is powerful because it:

- focuses on finitely many labels at a time (what we can actually manipulate/prove),
- guarantees that different windows fit together consistently,
- and allows us to talk rigorously about large-scale behavior (e.g., whether there is only one macroscopic law or many).

What is the infinite-volume Gibbs measure?

On a finite region, one can always write down an ordinary probability distribution for the labels inside that region (given an outside labeling). On the full lattice, there is no single "global density" to write down. Instead, we say that a probability measure μ on $(\Omega, \mathscr{F})$ is an **infinite-volume Gibbs measure** (for a given local specification) if, for *every* finite window Λ,

- the conditional distribution of the labels *inside* Λ, given the labels *outside* Λ, matches the prescribed local rule for that window.

In short: an infinite-volume Gibbs measure is any global law on labelings that *respects the same local recipe in every finite window.*

Two key outcomes can happen:

- **Uniqueness.** There is only one such global law. Then the large system behaves in one coherent way and distant regions are weakly related (correlations fade with distance).
- **Non-uniqueness.** There are multiple such global laws. Then different coherent global behaviors exist (often reflecting different large-scale patterns), and distant regions can remain strongly related.

6.2 Review: The DLR Conditions and Existence of Infinite-Volume Gibbs Measures

Motivation: What question are we trying to answer?

In finite regions $\Lambda \Subset \mathbb{Z}^d$, the Gibbs distribution is easy to define:

$$\mu_\Lambda^{\phi_{\Lambda^c}}(d\omega_\Lambda) \ \propto \ \exp[-\beta H_\Lambda(\omega_\Lambda \phi_{\Lambda^c})] \prod_{i\in\Lambda} \mu_0(d\omega_i),$$

where H_Λ is the finite-volume Hamiltonian and ϕ_{Λ^c} is a fixed boundary configuration. This distribution tells us how likely each spin configuration is *inside* Λ, given the fixed spins *outside*.

But what happens when we move to the entire infinite lattice $\mathbb{Z}^d$? We can no longer write down a global density, because the "total energy" $H_{\mathbb{Z}^d}(\omega)$ would be infinite and the partition function would not exist. So the question is:

How can we define a Gibbs measure for an infinite system where the Hamiltonian and partition function no longer make sense?

The Dobrushin–Lanford–Ruelle (DLR) conditions answer this by reformulating the problem in terms of **local consistency** rather than global energy. This is Olivia's point that we only need to consider some finite site of the lattice.

Formal definition

Let $\mathscr{F}_\Lambda$ be the σ–algebra generated by the spins inside Λ, and $\mathscr{F}_{\Lambda^c}$ the one generated by all spins outside. For each finite Λ, denote the finite-volume Gibbs law with boundary ω_{Λ^c} by $\mu_\Lambda^{\omega_{\Lambda^c}}$.

> **Theorem 6.1** (Dobrushin, Lanford, and Ruelle). *For every $f : \Omega \to \mathbb{R}$ measurable and bounded and for all $B \subset \Lambda \Subset \mathbb{Z}^d$,*
>
> $$\mathbb{E}_{\Lambda,\varphi}[f] = \mathbb{E}_{\Lambda,\varphi}[\mathbb{E}_{B,\cdot}[f]].$$

Interpretation. The DLR equations encode a simple but powerful idea:

Inside behavior = the same Gibbs rule you'd use in finite volume, given the outside.

If we know what the spins look like far away, the way the system randomizes the spins inside any finite box should still obey the usual Gibbs rule. Nothing "new" happens when we enlarge the lattice; the local equilibrium structure remains the same everywhere.

Existence of an infinite-volume Gibbs measure

It is not obvious that such a global measure μ actually exists — that there is some probability law on Ω that satisfies all these local DLR consistency conditions simultaneously.

However, Dobrushin, Lanford, and Ruelle proved that under very general assumptions on the interaction potential (such as finite range and summable strength), the family of finite-volume Gibbs measures

$$\{\mu_\Lambda^{\phi_{\Lambda^c}} : \Lambda \Subset \mathbb{Z}^d\}$$

is consistent in the sense of the Kolmogorov extension theorem. Hence, there exists at least one global measure μ on Ω whose finite-region conditionals are exactly those Gibbs distributions.

Such a μ is called an **infinite-volume Gibbs measure** or **DLR measure**.

Summary.

- A **global measure** is any probability law on Ω.
- A **DLR measure** (or infinite-volume Gibbs measure) is a global measure that satisfies the DLR equations for all finite regions.
- Existence means that at least one such measure μ exists for a given specification of local rules.

Next, we will re-express these DLR conditions in a more compact and abstract form using the language of *specification kernels.*

6.2.1 Gibbsian Specification

Definition 6.2. *(Specification and Specification Kernel) Consider the collection of probability kernel $\{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ where each $\pi_\Lambda : \mathscr{F} \times \Omega \to Prob(S^\Lambda)$ is defined by:*

$$\pi_\Lambda(A|\psi) \equiv \int_{\psi \in A} \frac{\exp(-\beta H_\Lambda(\psi))}{Z^{d,N}_{\beta,h,L,\varphi}} \left(\prod_{i \in \Lambda} \sigma \right) (d\psi) = \mu^{d,N}_{\beta,h,L,\varphi}(A). \tag{10}$$

Thus, $\{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ can effectively describe an entire model, and is thus called a specification.

Interpretation in Plain Language

This definition tells us how to construct the *Gibbsian specification kernel* $\pi_\Lambda(A|\psi)$, which assigns probabilities to different configurations inside a finite region Λ, given what happens outside that region (represented by the configuration ψ).

Intuitive idea. You can think of $\pi_\Lambda(A|\psi)$ as answering the question:

> *"If the spins outside Λ are fixed to the configuration ψ_{Λ^c}, what is the probability that the spins inside Λ fall into the set A?"*

The kernel computes this by assigning each possible inside configuration ψ_Λ a weight proportional to $\exp(-\beta H_\Lambda(\psi))$, where $H_\Lambda(\psi)$ is the total energy of the configuration (including its interactions with the fixed outside). Low-energy configurations are exponentially more likely. The constant $Z^{d,N}_{\beta,h,L,\varphi}$ in the denominator normalizes these weights so that they sum to 1, ensuring a valid probability measure.

- $\psi \in \Omega$: a configuration of spins across the entire lattice. The notation ψ_{Λ^c} refers to the spins *outside* the window Λ.
- $A \in \mathscr{F}$: an event (a subset of possible spin configurations) whose probability we are evaluating.
- $\mu^{d,N}_{\beta,h,L,\varphi}(A)$: the corresponding finite-volume Gibbs measure that results from this construction — it gives the probability of event A under the Boltzmann distribution.

Remark (parameter vs. evaluation). The specification kernel π_Λ is a fixed rule determined by the model parameters; it maps each outside configuration ω_{Λ^c} to a probability measure on S^Λ. When we write $\pi_\Lambda(\cdot\,|\,\omega)$ for a particular ω, we are *evaluating* that rule at the specific outside environment ω_{Λ^c}, obtaining a concrete inside distribution. But π_Λ itself is not fixed for one specific boundary condition; rather, it assigns a conditional law for every possible boundary condition.

With this tool, we can reformulate much of what we have before about Gibbs Measure to say, what Gibbs Measure(s) satisfy a specification.

> **Definition 6.3.** *(Single Site Specification Kernel) Define the single-site specification kernel*
>
> $$\pi_{\{i\}}(\cdot \mid \varphi) = \mu^{\varphi}_{\{i\}}(\cdot)$$
>
> *as the conditional distribution of coordinate i given the configuration outside i. In other words, "if the outside world looks like $\varphi : \Lambda^c \to \mathcal{S}$, then this specification kernel tells us how likely the inside looks like $\cdot$ (In this case, we are examining a single site).*

6.2.2 Single-site Ising kernel

Fix a site i. For any outside configuration $\omega_{\{i\}^c}$, define the local field

$$h_i(\omega) \;:=\; H \;+\; J\sum_{j\sim i} \omega_j.$$

Define the (unnormalized) weights

$$W_+(\omega) \;:=\; e^{\beta\, h_i(\omega)}, \qquad W_-(\omega) \;:=\; e^{-\beta\, h_i(\omega)}.$$

Then the single-site specification kernel (as a *rule* depending on ω) is

$$\pi_{\{i\}}(\sigma_i = +1 \mid \omega) = \frac{W_+(\omega)}{W_+(\omega) + W_-(\omega)}, \qquad \pi_{\{i\}}(\sigma_i = -1 \mid \omega) = \frac{W_-(\omega)}{W_+(\omega) + W_-(\omega)}.$$

Equivalently, in closed form,

$$\pi_{\{i\}}(\sigma_i = +1 \mid \omega) = \frac{e^{\beta h_i(\omega)}}{e^{\beta h_i(\omega)} + e^{-\beta h_i(\omega)}} = \tfrac{1}{2}\Big(1 + \tanh(\beta h_i(\omega))\Big), \quad \pi_{\{i\}}(\sigma_i = -1 \mid \omega) = \tfrac{1}{2}\Big(1 - \tanh(\beta h_i(\omega))\Big).$$

Remark (rule vs. evaluation). The kernel is a *rule* defined for every outside configuration $\omega_{\{i\}^c}$. If you plug in a specific ω^*, you just evaluate the same formulas with $\omega = \omega^*$.

6.3 Convex Structure of Gibbs Measures

Having now introduced specifications and Gibbs measures, we can study their *geometric structure*. The set of all Gibbs measures corresponding to a given specification π—denoted $\mathcal{G}(\pi)$—has a convex geometry that encodes how macroscopic phases arise and how uniqueness (or non-uniqueness) manifests.

Why this matters. Even before proving uniqueness, it is worth asking: if multiple Gibbs measures exist, how are they related? The answer is remarkably clean—every Gibbs measure can be expressed as a *mixture* (convex combination) of the most "pure" ones, called *extremal measures*. Hence, if uniqueness holds, the picture collapses: there is only one extremal point, and all possible mixtures coincide with it.

Formally, let $\mathcal{G}(\pi)$ denote the set of all Gibbs measures that satisfy the DLR equations for a given specification π.

Proposition 6.1 (Convexity and compactness)**.** *The set $\mathcal{G}(\pi)$ is convex and compact in the weak topology of probability measures on Ω.*

Sketch. Convexity follows immediately because the DLR equations are linear in the measure: if $\mu_1, \mu_2 \in \mathcal{G}(\pi)$, then $t\mu_1 + (1-t)\mu_2$ also satisfies the DLR equations for all $t \in [0,1]$. Compactness results from Prohorov's theorem: $\Omega = S^{\mathbb{Z}^d}$ is compact by Tychonoff's theorem, and the set of Gibbs measures is closed under weak convergence since the DLR relation is preserved under limits (the conditional expectations of local functions are continuous). See Friedli–Velenik, Chapter 6.8, for full details. ☐

Extremal measures and decomposition. Inside this convex set $\mathcal{G}(\pi)$, some elements are "pure"—they cannot be expressed as a nontrivial convex combination of other Gibbs measures. These are called *extremal Gibbs measures.*

Definition 6.2 (Extremality)**.** A measure $\mu \in \mathcal{G}(\pi)$ is *extremal* if for any $\mu_1, \mu_2 \in \mathcal{G}(\pi)$ and $t \in (0,1)$,

$$\mu = t\,\mu_1 + (1-t)\,\mu_2 \quad \Longrightarrow \quad \mu_1 = \mu_2 = \mu.$$

Theorem 6.3 (Choquet / extremal decomposition)**.** *Every $\mu \in \mathcal{G}(\pi)$ admits a unique representation*

$$\mu = \int_{\mathrm{ex}\,\mathcal{G}(\pi)} \nu \; w_\mu(d\nu), \tag{11}$$

where $\mathrm{ex}\,\mathcal{G}(\pi)$ denotes the set of extremal Gibbs measures, and w_μ is a probability measure on that set.

Comment. This is an application of the classical Choquet–Bishop–de Leeuw theorem, which ensures that every element of a compact convex set of probability measures admits such a unique integral representation over its extreme points. A full proof can be found in Friedli–Velenik, subsection 6.8, or standard texts on convex analysis. ☐

Intuition and physical meaning. Equation ((11)) tells us that every Gibbs measure is a "blend" of pure phases. Each extremal ν represents a distinct thermodynamic phase—e.g. the "mostly $+$" and "mostly $-$" phases in the low-temperature Ising model. A non-extremal μ then corresponds to a random mixture of these, governed by w_μ.

- If the model admits multiple extremal measures (e.g. spontaneous magnetization), then $\mathcal{G}(\pi)$ is not a single point; mixtures can reflect coexisting phases.
- If we later establish *uniqueness*, it means there is only *one* extremal measure. Then the entire convex set $\mathcal{G}(\pi)$ collapses to a single point: every Gibbs measure is the same, and every boundary condition produces the same infinite-volume behavior.

Summary. The convex structure of $\mathcal{G}(\pi)$ provides a geometric bridge between microscopic definitions and macroscopic phases. Proving uniqueness therefore has a direct physical interpretation: it means there is only one "pure" equilibrium phase, and no hidden mixtures can exist.

6.4 Clustering: Uniqueness ⇒ Decay of Correlations

In this subsection we record and prove (in the style of F&V, Ch. 6.5) that *uniqueness of the infinite-volume Gibbs measure implies clustering* (vanishing of long-range correlations). Uniqueness alone does not fix the rate (it could be polynomial, e.g. at criticality), but ensures decay to 0 as the separation grows.

Lemma 6.4 (Uniqueness $\Longleftrightarrow$ boundary–insensitive convergence (F&V Lem. 6.30))**.** *Let $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ be a specification and write*

$$\pi_\Lambda f(\omega) \;:=\; \int f(\eta)\,\pi_\Lambda(d\eta\,|\,\omega), \qquad f \text{ local.}$$

as the conditional expectation of f sampling the finite Λ_n. The following are equivalent:

(i) Uniqueness: $\mathcal{G}(\pi) = \{\mu\}$ *is a singleton.*

(ii) Boundary–insensitive convergence: *For every boundary configuration $\omega \in \Omega$, every increasing sequence $\Lambda_n \uparrow \mathbb{Z}^d$, and every local f,*

$$\pi_{\Lambda_n} f(\omega) \;\longrightarrow\; \mu(f).$$

Moreover, in (ii) *the convergence for* all ω *is essential.*

Intuitive proof. If there is only one Gibbs measure μ, then any finite–volume Gibbs distribution (with any boundary condition) must approach μ in the limit, because any weak subsequential limit of these finite–volume laws is itself a Gibbs measure, and uniqueness forces all such limits to coincide. Conversely, if the finite–volume expectations converge to the same limit for every boundary configuration, then the limiting measure is uniquely determined, so two distinct infinite–volume Gibbs measures cannot exist. Hence, uniqueness $\iff$ boundary–insensitive convergence. □

Theorem 6.4 (Clustering from uniqueness)**.** *Assume $\mathcal{G}(\pi) = \{\mu\}$ is a singleton. Let f, g be bounded local functions with* $\mathrm{supp}(f) = A$, $\mathrm{supp}(g) = B$ *where A, B are disjoint finite subsets of $\mathbb{Z}^d$. Then*

$$\lim_{d(A,B)\to\infty} \mathrm{Cov}_\mu(f, g) \;=\; 0.$$

Proof (detailed explanation). We restate what uniqueness gives us. By Lemma 6.30 of Friedli–Velenik, uniqueness of the Gibbs measure means the following:

For every boundary condition $\omega \in \Omega$ and every local observable $h : \Omega \to \mathbb{R}$, the finite-volume conditional expectations converge to the same infinite-volume value:

$$\pi_{\Lambda_n} h(\omega) \;\longrightarrow\; \mu(h) \quad \text{as} \quad \Lambda_n \uparrow \mathbb{Z}^d.$$

That is, when we enlarge the finite box Λ_n, the conditional expectation of h inside the box (given boundary condition ω) approaches the global Gibbs expectation $\mu(h)$. This convergence holds uniformly in the boundary ω.

Step 1. Choose two local observables and separate their supports.

Let f, g be bounded *local* functions. This means there exist finite subsets $A, B \subset \mathbb{Z}^d$ such that f depends only on the spins in A and g depends only on the spins in B:

$$\mathrm{supp}(f) = A, \qquad \mathrm{supp}(g) = B.$$

The distance between their supports is

$$d(A, B) := \min_{i\in A,\, j\in B} \|i - j\|.$$

Pick a finite region Λ containing A. Now let Λ_R be a larger box containing Λ and also the "buffer" of radius R around it, so that B lies outside Λ_R. Thus the annulus between Λ and Λ_R separates A and B by at least distance R.

Step 2. Write the finite-volume DLR identity.

For any boundary configuration ω, the DLR (consistency) property says:

$$\pi^{\omega}_{\Lambda_R}(fg) = \pi^{\omega}_{\Lambda_R}\Big(f \cdot \pi^{\cdot}_{\Lambda}(g)\Big).$$

Here:

- $\pi^{\omega}_{\Lambda_R}$ is the Gibbs specification on Λ_R given the boundary $\omega_{\Lambda_R^c}$.
- $\pi^{\cdot}_{\Lambda}(g)$ means we apply π_Λ to g while keeping the outer spins (beyond Λ) fixed.

Step 3. Use uniqueness to replace local conditional expectations by global ones.

Now, by uniqueness, for any exterior configuration ζ,

$$\pi^{\zeta}_{\Lambda}(g) \;\longrightarrow\; \mu(g) \quad \text{uniformly in } \zeta \quad \text{as} \quad R = d(A, B) \to \infty,$$

because the exterior of Λ (which fixes the boundary for g) is very far from B.

Therefore,

$$\pi^{\omega}_{\Lambda_R}(fg) = \pi^{\omega}_{\Lambda_R}\big(f \cdot \mu(g)\big) + o_R(1) = \mu(g)\, \pi^{\omega}_{\Lambda_R}(f) + o_R(1),$$

where $o_R(1) \to 0$ as $R \to \infty$.

Step 4. Let the box grow to the whole lattice.

Now send $\Lambda_R \uparrow \mathbb{Z}^d$. By uniqueness again,

$$\pi^{\omega}_{\Lambda_R}(f) \longrightarrow \mu(f).$$

Plugging this limit into the previous line gives

$$\mu(fg) = \mu(f)\mu(g) + o_R(1),$$

so the covariance $\mathrm{Cov}_\mu(f,g)$ tends to 0 as $R = d(A,B) \to \infty$.

Step 5. Interpretation. This shows that when there is a unique infinite-volume Gibbs measure, local observables depending on disjoint (and increasingly distant) regions become asymptotically independent. That is,

$$\boxed{\mathrm{Cov}_\mu(f,g) \longrightarrow 0 \quad \text{as} \quad d(A,B) \to \infty.}$$

Uniqueness thus implies the *clustering property*—the mathematical statement of decay of correlations. □

Intuition. Uniqueness means finite-volume expectations *forget* their boundaries as the volume grows. When the covariance $\mathrm{Cov}_\mu(f,g) = \mu(fg) - \mu(f)\mu(g)$ equals 0, it means that knowing the outcome of f tells you nothing about g: the two observables fluctuate completely independently under the Gibbs measure. In physical terms, there are no correlations left between distant spins.

6.5 The Dobrushin Influence Matrix

6.5.1 Definition

Definition 6.5 (Influence coefficients). *For distinct sites $i, j \in \mathbb{Z}^d$, define*

$$c_{ij}(\pi) := \sup_{\omega,\omega' :\ \omega_k = \omega'_k\ \ \forall k \neq j} \|\pi_{\{i\}}(\cdot|\omega) - \pi_{\{i\}}(\cdot|\omega')\|_{TV}.$$

The Dobrushin constant *is*

$$c(\pi) = \sup_i \sum_j c_{ij}(\pi).$$

Interpretation. c_{ij} measures the maximal change in the single-site distribution at i caused by flipping site j while keeping everything else fixed. The total influence $c(\pi)$ quantifies how strongly a site's environment controls it. If $c(\pi) < 1$, the cumulative effect of all neighbors is less than complete control—an idea that will yield a contraction argument.

From the single-site kernel to Dobrushin coefficients. Recall the single-site specification at i (Ising, $\sigma_i \in \{\pm 1\}$):

$$\pi_{\{i\}}(\sigma_i = +1 \mid \omega) = \frac{e^{\beta h_i(\omega)}}{e^{\beta h_i(\omega)} + e^{-\beta h_i(\omega)}} = \frac{1 + \tanh\big(\beta h_i(\omega)\big)}{2}, \qquad h_i(\omega) = H + \sum_{k \sim i} J_{ik}\, \omega_k.$$

Dobrushin's influence coefficient is

$$c_{ij} \ := \sup_{\omega,\omega' :\ \omega_k = \omega'_k\ \ \forall k \neq j} d_{\mathrm{TV}}\Big(\pi_{\{i\}}(\cdot \mid \omega),\ \pi_{\{i\}}(\cdot \mid \omega')\Big),$$

i.e. the worst-case total-variation change of the law at i when only site j is altered.

Binary case simplification. Since $\sigma_i \in \{\pm 1\}$, total variation equals the difference of "+1" probabilities:

$$d_{\mathrm{TV}}(\mathrm{Bern}(p), \mathrm{Bern}(q)) = |p - q|.$$

Thus

$$c_{ij} = \sup_{\omega,\omega' \text{ differ only at } j} \Big|\pi_{\{i\}}(+1 \mid \omega) - \pi_{\{i\}}(+1 \mid \omega')\Big|.$$

Field shift when flipping j. If ω and ω' differ only at j, then

$$h_i(\omega') - h_i(\omega) = J_{ij}\big(\omega'_j - \omega_j\big) \in \{-2J_{ij},\, 0,\, 2J_{ij}\}.$$

In the worst case (flip $+1 \leftrightarrow -1$), $|\Delta h| = 2|J_{ij}|$.

Maximizing the change. Write $p(h) := \frac{1+\tanh(\beta h)}{2}$. Then

$$\Big|p(h) - p(h-\Delta h)\Big| = \frac{1}{2}\Big|\tanh(\beta h) - \tanh\big(\beta(h-\Delta h)\big)\Big|.$$

A standard calculation shows that the supremum over $h \in \mathbb{R}$ occurs at $h = \frac{1}{2}\Delta h$ and yields

$$\sup_h \Big|p(h) - p(h-\Delta h)\Big| = \tanh\Big(\frac{\beta|\Delta h|}{2}\Big).$$

Plugging $|\Delta h| = 2|J_{ij}|$ gives

$$c_{ij} \;\le\; \tanh\big(\beta|J_{ij}|\big).$$

Normalization note. Some texts (or Hamiltonian conventions) absorb a factor 2 differently in the pair interaction, leading to $|\Delta h| = 4|J_{ij}|$ and hence the bound

$$c_{ij} \;\le\; \tanh\big(2\beta|J_{ij}|\big).$$

Use the version that matches your Hamiltonian normalization.

Summing influences. Dobrushin's constant is

$$c(\pi) := \sup_i \sum_{j\neq i} c_{ij}.$$

For nearest-neighbor Ising with uniform coupling J on $\mathbb{Z}^d$ (each i has $2d$ neighbors),

$$c(\pi) \;\le\; \sum_{j\sim i} \tanh\big(2\beta|J|\big) \;=\; (2d)\,\tanh\big(2\beta J\big).$$

Therefore, at high temperature (small β), $c(\pi) < 1$, which is Dobrushin's uniqueness condition.

6.6 Dobrushin Uniqueness Theorem

So we have that uniqueness is going to imply clustering - or the decay in correlations. We also have that uniqueness informs us of our mixture, that namely it is an extremal gibbs measure of one pure state, not a mix of multiple.

How now can we come up with some convenient condition that guarantee the uniqueness of our infinite volume Gibbs Measure for a certain specification? In answering this question, we will also find that the conveinent conditions also tell us something about the rate at which the correlations will decay.

We now state the central result and give a proof that follows Friedli–Velenik (Ch. 6.3–6.4) in structure and notation, with full details.

Theorem 6.5 (Dobrushin Uniqueness). *Let $\{\,\pi_{\{i\}}\,\}_{i\in\mathbb{Z}^d}$ be the single-site specification kernels associated with the model, and let $c_{ij}(\pi)$ and $c(\pi) = \sup_{i\in\mathbb{Z}^d}\sum_{j\in\mathbb{Z}^d} c_{ij}(\pi)$ be the Dobrushin influence coefficients and constants defined in the previous subsection. If*

$$c(\pi) \;<\; 1, \qquad \textit{(weak dependence condition)}$$

then the set of infinite-volume Gibbs measures compatible with π is a singleton:

$$\mathcal{G}(\pi) \;=\; \{\mu\}.$$

Roadmap of the proof.

1. *The Dobrushin Influence Matrix* Define the tool that we will use to measure sensitivities of sites to other sites.
2. *Oscillation seminorm and metric.* Define the single-site oscillations $\delta_i(f)$ and total oscillation $\Delta(f)$ of a function f; prove that $\Delta(f)$ controls the total range of f and serves as a Lipschitz seminorm.
3. *Dusting lemma (single-site update).* Show that applying the single-site kernel $\pi_{\{j\}}$ "kills" the dependence of f on site j and moves at most a c_{ji}–fraction of that dependence to other sites i.
4. *One-sweep contraction.* Compose single-site updates over a finite block Λ and bound the resulting total oscillation by a factor at most $c(\pi) < 1$ times the initial oscillation.
5. *Contraction of expectations and uniqueness.* By DLR, Gibbs measures are invariant under π_Λ; together with the one-sweep contraction this implies a strict contraction of differences of expectations for all Lipschitz local observables, which forces uniqueness by iteration.

6.6.1 Oscillation seminorm / Dobrushin Metric (and coupling representation)

For a bounded measurable function $f : \Omega \to \mathbb{R}$ define the *single-site oscillation* at site i by

$$\delta_i(f) := \sup_{\substack{\omega,\eta \\ \omega_k=\eta_k\ \forall k\neq i}} |f(\omega) - f(\eta)|. \tag{12}$$

The *total oscillation* of f is then

$$\Delta(f) := \sum_{i\in\mathbb{Z}^d} \delta_i(f). \tag{13}$$

Following Friedli–Velenik (Def. 6.2), the *oscillation seminorm* on functions is given by $\Delta(f)$, and the corresponding distance on probability measures by

$$\|\mu - \nu\|_{\text{osc}} := \sup\{\, |\mu(f) - \nu(f)| : \Delta(f) \le 1\,\}. \tag{14}$$

Intuition: what the oscillation seminorm measures. The quantity $\delta_i(f)$ measures how sensitive a function is to single-site changes. If all $\delta_i(f)$ are small then changing a single spin cannot change f very much: such functions are "smooth" or local in the sense that they do not react strongly to perturbations at any site.

This distance $\|\mu-\nu\|_{\text{osc}}$ is called the Dobrushin (or oscillation) metric, and is a way to measure how “different” two probability distributions over spin configurations are — not by comparing every possible detail, but by seeing how much they disagree on local observables. It's tuned so that convergence in this sense matches the physical idea of two states becoming indistinguishable in all local tests (this will be proven below).

Note. Coupling Interpretation:
Optional: This is not discussed in F&V and is just included for another intuitive way of interpreting this oscillation metric.

The oscillation distance admits a simple probabilistic interpretation: it is the smallest expected number of differing sites between two random configurations drawn optimally from μ and ν. This provides an intuitive geometric picture of how “far apart” the measures are in configuration space.

Lemma 6.6 (Coupling representation, cf. F&V Lemma 6.3)**.** *For any probability measures μ, ν on $(\Omega, \mathcal{F})$,*

$$\|\mu - \nu\|_{\text{osc}} = \inf_{\pi\in\mathcal{C}(\mu,\nu)} \mathbb{E}_\pi\left[\sum_{i\in\mathbb{Z}^d} \mathbf{1}\{\omega_i \neq \eta_i\}\right], \tag{15}$$

where $\mathcal{C}(\mu,\nu)$ denotes the set of couplings of μ and ν.

Proof idea. The idea is to show that the oscillation distance between two measures is the smallest possible average number of sites where two coupled configurations disagree. The proof is ommited as this is merely just to give another intuitive understanding. The focus of this lecture is uniqueness.

Next, with this Dobrushin (or oscillation) metric, we introduce the functional spaces used throughout the proof and establish some basic inequalities.

Definition 6.7 (Spaces of observables)**.** Let $\Omega = \mathcal{S}^{\mathbb{Z}^d}$ with product σ-algebra $\mathcal{F}$.

- The space of functions with *finite total oscillation* is
$$\mathcal{O}(\Omega) \;:=\; \{\, f : \Omega \to \mathbb{R} \text{ measurable} : \Delta(f) < \infty \,\}.$$
- The space of *continuous observables with finite oscillation* is
$$C(\Omega) \cap \mathcal{O}(\Omega) \;:=\; \{\, f \in C(\Omega) : \Delta(f) < \infty \,\}.$$

Lemma 6.8 (Basic oscillation bound (cf. F&V Lemma 6.32))**.** *For any bounded function $f : \Omega \to \mathbb{R}$,*

$$\sup_{\omega\in\Omega} f(\omega) \;-\; \inf_{\omega\in\Omega} f(\omega) \;\le\; \Delta(f).$$

Consequently, for any probability measures μ, ν on $(\Omega, \mathcal{F})$ and any $f \in C(\Omega) \cap \mathcal{O}(\Omega)$,

$$|\mu(f) - \nu(f)| \;\le\; \Delta(f).$$

Proof. Fix $\omega, \eta \in \Omega$ and enumerate $\mathbb{Z}^d$ as $(i_k)_{k\geq 1}$. Define intermediate configurations

$$\xi^{(0)} = \omega, \qquad \xi^{(m)} \text{ agrees with } \eta \text{ on } \{i_1, \dots, i_m\} \text{ and with } \omega \text{ elsewhere.}$$

Then $\xi^{(m)}$ and $\xi^{(m-1)}$ differ only at site i_m, so by definition of the single-site oscillations

$$\big|f(\xi^{(m)}) - f(\xi^{(m-1)})\big| \leq \delta_{i_m}(f).$$

Let $S_n := \sum_{k=1}^n \big[f(\xi^{(k)}) - f(\xi^{(k-1)})\big]$. Since $\sum_i \delta_i(f) = \Delta(f) < \infty$, we have $|S_n| \leq \sum_{k=1}^n \delta_{i_k}(f) \leq \Delta(f)$ and $(S_n)_n$ is a Cauchy sequence. Hence $S_n \to f(\eta) - f(\omega)$ and

$$|f(\omega) - f(\eta)| \;=\; \lim_{n\to\infty} |S_n| \;\leq\; \sum_{i\in\mathbb{Z}^d} \delta_i(f) \;=\; \Delta(f).$$

Next, fix $\varepsilon > 0$. By compactness of Ω and continuity of f, pick $\omega^\star, \eta^\star \in \Omega$ with

$$f(\omega^\star) \geq \sup f - \varepsilon, \qquad f(\eta^\star) \leq \inf f + \varepsilon.$$

Then

$$\sup f - \inf f \;\leq\; \big[f(\omega^\star) - f(\eta^\star)\big] + 2\varepsilon \;\leq\; \Delta(f) + 2\varepsilon.$$

Letting $\varepsilon \downarrow 0$ yields $\sup f - \inf f \leq \Delta(f)$.

Finally, for any probability measures μ, ν on $(\Omega, \mathscr{F})$, since $\inf f \leq f \leq \sup f$ pointwise, we have

$$\inf f \;\leq\; \mu(f), \nu(f) \;\leq\; \sup f,$$

and thus

$$|\mu(f) - \nu(f)| \;\leq\; \sup f - \inf f \;\leq\; \Delta(f).$$

Note: Because Ω is compact and $f \in C(\Omega)$, the function f is automatically bounded, uniformly continuous, and attains its supremum and infimum by the Weierstrass theorem. ☐

Lay intuition. The quantity $\Delta(f)$ measures the maximum cumulative sensitivity of f to local spin flips. Changing one spin can alter f by at most $\delta_i(f)$; summing these effects bounds the largest possible difference of f across the whole configuration space. Hence $\Delta(f)$ acts like a Lipschitz constant for f (describing the proportional change of the output of f given changes in the input) on the infinite product space.

So far, we've defined the oscillation seminorm and seen that it provides a natural way to measure how sensitive a function is to single-site changes in the configuration.

The next step is to understand how this "sensitivity" evolves when we start updating spins according to their local conditional distributions—that is, when we let the system dynamically adjust one site at a time. In other words: if we take a function that depends on the configuration and "resample" one site according to the model's local rule, how does the function's dependence on other sites change?

Answering this question leads us to the idea of the *single-site update operator*, which formalizes this resampling process, and to a key technical result that describes how local updates redistribute sensitivity across the lattice—the *Dusting Lemma.*

6.6.2 Single-site update on functions and the Dusting Lemma

Notation: We write

$$\omega^{(j,s)} := (\omega_1, \dots, \omega_{j-1}, s, \omega_{j+1}, \dots) \quad (:= s\,\omega_{j^c} \text{ in F\&V})$$

for the configuration obtained from ω by replacing its spin at site j with s.

For $j \in \mathbb{Z}^d$ and a bounded observable $f : \Omega \to \mathbb{R}$, define the *single-site update* of f at j by

$$(\pi_{\{j\}} f)(\omega) \;:=\; \sum_{s \in \mathcal{S}} f\big(\omega^{(j,s)}\big)\, \pi_{\{j\}}(s \mid \omega), \qquad \omega \in \Omega, \tag{16}$$

where $\pi_{\{j\}}(\cdot \mid \omega)$ is the single-site conditional distribution at site j.

Lemma 6.9 (Dusting Lemma (cf. F&V Lem. 6.34))**.** *Fix $j \in \mathbb{Z}^d$ and $f \in C(\Omega) \cap \mathcal{O}(\Omega)$. Then, for each $i \neq j$,*

$$\delta_i(\pi_{\{j\}} f) \;\le\; \delta_i(f) \;+\; c_{ji}\, \delta_j(f), \qquad \text{and} \qquad \delta_j(\pi_{\{j\}} f) \;=\; 0,$$

where c_{ji} is the Dobrushin influence coefficient.

Intuition. Resampling at j removes all dependence on j (so the updated function has zero sensitivity at j), while at most a c_{ji}-fraction of that lost dependence can be redistributed to other sites i.

Proof. Fix $i \neq j$ and take $\omega, \eta \in \Omega$ that differ only at site i. Then

$$\begin{aligned}(\pi_{\{j\}} f)(\omega) - (\pi_{\{j\}} f)(\eta) &= \sum_{s \in \mathcal{S}} \Big(f(\omega^{(j,s)})\, \pi_{\{j\}}(s \mid \omega) - f(\eta^{(j,s)})\, \pi_{\{j\}}(s \mid \eta) \Big) \\ &= \underbrace{\sum_{s \in \mathcal{S}} \big[f(\omega^{(j,s)}) - f(\eta^{(j,s)}) \big] \pi_{\{j\}}(s \mid \omega)}_{\text{(I)}} + \underbrace{\sum_{s \in \mathcal{S}} f(\eta^{(j,s)}) \Big[\pi_{\{j\}}(s \mid \omega) - \pi_{\{j\}}(s \mid \eta) \Big]}_{\text{(II)}}.\end{aligned}$$

For (I): for each s, the configurations $\omega^{(j,s)}$ and $\eta^{(j,s)}$ differ only at i, so

$$|f(\omega^{(j,s)}) - f(\eta^{(j,s)})| \;\le\; \delta_i(f),$$

and summing against a probability measure yields $|\text{(I)}| \le \delta_i(f)$.

For (II): Set $g(s) := f(\eta^{(j,s)})$. Then (II) $= \sum_s g(s)\, [\pi_{\{j\}}(s \mid \omega) - \pi_{\{j\}}(s \mid \eta)]$. Since changing s only modifies the spin at site j, we have

$$\max_s g(s) - \min_s g(s) \;\le\; \delta_j(f).$$

For probability vectors μ, ν on the finite set $\mathcal{S}$, the total variation distance is

$$\|\mu - \nu\|_{\text{TV}} := \tfrac{1}{2} \sum_{s \in \mathcal{S}} |\mu(s) - \nu(s)|\,.$$

If we set $\bar{g} := \frac{1}{2}(\max_s g(s) + \min_s g(s))$, then

$$\begin{aligned}\Big| \sum_s g(s)\, [\mu(s) - \nu(s)] \Big| &= \Big| \sum_s (g(s) - \bar{g})\, [\mu(s) - \nu(s)] \Big| \\ &\le \sup_s |g(s) - \bar{g}| \sum_s |\mu(s) - \nu(s)| \\ &= \big(\max_s g(s) - \min_s g(s) \big)\, \|\mu - \nu\|_{\text{TV}}.\end{aligned}$$

Applying this with $\mu = \pi_{\{j\}}(\cdot \mid \omega)$, $\nu = \pi_{\{j\}}(\cdot \mid \eta)$, and using $\max_s g(s) - \min_s g(s) \le \delta_j(f)$, we get

$$|\text{(II)}| \le \delta_j(f)\, \big\| \pi_{\{j\}}(\cdot \mid \omega) - \pi_{\{j\}}(\cdot \mid \eta) \big\|_{\text{TV}}.$$

Finally, since ω, η differ only at site i, by the definition of the Dobrushin influence coefficient,

$$\big\| \pi_{\{j\}}(\cdot \mid \omega) - \pi_{\{j\}}(\cdot \mid \eta) \big\|_{\text{TV}} \le c_{ji},$$

and therefore

$$|\text{(II)}| \;\le\; c_{ji}\, \delta_j(f).$$

Taking the supremum over all such pairs (ω, η) gives

$$\delta_i(\pi_{\{j\}} f) \;\le\; \delta_i(f) \;+\; c_{ji}\,\delta_j(f).$$

Finally, for $i = j$: if ω and η coincide outside of site j, then $\omega_{j^c} = \eta_{j^c}$ and $\pi_{\{j\}}(\cdot \mid \omega) = \pi_{\{j\}}(\cdot \mid \eta)$, which implies

$$(\pi_{\{j\}} f)(\omega) = (\pi_{\{j\}} f)(\eta),$$

and therefore

$$\delta_j(\pi_{\{j\}} f) = 0.$$

□

6.6.3 One-sweep contraction on oscillations

Let $\Lambda \Subset \mathbb{Z}^d$ be finite and fix an order $(j_1, \dots, j_m)$ of the sites in Λ. Define the sweep operator

$$\pi_\Lambda \;:=\; \pi_{\{j_m\}} \circ \cdots \circ \pi_{\{j_1\}}.$$

Introduce linear maps $R^{(j)} : \mathbb{R}^{\mathbb{Z}^d} \to \mathbb{R}^{\mathbb{Z}^d}$ acting on vectors of single-site oscillations $v = (v_i)_{i \in \mathbb{Z}^d}$ by

$$(R^{(j)} v)_i \;:=\; \begin{cases} v_i + c_{ji}\, v_j, & i \neq j, \\ 0, & i = j. \end{cases}$$

Equivalently, $R^{(j)}$ sets the j-th coordinate to zero and adds at most c_{ji} times the old j-mass to coordinate i, reflecting Lemma Theorem 6.9.

Proposition 6.10 (Oscillation transfer under a sweep (cf. F&V Prop. 6.33))**.** *For any bounded $f : \Omega \to \mathbb{R}$,*

$$\big(\,\delta_i(\pi_\Lambda f)\,\big)_{i \in \mathbb{Z}^d} \;\le\; \Big(R^{(j_m)} \cdots R^{(j_1)}\Big)\,\big(\,\delta_i(f)\,\big)_{i \in \mathbb{Z}^d} \quad \textit{coordinatewise.}$$

In particular,

$$\Delta(\pi_\Lambda f) \;\le\; \big\|R^{(j_m)} \cdots R^{(j_1)}\big\|_{1\to 1}\,\Delta(f), \qquad \big\|R^{(j_m)} \cdots R^{(j_1)}\big\|_{1\to 1} \;\le\; c(\pi),$$

where $c(\pi) := \sup_i \sum_k c_{ik}(\pi)$ is the Dobrushin constant.

Proof. Applying Lemma Theorem 6.9 successively, each update at site j_t acts linearly on the vector of single-site oscillations via $R^{(j_t)}$. Thus after m steps,

$$\delta(\pi_\Lambda f) \;=\; \delta(\pi_{\{j_m\}} \cdots \pi_{\{j_1\}} f) \;\le\; \big(R^{(j_m)} \cdots R^{(j_1)}\big)\,\delta(f),$$

which gives the coordinatewise inequality.
For the ℓ^1 bound, observe that each $R^{(j)}$ sets its j-th entry to zero and redistributes at most a total weight $\sum_{i \neq j} c_{ji} \le c(\pi)$ to other coordinates. Hence every column of $R^{(j)}$ has ℓ^1-sum $\le c(\pi)$, and the same holds for their product. Therefore

$$\big\|R^{(j_m)} \cdots R^{(j_1)}\big\|_{1\to 1} \le c(\pi),$$

yielding the claimed contraction. □

Intuition. A full sweep over Λ contracts the total oscillation $\Delta(f)$ by at least the uniform factor $c(\pi) < 1$, the maximal total incoming influence per site.

6.6.4 Contraction of expectations and proof of uniqueness

Proposition 6.11 (Contraction of differences under a sweep)**.** *Let $\Lambda \Subset \mathbb{Z}^d$ be finite and $(j_1, \dots, j_m)$ any order on Λ. Then for any probability measures μ, ν on $(\Omega, \mathscr{F})$ and any $f \in C(\Omega) \cap \mathcal{O}(\Omega)$,*

$$\big|\mu(\pi_\Lambda f) - \nu(\pi_\Lambda f)\big| \;\le\; \Delta(\pi_\Lambda f) \;\le\; c(\pi)\,\Delta(f).$$

Proof. By Lemma Theorem 6.8, $|\mu(g) - \nu(g)| \leq \Delta(g)$ for any $g \in \mathcal{O}(\Omega)$. Apply this to $g = \pi_\Lambda f$, then use Proposition Theorem 6.10 to get $\Delta(\pi_\Lambda f) \leq c(\pi)\,\Delta(f)$. ☐

Theorem 6.12 (Dobrushin Uniqueness (F&V Thm. 6.31)). *Assume $c(\pi) = \sup_i \sum_k c_{ik} < 1$. Then the set of infinite-volume Gibbs measures compatible with π is a singleton:*

$$\mathcal{G}(\pi) = \{\mu\}.$$

Proof. Let $\mu, \nu \in \mathcal{G}(\pi)$ and $f \in C(\Omega) \cap \mathcal{O}(\Omega)$. With out heavy lifting above, we can simply apply Proposition 5.7 repeatedly. Hence:

$$|\mu(f) - \nu(f)| \;=\; |\mu(\pi_\Lambda f) - \nu(\pi_\Lambda f)| \;\leq\; c(\pi)\,\Delta(f),$$

Replacing f by $\pi_\Lambda f$ and iterating the same argument n times yields

$$|\mu(f) - \nu(f)| \;\leq\; c(\pi)^n\,\Delta(f) \qquad \forall n \in \mathbb{N}.$$

Letting $n \to \infty$ (since $c(\pi) < 1$) gives $|\mu(f) - \nu(f)| = 0$ for all $f \in C(\Omega) \cap \mathcal{O}(\Omega)$, a separating class of observables. Therefore $\mu = \nu$. ☐

6.7 Physical Interpretation and High-Temperature Regime

At small β, spins fluctuate almost independently. The Dobrushin constant $c(\pi)$ becomes small, ensuring uniqueness. Increasing β strengthens correlations, enlarging $c(\pi)$, and at the critical point $c(\pi) = 1$ the system loses contractivity—correlations decay only polynomially and critical phenomena emerge.

Regime summary.

Condition	Physical regime	Behavior
$c(\pi) < 1$	High temperature	Unique Gibbs measure; exponential decay
$c(\pi) = 1$	Critical	Marginal uniqueness; polynomial decay
$c(\pi) > 1$	Low temperature	Multiple Gibbs measures (phase coexistence)

Correlation decay. For local f, g supported on disjoint A, B,

$$|\mathrm{Cov}_\mu(f, g)| \leq C(f, g)\, c(\pi)^{d(A,B)},$$

which expresses exponential clustering. In the Ising model, this corresponds to a finite correlation length $\xi = -1/\log c(\pi)$.

6.8 Summary and Conceptual Map

- The DLR framework formalizes local equilibrium in infinite systems.
- Single-site kernels $\pi_{\{i\}}$ provide a measurable description of local dependencies.
- Dobrushin coefficients c_{ij} quantify the strength of influence between sites.
- If the total influence $c(\pi) < 1$, local updates contract discrepancies, implying a unique global Gibbs measure.
- Uniqueness entails decay of correlations—physically, absence of long-range order.
- For the Ising model, this condition holds at sufficiently high temperature, matching physical intuition.

7 Jerry: The Aizenman-Simon proof of the disordered phase for all small β and other cluster expansions

7.1 Setup

Definition 7.1 (Space and spin dimensions). *We are given $d, N \in \mathbb{N}$ as the space and spin dimensions respectively.*

Definition 7.2 (Inverse temperature). *β is the inverse temperature.*

Definition 7.3 (Boxes). *Let $\Lambda := [-L, L]^d \cap \mathbb{Z}^d$ be a finite box within $\mathbb{Z}^d$.*

Definition 7.4 (Graph, adjacency matrix). *These boxes can be considered as a graph $G = (V, E)$, with and associated adjacency matrix A where*

$$A_{xy} = \begin{cases} 1 & x \sim y \\ 0 & \textit{otherwise} \end{cases}$$

Definition 7.5 (Spin configuration). *ψ is the spin configuration, $\psi : \Lambda \to \mathbb{S}^{N-1}$.*

Definition 7.6 (Discrete Laplacian, Hamiltonian). *The discrete Laplacian is defined as $-\Delta := D - A$, where D is the diagonal degree matrix of the graph. In our case, $-\Delta \equiv 2d\mathbb{1} - A$.*

The Hamiltonian $H : (\mathbb{S}^{N-1})^\Lambda \to \mathbb{R}$ is a functional that maps spin configurations to real numbers. In our case,

$$H = \frac{1}{2} \langle \psi, -\Delta\psi \rangle - \langle h, \psi \rangle$$

where h is the external magnetic field, which we will take to be equal to zero.

7.2 Local Ward Identities

The following proofs in this subsection are derived from [1].

Definition 7.7 (Expectation of an observable under the Gibbs measure). *The expectation of the observable Q under the Gibbs measure is defined by*

$$\mathbb{E}[Q(\psi)] \equiv \mathbb{E}_\beta[Q(\psi)] = \frac{\mathbb{E}_0[e^{-\beta H(\psi)} Q(\psi)]}{\mathbb{E}_0[e^{-\beta H(\psi)}]}$$

where $\mathbb{E}_0$ is an expectation under the a-priori measure.

Definition 7.8 (Automorphisms). *Consider a family of automorphisms γ_t satisfying $\gamma_t(QR) = \gamma_t(Q)\gamma_t(R)$ which preserve the a-priori expectation values, i.e.*

$$\mathbb{E}_0[\gamma_t(Q)] = \mathbb{E}_0[Q]$$

for all t. This also implies that $\partial_t \mathbb{E}_0[\gamma_t(Q)] = 0$.

Definition 7.9. *Define $\dot{Q} = \frac{d}{dt}\gamma_t(Q)\big|_{t=0}$*

Lemma 7.10 (Local Ward Identities)**.**

$$\mathbb{E}_\beta[\dot{Q}] = \beta\mathbb{E}_\beta[\dot{H}Q]$$

Proof. Definition 7.8 implies in particular that

$$\partial_t\big|_{t=0}\,\mathbb{E}_0[\gamma_t(e^{-\beta H}Q)] = 0$$

$$\partial_t\big|_{t=0}\,\mathbb{E}_0[\gamma_t(Q)e^{-\beta\gamma_t(H)}] = 0$$

Now evaluating the t-derivatives and noting that γ_0 is identity,

$$\mathbb{E}_0[\dot{Q}e^{-\beta H}] - \beta\mathbb{E}_0[\dot{H}Qe^{-\beta H}] = 0$$

$$\mathbb{E}_\beta[\dot{Q}] = \beta\mathbb{E}_\beta[\dot{H}Q]$$

□

7.3 Exponential decay of the two point function

Lemma 7.11 (Positive association in the Ising model)**.** *For any $f : \mathbb{R} \to \mathbb{R}$ which is odd and C^1, and such that f' is decreasing on $[0, \infty)$, we have*

$$\mathbb{E}\left[f\left(\sum_z a_z\psi_z\right)\psi_y\right] \leq f'(0)\sum_z a_z\mathbb{E}[\psi_x\psi_y]$$

where $\{a_z\}_z$ is some finite sequence of positive numbers.

Proof. Since f is odd, $f(0) = 0$. So we may rewrite for any $s \in \mathbb{R}$,

$$\begin{aligned} f(s) &= \int_{t=0}^{s} f'(t)\,\mathrm{d}t \\ &= \operatorname{sgn}(s)\int_{t=0}^{|s|} f'(t)\,\mathrm{d}t \\ &= \operatorname{sgn}(s)\int_{t=0}^{\infty} \chi_{[0,|s|]}(t)f'(t)\,\mathrm{d}t. \end{aligned}$$

Since $\operatorname{sgn}(s) = \chi_{[0,\infty)}(s) - \chi_{(-\infty,0)}(s)$ we get

$$\begin{aligned} f(s) &= \chi_{[0,\infty)}(s)\int_{t=0}^{\infty}\chi_{[0,|s|]}(t)f'(t)\,\mathrm{d}t - \chi_{(-\infty,0)}(s)\int_{t=0}^{\infty}\chi_{[0,|s|]}(t)f'(t)\,\mathrm{d}t \\ &= \chi_{[0,\infty)}(s)\int_{t=0}^{\infty}\chi_{[0,s]}(t)f'(t)\,\mathrm{d}t - \chi_{(-\infty,0)}(s)\int_{t=0}^{\infty}\chi_{[0,-s]}(t)f'(t)\,\mathrm{d}t. \end{aligned}$$

Plug this representation of f into $\mathbb{E}[f(S(\psi))\psi_y]$ (with $S(\psi) = \sum_z a_z\psi_z$) to get

$$\begin{aligned} \mathbb{E}[f(S(\psi))\psi_y] &= \mathbb{E}\left[\left(\chi_{[0,\infty)}(S(\psi))\int_{t=0}^{\infty}\chi_{[0,S(\psi)]}(t)f'(t)\,\mathrm{d}t - \chi_{(-\infty,0)}(S(\psi))\int_{t=0}^{\infty}\chi_{[0,-S(\psi)]}(t)f'(t)\,\mathrm{d}t\right)\psi_y\right] \\ &= \int_{t=0}^{\infty}\mathbb{E}\left[\chi_{[0,\infty)}(S(\psi))\chi_{[0,S(\psi)]}(t)\psi_y\right]f'(t)\,\mathrm{d}t - \int_{t=0}^{\infty}\mathbb{E}\left[\chi_{(-\infty,0)}(S(\psi))\chi_{[0,-S(\psi)]}(t)\psi_y\right]f'(t)\,\mathrm{d}t. \end{aligned}$$

Now note that

$$\mathbb{E}\left[\chi_{[0,\infty)}(S(\psi))\chi_{[0,S(\psi)]}(t)\psi_y\right] = \mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right]$$

whereas

$$\begin{aligned}\mathbb{E}\left[\chi_{(-\infty,0)}(S(\psi))\chi_{[0,-S(\psi)]}(t)\psi_y\right] &= \mathbb{E}\left[\chi_{(-\infty,-t)}(S(\psi))\psi_y\right] \\ &= -\mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right].\end{aligned}$$

In this last equality we have used the invariance of the probability measure under $\psi \mapsto -\psi$. We thus find the representation

$$\mathbb{E}[f(S(\psi))\psi_y] = 2\int_{t=0}^{\infty} \mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right] f'(t)\,\mathrm{d}t.$$

We claim that

$$\mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right] \geq 0.$$

To establish this we invoke *the FKG inequality* which requires the notion of functions $\psi \mapsto F(\psi)$ being increasing. Loosely speaking this is means that if we flip *any* spins from -1 to $+1$ the function can only go up. $\psi \mapsto \psi_y$ is such a function and $\psi \mapsto \chi_{[t,\infty)}(S(\psi))$ as well (using the fact a_z's are positive). The FKG inequality says that if F, G are two such increasing functions of ψ, then

$$\mathbb{E}[F(\psi)G(\psi)] \geq \mathbb{E}[F(\psi)]\mathbb{E}[G(\psi)].$$

Hence

$$\mathbb{E}[\chi_{[t,\infty)}(S(\psi))\psi_y] \geq \mathbb{E}[\chi_{[t,\infty)}(S(\psi))]\mathbb{E}[\psi_y] = 0.$$

In the last equality we use the fact that $\psi \mapsto \psi_y$ is an odd function (but $\mathbb{E}$ is invariant w.r.t. the flip). Hence in the integral,

$$\int_{t=0}^{\infty} \mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right] f'(t)\,\mathrm{d}t$$

the fact that f' is monotone decreasing and $\mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right] \geq 0$ implies

$$\begin{aligned}\int_{t=0}^{\infty} \mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right] f'(t)\,\mathrm{d}t &\leq f'(0)\int_{t=0}^{\infty} \mathbb{E}\left[\chi_{[t,\infty)}(S(\psi))\psi_y\right]\,\mathrm{d}t \\ &= f'(0)\mathbb{E}\left[\int_{t=0}^{\infty} \chi_{[0,S(\psi)]}(t)\psi_y\,\mathrm{d}t\right] \\ &= f'(0)\mathbb{E}\left[\left(\int_{t=0}^{\infty} \chi_{[0,S(\psi)]}(t)\,\mathrm{d}t\right)\psi_y\right] \\ &= f'(0)\mathbb{E}\left[S(\psi)\psi_y\right].\end{aligned}$$

□

Lemma 7.12 (Non-negativity of four point correlation function)**.**

$$\mathbb{E}_\beta[\psi_{x,1}\psi_{x,2}\psi_{y,1}(A\psi_2)_x] \geq 0$$

Proof. By linearity, it suffices to show that for any sites $x, y, z \in \Lambda$, the four-point correlation function is non-negative:

$$\mathbb{E}_\beta[\psi_{x,1}\psi_{x,2}\psi_{y,1}\psi_{z,2}] \geq 0.$$

We begin by writing the expectation in terms of the a priori measure $\mathbb{E}_0$. Let $F(\psi) = \psi_{x,1}\psi_{x,2}\psi_{y,1}\psi_{z,2}$.

The expectation is

$$\mathbb{E}_\beta[F(\psi)] = \frac{\mathbb{E}_0[F(\psi)e^{-\beta H(\psi)}]}{\mathbb{E}_0[e^{-\beta H(\psi)}]} = \frac{\mathbb{E}_0[F(\psi)e^{\beta \sum_{i\sim j}\psi_i\cdot\psi_j}]}{Z_\beta}.$$

Since the partition function Z_β is positive, it suffices to show the numerator is non-negative. We expand the Boltzmann factor as a power series:

$$\mathbb{E}_0\left[F(\psi)\sum_{k=0}^{\infty}\frac{\beta^k}{k!}\left(\sum_{i\sim j}\psi_i\cdot\psi_j\right)^k\right] = \sum_{k=0}^{\infty}\frac{\beta^k}{k!}\mathbb{E}_0\left[F(\psi)\left(\sum_{i\sim j}\psi_i\cdot\psi_j\right)^k\right].$$

The proof reduces to showing that each term in this series is non-negative. A generic term from the expansion of $(\sum\psi_i\cdot\psi_j)^k$ is a product of spin components where each component index appears an even number of times (since $\psi_i\cdot\psi_j = \sum_a \psi_{i,a}\psi_{j,a}$). Our function $F(\psi)$ also has this property (component 1 appears twice, component 2 appears twice). Therefore, the entire product inside the expectation $\mathbb{E}_0$ consists of a product of spin components where every component index appears an even number of times.

The non-negativity of these terms is formally guaranteed by the GKS (Griffiths-Kelly-Sherman) inequalities, generalized for the O(N) model. □

Theorem 7.13 (Subharmonicity)**.** *The $O(N)$ model has the following subharmonicity in space*

$$\mathbb{E}_\beta[\psi_x\cdot\psi_y] \le \frac{\beta}{N}\mathbb{E}[(A\psi)_x\cdot\psi_y(1-\delta_{xy})] + \delta_{xy}$$

Proof. Note that for $x = y$ this follows trivially as $\mathbb{E}_\beta[\psi_x\cdot\psi_x] = \mathbb{E}_\beta[\|\psi_x\|^2] = 1$. For the $N = 1$, $x \ne y$ case, using the law of total expectation,

$$\mathbb{E}_\beta[\psi_x\psi_y] = \mathbb{E}_\beta[\psi_y\cdot\mathbb{E}_\beta[\psi_x|\{\psi_z\}_{z\ne x}]]$$

The inner expectation is computed as

$$\mathbb{E}_\beta[\psi_x|\{\psi_z\}_{z\ne x}]] = \frac{\sum_{\psi_x=\pm1}\psi_x\exp(\beta\psi_x(A\psi)_x)}{\sum_{\psi_x=\pm1}\exp(\beta\psi_x(A\psi)_x)} = \frac{2\sinh(\beta(A\psi)_x)}{2\cosh(\beta(A\psi)_x)} = \tanh(\beta(A\psi)_x)$$

And so we derive that $\mathbb{E}_\beta[\psi_x\psi_y] = \mathbb{E}_\beta[\psi_y\tanh(\beta(A\psi)_x)]$. Finally, applying Lemma 7.11 with the function $f(x) = \tanh(\beta x)$ and $a_z = 1$ if $x\sim z$ we conclude that $\mathbb{E}_\beta[\psi_x\psi_y] \le \beta\mathbb{E}[(A\psi)_x\psi_y]$. We next prove this for $N \ge 2$, $x \ne y$. Consider the family of automorphisms

$$\gamma_t^x : (\mathbb{S}^{N-1})^\Lambda \to (\mathbb{S}^{N-1})^\Lambda$$

defined by

$$\gamma_t^x(\psi_{y,\lambda}) = \begin{cases}\psi_{y,\lambda} & y\ne x \text{ or } \lambda\ne 1,2\\ (\cos t)\psi_{x,1} - (\sin t)\psi_{x,2} & y = x, \lambda = 1\\ (\sin t)\psi_{x,1} + (\cos t)\psi_{x,2} & y = x, \lambda = 2\end{cases}$$

Geometrically, γ_t^x acts on the spin ψ_x by a rotation by t in the 1-2 plane. Clearly, γ_t^x obeys the constraints from above, and so we can apply the local Ward identities. Consider the observable $Q \equiv Q(\psi) = \psi_{x,2}\psi_{y,1}$.

First note that $\dot{Q} = \psi_{x,1}\psi_{y,1}$. Also note that

$$(\dot{\psi})_x = (-\psi_{x,2}, \psi_{x,1}, 0, \ldots, 0)^T.$$

We know that

$$\partial_t H_t = \frac{1}{2}\partial_t \langle \psi_t, -\Delta\psi_t\rangle = \langle \dot{\psi}_t, -\Delta\psi_t\rangle$$

And so evaluating at $t = 0$, we get $\dot{H} = \langle\dot{\psi}, -\Delta\psi\rangle$
The local Ward identities therefore yield

$$\begin{aligned}\mathbb{E}_\beta[\psi_{x,1}\psi_{y,1}] &= \beta\mathbb{E}_\beta[\psi_{x,2}\psi_{y,1}(\psi_{x,1}(-\Delta\psi_2)_x - \psi_{x,2}(-\Delta\psi_1)_x)] \\ &= \beta\mathbb{E}_\beta[-\psi_{x,2}^2\psi_{y,1}(-\Delta\psi_1)_x + \psi_{x,1}\psi_{x,2}\psi_{y,1}(-\Delta\psi_2)_x] \\ &= -2d\beta\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}\psi_{x,1}] + \beta\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}(A\psi_1)_x] \\ &\quad + 2d\beta\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}\psi_{x,1}] - \beta\mathbb{E}_\beta[\psi_{x,1}\psi_{x,2}\psi_{y,1}(A\psi_2)_x]\end{aligned}$$

With Lemma 7.12, we conclude that

$$\begin{aligned}\mathbb{E}_\beta[\psi_{x,1}\psi_{y,1}] &\le \beta\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}(A\psi_1)_x] \\ &= \frac{\beta}{N}\left(N\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}(A\psi_1)_x]\right) \\ &= \frac{\beta}{N}\left(\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}(A\psi_1)_x] + \sum_{j=2}^{N}\mathbb{E}_\beta[\psi_{x,j}^2\psi_{y,1}(A\psi_1)_x]\right)\end{aligned}$$

because we can replace the 2nd component of ψ_x with any component not equal to 1, by symmetry. Simon's inequality states that

$$\mathbb{E}_\beta[\psi_{x,2}^2\psi_{y,1}(A\psi_1)_x] \le \mathbb{E}_\beta[\psi_{x,1}^2\psi_{y,1}(A\psi_1)_x]$$

And so it reduces to

$$\mathbb{E}_\beta[\psi_{x,1}\psi_{y,1}] \le \frac{\beta}{N}\left(\sum_{j=1}^{N}\mathbb{E}_\beta[\psi_{x,j}^2\psi_{y,1}(A\psi_1)_x]\right)$$

Noting that $\|\psi_x\|^2 = 1$, this further reduces to

$$\mathbb{E}_\beta[\psi_{x,1}\psi_{y,1}] \le \frac{\beta}{N}\mathbb{E}_\beta[\psi_{y,1}(A\psi_1)_x]$$

By symmetry, this holds for every other component besides 1 too, so we finally conclude that

$$\mathbb{E}_\beta[\psi_x \cdot \psi_y] \le \frac{\beta}{N}\mathbb{E}_\beta[(A\psi)_x \cdot \psi_y]$$

□

Theorem 7.14 (Exponential decay)**.** *Denote* $\tau_{xy} = \mathbb{E}_\beta[\psi_x \cdot \psi_y]$ *and* $\epsilon = \beta/N$*. For* $\epsilon < \frac{1}{2d}$*, we get exponential decay*

$$\tau_{xy} \le \frac{1}{\epsilon m}\exp(-|\log(2d\epsilon)|\|x-y\|_1)$$

where $m = \frac{1}{\epsilon} - 2d$*.*

Proof. We know for a ferromagnetic $O(N)$ model that $\tau_{xy} \ge 0$. Theorem 7.13 tells us that

$$[(\mathbb{1} - \epsilon A)\tau]_{xy} \le \delta_{xy}$$

This reduces to

$$[(-\Delta + m\mathbb{1})\tau]_{xy} \leq \frac{\delta_{xy}}{\epsilon}$$

Since the singular values of $-\Delta$ satisfy $\sigma(-\Delta) \subset [0, 4d]$, and using the fact that $[(-\Delta+m\mathbb{1})^{-1}]_{xy} \geq 0$, we deduce

$$\tau_{xy} \leq \frac{1}{\epsilon}[(-\Delta + m\mathbb{1})^{-1}]_{xy} = [(1 - \epsilon A)^{-1}]_{xy}$$

Now consider the Neumann expansion

$$(\mathbb{1} - \epsilon A)^{-1} = \sum_{n=0}^{\infty} \epsilon^n A^n \Rightarrow [(\mathbb{1} - \epsilon A)^{-1}]_{xy} = \sum_{n=0}^{\infty} \epsilon^n (A^n)_{xy}$$

Since $A_{xy} = 1$ only if $x \sim y$, even though this series seems to start at $n = 0$, it can only be non-zero if n is sufficiently large. Indeed, the first positive term of the series can only start from $n \geq \|x - y\|_1$, so that there are enough steps to go from x to y. This series expansion also shows that $[(\mathbb{1} - \epsilon A)^{-1}]_{xy} \geq 0$. As such the series expansion reduces to

$$[(\mathbb{1} - \epsilon A)^{-1}]_{xy} = \sum_{n=\|x-y\|_1}^{\infty} \epsilon^n (A^n)_{xy}$$

Noting that $(A^n)_{xy} \leq (2d)^n$ (the total number of walks from x to y has this upper bound because the degree of each vertex is $2d$ at most) and $\epsilon < 1/2d$, we can observe exponential decay.

$$[(\mathbb{1} - \epsilon A)^{-1}]_{xy} \leq \sum_{n=\|x-y\|_1}^{\infty} (2d\epsilon)^n = \frac{(2d\epsilon)^{\|x-y\|_1}}{1 - 2d\epsilon} = \frac{1}{1 - 2d\epsilon} \exp(-|\log(2d\epsilon)|\|x - y\|_1)$$

Finally, conclude that

$$\tau_{xy} \leq \frac{1}{\epsilon m} \exp(-|\log(2d\epsilon)|\|x - y\|_1)$$

Note that as $m \to 0^+$, $|\log(2d\epsilon)| = |\log(1 - m\epsilon)| \to m\epsilon$, so the bound becomes

$$\tau_{xy} \leq \frac{1}{\epsilon m} \exp(-c_d(m)\|x - y\|_1)$$

with $c_d(m) \to m/2d$ as $m \to 0^+$. □

7.4 Optimal decay rate

Theorem 7.15 (Exponential decay, optimal decay rate)**.** *In fact for any $m > 0$ small,*

$$\tau_{xy} \leq \frac{1}{\epsilon}[(-\Delta + m\mathbb{1})^{-1}]_{xy} \leq \frac{1}{\epsilon m} \exp(-c_d(m)\|x - y\|_1)$$

with $c_d(m) \sim \sqrt{m}$ as $m \to 0^+$.

Proof. The key is to diagonalize A. Assume periodic boundary conditions on Λ. The discrete Fourier series is given by the following set of vectors in $\mathbb{C}^{|\Lambda|}$

$$\psi_{m,x} := \frac{1}{\sqrt{|\Lambda|}} \exp\left(\frac{2\pi i}{2L + 1} m \cdot x\right)$$

for $x, m \in \Lambda$. We can compute

$$A\psi_m = 2\sum_{j=1}^{d} \cos\left(\frac{2\pi m_j}{2L+1}\right)\psi_m$$

$\{\psi_m\}_m$ forms a complete eigenbasis that is orthonormal. The discrete Fourier series is the unitary $|\Lambda| \times |\Lambda|$ matrix whose columns are the ψ_ms. It is defined as

$$(\mathscr{F}\varphi)_m = \sum_{x\in\Lambda} \varphi_x \psi_{m,x} = \frac{1}{\sqrt{|\Lambda|}} \sum_{x\in\Lambda} \varphi_x \exp\left(\frac{2\pi i}{2L+1} m\cdot x\right)$$

The inverse is given by

$$(\mathscr{F}^*\hat{\varphi})_x = \frac{1}{\sqrt{|\Lambda|}} \sum_{m\in\Lambda} \hat{\varphi_m} \exp\left(-\frac{2\pi i}{2L+1} m\cdot x\right)$$

Then

$$A = \mathscr{F}^*\left(\left\{2\sum_{j=1}^{d}\cos\left(\frac{2\pi m_j}{2L+1}\right)\right\}_{m\in\Lambda}\right)\mathscr{F}.$$

As such clearly

$$(1-\varepsilon A)^{-1} = \mathscr{F}^*\left(\left\{\frac{1}{1-2\varepsilon\sum_{j=1}^{d}\cos\left(\frac{2\pi m_j}{2L+1}\right)}\right\}_{m\in\Lambda}\right)\mathscr{F}$$

and so

$$\begin{aligned}
[(\mathbb{1}-\varepsilon A)^{-1}]_{xy} &= \sum_{m\in\Lambda} \frac{\psi_m(x)\overline{\psi_m(y)}}{1-2\varepsilon\sum_{j=1}^{d}\cos\left(\frac{2\pi m_j}{2L+1}\right)} \\
&= \frac{1}{|\Lambda|}\sum_{m\in\Lambda} \frac{\exp\left(i\frac{2\pi m}{2L+1}\cdot(x-y)\right)}{1-2\varepsilon\sum_{j=1}^{d}\cos\left(\frac{2\pi m_j}{2L+1}\right)}.
\end{aligned}$$

Now, as $L\to\infty$, this becomes a Riemann sum approximation of a Riemann integral with which we may write

$$G_\varepsilon(x-y) := \lim_{L\to\infty}[(\mathbb{1}-\varepsilon A)^{-1}]_{xy} = \frac{1}{(2\pi)^d}\int_{k\in[-\pi,\pi]^d} dk \frac{\exp(ik\cdot(x-y))}{1-2\varepsilon\sum_{j=1}^{d}\cos(k_j)}.$$

Let us re-interpret each $k_j \in [-\pi,\pi]$ integral as a complex contour integral of the unit circle, i.e., $z_j = e^{ik_j}$, so we get

$$G_\varepsilon(x-y) = \frac{1}{(2\pi i)^d}\oint_{z_1\in S^1}\frac{dz_1}{z_1}\cdots\oint_{z_d\in S^1}\frac{dz_d}{z_d}\frac{\prod_{j=1}^{d} z_j^{x_j-y_j}}{1-\varepsilon\sum_{j=1}^{d}\left(z_j+\frac{1}{z_j}\right)}.$$

We may deform the contour integrations from the circle to slightly larger or smaller radii (for each j pick if we need to enlarge or shrink based on the sign of $x_j - y_j$). We can only change the radii as long as we don't create zeros in the denominator

$$z \mapsto 1-\varepsilon\sum_{j=1}^{d}\left(z_j+\frac{1}{z_j}\right).$$

Let us choose tentatively a sequence of changes $\{r_j\}_{j=1}^d$ (not change means $r_j = 1$). If $x_j - y_j > 0$ we choose $r_j \in (0,1)$ and if $x_j - y_j < 0$ we choose $r_j > 1$. Then, invoking Cauchy's theorem,

$$G_\varepsilon(x-y) = \prod_{j=1}^d r_j^{x_j - y_j} \frac{1}{(2\pi i)^d} \oint_{z_1 \in S^1} \frac{dz_1}{z_1} \cdots \oint_{z_d \in S^1} \frac{dz_d}{z_d} \frac{\prod_{j=1}^d z_j^{x_j-y_j}}{1 - \varepsilon \sum_{j=1}^d \left(r_j z_j + \frac{1}{r_j z_j} \right)}.$$

Assuming we haven't crossed any singularities. Also note that

$$\left| 1 - \varepsilon \sum_{j=1}^d \left(r_j z_j + \frac{1}{r_j z_j} \right) \right| \geq 1 - \varepsilon \sum_{j=1}^d \left| r_j z_j + \frac{1}{r_j z_j} \right| = 1 - \varepsilon \sum_{j=1}^d \left(r_j + \frac{1}{r_j} \right).$$

Hence

$$\begin{aligned} |G_\varepsilon(x-y)| &\leq \frac{\prod_{j=1}^d r_j^{x_j-y_j}}{\left| 1 - \varepsilon \sum_{j=1}^d \left(r_j + \frac{1}{r_j} \right) \right|} \frac{1}{(2\pi)^d} \left| \oint_{z_1 \in S^1} \cdots \oint_{z_d \in S^1} dz_1 \dots dz_d \right| \\ &= \frac{\prod_{j=1}^d r_j^{x_j-y_j}}{1 - \varepsilon \sum_{j=1}^d \left(r_j + \frac{1}{r_j} \right)}. \end{aligned}$$

Assuming we haven't crossed any singularities. Since the denominator (for all $r_j = 1$) is real and positive, a sufficient condition to never cross zeros is indeed

$$1 - \varepsilon \sum_{j=1}^d \left(r_j + \frac{1}{r_j} \right) > 0.$$

We simplify life by choosing isotropically: $r_j = \exp(\pm a)$ (depending on the sign of $x_j - y_j$ as described above) for some $a > 0$. Then

$$\prod_{j=1}^d r_j^{x_j-y_j} = \exp(-a\|x-y\|_1)$$

and

$$1 - \varepsilon \sum_{j=1}^d \left(r_j + \frac{1}{r_j} \right) = 1 - 2d\varepsilon \cosh(a)$$

which yields the condition

$$\cosh(a) < \frac{1}{2d\varepsilon}.$$

$$|G(x-y)| \leq \frac{\exp(-a\|x-y\|_1)}{1 - 2d\varepsilon \cosh(a)} \qquad \left(a < \operatorname{arccosh}\left(\frac{1}{2d\varepsilon} \right) \right).$$

Note that

$$\operatorname{arccosh}\left(\frac{1}{2d\varepsilon} \right) = \sqrt{2 - 4d\varepsilon} + \mathcal{O}\left((1-2d\varepsilon)^{3/2} \right)$$

which yields the optimal square root decay rate. ☐

8 Joshua: The Peierls solution for $N = 1$, $d \geq 2$

8.1 Setup

8.1.1 Objective

Let $\Lambda = [-L, L]^2 \cap \mathbb{Z}^2$ with a configuration of spins $\psi = \{\psi_x\}_{x \in \Lambda} \subseteq (\mathbb{S}^0)^{|\Lambda|} = \{\pm 1\}^{|\Lambda|}$. Our goal is to show that the two-point function under free boundary conditions, $\mathbb{E}_{\text{free}}[\psi_x \psi_y]$, does not decay as $L \to \infty$. Physically, this result demonstrates that there is some long-range order even as the lattice grows. A crucial intermediate step is to show that the spontaneous magnetization has a lower bound for sufficiently low temperatures, independent of L.

8.1.2 Borders

Every configuration admits a set of borders drawn along edges separating + and − spins (see Figure Figure 1). Note that each border is oriented. Our choice of orientation follows these rules:

1. Keep the − spins on the left and the + spins on the right.
2. If there is ambiguity, always turn to the left.

Borders are always closed. In Figure Figure 1a, the top left border is open, while the bottom border is closed. In Figure Figure 1b, there are three borders (due to the left-turn rule), each with a length of 4 unit segments. The right diagram exhibits *uniform boundary conditions*, in that all spins on the outer boundary are +.

+	−	+	+	+
−	−	+	+	+
+	+	+	+	+
+	+	+	−	+
+	+	+	+	+

(a) Mixed boundaries

+	+	+	+	+
+	+	+	−	+
+	+	−	+	+
+	−	+	+	+
+	+	+	+	+

(b) Uniform + boundaries

Figure 1: Visualizing borders in lattice configurations.

8.1.3 Energy and Probability

Recall that the Hamiltonian, or energy, can be rewritten as:

$$H(\psi) := -\sum_{x \sim y} \psi_x \psi_y$$

Note that (up to an additive constant) the Hamiltonian is twice the sum of the lengths of all borders, since a misaligned pair of spins (which is precisely what creates a unit of border) contributes +1 to the sum relative to −1 in the aligned case. The probability of any configuration ψ is then:

$$\mathbb{P}[\{\psi\}] = \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}c}(\psi) = \frac{\exp(-\beta H(\psi))}{\sum_{\psi'} \exp(-\beta H(\psi'))}$$

where the sum is over all (finitely many) configurations, and c is the counting measure.

8.2 Long-Range Order at Low Temperature

8.2.1 Motivation

Let $\Lambda_\psi^+, \Lambda_\psi^- \subseteq \Lambda$ be the vertices with positive and negative spins, respectively, in a given configuration. Then:

$$m_\Lambda := \frac{|\Lambda_\psi^+| - |\Lambda_\psi^-|}{|\Lambda|} = 1 - 2 \cdot \frac{|\Lambda_\psi^-|}{|\Lambda|} = -1 + 2 \cdot \frac{|\Lambda_\psi^+|}{|\Lambda|}$$

is the magnetization of the configuration, since $|\Lambda_\psi^+| + |\Lambda_\psi^-| = |\Lambda|$. Peierls showed that:

$$\frac{1}{|\Lambda|} \mathbb{E}_+[|\Lambda_\psi^-|] \leq C(\beta) \overset{\beta \to \infty}{\longrightarrow} 0$$

or equivalently, that the magnitude of magnetization does not vanish. Compare this with the free boundary case, which has:

$$\frac{1}{|\Lambda|}\mathbb{E}_{\text{free}}[|\Lambda_\psi^-|] = \frac{1}{2}$$

by symmetry.

Why is this bound important? Suppose it is true. Fix uniform + boundary conditions. We may equivalently write the magnetization as:

$$\begin{aligned}\mathbb{E}_+[\psi_x] = \mathbb{P}_+[\psi_x = +1] - \mathbb{P}_+[\psi_x = -1] &= \frac{1}{|\Lambda|}(\mathbb{E}_+[|\Lambda_\psi^+| - |\Lambda_\psi^-|]) \\ = \mathbb{E}_+[m_\Lambda] = 1 - 2\frac{\mathbb{E}_+[|\Lambda_\psi^-|]}{|\Lambda|} &\geq 1 - 2C(\beta) > 0\end{aligned}$$

for β large enough, thanks to the translation-invariance of the infinite-volume Gibbs state. In this case, we have:

$$\mathbb{E}_+[\psi_x\psi_y] \geq \mathbb{E}_+[\psi_x]\mathbb{E}_+[\psi_y] > 0$$

as well, by Griffiths' inequality. This result generalizes to free boundary conditions. For any zero external field, any infinite-volume Gibbs state is a convex combination of the uniform + boundary case and the uniform − boundary case. Furthermore:

$$\mathbb{E}_{\text{free}}[\psi_x\psi_y] = \frac{1}{2}\mathbb{E}_+[\psi_x\psi_y] + \frac{1}{2}\mathbb{E}_-[\psi_x\psi_y] > 0$$

where the last step follows by symmetry. Hence $\liminf_{\|x-y\|\to\infty} \mathbb{E}_{\text{free}}[\psi_x\psi_y] > 0$ as well, and the two-point function does not decay in the free boundary case either.

8.2.2 Intuition

Why should the magnetization bound be true in $d = 2$? With uniform + boundary conditions, a collection of − spins in a sea of + spins is surrounded by a *closed contour.*

- In 2D, the *energy* of such a contour grows linearly with its length, so the *probability* of seeing that contour is weighted by the Boltzmann factor $\exp(-\beta E(b)) = \exp(-2\beta b)$.
- Meanwhile, it turns out that the *entropy* (the number of such contours) grows only exponentially in the length with rate 3, i.e., on the order of 3^b up to polynomial factors.
- Hence, for β sufficiently large (low temperature), the expected "amount of border" is small, and so we will see net magnetization.

In contrast, this argument fails in $d = 1$ because every "border" is just a pair of endpoints, having constant energy which does not grow with size.

8.2.3 The Uniform Boundary Case

We consider the class of configurations Ω in which all $8L - 4$ spins on the outer boundary of Λ are +. Some simple observations we can make about configurations in Ω:

- All borders are closed, and therefore every − is enclosed by at least one closed border.
- Border lengths must be even and no less than 4, i.e., they may take values in $2\mathbb{N} + 2$.
- A border of length b encloses an area at most $b^2/16$ (thanks to the isoperimetric inequality; see Section **??** for proof), and thus contains at most that same number of spins.

The final observation is simply an argument for why the square offers the highest area to perimeter ratio among shapes on the lattice, but more formally it can be shown as:

Lemma 8.1 (Isoperimetric Inequality on $\mathbb{Z}^2$)**.** *A closed border of length b encloses at most $b^2/16$ vertices.*

Proof. Consider any border B. We define the *bounding box* of the region enclosed by B by its horizontal width W and vertical height H. Specifically, if the vertices enclosed range from $x_{\min}$ to $x_{\max}$ and $y_{\min}$ to $y_{\max}$, then $W = x_{\max} - x_{\min}$ and $H = y_{\max} - y_{\min}$ (representing the dimensions of the dual rectangle enclosing the region).

Consider the projection of the path B onto the horizontal axis. To span the width W, the path must travel at least W units to the right and W units to the left. Thus, the number of horizontal steps in B satisfies $b_{\text{horiz}} \geq 2W$. By the same logic for the vertical axis, the number of vertical steps satisfies $b_{\text{vert}} \geq 2H$.

The total length is $b = b_{\text{horiz}} + b_{\text{vert}}$. Therefore:

$$b \geq 2W + 2H \implies W + H \leq \frac{b}{2}$$

The number of spins (vertices) enclosed by the border is bounded by the area of this bounding box, $A \leq W \cdot H$. We maximize the product $W \cdot H$ subject to the constraint $W + H \leq b/2$. By the AM-GM inequality, the product is maximized when $W = H = b/4$. Thus:

$$A \leq \left(\frac{b}{4}\right)\left(\frac{b}{4}\right) = \frac{b^2}{16}$$

as desired. □

With these observations, we can bound the number of negative spins as:

$$|\Lambda_\psi^-| \leq \sum_{b \in 2\mathbb{N}+2} \frac{b^2}{16} \sum_{i=1}^{\alpha(b)} \chi_b^{(B_i)}$$

where $\alpha(b) \equiv \alpha_L(b)$ is the number of possible borders of length b, and $\chi_b^{(B_i)}$ is the indicator for the ith border of length b (according to some arbitrary indexing scheme) occurring in a configuration of Ω. We will bound this expression (in expectation) in two steps.

Lemma 8.2 (Border Bound)**.** *The number of possible closed, oriented borders of length b contained in a configuration in Ω is bounded by:*

$$\alpha(b) \leq \frac{16L^2 3^b}{3b} \propto L^2 b^{-1} 3^b$$

Proof. A border is a sequence of connected line segments on the lattice that separates + sites from − sites. To bound $\alpha(b)$, we make three observations:

1. To begin a path, we must first choose a starting location and direction. The lattice contains $4L^2$ vertices, each with 4 adjacent edges, for $16L^2$ possible first choices.

2. At each subsequent step, we arrive at a new vertex from which there are four possible directions for the next segment. However, one of those directions would simply reverse the direction just traversed, leaving at most 3 choices.

3. Every path constructed as such is overcounted by a factor of b, since it distinguishes between the b different possible starting points.

Combining these observations, the total number of directed paths of length b is:

$$\alpha(b) \leq 16L^2 \times \underbrace{3 \times 3 \times \cdots \times 3}_{b-1 \text{ times}} \times b^{-1} = \frac{16L^2 3^b}{3b}$$

which is precisely the desired bound. □

Lemma 8.3 (Probability Bound)**.** *For any $b \in 2\mathbb{N}+2$ and $1 \leq i \leq \alpha(b)$ integer, $\mathbb{E}_+ \chi_b^{(B_i)} \leq \exp(-2\beta b)$.*

Proof. If B is a (closed) border of length b and indexed by i, let $\Omega_B \subseteq \Omega$ be the set of configurations which contain B. Fix $\psi \in \Omega_B$ and observe that flipping all spins inside B yields $\psi^* \in \Omega$. Hence $\Omega_B^* = \{\psi^* : \psi \in \Omega_B\} \subseteq \Omega$. By definition:

$$\mathbb{E}_+ \chi_b^{(B_i)} = \frac{\sum_{\psi \in \Omega_B} \exp(-\beta H(\psi))}{\sum_{\psi \in \Omega} \exp(-\beta H(\psi))}$$

Then, because every misaligned spin in ψ which is on the inner boundary of B would be aligned in ψ^* (while all other contributions to the energy remain unchanged), flipping its contribution from $-\psi_x\psi_y = +1$ to $-\psi_x\psi_y = -1$, we have:

$$H(\psi^*) = H(\psi) - 2b$$

Now,

$$\begin{aligned}\sum_{\psi \in \Omega} \exp(-\beta H(\psi)) \geq \sum_{\psi^* \in \Omega_B} \exp(-\beta H(\psi^*)) &= \sum_{\psi \in \Omega_B} \exp(-\beta(H(\psi) - 2b)) \\ &= \exp(2\beta b) \sum_{\psi \in \Omega_B} \exp(-\beta H(\psi))\end{aligned}$$

so that from the first equation, $\mathbb{E}_+ \chi_b^{(B_i)} \leq \exp(-2\beta b)$ as desired. □

With these two results, we complete our argument: define $\kappa = 3\exp(-2\beta)$, so that by linearity of expectation:

$$\begin{aligned}|\Lambda|^{-1}\mathbb{E}_+ \left[|\Lambda_\psi^-|\right] &\leq \frac{1}{(2L+1)^2}\mathbb{E}_+ \left[\sum_{b \in 2\mathbb{N}+2} \frac{b^2}{16} \sum_{i=1}^{\alpha(b)} \mathbb{E}\chi_b^{(B_i)}\right] \\ &\leq \frac{1}{4L^2} \sum_{b \in 2\mathbb{N}+2} \frac{b^2}{16} \left(\frac{16L^2 3^b}{3b}\right)(\exp(-2\beta b)) \\ &= \frac{1}{12} \sum_{b \in 2\mathbb{N}+2} b \exp(-b(2\beta - \log 3)) =: C(\beta)\end{aligned}$$

which exhibits exponential decay (thus converging) insofar as $\beta > 2^{-1}\log 3$. In particular, $C(\beta) \to 0$ as $\beta \to \infty$, independently of L as desired.

9 Matias: Mermin-Wagner (Pfister)

9.1 Setup

Let $\Lambda_L := [-L,L]^d \cap \mathbb{Z}^d$ and $\psi : \Lambda \to \mathbb{S}^{N-1}$ with $N \geq 2$, be a random field distributed according to the Gibbs measure whose partition function is

$$Z_{\beta,\Lambda_L} := \int_{\psi:\Lambda_L\to\mathbb{S}^{N-1}} \exp\left(-\frac{1}{2}\beta \langle \psi, -\Delta\psi\rangle\right) \mathrm{d}\mu(\psi)$$

where μ is the a-priori measure, in this case the $|\Lambda|$-fold product of volume measures on $\mathbb{S}^{N-1}$ and

$$\begin{aligned}\langle \psi, -\Delta\psi\rangle &\equiv \langle \psi, -\Delta \otimes \mathbb{1}_N \psi\rangle_{\mathbb{R}^\Lambda\otimes\mathbb{R}^N} \\ &= \sum_{i=1}^{N} \sum_{x,y\in\Lambda_L} \psi_{x,i} (-\Delta)_{xy} \psi_{y,i} \,.\end{aligned}$$

With this setup, he goal of this section is to prove the following theorem.

Theorem 9.1 (Mermin-Wagner a la Pfister)**.** Let $N \geq 2$ and $d \leq 2$. Then any *infinite volume Gibbs measure is rotation invariant. I.e., let* $F : \left(\mathbb{S}^{N-1}\right)^{\mathbb{Z}^d} \to \mathbb{R}$ *be a measurable function which depends only on finitely-many spins. Then*

$$\mathbb{E}_\beta [F] = \mathbb{E}_\beta [F \circ T_R] \,.$$

where $T_R : \left(\mathbb{S}^{N-1}\right)^{\mathbb{Z}^d} \to \left(\mathbb{S}^{N-1}\right)^{\mathbb{Z}^d}$ *is the global constant rotation of all spins by some* $R \in O(N)$.

9.2 The proof

9.2.1 A finite box and a buffer

Let such $F : \left(\mathbb{S}^{N-1}\right)^{\mathbb{Z}^d} \to \mathbb{R}$ be a measurable function which depends only on finitely-many spins and pick L sufficiently large so that F only depends on the spins in Λ_L. Now pick some $B \in \mathbb{N}$ and consider the somewhat larger box Λ_{L+B}. Let $R \in O(N)$ and without loss of generality assume that R only rotates within the $1-2$ plane, say by some angle $\alpha \in [0, 2\pi)$.

Our general strategy will be to rotate the spins uniformly within Λ_L, and then gradually taper off the rotation from φ to nothing as we go from $\partial\Lambda_L$ to $\partial\Lambda_{L+B}$. Let us thus define now the *space dependent rotation*

$$\varphi : \Lambda_{L+B} \to [0, 2\pi)$$

where we'll have $\varphi_x = \alpha$ if $x \in \Lambda_L$ and $\varphi_x \to 0$ as x varies from $\partial\Lambda_L$ to $\partial\Lambda_{L+B}$ (just how this profile is taken we'll specify shortly).

Let ψ^φ be the rotated field, so that $\psi^\varphi_x = \psi_x$ rotated by α for all $x \in \Lambda_L$ and we simply have $\psi^\varphi_x = \psi_x$ for all $x \notin \Lambda_L \cup \Lambda_{L+B}$.

We will use the following

Lemma 9.2. *We have*

$$\frac{1}{2}\langle \psi^\varphi, -\Delta\psi^\varphi\rangle + \frac{1}{2}\left\langle \psi^{-\varphi}, -\Delta\psi^{-\varphi}\right\rangle - \frac{1}{2}\langle \varphi, -\Delta\varphi\rangle \leq \langle \psi, -\Delta\psi\rangle \,.$$

Proof. We have

$$\langle \psi, -\Delta\psi\rangle = \sum_{x,y\in\Lambda_{L+B}} (-\Delta)_{xy} \langle \psi_x, \psi_y\rangle_{\mathbb{R}^N}$$

so we may write

$$\langle \psi^\varphi, -\Delta\psi^\varphi\rangle = \sum_{x,y\in\Lambda_{L+B}} (-\Delta)_{xy} \left\langle \psi_x, \psi_y^{\varphi_y-\varphi_x}\right\rangle_{\mathbb{R}^N}$$

and if we parametrize $\psi_x^\alpha \equiv \begin{bmatrix} \cos(\alpha)\psi_{x1} - \sin(\alpha)\psi_{x2} \\ \sin(\alpha)\psi_{x1} + \cos(\alpha)\psi_{x2} \\ \psi_{x3} \\ \dots \end{bmatrix}$ (i.e. since we are just rotating in the $1-2$ plane) then

$$\begin{aligned} \left\langle \psi_x, \psi_y^\alpha \right\rangle_{\mathbb{R}^N} &= \langle \psi_x, \psi_y \rangle - \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} \psi_{y1} \\ \psi_{y2} \end{bmatrix} \right\rangle + \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} \cos(\alpha)\psi_{y1} - \sin(\alpha)\psi_{y2} \\ \sin(\alpha)\psi_{y1} + \cos(\alpha)\psi_{y2} \end{bmatrix} \right\rangle \\ &= \langle \psi_x, \psi_y \rangle + \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} (\cos(\alpha) - 1)\psi_{y1} - \sin(\alpha)\psi_{y2} \\ \sin(\alpha)\psi_{y1} + (\cos(\alpha) - 1)\psi_{y2} \end{bmatrix} \right\rangle \end{aligned}$$

So

$$\begin{aligned} \frac{1}{2}\left\langle \psi_x, \psi_y^\alpha \right\rangle_{\mathbb{R}^N} + \frac{1}{2}\left\langle \psi_x, \psi_y^{-\alpha} \right\rangle_{\mathbb{R}^N} - \langle \psi_x, \psi_y \rangle &= \frac{1}{2}\left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} (\cos(\alpha) - 1)\psi_{y1} \\ (\cos(\alpha) - 1)\psi_{y2} \end{bmatrix} \right\rangle + \frac{1}{2}\left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} (\cos(\alpha) - 1)\psi_{y1} \\ (\cos(\alpha) - 1)\psi_{y2} \end{bmatrix} \right\rangle \\ &= \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} \psi_{y1} \\ \psi_{y2} \end{bmatrix} \right\rangle (\cos(\alpha) - 1) \end{aligned}$$

Hence

$$\begin{aligned} \frac{1}{2}\langle \psi^\varphi, -\Delta\psi^\varphi \rangle + \frac{1}{2}\langle \psi^{-\varphi}, -\Delta\psi^{-\varphi} \rangle - \langle \psi, -\Delta\psi \rangle &= \sum_{x,y \in \Lambda_{L+B}} (-\Delta)_{xy} \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} \psi_{y1} \\ \psi_{y2} \end{bmatrix} \right\rangle (\cos(\varphi_y - \varphi_x) - 1) \\ &\leq \sum_{x,y \in \Lambda_{L+B}} (-\Delta)_{xy} \left\langle \begin{bmatrix} \psi_{x1} \\ \psi_{x2} \end{bmatrix}, \begin{bmatrix} \psi_{y1} \\ \psi_{y2} \end{bmatrix} \right\rangle \frac{1}{2}(\varphi_y - \varphi_x)^2 \\ &\leq \sum_{x,y \in \Lambda_{L+B} : x \sim y} \frac{1}{2}(\varphi_y - \varphi_x)^2 \\ &= \frac{1}{2}\langle \varphi, -\Delta\varphi \rangle \, . \end{aligned}$$

□

Lemma 9.3 (Optimal energy)**.** *In $d \leq 2$ there is a choice of field $\varphi : \Lambda_{L+B} \to [0, 2\pi)$ so that $\varphi = \alpha$ within Λ_L, $\varphi = 0$ on $\partial\Lambda_{L+B}$ and its energy is bounded by*

$$K := \frac{C}{\log\left(1 + \frac{B}{L}\right)} \quad \exists\, C \text{ independent of } L \text{ and } B.$$

Proof. We are trying to minimize the energy of a field which is constant within Λ_L and then tapers off to zero as x varies from $\partial\Lambda_L$ to $\partial\Lambda_{L+B}$. To do so, we consider the gradual rotation:

$$\varphi_x \equiv \varphi(\|x\|) \quad := \quad \begin{cases} \alpha & \|x\| \leq L \\ \alpha\frac{\log\left(\frac{L+B+1}{\|x\|+1}\right)}{\log\left(\frac{L+B+1}{L+1}\right)} & L < \|x\| < L + B \\ 0 & \|x\| \geq L + B \end{cases} \qquad (x \in \Lambda_{L+B}) \, .$$

Such a rotation causes the energy to change by:

$$E(\varphi) := H(\psi^\varphi) - H(\psi) = \tfrac{1}{2}\left(\langle \psi^\varphi, -\Delta\psi^\varphi \rangle - \langle \psi, -\Delta\psi \rangle\right)$$

Observe the following by our earlier lemma:

$$\begin{aligned} \frac{1}{2}\langle \psi^\varphi, -\Delta\psi^\varphi \rangle + \frac{1}{2}\langle \psi^{-\varphi}, -\Delta\psi^{-\varphi} \rangle - \langle \psi, -\Delta\psi \rangle &= \frac{1}{2}\left(\langle \psi^\varphi, -\Delta\psi^\varphi \rangle - \langle \psi, -\Delta\psi \rangle\right) + \frac{1}{2}\left(\langle \psi^{-\varphi}, -\Delta\psi^{-\varphi} \rangle - \langle \psi, -\Delta\psi \rangle\right) \\ &\leq \frac{1}{2}\langle \varphi, -\Delta\varphi \rangle \end{aligned}$$

Hence we get either $\frac{1}{2}\left(\langle\psi^{\varphi},-\Delta\psi^{\varphi}\rangle-\langle\psi,-\Delta\psi\rangle\right)=H(\psi^{\varphi})-H(\psi)\leq\frac{1}{2}\langle\varphi,-\Delta\varphi\rangle$ or $H(\psi^{-\varphi})-H(\psi)\leq\frac{1}{2}\langle\varphi,-\Delta\varphi\rangle$. Below we will assume WLOG that the former is true (if the latter were instead true just replace φ by $-\varphi$ below).

For any $L\leq\|x\|\leq L+B$, we have:

$$\frac{\partial\varphi_x}{\partial\|x\|}\equiv\varphi'(\|x\|)=-\frac{\alpha}{\log\left(\frac{L+B+1}{L+1}\right)}\frac{1}{\|x\|+1}$$

Thus for $x\sim y,\ x\in\Lambda_{L+B}$, we have the following bound by the mean value theorem:

$$|\varphi_x-\varphi_y|\leq\sup_{t\in[\|x\|,\|y\|]}|\varphi_x'(t)|\times|\|x\|-\|y\||\leq\frac{\alpha}{\log\left(\frac{L+B+1}{L+1}\right)}\frac{1}{\min\{\|x\|,\|y\|\}+1},$$

where we used the fact that $x\sim y\implies|\|x\|-\|y\||\leq 1$. Hence we see:

$$\begin{aligned}E(\varphi)\leq\frac{1}{2}\langle\varphi,-\Delta\varphi\rangle=\frac{1}{2}\sum_{x\sim y}(\varphi_x-\varphi_y)^2\leq\frac{1}{2}&\sum_{\substack{x\sim y\\ \min(\|x\|,\|y\|)\in[L,L+B]}}\frac{\alpha^2}{\left(\log\left(\frac{L+B+1}{L+1}\right)\right)^2}\frac{1}{(\min\{\|x\|,\|y\|\}+1)^2}\\&=C'\frac{\alpha^2}{\left(\log\left(\frac{L+B+1}{L+1}\right)\right)^2}\sum_{r=L}^{L+B}\frac{r^{d-1}}{(r+1)^2},\end{aligned}$$

where in the final step we reorganized the sum by collecting all edges whose endpoint closer to the origin has distance $r=\min(\|x\|,\|y\|)$ from the center. For a lattice in $\mathbb{Z}^d$ and a fixed distance from the center, r, there are $\mathcal{O}(r^{d-1})$ points with $r=\min\{\|x\|,\|y\|\}$ (this is clear from recognizing that the number of the these points is proportional to the lattice surface area of the sphere of radius r and a sphere in dimension d has surface area $\mathcal{O}(r^{d-1})$). Thus for $d=2$, we see that this sum behaves like a harmonic sum, i.e.:

$$\int_L^{L+B}\frac{r}{(r+1)^2}\,dr\lesssim\log\frac{L+B+1}{L+1}\implies E(\varphi)\leq\widetilde{C}\frac{\alpha^2}{\left(\log\left(\frac{L+B+1}{L+1}\right)\right)^2}\log\frac{L+B+1}{L+1}=\frac{\widetilde{C}\alpha^2}{\log\left(\frac{L+B+1}{L+1}\right)}.$$

So, by defining $C:=\widetilde{C}\alpha^2$, we obtain the desired inequality.

We note that this is precisely the part of the proof that relies heavily on the assumption that $d\leq 2$. Following the above analysis, one sees that for $d=1$, we can get a bound $C\frac{\alpha^2}{B}$, which is a stronger bound than we had in the $d=2$ case above. On the other hand, for $d\geq 3$, one sees that the above analysis fails and we instead find a bound like $\frac{\alpha^2}{L^{d-2}}$. □

Let us finish the proof the the theorem. Assume WLOG $F \geq 0$ (see below for why this doesn't lose generality). Then:

$$
\begin{aligned}
Z_{\beta,\Lambda_{L+B}}\,\mathbb{E}_{\beta,\Lambda_{L+B}}[F\circ T_R] &= \int_{\psi:\Lambda_{L+B}\to\mathbb{S}^{N-1}} \exp\left(-\tfrac{1}{2}\beta\,\langle\psi,-\Delta\psi\rangle\right) F(T_R(\psi))\;\mathrm{d}\mu(\psi)\\
&\leq \exp\left(\tfrac{1}{2}\beta K\right)\int_\psi \exp\Big(-\tfrac{1}{2}\beta\Big(\tfrac{1}{2}\langle\psi^{\varphi},-\Delta\psi^{\varphi}\rangle+\tfrac{1}{2}\langle\psi^{-\varphi},-\Delta\psi^{-\varphi}\rangle\Big)\Big)\, F\left(\psi^{\pm\varphi}\right)\mathrm{d}\mu(\psi)\\
&= \exp\left(\tfrac{1}{2}\beta K\right)\int_\psi \exp\Big(-\tfrac{1}{2}\beta\Big(\tfrac{1}{2}\langle\psi^{\varphi},-\Delta\psi^{\varphi}\rangle+\tfrac{1}{2}\langle\psi^{-\varphi},-\Delta\psi^{-\varphi}\rangle\Big)\Big)\,\sqrt{F(\psi^{\varphi})\,F(\psi^{-\varphi})}\,\mathrm{d}\mu(\psi)\\
&\overset{\text{C.S.}}{\leq} \exp\left(\tfrac{1}{2}\beta K\right)\sqrt{\int_\psi \exp\left(-\tfrac{1}{2}\beta\langle\psi^{\varphi},-\Delta\psi^{\varphi}\rangle\right)F(\psi^{\varphi})\,\mathrm{d}\mu(\psi)}\\
&\qquad\times\sqrt{\int_\psi \exp\left(-\tfrac{1}{2}\beta\langle\psi^{-\varphi},-\Delta\psi^{-\varphi}\rangle\right)F(\psi^{-\varphi})\,\mathrm{d}\mu(\psi)}\\
&= \exp\left(\tfrac{1}{2}\beta K\right)\sqrt{\int_\psi \exp\left(-\tfrac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)F(\psi)\,\mathrm{d}\mu(\psi^{-\varphi})}\times\sqrt{\int_\psi \exp\left(-\tfrac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)F(\psi)\,\mathrm{d}\mu(\psi^{\varphi})}\\
&\overset{\mathrm{d}\mu(\psi^{\pm\varphi})=\mathrm{d}\mu(\psi)}{=} \exp\left(\tfrac{1}{2}\beta K\right)\int_\psi \exp\left(-\tfrac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)F(\psi)\,\mathrm{d}\mu(\psi) = \exp(\tfrac{1}{2}\beta K)Z_{\beta,\Lambda_{L+B}}\,\mathbb{E}_{\beta,\Lambda_{L+B}}[F]\;.
\end{aligned}
$$

After performing the same analysis with a rotation in the other direction, we see:

$$
\exp\left(-\frac{1}{2}\beta K\right)\mathbb{E}_{\beta,\Lambda_{L+B}}\left[F\right] \leq \mathbb{E}_{\beta,\Lambda_{L+B}}\left[F\circ T_R\right] \leq \exp\left(\frac{1}{2}\beta K\right)\mathbb{E}_{\beta,\Lambda_{L+B}}\left[F\right]\;.
$$

Since F only depends on spins within Λ_L this now implies for *any* infinite volume Gibbs measure that

$$
\exp\left(-\frac{1}{2}\beta K\right)\mathbb{E}_{\beta}\left[F\right] \leq \mathbb{E}_{\beta}\left[F\circ T_R\right] \leq \exp\left(\frac{1}{2}\beta K\right)\mathbb{E}_{\beta}\left[F\right]\;. \tag{17}
$$

If $\mathbb{E}_\beta[F]=0$, then we immediately see by the above inequality that $\mathbb{E}_\beta[F\circ T_R]=0=\mathbb{E}_\beta[F]$. If $\mathbb{E}_\beta[F]\neq 0$, then we observe the following:

$$
\exp\left(-\frac{1}{2}\beta K\right)\leq\frac{\mathbb{E}_\beta[F\circ T_R]}{\mathbb{E}_\beta[F]}\leq\exp\left(\frac{1}{2}\beta K\right)\to 1\leq\frac{\mathbb{E}_\beta[F\circ T_R]}{\mathbb{E}_\beta[F]}\leq 1 \implies \mathbb{E}_\beta[F\circ T_R]=\mathbb{E}_\beta[F],
$$

where we used $K\to 0$ as $B\to\infty$. Thereby proving Theorem 9.1.

We note that assuming $F\geq 0$ did not lose generality because we can always write $F=F^+-F^-$ where $F^+(x):=\max\{F(x),0\}$ & $F^-(x):=\max\{-F(x),0\}\implies F^+,F^-\geq 0$. Meaning, that by the above argument:

$$
\mathbb{E}_\beta[F]=\mathbb{E}_\beta[F^+]-\mathbb{E}_\beta[F^-]=\mathbb{E}_\beta[F^+\circ T_R]-\mathbb{E}_\beta[F^-\circ T_R]=\mathbb{E}_\beta[F\circ T_R]
$$

9.3 Consequences

One immediate consequence of the above result is that for $N\geq 2,\ d=2$, there is no spontaneous magnetization. This is because by Theorem 9.1 with $F=\psi_x$ (a spin state at $x\in\Lambda$), we must have:

$$
\mathbb{E}_\beta[\psi_x]=\mathbb{E}_\beta\left[T_R\circ\psi_x\right]=\mathbb{E}_\beta\left[R\psi_x\right]=R\mathbb{E}_\beta\left[\psi_x\right]\quad\forall\ R\in O(N).
$$

This implies that $\mathbb{E}_\beta[\psi_x]=0$, which is precisely what it means for there to be no spontaneous magnetization (or in other words there is no spontaneous breaking of continuous symmetry).

We also know that by a standard argument (refer to section 3 of [2]), we can pin down extremal Gibbs measures by only probing events that have $\{0,1\}$ probability, so that the infinite volume Gibbs measures actually agree. Concretely, two measures that agree on these events with probability $\{0,1\}$ for any F measurable must agree on the entire σ-algebra. In our case, we have:

$$
\mathbb{E}_\beta[F]=\mathbb{E}_\beta[F\circ T_R]\implies\mathbb{E}_{\mathbb{P},\beta}[F]=\mathbb{E}_{\mathbb{P}_R,\beta}[F]\quad\forall\ F\text{ depending on finitely many spins and measurable}\implies\mathbb{P}=\mathbb{P}_R,
$$

where $\mathbb{P}_R$ and $\mathbb{P}$ denote the rotated and unrotated infinite-volume Gibbs measures respectively. Thus, after using this to show that the extremal Gibbs measures are equal, we may conclude that the unrotated and rotated infinite-volume Gibbs measures are equal, as any infinite-volume Gibbs measure is a mixture of the extremal ones.

Finally, Theorem 9.1 implies that there is no long range order (LRO) for $N \geq 2$ and $d = 2$. Recall that we say there is LRO if

$$\lim_{\|x-y\|\to\infty} \mathbb{E}_\beta \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N}\right] \neq 0\,.$$

To show that in our case this limit $= 0$, fix a lattice Λ_L with x at the origin and choose $\|y\|$ large enough so that $y \notin \Lambda_L$. Thus we see the following

$$\mathbb{E}_\beta \left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N}\right] = \mathbb{E}_\beta \left[\mathbb{E}_\beta\left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N} | \Lambda_L^c\right]\right] = \mathbb{E}_\beta \left[\langle \mathbb{E}_\beta\left[\psi_x | \Lambda_L^c\right], \psi_y \rangle_{\mathbb{R}^N}\right] = \mathbb{E}_\beta \left[\langle \mathbb{E}_{\beta,\Lambda_L}[\psi_x], \psi_y \rangle_{\mathbb{R}^N}\right]\,,$$

where in the first equality we used the law of iterated expectations, in the second we used the fact that $y \notin \Lambda_L$, and in the third we wrote $\mathbb{E}_\beta[\psi_x|\Lambda_L^c] \equiv \mathbb{E}_{\beta,\Lambda_L}[\psi_x]$ because $\mathbb{E}_{\beta,\Lambda_L}[\psi_x]$ denotes the average of ψ_x in Λ_L given fixed boundary conditions, which is precisely what $\mathbb{E}_\beta[\psi_x|\Lambda_L^c]$ denotes. Hence we see:

$$\mathbb{E}_{\beta,\Lambda_L}[\psi_x] \to \mathbb{E}_\beta[\psi_x] = 0 \text{ as } L \to \infty \implies \boxed{\lim_{\|x-y\|\to\infty} \mathbb{E}_\beta[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N}] = 0}\,,$$

meaning that there is no LRO for $N \geq 2$, $d = 2$.

10 Ken: Brydges–Fröhlich–Spencer graphical expansion of $O(N)$ model

In this section we explain how classical lattice spin systems can be rewritten as gases of random walks or "polymers" and how this representation gives us an upper bounds on critical temperatures. Our exposition follows the ideas of the paper by Brydges, Fröhlich, and Spencer [3] as well as the note by Ueltschi [16].

10.1 Random Walk Representations

I will first introduce a somewhat general and truly amazing result from linear algebra: we can express a matrix inverse as a sum over random walks, and a matrix determinant as a sum over closed random walks.

Definition 10.1 (Random walk). *A (finite) random walk on a lattice Λ is a finite sequence of ordered pairs of sites*

$$\omega = \{(x_1, x_2), (x_2, x_3), \dots, (x_{N-1}, x_N)\}, \qquad x_k \in \Lambda,$$

We write $|\omega| = N$ for the length. We write $\omega : x \to y$ if $x_1 = x$ and $x_N = y$. For given $k \in \Lambda$ we can ask how many times ω visits k; this number we denote as $n(k,\omega)$ and is just a number of elements in $\{x_1, x_2, ..., x_N\}$ that is equal to k.

Lemma 10.2 (Random walk representation of matrix inverse). *Let J be a symmetric matrix whose diagnoal entries are zero, and A be a diagonal matrix with non-zero diagonal entries. Entries of both matrices are specified by the site index $x \in \Lambda$. For any $x, y \in \Lambda$*

$$((A-J)^{-1})_{xy} = \sum_{\omega : x \to y} \left(\prod_{(x_i, x_j) \in \omega} J_{x_i x_j} \right) \left(\prod_{k \in \Lambda} A_k^{-n(k,\omega)} \right), \tag{18}$$

provided that the series converge. The first product runs over all random walks that start at x and end at y

Proof. Expand the inverse in a Neumann series,

$$(A-J)^{-1} = A^{-1} \sum_{m \geq 0} (JA^{-1})^m,$$

and identify the matrix element (x, y) of $(JA^{-1})^m$ with a sum over all m-step paths from x to y. For example, the term $m = 2$ gives the sum over random walks of length 2:

$$\begin{aligned}\left(\Lambda^{-1} J \Lambda^{-1} J \Lambda^{-1}\right)_{xy} &= \sum_{\substack{i_1, i_2, i_3 \in \Lambda \\ i_1 = x,\ i_3 = y}} A_{i_1}^{-1} J_{i_1 i_2} A_{i_2}^{-1} J_{i_2 i_3} A_{i_3}^{-1} \\ &= \sum_{\substack{\omega : x \to y \\ |\omega| = 2}} \left(\prod_{s \in \omega} J_s \right) \prod_{k \in L} A_k^{-n(k,\omega)} \end{aligned} \tag{19}$$

Summing over m gives the result. □

The same combinatorial idea leads to a representation of $\det(A - J)$ as an exponential of a sum over *random loops*, which are random walks that start and end at the same site, modulo cyclic permutations of the steps.

Lemma 10.3 (Random loop representation of matrix determinant).

$$\det(A-J)^{-1} = \left(\prod_x A_x \right)^{-1} \exp\left(\sum_{\omega^*} J_{\omega^*} \prod_x A_x^{-n(x,\omega^*)} \right),$$

provided the sum converges.

Proof. The idea is the same as Lemma 10.2:

$$\begin{aligned}
\det(A-J)^{-1} &= \det A^{-1}\ \det(1-A^{-1}J)^{-1} \\
&= \det A^{-1}\ \exp\big(-\operatorname{tr}\log(1-A^{-1}J)\big) \\
&= \det A^{-1}\ \exp\left(\sum_{k=1}^{\infty}\frac{1}{k}\ \operatorname{tr}\big((A^{-1}J)^k\big)\right) \\
&= \det A^{-1}\ \exp\left(\sum_{k=1}^{\infty}\frac{1}{k}\sum_{i\in\Lambda}\sum_{\substack{\omega:\, i\to i\\ |\omega|=k}}\Big(\prod_{s\in\omega}J_s\Big)\prod_{j\in\Lambda}A_j^{-n(j,\omega)}\right) \\
&= \Big(\prod_x A_x\Big)^{-1}\exp\left(\sum_{\omega^*}\Big(\prod_{s\in\omega^*}J_s\Big)\prod_{j\in\Lambda}A_j^{-n(j,\omega)}\right).
\end{aligned}$$

Define $J_{\omega^*} := \prod_{s\in\omega^*} J_s$ to get the result. □

10.2 Symanzik's Polymer Representation of Two-Point Correlation

At each site in lattice $x\in\Lambda$ we place a spin $\psi_x\in\mathbb{S}^{N-1}\subset\mathbb{R}^N$. We collect all spins into a configuration $\psi=\{\psi_x\}_{x\in\Lambda}\in\mathbb{S}^{(N-1)|\Lambda|}$. The ferromagnetic interaction is encoded in a symmetric, zero-diagonal coupling matrix $J=(J_{xy})_{x,y\in\Lambda}$, for example a nearest-neighbor coupling:

$$J_{xy}=J_{yx}=\beta\quad\text{for } x\sim y,\qquad J_{xx}=0, \tag{20}$$

and the Hamiltonian is

$$H\ =\ -\frac{1}{2}\sum_{x,y\in L}\sum_{\alpha=1}^{N}J_{xy}\psi_x^\alpha\psi_y^\alpha \tag{21}$$

The expectation value of an observable $F:\mathbb{R}^{N|\Lambda|}\to\mathbb{R}$ is

$$\mathbb{E}[F]\ :=\ \frac{1}{Z}\int_{\psi:\Lambda\to\mathbb{S}^{N-1}}F(\psi)\,e^{-H}\mathrm{d}\mu(\psi) \tag{22}$$

where Z is the partition function.

Theorem 10.4.

$$\mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}]=\frac{1}{Z}\sum_{\omega:x\to y}J_\omega\,Z(\omega),$$

where ω ranges over nearest–neighbour random walks from x to y, $J_\omega=\prod_{(u,v)\in\omega}J_{uv}$, and $Z(\omega)$ is the polymer partition function to be defined later. Note that this result is valid for any J as long as it is symmetric and zero in diagonal.

Proof. **Step 1: Replace the Sphere Constraint by a Fourier Representation**

Using the identity

$$\int_{S^{N-1}}d\mu(\psi_x)=\int_{\mathbb{R}^N}d\psi_x\,\delta\big(|\psi_x|^2-1\big),$$

we rewrite the partition function as

$$Z=\int\prod_{x\in\Lambda}d\psi_x\,\delta(|\psi_x|^2-1)\ \exp\left(\sum_{x,y}J_{xy}\,\psi_x\cdot\psi_y\right).$$

Introduce a Fourier representation of the constraint

$$\delta(|\psi_x|^2 - 1) = \int_{\mathbb{R}} da_x \, \widehat{g}(a_x) \, e^{-ia_x|\psi_x|^2},$$

where

$$\widehat{g}(a_x) = C_N (2ia_x)^{-N/2}$$

for a constant C_N depending only on N. Then

$$Z = \int \prod_{x\in\Lambda} da_x \, \widehat{g}(a_x) \int \prod_x d\psi_x \exp\left(-\tfrac{1}{2}\langle \psi, M\psi\rangle\right),$$

with

$$M = 2iA - J, \qquad A_{xx} = a_x, \qquad J = (J_{xy}).$$

Step 2: Gaussian Evaluation

The Gaussian integral gives

$$Z = \int \prod_{x\in\Lambda} da_x \, \widehat{g}(a_x) \, (2\pi)^{N|\Lambda|/2} \det(M)^{-N/2}.$$

Similarly, inserting sources and differentiating, one obtains

$$\mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}] = \frac{1}{Z} \int \prod_z da_z \, \widehat{g}(a_z) \, \det(M)^{-N/2} \, M_{xy}^{-1}. \tag{23}$$

Step 3: Random Walk Expansion for M^{-1}

Following Lemma 10.2, decompose $M = D - J$ with $D = 2iA$ diagonal. Then

$$M_{xy}^{-1} = \sum_{\omega: x\to y} J_\omega \prod_{k\in\Lambda} (2ia_k)^{-n(k,\omega)},$$

where $n(k,\omega)$ is the number of visits of ω to site k.

Step 4: Loop Expansion for $\det(M)^{-1}$

Lemma 10.3 yields

$$\det(M)^{-1} = \left(\prod_{k\in\Lambda} (2ia_k)^{-1}\right) \exp\left(\sum_{\omega^*} \frac{J^*_\omega}{|\omega^*|} \prod_{k\in\Lambda} (2ia_k)^{-n(k,\omega^*)}\right),$$

where the sum is over random loops ω^* and J_{ω^*} is the product of couplings along ω^*. Raising to the power $N/2$ and expanding leads to

$$\det(M)^{-N/2} = \sum_{n\geq 0} \frac{(N/2)^n}{n!} \sum_{\omega_1^*,\dots,\omega_n^*} \frac{J_{\omega_1^*}\cdots J_{\omega_n^*}}{|\omega_1^*|\cdots|\omega_n^*|} \prod_{k\in\Lambda} (2ia_k)^{-m_k - N/2},$$

where $m_k = n(k,\omega_1^*) + \cdots + n(k,\omega_n^*)$.

Step 5: Insert Random Walk expansions

Inserting this random–walk expansion, we obtain

$$\mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}] = \frac{1}{Z} \sum_{\omega:x\to y} J_\omega \sum_{n\geq 0} \frac{(N/2)^n}{n!} \sum_{\omega^1,\dots,\omega^n} \frac{J_{\omega^1}\cdots J_{\omega^n}}{|\omega^1|\cdots|\omega^n|}$$

$$\times \prod_{k\in\Lambda} \left(\int_{\mathbb{R}} da_k \, \widehat{g}(a_k) \, (2ia_k)^{-m_k - n(k,\omega) - N/2}\right).$$

Similarly, the partition function Z corresponds to the same expression but without the open walk ω (i.e., with $n(k,\omega)=0$). We write the product over $k \in \Lambda$ as

$$\prod_{k\in\Lambda}\left(\int_{\mathbb{R}} da_k\, \widehat{g}(a_k)\,(2ia_k)^{-m_k-n(k,\omega)-N/2}\right) =: \exp\bigl(-U(\omega,\omega^1,\dots,\omega^n)\bigr).$$

In the case without the open walk ω we obtain

$$\prod_{k\in\Lambda}\left(\int_{\mathbb{R}} da_k\, \widehat{g}(a_k)\,(2ia_k)^{-m_k-N/2}\right) =: \exp\bigl(-U(\omega^1,\dots,\omega^n)\bigr).$$

This somewhat brutal simplifiction in notation and artificial choice of letter U (as if it were internal energy!) will in a minute reveal to us the reason $Z(\omega)$ is called "ploymer gas partition function." We can therefore write the partition function as

$$Z = \sum_{n\geq 0}\frac{(N/2)^n}{n!}\sum_{\omega^1,\dots,\omega^n}\frac{J_{\omega^1}\cdots J_{\omega^n}}{|\omega^1|\cdots|\omega^n|}e^{-U(\omega^1,\dots,\omega^n)}.$$

For a fixed open random walk $\omega : x \to y$, we define the *polymer partition function in the presence of* ω by

$$Z(\omega) := \sum_{n\geq 0}\frac{(N/2)^n}{n!}\sum_{\omega^1,\dots,\omega^n}\frac{J_{\omega^1}\cdots J_{\omega^n}}{|\omega^1|\cdots|\omega^n|}e^{-U(\omega,\omega^1,\dots,\omega^n)}.$$

Voila! This is the grand partition function of polymer gas! Combining everything, we obtain

$$\mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}] = \frac{1}{Z}\sum_{\omega:x\to y} J_\omega\, Z(\omega),$$

□

10.3 Estimates on the Critical Temperature

We consider the hard–spin $O(N)$ model on $\Lambda \subset \mathbb{Z}^v$ with nearest–neighbour interactions:

$$\psi_x \in S^{N-1} \subset \mathbb{R}^N, \quad |\psi_x| = 1,$$

$$H_\Lambda(\psi) = -\beta\sum_{\langle x,y\rangle}\psi_x\cdot\psi_y, \qquad J_{xy} = \begin{cases}\beta, & x,y \text{ nearest neighbours,}\\ 0, & \text{otherwise.}\end{cases}$$

Theorem 10.5 (Critical temperature bound, nearest–neighbour case)**.** *Let β_c the smallest inverse temperature for which long–range order can occur. If*

$$\beta\left(\frac{2v-1}{N}+\frac{1}{N+2}\right) < 1, \tag{24}$$

then there exist constants $C, c > 0$, independent of Λ, such that

$$0 \leq \mathbb{E}[\psi_0^{(1)}\psi_x^{(1)}] \;\leq\; Ce^{-c|x|} \qquad \textit{for all } x \in \Lambda.$$

This in particular gives the lower bound for critical temperature

$$\beta_c \;\geq\; \frac{1}{\dfrac{2v-1}{N}+\dfrac{1}{N+2}}.$$

Proof. We start at the result 10.4 in the case of nearest neighbor coupling:

$$\mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}] = \sum_{\omega:x\to y}\beta^{|\omega|}\frac{Z(\omega)}{Z}$$

We wish to bound the ratio $Z(\omega)/Z$. We stare at the expression for $Z(\omega)$ above, realize that the only important factor that does not cancel is $e^{-U(\Omega)}$, whose site-wise component, in the case of nearest neighbor coupling, is known to be

$$F(m_y(\Omega)) = \frac{C_N}{N(N+2)\cdots(N+2m-2)}$$

where $\Omega = (\omega^1, \ldots, \omega^n)$ be a loop configuration and $m_y(\Omega)$ be the total number of visits of the loops in Ω to site y. If we add an open random walk ω to Ω, the number of visits at y increases by $n(y,\omega)$. Since the single–site factor in the loop gas is$C_N > 0$ is independent of m, we have

$$\frac{F(m_y(\Omega) + n(y,\omega))}{F(m_y(\Omega))} = \prod_{r=0}^{n(y,\omega)-1} \frac{1}{N + 2(m_y(\Gamma) + r)}.$$

But each term satisfies

$$\frac{1}{N + 2(m_y(\Omega) + r)} \leq \begin{cases} \dfrac{1}{N}, & \text{for the first visit of } \omega \text{ at } y, \\ \dfrac{1}{N+2}, & \text{for subsequent visits at } y, \end{cases}$$

So define $\chi(y,\omega) \in \{0,1\}$ to be 1 if ω ever visits y and 0 otherwise. We then have

$$\frac{F(m_y(\Omega) + n(y,\omega))}{F(m_y(\Omega))} \leq N^{-\chi(y,\omega)}(N+2)^{-n(y,\omega)+\chi(y,\omega)}.$$

Taking the product over y and noting that loop factors cancel in the ratio $Z(\omega)/Z$, we obtain

$$\frac{Z(\omega)}{Z} = \prod_{y \in \Lambda} \frac{F(m_y(\Omega) + n(y,\omega))}{F(m_y(\Omega))} \leq \prod_{y \in \Lambda} N^{-\chi(y,\omega)}(N+2)^{-n(y,\omega)+\chi(y,\omega)}$$

We can interpret this in terms of how much it "costs" to add one polymer to the gas. Namely, for a given random walk ω its step can contribute a factor of either $\leq 1/(N+2)$ (when it visits a site it has already been before) or $\leq 1/N$ (otherwise). Then we can write

$$\frac{Z(\omega)}{Z} \leq N^{-n_{\text{new}}(\omega)} (N+2)^{-n_{\text{old}}(\omega)}.$$

This leads to

$$0 \leq \mathbb{E}[\psi_x^{(1)}\psi_y^{(1)}] \leq \sum_{\omega: x \to y} \beta^{|\omega|} N^{-n_{\text{nb}}(\omega)} (N+2)^{-n_{\text{b}}(\omega)}. \tag{25}$$

Fix a length ℓ and consider all walks ω from x to y of length ℓ. At each step of a walk on $\mathbb{Z}^d$ there are exactly $2d$ possible nearest–neighbour moves, at most one of which "backtracks" the previous step and at most $2d-1$ of which does not. Thus, if we sum the weights step by step, the total contribution of all length–ℓ walks is bounded by

$$\left[(2d-1)\frac{\beta}{N} + \frac{\beta}{N+2}\right]^\ell.$$

Define

$$\alpha := \beta\left(\frac{2v-1}{N} + \frac{1}{N+2}\right).$$

Then

$$\sum_{\substack{\omega: x \to y \\ |\omega| = \ell}} \beta^{|\omega|} N^{-n_{\text{new}}(\omega)} (N+2)^{-n_{\text{old}}(\omega)} \leq \alpha^\ell,$$

and hence $\sum_{\ell \geq |x|} \alpha^\ell$ bounds the two-point correlation if $\alpha < 1$. This completes the proof. □

In their original paper Brydges, Fröhlich, and Spencer [3] provide a more general result that does not require J to be nearest-neighbor coupling.

11 Rees: The Fröhlich-Simon-Spencer proof of LRO for $N \geq 2$, $d \geq 3$, chessboard inequalities and reflection positivity

11.1 Setup

Let $\Lambda := [-L, L]^d \cap \mathbb{Z}^d$ and $\psi : \Lambda \to \mathbb{S}^{N-1}$ with $N \geq 2$, be a random field distributed according to the Gibbs measure whose partition function is

$$Z_{\beta,\Lambda} := \int_{\psi\in\Omega} \exp\left(-\frac{1}{2}\beta \langle \psi, -\Delta\psi\rangle\right) \mathrm{d}\mu(\psi)$$

where μ is the a-priori measure on $\Omega \equiv \left(\mathbb{S}^{N-1}\right)^\Lambda$, in this case the $|\Lambda|$-fold product of volume measures on $\mathbb{S}^{N-1}$ and

$$\begin{aligned}\langle \psi, -\Delta\psi\rangle &\equiv \langle \psi, -\Delta \otimes \mathbb{1}_N \psi\rangle_{\mathbb{R}^\Lambda\otimes\mathbb{R}^N} \\ &= \sum_{i=1}^{N}\sum_{x,y\in\Lambda} \psi_{x,i} (-\Delta)_{xy} \psi_{y,i} \,.\end{aligned}$$

We are interested in a lower bound on

$$\mathbb{E}_{\beta,\Lambda}\left[\langle \psi_u, \psi_v\rangle_{\mathbb{R}^N}\right]$$

for some fixed $u, v \in \Lambda$, for all sufficiently large β.

11.1.1 Discrete Fourier series

We take *periodic* boundary conditions here, so really, $(-\Delta)_{xy} = -1$ if $x \sim y$ where now $x \sim y$ if either $\|x - y\|_1 = 1$ or $x_{j_0} = -L$ and $y_{j_0} = -L$ (or reversed) for some $j_0 = 1, \cdots, d$ and $x_j = y_j$ for all $j \in \{1, \cdots, d\} \setminus \{j_0\}$.

The benefit of taking periodic boundary conditions is that now we may employ the discrete Fourier series. It should be considered a unitary matrix $\mathcal{F} : \mathbb{C}^{|\Lambda|} \to \mathbb{C}^{|\Lambda|}$ whose matrix elements are

$$(\mathcal{F})_{\xi,x} \equiv \frac{1}{(2L+1)^{\frac{d}{2}}} \exp\left(\mathrm{i}2\pi\frac{x\cdot\xi}{2L+1}\right) \qquad (x, \xi \in \Lambda)\,.$$

Note that sometimes it is customary to take a bijective indexing set

$$\Lambda^* \equiv \left\{ 2\pi\frac{\xi}{2L+1} \,\middle|\, \xi \in \Lambda \right\} \subseteq [-\pi, \pi]^d$$

(called the Brillouin zone) whence we have

$$(\mathcal{F})_{p,x} \equiv \frac{1}{(2L+1)^{\frac{d}{2}}} \exp(\mathrm{i}x\cdot p) \qquad (x \in \Lambda, p \in \Lambda^*)\,.$$

This does not change anything. We take the first convention.

Then note that

$$(\mathcal{F}(-\Delta)\mathcal{F}^*)_{\xi,\tilde{\xi}} = \delta_{\xi,\tilde{\xi}} 2\sum_{j=1}^{d}\left(1 - \cos\left(\frac{2\pi\xi_j}{2L+1}\right)\right) =: \delta_{\xi,\tilde{\xi}}\mathcal{E}(\xi)\,.$$

We now define the transformed spin-field in momentum space,

$$\hat{\psi} := \mathcal{F}\psi \in \left(\mathbb{C}^N\right)^\Lambda .$$

In particular we need to be aware that now $\hat{\psi}_{\xi,j} \in \mathbb{C}$! The fact that ψ is real-valued means that $\hat{\psi}_{-\xi,j} = \overline{\hat{\psi}_{\xi,j}}$.

We have, by unitarity of $\mathcal{F}$,

$$\langle \psi, -\Delta \otimes \mathbb{1}_N \psi\rangle_{\mathbb{R}^\Lambda\otimes\mathbb{R}^N} = \sum_{\xi\in\Lambda} \mathcal{E}(\xi) \left\|\hat{\psi}_\xi\right\|^2_{\mathbb{C}^N}.$$

Beware that even though $\langle \psi, -\Delta\psi\rangle$ becomes diagonal, the whole integral *does not factorize* since the volume measure on the sphere mixes all momentum components! I.e.,

$$\begin{aligned}1 \overset{!}{=} \|\psi_x\|^2_{\mathbb{R}^N} &= \sum_{j=1}^{N}\psi^2_{x,j} = \sum_{j=1}^{N}\left(\mathcal{F}^*\hat{\psi}\right)^2_{x,j} = \sum_{j=1}^{N}\sum_{\xi,\tilde{\xi}\in\Lambda} (\mathcal{F}^*)_{x,\xi} (\mathcal{F}^*)_{x,\tilde{\xi}} \hat{\psi}_{\xi,j}\hat{\psi}_{\tilde{\xi},j} \\ &= \sum_{j=1}^{N}\sum_{\xi,\tilde{\xi}\in\Lambda} |\Lambda|^{-1} \exp\left(-\frac{2\pi}{2L+1}\mathrm{i}\left(\xi+\tilde{\xi}\right)\cdot x\right) \hat{\psi}_{\xi,j}\hat{\psi}_{\tilde{\xi},j}\end{aligned}$$

so we see that *all modes* of $\hat{\psi}$ are involved in the unit-vector constraint.

We now turn to the two-point function

$$\tau : \Lambda^2 \to [0,1]$$

given by

$$\tau_{xy} \quad := \quad \mathbb{E}\left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N}\right] \,.$$

We think of τ as a matrix on $\mathbb{C}^{|\Lambda|}$ and as such, we may be interested in $\hat{\tau} := \mathscr{F}\tau\mathscr{F}^*$. First we note that thanks to our periodic boundary conditions, τ_{xy} is only a function of $x-y$ and not of x,y alone (i.e., it is a Toeplitz matrix). As such we know that $\mathscr{F}$ diagonalizes it so that

$$\hat{\tau}_{\xi,\tilde{\xi}} \quad \equiv \quad (\mathscr{F}\tau\mathscr{F}^*)_{\xi,\tilde{\xi}} =: \delta_{\xi,\tilde{\xi}}\sigma(\xi)$$

for some *symbol* $\sigma : \Lambda \to \mathbb{C}$. Let us calculate σ:

$$\begin{aligned}
\hat{\tau}_{\xi,\tilde{\xi}} \quad &\equiv \quad \sum_{x,y\in\Lambda} (\mathscr{F})_{\xi x}\, \tau_{xy}\, (\mathscr{F}^*)_{y\tilde{\xi}} \\
&= \quad \sum_{x,y\in\Lambda} |\Lambda|^{-1} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot x - \mathrm{i}\frac{2\pi}{2L+1}\tilde{\xi}\cdot y\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{x,j}\psi_{y,j}\right] \\
&= \quad \sum_{x,y\in\Lambda} |\Lambda|^{-1} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot x - \mathrm{i}\frac{2\pi}{2L+1}\tilde{\xi}\cdot y\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{x-y,j}\psi_{0,j}\right] \\
&\overset{z:=x-y}{=} \quad \sum_{z,y\in\Lambda} |\Lambda|^{-1} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot (z+y) - \mathrm{i}\frac{2\pi}{2L+1}\tilde{\xi}\cdot y\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{0,j}\right] \\
&= \quad \delta_{\xi,\tilde{\xi}} \sum_{z\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot z\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{0,j}\right] \,.
\end{aligned}$$

Hence we identify

$$\sigma(\xi) \quad := \quad \sum_{z\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot z\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{0,j}\right] \,.$$

Actually the symbol may be further simplified into $\sigma(\xi) = \mathbb{E}\left[\left\|\hat{\psi}_\xi\right\|^2_{\mathbb{C}^N}\right]$. Indeed,

$$\begin{aligned}
\sigma(\xi) \quad &= \quad \sum_{z\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot z\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{0,j}\right] \\
&= \quad |\Lambda|^{-1} \sum_{z,y\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot z\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{0,j}\right] \\
&= \quad |\Lambda|^{-1} \sum_{z,y\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot (z-y)\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z-y,j}\psi_{0,j}\right] \\
&= \quad |\Lambda|^{-1} \sum_{z,y\in\Lambda} \exp\left(\mathrm{i}\frac{2\pi}{2L+1}\xi\cdot (z-y)\right) \mathbb{E}\left[\sum_{j=1}^N \psi_{z,j}\psi_{y,j}\right] \\
&= \quad \mathbb{E}\left[\sum_{j=1}^N \hat{\psi}_{\xi,j}\overline{\hat{\psi}_{\xi,j}}\right] \\
&= \quad \mathbb{E}\left[\left\|\hat{\psi}_\xi\right\|^2_{\mathbb{C}^N}\right] \,.
\end{aligned}$$

11.2 The Gaussian domination bound

11.2.1 Reflection positivity

Let $R:\Lambda\to\Lambda$ be a reflection about a plane. For example, $x\mapsto -x$. This divides Λ into $\Lambda_\pm$ such that R yields a bijection $\Lambda_+\cong\Lambda_-$. Each observable $F:\Omega\to\mathbb{C}$ then has its reflected version $F_R := F\circ R$ given by

$$F_R(\psi) \quad\equiv\quad F(\psi\circ R)\,.$$

Definition 11.1 (Reflection positive measures)**.** *We say that $\langle\cdot\rangle:\Omega^{\mathbb{C}}\to\mathbb{C}$ is reflection positive iff for any $F,G:\Omega\to\mathbb{C}$ and any reflection $R:\Lambda\to\Lambda$ we have*

$$\left\langle\overline{F}G_R\right\rangle \quad=\quad \left\langle G\overline{F_R}\right\rangle$$

and

$$\left\langle\overline{F}F_R\right\rangle\geq 0\,.$$

Claim 11.2. *The classical O(N) Gibbs measure is reflection positive.*

Proof. Consider the Hamiltonian of the classical $O(N)$ model w.r.t. any reflection R. We have

$$\begin{aligned}\langle\psi,-\Delta\psi\rangle &= \sum_{x,y\in\Lambda_+:x\sim y}\|\psi_x-\psi_y\|^2_{\mathbb{R}^N}+\sum_{x,y\in\Lambda_-:x\sim y}\|\psi_x-\psi_y\|^2_{\mathbb{R}^N}+\sum_{x\in\Lambda_+,y\in\Lambda_-:x\sim y}\|\psi_x-\psi_y\|^2_{\mathbb{R}^N}\\ &=: H_+(\psi_+)+RH_+(\psi_-)+I(\psi_+)RI(\psi_-)\,.\end{aligned}$$

Then

$$\exp\left(-\frac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)=\exp\left(-\frac{1}{2}\beta H_+(\psi_+)\right)\left(R\exp\left(-\frac{1}{2}\beta H_+(\psi_-)\right)\right)\exp\left(-\frac{1}{2}\beta I(\psi_+)RI(\psi_-)\right)\,.$$

We further rewrite

$$\exp\left(-\frac{1}{2}\beta I(\psi_+)RI(\psi_-)\right)=\sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{1}{2}\beta\right)^n\left(I(\psi_+)RI(\psi_-)\right)^n$$

so that

$$\exp\left(-\frac{1}{2}\beta\langle\psi,-\Delta\psi\rangle\right)=A(\psi_+)RA(\psi_-)$$

for some function A. However, the a-priori measure μ certainly obeys reflection positivity so the full measure $\mathbb{E}_{\beta,\Lambda}$ also does. □

Corollary 11.3 (Cauchy-Schwarz-Reflection)**.** *For any $F,G:\Lambda\to\mathbb{C}$ and any reflection R we have*

$$\left|\left\langle\overline{F}G\right\rangle\right|^2 \quad\leq\quad \left\langle\overline{F}F_R\right\rangle\left\langle\overline{G}G_R\right\rangle\,.$$

As a result, for any $F_\pm:\Lambda_\pm\to\mathbb{C}$ we have

$$\left|\left\langle\overline{F_+}F_-\right\rangle\right|^2 \quad\leq\quad \left\langle\overline{F_+}(F_+)_R\right\rangle\left\langle\overline{F_-}(F_-)_R\right\rangle\,.$$

Corollary 11.4 (Chess-board inequality)**.** *Now suppose that we consider* all *possible reflections in* Λ*, and we apply the above reflection Cauchy-Schwarz successively. Say that we for any* $x,y\in\Lambda$*,* $x\sim y$*, we have* $G_{xy}:\left(\mathbb{S}^{N-1}\right)^{2}\to\mathbb{R}$ *symmetric. Then successively applying the above we learn*

$$\left\langle\prod_{x\sim y}\exp\left(G_{xy}\left(\psi_{x},\psi_{y}\right)\right)\right\rangle\leq\prod_{x\sim y}\left(\left\langle\prod_{x'\sim y'}\exp\left(G_{xy}\left(\psi_{x'},\psi_{y'}\right)\right)\right\rangle\right)^{\frac{1}{|E|}}\tag{26}$$

where $|E|$ *is the number of nearest-neighbor edges in* Λ*.*

11.2.2 The Gaussian domination bound

Theorem 11.5. *We claim that*

$$\sigma\left(\xi\right)\leq\frac{1}{\beta\mathcal{E}\left(\xi\right)}\qquad\left(\xi\in\Lambda\setminus\left\{\,0\,\right\}\right)\,.$$

Proof. We have

$$\sigma\left(\xi\right)\equiv\mathbb{E}\left[\left\|\hat{\psi}_{\xi}\right\|_{\mathbb{C}^{N}}^{2}\right]$$

so it is obvious that without the interaction-constraints

$$1=\sum_{j=1}^{N}\sum_{\xi,\tilde{\xi}\in\Lambda}\left|\Lambda\right|^{-1}\exp\left(-\frac{2\pi}{2L+1}\mathrm{i}\left(\xi+\tilde{\xi}\right)\cdot x\right)\hat{\psi}_{\xi,j}\hat{\psi}_{\tilde{\xi},j}$$

we would actually have an equality rather than inequality. Moreover, since $\left\|\hat{\psi}_{\xi}\right\|_{\mathbb{C}^{N}}^{2}$ is really a quadratic function of ψ it is enough to prove the inequality at the level of the moment generating function, i.e., to prove

$$\mathbb{E}_{\beta}\left[\exp\left(\left\langle u,\psi\right\rangle\right)\right]\quad\leq\quad\exp\left(\frac{1}{2\beta}\left\langle u,-\Delta^{-1}u\right\rangle\right)\qquad\left(u\in\mathbb{R}^{\Lambda}\setminus\ker\left(-\Delta\right)\right)$$

(since we can take Hessian w.r.t. u to get an upper bound on the two point function, but if we have an upper bound on any two point function, then we have the necessary bound). Hence our goal is to show

$$\mathbb{E}_{\beta}\left[\exp\left(\left\langle u,\psi\right\rangle\right)\exp\left(-\frac{1}{2\beta}\left\langle u,-\Delta^{-1}u\right\rangle\right)\right]\leq1\,.$$

By completing the square this is equivalent to

$$\int_{\psi\in\Omega}\exp\left(-\frac{1}{2}\beta\left\langle\left(\psi-\frac{1}{\beta}\Delta^{-1}u\right),-\Delta\left(\psi-\frac{1}{\beta}\Delta^{-1}u\right)\right\rangle\right)\mathrm{d}\mu\left(\psi\right)\leq\int_{\psi\in\Omega}\exp\left(-\frac{1}{2}\beta\left\langle\psi,-\Delta\psi\right\rangle\right)\mathrm{d}\mu\left(\psi\right).$$

Let us define

$$Z\left(v\right):=\int_{\psi\in\Omega}\exp\left(-\frac{1}{2}\beta\left\langle\left(\psi-v\right),-\Delta\left(\psi-v\right)\right\rangle\right)\mathrm{d}\mu\left(\psi\right)\qquad\left(v\in\mathbb{R}^{\Lambda}\right)\,.$$

Then it would suffice to show that

$$Z\left(v\right)\leq Z\left(0\right)\qquad\left(v\in\mathbb{R}^{\Lambda}\right)\,.$$

This however is a consequence of (26) since we would take

$$G_{xy}\left(\psi_{x},\psi_{y}\right)\quad:=\quad-\frac{1}{2}\beta\|\psi_{x}-v_{x}-\psi_{y}+v_{y}\|_{\mathbb{R}^{N}}^{2}$$

but if $v:\Lambda\to\mathbb{R}^{N}$ is constant it does not influence this function and we just get $Z\left(0\right)$. □

11.2.3 Proof of Long Range Order Given the Gaussian Domination Bound

By general arguments we know that

$$\tau : \Lambda^2 \quad \to \quad [0,1]$$

converges, as $L \to \infty$, to a function

$$\tau^\infty : \mathbb{Z}^d \times \mathbb{Z}^d \quad \to \quad [0,1]$$

given by

$$\tau^\infty_{xy} \quad := \quad \lim_{L\to\infty} \mathbb{E}_{\beta,\Lambda}\left[\langle \psi_x, \psi_y \rangle_{\mathbb{R}^N}\right] \, .$$

As we have seen, this function actually only depends on $\|x-y\|$ and not on x, y individually. Now, we *define* long range order as the non-decay of that function at infinity.

Theorem 11.6. *If σ obeys the Gaussian domination bound*

$$\sigma(\xi) \leq \frac{1}{\beta \mathcal{E}(\xi)} \qquad (\xi \in \Lambda \setminus \{ 0 \})$$

then τ^∞_{xy} has no off-diagonal decay, i.e.

$$\liminf_{\|x-y\|\to\infty} \tau^\infty_{x,y} > 0 \, .$$

Proof. Our general strategy will be to study σ^∞ instead of τ^∞. But what is σ^∞? First, we make a change of variables. In the infinite volume limit, we rather want a function (or measure, as we will see) on $\mathbb{T}^d \equiv [-\pi,\pi]^d$ as opposed to Λ or $\mathbb{Z}^d$. As such we naively would have define

$$\sigma^\infty(p) := \lim_{L\to\infty} \sigma\left(\frac{(2L+1)}{2\pi}p\right) \qquad \left(p \in \mathbb{T}^d\right) \, .$$

The problem with this definition is that σ^∞ may not define a pointwise function but rather a measure. Intuitively, what we mean is that σ^∞, on top of having an L^1 density on $\mathbb{T}^d$ with respect to the Lebesgue measure, also has a point mass at $p = 0$. Note that thanks to the Gaussian domination bound, this indeed may *only* happen at $p = 0$ and nowhere else. Anticipating this, we rather define σ^∞ as a *measure* on $\mathbb{T}^d$ via

$$\sigma^\infty(A) := \lim_{L\to\infty} \frac{1}{|\Lambda|} \sum_{\xi\in\Lambda \wedge \frac{2\pi\xi}{2L+1} \in A} \sigma(\xi) \qquad \left(A \subseteq \mathbb{T}^d \text{ msrbl.}\right) \, .$$

The right hand side is supposed to be meant as an approximation of the Riemann integral, which converges as $L \to \infty$. As we said, since this measure must be regular away from $p = 0$ (thanks to the Gaussian domination bound) we make the Ansatz

$$\sigma^\infty = m^2 \delta_0 + f \mathrm{d}\lambda$$

where $m \geq 0$, δ_0 is the Dirac delta measure at $p = 0$ and λ is the Lebesgue measure on $\mathbb{T}^d$, and finally $f \in L^1\left(\mathbb{T}^d\right)$. Then we have

$$\sigma^\infty\left(\mathbb{T}^d\right) \equiv \int_{\mathbb{T}^d} \mathrm{d}\sigma^\infty = m^2 + \int_{\mathbb{T}^d} f \mathrm{d}\lambda \, .$$

On the other hand, at finite volume, we have thanks to Parseval

$$\sum_{\xi\in\Lambda} \sigma(\xi) = \mathbb{E}\left[\sum_{\xi\in\Lambda} \left\|\hat{\psi}_\xi\right\|^2_{\mathbb{C}^N}\right] = \mathbb{E}\left[\sum_{x\in\Lambda} \|\psi_x\|^2_{\mathbb{C}^N}\right] = |\Lambda| \, .$$

Dividing by $|\Lambda|$ and taking the limit $L \to \infty$ we find

$$\sigma^\infty\left(\mathbb{T}^d\right) = m^2 + \int_{\mathbb{T}^d} f \mathrm{d}\lambda = 1$$

and moreover,

$$\int_{\mathbb{T}^d} f \mathrm{d}\lambda \equiv \lim_{L\to\infty} \frac{1}{|\Lambda|} \sum_{\xi\in\Lambda:\xi\neq 0} \sigma(\xi) \leq \frac{1}{2\beta} \int_{\mathbb{T}^d} \frac{1}{2\sum_{j=1}^d (1-\cos(p_j))} \mathrm{d}p\,.$$

Hence if $d \geq 3$ and β is sufficiently large, $m > 0$ so that indeed σ^∞ has a point mass at $p = 0$. □

Lemma 11.7 (The Riemann-Lebesgue lemma)**.** *If*

$$\sigma^\infty = m^2 \delta_0 + f \mathrm{d}\lambda$$

for some $m > 0$ and $f \in L^1(\mathbb{T}^d)$ then τ^∞ has no off-diagonal decay at infinity.

Proof. We have, for any $x \in \mathbb{Z}^d$,

$$\begin{aligned} \tau^\infty_{x,0} &\equiv \int_{p\in\mathbb{T}^d} \exp(\mathrm{i}p\cdot z)\,\mathrm{d}\sigma^\infty(p) \\ &= m^2 + \underbrace{\int_{p\in\mathbb{T}^d} \exp(\mathrm{i}p\cdot z) f(p)\,\mathrm{d}\lambda(p)}_{\to 0 \text{ as } \|z\|\to\infty}\,. \end{aligned}$$

Indeed, to see this last implication, let us approximate f by a trigonometric *polynomial* $f(p) = \sum_n c_n \mathrm{e}^{\mathrm{i}n\cdot p}$ and then pass to the limit thanks to density of polynomials in L^1. □

11.3 The Spin-Wave Perspective

When discussing the finite-volume Gibbs states, whose infinite-volume limit is expected to be translation invariant it is natural to start with finite systems in

$$\Lambda(L) := (-L/2, -L/2+1, \ldots, L/2]^d, \tag{27}$$

under periodic boundary conditions.

11.3.1 Fourier Representation

Fourier's representation of any function defined on Λ—in our case the spin configuration $\psi = \{\psi_x\}_{x\in\Lambda}$—is a decomposition into "spin waves" with amplitudes ψ_p:

$$\psi_x := \frac{1}{\sqrt{L^d}} \sum_{p\in\Lambda_L^*} \psi_p\, e^{-ip\cdot x}, \tag{28}$$

where the allowed momenta are those for which the wave is periodic on $\Lambda(L)$, i.e. the points of the Brillouin zone

$$\Lambda_L^* := (-\pi,\pi]^d \cap \frac{\pi}{L}\mathbb{Z}^d. \tag{29}$$

The inverse transform, by which $\{\psi_p\}$ is computed from the spin configuration, is given by

$$\psi_p := \frac{1}{\sqrt{L^d}} \sum_{x\in\Lambda(L)} e^{ip\cdot x}\,\psi_x. \tag{30}$$

Under the normalization chosen above, the Parseval–Plancherel identity takes the form of the sum rule

$$\sum_{p\in\Lambda_L^*} |\psi_p|^2 = \sum_{x\in\Lambda(L)} |\psi_x|^2 = |\Lambda(L)|, \tag{31}$$

with the last equality holding for unit-length spins, $|\psi_x| = 1$.

11.3.2 Hamiltonian Under Periodic Boundary Conditions

The spin-wave decomposition is particularly convenient when the Hamiltonian is taken with periodic boundary conditions. For spins of unit length, $\|\psi_x\| = 1$, and pair interactions, it is convenient to shift the Hamiltonian by a constant and symmetrize the couplings (neither affecting the Gibbs state). This allows one to present the Hamiltonian as

$$H_\Lambda^{\text{per}}(\psi) = \frac{1}{2} \sum_{\{x,y\}\subset\Lambda} J_{x-y} \, \|\psi_x - \psi_y\|^2 \; - \; h \sum_{x\in\Lambda} \psi_x \; = \; \frac{1}{2} \sum_{x,y\in\Lambda} J_{x-y} \, \psi_x \psi_y \; - \; h \sum_{x\in\Lambda} \psi_x, \tag{32}$$

with

$$J_u = J_{-u}, \qquad J_0 = -\sum_{\|n\|=1} J_n. \tag{33}$$

Formulated in this manner, the interaction energy is a non-negative quadratic form, invariant under periodic shifts, and it possesses a family of zero-energy modes corresponding to constant spin configurations.

Since plane waves form an orthonormal basis in $\ell^2(\Lambda)$ and are eigenfunctions of the shift operator (and hence also of the convolution operator with kernel J_{x-y}), the Hamiltonian diagonalizes in the Fourier basis. Using either the spectral representation of the quadratic form or the explicit Fourier transform ((30)), one finds

$$H_\Lambda^{\text{per}}(\psi) = \sum_{p\in\Lambda_L^*} E(p) \, |\psi_p|^2 \tag{34}$$

with

$$E(p) := \frac{1}{2} \sum_{u\in\Lambda(L)} e^{ip\cdot u} \, J_u. \tag{35}$$

11.3.3 Nearest-Neighbor Case

For the nearest-neighbor interaction,

$$J_u^{(\text{n.n.})} = \begin{cases} -1, & \|u\| = 1, \\ 2d, & u = 0, \\ 0, & \text{otherwise,} \end{cases} \tag{36}$$

one finds

$$E(p) = \sum_{j=1}^{d} \big(1 - \cos p_j\big) = 2 \sum_{j=1}^{d} \sin^2\Big(\frac{p_j}{2}\Big) \; \approx \; \frac{1}{2}\|p\|^2 \qquad (p \text{ small}). \tag{37}$$

Here one can recognize the discrete Laplacian (the second-difference operator) in the small-p behavior.

11.3.4 Fourier Representation of Correlations

The change of perspective—from the spin configuration $\{\psi_x\}$ to its plane-wave amplitudes $\{\psi_p\}$—is completed by noting that the Fourier transform of the spatial spin–spin correlation function of any shift-invariant state $\rho_\Lambda^{(L)} = \langle\cdot\rangle_\Lambda$ reappears as the intensity of the mode p:

$$S_\rho^{(L)}(p) := \sum_{x\in\Lambda(L)} e^{ip\cdot x} \, \mathbb{E}[\psi_0 \psi_x]_\Lambda \; = \; \mathbb{E}\big[|\psi_p|^2\big]_\Lambda. \tag{38}$$

Proof. Begin with the definition of variance in momentum space

$$\psi_p = \frac{1}{\sqrt{L^d}} \sum_{x\in\Lambda} e^{-ip\cdot x} \psi_x \tag{39}$$

$$|\psi_p|^2 = \psi_p \psi_p^* = \frac{1}{L^d} \sum_{x,y\in\Lambda} e^{-ip\cdot x} e^{ip\cdot y} \psi_x \psi_y^* \tag{40}$$

Taking the expectation gives

$$\mathbb{E}[|\psi_x|^2] = \frac{1}{L^d} \sum_{x\in\Lambda} \sum_{y-x\in\Lambda} e^{ip\cdot(y-x)} \mathbb{E}[\psi_0 \psi_{y-x}] \tag{41}$$

Let $z = y - x \in \Lambda$

$$\mathbb{E}[|\psi_p|^2] = \sum_{z\in\Lambda} e^{ip\cdot z}\mathbb{E}[\psi_0\psi_z] \tag{42}$$

□

This identity expresses the structure factor as the expected power in each spin-wave mode.

11.4 A Criterion for Continuous Symmetry Breaking in $d \geq 3$

A heuristic explanation of the possibility of continuous symmetry breaking in dimensions $d \geq 2$ is provided by the combination of:

- **The equipartition law**, which—while not a theorem—is presented in physics courses as a rule of thumb: each quadratic mode is expected to carry about $\frac{1}{2}kT = \frac{1}{2}\beta^{-1}$ of energy.
- **The sum rule** given by the Fourier representation of the correlation function (cf. equation (8.6) in the main text).

Their combination leads to a sufficient condition for symmetry breaking in which one may discern an analogy to the mechanism underlying the Bose–Einstein condensation phenomenon (macroscopic occupation of the ground state in a system of bosons).

Proposition 11.1 (Condition implying symmetry breaking)**.** *Let $d > 2$ and consider a system of bounded spins on $\mathbb{Z}^d$ with the nearest–neighbor interaction (8.7). Assume that the following* Gaussian domination bound *holds:*

$$S^{(L)}_{\rho,\beta}(p) \;\leq\; \frac{2}{\beta E(p)}, \tag{43}$$

and let

$$C_d \;:=\; \frac{1}{(2\pi)^d}\int_{[-\pi/2,\pi/2]^d} \frac{dp}{E(p)}. \approx \frac{1}{|\Lambda|}\sum_{p\neq 0}\frac{1}{E(p)} \qquad (L\to\infty) \tag{44}$$

Then for any $\beta > C_d/2$:

(1) The magnetization satisfies

$$\liminf_{L\to\infty} \mathbb{E}\Big[\Big|\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\psi_x\Big|^2\Big] \;\geq\; 1 - \frac{C_d}{2\beta} \;>\; 0. \tag{45}$$

Note here: the expectation of the total magnetization will be zero, but the variance of the magnetization will not be (as will be proven in a moment). Of course it is less than one, so we can write the inequality in the form of Eq. (45).

The dimension restriction $d > 2$ (for finite–range interactions) arises through the requirement that $1/|p|^2$ be locally integrable.

Proof. We begin with the identity valid for unit spins $\|\psi_x\| = 1$:

$$\frac{1}{|\Lambda|}|\hat{\psi}_0|^2 + \frac{1}{|\Lambda|}\sum_{p\in\Lambda^*\setminus\{0\}}|\hat{\psi}_p|^2 \;=\; 1 \tag{46}$$

which follows from the Parseval–Plancherel identity. The $p = 0$ term is singled out because:

(i) it contains the bulk magnetization,

$$\frac{1}{|\Lambda|}|\hat{\psi}_0|^2 = \Big|\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\psi_x\Big|^2 \tag{47}$$

(ii) the Gaussian domination bound gives no information about it since $E(0) = 0$.

Taking the expectation of Eq. (46) yields

$$\mathbb{E}\Big[\Big|\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\sigma_x\Big|^2\Big] \;=\; 1 \;-\; \frac{1}{|\Lambda|}\sum_{p\in\Lambda^*\setminus\{0\}}\mathbb{E}[\,|\hat{\psi}_p|^2\,] \tag{48}$$

Using the Gaussian domination bound Eq. (45),

$$\mathbb{E}\Big[\Big|\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\psi_x\Big|^2\Big] \;\geq\; 1-\frac{1}{|\Lambda|}\sum_{p\neq 0}\frac{2}{\beta}\,\frac{1}{E(p)}. \tag{49}$$

For large L, the sum is a Riemann approximation to the integral defining C_d:

$$\lim_{L\to\infty}\frac{1}{|\Lambda|}\sum_{p\neq 0}\frac{1}{E(p)} = C_d. \tag{50}$$

Care is required here because $1/E(p)$ is not uniformly continuous on the Brillouin zone. One splits the sum into $|p|\geq\varepsilon$ (handled by standard theorems) and $|p|<\varepsilon$, for which elementary bounds show that the contribution is $O(\varepsilon^{d-2})\to 0$ as $\varepsilon\to 0$. Altogether this yields Eq. (50), completing the proof of item (1). ☐

11.5 Reflection Positivity: Definition and Examples

The criterion for symmetry breaking stated in Proposition 8.1 is based on elementary but nontrivial considerations. However, verifying its key assumption (8.13) is rather challenging. At present, this verification has been achieved only for systems that possess *reflection positivity*. The purpose of this section is to introduce that concept and derive the chessboard inequality that follows from it.

Reflection positivity applies to systems defined either on $\mathbb{R}^d$ or on a graph that is symmetric with respect to reflection across a hyperplane (or a family of hyperplanes). Examples include the lattice $\mathbb{Z}^d$ and the finite periodic boxes

$$\Lambda = [-L/2, L/2)^d\cap\mathbb{Z}^d,$$

with periodic boundary conditions. Reflections may be taken across planes passing either through midpoints of edges or through vertices.

Let the reflection plane divide Λ, excluding the plane itself, into two regions, which we denote by $\Lambda^{\pm}$. Whenever vertices lie on the symmetry plane, we include those vertices in both $\Lambda^{\pm}$ for convenience.

Let

$$R:\Lambda\to\Lambda$$

be the geometric reflection. This is an involution ($R^2=1$). It induces an action on configurations $\sigma\in\Omega$ by

$$(R\sigma)_x := \sigma_{R(x)},$$

and hence on observables $F:\Omega\to\mathbb{C}$ by

$$(RF)(\sigma) := F(R\sigma).$$

Definition 11.2 (Reflection Positivity)**.** A state $\langle\cdot\rangle$ is called *reflection positive* (RP) with respect to the reflection R if for all observables $F,G:\Omega\to\mathbb{C}$ supported in Λ^+,

$$\langle\,\overline{F}\,RF\,\rangle\;\geq\;0,\qquad\langle\,\overline{F}\,RG\,\rangle\;=\;\langle\,GR\overline{F}\,\rangle. \tag{51}$$

A key example is the nearest–neighbor $O(N)$ model. On Λ_L with periodic boundary conditions, the corresponding Gibbs state is reflection positive with respect to reflections about any of the symmetry hyperplanes of the graph (whether passing through vertices or through edges). This follows from the following structural criterion for RP.

Proposition 11.3 (Sufficient Condition for Reflection Positivity)**.** *A Gibbs state on Λ_L is reflection positive with respect to a symmetry reflection R provided its Hamiltonian can be written in the form*

$$-H \;=\; A+RA\;+\;\sum_{j=1}^{k}B_j\,RB_j, \tag{52}$$

where A and each B_j depend only on the spins in Λ^+.

Proof. First note that, by independence of spins in the a priori measure, for any collection of observables $A_1, \ldots, A_m$ depending only on spins in Λ^+,

$$\Big\langle \prod_{j=1}^{m} A_j \, RA_j \Big\rangle_0 \;\geq\; 0, \tag{53}$$

where $\langle\cdot\rangle_0$ denotes the $\beta = 0$ Gibbs state (equivalently, the a priori product measure).

Next, from the assumed decomposition Prop. (Theorem 11.3) we have

$$e^{-\beta H} = e^{\beta A} \, R\big(e^{\beta A}\big) \prod_{j=1}^{k} e^{\beta B_j RB_j}. \tag{54}$$

Expanding each factor using

$$e^{\beta B_j RB_j} = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} (B_j RB_j)^n, \tag{55}$$

we see that the Gibbs weight is a linear combination (with positive coefficients) of terms of the form

$$\big(\prod A_j RA_j\big)\big(\prod B_j RB_j\big).$$

The positivity property Eq. (53) then implies the reflection positivity conditions ((51)). ☐

11.6 The Chessboard Inequality

Reflection positivity leads to powerful inequalities, the first of which is a direct consequence of the Cauchy–Schwarz inequality.

Let $\mathscr{B}$ denote the linear space of observables $F : \Omega \to \mathbb{C}$. Given two observables $F, G \in \mathscr{B}$, define

$$[F, G] := \mathbb{E}[\overline{F} \, RG] \tag{56}$$

which is a nonnegative sesquilinear form by reflection positivity. By Cauchy–Schwarz,

$$|[F, G]|^2 \leq [F, F] \, [G, G].$$

In our setting, this yields

$$|\mathbb{E}[FG]|^2 \;\leq\; \mathbb{E}[\overline{F} \, RF] \, \mathbb{E}[\overline{G} \, RG]. \tag{57}$$

Factorization across a reflection plane

Let R be a reflection decomposing the domain as $\Lambda = \Lambda^+ \cup \Lambda^-$. Let $\mathscr{B}_\alpha$ denote the class of observables depending only on the spins in Λ^α, for $\alpha \in \{+, -\}$. Applying (**??**) to $F_+ \in \mathscr{B}_+$ and $F_- \in \mathscr{B}_-$ gives

$$|\mathbb{E}[F_+ F_-]|^2 \;\leq\; \mathbb{E}[F_+ RF_+] \, \mathbb{E}[F_- RF_-]. \tag{58}$$

In integral notation the inequality becomes

$$\left| \int F_+(\psi^+) \, F_-(\psi^-) \, \rho(d\psi) \right|^2 \;\leq\; \left(\int F_+(\psi^+) \, F_+(\psi^-) \, \rho(d\psi) \right) \left(\int F_-(\psi^+) \, F_-(\psi^-) \, \rho(d\psi) \right) \tag{59}$$

Geometrically, the right-hand side involves "tiling" the domain by reflections of a single observable across the symmetry hyperplane.

Decomposition into boxes

Consider now a spin system on Λ_L whose state is invariant under periodic shifts and is reflection positive with respect to reflections across all hyperplanes perpendicular to coordinate axes. These reflections partition Λ into $K = 2^d$ boxes

$$\Lambda = \bigcup_{\alpha=1}^{K} \Lambda_\alpha,$$

with overlaps only along boundaries. Let $\mathscr{B}_\alpha$ denote the observables depending on spins in Λ_α.

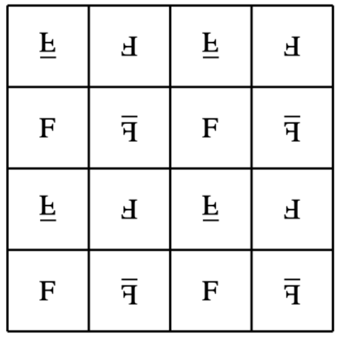


Figure 2: Decomposition into Boxes, including reflections and complex conjugations

Theorem 11.4 (Chessboard Inequality)**.** *Let* $\{F_\alpha\}_{\alpha=1}^K$ *be observables with* $F_\alpha \in \mathcal{B}_\alpha$*. Assume the Gibbs state is reflection positive with respect to each of the reflection planes defining the boxes. Then*

$$\left|\mathbb{E}\left[\prod_{\alpha=1}^{K} F_\alpha(\sigma|_{\Lambda_\alpha})\right]\right| \le \prod_{\alpha=1}^{K}\left(\mathbb{E}\left[\prod_{\beta=1}^{K} F_\alpha^{\#}(\sigma|_{\Lambda_\beta})\right]\right)^{1/K} \tag{60}$$

Here $F_\alpha^{\#}$ *denotes the function obtained by reflecting and complex conjugating* F_α *repeatedly in all directions with the reflection pattern of period* 2 *("the chessboard tiling"). See Figure 2*

Proof. Since (59) is homogeneous of degree 1 in each F_α, it suffices to assume the normalization

$$\mathbb{E}\left[\prod_{\beta=1}^{K} F_\alpha^{\#}(\sigma|_{\Lambda_\beta})\right] = 1, \tag{61}$$

for each α.

With this reduction, the inequality becomes equivalent to the following:

Let $S = \{F_j\}_{j=1}^K \subset \mathcal{B}_{\alpha_0}$ be a collection of functions measurable in a common box Λ_{α_0}, each normalized as in (**??**). For an assignment

$$\kappa : \{1,\dots,K\} \longrightarrow \{1,\dots,K\},$$

consider

$$M(\kappa) := \left|\mathbb{E}\left[\prod_{\alpha=1}^{K} F_{\kappa(\alpha)}^{\#}(\sigma|_{\Lambda_\alpha})\right]\right|. \tag{62}$$

We must show that the maximum of $M(\kappa)$ is attained for an assignment κ that is constant (i.e. $\kappa(\alpha)$ is the same for all α).

By the Cauchy–Schwarz inequality, if κ is a maximizer, then so is the assignment obtained by symmetrizing κ with respect to any reflection plane. Such symmetrizations monotonically reduce the number of nearest–neighbor disagreements without decreasing $M(\kappa)$. Thus maximizers exist with no disagreements between neighboring boxes. Any such configuration has $\kappa(\alpha)$ constant across all α, proving the claim. □

11.7 The Gaussian Domination Bound

Armed with the chessboard inequality, we proceed toward the proof of the Gaussian domination bound (8.13), originally due to [**?**]. Throughout, we consider the $O(N)$ spin model with interaction (8.7), and we assume that the corresponding Gibbs state is *reflection positive* (which is the case for the nearest–neighbor model discussed in Section 8.2).

A key step is to analyze a *shifted partition function.* For a field $\eta : \Lambda \to \mathbb{R}^N$, define

$$Z(\eta) := \int \exp\left\{-\frac{\beta}{2}\sum_{x,y} J_{x,y}\,\|(\psi_x+\eta_x)-(\psi_y+\eta_y)\|^2\right\}\prod_{u\in\Lambda}\rho_0(d\psi_u), \tag{63}$$

where ρ_0 is the *a priori* measure.

Lemma 11.5 (Monotonicity of the Shifted Partition Function)**.** *If the Gibbs state corresponding to the Hamiltonian (8.7) is reflection positive, then for every field* $\eta : \Lambda \to \mathbb{R}^N$*,*

$$Z(\eta) \;\le\; Z(0). \tag{64}$$

Proof. It is convenient to present the a priori measure ρ_0 (supported on the sphere) as the weak limit

$$\rho_0(d\psi) = \lim_{\varepsilon \downarrow 0} \rho_\varepsilon(d\psi), \tag{65}$$

where ρ_ε is a probability measure on all of $\mathbb{R}^N$ with strictly positive smooth density $g_\varepsilon(\sigma)$. For example,

$$g_\varepsilon(\sigma) = C_\varepsilon \exp\biggl(-\frac{\|\sigma\|^2 - 1}{\varepsilon}\biggr).$$

By continuity it suffices to prove that, for each $\varepsilon > 0$,

$$Z_\varepsilon(\eta) \le Z_\varepsilon(0). \tag{66}$$

For $\varepsilon > 0$ the shift $\sigma_x \mapsto \sigma_x + \eta_x$ can be absorbed by a change of variables, yielding

$$Z_\varepsilon(\eta) = Z_\varepsilon(0)\,\Bigl\langle \prod_{x\in\Lambda} T_{\eta_x}(\psi_x)\Bigr\rangle, \qquad T_{\eta_x}(\psi_x) := \frac{g_\varepsilon(\psi_x - \eta_x)}{g_\varepsilon(\psi_x)}. \tag{67}$$

For any constant shift $\eta_x \equiv y$, the factor in brackets equals 1, because the Hamiltonian is invariant under global spin shifts. Applying the chessboard inequality to the product $\prod_x T_{\eta_x}(\psi_x)$—with the decomposition $\Lambda = \bigcup_x \{x\}$ into singletons—implies that

$$\mathbb{E}\Bigl[\prod_x T_{\eta_x}(\psi_x)\Bigr] \;\le\; 1,$$

which gives ((66)). The claim for $Z(\eta)$ follows by the limiting procedure $\varepsilon \downarrow 0$. ☐

Theorem 11.6 (Gaussian Domination Bound)**.** *For the* $O(N)$ *spin model on* $\mathbb{Z}^d$ *with interaction, whenever the Gibbs state is reflection positive, the two–point function satisfies*

$$S^{(L)}_\rho(p) \;\le\; \frac{1}{2\beta\, E(p)}, \qquad p \neq 0, \tag{68}$$

where $E(p)$ *is the dispersion relation appearing in (8.7).*

Proof. By Lemma Theorem 11.5, for any field η and any $\epsilon > 0$,

$$\frac{Z(\epsilon\eta)}{Z(0)} = \mathbb{E}\Biggl[\exp\Biggl(\epsilon \sum_{x,y} \psi_x J_{x,y}\eta_y - \frac{\epsilon^2}{2\beta}\sum_{x,y} \eta_x J_{x,y}\eta_y\Biggr)\Biggr] \;\le\; 1. \tag{69}$$

Equivalently,

$$\mathbb{E}\Biggl[\exp\Biggl(\epsilon \sum_{x,y} \psi_x J_{x,y}\eta_y\Biggr)\Biggr] \;\le\; \exp\Biggl(\frac{\epsilon^2}{2\beta}\sum_{x,y} \eta_x J_{x,y}\eta_y\Biggr). \tag{70}$$

Expand both sides in a Taylor series in ϵ. The linear terms vanish (by symmetry), and comparing coefficients of ϵ^2 gives, for any real η,

$$\beta\mathbb{E}\Biggl[\Bigl|\sum_{x,y} \psi_x J_{x,y}\eta_y\Bigr|^2\Biggr] \;\le\; \sum_{x,y} \eta_x J_{x,y}\eta_y. \tag{71}$$

Use the Fourier decomposition

$$\sum_{x,y} \psi_x J_{x,y}\eta_y = \sum_{p\in\Lambda^*} \widehat{\psi}_p\, E(p)\, \widehat{\eta}_p,$$

and

$$\mathbb{E}[\widehat{\psi}_{p_1}\,\overline{\widehat{\psi}_{p_2}}] = \delta_{p_1,p_2}\, S^{(L)}_\rho(p).$$

Equation ((71)) becomes

$$\sum_p E(p)^2\, S_\rho^{(L)}(p)\, |\widehat{\eta}_p|^2 \;\le\; \frac{1}{2\beta} \sum_p E(p)\, |\widehat{\eta}_p|^2.$$

Since this holds for every test field η, we conclude the pointwise bound

$$S_\rho^{(L)}(p) \;\le\; \frac{1}{2\beta\, E(p)},$$

which is precisely ((68)). □

The implications of this bound for symmetry breaking were stated earlier in this section.

12 Kevin: The Lee-Yang theorem a la Newman

12.1 Definitions and Notation

Let Λ be a finite set of sites, and let $|\Lambda| = k$.

A signed measure μ on a measure space $(X, \mathscr{F})$ is a measure that can take positive or negative values. On $\mathbb{R}$, we say that μ is even if for all $A \in \mathscr{F}$:

$$\mu(A) = \mu(-A), \quad -A \coloneqq \{-x : x \in A\}.$$

Similarly, we say μ is odd if for all $A \in \mathscr{F}$,

$$\mu(-A) = -\mu(A).$$

Given a system of one-dimensional spin random variables $\{\psi_j : j \in [N]\}$ *Jacob*: Haven't defined what is N. with coupling matrix J and external field $\mathbf{h}$, we define the Hamiltonian

$$H = H(\boldsymbol{\psi}, \mathbf{h}) = -\sum_{k \le j = 1}^{N} J_{kj}\psi_k\psi_j - \sum_{j=1}^{N} h_j\psi_j.$$

For a vector $\mathbf{h}$, we use the notation

$$\mathrm{Re}(\mathbf{h}) = (\mathrm{Re}\, h_1, \mathrm{Re}\, h_2, \ldots, \mathrm{Re}\, h_N).$$

We define $\mathcal{V}$ to be a vectorization of the lower triangle of J:

$$\mathcal{V} = (J_{kj})_{k \le j = 1}^{N}.$$

Conversely, we could define J from $\mathcal{V}$. We also define the quadratic form

$$\mathcal{V}(\boldsymbol{\psi}, \boldsymbol{\psi}) = \sum_{k \le j = 1}^{N} J_{kj}\psi_k\psi_j.$$

12.2 Objective

Our goal is to prove the following theorem.

Theorem 12.1. *Suppose each μ_i, $i \in [N]$, is a signed even or odd measure on $\mathbb{R}$ satisfying $\mu_i(\psi) \ne \delta(\psi)$. Also suppose:*

(A)

$$\int \exp(b\psi^2)\, d|\mu_i(\psi)| < \infty \quad \textit{for all } b \ge 0,$$

(B) For $h \in \mathbb{C}$ with $\mathrm{Re}(h) > 0$:

$$\int e^{h\psi}\, d\mu_i(\psi) \ne 0.$$

and that J satisfies $J_{kj} \ge 0$. If $\beta > 0$ and $\mathrm{Re}(\mathbf{h}) > 0$, then the partition function

$$Z \coloneqq \int_{\mathbb{R}^N} e^{-\beta H}\, d\mu_1(\psi_1) \cdots d\mu_N(\psi_N)$$

and the correlation functions

$$\langle \psi_{i_1} \cdots \psi_{i_m} \rangle \coloneqq \frac{1}{Z} \int_{\mathbb{R}^N} \psi_{i_1} \cdots \psi_{i_m} e^{-\beta H}\, d\mu_1(\psi_1) \cdots d\mu_N(\psi_N)$$

do not vanish.

We define an **Ising system** to be a pair (N, ρ^N) where $N \in \mathbb{Z}_+$ and ρ^N is a positive measure on $\mathbb{R}^N$ of the form

$$\rho^N(\boldsymbol{\psi}) = e^{\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi})} \prod_{j=1}^{N} \frac{1}{2}\left(\delta(\psi_j - 1) + \delta(\psi_j + 1)\right).$$

Let $\mathcal{M}$ be a finite signed measure on $\mathbb{R}$. We say that $\mathcal{M}$ has the **Lee–Yang property** if for every choice of $N \in \mathbb{Z}_+$, nonzero measures $\mu_1, \ldots, \mu_N \in \mathcal{M}$, and $\mathcal{V} \geq 0$, the following hold:

(A') The signed measure

$$\mu^N(\boldsymbol{\psi}) := e^{\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi})} \prod_{j=1}^{N} \mu_j(\psi_j)$$

is finite.

(B') When $\mathrm{Re}(\mathbf{h}) > 0$:

$$Z(\mathbf{h}) := \int e^{\mathbf{h}\cdot\boldsymbol{\psi}}\, d\mu^N(\boldsymbol{\psi}) \neq 0.$$

A single measure μ is said to have the Lee–Yang property if $\{\mu\}$ does.

We define $\mathcal{N}$ to be the set of all even or odd signed measures on $\mathbb{R}$ satisfying conditions (A) and (B). To prove our goal theorem, it suffices to prove:

Theorem 12.2. *$\mathcal{N}$ has the Lee–Yang property.*

Theorem 12.3. *If each $Q_i(\psi)$ is an even or odd polynomial with only purely imaginary zeros, then when* $\mathrm{Re}(\mathbf{h}) > 0$*:*

$$\langle Q_1(\psi_1) \cdots Q_N(\psi_N)\rangle_{\mathbf{h}} \neq 0.$$

To prove that these theorems are sufficient, suppose $\mu_i, i \in [N]$ satisfy (A) and (B). By Theorem 12.2, $\mathcal{N}$ has the Lee-Yang property, so

$$\mu^N(\boldsymbol{\psi}) = e^{\beta\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi})} \prod_{j=1}^{N} \mu_j(\psi_j)$$

is finite by (A'), and by (B'):

$$\int e^{\beta h\cdot\boldsymbol{\psi}} d\mu^N(\boldsymbol{\psi}) = \int e^{\beta[\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi}) + h\cdot\boldsymbol{\psi}]} \prod_{j=1}^{N} d\mu_j(\psi_j) = \int e^{-\beta H} \prod_{j=1}^{N} d\mu_j(\psi_j) \neq 0.$$

If we take

$$Q_j(\psi_j) = \begin{cases} \psi_j, & j \in \{i_1, \cdots, i_m\} \\ 1, & \text{otherwise} \end{cases}.$$

Then each Q_j is an even or odd polynomial in ψ with only purely imaginary zeros (0 is purely imaginary), so by Theorem 12.3:

$$\langle Q_1(\psi_1) \cdots Q_N(\psi_N)\rangle_{\mathbf{h}} = \langle \psi_{i_1} \cdots \psi_{i_m}\rangle \neq 0.$$

12.3 $\mathcal{V}$ has the Lee-Yang Property

Let $\mathcal{V}$ be the set of signed measures μ on $\mathbb{R}$ such that, for some Ising system (N, ρ^N), $\boldsymbol{\lambda} \in \mathbb{R}_+^N$, $\mathbf{n} \in \{0,1\}^N$, and $K \in \mathbb{R}$:

$$E_\mu(h) = \int e^{h\psi}\, d\mu(\psi) = K \exp\left(-\frac{\pi i}{2} \sum_{k=1}^{N} n_k\right) \int \exp\left[h(\boldsymbol{\lambda}\cdot\boldsymbol{\psi}) + \frac{\pi i}{2}\mathbf{n}\cdot\boldsymbol{\psi}\right] d\rho^N(\boldsymbol{\psi}).$$

By the definition of ρ^N in an Ising system, ρ^N only puts mass on the set of points where

$$\boldsymbol{\psi} \in \{-1, 1\}^N$$

which is a finite set. Hence, the RHS above is just a finite sum over those points:

$$\sum_{\psi\in\{-1,1\}^N} A_\psi e^{h(\lambda\cdot\psi)}.$$

where A_ψ is the remaining factor. Since the equality holds for all h, by the linear independence of exponentials, it must be that μ only puts mass on the finite set of points $F := \{\lambda\cdot\psi : \psi\in\{-1,1\}^N\}$. Then

$$\int \exp(b\psi^2)d|\mu(\psi)| = \sum_{\psi\in F}|\mu(\psi)|e^{b\psi^2} < \infty.$$

So every $\mu\in\mathcal{V}$ satisfies (A). To show that every $\mu\in\mathcal{V}$ satisfies (B), we introduce the following:

Theorem 12.4 (Lee-Yang Theorem)**.** *If (N,ρ^N) is an Ising system, then when* $\mathrm{Re}(\mathbf{h})>0$*:*

$$\int e^{\mathbf{h}\cdot\boldsymbol{\psi}}d\rho^N(\boldsymbol{\psi}) \neq 0.$$

We use this theorem without proof. Defining $h_k(h) := h\lambda_k + \frac{\pi i}{2}n_k$, we can write

$$h(\boldsymbol{\lambda}\cdot\boldsymbol{\psi}) + \frac{\pi i}{2}\mathbf{n}\cdot\boldsymbol{\psi} = \sum_{k=1}^N\left(h\lambda_k + \frac{\pi i}{2}n_k\right)\psi_k = \sum_{k=1}^N h_k(h)\psi_k.$$

Note that if $\mathrm{Re}(h)>0$, since $\boldsymbol{\lambda}\in\mathbb{R}_+^N$, for each k:

$$\mathrm{Re}(h_k(h)) = h\lambda_k > 0.$$

so by Theorem 12.4:

$$\int \exp\left[h(\boldsymbol{\lambda}\cdot\boldsymbol{\psi}) + \frac{\pi i}{2}\mathbf{n}\cdot\boldsymbol{\psi}\right]d\rho^N(\boldsymbol{\psi}).$$

and therefore $E_\mu(h)\neq 0$. Hence, condition (2) is satisfied.

Lemma 12.5. *$\mathcal{V}$ has the Lee-Yang property.*

Proof. We've shown that every $\mu\in\mathcal{V}$ satisfies (A). Let $\mu_1,\dots,\mu_N\in\mathcal{V}$. Define $M := \max_{1\le i,j\le N}|J_{ij}|$. Then, for every $\boldsymbol{\psi}$:

$$\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi}) \le \sum_{i,j}|J_{ij}||\psi_i||\psi_j| \le M\left(\sum_{i=1}^N|\psi_i|\right)^2 \le MN\sum_{i=1}^N\psi_i^2.$$

Where the last inequality follows from Cauchy-Schwarz. Let $b := MN$. Then

$$\int e^{\mathcal{V}(\boldsymbol{\psi},\boldsymbol{\psi})}\prod_{j=1}^N d\mu_j(\psi_j) \le \int\prod_{j=1}^N e^{b\psi_i^2}\prod_{j=1}^N d\mu_j(\psi_j) = \prod_{j=1}^N\int e^{b\psi_j^2}d\mu_j(\psi_j).$$

where the equality follows from Fubini's theorem, since the integrand is non-negative. By condition (A), each factor on the RHS is finite, so (A') is satisfied.

By the definition of $\mathcal{V}$, for each μ_j there exists an Ising system $(M_j,\rho_j(\boldsymbol{\psi}_j))$ such that

$$E_j(h_j) = \int e^{h_j\psi}d\mu_j(\psi) = D_j\int\exp\{h_j(\boldsymbol{\lambda}_j\cdot\boldsymbol{\psi}_j) + \frac{i\pi}{2}\boldsymbol{n}_j\cdot\boldsymbol{\psi}_j\}d\rho_j(\boldsymbol{\psi}_j)$$

where D_j is the remaining factor. We define a new Ising system $(M, \tilde{\rho}^M)$ with $M = \sum_{i=1}^N M_i$ and

$$\tilde{\rho}^M(\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_N) = \exp\left\{\sum_{k \leq j=1}^N J_{kj}(\boldsymbol{\lambda}_j \cdot \boldsymbol{\psi}_j)(\boldsymbol{\lambda}_k \cdot \boldsymbol{\psi}_k)\right\}\prod_{j=1}^N \rho_j(\boldsymbol{\psi}_j).$$

Then

$$\begin{aligned} Z(\mathbf{h}) &= \int \exp\left[\mathbf{h} \cdot \boldsymbol{\psi} + \mathcal{V}(\boldsymbol{\psi}, \boldsymbol{\psi})\right] \prod_{k=1}^N \mu_k(\psi_k) \\ &= \int e^{\mathcal{V}(\boldsymbol{\psi}, \boldsymbol{\psi})} \prod_{k=1}^N \left[e^{h_k \psi_k} d\mu_k(\psi_k)\right] \\ &= \left(\prod_{k=1}^N D_k\right) \int_{\mathbb{R}^M} \exp\left\{\sum_{k=1}^N \left[h_k(\boldsymbol{\lambda}_k \cdot \boldsymbol{\psi}_k) + \frac{i\pi}{2}\mathbf{n}_k \cdot \boldsymbol{\psi}_k\right]\right\} d\tilde{\rho}^M \end{aligned}$$

Since $\boldsymbol{\lambda}_k > 0$ and $\mathbf{n}_k \in \{0,1\}^N$, if $\text{Re}(\mathbf{h}) > 0$, then for all k:

$$\text{Re}\left[h_k \boldsymbol{\lambda}_k + \frac{i\pi}{2}\mathbf{n}_k\right] > 0$$

so by Theorem 12.4 applied to $(M, \tilde{\rho}^M)$, $Z(\mathbf{h}) \neq 0$. Therefore, condition (B') is satisfied, which completes the proof. □

12.4 $\mathcal{N} = \overline{\mathcal{V}}$

We define $\overline{\mathcal{V}}$ as the set of signed measure on $\mathbb{R}$ that satisfy (A) and for which there exists a sequence $\mu_n \in \mathcal{V}$ such that

1. With C and c independent of n:

$$|E_{\mu_n}(h)| = \left|\int e^{h\psi} d\mu_n(\psi)\right| \leq C \exp\{c|h|^2\} \text{ for all } h \in \mathbb{C}$$

2. On compact subsets of $\mathbb{C}$:

$$E_{\mu_n}(h) \to E_\mu(h) \text{ uniformly}$$

Since measures in $\mathcal{V}$ are even or odd, and limits of analytic functions preserve parity, the same is true for $\overline{\mathcal{V}}$. Consider the following, which we use without proof.

Theorem 12.6 (Hurwitz's Theorem)**.** *Let f_n be a sequence of non-constant, analytic functions on a $D \subseteq \mathbb{C}$ converging uniformly on compact sets to an analytic function f. If each f_n has no zeros on D, then the limit f either has no zeros or is identically zero.*

By the condition of uniform convergence on compact sets in the definition of $\overline{\mathcal{V}}$, we can apply Theorem 12.6 to determine that measure in $\overline{\mathcal{V}}$ satisfy condition (2) of our goal theorem (E_μ cannot be identically 0 because of our definition of $\mathcal{V}$). Hence, $\overline{\mathcal{V}} \subseteq \mathcal{N}$. Next, we establish several key lemmas.

Lemma 12.7. *$E_\mu(h)$ can be represented as*

$$E_\mu(h) = Kh^m \prod_{j=1}^\infty \left(1 + \left(\frac{h}{\alpha_j}\right)^2\right)$$

for some $K \in \mathbb{R}$, $m \in \mathbb{Z}_{\geq 0}$, and $0 < \alpha_1 \leq \alpha_2 \leq \dots$ where $\sum_{j=1}^\infty \left(\frac{1}{\alpha_j}\right)^2 < \infty$.

Proof. Condition (A) restricts the growth of $E_\mu(h)$ at infinity. Since μ is even or odd, E_{μ_n} is also even or odd, and condition (B) states that $E_\mu(h)$ is non-zero for $\mathrm{Re}(h) > 0$. Hence, the zeros of E_μ occur in conjugate pairs, and they must be purely imaginary. Expressing the nonzero zeros as $h = \pm i\alpha_j$ with $0 < \alpha_1 \le \alpha_2 \le \dots$, we use the following:

Hadamard Factorization Theorem. If $f(h)$ is an entire function, we can write

$$f(h) = e^{P(h)} \prod_j E_p\left(\frac{h}{h_j}\right)$$

where $P(h)$ is a polynomial, h_j are the nonzero zeros of $f(h)$, and E_p are correction factors to ensure convergence.

Applying this to E_μ, the terms corresponding to conjugate pairs $\pm i\alpha_j$ can be combined, and the evenness/oddness of E_μ, as well as its controlled growth, eliminates the other factors and simplifies the factorization to

$$E_\mu(h) = Kh^m \prod_{j=1}^{\infty} \left(1 + \left(\frac{h}{\alpha_j}\right)^2\right)$$

where the controlled growth by (A) ensures that $\sum_{j=1}^{\infty} \left(\frac{1}{\alpha_j}\right)^2 < \infty$. $\square$

Lemma 12.8. *If $\mu_1, \mu_2 \in \mathcal{V}$, then $\mu_1 * \mu_2 \in \mathcal{V}$, where $\mu_1 * \mu_2$ denotes the convolution of μ_1 and μ_2:*

$$(\mu_1 * \mu_2)(A) = \int \mu_1(A - \psi) d\mu_2(\psi)$$

Proof. Convolutions have the property

$$E_{\mu_1 * \mu_2}(h) = E_{\mu_1}(h) \cdot E_{\mu_2}(h)$$

Recall by our definition of $\mathcal{V}$ that each μ_i is associated with an Ising system $(N_i, \rho_i^{N_i})$ by

$$E_{\mu_i}(h) = K \exp\left(-\frac{\pi i}{2} \sum_{k=1}^{N_i} n_k\right) \int \exp\left[h(\boldsymbol{\lambda}_i \cdot \boldsymbol{\psi}) + \frac{\pi i}{2} \mathbf{n}_i \cdot \boldsymbol{\psi}\right] d\rho_i^N(\boldsymbol{\psi}).$$

It is clear that by multiplying E_{μ_1} and E_{μ_2}, we would see that $\mu_1 * \mu_2$ is similarly associated with the Ising system $(N_1 + N_2, \rho)$, where ρ is the direct product (product measure) of ρ_1 and ρ_2 on $\mathbb{R}^{N_1+N_2}$. Hence, $\mu_1 * \mu_2 \in \mathcal{V}$. $\square$

Lemma 12.9. *Consider the α_j in the factorization above. For any $n \in \mathbb{Z}_+$ and any $a \ge \sqrt{2}\alpha_n$, there exists a measure $\mu_{n,a} \in \mathcal{V}$ such that*

$$E_{\mu_{n,a}}(h) = K \left[a \sinh\left(\frac{h}{a}\right)\right]^m \prod_{j=1}^{n} \frac{1}{2} \left(\frac{a}{\alpha_j}\right)^2 \left[\cosh\left(\frac{2h}{a}\right) - \left(1 - 2\left(\frac{\alpha_j}{a}\right)^2\right)\right]$$

Proof. We use the following measures in $\mathcal{V}$. For $N = 1$, $\lambda_1 = \frac{1}{a}$, $n_1 = 1$, there exists $\mu_1 \in \mathcal{V}$ such that:

$$E_{\mu_1}(h) = K \sinh\left(\frac{h}{a}\right)$$

For $N = 2$, $\lambda_1 = \lambda_2 = \frac{1}{a}$, and $n_1 = n_2 = 1$, there exists $\mu_2 \in \mathcal{V}$ such that:

$$E_{\mu_2}(h) = \frac{1}{2} K \exp(J_{12}) \left(\cosh\left(\frac{2h}{a}\right) - \exp(-2J_{12})\right)$$

Convoluting μ_1 and μ_2 gives us the desired result. □

Finally, we can prove the following:

Theorem 12.10. $\mathcal{N} = \overline{\mathcal{V}}$.

Proof. Let $\mu \in \mathcal{N}$ be arbitrary and $a_n = \sup\left(1, \sqrt{2}\alpha_n, 2n + m + 2\right)$, where m and the α_j are the same those in the expansion of E_μ from Lemma 12.7. We define $E_n(h)$ as the RHS of Lemma 12.9 with $a = a_n$, and we define

$$\tilde{E}_n(h) = Kh^m \prod_{j=1}^{n} \left(1 + \left(\frac{h}{\alpha_j}\right)^2\right)$$

This is just the truncation of the factorization of E_μ in Lemma 12.7. Our goal is to show that $E_n \to E_\mu$ in the sense of the definition of $\overline{\mathcal{V}}$. By using some basic inequalities:

$$\left|\prod_{j=1}^{n} a_j - \prod_{j=1}^{n} b_j\right| \le \left(\sum_{j=1}^{n} |a_j - b_j|\right) \prod_{j=1}^{n} \sup(1, |a_j|, |b_j|)$$
$$\sup(1, |a_j|, |b_j|) \le (1 + |b_j - 1| + |a_j - b_j|)$$

and common estimates for trigonometric functions with $a, \varepsilon \ge 1$, $w = \frac{h}{a_n}$, and $\varepsilon = \frac{1}{2}\left(\frac{a_n}{\alpha_j}\right)^2$ derived from series:

$$|\sinh w - w| \le |w|^3 e^{|w|}$$
$$|\cosh(2w) - 1 - 2w^2| \le 2|w|^4 e^{2|w|}$$
$$\sup\left(1, \left|a \sinh\left(\frac{h}{a}\right)\right|, |h|\right) \le e^{|h| + |h/a|}$$
$$1 + \varepsilon|2w^2| + \varepsilon|\cosh(2w) - 1 - 2w^2| \le \exp\{2\varepsilon|w|^2 + 2|w|\}$$

we can estimate the following bounds:

$$|E_n(h) - \tilde{E}_n(h)| \le \left|\frac{K}{a_n}\right|^2 (m|h|^3 + D|h|^4) \exp\left[(m+1)|h| + D|h|^2\right]$$
$$|\tilde{E}_n(h) - E_\mu(h)| \le \left[\sum_{j=n+1}^{\infty} \left(\frac{1}{\alpha_j}\right)^2\right] |K||h|^{m+2} \exp(D|h|^2)$$

where $D = \sum_{j=1}^{\infty} \left(\frac{1}{\alpha_j}\right)^2$. From Lemma 12.7 we also easily obtain

$$|E_\mu(h)| \le |K||h|^m \exp[D|h|^2]$$

Combining these three bounds, we get that for some fixed polynomial Q:

$$\begin{aligned} |E_n(h)| &\le |E_n(h) - \tilde{E}_n(h)| + |\tilde{E}_n(h) - E_\mu(h)| + |E_\mu(h)| \\ &\le Q(|h|) \exp[(m+1)|h| + D|h|^2] \end{aligned}$$

Importantly, Q, m, and D do not depend on n. Then, by taking $c > D$ and C sufficiently large, we obtain the bound

$$|E_n(h)| \le C \exp c|h|^2$$

for all $h \in \mathbb{C}$, which is condition (1) of $\overline{\mathcal{V}}$. Additionally:

$$\begin{aligned}|E_n(h) - E_\mu(h)| &\leq |E_n(h) - \tilde{E}_n(h)| + |\tilde{E}_n(h) - E_\mu(h)| \\ &\leq P_n(|h|)\exp\{(m+1)|h| + D|h|^2\}\end{aligned}$$

where $P_n(|h|)$ is the polynomial

$$P_n(|h|) = \frac{|K|}{|a_n|^2}(m|h|^3 + D|h|^4) + \left(\sum_{j=n+1}^{\infty} \frac{1}{|\alpha_j|^2}\right)|K||h|^{m+2}$$

Let $\mathcal{K} \subset \mathbb{C}$ be a compact set. Taking $R = \sup\{|h| : h \in \mathcal{K}\}$, we see that $P_n(|h|) \leq P_n(R)$, and $P_n(R)$ is a constant. Since $a_n \to \infty$, we know $\frac{1}{|a_n|^2} \to 0$, and since D converges, the tail must vanish in the limit, i.e. $\sum_{j=n+1}^{\infty} \frac{1}{|\alpha_j|^2} \to 0$. Therefore, $P_n(R) \to 0$. Similarly, we can bound

$$\exp\{(m+1)|h| + D|h|^2\} \leq \exp\{(m+1)R + DR^2\}$$

where the RHS is a constant. Then

$$|E_n(h) - E_\mu(h)| \leq P_n(R)\exp\{(m+1)R + DR^2\} \to 0$$

and since the RHS of the inequality is a constant, this is uniform convergence, which gives us condition (2) of $\overline{\mathcal{V}}$. Hence, $\mathcal{N} \subseteq \overline{\mathcal{V}}$. We showed earlier that $\overline{\mathcal{V}} \subseteq \mathcal{N}$, so this completes the proof. □

12.5 Conclusion

First, we establish the following.

Lemma 12.11. *If $\mu(\psi) \in \mathcal{N}$, then $\exp\{b\psi^2\}\mu(\psi), Q(\psi)\mu(\psi) \in \mathcal{N}$ for any $b \geq 0$ and any polynomial Q that is even or odd and has purely imaginary zeros.*

Proof. Since $\mu(\psi)$ is even or odd, $e^{b\psi^2}\mu(\psi), Q(\psi)\mu(\psi)$ are also even or odd. Additionally, satisfaction of (A) is preserved, as the Gaussian/polynomial factor can be bounded by $e^{b_0\psi^2}$ for some sufficiently large b_0, and

$$\int \exp(b\psi^2)d|e^{b_0\psi^2}\mu(\psi)| = \int \exp[(b+b_0)\psi^2]d|\mu(\psi)|$$

still satisfies (A). Next we let $Q(\psi) = \sum_k a_k\psi^k$ and use the fact that $\psi^k e^{h\psi} = \frac{d^k}{dh^k}e^{h\psi}$ to write

$$\begin{aligned}E_{Q\mu}(h) &:= \int Q(\psi)e^{h\psi}d\mu(\psi) = \sum_k a_k \int \psi^k e^{h\psi}d\mu(\psi) \\ &= \sum_k a_k \frac{d^k}{dh^k}\int e^{h\psi}d\mu(\psi) = Q\left(\frac{d}{dh}\right)E_\mu(h)\end{aligned}$$

Next we recall that

$$Q\left(\frac{d}{dh}\right)\tilde{E}_n(h) \to Q\left(\frac{d}{dh}\right)E_\mu(h)$$

uniformly on compact sets in $\mathbb{C}$. Since, using standard results on polynomials, $Q(\frac{d}{dh})\tilde{E}_n(h)$ has purely imaginary zeros, by Theorem 12.6, $E_{Q\mu}(h)$ has purely imaginary zeros, so $Q(\psi)\mu(\psi)$ satisfies (B) and hence $Q(\psi)\mu(\psi) \in \mathcal{N}$. By this result, we know that

$$\mu_n := \left(1 + \frac{b\psi^2}{n}\right)^n \mu \in \mathcal{N}$$

and we observe that, pointwise:

$$\left(1+\frac{b\psi^2}{n}\right)^n \to e^{b\psi^2}$$

Then, letting $\mu' = e^{b\psi^2}\mu$:

$$E_{\mu'}(h) = \int e^{h\psi}e^{b\psi^2}d\mu(\psi) = \int \lim_{n\to\infty}\left[e^{h\psi}\left(1+\frac{b\psi^2}{n}\right)^n\right]d\mu(\psi)$$
$$= \lim_{n\to\infty}\int e^{h\psi}\left(1+\frac{b\psi^2}{n}\right)^n d\mu(\psi) = \lim_{n\to\infty}E_{\mu_n}(h)$$

where the integral and limit can be interchanged due to condition (A) and the Dominated Convergence Theorem. On a compact set $K \subset \mathbb{C}$, this convergence is uniform, as, letting $R = \sup\{|h| : h \in K\}$:

$$\left|e^{h\psi}\left(1+\frac{b\psi^2}{n}\right)^n\right| \le \exp\{R|\psi| + b\psi^2\}$$

is a uniform bound for all $h \in K$. Hence, since $E_{\mu_n}(h)$ is non-zero on compact sets of the half-plane $\{\mathrm{Re}(h) > 0\}$, by Theorem 12.6, $E_{\mu'}(h) \neq 0$ whenever $\mathrm{Re}(h) > 0$. Thus, $e^{b\psi^2}\mu(\psi)$ satisfies (B), and therefore $e^{b\psi^2}\mu(\psi) \in \mathcal{N}$. □

We can finally revisit Theorem 12.2.

Proof of Theorem 12.2. We choose nonzero $\mu_k \in \overline{\mathcal{V}}$, $k = 1, \dots, N$ and define $\mu^N(\boldsymbol{\psi})$ and $Z^N(\mathbf{h})$ as in (A') and (B'). By the properties of $\overline{\mathcal{V}}$, (A') is already satisfied, so we seek to prove that (B') is satisfied. We may assume WLOG that along the diagonal of J, we have $J_{kk} = 0$. Otherwise, by Lemma 12.11, we could replace $\mu_k(\psi_k)$ with $\exp(J_{kk}\psi_k^2)\mu_k(\psi_k)$ to absorb the nonzero diagonal term. We choose $\mu_{j,i} \in \mathcal{V}$ according to the definition of $\overline{\mathcal{V}}$ so that, on compact sets in $\mathbb{C}$:

$$E_{j,i}(h) \coloneqq E_{\mu_{j,i}}(h) \to E_j(h) \coloneqq E_{\mu_j}(h)$$

uniformly as $i \to \infty$, and

$$|E_{j,i}(h)| \le Ce^{c|h|^2}$$

with C and c independent of i and j. For $\mathbf{i} = (i_1, \dots, i_N)$, we define $\mu_{\mathbf{i}}(\boldsymbol{\psi})$ and $Z_{\mathbf{i}}(\mathbf{h})$ to be the same as μ^N and Z^N in (A') and (B'), but with each μ_j replaced by $\mu_{j,i}$. Since each $\mu_{j,i} \in \mathcal{V}$ has the Lee-Yang property, by (B'), we know $Z_{\mathbf{i}}(\mathbf{h}) \neq 0$ when $\mathrm{Re}(\mathbf{h}) > 0$. We also define $\mu_{\mathbf{i},k}$ and $Z_{\mathbf{i},k}$ to be defined analogously to $\mu_{\mathbf{i}}$ and $Z_{\mathbf{i}}$, but with μ_j only replaced for $j \ge k$. For each j, we can treat $Z_{\mathbf{i}}$ as an entire function of h_j, so by Theorem 12.6, it suffices to show that $Z_{\mathbf{i}}(\mathbf{h}) \to Z^N(\mathbf{h})$ uniformly on compact subsets of $h_j \in \mathbb{C}$ as $\mathbf{i} \to \infty$ to conclude that $Z^N \neq 0$ when $\mathrm{Re}(\mathbf{h}) > 0$. In other words, we want to show that on compact subsets of $\mathbb{C}$:

$$\lim_{i_n\to\infty}\cdots\lim_{i_1\to\infty} Z_{\mathbf{i}}(\mathbf{h}) = Z^N(\mathbf{h})$$

uniformly. It suffices to show that for each $j, k \in [N]$:

$$Z_{\mathbf{i},k}(\mathbf{h}) \to Z_{\mathbf{i},k+1}(\mathbf{h})$$

uniformly on compacts of h_j as $i_k \to \infty$.

First, we define

$$\boldsymbol{\psi}_k = (\psi_1, \dots, \psi_{k-1}, \psi_{k+1}, \dots, \psi_N)$$
$$\mathbf{h}_k = (h_1, \dots, h_{k-1}, h_{k+1}, \dots, h_N)$$
$$\mathbf{J}_k = (J_{1k}, \dots, J_{(k-1)k}, J_{(k+1)k}, \dots, J_{Nk})$$
$$\tilde{\mu}_{\mathbf{i},k}(d\boldsymbol{\psi}_k) = \exp\left(\sum_{\substack{l<j\\ l,j\neq k}} J_{lj}\,\psi_l\psi_j\right)\prod_{j\neq k} d\mu_{j,i_j}(\psi_j).$$

Then, integrating with respect to ψ_k:

$$\begin{aligned}Z_{\mathbf{i},k}(\mathbf{h}) &= \int_{\mathbb{R}^N} \exp\{\mathbf{h}\cdot\boldsymbol{\psi} + \psi_k(\mathbf{J}_k\cdot\boldsymbol{\psi}_k)\} d\mu_{k,\mathbf{i}_k}(\psi_k)\, d\tilde{\mu}_{\mathbf{i},k}(\boldsymbol{\psi}_k) \\ &= \int_{\mathbb{R}^{N-1}} \exp\{\mathbf{h}_k\cdot\boldsymbol{\psi}_k\} E_{k,i_k}(h_k + \mathbf{J}_k\cdot\boldsymbol{\psi}_k)\, d\tilde{\mu}_{\mathbf{i},k}(\boldsymbol{\psi}_k).\end{aligned}$$

Since$|E_{j,i}(h)| \leq Ce^{c|h|^2}$ and E_{k,i_k} converges to E_k uniformly, we can take the limit of the RHS above as $i_k \to \infty$ for either $k = j$ or $k \neq j$ to get that, uniformly in h_j:

$$\lim_{i_k\to\infty} Z_{\mathbf{i},k}(\mathbf{h}) = \int_{\mathbb{R}^{N-1}} \exp\{\mathbf{h}_k\cdot\psi_k\} E_k(h_k + \mathbf{J}_k\cdot\psi_k)\, d\tilde{\mu}_{\mathbf{i},k}(\psi_k) = Z_{\mathbf{i},k+1}(\mathbf{h}).$$

To summarize, $j \in [1:N]$ indexes over the coordinates of the system, $i_j \in \mathbb{N}$ indexes over the approximation level of measure μ_j, and $k \in [1:N]$ indexes which coordinate's measure is being replaced. By taking limits to replace the measures one-at-a-time, we eventually recover Z^N.

Step k	**Limit**	$Z_{\mathbf{i},k}$	**Measures** $(\mu_1,\ldots,\mu_N)$
1	$i_1\to\infty$	$Z_{\mathbf{i},1} = Z_{\mathbf{i}}$	$(\mu_{1,i_1},\ \mu_{2,i_2},\ \mu_{3,i_3},\ \ldots,\ \mu_{N,i_N})$
2	$i_2\to\infty$	$Z_{\mathbf{i},2}$	$(\mu_1,\ \mu_{2,i_2},\ \mu_{3,i_3},\ \ldots,\ \mu_{N,i_N})$
3	$i_3\to\infty$	$Z_{\mathbf{i},3}$	$(\mu_1,\ \mu_2,\ \mu_{3,i_3},\ \ldots,\ \mu_{N,i_N})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$i_k\to\infty$	$Z_{\mathbf{i},k}$	$(\mu_1,\ \mu_2,\ \ldots,\ \mu_{k-1},\ \mu_{k,i_k},\ \mu_{k+1,i_{k+1}},\ \ldots,\ \mu_{N,i_N})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	$i_N\to\infty$	$Z_{\mathbf{i},N}$	$(\mu_1,\ \mu_2,\ \mu_3,\ \ldots,\ \mu_{N-1},\ \mu_{N,i_N})$
$N+1$	(end)	Z^N	$(\mu_1,\ \mu_2,\ \mu_3,\ \ldots,\ \mu_N)$

□

Lastly, we revisit Theorem 12.3.

Proof of Theorem 12.3. Because each Q_j is even/odd with purely imaginary zeros, by Lemma 12.11:

$$Q_j(\psi_j)\mu_j(\psi_j) \in \mathcal{n}$$

Then by Theorem 12.2, we can apply (B') from the Lee-Yang property to the correlation functions, which recovers Theorem 12.3. □

13 Neill: The Lee-Yang theorem via Lieb-Sokal

In this section, we take a different approach from the previous considerations, as we now examine the zeros of the partition function following the methods of [9]. To set the stage more concretely, we shall work on a finite lattice $\Lambda \subset \mathbb{Z}^d$ with $|\Lambda| = n$ (to each site i in the lattice Λ, we associate a spin ψ_i), and we consider the following Ising-like partition function:

$$Z(h_1, ..., h_n) = \int_{\psi:\Lambda\to\mathbb{R}} \exp\left[\beta\left(\sum_{i,j=1}^{n} J_{ij}\psi_i\psi_j + \sum_{i=1}^{n} h_i\psi_i\right)\right] d\mu(\psi)$$

Some remarks are in order for this expression for the partition function. This is a considerable generalization of the Ising model which we have considered earlier. First, we allow each spin in the lattice ψ_i to take values in $\mathbb{R}$, not just ± 1. Correspondingly, μ is now a measure on $\mathbb{R}^n$. We also allow the inter-spin interaction J_{ij} to couple more than just nearest neighbor spins, and the coupling strength can depend on the pair. Moreover, we have introduced a site dependent magnetic field h_i into the Hamiltonian, and we consider the partition function to be a function of this magnetic field. We can obtain the typical Ising model partition function by letting $\mu(\psi_i) = \delta(\psi_i - 1) + \delta(\psi_i + 1)$, setting $h_i = 0$, and choosing $J_{ij} = J$ if $i \sim j$ and zero otherwise. As one final remark, since we will not be concerned with the explicit temperature dependence of the system, henceforth, we will absorb β into our definition of J_{ij} and h_i. With this model in mind, the essence of the Lee-Yang theorem is as follows: if the measure μ possesses the Lee-Yang property (we will define this more explicitly below), then, for any collection of $J_{ij} \geq 0$ (i.e. ferromagnetic interactions), the partition function $Z(h_1, ..., h_n)$ does not vanish whenever $\mathrm{Re}(h_i) > 0$ for all sites i. As a further extension, in the case that μ is also an even measure (as is the case for the typical Ising model), it follows that $Z(h_1, ..., h_n)$ can only have zeros on the imaginary axis. Thus, the Lee-Yang theorem reveals a significant amount of structure to the zeros of the partition function for very general ferromagnetic systems. Since thermodynamic phase transitions often present themselves through singularities in thermodynamic potentials (which are obtained through logarithms of the partition function), this study of the structure of zeros of the partition function also sheds light on the conditions for phase transitions in this system.

To motivate our approach for proving the Lee-Yang theorem, we make the following informal observations:

$$Z(h_1, ..., h_n) = \exp\left(\sum_{i,j} J_{ij}\frac{\partial}{\partial h_i}\frac{\partial}{\partial h_j}\right)\int \exp\left(\sum_{i=1}^{n} h_i\psi_i\right)d\mu(\psi)$$

$$\exp\left(\sum_{i,j} J_{ij}\psi_i\psi_j\right) = \lim_{N\to\infty}\prod_{i,j=1}^{n}\left(1 + \frac{J_{ij}\psi_i\psi_j}{N}\right)^N$$

The idea of the above observations is to say that our desired partition function can be obtained by applying polynomial differential operators (more precisely, it is the limit of polynomial differential operators) to the non-interacting partition function $Z_0(h_1, ..., h_n) = \int \exp\left(\sum_{i=1}^{n} h_i\psi_i\right)d\mu(\psi)$. Since it is relatively straightforward to work with the non-interacting partition function, the problem essentially boils down to understanding the behavior of the zeros of a function when it is acted on by a polynomial differential operator.

13.1 Zeros of Polynomials

We state the following preliminary results about polynomials without proof.

Lemma 13.1. *Let P_0, P_1 be single variable complex polynomials. Suppose $P_0(z)+vP_1(z) \neq 0$ whenever $Re(z), Re(v) \geq 0$. Then $P_0(z) + P_1'(z) \neq 0$ whenever $Re(z) \geq 0$.*

Theorem 13.2 (Hurwitz's Theorem)**.** *Let $D \subset \mathbb{C}^n$ be open. Suppose $f_k : D \to \mathbb{C}$ is a sequence of non-vanishing holomorphic functions which converge to f uniformly on compact subsets of D. Then, either f does not vanish in D or f is identically zero.*

Theorem 13.3 (Grace's Theorem)**.** *Let $K \subset \mathbb{C}$ be a closed half-plane. Let $F(z) = \sum_{m=0}^{N} a_m z^m$ be a polynomial that does not vanish for $z \in K$. Let $z_1, ..., z_N$ be complex variables and define the functions $E_0, ..., E_N$ according to $E_0 = 1$ and:*

$$E_m(z_1, ..., z_N) = \sum_{i_1 < i_2 < ... < i_m} z_{i_1} z_{i_2} ... z_{i_m}$$

for $1 \leq m \leq N$. Then the polynomial:

$$\tilde{F}(z_1, ..., z_N) = \sum_{m=0}^{N} a_m \binom{N}{m}^{-1} E_m(z_1, ..., z_N)$$

does not vanish whenever $z_m \in K$ for all m.

The purpose of Grace's theorem for our considerations is that it allows us to make a multi-variable polynomial from a single variable one that is linear in each variable and has the same vanishing properties as the single variable polynomial. Moreover, we can recover the original polynomial by setting all the variables in the multi-variable polynomial to be equal to each other. Before proceeding, we introduce some convenient notation:

Definition 13.4. *Let $\mathbf{z} = (z_1, ..., z_n)$ denote an n-tuple of complex variables. For $\lambda \in \mathbb{R}$, we say $Re(\mathbf{z}) \geq \lambda$ iff $Re(z_k) \geq \lambda$ for all k. We also define $\frac{\partial}{\partial \mathbf{z}} \equiv (\frac{\partial}{\partial z_1}, ..., \frac{\partial}{\partial z_n})$.*

Definition 13.5. *For $\mathbf{z} \in \mathbb{C}^n$, let $P(\mathbf{z})$ denote a polynomial of n complex variables. We define $P(\frac{\partial}{\partial \mathbf{z}})$ as the differential operator obtained by replacing each instance of z_k in the definition of $P(\mathbf{z})$ with $\frac{\partial}{\partial z_k}$ for all k.*

Definition 13.6. *We define the set $H^n \subset \mathbb{C}^n$ by $H^n \equiv \{\mathbf{z} \in \mathbb{C}^n : Re(\mathbf{z}) \geq 0\}$.*

Using the above results, we prove the following property of polynomial differential operators.

Lemma 13.7. *Let $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$, and suppose $P(\mathbf{v}), Q(\mathbf{w})$ are polynomials. Define $R(\mathbf{v}, \mathbf{w}) = P(\mathbf{v})Q(\mathbf{w})$ and $S(\mathbf{z}) = P(\frac{\partial}{\partial \mathbf{z}})Q(\mathbf{z})$. Then:*

1. *If $R(\mathbf{v}, \mathbf{w}) \neq 0$ whenever $\mathbf{v}, \mathbf{w} \in \overline{H^n}$, then $S(\mathbf{z}) \neq 0$ whenever $\mathbf{z} \in \overline{H^n}$.*
2. *If $R(\mathbf{v}, \mathbf{w}) \neq 0$ whenever $\mathbf{v}, \mathbf{w} \in H^n$, then either $S(\mathbf{z}) \neq 0$ whenever $\mathbf{z} \in H^n$ or $S(\mathbf{z})$ is identically zero.*

Proof. 1. Let N denote the largest degree term of $R(\mathbf{v}, \mathbf{w})$ in the variable $\mathbf{v}$. Let us focus our attention to some v_k by fixing each v_j for $j \neq k$ in $\overline{H^1}$ and fixing $\mathbf{w}$ in $\overline{H^n}$. Then we can view R as a polynomial of only v_k (i.e. $R(v_k) = \sum_{m=0}^{N} b_m v_k^m$), and this polynomial does not vanish for $v_k \in \overline{H^1}$. We now introduce N complex variables $v_k^{(1)}, ..., v_k^{(N)}$. By Grace's Theorem, replacing $v_k^m \to \binom{N}{m}^{-1} E_m(v_k^{(1)}, ..., v_k^{(N)})$ in the expression for R ensures that the transformation of $R(v_k) \to \tilde{R}(v_k^{(1)}, ..., v_k^{(N)})$ is non-vanishing when $v_k^{(j)} \in \overline{H^1}$ for all j and $\mathbf{w} \in \overline{H^n}$. Repeat this replacement process for all k to obtain a polynomial $\tilde{\mathcal{R}}(\{v_k^{(j)}\}, \mathbf{w})$ that satisfies the following properties: it is non-vanishing when $v_k^{(j)} \in \overline{H^1}$ for all k, j and $\mathbf{w} \in \overline{H^n}$, it is linear in each variable $v_k^{(j)}$, and setting $v_k^{(j)} = v_k$ for all j and for each k recovers the original polynomial $R(\mathbf{v}, \mathbf{w})$ (see the remark after Grace's theorem for why this last property is true).

Now, for a given k, j, fix all variables of $\tilde{\mathcal{R}}$ in $\overline{H^1}$ so that we can effectively view $\tilde{\mathcal{R}}$ as a function of only $v_k^{(j)}$ and w_k. By the linearity of $\tilde{\mathcal{R}}$ in each variable $v_k^{(j)}$, we can say that $\tilde{\mathcal{R}}(v_k^{(j)}, w_k) = T_0(w_k) + v_k^{(j)} T_1(w_k)$ where T_0, T_1 are polynomials. We also know that $\tilde{\mathcal{R}}(v_k^{(j)}, w_k)$ does not vanish whenever $v_k^{(j)}, w_k \in \overline{H^1}$. According to Lemma 13.1, we can replace $v_k^{(j)} \to \frac{\partial}{\partial w_k}$ and the resulting polynomial of w_k does not vanish whenever $w_k \in \overline{H^1}$. Repeat this replacement process for all $v_k^{(j)}$ to obtain a polynomial of only $\mathbf{w}$ that does not vanish whenever $\mathbf{w} \in \overline{H^n}$. Notice that, in doing all of these replacements, we have effectively made the transformation $R(\mathbf{v}, \mathbf{w}) \to R(\frac{\partial}{\partial \mathbf{w}}, \mathbf{w}) = S(\mathbf{w})$, which concludes the proof.

2. Let $\epsilon > 0$, and define $R^{(\epsilon)}(\mathbf{v}, \mathbf{w}) = R(v_1 + \epsilon, ..., v_n + \epsilon, w_1 + \epsilon, ..., w_n + \epsilon)$. Define $S^{(\epsilon)}(\mathbf{z})$ similarly. It follows that $R^{(\epsilon)}(\mathbf{v}, \mathbf{w})$ does not vanish whenever $\mathbf{v}, \mathbf{w} \in \overline{H^n}$. By the above result, this implies that $S^{(\epsilon)}(\mathbf{z})$ does not vanish for $\mathbf{z} \in \overline{H^n}$. Since $S^{(\epsilon)}(\mathbf{z})$ converges to $S(\mathbf{z})$ uniformly on compact subsets as $\epsilon \to 0$, Hurwitz's Theorem allows us to conclude that either $S(\mathbf{z})$ does not vanish for $\mathbf{z} \in H^n$ or $S(\mathbf{z})$ is identically zero, as desired. □

13.2 Extension to Entire Functions

We now aim to extend the result of Lemma 13.7 to more general entire functions. In particular, for entire functions f, g, we wish to make sense of expressions of the form $f(\frac{\partial}{\partial \mathbf{z}})g(\mathbf{z})$ and then study their zeros. This motivates the following definitions.

Definition 13.8. *For a multi-index* $\mathbf{m} = (m_1, ..., m_n) \subset \mathbb{N}^n$ *and* $\mathbf{z} \in \mathbb{C}^n$*, we define* $\mathbf{z}^{\mathbf{m}} \equiv \prod_{i=1}^n z_i^{m_i}$ *and* $(\frac{\partial}{\partial \mathbf{z}})^{\mathbf{m}} \equiv (\frac{\partial}{\partial z_i})^{m_i} ... (\frac{\partial}{\partial z_n})^{m_n}$.

Definition 13.9. *For entire functions* $f, g : \mathbb{C}^n \to \mathbb{C}$*, consider their power series expansions* $f(\mathbf{z}) = \sum_{\mathbf{m}} \alpha_{\mathbf{m}} \mathbf{z}^{\mathbf{m}}$ *and* $g(\mathbf{z}) = \sum_{\mathbf{m}} \beta_{\mathbf{m}} \mathbf{z}^{\mathbf{m}}$*. Then, we interpret* $f(\frac{\partial}{\partial \mathbf{z}})g(\mathbf{z})$ *as the following formal power series which may or may not converge:*

$$f(\frac{\partial}{\partial \mathbf{z}})g(\mathbf{z}) = \sum_{\mathbf{k},\mathbf{m}} \alpha_{\mathbf{k}} \beta_{\mathbf{m}} (\frac{\partial}{\partial \mathbf{z}})^{\mathbf{k}} \mathbf{z}^{\mathbf{m}}$$

Our goal is to find a sufficiently nice class of entire functions for which the above power series is well defined. It turns out that the following space works.

Definition 13.10. *Given an entire function* $f : \mathbb{C}^n \to \mathbb{C}$ *and* $\lambda > 0$*, we define*

$$||f||_\lambda = \sup_{z \in \mathbb{C}^n} [\exp(-\lambda \sum_{i=1}^n |z_i|^2) |f(z)|]$$

Then, for $a \geq 0$*, we let* $\mathcal{A}_a^n$ *denote the space of entire functions* f *such that* $||f||_\lambda < \infty$ *for all* $\lambda > a$*. The topology of this space is induced by the countable family of norms* $|| \cdot ||_{a+1/k}$ *for* $k \in \mathbb{N}$*.*

The Cauchy integral formula implies that $\mathcal{A}_a^n$ is closed under differentiation. Moreover, one can show that a sequence $(f_k) \subset \mathcal{A}_a^n$ converges in the topology of $\mathcal{A}_a^n$ iff they converge pointwise.

Lemma 13.11. *Let* $a, b, c \geq 0$ *with* $ab < \frac{1}{4}$ *and* $c = \frac{b}{1-4ab}$*. Suppose* $f \in \mathcal{A}_a^n$ *and* $g \in \mathcal{A}_b^n$*. Then* $h(\mathbf{z}) \equiv f(\frac{\partial}{\partial \mathbf{z}})g(\mathbf{z})$ *is a well-defined entire function and* $h \in \mathcal{A}_c^n$*.*

Proof. We start by obtaining bounds on the power series coefficients for f and g (we follow the notation used in Definition 13.10). By the Cauchy Integral Formula, for disks $D_k \subset \mathbb{C}$ with $|\partial D_k| = 2\pi r_k$, we have that:

$$\alpha_{\mathbf{k}} = \frac{1}{(2\pi i)^n} \int_{\partial D_1} ... \int_{\partial D_n} \frac{f(\mathbf{z})}{z_1^{k_1+1} ... z_n^{k_n+1}} dz_1 ... dz_n$$

Choosing $a' > a$ and $b' > b$, we know that $||f||_{a'} < \infty$ and $||g||_{b'} < \infty$. This allows us to say that:

$$\begin{aligned} |\alpha_{\mathbf{k}}| &\leq \frac{1}{r_1^{k_1} ... r_n^{k_n}} \sup_{\mathbf{z} \in \partial D_1 \times ... \times \partial D_n} |f(\mathbf{z})| \\ &\leq \frac{1}{r_1^{k_1} ... r_n^{k_n}} ||f||_{a'} \prod_{i=1}^n e^{a' r_i^2} \\ &\leq ||f||_{a'} \prod_{i=1}^n (\frac{2ae}{k_i})^{k_i/2} \end{aligned}$$

where we set $r_i = \sqrt{\frac{k_i}{2a'}}$ to get the final inequality. Using the identity $(\frac{k}{2e})^{k/2} \geq C\Gamma((k+1)/2)$ for $C > 0$, where Γ is the gamma function, we can say that

$$|\alpha_{\mathbf{k}}| \leq C'||f||_{a'} \prod_{i=1}^{n} \frac{a^{k_i/2}}{\Omega(k_i)}$$

for some $C' > 0$, where we have defined $\Omega(0) = 1, \Omega(2k+1) = \Omega(2k+2) = k!$.

Based on the above reasoning, we have reduced this problem to the case in which the power series coefficients are given by $\alpha_{\mathbf{k}} = \prod_{i=1}^{n} \frac{a'^{k_i/2}}{\Omega(k_i)}$. One can check that these coefficients can be bounded (up to multiplicative constant) by the power series coefficients for the entire function given by $\prod_{i=1}^{n}(1 + (z_i + z_i^2)e^{az_i^2})$. Since the contributions of the different components of $\mathbf{z}$ have factored like this, we can set $n = 1$ (we can also drop the constant term 1). Thus, after putting everything together, we find that we only need to evaluate the case in which

$$f(z) = (z + z^2)e^{a'z^2}$$

$$g(z) = (z + z^2)e^{b'z^2}$$

and we can also take z to be real and positive. We now sketch the computation in this specific case. Using Gaussian integration, we can write:

$$e^{a'x^2} = \frac{1}{\sqrt{\pi a'}} \int_{-\infty}^{\infty} \exp\left(\frac{-t^2}{a'} + 2tx\right) dt$$

We also have the formal identity

$$\exp\left(2t\frac{\partial}{\partial x}\right)g(x) = g(x + 2t)$$

This identity comes from the fact that, if we expand the exponential as a power series, we obtain the formula for the Taylor expansion of g around x evaluated at $2t$. Therefore, we can say that:

$$\begin{aligned}
f\left(\frac{\partial}{\partial x}\right)g(x) &= \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right)\exp\left(a'\frac{\partial^2}{\partial x^2}\right)g(x) \\
&= \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right)\frac{1}{\sqrt{\pi a'}}\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{a'}\right)\exp\left(2t\frac{\partial}{\partial x}\right)g(x) \\
&= \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right)\frac{1}{\sqrt{\pi a'}}\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{a'}\right)[(x+2t) + (x+2t)^2]\exp(b'(x+2t)^2)dt \\
&= \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right)P^{(1)}_{a',b'}(x)\exp\left(\frac{b'}{1-4a'b'}x^2\right) \\
&= P^{(2)}_{a',b'}(x)\exp\left(\frac{b'}{1-4a'b'}x^2\right)
\end{aligned}$$

where $P^{(1)}_{a',b'}(x), P^{(2)}_{a',b'}(x)$ are a polynomials. From this, we can see that $f(\frac{\partial}{\partial z})g(z) \in \mathcal{A}^n_c$, which concludes the proof. □

While the above lemma shows that $\mathcal{A}^n_a$ is sufficiently nice class of functions so that we can define our desired entire differential operators, this space does not make any mention to the zeros of the entire functions. We remedy this with the following definition.

Definition 13.12. *We define $\mathcal{P}(H^n)$ as the space of polynomials on $\mathbb{C}^n$ that have no zeros in H^n. Furthermore, let $\overline{\mathcal{P}_a}(H^n)$ denote the closure of $\mathcal{P}(H^n)$ in the topology of $\mathcal{A}^n_a$.*

By Hurwitz's Theorem, we can conclude that any entire function $f \in \overline{\mathcal{P}_a}(H^n)$ either does not vanish in H^n or is identically zero. Thus, $\overline{\mathcal{P}_a}(H^n)$ defines the space of functions we wish to work with from now on. Importantly, ferromagnetic interactions are a part of this space.

Lemma 13.13. *For $\mathbf{z} \in \mathbb{C}^n$, let $f(\mathbf{z}) = \exp(\sum_{i,j=1}^{n} J_{ij}z_iz_j)$. Then, if $J_{ij} \geq 0$ for all i, j, it follows that $f \in \overline{\mathcal{P}}_{||J||}(H^n)$ where $||J||$ refers to the operator norm of J if we view the terms J_{ij} as defining the components of a matrix.*

Proof. Recall the identity we introduced earlier

$$f(\mathbf{z}) = \lim_{N\to\infty} \prod_{i,j} (1 + \frac{J_{ij} z_i z_j}{N})^N$$

Due to the non-negativity of J_{ij}, one can see that the polynomial $\prod_{i,j}(1 + \frac{J_{ij}z_iz_j}{N})^N$ for $N \in \mathbb{N}$ is non-vanishing when $\mathbf{z} \in H^n$. Moreover, we also have the bound

$$\begin{aligned} |\prod_{i,j}(1 + \frac{J_{i,j}z_iz_j}{N})^N| &\leq \exp(\sum_{i,j} |z_i| J_{ij} |z_j|) \\ &\leq \exp(||J|| \sum_i |z_i|^2) \end{aligned}$$

This implies that $f \in \overline{\mathscr{P}}_{||J||}(H^n)$. ☐

We now formulate Lemma 13.11 in terms of the space $\overline{\mathscr{P}_a}(H^n)$.

Lemma 13.14. *Let $a, b, c \geq 0$ with $ab < \frac{1}{4}$ and $c = \frac{b}{1-4ab}$. Suppose $f \in \overline{\mathscr{P}_a}(H^n)$ and $g \in \overline{\mathscr{P}_b}(H^n)$. Then $h(\mathbf{z}) \equiv f(\frac{\partial}{\partial \mathbf{z}})g(\mathbf{z})$ is a well-defined entire function and $h \in \overline{\mathscr{P}_c}(H^n)$.*

Proof. This follows from combining Lemma 13.7 and Lemma 13.11 along with the fact that there exist polynomials f_n, g_n which are non-zero in H^n and $f_n g_n \to fg$ in $\mathscr{A}_c^n$ ☐

The above lemma concludes our investigations into the zeros of entire functions. We now turn our attention to specifying the types of measures we wish to work with.

13.3 Specifying the Space of Measures

Definition 13.15. *Let $\mathfrak{M}^n$ denote the set of measures in the space of tempered distributions on $\mathbb{R}^n$ (i.e. the space of continuous linear functionals of Schwartz functions on $\mathbb{R}^n$). Given a Schwartz function f, we think of a measure in this space as yielding the functional given by $\mu(f) \equiv \int f(\mathbf{x}) d\mu(\mathbf{x})$. For $a > 0$, let $\mathscr{T}_a^n$ denote the space of $\mu \in \mathfrak{M}^n$ such that*

$$\mu = \exp(-a \sum_{i=1}^n x_i^2) \mu_a$$

for some $\mu_a \in \mathfrak{M}$.

The purpose for considering this space is that it behaves well under Laplace transformations.

Lemma 13.16. *Let $a > 0$. Given a measure $\mu \in \mathscr{T}_a^n$, the Laplace transform μ, defined as the function from $\mathbb{C}^n \to \mathbb{C}$ given by*

$$\hat{\mu}(\mathbf{z}) = \int e^{\mathbf{z}\cdot\mathbf{x}} d\mu(\mathbf{x})$$

is well defined and an element of $\mathscr{A}_{1/4a}^n$.

Proof. The key observation is that $\hat{\mu}(\mathbf{z}) = \mu_a(f_z)$ where μ_a is as defined in Definition 13.15 and $f_z(\mathbf{x}) = \exp(\mathbf{z}\cdot\mathbf{x} - a\sum_i |x_i|^2)$. Defining $|\mathbf{m}| = \sum_i m_i$ for any multi-index $\mathbf{m}$, it is a property of any tempered distribution T and any Schwartz function f that

$$|T(f)| \leq C \sup_{\mathbf{x}\in\mathbb{R}^n} (1 + |\mathbf{x}|^M) \sum_{|\mathbf{m}|\leq N} |(\frac{\partial}{\partial \mathbf{x}})^{\mathbf{m}} f(\mathbf{x})|$$

for some choice of C, M, N. Using the explicit form of f_z, this implies that

$$|\mu_a(f_z)| \leq C' \sup_{\mathbf{x}\in\mathbb{R}^n} (1 + |\mathbf{x}|^{M+N}) \exp(\frac{|\mathbf{z}|^2}{4a})$$

This proves that $\hat{\mu} \in \mathcal{A}^n_{1/4a}$, as desired. $\square$

The reason for looking at Laplace transforms is because $\hat{\mu}(\mathbf{h})$ is precisely the non-interacting partition function in a magnetic field $\mathbf{h}$. Thus, the above lemma shows that the space of measures $\mathcal{T}^n_a$ yields well-behaving non-interacting partition functions. However, as discussed above, we want these Laplace transforms to be in $\overline{\mathcal{P}_a}(H^n)$, not just $\mathcal{A}^n_a$. If this is indeed the case, we get a very nice property.

Lemma 13.17. *Let $0 \leq a \leq b$, let $\mu \in \mathcal{T}^n_b$ such that $\hat{\mu} \in \overline{\mathcal{P}}_{1/4b}(H^n)$, and let $f \in \overline{\mathcal{P}}_a(H^n)$. Then, the measure $f\mu \in \mathcal{T}^n_c$ for all $0 \leq c < b - a$ and its Laplace transform is in $\overline{\mathcal{P}}_{1/4c}(H^n)$.*

Proof. We first show that $f\mu \in \mathcal{T}^n_c$. Write $c = b - a - \epsilon$ for some $\epsilon > 0$, and note that $\mu = \exp(-b\sum_i x_i^2)\mu_b$ for some $\mu_b \in \mathcal{M}$. Observe that

$$\begin{aligned} \exp(c\sum_i x_i^2) f\mu &= \exp((c-b)\sum_i x_i^2) f\mu_b \\ &= \exp(-(a+\epsilon)\sum_i x_i^2) f\mu_b \end{aligned}$$

Since the growth of $\exp(-(a+\epsilon)\sum_i x_i^2) f(\mathbf{x})$ at infinity is bounded, one can see that $\exp(-(a+\epsilon)\sum_i x_i^2) f\mu_b \in \mathcal{M}$, which proves that $f\mu \in \mathcal{T}^n_c$.

To show that $\hat{f\mu} \in \overline{\mathcal{P}}_{1/4b}(H^n)$, first suppose that f is a polynomial. Much like with the Fourier transform, one can verify that $\hat{f\mu}(\mathbf{z}) = f(\frac{\partial}{\partial \mathbf{z}})\hat{\mu}(\mathbf{z})$. Using Lemma 13.14 proves the result in this case. In general, since $f \in \overline{\mathcal{P}}_a(H^n)$, we can consider a sequence of polynomials f_n converging to f in $\mathcal{A}^n_a$. The result then follows from the previous case by taking a limit (we omit some of the technicalities in comparing convergence in the space of distributions to convergence of the Laplace transforms in $\mathcal{A}^n_{1/4c}$; the interested reader can refer to the paper [9]). $\square$

13.4 The Lee-Yang Theorem

We now have all the tools we need to properly state and prove the Lee-Yang theorem. We first define what we meant by the Lee-Yang property which we mentioned in the beginning.

Definition 13.18. *A finite, positive measure μ on $\mathbb{R}^n$ (with $\mu \neq 0$) possess the Lee-Yang property with falloff γ iff $\mu \in \mathcal{T}^n_\gamma$ and $\hat{\mu} \in \overline{\mathcal{P}}_{1/4\gamma}(H^n)$.*

Theorem 13.19 (Lee-Yang Theorem)**.** *Let μ be measure possessing the Lee-Yang property with falloff γ. Let $f \in \overline{\mathcal{P}}_\sigma(H^n)$ with $\sigma < \gamma$ such that $f \geq 0$ on the support of μ and is strictly positive on a set of non-zero μ-measure. Then, for $\eta < \gamma - \sigma$, $f\mu$ has the Lee-Yang property with falloff η.*

Proof. The conditions on f are chosen precisely to ensure that $f\mu \neq 0$ is a positive measure. The proof of this theorem follows immediately from Lemma 13.17, just repackaged into more convenient language. $\square$

To more clearly elucidate why this result is exactly what we wanted, we present the following corollary.

Corollary 13.20. *Let μ be measure possessing the Lee-Yang property with falloff γ. For $J_{ij} \geq 0$ with $\sigma = ||J|| < \gamma$ (where $||J||$ is the operator norm of the matrix defined by the elements J_{ij}), the function defined by*

$$Z(h_1, ..., h_n) = \int \exp(\sum_{i,j=1}^{n} J_{ij}\psi_i\psi_j + \sum_{i=1}^{n} h_i\psi_i)d\mu(\psi)$$

does not vanish whenever $\mathbf{h} \in H^n$.

Proof. Combining Lemma 13.13 and Theorem 13.19 tells us that $\hat{f\mu} \in \overline{\mathscr{P}}_\eta(H^n)$ for $\eta < \gamma - \sigma$. In particular, since this Laplace transform is not identically zero, this implies that $\hat{f\mu}(\mathbf{h})$ does not vanish in H^n. Simply plugging into the definition of the Laplace transform yields

$$\hat{f\mu}(\mathbf{h}) = \int \exp(\sum_{i,j=1}^{n} J_{ij}\psi_i\psi_j + \sum_{i=1}^{n} h_i\psi_i)d\mu(\psi)$$

This concludes the proof. □

While this result is very general, it is not hard to verify that it also holds for the typical Ising model measure. In fact, that measure possess the Lee-Yang property for any falloff γ, so we are free to choose any non-negative ferromagnetic interaction terms J_{ij}.

Appendix

A Conditional Expectation and Probability

In this section, we define conditional expectation and probability, necessary in Section 3. For more details, look at the appendices of [7] Sections B.9.5, [5] Section 4.1, as well as [15] Section 7.8. Throughout the section, Ω is the space that we are working on, $\mathcal{F}$ is σ-algebra on Ω paired with the probability $\mathbb{P}$.

Consider the familiar definition from elementary probability:

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{72}$$

In school, we learned that this is the conditional probability of an event A happening with respect to an event B where $\mathbb{P}(B) > 0$.

Notice that (72) defines a new probablity meausure $\mathbb{P}(\cdot|B)$, allowing us to define *conditional expectation given* B of $X \in L^1(\Omega, \mathbb{P})$

$$E[X|B] := \sum_{\psi \in \Omega} X(\psi)\mathbb{P}(\psi|B). \tag{73}$$

Remark A.1. *We remark that one can generalize (72) by writing* $E[X|B] := \int_\Omega X(\psi)\mathbb{P}(d\psi|B)$.

Example A.2. *Suppose we are rolling two dice (Random Variables X_1, X_2) and want to know the expectation of the sum $S = X_1 + X_2$. Suppose we are given $E[X_1|S > 5]$ and $E[X_1|S \leq 5]$. Then, we can naturally consider the random variable*

$$\psi \mapsto E[X_1|S > 5]\mathbb{1}_{\{S>5\}}(\psi) + E[X_1|S \leq 5]\mathbb{1}_{\{S\leq 5\}}(\psi).$$

Like this example, if $(B_k) \subset \mathcal{F}$ be countable measurable partition of Ω and suppose $\mathcal{B} \subset \mathcal{F}$ is sub σ-algebra containing all (B_k). Then, the occurrence of some $B \in \mathcal{B}$ provides some information on the occurrence of some events B_k i.e. if for all $\mathbb{P}(B_k) > 0$ for all k, for all $X \in L^1(\mathbb{P})$, define the random variable

$$E[X|\mathcal{B}](\psi) := \sum_k E[X|B_k]\mathbb{1}_{B_k}(\psi). \tag{74}$$

Two defining characteristics of the above definition (74) is that for $X \in L^1(\mathbb{P})$:

(a) $\psi \mapsto E[X|\mathcal{B}](\psi)$ is $\mathcal{B}$-measurable function,

(b)$E[E[X|\mathcal{B}]\mathbb{1}_B] = E[X\mathbb{1}_B]$ for all $B \in \mathcal{B}$.

In particular, letting $B = \Omega$, we get that $E[E[X|\mathcal{B}]] = E[X]$. One can easily see that part (a) follows from the fact that $B_k \in \mathcal{B}$ and part (b) above can be proven by writing out the definition and using the countable additivity property of measures.

However, notice that we had a big assumption in our definition: $\mathbb{P}(B_k) > 0$ for all k. When defining infinite volume Gibbs measure, for example, we condition on finite alternating configurations, and fix an infinite number of them. This means all of our conditions (in the above case B_k's) will happen with probability 0; this requires us to change the definition above! But we are in luck! It turns out that (a) and (b) above are the key properties we need, and we can always find a function $Y \in L^1(\mathbb{P})$ satisfying exactly (a) and (b). To put it formally,

Theorem A.3 (Well-definedness of Conditional Expectation)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Consider $X \in L(\mathbb{P})$ and a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$. There exists a random variable $Y \in L^1(\mathbb{P})$, unique up to measure zero sets, for which the following conditions hold:*

- *Y is $\mathcal{G}$-measruable.*
- *For all $G \in \mathcal{G}$, $E[Y\mathbb{1}_G] = E[X\mathbb{1}_G]$.*

Please refer to [5] for the proof.

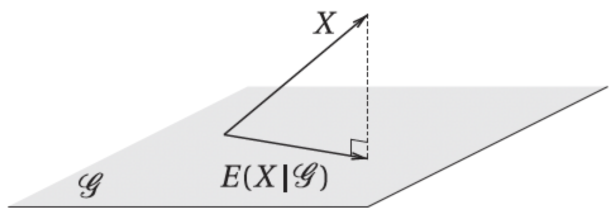
Definition A.4. *We use $E[X|\mathcal{G}]$ to denote above Y.*

Because we can approximate any measurable function by indicator functions, we have this theorem:

Corollary A.5. *Let $X \in L^1(\mathbb{P})$ and $\mathcal{G} \subset \mathcal{F}$ a sub-σ algebra. Then for all $\mathcal{G}$-measurable $Z \in L^1(\mathbb{P})$,*

$$E[XZ] = E[E[X|\mathcal{G}]Z].$$

Remark A.6. *Notice that this means $E[(X - E[X|\mathcal{G}])Z] = 0$ for all $\mathcal{G}$-measurable $Z \in L^1(\mathbb{P})$. So if we decide to just stay in real $L^2(\mathbb{P})$ instead, we can define inner product by $\langle X, Y\rangle := E[XY]$, then it means that $E[X|\mathcal{G}]$ defines the orthogonal projection of X onto the subspace $\{Z \in L^2(\mathbb{P}) : Z$ if $\mathcal{G}$-measurable$\}$. Below figure is from [7].*



Now, we can define conditional probability with respect to a sub-σ algebra.

Definition A.7. *Let $\mathcal{G} \subset \mathcal{F}$ sub-σ algebra. The conditional probability of $A \in \mathcal{F}$ with respect to $\mathcal{G}$ is given by:*

$$\mathbb{P}(A|\mathcal{G})(\psi) := E[\mathbb{1}_A|\mathcal{G}](\psi) \tag{75}$$

Then, by up to measure zero equivalence, $\mathbb{P}(A|\mathcal{G})(\psi)$ is the unique function that satisfies, for every $G \in \mathcal{G}$, the following identity:

$$\mathbb{P}(A \cap G) = \int_G \mathbb{P}(A|\mathcal{G})(\psi) d\mathbb{P}(\psi). \tag{76}$$

Note that every fixed $\psi \in \Omega$, except for a measure zero set of them, defines a probability measure by $\mathbb{P}(\cdot|\mathcal{G})$. This leads to a definition.

Definition A.8. *We say map $\hat{\mathbb{P}}(\cdot|\mathcal{G})(\cdot) : \mathcal{G} \times \Omega \to [0, 1]$ is called a regular conditional probability with respect to $\mathcal{G}$ if*

1. *For every fixed $\psi \in \Omega$, $\hat{\mathbb{P}}(\cdot|\mathcal{G})(\psi)$ is a probability measure on $(\Omega, \mathcal{F})$.*
2. *For every $A \in \mathcal{F}$, $\hat{\mathbb{P}}(A|\mathcal{G})(\cdot)$ is a version of $\mathbb{P}(A|\mathcal{G})$, i.e. $\mathbb{P}(A|\mathcal{G})(\cdot) = E[\mathbb{1}_A|\mathcal{G}](\cdot)$ $\mathbb{P}$- almost surely.*

B Probability Kernel and Specification

The goal of this section is to introduce the readers to probability kernel and specification as well as to relate them to conditional probability. The theory developed in this section is heavily used in Section 3.

In this seciton, we denote $\Omega = \prod_{i\in\mathbb{N}} K$ where $K \subset \mathbb{R}^n$ is a compact set. Let us endow Ω with a product topology. Notice that by Tychonoff's Theorem, Ω is a compact space. Let us call $\mathcal{F}$ the Borel σ-algbera on Ω. By definition of product topology, $\mathcal{F}$ is σ-algbera generated by sets inverse images of finite-subset projections, i.e.

$$\mathcal{F} = \sigma\left(\rho_\Lambda^{-1}(A) : A \in \mathcal{P}\left(\prod_{i\in\Lambda\Subset\mathbb{N}} K\right)\right)$$

where $\rho_\Lambda : \Omega \to \prod_{i\in\Lambda\Subset\mathbb{N}} K$ is defined by projection onto Λ coordinates.

In addition, for every $S \subset \mathbb{N}$ (here S is not necessarily fintie), we denote $\mathcal{F}_S$ as the Borel σ-algebra on $\Omega_S \equiv \prod_{i\in S} K$. One can observe easily that for every $S \Subset \mathbb{N}$, $\mathcal{F}_S$ is a sub-σ algebra of $\mathcal{F}$ by the definition of product topology.

Now we introduce the main definition of this section.

Definition B.1 (Probability Kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$.). *Let $\Lambda \Subset \mathbb{N}$. A probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$ is a map $\pi_\Lambda : \mathcal{F} \times \Omega \to [0,1]$ such that*

- *For every $\psi \in \Omega$, $\pi_\Lambda(\cdot|\psi)$ is a probability measure on $(\Omega, \mathcal{F})$.*
- *For every $A \in \mathcal{F}$, $\pi_\Lambda(A|\cdot)$ is $\mathcal{F}_{\Lambda^c}$-measurable function.*

One may find other texts refer to probability kernel as Markov kernel, transition kernel, or stochastic kernel.

Definition B.2 (Proper Probability Kernel). *A probability kernel $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$, π_Λ, is said to be proper if for every $B \in \mathcal{F}_{\Lambda^c}$, and for every $\psi \in \Omega$,*

$$\pi_\Lambda(B|\psi) = \mathbb{1}_B(\psi). \tag{77}$$

For every measurable $f : \Omega \to \mathbb{R}$, we can define another $\mathcal{F}_{\Lambda^c}$-function via:

$$\pi_\Lambda f(\psi) \equiv \int_\Omega f(\eta)\pi_\Lambda(d\eta|\psi). \tag{78}$$

Next we introduce a key fact about proper probability kernel.

Theorem B.3. *If π_Λ is a proper probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$, for every $A \in \mathcal{F}$,*

$$\pi_\Lambda(A|\psi) = \int_{\eta_\Lambda \in \Omega_\Lambda} \mathbb{1}_A(\eta_\Lambda \psi_{\Lambda^c})\pi_\Lambda(d(\eta_\Lambda \psi_{\Lambda^c})|\psi). \tag{79}$$

Consequently, if $f : \Omega \to \mathbb{R}$ is a bounded $\mathcal{F}$-measurable function,

$$\pi_\Lambda f(\psi) = \int_{\eta_\Lambda \in \Omega_\Lambda} f(\eta_\Lambda \psi_{\Lambda^c})\pi_\Lambda(d(\eta_\Lambda \psi_{\Lambda^c})|\psi). \tag{80}$$

The proof follows from observing that being a proper kernel implies that π_Λ is determined by $\{\pi_\Lambda(\{\eta_\Lambda \psi_{\Lambda^c}\}|\psi)\}_{\eta \in \Omega_\Lambda}$.

Next, we define what a specification is.

Definition B.4 (Specification). *We say family $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{N}}$ is a specification if*

- *For every $\Lambda \Subset \mathbb{N}$, π_Λ is proper probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$.*
- *For every $\Delta \subset \Lambda \Subset \mathbb{N}$, we have*

$$\pi_\Lambda \pi_\Delta = \pi_\Lambda$$

where we define composition of probability kernels via

$$\pi_\Lambda \pi_\Delta(A|\psi) \equiv \int_\Omega \pi_\Delta(A|\eta)\pi_\Lambda(d\eta|\psi). \tag{81}$$

One can manually for every proper probability kernel π_Δ and π_Λ, $\pi_\Lambda \pi_\Delta$ as defined above in (81) is a proper probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$. We can moreover compose a measure on $(\Omega, \mathcal{F})$ with a proper probability kernel.

Definition B.5. *For every measure μ on $(\Omega, \mathcal{F})$ and probability kernel π_Λ, define $\mu\pi_\Lambda$ via for $A \in \mathcal{F}$,*

$$\mu\pi_\Lambda(A) \equiv \int_\Omega \pi_\Lambda(A|\psi)\mu(d\psi) \tag{82}$$

and one can check that $\mu\pi_\Lambda$ is a measure on $(\Omega, \mathcal{F})$.

Lemma B.6. *For every bounded measurable function f, measure μ on $(\Omega, \mathscr{F})$, and proper probability kernel π_Λ, we have that*

$$\mu\pi_\Lambda(f) \equiv \int_\Omega f(\psi)\mu\pi_\Lambda(d\psi) = \int_\Omega \pi_\Lambda f(\psi)\mu(d\psi) \equiv \mu(\pi_\Lambda f). \tag{83}$$

Proof. One can see that when $f = \mathbb{1}_A$ for some $A \in \mathscr{F}$, above follows from the definition. Then, because each Lebesgue integral of a bounded measurable function is given as a limit of a linear combination of indicator functions, we have the desired equality. □

Definition B.7 (Compatible Measure to a Specification)**.** *Let $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{N}}$ is a specification. A measure μ on $(\Omega, \mathscr{F})$ is compatible with π if for every $\Lambda \Subset \mathbb{N}$,*

$$\mu = \mu\pi_\Lambda. \tag{84}$$

The set of all measures compatible with π is denoted $\mathcal{G}(\pi)$.

Now, we try to relate the probability kernel to conditional probability.

Theorem B.8 (Relating Probability Kernel with Conditional Probability.)**.** *A measure μ is compatible with a specification $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{N}}$ if and only if each kernel π_Λ provides a regular version of $\mu(\cdot|\mathscr{F}_c)$, the conditional probability with respect to $\mathscr{F}_{\Lambda^c}$, i.e. for every $A \in \mathscr{F}$, $\mu(A|\mathscr{F}_{\Lambda^c})(\cdot) = \pi_\Lambda(A|\cdot)$ μ-almost surely.*

To prove this theorem, we introduce a small lemma.

Lemma B.9. *Assume that π_Λ is proper. Then for every $A \in \mathscr{F}$ and for every $B \in \mathscr{F}_{\Lambda^c}$, we have that*

$$\pi_\Lambda(A \cap B|\cdot) = \pi_\Lambda(A|\cdot)\mathbb{1}_B(\cdot). \tag{85}$$

Proof. We consider two cases: when $\psi \in B$ and when $\psi \notin B$. When $\psi \in B$, Since π_Λ is proper, we have that $\pi_\Lambda(B|\psi) = \mathbb{1}_B(\psi) = 1$. Then, we get that

$$\pi_\Lambda(A \cap B|\psi) = \pi_\Lambda(A|\psi) - \pi_\Lambda(A \cap B^c|\pi) = \pi_\Lambda(A|\psi) = \pi_\Lambda(A|\psi)$$

where the first equality follows from additivity of measure and the second from the fact that $\pi_\Lambda(\cdot|\psi)$ is a probability measure. For the second case when $\psi \notin B$, we have that $\pi_\Lambda(B|\psi) = 0$ and thus

$$\pi_\Lambda(A \cap B|\psi) = 0 = \pi_\Lambda(A|\psi)\mathbb{1}_B(\psi)$$

where the first equality follows from monotonicity of measure. □

Proof of Theorem B.8. First suppose μ is compatible to $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{N}}$. Then for every $A \in \mathscr{F}_\Lambda$ and $B \in \mathscr{F}_{\Lambda^c}$, we have that

$$\int_B \pi_\Lambda(A|\psi)\mu(d\psi) = \int_\Omega \pi_\Lambda(A \cap B|\psi)\mu(d\psi) = \mu\psi_\Lambda(A \cap B) = \mu(A \cap B).$$

Recall the definition of conditional expectation:

$$\mu(A \cap B) = \int_B \mu(A|\mathscr{F}_{\Lambda^c})(\psi)\mu(d\psi).$$

By the uniqueness of conditional probability (c.f. Theorem A.3) we have that μ-almost surely, we have

$$\mu(A|\mathscr{F}_{\Lambda^c})(\cdot) = \pi_\Lambda(A|\cdot), \tag{86}$$

showing that each kernel π_Λ provides a regular version of $\mu(\cdot|\mathscr{F}_{\Lambda^c})$.

Conversely, suppose π_Λ provides a regular version of $\mu(\cdot|\mathscr{F}_{\Lambda^c})$ for every $\Lambda \Subset \mathbb{N}$. Then, for every $A \in \mathscr{F}$, we have that

$$\mu\pi_\Lambda(A) = \int_\Omega \pi_\Lambda(A|\psi)\mu(d\psi) = \int_\Omega \mu(A|\mathscr{F}_{\Lambda^c})(\psi)\mu(d\psi) = \mu(A \cap \Omega) = \mu(A).$$

Thus, we have that μ is compatible with π. □

References

[1] Michael Aizenman and Barry Simon. Local ward identities and the decay of correlations in ferromagnets. *Communications in Mathematical Physics*, 77:137–143, 1980.

[2] Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., third edition, 1995.

[3] David Brydges, Jürg Fröhlich, and Thomas Spencer. The random walk representation of classical spin systems and correlation inequalities. *Communications in Mathematical Physics*, 83(1):123–150, 1982.

[4] R. L. Dobrushin and S. B. Shlosman. Completely analytical interactions: Constructive description. *Journal of Statistical Physics*, 46(5):983–1014, Mar 1987.

[5] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, 5th edition, 2019.

[6] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, 2nd edition, 1999.

[7] Sandro Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, Cambridge, 2017.

[8] G.S. Joyce. Classical heisenberg model. *Physical Review*, 155(2), 1967.

[9] Elliott H Lieb and Alan D Sokal. A general lee-yang theorem for one-component and multicomponent ferromagnets. *Communications in Mathematical Physics*, 80(2):153–179, 1981.

[10] James Munkres. *Topology (2nd Edition)*. Pearson, 2000.

[11] Ron Peled and Yinon Spinka. Lectures on the spin and loop $O(n)$ models. *arXiv preprint*, arXiv:1708.00058, 2017. Version v3, revised 3 Jul 2019.

[12] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics I: Functional Analysis*. Academic Press Inc., 1980.

[13] Walter Rudin. *Real and Complex Analysis (Higher Mathematics Series)*. McGraw-Hill Education, 1986.

[14] Walter Rudin. *Functional Analysis*. McGraw-Hill Science/Engineering/Math, 1991.

[15] Jacob Shapiro. Lecture notes on measure theory, May 2025. [Online; accessed October 13, 2025].

[16] Daniel Ueltschi. Notes on the symanzik representation (corrected). Unpublished manuscript, 2025.