Statistical Field Theory 1

1D Ising Model to Quantum Spin

Benjamin Gerraty

30 November 2023

1 Introduction

The overarching topic of this seminar is studying statistical mechanical systems, with a goal of looking towards neural networks. This talk aims to be an introduction to studying statistical mechanical systems in their thermodynamic limit ($N \to \infty$) as *field theories*. Statistical field theory (SFT) [Mus20, Sha17] deals with the behaviour or classical or quantum systems consisting of an enormous number of degrees of freedom, systems that have different phases, and we aim to study primarily the phase transitions between then, and primarily what happens at the critical points these transitions occur.

We will be using the Ising Model as our pathfinder through this story from Statistical Mechanics to Conformal Field Theory, which is a type of statistical field theory where the theory may be solved exactly. There are several reasons to choose the Ising Model, namely is that is demonstrates the features of phase transitions, while having simple Boolean degrees of freedom $s_i = \pm 1$ defined on a lattice, and that the model may be solved exactly in the $d = 1$ and $d = 2$ dimensions.

Using the path integral formulation of quantum mechanics, as the lattice spacing of the 2D Ising model goes to zero, one can quantum field theory (QFT). Indeed, we will see that slightly away from the critical temperature T_c and in the absence of an externel magnetic field, the two-dimensional Ising model is equivalent to a fermionic neutral particle (Majorana fermion) that satisfies a Dirac equation. The main advantage of QFT is that it embodies a strong set of contrains coming from the compatibility of quantum mechanics with special relativity, which turns into general relations such as the completeness relation of multiparticle states or the unitarity of their scatttering process. These general properties make QFT help us understand the underlying aspect of phase transitions that may appear mysterious in the discete formulation of the statistical model.

At the critical points of interest, the QFT's are scale invariant and the correlation length ξ will diverge, and such theories are invariant under a larger symmetry group, which we will see is the conformal group: the set of geometric transformations that implement the scaling of the length of the vectors while preserving their relative angle. Quantum field theories that are invariant under conformal transformations are called conformal field theories (CFTs) [Rid13, PDF97]. In two dimensions, the conformal transformations coincide with the mappings by analytic functions of a complex variable, characterised by an infinite dimensional algebra known as the Virasoro algebra $\mathfrak{Vir}$. This enables us to determine the exact expressions of the correlation functions of by solving certain differential equations. Using these tools from CFT, our primary goal of these talks is to build towards computing the correlation functions for the 2D Ising Model at the critical point, but you can see [Ger21] for spoilers.

Its worth mentioning that away from the critical points, the quantum field theories are, in general, quite massive, and analysis is often carried out through perturbation approaches. There are some favourable cases that give rise to integrable models, but we won't consider these situations in these SFT talks. At some point in the near future the hope is that the tools and formalism combined with singular learning theory (SLT) will have something interesting to say about the loss functions $L(w)$ of toy neural networks such as the toy model of superposition (TMS) [CLM+23, EHO+22].

2 Statistical Mechanics and Scaling Laws

With a goal of looking towards field theories of statistical models, since these field theories are described in language of Euclidean Quantum Field Theories, it will be necessary to build towards the the path integral approach for a quantum system. We will instead start with a $d = 1$ Ising model, and show how this corresponds to a *single* spin$-\frac{1}{2}$ particle evolving in time. While this example won't be associated to a field theory, it will serve nicely as a refresher of quantum mechanics, and make the analysis we care about ($d = 2$ Ising Model) easier. This is the simplest way to to learn about how a d-dimensional classical problem may be mapped onto a $d-1$-dimensional quantum problem.

In general, the allowed values for a classical variables ($s_i = \pm 1$ for the Ising model) will correspond to the maximal set of eigenvalues of the operators in the quantum problem (the third Pauli matrix σ^3 in our example). The correlation functions will become the expectaton values in the ground state of a certain *transfer matrix*. The different sites in the $d = 1$ lattice will correspond to different, discrete times in the life a quantum degree of freedom.

The Ising model [Ton17, Ton12] is easy to state, unfortunately, it is hard to solve. We have a lattice in d spatial dimensions with N lattice sites. On each lattice site $i = 1, \ldots, N$ there is a spin s_i which can either be spin up $s_i = 1$ or spin down $s_i = -1$. The collection of spins $\{s_i\}$ has energy

$$\mathcal{H}[s_i] = -B \sum_i s_i - J \sum_{\langle ij \rangle} s_i s_j. \tag{2.1}$$

The first term is due to an external magnetic field B that we impose on the system, where if $B > 0$ the all the spins will want to be up ($\uparrow$) to lower their energy. The second term is the interactions between the neighbouring spins, where $\langle ij \rangle$ means that we sum over all the "nearset neighbour" pairs in the lattice. The number of such pairs depends on the dimension d and the type of lattice.

If $J > 0$, neighbouring spins prefer to be aligned ($\uparrow\uparrow$ or $\downarrow\downarrow$) and we refer to the system as a *ferromagnet*. If $J < 0$ the spins want to be opposing ($\downarrow\uparrow$ or $\uparrow\downarrow$) which makes the system an *anti-ferromagnet*. Going forward we will assume $J > 0$, although the differences between the systems are minor.[1]

We are interested in the physics of the Ising model at finite temperature T. The interaction energy encourages the spins to align the same way, and the external magnetic field encourages the spins to align in the same direction. Meanwhile, the temperature encourages the spins to ignore both the interactions and external magnetic as is increases. Since there are many more random configurations than aligned configurations, the temperature will mess up the nicely ordered states the interactions and magnetic field prepared.

In the canonical ensemble, the probability of the system sitting in a certain configuration of spins $\{s_i\}$ is given by

$$p[s_i] = \frac{1}{Z} e^{-\beta \mathcal{H}[s_i]}, \tag{2.2}$$

where $\beta = 1/T$ and Z is the *partition function* given by

$$Z(T, J, B) = \sum_{\{s_i\}} e^{-\beta \mathcal{H}[s_i]}. \tag{2.3}$$

If we're able to perform the sum to compute Z, we can extract any information about the system we want to know. For example, the thermodynamic free energy

$$F_{\text{thermo}}(T, B) = -T \log Z. \tag{2.4}$$

Here the free energy $F_{\text{thermo}}(T, B)$ is a function of the thermodynamic variables temperature T and external magnetic field strength B.

[1] In the anti-ferromagnetic case ($J < 0$) on certain lattices it is not possible for spins to be opposite to all their neighbours, such as with $d = 2$ on a triangular lattice.

One quantity we will be particularly interested in will be the average spin of the configuration, we will we refer to as the equilibrium *magnetisation*

$$m = \frac{1}{N}\left\langle \sum_i s_i \right\rangle. \tag{2.5}$$

This quantity will take values in the range $m \in [-1, 1]$. From our discussion above, we would expect that for $B > 0$ we will have $m \to 1$ at low temperatures where the spins are ordered, and $m \to 0$ at high temperatures where spins are arranged randomly. To make this intuition more precise, using (2.2) we can check that the magnetisation can be written as

$$m = \frac{1}{N}\sum_{\{s_i\}} \frac{e^{\beta \mathcal{H}[s_i]}}{Z} \sum_i s_i = \frac{1}{N\beta}\frac{\partial \log Z}{\partial B}, \tag{2.6}$$

which can be seen via

$$\begin{aligned} \frac{1}{N\beta}\frac{\partial \log Z}{\partial B} &= \frac{1}{N\beta}\frac{\partial}{\partial B}\log\left(\sum_{\{s_i\}} e^{-\beta(-B\sum_i s_i - J\sum_{\langle ij\rangle} s_i s_j)}\right) \\ &= \frac{1}{N\beta}\frac{1}{Z}\sum_{\{s_i\}} \beta\left(\sum_i s_i\right) e^{(\beta B\sum_i s_i + \beta J\sum_{\langle ij\rangle} s_i s_j)} \\ &= \frac{1}{N}\sum_{\{s_i\}} \frac{e^{\beta\mathcal{H}[s_i]}}{Z}\sum_i s_i \\ &= m. \end{aligned} \tag{2.7}$$

Taking further derivatives allow us to compute higher moment of the distribution, which we will see later. From this point onwards we are going to consider the partition function to be

$$Z = \sum_{\{s_i\}} \exp\left(\sum_{i=0}^{N-1} K(s_i s_{i+1} - 1)\right) \tag{2.8}$$

where $K = \beta J > 0$ and the subtraction merely adds an addition spin-independent constant of $-K$ to every site for coveience, which will shift the free energy βF by NK. Another observable of interest will be the correlation functions of different spins

$$G^{(2)}(i,j) = \langle s_i s_j\rangle = \frac{1}{Z}\sum_{\{s_i\}} s_i s_j \exp\left(\sum_k K(s_k s_{k+1} - 1)\right). \tag{2.9}$$

When there is translation invariance, One would expect the correlation of two observables to be a function of their *correlation length* ξ

$$G^{(2)}(i,j) \sim \frac{e^{-|j-i|/\xi}}{|j-i|^{d-2+\eta}}, \tag{2.10}$$

which highlights that we expect two spins i and j to have exponential decay as a function of their distance except when the correlation length diverges $\xi \to \infty$ at a critical point and we expect power law behaviour. This formula and reason for expected power law behaviour should not be obvious, it is a long story that we will not go into today but [Ola20] are some great lecture notes on this topic. To clarify, we will be interested in the case when $\xi(T = T_c) \to \infty$ and $G^{(2)}(i,j) \sim |j-i|^{-\eta}$ when $d = 2$ using CFT to compute η for different observables.

3 Interlude: Notation and Pictures

Before discussing how this classical $d = 1$ Ising model can be mapped onto a $d = 0$ Quantum spin-1/2 particle, we need to quickly review some notation. Quantum mechanics, quantum field

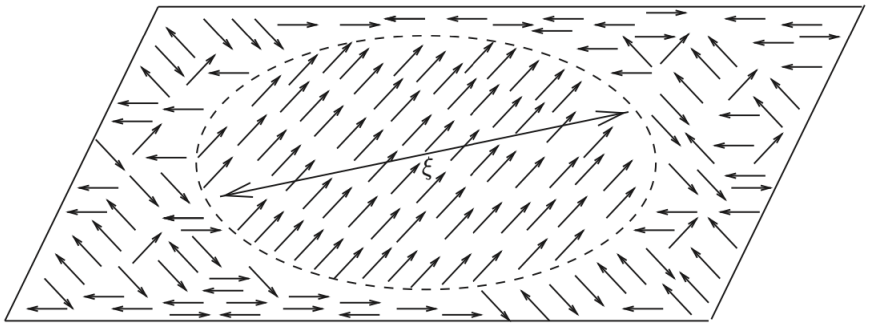

Figure 2.1: Correlation Length $\xi(T)$

theory and conformal field theory all use the *Dirac notation* or *bra-ket notation* introduced by Dirac. Much of this notation explanation can be found in [Hal13, Vol21]. This notation is is cumbersome to get used to, but pretty much unavoidable when reading any reference in quantum physics or the statistical field theory we will be interested in due to their formulation as QFTs and CFTs.

3.1 Dirac (Bra-Ket) Notation

Notation 3.1. *A vector ψ in $\mathbf{H}$ (Hilbert space or Quantum state space) is referred to as a **ket** and is denoted by $|\psi\rangle$. A continuous linear functional on $\mathbf{H}$ is called a **bra** such that for any $\phi \in \mathbf{H}$, let $\langle\phi|$ denote the bra. This is function such that*

$$\langle\phi| : \mathbf{H} \to \mathbb{C}, \quad \langle\phi|(\psi) = \langle\phi, \psi\rangle. \tag{3.1}$$

The bracket (or bra-ket) of two vectors $\phi, \psi \in \mathbf{H}$ is the result of applying the bra $\langle\phi|$ to the ket $|\psi\rangle$, namely the inner product of ϕ and ψ, denoted $\langle\phi|\psi\rangle$.

In quantum mechanics, observables are *operators*.

Postulate 3.1 (Measurement)**.** *1. All physical quantities or **observables** (e.g. position x, momentum p, spin s, ...) are associated with linear, Hermitian **operators** (e.g (e.g. position $\hat{x}$, momentum $\hat{p}$, spin $\hat{\sigma}$, ...), so*

$$\textit{quantity } A \longleftrightarrow \textit{operator } \hat{A} = \hat{A}^\dagger. \tag{3.2}$$

*2. When an ideal **measurement** of the quantity A is performed, the experimental result is always an eigenvalue of $\hat{A}$.*

If A is an operator, it is often denoted by $\hat{A}$ to indicate so. Usually the hats are not included when reading the literature, but we will use it only when confusion could arise. For the operator A acting on $\mathbf{H}$ and $\phi \in \mathbf{H}$ a vector, we can form a linear functional $\langle\phi| A$, meaning the linear map $\psi \mapsto \langle\phi|A\psi\rangle$. Physicists generally write can expression of this form as

$$\langle\phi| A |\psi\rangle . \tag{3.3}$$

This notation emphasises that there are two different ways about thinking about this quantity. We may think of $\langle\phi| A |\psi\rangle$ either as the linear functional $\langle\phi| A$ applied to the vector $|\psi\rangle$ or as the the linear functional $\langle\phi|$ applied to the vector $A|\psi\rangle$.

Notation 3.2. *For any ϕ and ψ in $\mathbf{H}$, the expression $|\phi\rangle\langle\psi|$ denotes the linear operator on $\mathbf{H}$ given by*

$$(|\phi\rangle\langle\psi|)(\chi) = |\phi\rangle \langle\psi|\chi\rangle = \langle\psi|\chi\rangle |\phi\rangle . \tag{3.4}$$

That is, in mathematics notation, $|\phi\rangle\langle\psi|$ is the operator sending χ to $\langle\psi, \chi\rangle \phi$.

Notation 3.3. *Given a family of vectors in* $\mathbf{H}$ *labelled by, say, three indices* n, l *and* m*, rather than denoting these vectors as* $|\psi_{n,l,m}\rangle$*, a physicist will denote them simply as* $|n, l, m\rangle$*.*

Notation 3.3 does have a problem, if we have two different sets of vector labelled by the same set of indices, a mathematician can simply label them as $\phi_{n,l,m}$ and $\psi_{n,l,m}$ whereas a physicist has a problem.

As an example of the Dirac notation, suppose that an operator $\hat{H}$ has an orthonormal set of basis eigenvectors ψ_n. A physicist would express the decomposition of a general vector in terms of this basis as be

$$I = \sum_n |n\rangle\langle n|, \tag{3.5}$$

where ψ_n is represented simply as $|n\rangle$ and $|n\rangle\langle n|$ is (given that $|n\rangle$ is a unit vector) the orthogonal projection onto the one-dimensional subspace spanned by the vector $|n\rangle$. This is justified by the completeness of the Hilbert space $\mathbf{H}$ giving the decomposition for a vector ψ. In other words

$$\psi = \sum_n \langle\psi|\psi_n\rangle\psi_n \tag{3.6}$$

for a complete set of orthonormal basis vectors in $\mathbf{H}$. Therefore,

$$\langle\psi|\psi\rangle = \sum_n \langle\psi|\psi_n\rangle\langle\psi_n|\psi\rangle \tag{3.7}$$

which implies (3.5). It is common practice for a physicist to refer to this as inserting "a complete set of states" into an inner product to compute observables.

Notation 3.4. *In the physics literature, the complex conjugate of a complex number* z *is denoted as* z^* *rather than* $\overline{z}$*. What a mathematician calls the adjoint of an operator and denoted by* A^**, a physicist calls the Hermitian conjugate of* A *and denotes it by* $A^\dagger$*. Physicists refer to self-adjoint operators as Hermitian.*

We may express the concept of an adjoint of an operator (or Hermitian conjugate) of an operator using Dirac notation. If A is a bounded operator on $\mathbf{H}$. then $A^\dagger$ is the unique bounded operator such that

$$\langle\psi|\, A = \langle A^\dagger\psi|\,. \tag{3.8}$$

Notation 3.5. *Using Dirac notation, given an operator* $\hat{O}$*, matrix elements are expressed as*

$$\begin{aligned}
\hat{O} &= 1\hat{O}1 \\
&= \left(\sum_{n=1}^{N} |n\rangle\langle n|\right)\hat{O}\left(\sum_{m=1}^{N} |m\rangle\langle m|\right) \\
&= \sum_{n=1}^{N} |n\rangle\langle n| \sum_{m=1}^{N} |n\rangle\langle n|\hat{O}|m\rangle\langle m| \\
&= \sum_{n=1}^{N}\sum_{m=1}^{N} O_{nm}|n\rangle\langle m|
\end{aligned} \tag{3.9}$$

where we matrix elements $\hat{O}$ *with respect to the basis* $|n\rangle$ *are* O_{nm}*.*

3.2 Schrödinger v. Heisenberg Picture of Quantum Mechanics

In the Schrödinger picture, state vectors evolve in time $|\psi(t)\rangle$ through the Schrödinger equation

$$i\hbar\frac{d\,|\psi(t)\rangle}{dt} = H\,|\psi(t)\rangle \tag{3.10}$$

but the operators of fundamental observables such as momentum $\hat{p}$ and position $\hat{x}$ do not. Given any inital state $|\psi(0)\rangle$ we can find its future evolution by solving this equation. If we think of the vector $|\psi(t)\rangle$ and starting at $|\psi(0)\rangle$ evolving via some time operator $U(t)$ so that

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle\,, \tag{3.11}$$

then the Schrödinger equation

$$i\hbar\frac{dU(t)}{dt} = HU(t) \tag{3.12}$$

gives a formal solution $U(t) = e^{-i\frac{H}{\hbar}t}$, which is unitary $U^\dagger U = I$.

In the Heisenberg picture, operators evolve in time via time evolution operator $U(t)$. Therefore we define the *Heisenberg Operators* to be

$$\Omega(t) = U(t)^\dagger \Omega U(t) \tag{3.13}$$

where Ω is the Schrödinger picture operator. The upshot here is that either picture will compute the same observables

$$\langle\psi(t)|\,\Omega\,|\psi(t)\rangle = \langle\psi(0)|\,U(t)^\dagger\Omega U(t)\,|\psi(0)\rangle = \langle\psi(0)|\,\Omega(t)\,|\psi(0)\rangle\,. \tag{3.14}$$

One can perform a *Wick* rotation $t = -i\tau$ to consider Imaginary time (or Euclidean Quantum Mechanics) in all the results above, which is what corresponds to statistical mechanics. If we make such a rotation and set $t = -i\tau$ then set $\hbar = 1$, we get

$$U(\tau) = e^{-\tau H}. \tag{3.15}$$

4 Statistical to Quantum Mechanics

Now for the derivation of the Euclidean quantum problem underlying the classical $d = 1$ Ising model. We consider a model with periodic boundary conditions, so site N is the same as site 0. Consider the partition function

$$Z = \sum_{\{s_i\}} e^{-\beta\sum_{i=1}^n -J(s_i s_{i+1}-1)} = \sum_{\{s_i\}}\prod_i e^{K(s_i s_{i+1}-1)} \tag{4.1}$$

where $K = \beta J$ and we have added a spin-independent constant of $-K$ added to each site for convenience later. This will merely shift the free energy βF by NK. Each exponential factor $e^{K(s_i s_{i+1}-1)}$ has two degrees of freedom s_i, s_{i+1} which can both two values, so each factor can have 4 values. We will now introduce a 2×2 *transfer matrix* whose rows and columns are labelled by these spins s', s will have matrix elements

$$T_{s's} = \langle s'|\,T\,|s\rangle = e^{K(s's-1)}, \tag{4.2}$$

In other words we will have elements

$$T_{++} = T_{--} = 1, \quad T_{+-} = T_{-+} = \exp(-2K). \tag{4.3}$$

The transfer matrix describes if there is a spin So

$$T = \begin{pmatrix} 1 & e^{-2K} \\ e^{-2K} & 1 \end{pmatrix} = I + e^{-2K}\sigma_1 \tag{4.4}$$

where I is the 2×2 identity and σ_1 is the first Pauli matrix. The transfer matrix T here is both real and Hermitian, so $T^\dagger = T$.

Definition 4.1 (Pauli Matrices)**.** We have seen that (at least classically) electrons have this property called spin, and it something to do with the *Pauli Matrices*, given below

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.5}$$

with commutation ($[A,B] = AB - BA$) relations

$$[\sigma_1,\sigma_2] = 2i\sigma_3, \quad [\sigma_2,\sigma_3] = 2i\sigma_1, \quad [\sigma_3,\sigma_1] = 2i\sigma_2. \tag{4.6}$$

We can then use the transfer matrix to rewrite out partition function (4.1) as

$$\begin{aligned}
Z &= \sum_{\{s_i\}_{i=0}^{N}} \prod_{i=0}^{N-1} e^{K(s_i s_{i+1}-1)} \\
&= \sum_{s_0}\sum_{s_2}\cdots\sum_{s_N}\prod_{i=0}^{N-1} \langle s_i|\, T\, |s_{i+1}\rangle \\
&= \sum_{s_0=\pm 1}\sum_{s_2=\pm 1}\cdots\sum_{s_N=\pm 1} \langle s_0|\, T\, |s_1\rangle \langle s_1|\, T\, |s_2\rangle \cdots \langle s_{N-1}|\, T\, |s_N\rangle \qquad (4.7)\\
&= \sum_{s_0}\sum_{s_N} \langle s_0|\, T \left(\sum_{s_1} |s_1\rangle\langle s_1|\right) T \left(\sum_{s_2} |s_2\rangle\langle s_2|\right)\cdots\left(\sum_{s_{N-1}} |s_{N-1}\rangle\langle s_{N-1}|\right) T\, |s_N\rangle\,, \quad (\dagger)\\
&= \sum_{s_0}\sum_{s_N} \langle s_0|\, T^{N-1}\, |s_N\rangle\,.
\end{aligned}$$

In the step (†) we are merely emphasising that only those operators are summed or and the sum is brought in to highlight the connection to the identity operator. If we are considering the case of periodic boundary conditions, where $s_0 = s_N$, then this becomes the trace of the partition function

$$Z = \operatorname{Tr} T^N \tag{4.8}$$

Let us return now to

$$T = I + e^{-2K}\sigma_1 \tag{4.9}$$

We would like to combine this transfer matrix into a matrix exponential. Let us now define $K^*(K)$ (note K^* is not the complex conjugate of K) as

$$\tanh K^*(K) = e^{-2K}. \tag{4.10}$$

. Now since all the pauli matrices square to the identity matrix, we can make us of

$$\begin{aligned}
e^{K^*\sigma_1} &= \sum_{k=0}^{\infty} \frac{(K^*)^k \sigma_1^k}{k!} \\
&= \sum_{k \text{ even}} \frac{(K^*)^k}{k!} + \sum_{k \text{ odd}} \frac{(K^*)^k}{k!}\sigma_1 \qquad (4.11)\\
&= \cosh K^* + \sinh K^* \sigma_1 \\
&= \cosh K^*(I + \tanh K^* \sigma_1).
\end{aligned}$$

Therefore we have

$$T = \frac{e^{K^*(K)\sigma_1}}{\cosh K^*(K)}, \tag{4.12}$$

where we will often drop the denominator, since it is not relevant in all averages and correlation functions we will be interested in. We refer to K^* as the *dual* of K, and if you invert (4.10) you get

$$\tanh K = e^{-2K^*} \tag{4.13}$$

so K is the *dual* of K^* too. Notice that when one is small, the other is large. Calculating the eigenvalues of (4.12) one obtains

$$\lambda_0 = e^{K^*}, \quad \lambda_1 = e^{-K^*}, \quad \frac{\lambda_1}{\lambda_0} = e^{-2K^*}, \tag{4.14}$$

with the corresponding eigenvectors

$$|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, |1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}. \tag{4.15}$$

It will be useful in computing the correlation functions that

$$\sigma_3 |0\rangle = |1\rangle , \quad \sigma_3 |1\rangle = |0\rangle \tag{4.16}$$

which can be checked using the Pauli matrices and the eigenvectors. If we have a basis of eigenvectors for the third Pauli matrix σ_3 instead, we will then have

$$\sigma_3 |+\rangle = + |+\rangle , \quad \sigma_3 |-\rangle = - |-\rangle , \tag{4.17}$$

where

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , |-\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{4.18}$$

4.1 Correlation Functions

Now lets consider the correlation function we are primarily interested in computing. Recall (2.9) which was

$$\langle s_j s_i \rangle = \frac{1}{Z} \sum_{\{s_i\}} s_j s_i \exp \left(\sum_k K(s_k s_{k+1} - 1) \right) \tag{4.19}$$

Now we claim that if the boundary spins are fixed at values s_0 and s_N (so we don't sum over them) that

$$\langle s_j s_i \rangle = \frac{\langle s_N | T^{N-j} \sigma_3 T^{j-i} \sigma_3 T^i | s_0 \rangle}{\langle s_N | T^N | s_0 \rangle} \tag{4.20}$$

To see why this is correct, we can insert a complete set of σ_3 eigenstates between each factor of T, and retrace the calculation (4.7), which gives

$$\begin{aligned}
\frac{\langle s_N | T^{N-j} \sigma_3 T^{j-i} \sigma_3 T^i | s_0 \rangle}{\langle s_N | T^N | s_0 \rangle} &= \frac{\langle s_N | T^{N-j} \sigma_3 T^{j-i} \sigma_3 \left(\sum_{s_i} |s_i\rangle \langle s_i| \right) T \cdots T \left(\sum_{s_1} |s_1\rangle \langle s_1| \right) T |s_0\rangle}{\langle s_N | T^N | s_0 \rangle} \\
&= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_{N-1}=\pm 1} \frac{1}{Z} \langle s_N | T | s_{N-1} \rangle \cdots \langle s_{j+1} | T | s_j \rangle \\
&\quad s_j \langle s_j | T | s_{j-1} \rangle \cdots \langle s_{i+1} | T | s_i \rangle \, s_i \langle s_i | T | s_{i-1} \rangle \cdots \langle s_1 | T | s_0 \rangle \\
&= \frac{1}{Z} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_N} s_j s_i \prod_{i=0}^{N-1} \langle s_{i+1} | T | s_i \rangle \\
&= \frac{1}{Z} \sum_{\{s_i\}_{i=1}^{N-1}} \prod_{i=0}^{N-1} s_j s_i \, e^{K(s_i s_{i+1} - 1)} \\
&= \langle s_j s_i \rangle .
\end{aligned} \tag{4.21}$$

which is the classical two-point correlation function. You should read (4.20) as, reading from right to left with the complete set of states inserted, T^i gives a sum of the Boltzman weights until we get to site i, then σ_3 acts on this eigenstate and gives the eigenvalue (spin) s_i, then we process via transfer matrix to site j and pull out s_j, then continue to the N-th site. The denominator is just the partition function. Let's rewrite this is another way, Heisenberg operators are defined via (3.13). So we will define analagously

$$\sigma_3(n) = T^{-n} \sigma_3 T^n , \tag{4.22}$$

where the index n plays the role of discrete integer-valued time, and T is thought of as the time-evolution operator for one unit of Euclidean time. In terms of these operators the correlation function becomes

$$\langle s_j s_i \rangle = \frac{\langle s_N | T^N \sigma_3(j) \sigma_3(i) | s_0 \rangle}{\langle s_N | T^N | s_0 \rangle} . \tag{4.23}$$

Notice dropping the $\cosh K^*$ does not affect $\langle s_j s_i \rangle$ since it cancels out. Now we can also expand the transfer matrix using the spectral theorem as the operator

$$T = \lambda_0 |0\rangle\langle 0| + \lambda_1 |1\rangle\langle 1| . \tag{4.24}$$

and

$$T^N = \lambda_0^N |0\rangle\langle 0| + \lambda_1^N |1\rangle\langle 1|. \tag{4.25}$$

Now since both λ_0 and λ_1 are nonzero (4.14), we can use the *Perron-Frobenius Theorem* which states that a square matrix with non-zero eigenvalues will have a largest eigenvalue and a corresponding eigenvector with strictly positive components. Assuming λ_0 is the biggest one, we can approximate

$$\lim_{N\to\infty} T^N \approx \lambda_0^N \left[|0\rangle\langle 0| + O\left(\frac{\lambda_1}{\lambda_0}\right)\right] \approx \lambda_0^N |0\rangle\langle 0|. \tag{4.26}$$

Consider now the limit as $N \to \infty$ (called the thermodynamic limit) for (4.23), with i and j fixed at values far from 0 and N (the end points), so $N - j$ and i are both large. We may then approximate the decomposition (4.25) by

$$T^\alpha \approx \lambda_0^\alpha |0\rangle\langle 0|, \quad \alpha = N, N-j, i. \tag{4.27}$$

In this limit we have

$$\langle s_j s_i\rangle = \frac{\langle s_N|0\rangle\, \lambda_0^{N-j} \sigma_3 T^{j-i} \sigma_3 \lambda_0^i\, |0\rangle\, \langle 0|s_0\rangle}{\langle s_N|0\rangle\, \lambda_0^N\, \langle 0|s_0\rangle} = \langle 0|\, \sigma_3(j)\sigma_3(i)\, |0\rangle\,. \tag{4.28}$$

If $i > j$, we will get the operators in the reverse order. If we define the *time-ordering symbol* $\mathcal{T}$ by

$$\mathcal{T}(\sigma_3(i)\sigma_3(j)) = \begin{cases} \sigma_3(i)\sigma_3(j), & \text{if } i > j \\ \sigma_3(j)\sigma_3(i), & \text{if } j < i. \end{cases} \tag{4.29}$$

so that in general we have

$$\langle s_j s_i\rangle = \langle 0|\, \mathcal{T}(\sigma_3(j)\sigma_3(i))\, |0\rangle \tag{4.30}$$

This operator basically just ensures that events that occur first, act on the state variable first. The reason for connect i and j with time will be clarified when we get to the Hamiltonian later. Let's finish off this discussion on the correlation function but computing

4.2 The Hamiltonian

We have seen that the transfer matrix T plays the role of the time-evolution operator, given that the Heisenberg operators that arose naturally were defined as

$$\sigma_3(j) = T^{-j}\sigma_3 T^j. \tag{4.31}$$

If we really are to identify T with $U(\tau) = e^{-\tau H}$, what is τ? It is one step on our discrete time lattice. Setting the lattice spacing to $\tau = 1$, we can introduce a Hamiltonian via

$$T = e^{-H}. \tag{4.32}$$

Now we see the reason for wanting to define our transfer matrix as an exponential, and we can read off

$$H = -K^*\sigma_1 \tag{4.33}$$

which describes a single spin-$\frac{1}{2}$ particle in the x-plane, as promised. We can write this with our initial parameters with $K = \beta J$ and $\tanh K^*(K) = e^{-2K}$ as

$$H = \operatorname{arctanh}(e^{-2\beta J}) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.34}$$

5 Summary: Classical to Quantum

We have seen that the partition function and correlation functions of the classical Ising problem can be restated in terms of a two-dimensional Hilbert space (the kinematic homeland of a single quantum spin-$\frac{1}{2}$ degree of freedom). The Schrödinger operators of this theory and σ_1 (and σ_3 if

we were to include to external magnetic field), the operator σ_2 does not enter since there are no complex numbers in the classical problem. The Ising spins s are associated with eigenvalues of σ_3.

The Transfer matrix T plays the role of the Euclidean time-evolution operator for one unit of discrete time $\Delta\tau$

$$T \iff U(\Delta\tau), \tag{5.1}$$

where we have been using $\Delta\tau = 1$. The $N+1$ points in the spatial lattice correspond to the $N+1$ instants of Euclidean time $\tau = j\Delta\tau$. The finite Ising chain corresponds to the time evolution over a finite time. Next the partiton function Z with $s_0 = s_i$ and $s_N = s_f$ where i and f stand for initial and final

$$\langle s_N = s_f | \, T^N \, | s_0 = s_i \rangle \iff \langle s_f | \, U(N\Delta\tau) \, | s_i \rangle \,, \tag{5.2}$$

corresponds to the matrix elements of the propagator U for imaginary time $N\Delta\tau$ between the states $\langle s_f|$ and $|s_i\rangle$. The Heisenberg operators are related via the Schrödinger operators as expected

$$\sigma_3(j) = T^{-j}\sigma_3 T^j \iff U^{-1}(j\Delta\tau)\sigma_3 U(j\Delta\tau) = \sigma_3(\tau = j\Delta\tau). \tag{5.3}$$

If we formally define a Hamiltonian by

$$T = e^{-H\Delta\tau} \tag{5.4}$$

then the dominant eigenvector of T is the ground state eigenvector $|0\rangle$ of H. In our $d = 1$ case, H was a simple operator $-K^*\sigma_1$. In general the T coming from a sensible statistical mechanics problem will be a nice operator, but its logarithm H need not be, meaning may involve arbitrary products of operators. The correlation function of the Ising model in the thermodynamic limit $N \to \infty$ is the ground state expectation value of the time-ordered product of the corresponding Heisenberg operators

$$\langle s_j s_i \rangle \iff \langle 0 | \, \mathcal{T}(\sigma_3(j)\sigma_3(i)) \, | 0 \rangle \,. \tag{5.5}$$

Next time we'll see how the two dimensional Ising Model reduces to a one dimensional quantum problem, with a path integral that will lead us into it's field theory.

References

[CLM+23] Zhongtian Chen, Edmund Lau, Jake Mendel, Susan Wei, and Daniel Murfet. Dynamical versus bayesian phase transitions in a toy model of superposition, 2023.

[EHO+22] Nelson Elhage, Tristan Hume, Catherine Olsson, Nicholas Schiefer, Tom Henighan, Shauna Kravec, Zac Hatfield-Dodds, Robert Lasenby, Dawn Drain, Carol Chen, Roger Grosse, Sam McCandlish, Jared Kaplan, Dario Amodei, Martin Wattenberg, and Christopher Olah. Toy models of superposition, 2022.

[Ger21] Benjamin Gerraty. The Structure Constants of the Minimal Models, 2021.

[Hal13] Brian C. Hall. *Quantum Theory for Mathematicans*. Springer, 2013.

[Mus20] G. Mussardo. *Statistical Field Theory: An Introduction to Exactly Solved Models in Statistical Physics*. Oxford University Press, 2020.

[Ola20] Enzo Olandini. Lecture Notes of Statistical Mechanics, 2020.

[PDF97] David Sénéchal Philippe Di Francesco, Pierre Mathieu. *Conformal Field Thoery*. Springer, 1997.

[Rid13] David Ridout. An Introduction to Conformal Field Theory, 2013.

[Sha17] R. Shankar. *Quantum Field Theory and Condensed Matter*. Cambridge University Press, 2017.

[Ton12] David Tong. Lectures on Statistical Physics, 2012.

[Ton17] David Tong. Lectures on Statistical Field Theory, 2017.

[Vol21] Ray Volkas. PHYC90007 Quantum Mechanics, 2021.