Phys 229ab Advanced Mathematical Methods: Conformal Field Theory

David Simmons-Duffin

Caltech, dsd@caltech.edu

Notes for Physics 229, 2017-2018. These notes are in progress. They were last updated on January 8, 2018.

Contents

1. Introduction

Resources

These introductory notes are heavily based on Silviu Pufu's Bootstrap 2017 lectures [] and John McGreevy's lectures on QFT [].

1.1. *QFT and emergent symmetry*

Quantum Field Theory is a universal language for theoretical physics. It shows up in many different settings, for example

- statistical physics,
- condensed matter physics,
- particle physics (SM and beyond),
- string theory/holography.

The microscopic physics in all of these settings can be quite complicated. However, often the macroscopic physics displays extra "emergent" symmetries that can help us do computations.

For example, in condensed matter physics, we are interested in describing a material made up of atoms with some lattice spacing a.[a] We refer to the detailed lattice system as the "microscopic theory." Quantum field theory is a good description at distances much larger than the lattice spacing $x \gg a$, or equivalently energy/momenta much lower than the UV cutoff $\Lambda_{UV} = 1/a$.[b] At these large scales, the discrete translation symmetry of the lattice becomes a continuous symmetry.

Similarly, in statistical physics, particle physics, and string theory, QFT can describe distance scales much larger than the characteristic scales of the microscopic theory.

Statistical systems are described by QFTs in Euclidean signature, e.g. on $\mathbb{R}^d$. Such QFTs capture properties of the equilibrium state. By contrast, condensed matter and particle systems are described by QFTs in Lorentzian signature, e.g. on $\mathbb{R}^{d-1,1}$. Such QFTs encode time-dependent quantum dynamics.

We will be interested in QFTs with rotational symmetry, by which we mean $\mathrm{SO}(d)$ symmetry in Euclidean signature and $\mathrm{SO}(d-1,1)$ symmetry in Lorentzian signature. In particle physics and string theory, this symmetry is built into the microscopic theory. However in lattice systems, rotational

[a] We ignore the possibility of lattice defects for the moment.
[b] QFTs that are valid below a UV cutoff are often called "effective field theories" (EFTs).

symmetry must be emergent. This means that correlation functions become rotationally invariant in the limit of large distances, even though microscopic correlation functions are not rotationally-invariant.[c] In particular, for condensed matter systems, the effective "speed of light" associated with $\mathrm{SO}(d-1,1)$-invariance is an emergent property and has nothing to do with the speed of actual light. (We will see some explicit examples later.)

Under general conditions, $\mathrm{SO}(d)$-invariant Euclidean QFTs are in one-to-one correspondence with $\mathrm{SO}(d-1,1)$-invariant Lorentzian QFTs. The map between them is called Wick rotation, and we will discuss it in detail. Because of this correspondence, we can focus mostly on Euclidean QFTs, and later understand Lorentzian QFTs by Wick rotating what we learned in Euclidean signature.

1.2. *The mass gap and critical points*

So far, we are interested in theories with Poincare symmetry

$$G_{\mathrm{Poincare}} = \mathbb{R}^d \rtimes \mathrm{SO}(d). \tag{1}$$

From the point of view of long-distance physics, the most important property of a Poincare-invariant theory is its mass gap

$$m_{\mathrm{gap}} = E_1 - E_0, \tag{2}$$

where E_0, E_1 are the energies of the ground state and first excited state, respectively.[d] Theories with $m_{\mathrm{gap}} > 0$ are called "gapped."

To understand why m_{gap} is important, let us study a two-point function of a scalar local operator $\phi(x) = \phi(x^0, \mathbf{x})$. We demand that $\phi(0)$ have vanishing vacuum expectation value by subtracting off an appropriate multiple of the unit operator. By Poincare invariance, it suffices to consider $\langle 0|\phi(x^0, \mathbf{0})\phi(0)|0\rangle$ with $x^0 > 0$. The two-point function is then given by

$$\begin{aligned}\langle 0|\phi(x^0, \mathbf{0})\phi(0)|0\rangle &= \langle 0|\phi(0)e^{-Hx^0}\phi(0)|0\rangle \\ &= \sum_\psi |\langle 0|\phi(0)|\psi\rangle|^2 e^{-E_\psi x^0}.\end{aligned} \tag{3}$$

Here, H is the Hamiltonian with the vacuum energy subtracted off, and ψ runs over an orthonormal basis of eigenstates of H. To get the right-hand

[c] Emergent rotational symmetry is very familiar: we often cannot determine the orientation of a microscopic lattice using macroscopic observations. Some examples of materials *without* emergent $\mathrm{SO}(d)$ symmetry are crystals like salt. A very exotic example is the Haah code [].

[d] Some QFTs with topological order can have multiple degenerate ground states.

side, we have used $\phi(x^0, \mathbf{0}) = e^{Hx^0}\phi(0,\mathbf{0})e^{-Hx^0}$ and $H|0\rangle = 0$.[e]

The key point is that the operator e^{-Hx^0} exponentially damps states with energy $E_\psi \gg 1/x^0$. At large x^0, the correlator is dominated by ψ with the smallest nonzero eigenvalue of H, which is m_{gap}.[f]

Thus, when m_{gap} is nonzero, correlation functions of local operators fall off at least as fast as $e^{-|x|/\xi}$, where $\xi \equiv 1/m_{\text{gap}}$ is called the "correlation length." Generic statistical and condensed matter systems have microscopic correlation lengths $\xi \sim a$, or equivalently $m_{\text{gap}} \sim \Lambda_{UV}$. At long distances, they are described by QFTs whose local correlation functions vanish, called topological quantum field theories (TQFTs).

However, sometimes by tuning parameters in the microscopic Hamiltonian, we can make m_{gap} much smaller than Λ_{UV}, and even arrange for m_{gap} to vanish. Points in parameter space where $m_{\text{gap}} = 0$ are called critical points. At a critical point, the system experiences a phase transition, and develops nonzero correlations at arbitrarily long distances.[g]

Long-distance correlation functions at critical points have no intrinsic length-scale because all memory of dimensionful microscopic quantities (like the lattice spacing a) disappears when distances become arbitrarily large. For example, critical two-point functions behave as pure power laws

$$\langle\phi(x)\phi(0)\rangle = \frac{C}{|x|^{2\Delta}} \qquad (\text{critical point}, x \gg a), \tag{4}$$

where C and Δ are constants depending on ϕ. The quantity Δ is called the scaling dimension of ϕ.

1.3. *Scaling and conformal symmetry*

A more precise way to state the lack of an intrinsic length scale is to say that theories with $m_{\text{gap}} = 0$ have an emergent symmetry under rescaling

$$x^\mu \to \lambda x^\mu \qquad (\lambda > 0). \tag{5}$$

Under very general conditions (that we will discuss), critical points also display less obvious emergent symmetries called conformal transformations.

[e]Note that the Euclidean time-evolution operator is e^{-Hx^0} as opposed to the familiar e^{-iHt} in Lorentzian signature. They are related by Wick rotation $x^0 = it$. We will discuss this in much more detail in later sections.

[f]Note that the vacuum does not contribute as an intermediate state because we have demanded $\langle 0|\phi(0)|0\rangle = 0$.)

[g]If the Standard Model were like a generic condensed matter system, we might expect m_{gap} to be close to the UV cutoff, which is perhaps the GUT scale 10^{15}GeV or Planck scale 10^{18}GeV. The hierarchy problem is the problem of explaining why the Standard Model is so close to a critical point.

A conformal transformation $x \to x'(x)$ is a map that looks like a rotation and rescaling near each point,

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \Omega(x) R^{\mu}{}_{\nu}(x), \qquad R^{\mu}{}_{\nu} \in \mathrm{SO}(d). \tag{6}$$

An example is a special conformal transformation,

$$x^{\mu} \to \frac{x^{\mu} - b^{\mu} x^2}{1 - 2b \cdot x + b^2 x^2} \qquad (b \in \mathbb{R}^d). \tag{7}$$

In 2-dimensions, there are more exotic examples, like the one pictured in figure **??**.[h]

Here is some rough intuition for why critical points display conformal symmetry. We know that a critical theory is invariant under rescalings and rotations. If the theory is also local, in the sense that degrees of freedom at a point only interact directly with other degrees of freedom at nearby points, then the theory should also be invariant under transformations that locally look like a rescaling and rotation. This is the defining property of a conformal transformation. Turning this rough intuition into a theorem is a difficult problem (that we will discuss in more detail later). However, it seems to be true in a wide class of systems.

QFTs that are invariant under conformal symmetry are called conformal field theories (CFTs). To summarize, CFTs describe critical points where $m_{\mathrm{gap}} = 0$. One can also understand the neighborhood of a critical point where m_{gap} is nonzero (but still $m_{\mathrm{gap}} \ll \Lambda_{UV}$) by studying perturbations of the associated CFT.

1.4. *Examples of critical points*

So far our discussion has been very abstract, so let us introduce some examples. One of our goals will be to infer from examples a set of axioms that CFTs should satisfy. We will study these axioms in the next part of the course. Another goal will be to be more precise about how and why statistical and condensed matter systems are described by quantum field theory, and how critical points come about.

1.4.1. *Magnets*

Our first examples of critical points occur in magnets. Given a magnet with temperature T, we can apply a magnetic field $\vec{H}$ and measure the

[h] We will give a precise definition for what it means for a theory to be invariant under transformations like (5) and (7) later in the course.

magnetization $\vec{M}$. There are three main types of magnets in 3-dimensions, which are distinguished by their symmetries:[i]

- *Uni-axial magnet*: individual magnetic moments $\vec{\mu}$ are confined to lie along a fixed axis. Uni-axial magnets have an $O(1) = \mathbb{Z}_2$ symmetry under which $\vec{H} \to -\vec{H}$ and $\vec{M} \to -\vec{M}$.
- *XY magnet*: magnetic moments $\vec{\mu}$ are oriented in a plane. Such magnets have an $O(2)$ symmetry under which $\vec{H}$ and $\vec{M}$ rotate in the plane.
- *Heisenberg magnet*: magnetic moments $\vec{\mu}$ are unconstrained. Such magnets have $O(3)$ symmetry under which $\vec{H}$ and $\vec{M}$ transform in the vector representation.

For the moment, we will focus on the simplest case of uni-axial magnets. We denote the projections of $\vec{H}, \vec{M}$ onto the appropriate axis by H, M.

In experiments, we observe the following. There exists a "critical temperature" T_c, such that

- For $T < T_c$, the preferred state of the magnet has nonzero magnetization $M \neq 0$ when $H = 0$. In other words, the $\mathbb{Z}_2$ symmetry is spontaneously broken.
- For $T > T_c$, the magnet has $M = 0$ when $H = 0$, i.e. the $\mathbb{Z}_2$ symmetry is unbroken.

The corresponding phase diagram is pictured in figure **??**. The point

$$H = 0,\ T = T_c \tag{8}$$

is a critical point, and is described by a CFT at long distances. To reach the critical point, we must tune two parameters: H and T. Tuning $H = 0$ is easy because that is where the microscopic theory has $\mathbb{Z}_2$ symmetry. However, the value of T_c depends on the specific material.

In more detail, the behavior of the magnetization in different phases is shown in figure **??**. Close to T_c, observables exhibit so-called "scaling" behavior, characterized by various critical exponents. Let us define the dimensionless couplings

$$t \equiv \frac{T - T_c}{T_c}, \qquad h = \frac{H}{k_B T}. \tag{9}$$

Some examples of critical exponents are

[i] Here, we mean non-spacetime symmetries, usually called "global" or "flavor" symmetries. The emergent spacetime symmetry group is still the Poincare group, or the conformal group at the critical point.

- α: the heat capacity at $h = 0$ behaves as

$$C = \frac{\partial^2 F}{\partial T^2} \propto |t|^{-\alpha}. \tag{10}$$

  (Here F is the free-energy.)
- β: the spontaneous magnetization behaves as

$$\lim_{H \to 0^+} M \propto (-t)^{\beta}. \tag{11}$$

- γ: the zero-field susceptibility behaves as

$$\chi = \left.\frac{\partial M}{\partial H}\right|_{H=0} \propto |t|^{-\gamma}. \tag{12}$$

- δ: the magnetization at $T = T_c$ behaves as

$$|M| \propto |h|^{1/\delta}. \tag{13}$$

- ν and η: the correlation length ξ can be measured by studying a two-point correlation functions of spins

$$G(x) = \langle s(x)s(0)\rangle - \langle s(0)\rangle^2. \tag{14}$$

  Away from the critical point, $G(x) \sim e^{-|x|/\xi}$ decays exponentially. However, as $t \to 0$, the correlation length diverges as

$$\xi \propto |t|^{-\nu}. \tag{15}$$

  Equivalently, the mass-gap goes to zero as $m_{\text{gap}} \propto |t|^{\nu}$. Precisely at $t = 0$, the two-point function takes the form

$$G(x) \propto \frac{1}{|x|^{d-2+\eta}}, \tag{16}$$

  i.e. the spin operator has dimension $\Delta_s = \frac{d-2}{2} + \frac{\eta}{2}$.

Don't worry, I can't keep track of all these critical exponents either. We will see shortly that all of this behavior can be explained using effective field theory, scaling symmetry, and dimensional analysis.

Now, here is an amazing fact:

We find the same critical exponents in many different uni-axial magnets, regardless of what material they're made of.

In fact, critical uni-axial magnets are all described by the same scale-invariant QFT at long distances. This phenomenon is called "critical universality."

1.4.2. *Liquid-vapor transitions*

Other critical points appear in liquid-vapor transitions. For example, the phase-diagram of water is pictured in figure **??**. Near room temperatures and pressures, there is a sharp distinction between the liquid and gas phases. However, at higher temperatures and pressures, the distinction between liquid and gas disappears at a critical point (T_c, P_c). For example, in water $T_c = 647\,\mathrm{K}$, $P_c = 374\,\mathrm{Atm}$.

Note that the critical points of magnets and water are both obtained by tuning two parameters. Comparing neighborhoods of the critical points in figures **??** and **??**, we can make the following rough analogy between water and magnets:

$$\begin{aligned} P - P_c &\sim H, \\ \rho - \rho_c &\sim M. \end{aligned} \tag{17}$$

where ρ is the density and ρ_c is the critical density.[j]

In measurements of critical water, we again observe scaling behavior. For example, the heat capacity behaves as

$$C \sim |t|^{-\alpha}, \tag{18}$$

where t is again given by (9). Additionally, the difference in density between the liquid and gas phases behaves as

$$\rho_{\text{liquid}} - \rho_{\text{gas}} \sim (-t)^{\beta}. \tag{19}$$

Amazingly,

> **Water and other liquid-vapor transitions have precisely the same critical exponents as uni-axial magnets.**

We say that liquid-vapor transitions are in the same "universality class" as uni-axial magnets.

1.4.3. *The Ising model*

The Ising model is a simplified model of a magnet that is still rich enough to capture its critical behavior. Its degrees of freedom are classical spins

[j] Water does not have a microscopic $\mathbb{Z}_2$ symmetry, but it turns out that one emerges near the critical point. Roughly speaking, the $\mathbb{Z}_2$ switches the liquid (high-density) and gas (low-density) phases. To make a more precise analogy, we should identify M with the combination of ρ and P that flips sign under the emergent $\mathbb{Z}_2$.

taking values ± 1, with one spin for each site on a cubic lattice[k]

$$s_i \in \{\pm 1\}, \qquad i \in \mathbb{Z}^d. \tag{20}$$

The partition function is a sum over all configurations of spins, with Boltzmann weights that mimic the interactions between physical spins in a magnet,

$$\begin{aligned} Z(K,h) &= \sum_{\{s_i\}} e^{-S[s]}, \\ S[s] &= -K \sum_{\langle ij \rangle} s_i s_j - h \sum_j s_j. \end{aligned} \tag{21}$$

The notation $\langle ij \rangle$ means that i, j are neighboring lattice sites. (You can think of K as the product βJ where β is the inverse temperature and J is the spin-spin interaction.) We can also compute correlation functions by inserting spins into the sum

$$\langle s_{i_1} \cdots s_{i_n} \rangle = \frac{1}{Z} \sum_{\{s_i\}} s_{i_1} \cdots s_{i_n} e^{-S[s]}. \tag{22}$$

The "magnetization" is proportional to the one-point function $\langle s_i \rangle$ (which is independent of i).

When K is positive, the term in $S[s]$ proportional to K makes spins want to align. This term has a $\mathbb{Z}_2$ symmetry under which $s_i \to -s_i$ for all i. The term proportional to h breaks this $\mathbb{Z}_2$ symmetry and causes spins to preferentially have the same sign as h. Both of these effects compete against statistical fluctuations.

In dimension $d \geq 2$, the Ising model famously exhibits a critical point at a special value $K = K_c$ and $h = 0$. For $K > K_c$, spins spontaneously align and break the $\mathbb{Z}_2$ symmetry. For $K < K_c$, statistical fluctuations cause the spins to randomize and the $\mathbb{Z}_2$ is unbroken.

The Ising critical point displays precisely the same long-distance correlation functions and critical exponents as real uni-axial magnets and liquid-vapor systems. In fact, for our purposes, we can think of the Ising model as yet another physical system in the same universality class as these other systems. Its main distinguishing feature is that it is much simpler than e.g. actual water, and thus much easier to study.[l] By studying the mathematics of the Ising model, we will be able to understand gaplessness and a few

[k] The Ising model can also be formulated on other lattices, and in many cases these models lie in the same universality class.

[l] There are many other abstract statistical lattice models in the same universality class as the Ising model, for example compass models `arXiv:1303.5922` and the Blume-Capel

QFT/CFT axioms in a simple way. By the power of critical universality, quantities we compute in the abstract Ising model at $K = K_c$ agree exactly with interesting quantities in real physical systems.

1.4.4. *Continuum ϕ^4 theory*

The Ising lattice model is a good starting point for understanding why classical statistical systems are described by QFT at long distances. We can think of each configuration of spins as a map

$$s : \mathbb{Z}^d \to \{\pm 1\}, \tag{23}$$

and the partition function is a sum over all such maps. This is a discrete version of a path integral.

Typically, in QFT we integrate over continuous functions on a continuous space, e.g.

$$\phi : \mathbb{R}^d \to \mathbb{R}. \tag{24}$$

If we study correlation functions of the Ising model at long distances (much larger than the lattice spacing), it's not hard to imagine that we can approximate the lattice as continuous $i \in \mathbb{Z}^d \to x \in \mathbb{R}^d$. Similarly, we can replace individual spins s_i with average densities of spins $\phi(x)$ in small neighborhoods, so that the effective spin at a point becomes a real number.

This suggests that the partition sum of the Ising model might be related to the path integral for a scalar field on $\mathbb{R}^d$, at least at long distances. That is, perhaps we can take

$$\begin{aligned} \mathbb{Z}^d &\to \mathbb{R}^d \\ s_i &\to \phi(x) \\ \sum_{\{s_i\}} &\to \int D\phi \\ e^{-S[s]} &\to e^{-S[\phi]}, \end{aligned} \tag{25}$$

for some continuum action $S[\phi]$, without changing the long-distance behavior of correlation functions too much.

There is a non-rigorous procedure called "taking the continuum limit" of a lattice model that tries to justify these replacements. However, it involves several uncontrolled approximations, and it is perhaps more honest to just

model `arXiv:1711.10946`. The Ising model on a cubic lattice is arguably the simplest, however the Blume-Capel model has some features that make it even better for simulation.

study a continuum scalar theory and compare its behavior with the Ising model.

For concreteness, let us focus on 3-dimensions. The simplest interacting theory of a scalar field in 3d has Euclidean action

$$S[\phi] = \int d^3x \left(\frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + g\phi^4 \right). \tag{26}$$

Note that ϕ has mass-dimension $1/2$, so that m and g both have mass-dimension 1. Thus, ϕ^2 and ϕ^4 are relevant interactions, and we can think of this theory as an RG flow from the free-boson at high energies to something else at low energies.

Let us map out the phase diagram of this theory. Because m and g are the only dimensionful parameters, only their ratio matters.[m] Suppose first that $m^2/g^2 \gg 0$ is large and positive. In this limit, the theory has a single massive vacuum with mass approximately $m_{\text{gap}} \approx m$. Now suppose $m^2/g^2 \ll 0$ is very negative. In this case, the theory has two very massive vacuua that spontaneously break the $\mathbb{Z}_2$ symmetry, with $m_{\text{gap}} \approx \sqrt{-2m^2}$.

If we start at very large $|m^2|$ (in either phase) and decrease $|m^2|$, m_{gap} decreases too. There exists a critical ratio $m^2/g^2 = r_c$ where $m_{\text{gap}} \to 0$ and the theory is described by a nontrivial CFT. One can justify this claim using the ϵ-expansion and other perturbative techniques, as we will discuss, or from numerical simulations.

The phase diagram we have just described is the same as the Ising phase diagram. Indeed ϕ^4-theory is also in the Ising universality class.

Near the critical point, the ϕ^4 interaction cannot be treated as a small perturbation. More explicitly, because the coupling constant g has mass-dimension 1, perturbation theory is really an expansion in gx, where x is the characteristic length scale of the observable we are computing. In the long-distance limit $x \to \infty$ (where the CFT emerges), this expansion breaks down. Thus, the Ising CFT cannot be studied with traditional perturbation theory. We will later discuss two different perturbative (but uncontrolled) expansions for the Ising CFT: the ϵ-expansion and the large-N expansion.

The theory with action (26) and $m^2/g^2 = r_c$ is equivalent to the free boson at short distances (high energies/UV) and the Ising CFT at long

[m] To be more precise, there is also a UV cutoff Λ that depends on how the theory is regularized. The quartic term gives rise to a linear divergence proportional to the ϕ^2 term, so physical masses are given by $m^2_{\text{phys}} = m^2 + \alpha g\Lambda$, with α a dimensionless number. (In some regularization schemes, like dimensional regularization, the linear divergence vanishes.) However, shifting m^2 is the only way the cutoff can enter a physical observable. Thus, our discussion becomes correct after the replacement $m^2 \to m^2_{\text{phys}}$.

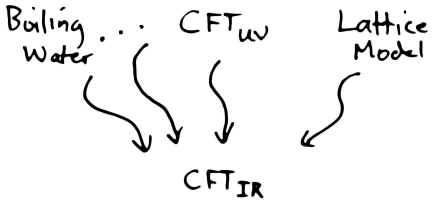

Fig. 1. Many microscopic theories can flow to the same IR CFT. We say that the theories are IR equivalent, or IR dual.

distances (low energies/IR). The free boson is itself a CFT, so we have an RG flow from one CFT to another (figure 1). This construction, where we start with a UV CFT and perturb it so that it flows to an IR CFT, is one possible definition of a general QFT. In this definition, we must allow for the possibility that the IR theory is gapped, in which case $\mathrm{CFT}_{\mathrm{IR}}$ is a TQFT which is technically a special case of a CFT (where all local correlation functions are zero).

An RG flow from the free boson is perhaps the cleanest theoretical construction of the Ising CFT. It shows that the Ising CFT inherits properties of a continuum quantum field theory. For example, ϕ^4-theory has rotational invariance and $\mathbb{Z}_2$ symmetry, so the Ising CFT does too. The Ising lattice model is easier to simulate, but it does not have microscopic rotational symmetry. Meanwhile, water has rotational symmetry but no microscopic $\mathbb{Z}_2$ symmetry. We will learn a lot about the Ising CFT by going back and forth between different microscopic realizations.

1.4.5. *Other universality classes*

We have seen several different systems that fall into the Ising universality class. However, not every critical point is described by the Ising CFT. Another important class of 3d CFTs are the $O(N)$ models. These can be described as the critical point of a theory of N bosons ϕ_i $(i = 1, \ldots, N)$ with a quartic interaction that respects an $O(N)$ global symmetry

$$S = \int d^3x \left(\frac{1}{2} \sum_i \partial_\mu \phi_i \partial^\mu \phi_i + \frac{1}{2} m^2 \sum_i \phi_i \phi_i + g \left(\sum_i \phi_i \phi_i \right)^2 \right). \quad (27)$$

XY magnets are in the same universality class of the $O(2)$ model and Heisenberg magnets are in the same universality class of the $O(3)$ model.

But this is just the tip of the iceberg — there is a huge zoo of different types of critical points that we can construct in a variety of ways. One of our goals will be to better understand this zoo.

2. Path integrals, quantization, and the quantum-classical correspondence

In this section, we explain in more detail why classical statistical systems are described by Euclidean quantum field theory, and also how they are related to quantum condensed-matter systems via Wick rotation. Along the way, we will introduce some important technical concepts in QFT. This section is heavily based on John McGreevy's lectures [].

2.1. *The 1d Ising model and the transfer matrix*

Let us start with the Ising lattice model in 1-dimension. For concreteness, we will study the theory on a periodic lattice with length M, so the spins s_i are labeled by $i \in \mathbb{Z}/M\mathbb{Z}$. The partition function is given by

$$Z_1 = \sum_{\{s_i=\pm 1\}} e^{-S[s]}$$
$$S[s] = -K\sum_{i=1}^{M} s_i s_{i+1} - h\sum_{i=1}^{M} s_i. \tag{28}$$

We refer to $S[s]$ as the "action," even though it is equal to βH, where H is the classical Hamiltonian. This is because we would like to reserve the word Hamiltonian for a completely different object that will appear shortly.

As mentioned in the introduction, the partition sum should be thought of as a discrete version of a 1-dimensional path-integral. This 1-dimensional path integral can be computed by relating it to a 0-dimensional quantum theory. This is an example of the notion of *quantization*.

The key idea is to build up the partition sum by moving along the lattice site-by-site. Forget about periodicity for the moment, and consider the contribution to the partition function from spins $j < i$ for some fixed i,

$$Z_{\text{partial}}(i, s_i) = \sum_{\{s_j : j<i\}} e^{K\sum_{j<i} s_j s_{j+1} + h\sum_{j<i} s_j}. \tag{29}$$

Because of the interaction term $s_{i-1}s_i$, we cannot do the sum over $\{s_1, \ldots, s_{i-1}\}$ without specifying the spin s_i. Thus, we have a function

of s_i. In short, $Z_{\text{partial}}(i, s)$ is the partition function of the theory on the lattice $1 \dots i$, with fixed boundary condition s at site i.

Note that $Z_{\text{partial}}(i+1, s_{i+1})$ can be related to $Z_{\text{partial}}(i, s_i)$ by inserting the remaining Boltzmann weights depending on s_i and performing the sum over $s_i = \pm 1$,

$$Z_{\text{partial}}(i+1, s_{i+1}) = \sum_{s_i = \pm 1} T(s_{i+1}, s_i) Z_{\text{partial}}(i, s_i), \tag{30}$$

where

$$T(s_{i+1}, s_i) \equiv e^{K s_i s_{i+1} + h s_i}. \tag{31}$$

The key step is to recognize (30) as a discrete version of the Schrodinger equation in a 2-dimensional Hilbert space $\mathcal{H}$. This Hilbert space has basis $|s\rangle = |\pm 1\rangle$. The $T(s', s)$'s are elements of a 2×2 matrix $\hat{T}$ acting on $\mathcal{H}$

$$T(s', s) = \langle s' | \hat{T} | s \rangle, \qquad \hat{T} = \begin{pmatrix} e^{K+h} & e^{-K-h} \\ e^{-K+h} & e^{K-h} \end{pmatrix}, \tag{32}$$

and $Z_{\text{partial}}(i, s)$ are the components of a vector $|\Psi_i\rangle \in \mathcal{H}$,

$$Z_{\text{partial}}(i, s) = \langle s | \Psi_i \rangle. \tag{33}$$

In this notation, (30) becomes

$$|\Psi_{i+1}\rangle = \hat{T} |\Psi_i\rangle. \tag{34}$$

The matrix $\hat{T}$ is called the "transfer matrix", and it plays the role of a discrete time-translation operator. Here, i should be thought of as a discrete Euclidean time coordinate.

To be explicit, the (integrated) Schrodinger equation in a quantum theory in Euclidean time is

$$|\Psi(\tau + \Delta\tau)\rangle = e^{-\Delta\tau \hat{H}} |\Psi(\tau)\rangle, \tag{35}$$

where τ is the Euclidean time coordinate, $\Delta\tau$ is some time-step, and $\hat{H}$ is the quantum Hamiltonian. Thus, the 1-dimensional Ising lattice model is precisely equivalent to a 2-state quantum theory with Hamiltonian

$$\hat{H} = -\frac{1}{\Delta\tau} \log \hat{T}. \tag{36}$$

When the lattice is periodic, the partition function is related to the transfer matrix by

$$\begin{aligned} Z_1 &= \sum_{\{s_i\}} \langle s_M | \hat{T} | s_{M-1} \rangle \langle s_{M-1} | \hat{T} | s_{M-2} \rangle \cdots \langle s_1 | \hat{T} | s_M \rangle \\ &= \operatorname{Tr}(\hat{T}^M). \end{aligned} \tag{37}$$

This is now easy to evaluate by diagonalizing $\hat{T}$,

$$\mathrm{Tr}(\hat{T}^M) = \lambda_+^M + \lambda_-^M, \tag{38}$$

where

$$\begin{aligned}\lambda_\pm &= e^K \cosh h \pm \sqrt{e^{2K}\sinh^2 h + e^{-2K}} \\ &\to \begin{cases} 2\cosh K \\ 2\sinh K \end{cases} \qquad (\text{when } h = 0).\end{aligned} \tag{39}$$

In the thermodynamic limit $M \to \infty$, the partition function is dominated by the larger eigenvalue of $\hat{T}$

$$Z_1 = \lambda_+^M \left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^M\right) \approx \lambda_+^M. \tag{40}$$

In quantum mechanical language, the state with the largest eigenvalue of $\hat{T}$ has the smallest eigenvalue of $\hat{H}$ — i.e. it is the ground state, and we should call it $|0\rangle$. We have shown that the ground state dominates the thermodynamic limit, and contributions from the excited state are exponentially suppressed in the energy gap

$$m_{\mathrm{gap}} \equiv -\frac{1}{\Delta\tau}\log(\lambda_-/\lambda_+). \tag{41}$$

We can also use the transfer matrix to compute correlation functions. For example, consider the two-point function $\langle s_{i_1} s_{i_2}\rangle$ and suppose $i_1 > i_2$. We have

$$\begin{aligned}\langle s_{i_1} s_{i_2}\rangle &= \frac{1}{Z_1}\sum_{\{s_i\}} \langle s_M|\hat{T}|s_{M-1}\rangle \cdots \langle s_{i_1+1}|\hat{T}|s_{i_1}\rangle s_{i_1} \langle s_{i_1}|\hat{T}|s_{i_1-1}\rangle \cdots \\ &\qquad \times \cdots \langle s_{i_2+1}|\hat{T}|s_{i_2}\rangle s_{i_2}\langle s_{i_2}|\hat{T}|s_{i_2-1}\rangle \cdots \langle s_1|\hat{T}|s_M\rangle \\ &= \frac{1}{Z_1}\mathrm{Tr}(\hat{T}^{M-i_1}\hat{s}\hat{T}^{i_1-i_2}\hat{s}\hat{T}^{i_2}) \qquad (i_1 > i_2).\end{aligned} \tag{42}$$

Here, we introduced the spin operator $\hat{s}$, which has matrix elements

$$\langle s|\hat{s}|s'\rangle = s\delta_{ss'}. \tag{43}$$

It is easy to compute the correlation function (42) by expressing $\hat{s}$ in the eigenbasis of $\hat{T}$.

Exercise 2.1. *Show that in the limit of large M and large "distance" $i_1 - i_2$, the correlator factorizes into a product of expectation values $\langle 0|\hat{s}|0\rangle$, plus exponential corrections from the excited state*

$$\langle s_{i_1} s_{i_2}\rangle = \langle 0|\hat{s}|0\rangle^2 + O(e^{-m_{\mathrm{gap}}\tau}, e^{-m_{\mathrm{gap}}(L-\tau)}), \tag{44}$$

where $\tau \equiv (i_1 - i_2)\Delta\tau$, $L \equiv M\Delta\tau$.

Let us write (42) in a slightly different way by introducing "Heisenberg picture" operators

$$\hat{s}(i) = \hat{T}^{-i}\hat{s}\hat{T}^{i}. \tag{45}$$

Equation (42) is equivalent to

$$\langle s_{i_1} s_{i_2} \rangle = \frac{1}{Z_1}\mathrm{Tr}(\hat{s}(i_1)\hat{s}(i_2)\hat{T}^M) \qquad (i_1 > i_2). \tag{46}$$

Note that in deriving (42, 46), we used that $i_1 > i_2$. If instead $i_2 > i_1$, the path integral would give a product of operators in the opposite order

$$\langle s_{i_1} s_{i_2} \rangle = \frac{1}{Z_1}\mathrm{Tr}(\hat{s}(i_2)\hat{s}(i_1)\hat{T}^M) \qquad (i_2 > i_1). \tag{47}$$

The general statement is that the path integral becomes a time-ordered product of quantum operators[n]

$$\langle s_{i_1} \cdots s_{i_n} \rangle = \mathrm{Tr}(T\{\hat{s}(i_1)\cdots\hat{s}(i_n)\}\hat{T}^M), \tag{48}$$

where the definition of the time-ordering symbol is

$$T\{\hat{s}(i_1)\cdots\hat{s}(i_n)\} \equiv \hat{s}(i_1)\cdots\hat{s}(i_n)\theta(i_1 > \cdots > i_n) + \text{permutations}. \tag{49}$$

Here $\theta(i_1 > \cdots > i_n)$ is 1 if the i_k are in the specified order and zero otherwise.

2.2. *Quantization in quantum mechanics*

The procedure of turning a path integral into a product of quantum operators is called quantization. It is an extremely general procedure that ultimately stems from "cutting and gluing" rules of the path integral.

For a more familiar example, let us review quantization for the path integral of a quantum mechanical particle. This theory has Euclidean action

$$S[x] = \int d\tau \left(\frac{1}{2}\dot{x}^2 + V(x)\right), \tag{50}$$

where $x(\tau)$ is a map from Euclidean time τ to $\mathbb{R}$. Consider the path integral on the interval $\tau \in [\tau_a, \tau_b]$ with fixed boundary conditions

$$U(x_b, \tau_b; x_a, \tau_a) \equiv \int_{\substack{x(\tau_b)=x_b \\ x(\tau_a)=x_a}} Dx(\tau) e^{-S[x]}. \tag{51}$$

[n] We hope that the time-ordering symbol $T\{\cdots\}$ will not be confused with the transfer matrix.

The integrals $U(\cdots)$ satisfy simple cutting and gluing rules that come from grouping paths according to their positions at a fixed time $\tau_c \in (\tau_a, \tau_b)$,

$$U(x_b, \tau_b; x_a, \tau_a) = \int_{-\infty}^{\infty} dx_c U(x_b, \tau_b; x_c, \tau_c) U(x_c, \tau_c; x_a, \tau_a). \tag{52}$$

To quantize the theory, we build up the path integral in small time-increments using the gluing rule. Suppose we have already computed the path integral $U(x_0, \tau_0; x_a, \tau_a)$ on the interval $[\tau_a, \tau_0]$. To extend $\tau_0 \to \tau_0 + \epsilon$, we have

$$U(x_1, \tau_0 + \epsilon; x_a, \tau_a) = \int dx_0 \, U(x', \tau_0 + \epsilon; x_0, \tau_0) U(x_0, \tau_0; x_a, \tau_a). \tag{53}$$

The quantity

$$U(x_1, \tau_0 + \epsilon; x_0, \tau_0) = U(x_1, \epsilon; x_0, 0). \tag{54}$$

plays the role of a transfer matrix. To see this, introduce an orthonormal basis vector $|x\rangle$ for each $x \in \mathbb{R}$. The $|x\rangle$ are analogs of the basis states $|s = \pm 1\rangle$ in the Ising case. Writing

$$\begin{aligned} \langle x_1 | \hat{T}_\epsilon | x_0 \rangle &= U(x_1, \epsilon; x_0, 0), \\ \langle x_0 | \Psi(\tau_0) \rangle &= U(x_0, \tau_0; x_a, \tau_a), \end{aligned} \tag{55}$$

equation (53) becomes

$$|\Psi(\tau_0 + \epsilon)\rangle = \hat{T}_\epsilon |\Psi(\tau_0)\rangle. \tag{56}$$

To recover the Schrodinger equation, it remains to show that $\hat{T}_\epsilon = 1 - \epsilon \hat{H} + O(\epsilon^2)$, where

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x) \tag{57}$$

is the usual quantum-mechanical Hamiltonian. This is a standard argument, and we give a quick version here for completeness. We have

$$\begin{aligned} \langle x_1 | \hat{T} | x_0 \rangle &= \int_{\substack{x(\epsilon) = x_1 \\ x(0) = x_0}} Dx(\tau) e^{-S[x]}, \\ S[x] &= \int_0^\epsilon d\tau \left(\frac{1}{2} \dot{x}^2 + V(x) \right). \end{aligned} \tag{58}$$

Because the time interval is so short, if the distance traveled is larger than $O(\epsilon)$, the amplitude will be highly suppressed. Thus, let us assume $|x(\tau) -$

$x_0| = O(\epsilon)$. This means we can replace $V(x) \to V(x_0)$ (up to subleading corrections in ϵ) and the potential factors out of the path integral

$$e^{-S[x]} = e^{-\epsilon V(x_0)} \exp\left(\int_0^\epsilon d\tau \frac{\dot{x}^2}{2}\right). \tag{59}$$

We can now split $x(\tau)$ into a classical term and a fluctuation term $\xi(\tau)$ with boundary conditions $\xi(0) = \xi(\epsilon) = 0$,

$$\begin{aligned} x(\tau) &= x_0\left(1 - \frac{\tau}{\epsilon}\right) + x_1\frac{\tau}{\epsilon} + \xi(\tau), \\ \int_0^\epsilon d\tau \frac{\dot{x}^2}{2} &= \frac{(x_1 - x_0)}{2\epsilon} + \int_0^\epsilon d\tau \frac{\dot{\xi}^2}{2}. \end{aligned} \tag{60}$$

The path integral over ξ contributes a constant A_ϵ that depends on how the theory is regulated.[o] We thus find

$$\langle x'|\hat{T}|x_0\rangle = A_\epsilon e^{-\epsilon V(x_0)} e^{-\frac{(x_1-x_0)^2}{2\epsilon}}(1 + O(\epsilon)) \tag{61}$$

Finally, the Gaussian factor can be expanded in ϵ using

$$e^{-\frac{x^2}{2\epsilon}} = \sqrt{2\pi\epsilon}\left(\delta(x) + \frac{\epsilon}{2}\delta''(x) + O(\epsilon^2)\right), \tag{62}$$

to get

$$\begin{aligned} \langle x_1|\hat{T}|x_0\rangle &= A_\epsilon\sqrt{2\pi\epsilon}\left(\delta(x_1 - x_0) - \epsilon\left(-\delta''(x_1 - x_0) + V(x_0)\right) + O(\epsilon^2)\right) \\ &= A_\epsilon\sqrt{2\pi\epsilon}\langle x_1|1 - \epsilon\hat{H}|x_0\rangle \end{aligned} \tag{63}$$

The prefactor $A_\epsilon\sqrt{2\pi\epsilon}$ is an overall regulator-dependent constant that must be renormalized by adding a "cosmological constant" term to the action.

2.3. *The 2d Ising model*

[o] A very simple regulator is to approximate $x(\tau)$ as a piecewise linear path over segments of length ϵ. This is equivalent to simply setting $\xi = 0$ and not doing the integral.