# 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 23, 2018. In a few places, there are reminders in blue to the author to add additional discussion/comments. There may also be empty citations and missing figures.

## Contents

# 1. Introduction

### *Resources*

These introductory notes are heavily based on Silviu Pufu's Bootstrap 2017 lectures [], John Cardy's book *Scaling and Renormalization in Statistical Physics* 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

[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).

is built into the microscopic theory. However in lattice systems, rotational 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

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

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

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

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

To understand why $m_{\text{gap}}$ 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

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

[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.

Here, $H$ is the Hamiltonian with the vacuum energy subtracted off, and $\psi$ runs over an orthonormal basis of eigenstates of $H$. To get the right-hand 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 $\psi$ with the smallest nonzero eigenvalue of $H$, which is $m_{\text{gap}}$.[f]

Thus, when $m_{\text{gap}}$ 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 $m_{\text{gap}}$ much smaller than $\Lambda_{UV}$, and even arrange for $m_{\text{gap}}$ 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 $\Delta$ are constants depending on $\phi$. The quantity $\Delta$ is called the scaling dimension of $\phi$.

### 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

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

[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 $m_{\text{gap}}$ 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.

Under very general conditions (that we will discuss), critical points also display less obvious emergent symmetries called conformal transformations. 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 $m_{\mathrm{gap}}$ 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.

[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.

### 1.4.1. *Magnets*

Our first examples of critical points occur in magnets. The discussion in this section borrows heavily from Cardy []. Given a magnet with temperature $T$, we can apply a magnetic field $\vec{H}$ and measure the 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"

[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.

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

- $\alpha$: 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.)
- $\beta$: the spontaneous magnetization behaves as
$$\lim_{H \to 0^+} M \propto (-t)^{\beta}. \tag{11}$$
- $\gamma$: the zero-field susceptibility behaves as
$$\chi = \left.\frac{\partial M}{\partial H}\right|_{H=0} \propto |t|^{-\gamma}. \tag{12}$$
- $\delta$: the magnetization at $T = T_c$ behaves as
$$|M| \propto |h|^{1/\delta}. \tag{13}$$
- $\nu$ and $\eta$: the correlation length $\xi$ 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:

**Claim 1.** *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 $\rho$ is the density and $\rho_c$ 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,

**Claim 2.** *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.

*End of lecture 1*

[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 $\rho$ and $P$ that flips sign under the emergent $\mathbb{Z}_2$.

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 taking values $\pm 1$, 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 $\beta J$ where $\beta$ 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.

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

actual water, and thus much easier to study.[1] By studying the mathematics of the Ising model, we will be able to understand gaplessness and a few 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 $\phi^4$ 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.

[1] 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 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.

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 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 $\phi$ has mass-dimension 1/2, so that $m$ and $g$ both have mass-dimension 1. Thus, $\phi^2$ and $\phi^4$ 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 $|m^2|$, $m_{\text{gap}}$ 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 $\epsilon$-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 $\phi^4$-theory is also in the Ising universality class.

Near the critical point, the $\phi^4$ 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)

[m] To be more precise, there is also a UV cutoff $\Lambda$ that depends on how the theory is regularized. The quartic term gives rise to a linear divergence proportional to the $\phi^2$ term, so physical masses are given by $m^2_{\text{phys}} = m^2 + \alpha g\Lambda$, with $\alpha$ 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 $m^2 \to m^2_{\text{phys}}$.

expansions for the Ising CFT: the $\epsilon$-expansion and the large-$N$ expansion.

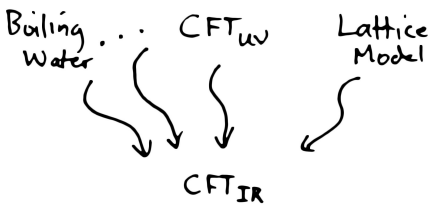


Fig. 1. Many microscopic theories can flow to the same IR CFT. We say that the theories are IR equivalent, or IR dual.

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 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 $\text{CFT}_{\text{IR}}$ 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, $\phi^4$-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 $\phi_i$ $(i = 1, \ldots, N)$

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.

Discuss non-uniqueness of UV Lagrangian description. An example is particle-vortex duality between the gauged $O(2)$ model and the $O(2)$ model. Distinguish between IR duality and exact duality.

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, like cutting-and-gluing rules for path integrals and the idea of quantization. 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 $i \in \mathbb{Z}_M$. The partition function is given by

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

We refer to $S[s]$ as the "action," even though it is equal to $\beta H$, 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 $1 \ldots i$, with fixed boundary condition $s$ at site $i$.[n]

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 that depend 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 $2 \times 2$ matrix $\widehat{T}$ acting on $\mathcal{H}$

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

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

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

The matrix $\widehat{T}$ 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.

[n] We must also impose some boundary condition at site 1. The precise choice is not important for this discussion, so we have left it implicit.

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

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

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

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

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

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

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

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

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

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

In quantum mechanical language, the state with the largest eigenvalue of $\widehat{T}$ has the smallest eigenvalue of $\widehat{H}$ — 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. Contributions from the excited state are exponentially suppressed in the energy gap times the size of the system

$$\begin{aligned} \left( \frac{\lambda_-}{\lambda_+} \right)^M &= e^{-(M\Delta\tau) m_{\mathrm{gap}}}, \\ m_{\mathrm{gap}} &\equiv -\frac{1}{\Delta\tau} \log(\lambda_-/\lambda_+). \end{aligned} \tag{41}$$

*End of lecture 2*

We can also use the transfer matrix to compute correlation functions. For example, consider the two-point function $\langle s_{i_1} s_{i_2}\rangle_{\mathbb{Z}_M}$ where the subscript $M$ indicates that we are on a periodic lattice with length $M$. Suppose $i_1 > i_2$. We have

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

Here, we introduced the Pauli spin operator $\sigma^z$ that measures the spin of a state

$$\sigma^z|s\rangle = s|s\rangle. \tag{43}$$

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

**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 $\langle 0|\sigma^z|0\rangle$, plus exponential corrections from the excited state*

$$\langle s_{i_1} s_{i_2}\rangle_{\mathbb{Z}_M} = \langle 0|\sigma^z|0\rangle^2 + O(e^{-m_{\mathrm{gap}}\tau}, e^{-m_{\mathrm{gap}}(L-\tau)}), \tag{44}$$

*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

$$\widehat{s}(i) = \widehat{T}^{-i}\sigma^z\widehat{T}^i. \tag{45}$$

Equation (42) is equivalent to

$$\langle s_{i_1} s_{i_2}\rangle_{\mathbb{Z}_M} = \frac{1}{Z_1}\mathrm{Tr}(\widehat{s}(i_1)\widehat{s}(i_2)\widehat{T}^M) \qquad (i_1 > i_2). \tag{46}$$

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

$$\langle s_{i_1} s_{i_2}\rangle_{\mathbb{Z}_M} = \frac{1}{Z_1}\mathrm{Tr}(\widehat{s}(i_2)\widehat{s}(i_1)\widehat{T}^M) \qquad (i_2 > i_1). \tag{47}$$

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

$$\langle s_{i_1} \cdots s_{i_n} \rangle_{\mathbb{Z}_M} = \frac{1}{Z_1} \mathrm{Tr}(T\{\widehat{s}(i_1) \cdots \widehat{s}(i_n)\} \widehat{T}^M), \tag{48}$$

where the definition of the time-ordering symbol is

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

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

Finally, in the limit of large-$M$, the factor $\frac{\widehat{T}^M}{\lambda_+^M}$ projects onto the ground state, so a path integral on the full lattice $\mathbb{Z}$ becomes a vacuum expectation value of a time-ordered product

$$\langle s_{i_1} \cdots s_{i_n} \rangle_{\mathbb{Z}} = \langle 0 | T\{\widehat{s}(i_1) \cdots \widehat{s}(i_n)\} | 0 \rangle. \tag{50}$$

**Exercise 2.2.** *Consider adding a next-to-nearest neighbor interaction to the 1d Ising model*

$$S[s] = \sum_i K s_i s_{i+1} + K' s_i s_{i+2}. \tag{51}$$

*Find the associated quantum Hilbert space and transfer matrix, and compute the partition function on $\mathbb{Z}_M$ in the limit of large $M$.*

### 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 particle on the real line. This theory has Euclidean action

$$S[x] = \int d\tau \left( \frac{1}{2} \dot{x}^2 + V(x) \right), \tag{52}$$

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

$$U(x_b, \tau_b; x_a, \tau_a) \equiv \int_{\substack{x(\tau_b)=x_b \\ x(\tau_a)=x_a}} Dx(\tau) e^{-S[x]}. \tag{53}$$

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

By grouping paths according to their positions at a fixed time $\tau_c \in (\tau_a, \tau_b)$, we obtain a simple "gluing" rule for $U$,

$$U(x_b, \tau_b; x_a, \tau_a) = \int_{-\infty}^{\infty} dx_c U(x_b, \tau_b; x_c, \tau_c) U(x_c, \tau_c; x_a, \tau_a). \tag{54}$$

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 write

$$U(x_1, \tau_0 + \epsilon; x_a, \tau_a) = \int dx_0 \, U(x', \tau_0 + \epsilon; x_0, \tau_0) U(x_0, \tau_0; x_a, \tau_a). \tag{55}$$

The quantity

$$U(x_1, \tau_0 + \epsilon; x_0, \tau_0) = U(x_1, \epsilon; x_0, 0). \tag{56}$$

plays the role of a transfer matrix, and $U(x_0, \tau_0; x_a, \tau_a)$ plays the role of a state at time $\tau_0$. To see this, we introduce a Hilbert space

$$\mathcal{H} = \text{Span}\{|x\rangle : x \in \mathbb{R}\} \tag{57}$$

with inner product $\langle x'|x\rangle = \delta(x - x')$. The $|x\rangle$ are analogs of the basis states $|s = \pm 1\rangle$ in the Ising model. Defining $\widehat{T}_\epsilon$ and $|\Psi(\tau_0)\rangle$ by

$$\begin{aligned} \langle x_1|\widehat{T}_\epsilon|x_0\rangle &= U(x_1, \epsilon; x_0, 0), \\ \langle x_0|\Psi(\tau_0)\rangle &= U(x_0, \tau_0; x_a, \tau_a), \end{aligned} \tag{58}$$

equation (55) becomes

$$|\Psi(\tau_0 + \epsilon)\rangle = \widehat{T}_\epsilon |\Psi(\tau_0)\rangle. \tag{59}$$

To recover the Schrodinger equation, we must show that

$$\widehat{T}_\epsilon = 1 - \epsilon \widehat{H} + O(\epsilon^2), \tag{60}$$

where

$$\widehat{H} = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + V(x) \tag{61}$$

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

$$\begin{aligned} \langle x_1|\widehat{T}_\epsilon|x_0\rangle &= \int_{\substack{x(\epsilon)=x_1 \\ x(0)=x_0}} Dx(\tau) e^{-S[x]}, \qquad (\epsilon \ll 1) \\ S[x] &= \int_0^\epsilon d\tau \left(\frac{1}{2}\dot{x}^2 + V(x)\right). \end{aligned} \tag{62}$$

Because the time interval is so short, if $|x(\tau)-x_0|$ ever becomes larger than $O(\epsilon)$, the kinetic term will cause the amplitude to be highly suppressed. Thus, let us assume $|x(\tau)-x_0|$ is of order $\epsilon$. This means we can replace $V(x)\to V(x_0)$ (up to subleading corrections in $\epsilon$), so the potential factors out of the integrand

$$e^{-S[x]} = e^{-\epsilon V(x_0)} \exp\left(\int_0^\epsilon d\tau \frac{\dot{x}^2}{2}\right). \tag{63}$$

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

$$\begin{aligned} x(\tau) &= x_0\left(1-\frac{\tau}{\epsilon}\right) + x_1\frac{\tau}{\epsilon} + \zeta(\tau), \\ \int_0^\epsilon d\tau\frac{\dot{x}^2}{2} &= \frac{(x_1-x_0)}{2\epsilon} + \int_0^\epsilon d\tau\frac{\dot{\zeta}^2}{2}. \end{aligned} \tag{64}$$

The path integral over $\zeta$ contributes a constant $A_\epsilon$ that depends on how the theory is regulated.[P] We thus find

$$\langle x_1|\widehat{T}|x_0\rangle = A_\epsilon e^{-\epsilon V(x_0)} e^{-\frac{(x_1-x_0)^2}{2\epsilon}}(1+O(\epsilon)). \tag{65}$$

Finally, the Gaussian factor can be expanded in $\epsilon$ using

$$e^{-\frac{x^2}{2\epsilon}} = \sqrt{2\pi\epsilon}\left(\delta(x)+\frac{\epsilon}{2}\delta''(x)+O(\epsilon^2)\right), \tag{66}$$

which gives

$$\begin{aligned} \langle x_1|\widehat{T}|x_0\rangle &= A_\epsilon\sqrt{2\pi\epsilon}\left(\delta(x_1-x_0)-\epsilon\left(-\frac{1}{2}\delta''(x_1-x_0)+V(x_0)\delta(x_1-x_0)\right)+O(\epsilon^2)\right) \\ &= A_\epsilon\sqrt{2\pi\epsilon}\langle x_1|1-\epsilon\widehat{H}+O(\epsilon^2)|x_0\rangle. \end{aligned} \tag{67}$$

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

Thus, in the limit $\epsilon\to 0$, equation (59) becomes a Euclidean version of the Schrodinger equation

$$\frac{d}{d\tau}|\Psi(\tau)\rangle = -\widehat{H}|\Psi(\tau)\rangle. \tag{68}$$

Integrating using the initial condition $U(x',0,x,0)=\delta(x-x')$, we find

$$U(x_b,\tau_b;x_a,\tau_a) = \langle x_b|e^{-\widehat{H}(\tau_b-\tau_a)}|x_b\rangle. \tag{69}$$

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

Path integrals on different geometries correspond to different quantum mechanical observables. For example, the path integral on a circle $S^1_\beta$ of length $\beta$ with periodic boundary conditions is equal to

$$\int_{-\infty}^{\infty} dx\, U(x,\beta;x,0) = \mathrm{Tr}(e^{-\beta\widehat{H}}), \tag{70}$$

which is the quantum partition function at inverse temperature $\beta$. This is the continuum analog of equation (37). Inserting observables into the path integral gives a time-ordered product of the associated quantum operators. For example, with periodic boundary conditions,

$$\begin{aligned}\langle x(\tau_1)\cdots x(\tau_n)\rangle_{S^1_\beta} &= \int_{x(\beta)=x(0)} Dx\, x(\tau_1)\cdots x(\tau_n)e^{-S[x]} \\ &= \mathrm{Tr}(T\{\widehat{x}(\tau_1)\cdots\widehat{x}(\tau_n)\}e^{-\beta\widehat{H}}),\end{aligned} \tag{71}$$

where we introduced the operator

$$\begin{aligned}\widehat{x}(\tau) &= e^{\tau\widehat{H}}\widehat{x}e^{-\tau\widehat{H}}, \\ \widehat{x}|x\rangle &= x|x\rangle.\end{aligned} \tag{72}$$

To derive (71), we cut the interval $[0,\beta]$ into segments $[0,\tau_1]\cup[\tau_1,\tau_2]\cup\cdots[\tau_n,\beta]$ (assuming $\tau_n > \cdots > \tau_1$) and use (69) for each segment. The operator $\widehat{x}$ is the analog of $\widehat{s}$ and equation (71) is the analog of (48).

Finally, in the limit $\beta\to\infty$, the factor $e^{-\beta\widehat{H}}$ projects onto the ground state. Thus the Euclidean path integral on $\mathbb{R}$ is equal to a time-ordered correlation function in the ground state,

$$\langle x(\tau_1)\cdots x(\tau_n)\rangle_{\mathbb{R}} = \langle 0|T\{\widehat{x}(\tau_1)\cdots\widehat{x}(\tau_n)\}|0\rangle. \tag{73}$$

This is the continuum analog of (50).

*End of lecture 3*

### 2.3. *The 2d Ising model*

Let us now consider a slightly more complicated case: the 2d Ising model. This will be a good toy example for how cutting, gluing, and quantization work in a general QFT. For simplicity, we set $h = 0$. We consider the partition function on the doubly-periodic lattice $\mathbb{Z}_M\times\mathbb{Z}_N$ and label spins $s_{i,j}$ by a pair $(i,j)\in\mathbb{Z}_M\times\mathbb{Z}_N$.

The action is given by

$$\begin{aligned} S[s] &= -K\sum_{i,j}(s_{i,j}s_{i+1,j} + s_{i,j}s_{i,j+1}) \\ &= K\sum_{i,j}\left(\frac{1}{2}(s_{i,j+1} - s_{i,j})^2 - 1\right) - K\sum_{i,j} s_{i,j}s_{i+1,j} \\ &= \text{const.} + \sum_{j=1}^{N} L(\mathbf{s}_{j+1}, \mathbf{s}_j). \end{aligned} \tag{74}$$

In the last line, we split the action into contributions from pairs of neighboring rows. The notation $s_j$ represents the configuration of spins in the $j$-th row,

$$(\mathbf{s}_j)_i = s_{i,j}. \tag{75}$$

The action associated with a pair of neighboring rows is given by

$$L(\mathbf{s}', \mathbf{s}) = \frac{1}{2}K\sum_{i=1}^{M}(\mathbf{s}'_i - \mathbf{s}_i)^2 - \frac{1}{2}K\sum_{i=1}^{M}(\mathbf{s}_{i+1}\mathbf{s}_i + \mathbf{s}'_{i+1}\mathbf{s}'_i). \tag{76}$$

(The constant in (74) gives an unimportant multiplicative constant in the path integral that will disappear in normalized correlation functions, so we will ignore it.)

To quantize the theory, we can think of the $j$ direction as time, so $\mathbf{s}_j$ is interpreted as a classical configuration on a fixed time-slice. The Hilbert space has an orthonormal basis vector for each such configuration,

$$\begin{aligned} \mathcal{H}_M &= \text{Span}\,\{|\pm 1, \pm 1, \cdots, \pm 1\rangle\} \\ &= \bigotimes_{i=1}^{M} \mathcal{H}_i, \end{aligned} \tag{77}$$

where $\mathcal{H}_i$ is a 1-qubit Hilbert space for each site $i$. $\mathcal{H}_M$ is the quantum Hilbert space of $M$ qubits, and is $2^M$-dimensional.

The transfer matrix between successive time slices is a $2^M \times 2^M$ matrix with entries

$$\langle \mathbf{s}'|\widehat{T}|\mathbf{s}\rangle = e^{-L(\mathbf{s}',\mathbf{s})}. \tag{78}$$

The partition function on $\mathbb{Z}_M \times \mathbb{Z}_N$ is then

$$Z(\mathbb{Z}_M \times \mathbb{Z}_N) = \text{Tr}_{\mathcal{H}_M}(\widehat{T}^N). \tag{79}$$

To compute correlation functions, we need an operator that measures the spin at site $i$. This is simply the Pauli spin matrix $\sigma_i^z$ associated with the $i$-th site

$$\sigma_i^z|s_1,\dots,s_j,\dots,s_M\rangle = s_i|s_1,\dots,s_i,\dots,s_M\rangle. \tag{80}$$

We also have Heisenberg picture operators

$$\widehat{s}_{i,j} = \widehat{T}^{-j}\sigma_i^z\widehat{T}^j. \tag{81}$$

Correlation functions become traces of time-ordered products, e.g.

$$\begin{aligned}\langle s_{i_1,j_1}s_{i_2,j_2}\rangle &= \mathrm{Tr}_{\mathcal{H}_M}(\widehat{T}^{N+j_2-j_1}\sigma_{i_1}^z\widehat{T}_2^{j_1-j_2}\sigma_{i_2}^z)\theta(j_1-j_2)+(1\leftrightarrow 2)\\ &= \mathrm{Tr}_{\mathcal{H}_M}(T\{\widehat{s}_{i_1,j_1}\widehat{s}_{i_2,j_2}\}\widehat{T}^N).\end{aligned} \tag{82}$$

Let us now emphasize an important new ingredient in the 2-dimensional case compared to the 1-dimensional case. To arrive at (79), we had to choose $j$ as the time direction. We then cut the path integral along rows of constant $j$. However, we could just as well have chosen $i$ as the time direction and cut the path integral along columns of constant $i$. This would give a different Hilbert space $\mathcal{H}_N$ with dimension $2^N$, a new transfer matrix $\widehat{T}'$ (acting on $\mathcal{H}_N$), and a different formula for the same path integral,

$$Z(\mathbb{Z}_M\times\mathbb{Z}_N) = \mathrm{Tr}_{\mathcal{H}_N}(\widehat{T}'^M) = \mathrm{Tr}_{\mathcal{H}_M}(\widehat{T}^N). \tag{83}$$

In this new quantization, an insertion of $s_{i,j}$ in the path integral becomes an insertion of

$$s_{i,j}\to\widehat{T}'^{-i}\sigma_j^z\widehat{T}'^i \tag{84}$$

in a time-ordered product of quantum operators. Let us emphasize that the operators (81) and (84) truly are different, even though they represent the same path integral variable. They even act on different-dimensional Hilbert spaces ($2^M$ vs. $2^N$)! Furthermore, even the notion of "time"-ordering is different in this different quantization: in one case, operators are ordered according to their $i$ coordinates, and in the other they are ordered according to their $j$ coordinates.

**Exercise 2.3.** *Show how to quantize the 2d Ising model in yet another way, using $i+j$ as the time coordinate and $i-j$ as space coordinate. What is the Hilbert space? What is the transfer matrix?*

### 2.4. *Atiyah-Segal axioms*

We are now ready to understand some axioms of continuum QFT. We will simultaneously state them and give examples in a Lagrangian theory of a real scalar field $\phi$. A version of these axioms in TQFTs is due to Atiyah and a version in 2d CFTs is due to Segal. We will be interested in QFTs that depend on the geometry of the space in which they live, so when we say "manifold" below, we mean "Riemannian manifold with metric."

Consider a $d$-dimensional QFT $Q$.

(1) To every $(d-1)$-manifold $N$ without boundary, $Q$ assigns a Hilbert space $\mathcal{H}_N$ called the space of states on $N$.
For example, in theories with an explicit path integral over some fields ("Lagrangian theories"), $\mathcal{H}_N$ has an orthonormal basis state for every field configuration on $N$. In scalar field theory,

$$\mathcal{H}_N = \mathrm{Span}\{|\phi_b\rangle : \phi_b \in C(N,\mathbb{R})\}. \tag{85}$$

with inner product $\langle\phi_{b1}|\phi_{b2}\rangle = \delta(\phi_{b1}-\phi_{b2})$ (a functional $\delta$-function). Here $C(N,\mathbb{R})$ is a space of real-valued functions on $N$.[q] More formally, we can write

$$\mathcal{H}_N = L^2(C(N,\mathbb{R}),\mathbb{C}), \tag{86}$$

where $L^2(X)$ denotes the square-integrable complex functions on $X$. The space $\mathcal{H}_N$ is the analog of the Hilbert spaces (57) and (77) for quantum mechanics and the 2d Ising model. Let us discuss the example of quantum mechanics more explicitly. Quantum mechanics is a 1-dimensional QFT. In this case $d-1=0$, and the only possible 0-manifolds are disjoint unions of points. The space of field configurations on a point is $C(\bullet,\mathbb{R})=\mathbb{R}$, so the Hilbert space associated to a point is $\mathcal{H}_\bullet = L^2(C(\bullet,\mathbb{R}),\mathbb{C}) = L^2(\mathbb{R},\mathbb{C})$. This is spanned by the usual basis states $|x\rangle$ with inner product $\delta(x-x')$.[r]

(2) The space of states on a disconnected union is a tensor product

$$\mathcal{H}_{N_1\sqcup N_2} = \mathcal{H}_{N_1}\otimes\mathcal{H}_{N_2}. \tag{87}$$

[q] We are being deliberately vague about *which* space of functions on $N$, e.g. $C^1(N)$, $C^\infty(N)$, etc.. In reality, the path integral is defined by introducing a regulator and taking a limit as the regulator is removed. Different regulators will result in different spaces of functions, and we prefer to be agnostic about which regulator is being used.

[r] Note in particular that $L^2(C(\bullet,\mathbb{R}),\mathbb{C})$ is totally different from $L^2(\bullet,\mathbb{C})=\mathbb{C}$. In general, don't confuse the Hilbert space $\mathcal{H}_N$ with the space of square-integrable functions $L^2(N,\mathbb{C})$.

This should be clear in scalar theory from the definition (85). It will be useful to assign a Hilbert space to the empty manifold, and the only reasonable choice consistent with (87) is

$$\mathcal{H}_\varnothing = \mathbb{C}. \tag{88}$$

(3) To every $d$-manifold $M$ with incoming boundary $N$ and outgoing boundary $N'$, $Q$ assigns a transition amplitude[s]

$$Z_M : \mathcal{H}_N \to \mathcal{H}_{N'}. \tag{89}$$

In Lagrangian theories, $Z_M$ is obtained by fixing boundary conditions on $N, N'$ and performing the path integral over the interior. For example in a scalar theory, the matrix elements of $Z_M$ are given by

$$\langle \phi_b' | Z_M | \phi_b \rangle = \int_{\substack{\phi|_N = \phi_b \\ \phi|_{N'} = \phi_b'}} D\phi(x) e^{-S[\phi]}. \tag{90}$$

Here, the notation $\phi|_N$ means the restriction of $\phi$ to $N$.

In the special case where $M$ is a closed manifold (without boundary), $Z_M$ is a linear map

$$Z_M : \mathbb{C} \to \mathbb{C}. \tag{91}$$

A linear map $\mathbb{C} \to \mathbb{C}$ is simply multiplication by complex number. That number is the "partition function" of the QFT on $M$.

Another important case is where $M$ has boundaries $(\varnothing, N)$. (The first element of the ordered pair is the incoming boundary and the second is the outgoing boundary.) In this case, $Z_M$ gives a linear map

$$Z_M : \mathbb{C} \to \mathcal{H}_N. \tag{92}$$

Any such map is uniquely determined by its action on $1 \in \mathbb{C}$, so this is equivalent to a state $|\Psi\rangle = Z_M(1) \in \mathcal{H}_N$. Thus, we can "prepare a state" by performing the path integral with fixed boundary conditions. We saw several examples in the previous sections, for example the definition of $Z_{\text{partial}}(i, s)$ in the 1d Ising model.

Finally, consider the case where $M$ has boundaries $(N, \varnothing)$. Then we have a linear map

$$Z_M : \mathcal{H}_N \to \mathbb{C}, \tag{93}$$

which is equivalent to a bra-state $\langle \Psi | \in \mathcal{H}_N^*$.

The usefulness of assigning $\mathcal{H}_\varnothing = \mathbb{C}$ is that all these special cases (91, 92, 93) are covered by one axiom (89).

[s] By incoming vs. outgoing boundary, we assume that each boundary component $N$ comes equipped with a nonzero infinitesimal vector field pointing into or out of $M$.

(4) Suppose $M_1$ has boundaries $(N, N')$ and $M_2$ has boundaries $(N', N'')$. We can form a new $d$-manifold $M_1 \cup_{N'} M_2$ by gluing $M_1$ and $M_2$ along $N'$. This manifold has boundaries $(N, N'')$. The total transition amplitude is the composition

$$Z_{M_1 \cup_{N'} M_2} = Z_{M_2} Z_{M_1} : \mathcal{H}_N \to \mathcal{H}_{N''}. \tag{94}$$

In scalar field theory, this follows by cutting the path integral along $N'$ by fixing boundary conditions $\phi'_b \in C(N', \mathbb{R})$, and then finally performing the $(d-1)$-dimensional path integral over $\phi'_b$,

$$\begin{aligned} &\int_{\substack{\phi|_N = \phi_b \\ \phi|_{N''} = \phi''_b}} D\phi(x) e^{-S[\phi]} \\ &= \int_{y \in N'} D\phi'_b(y) \int_{\substack{\phi_2|_{N'} = \phi'_b \\ \phi_1|_{N''} = \phi''_b \\ x \in M_2}} D\phi_2(x) e^{-S[\phi_2]} \int_{\substack{\phi_1|_N = \phi_b \\ \phi_1|_{N'} = \phi'_b \\ x \in M_1}} D\phi_1(x) e^{-S[\phi_1]}. \end{aligned} \tag{95}$$

This is the field-theory analog of our gluing rule for transition amplitudes in quantum mechanics (54). In the notation (90), we have simply inserted a complete set of states on $N'$,

$$\langle \phi''_b | Z_{M_2} Z_{M_1} | \phi_b \rangle = \int_{y \in N'} D\phi'_b(y) \langle \phi''_b | Z_{M_2} | \phi'_b \rangle \langle \phi'_b | Z_{M_1} | \phi_b \rangle. \tag{96}$$

(5) Describe generalization to correlation functions.

(6) Discuss orientation, orientation reversal, conjugation.

Discuss the possibility of giving the manifolds different amounts of structure, for instance topological structure, a Riemannian metric, etc. In these notes, we will initially be interested in manifolds with a metric. Later, we will see that in certain cases we can relax this structure to only a conformal class of metrics.

Although we have given explicit prescriptions in the case of a continuum theory, all of these axioms have analogs in the case of the 2d Ising model. It is not hard to imagine how the continuum axioms might arise in the continuum limit of the Ising model, at least in the case where the "manifolds" $M, N, \ldots$ are regions of flat space.

*End of lecture 4*

### 2.5. *Quantization in continuum QFT*

Let us use the above axioms to describe the quantization procedure one more time, now in continuum QFT. Consider the $d$-manifold $M_L = I_L \times N$,

where $I_L = [0, L]$ is an interval of length $L$ and $N$ is a $(d-1)$-manifold, and suppose $M_L$ is endowed with the product metric. The composition rule (94) implies that

$$Z_{M_L} = e^{-L\widehat{H}} \tag{97}$$

for some operator

$$\widehat{H} : \mathcal{H}_N \to \mathcal{H}_N, \tag{98}$$

namely the Hamiltonian.

Consider a correlation function on $S^1_\beta \times N$

$$\langle \phi(\tau_1, y_1) \cdots \phi(\tau_n, y_n) \rangle_{S^1_\beta \times N} = \int_{\phi(\beta,y)=\phi(0,y)} D\phi(x)\, \phi(\tau_1, y_1) \cdots \phi(\tau_n, y_n) e^{-S[\phi]}. \tag{99}$$

Here, $\tau, y$ are coordinates on $S^1_\beta, N$, respectively. As before, we introduce operators $\widehat{\phi}(y)$ that measure the field configuration on a time-slice

$$\widehat{\phi}(y)|\phi_b\rangle = \phi_b(y)|\phi_b\rangle. \tag{100}$$

To compute (99), we cut the path integral at times $\tau_1, \ldots, \tau_n$ and repeatedly use (97). By the usual logic, this gives

$$\langle \phi(\tau_1, y_1) \cdots \phi(\tau_n, y_n) \rangle_{S^1_\beta \times N} = \mathrm{Tr}_{\mathcal{H}_N}(e^{-\beta\widehat{H}} T\{\widehat{\phi}(\tau_1, y_1) \cdots \widehat{\phi}(\tau_n, y_n)\}), \tag{101}$$

where

$$\widehat{\phi}(\tau, y) \equiv e^{\tau\widehat{H}} \widehat{\phi}(y) e^{-\tau\widehat{H}}, \tag{102}$$

and $T\{\cdots\}$ represents time ordering in $\tau$. Taking the limit of a long interval, the factor $e^{-\beta\widehat{H}}$ projects onto the ground state, and we obtain[t]

$$\langle \phi(\tau_1, y_1) \cdots \phi(\tau_n, y_n) \rangle_{\mathbb{R} \times N} = \langle 0 | T\{\widehat{\phi}(\tau_1, y_1) \cdots \widehat{\phi}(\tau_n, y_n)\} | 0 \rangle. \tag{103}$$

For a QFT in $\mathbb{R}^d$, we can quantize the theory in many different ways by choosing different directions to play the role of "time." In every quantization of the theory, we have

$$\langle \phi(x_1) \cdots \phi(x_n) \rangle_{\mathbb{R}^d} = \langle 0 | T\{\widehat{\phi}(x_1) \cdots \widehat{\phi}(x_n)\} | 0 \rangle. \tag{104}$$

However, in different quantizations, the objects appearing on the right-hand side are different. The Hilbert spaces are different (though isomorphic if the

[t] In (103), we are assuming that the theory has a unique ground state on $N$. More generally, there might be multiple ground states on $N$, in which case the limit $\beta \to \infty$ becomes a trace in the space of ground states.

theory is SO($d$)-invariant), the ground states $|0\rangle$ are different, the quantum operators $\widehat{\phi}(x)$ are different, and the time-ordering symbols means different things. However, because they originate from a single path integral, the resulting expectation values are the same. It is sometimes useful to think of expressions like the RHS above as "interpretations" of one underlying path integral.

Discuss boundary conditions and whether we should think of $\mathbb{R}^d$ as having a boundary or not.

### 2.6. *The quantum transverse-field Ising model*

After all this abstract nonsense, let's return to more concrete things and try to understand the critical point of the 2d Ising model. The transfer matrix $\widehat{T}$ of the 2d Ising lattice model was diagonalized in 1944 by Onsager. Unfortunately, we won't have time to describe his solution. Instead, we will study a closely related model in the same universality class as the lattice model.

Recall that the transfer matrix has matrix elements $\langle \mathbf{s}'|\widehat{T}|\mathbf{s}\rangle = e^{-L(\mathbf{s}',\mathbf{s})}$. Let us write $\widehat{T}$ in a more familiar way as an operator on a spin chain. First split $L$ into contributions from horizontal and vertical bonds

$$\begin{aligned} L(\mathbf{s}',\mathbf{s}) &= L_h(\mathbf{s}') + L_h(\mathbf{s}) + L_v(\mathbf{s}',\mathbf{s}). \\ L_h(\mathbf{s}) &= -\frac{1}{2}K\sum_i \mathbf{s}_{i+1}\mathbf{s}_i, \\ L_v(\mathbf{s}',\mathbf{s}) &= \sum_i \frac{1}{2}K(\mathbf{s}'_i - \mathbf{s}_i)^2. \end{aligned} \tag{105}$$

Note that

$$e^{\frac{1}{2}K\sum_i \sigma_i^z\sigma_{i+1}^z}|\mathbf{s}\rangle = e^{-L_h(\mathbf{s})}|\mathbf{s}\rangle. \tag{106}$$

Meanwhile, $L_v$ only involves spins at a single site, so let us imagine that we have only one site. Note that

$$\langle s'|(1+e^{-2K}\sigma^x)|s\rangle = e^{-\frac{1}{2}K(s'-s)^2}. \tag{107}$$

We also have

$$\begin{aligned} 1+e^{-2K}\sigma^x &= e^{A+K'\sigma^x}, \qquad \text{where} \\ \tanh K' &= e^{-2K}, \\ e^A &= \sqrt{1-e^{-4K}}, \end{aligned} \tag{108}$$

which follows by expanding out the Taylor series for $e^{A+K'\sigma^x}$ and matching the coefficients of $1, \sigma^x$. Thus,

$$e^{AM}\langle \mathbf{s}'|e^{K'\sum_i \sigma_i^x}|\mathbf{s}\rangle = e^{-L_v(\mathbf{s}',\mathbf{s})}. \tag{109}$$

The constant $e^{AM}$ will cancel in correlation functions, so we will ignore it. Putting everything together, we find

$$\widehat{T} \propto \exp\left(\frac{1}{2}K\sum_i \sigma_i^z\sigma_{i+1}^z\right)\exp\left(K'\sum_i \sigma_i^x\right)\exp\left(\frac{1}{2}K\sum_i \sigma_i^z\sigma_{i+1}^z\right). \tag{110}$$

In a quantum-mechanical interpretation, we would write $\widehat{T} = e^{-\Delta\tau\widehat{H}}$, but the resulting $\widehat{H}$ would be very complicated.

To get the theory we'll actually study, we first allow $K', K$ to be independent parameters, and then take a limit where $K, K'$ become small with fixed ratio $K'/K = g$. These modifications may seem somewhat drastic, but it turns out that, at the critical point, they can be compensated by simply rescaling the time coordinate. (We will not prove this, unfortunately.) This results in

$$\widehat{T} \to e^{-\Delta\tau\widehat{H}}, \tag{111}$$

where $\Delta\tau$ is small and the Hamiltonian $\widehat{H}$ is relatively simple,

$$\widehat{H} = \sum_i \sigma_i^z\sigma_{i+1}^z + g\sum_i \sigma_i^x. \tag{112}$$

This is the Hamiltonian of the "quantum transverse-field Ising model" (TFIM). It describes a 1-dimensional chain of quantum spins with a nearest-neighbor interaction in the $z$-direction and an applied "transverse" magnetic field in the $x$-direction.[u,v]

### 2.6.1. *Solution via the Jordan-Wigner transformation*

In the remainder of this section, let us solve the TFIM by diagonalizing $\widehat{H}$. Since our Hilbert space is a tensor product of two possible states for each

[u] Don't confuse the transverse field in the quantum Ising model with the applied magnetic field $h$ in the thermodynamic Ising model. At the beginning of this discussion, we set $h = 0$.

[v] A similar procedure starting from the $d$-dimensional classical Ising model gives the quantum TFIM on a $(d-1)$-dimensional spatial lattice, which is again in the same universality class.

site, it's tempting to think of it as a fermionic Fock space, with creation and annihilation operators

$$\sigma_n^{\pm} \equiv \frac{1}{2}(\sigma_n^y \pm i\sigma_n^z). \tag{113}$$

This is not correct, since these operators commute rather than anticommute at different sites. However, this can be fixed with a classic trick called the Jordan-Wigner transformation. We define fermionic creation and annihilation operators

$$c_n^\dagger = \left(\prod_{i=1}^{n-1} \sigma_i^x\right) \sigma_n^+, \qquad c_n = \left(\prod_{i=1}^{n-1} \sigma_i^x\right) \sigma_n^-, \tag{114}$$

which now satisfy canonical anticommutation relations

$$\{c_n^\dagger, c_m\} = \delta_{nm}, \qquad \{c_n, c_m\} = \{c_n^\dagger, c_m^\dagger\} = 0. \tag{115}$$

These new creation and annihilation operators are *nonlocal* on the spin chain. We can think of them as creating and destroying fermionic solitons.

**Exercise 2.4.** *Verify the anticommutation relations (115).*

It's perhaps not surprising that we can write nonlocal variables that behave like fermions. The surprise is that the Hamiltonian in these new variables is still local, and actually quite simple

$$H = \sum_{n=1}^{M-1} (c_n^\dagger + c_n)(c_{n+1}^\dagger - c_{n+1}) - P(c_M^\dagger + c_M)(c_1^\dagger - c_1) + g\sum_{n=1}^{M}(2c_n^\dagger c_n - 1) \tag{116}$$

where $P = \prod_{n=1}^{M} \sigma_n^x = (-1)^F$ is the parity operator. Within each parity eigenspace, the Hamiltonian becomes simply

$$H = \sum_{n=1}^{N} (c_n^\dagger + c_n)(c_{n+1}^\dagger - c_{n+1}) + g\sum_{n=1}^{N}(2c_n^\dagger c_n - 1), \tag{117}$$

where when $P = -1$ we must impose periodic boundary conditions $c_{M+1} = c_1$, and when $P = 1$ we must impose antiperiodic boundary conditions $c_{M+1} = -c_1$.[w] Our Hamiltonian is now translationally invariant and

[w] The choice of which eigenspace $P = \pm 1$ to consider is equivalent to a choice of spin structure. The TFIM is a bosonic theory and can be formulated without choosing a spin structure. However, once we introduce fermions, we are required to choose a spin structure.

quadratic in creation and annihilation operators, so it can be diagonalized via a Fourier transform and Bogoliubov transformation:

$$H = \sum_k \left[ -(2\cos k - 2g) c_k^\dagger c_k - i \sin k \left( c_{-k}^\dagger c_k^\dagger + c_{-k} c_k \right) \right] - Mg \tag{118}$$

$$= \sum_k \epsilon(k) \left( b_k^\dagger b_k - \frac{1}{2} \right), \tag{119}$$

where $b_k^\dagger$ and $b_k$ are new canonically normalized creation and annihilation operators, and

$$\epsilon(k) = 2\sqrt{g^2 - 2g\cos(k) + 1} \tag{120}$$

is the dispersion relation for the free fermionic quasiparticles created by $b_k^\dagger$. Because of the boundary conditions imposed by parity, the quasimomenta must take the values

$$k = \begin{cases} \frac{2m\pi}{M} & P = -1, \\ \frac{(2m+1)\pi}{M} & P = 1. \end{cases} \tag{121}$$

for $m = 0, 1, \ldots, M-1$.

Write out the Bogoliubov transformation in more detail.

**Exercise 2.5.** *Derive expressions (116), (118), and (119).*

In the continuum limit $M \to \infty$, $k$ becomes continuous. Correlation functions are dominated by the smallest values of $\epsilon(k)$, which (assuming $g > 0$) occur near $k = 0$. Expanding around this point, we have

$$\epsilon(k)^2 = 4(1-g)^2 + 4gk^2 + O(k^4). \tag{122}$$

Thus, the mass gap is given by $m_{\text{gap}} = |g - 1|$. At the special point $g = g_c = 1$, the mass gap goes to zero and we have a critical point. Long-distance correlation functions are dominated by states with arbitrarily low energy, which requires $k \to 0$. Fur such states, we can drop the $O(k^4)$ term, and we obtain a relativistic dispersion relation[x]

$$\epsilon(k)^2 = 4k^2 \qquad (g = g_c,\ k \ll 1). \tag{123}$$

To summarize, in the continuum limit, at the critical coupling, and at long distances, the quantum TFIM becomes a free relativistic fermion. This is yet another example of IR equivalence/duality. The IR theory has an emergent $\mathrm{SO}(1,1)$ Lorentz symmetry that was not present in the original

[x] To obtain the usual dispersion relation $\epsilon = |k|$, we need to redefine either time or space by a factor of 2.

spin system. In fact, it is also has emergent conformal symmetry, as we'll see later.

The fact that the critical point of the 1+1d quantum TFIM (and consequently other theories in the same universality class, like the 2d Ising lattice model) is equivalent to a free theory is very special to 2-dimensions. Our derivation relied in a fundamental way on the Jordan-Wigner transformation and does not work, for example in 2+1d.[y]

*End of lecture 5*

## 3. CFTs in perturbation theory

### 3.1. *Scaling and renormalization*

In the quantum TFIM, we found $m_{\text{gap}} = |g - g_c|$, which means that the correlation length critical exponent $\nu$ (defined in equation (15)) is $\nu = 1$ for this theory. In this section, we explain how critical exponents are related to operator dimensions in the scale-invariant QFT (SFT) that describes the critical point.

Consider a microscopic system that has a long-distance description in terms of a quantum field theory $\text{QFT}_{\text{IR}}$. We will think of $\text{QFT}_{\text{IR}}$ in the language of effective field theory: it has a length cutoff $a$ and a set of coupling constants $g_A$ determined by running and matching.

> **Matching:** First, we choose a cutoff $a$ near the characteristic length scales of the microscopic theory. We fix the coupling constants $g_A(a)$ at this scale by computing physical observables in both the UV and IR theories at distances $x \gtrsim a$ and demanding that they agree term-by-term in a power series in $a/x$. To match more terms, we must adjust more coupling constants.
>
> An important point is that this matching procedure is analytic in the UV parameters. Specifically, if the action of the UV theory has a parameter $t$, then $g_A(a)$ will generically have a regular power-series expansion in $t$
>
> $$g_A(a) = \gamma_{A0} + \gamma_{A1} t + \gamma_{A2} t^2 + \dots . \tag{124}$$
>
> Roughly speaking this is because the matching happens at a fixed length scale, while nonanlyticities come from summing up contributions over

[y] There are newly-discovered versions of the Jordan-Wigner transformation in higher dimensions (`arXiv:1711.00515` — one of your classmates is a coauthor!), but they do not relate the 2+1d TFIM to a free theory.

many scales, as we will see shortly.[z]

**Running:** Now we perform a "coarse-graining" operation where we rescale the cutoff $a \to ba$ with $b = 1 + \delta\ell$, $\delta\ell \ll 1$ and adjust the coupling constants $g_A$ to leave the long-distance physics invariant.[aa] The adjustment takes the form

$$a\frac{dg_A}{da} = -\beta_A(g), \tag{125}$$

where $\beta_A(g)$ is a vector-field in the space of couplings, called the "beta function." The trajectories defined by (125) are "renormalization group (RG) flows."[bb]

An example RG flow is plotted in figure **??**. We will be interested in fixed-points where $\beta_A(g_*) = 0$. At a fixed-point, observables become scale-invariant. To see this explicitly, note that by dimensional analysis, any dimensionless observable (for example a dimensionless ratio of correlation functions) must have the form

$$G\left(\frac{x_{ij}}{a}; g_A\right), \tag{126}$$

where $x_{ij}$ are various distances and $g_A$ are coupling constants. The $\beta$-function is defined by demanding that physical observables are invariant under simultaneously rescaling $a$ and modifying the $g_A$, i.e.

$$\left(a\frac{\partial}{\partial a} + \beta_A(g)\frac{\partial}{\partial g_A}\right) G\left(\frac{x_{ij}}{a}; g_A\right) = 0. \tag{127}$$

Evaluating this equation at a fixed point $g_*$, we find

$$a\frac{\partial}{\partial a} G\left(\frac{x_{ij}}{a}; g_*\right) = 0 \qquad \Longrightarrow \qquad G\left(\frac{x_{ij}}{a}; g_*\right) = f\left(\frac{x_{ij}}{x_{kl}}\right), \tag{128}$$

[z] Put another way, the matching computation depends on degrees of freedom at some high energy scale. These degrees of freedom don't care about whether the theory will eventually reach a fixed-point at low energies (that is a low-energy question). From the point of view of the UV degrees of freedom, the point $t = 0$ is just some generic point in the space of couplings and there is no reason to have a singularity there.

[aa] In practice, we usually use a modified version of this condition, where we keep the physics invariant up to corrections of the form $(a/x)^n$ for some power $n$ that depends on the desired accuracy. This goes hand-in-hand with the matching procedure: first we decide on how many power-law corrections $(a/x)^n$ we want to keep, we then match coupling constants up to this order and find RG equations that preserve observables up to this order. A common simplification is to throw away all corrections $(a/x)^n$ with $n > 0$ — i.e. take the limit as the cutoff is completely removed $a \to 0$. This is the version of RG you find in most textbooks.

[bb] The beta function $\beta_A$ is analytic in $g_A$ by a similar argument to the one above: it is computed by comparing two nearby length/energy scales.

i.e. $G$ is a function of ratios of distances alone.

A fixed-point of the RG flow must be a critical point. The reason is that the scaling behavior (128) is inconsistent with the long-distance behavior

$$\langle \mathcal{O}(x)\mathcal{O}(y)\rangle \sim e^{-m_{\text{gap}}|x-y|} \qquad (|x-y| \gg a). \tag{129}$$

The only exceptions are $m_{\text{gap}} = 0$ (so that we have a nontrivial scale-invariant theory) or $m_{\text{gap}} = \infty$ (so that we have a TQFT).

Very close to a fixed point, $\beta_A$ takes the form

$$\beta_A = -\sum_B Y_{AB} u_B + O(u^2), \tag{130}$$

where $u = g - g_*$. Redefining the $u_A$ to diagonalize $Y_{AB}$,[cc] we have

$$\beta_A \approx -y_A u_A, \tag{131}$$

which integrates to

$$u_A \sim b^{y_A}. \tag{132}$$

We call the $u_A$ "scaling variables."

### 3.1.1. *Relevant, marginal, and irrelevant variables*

From here, we can distinguish three cases

- $y_A > 0$: the coupling $u_A$ is relevant and grows at long distances. An RG flow in the $u_A$-direction takes us away from the fixed-point.
- $y_A = 0$: the coupling $u_A$ is marginal. The fate of an RG flow in the $u_A$-direction depends on higher-order terms in the $\beta$-function.
- $y_A < 0$: the coupling $u_A$ is irrelevant. An RG flow in the $u_A$-direction takes us towards the fixed-point.

Relevant couplings must be tuned to stay close to the fixed point. From our discussion of the phase diagram of the Ising model, the fixed-point of the critical Ising model must have two relevant couplings: a thermal scaling variable $u_t$ and a magnetic scaling variable $u_h$. In addition, there are an infinite number of irrelevant variables $u_3, \ldots$.[dd] By our discussion of matching, the scaling variables $u_t, u_h$ should be analytic in the parameters

[cc] It is possible to have situations where $Y_{AB}$ cannot be diagonalized, but rather has a nontrivial Jordan block form. This occurs, for example in logarithmic CFTs, which we may discuss briefly later. However, logarithmic CFTs are not unitary. We will prove later that in unitary CFTs, the matrix $Y_{AB}$ is diagonalizable. For now, let us assume it.

[dd] When relevant variables respect a symmetry but irrelevant variables do not, that symmetry will be emergent at long-distances. The irrelevant variables that are initially nonzero depend on the microscopic theory. For example, in magnets with no applied

of the UV theory. They must also vanish when $t = h = 0$ and be consistent with $\mathbb{Z}_2$ symmetry. Thus,

$$\begin{aligned} u_t &= t/t_0 + O(t^2, h^2) \\ u_h &= h/h_0 + O(th), \end{aligned} \tag{133}$$

where $t_0$ and $h_0$ are non-universal constants.[ee]

In the action, scaling variables $u_A$ multiply local operators $\Phi_A$ that transform in a simple way under coarse-graining. The action linearized around the fixed-point is given by

$$S = S_0 + \int \frac{d^d x}{a^d} \sum_A u_A \Phi_A(x). \tag{134}$$

Here factors of the length-cutoff $a^{-d}$ are needed to make $S$ dimensionless. For example, if the UV theory is a lattice model, $a^{-d}$ comes from writing the sum over lattice sites as an integral over space

$$\sum_{i \in \mathbb{Z}^d} \to \int \frac{d^d x}{a^d}. \tag{135}$$

In our conventions, the $\Phi_A$ have classical dimension zero. (The word "classical" is to distinguish from another notion of dimension that we will introduce shortly.)

In order for the action to be invariant under coarse-graining, we must change the operators according to

$$\begin{aligned} a &\to ba, \\ u_A &\to b^{y_A} u_A, \\ \Phi_A &\to b^{\Delta_A} \Phi_A, \end{aligned} \tag{136}$$

where

$$y_A + \Delta_A = d. \tag{137}$$

magnetic field, all $\mathbb{Z}_2$-odd irrelevant variables will vanish by symmetry. However, in liquids the $\mathbb{Z}_2$-symmetry is emergent, so there will in general be nonzero $\mathbb{Z}_2$-odd irrelevant variables. In systems where rotational invariance is emergent, there will also be irrelevant variables that transform nontrivially under $\mathrm{SO}(d)$, e.g. a tensor $u_{\mu_1 \cdots \mu_j}$.

[ee] Generically, the map between UV and IR parameters is non-degenerate (i.e. differentiable with invertible derivative) at the fixed-point, so e.g. $t_0, h_0$ are not 0 or $\infty$. To get a nondegenerate map, we might need to find the right parameters in the microscopic theory to wiggle. For example, if we had a magnet such that $u_t \sim t^2$ (I don't know of any case where this actually happens), then in addition to dialing the temperature and magnetic field, we could also compress the magnet with some pressure $P$. The three-dimensional map between the UV parameters $(T, H, P)$ and scaling variables $u_t, u_h, u_3$ will then generically be nondegenerate. If it still isn't, we can add an additional parameter, etc..

The number $\Delta_A$ is called the "scaling dimension" of $\Phi_A$.

At a fixed-point, our coarse-graining rules fix the two-point function

$$\langle \Phi_A(x_1)\Phi_A(x_2)\rangle \propto \frac{a^{2\Delta_A}}{|x_1-x_2|^{2\Delta_A}}. \tag{138}$$

The right hand side is the only quantity consistent with Poincare invariance, with having classical dimension zero, and with the transformations (136).

We can simplify this argument by introducing a version of dimensional analysis specialized to our fixed-point. We define operators and coupling constants with nontrivial length dimension by

$$\begin{aligned} \mathcal{O}_A &= a^{-\Delta_A}\Phi_A, \\ \lambda_A &= a^{-y_A}u_A = a^{\Delta_A-d}u_A. \end{aligned} \tag{139}$$

These combinations are invariant under the transformation (136). Thus, invariance under coarse-graining means that correlators in terms of $\mathcal{O}_A, \lambda_A$ are independent of the cutoff $a$. The tradeoff is that correlators must now also be consistent with dimensional analysis under which $\mathcal{O}_A$ has length-dimension $-\Delta_A$ and $\lambda_A$ has length-dimension $-y_A = \Delta_A - d$. For example, at the fixed-point $\lambda_A = 0$, a two-point function of $\mathcal{O}_A$ must be given by

$$\langle \mathcal{O}_A(x_1)\mathcal{O}_A(x_2)\rangle \propto \frac{1}{|x_1-x_2|^{2\Delta_A}}. \tag{140}$$

The right-hand side is the only quantity consistent with Poincare symmetry and dimensional analysis. (Of course, it also follows trivially from (138) and (139).)

We can now analyze the action using this specially-adapted version of dimensional analysis. We have

$$S = S_0 + \int d^dx \sum_A \lambda_A \mathcal{O}_A(x). \tag{141}$$

Let us split the sum into relevant and irrelevant contributions

$$S = \left(S_0 + \int d^dx \sum_{\Delta_A<d} \lambda_A\mathcal{O}_A(x)\right) + \int d^dx \sum_{\Delta_A>d} \lambda_A\mathcal{O}_A(x) \tag{142}$$

(for simplicity we assume there are no marginal operators). The irrelevant interactions scale towards the fixed-point, so at low energies we will assume they can be treated as small perturbations of the action in parentheses.[ff]

[ff] This assumption is not always correct — it can happen that the effects of an irrelevant interaction initially decrease near the fixed-point, but begin to grow when the effect of relevant interactions gets strong. This is called a "dangerously-irrelevant" operator.

For concreteness, let us focus on the Ising model. We first set the irrelevant $\lambda_A$'s to zero so we can later treat them as small perturbations. Let us also tune parameters in the UV so that $\lambda_h = 0$ (in the case of a magnet, we switch off the external field). There is now exactly one relevant parameter turned on, namely $\lambda_t$. We can associate a mass scale to it given by

$$m_t = |\lambda_t|^{1/y_t} = |u_t|^{1/y_t} a^{-1}. \tag{143}$$

By dimensional analysis, any other dimensionful quantity must be proportional to this mass scale. For example, the mass gap of the theory in the IR must be

$$m_{\text{gap}} \propto m_t \propto |u_t|^{1/y_t} \propto |t|^{1/y_t}. \tag{144}$$

This determines the critical exponent

$$\nu = \frac{1}{y_t} = \frac{1}{d - \Delta_t}. \tag{145}$$

Note that the constant of proportionality in (144) can depend on the sign of $t$. The reason is that, depending on the sign of $t$, the theory may flow to different theories in the IR (figure **??**). For example, in the Ising model, when $t > 0$ the theory flows to a gapped phase where $\mathbb{Z}_2$ is unbroken, whereas when $t < 0$, the theory flows to a gapped phase that spontaneously breaks $\mathbb{Z}_2$. No matter where the theory flows, $m_t$ is still the only mass-scale around, so $m_{\text{gap}}$ must be proportional to $m_t$. However, the actual dynamics that determines $m_{\text{gap}}$ can be arbitrarily complicated and generally depends in a nontrivial way on the actual IR theory.

*End of lecture 6*