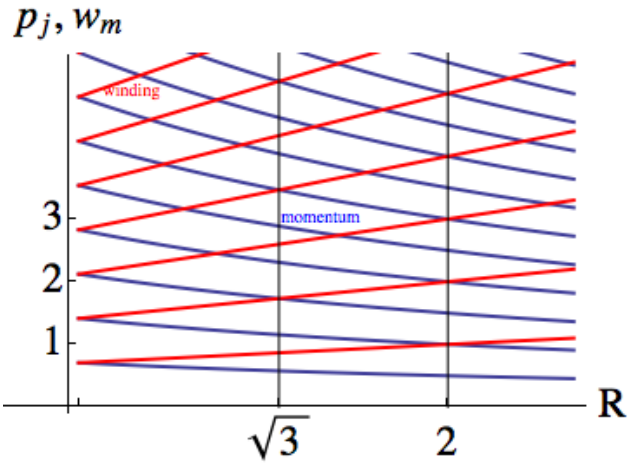
Physics 239a:
Where do quantum field theories come from?

Lecturer: McGreevy



Last updated: 2024/05/17, 12:49:47

0.1 Introductory remarks

I begin with some discussion of my goals for this course. This is a special topics course directed at graduate students in theoretical physics; this includes high-energy theory and condensed matter theory and maybe some other areas, too.

The subject of the course is regulated quantum field theory (QFT): we will study quantum field theories which can be constructed by starting from systems with finitely many degrees of freedom per unit volume, with local interactions between them. Often these degrees of freedom will live on a lattice and perhaps could arise as a model of a solid. So you could say that the goal is to understand how QFT arises from condensed matter systems. (However, I do not plan to discuss any chemistry or materials science.)

Why do I make a big deal about 'regulated'? For one thing, this is how it comes to us (when it does) in condensed matter physics. For another, such a description is required if we want to know what we're talking about. For example, we need it if we want to know what we're talking about well enough to explain it to a computer. Many QFT problems are too hard for our brains. A related but less precise point is that I would like to do what I can to erase the problematic perspective on QFT which 'begins from a classical lagrangian and quantizes it' etc, and leads to a term like 'anomaly'.

The crux of many problems in physics is the correct choice of variables with which to label the degrees of freedom. Often the best choice is very different from the obvious choice; a name for this phenomenon is 'duality'. We will study many examples of it (Kramers-Wannier, Jordan-Wigner, bosonization, Wegner, particle-vortex, perhaps others). This word is dangerous (it is one of the forbidden words on my blackboard) because it is *about* ambiguities in our (physics) language. I would like to reclaim it.

An important bias in deciding what is meant by 'correct' or 'best' in the previous paragraph is: we will be interested in low-energy and long-wavelength physics, near the groundstate. For one thing, this is the aspect of the present subject which is like 'elementary particle physics'; the high-energy physics of these systems is of a very different nature and bears little resemblance to the field often called 'high-energy physics' (for example, there is volume-law entanglement).

Also, notice that above I said 'a finite number of degrees of freedom *per unit volume*'. This last is important, because we are going to be interested in the thermodynamic limit. Questions about a finite amount of stuff (this is sometimes called 'mesoscopics') tend to be much harder.

An important goal for the course is demonstrating that many fancy phenomena precious to particle physicists can emerge from very humble origins in the kinds of (completely well-defined) local quantum lattice models we will study. Here I have in mind: fermions, gauge theory, photons, anyons, strings, topological solitons, CFT, and many other sources of wonder I'm forgetting right now.

Coarse classification of translation-invariant states of matter

So the most basic question we can ask is: how many degrees of freedom are there at the lowest energies? There are essentially three possibilities:

1. **None.**
2. **Some.**
3. **A lot.**

A more informative tour through that list goes like this. First let me make the assumption that the system has translation invariance, so we can label the excitations by momentum. (Relaxing this assumption is interesting but we won't do it this quarter.)

1. **None:** Such a system has an energy gap ('is gapped'): the energy difference $\Delta E = E_1 - E_0$ between the first excited state and the groundstate is nonzero, even in the thermodynamic limit. Note that ΔE is almost always nonzero in finite volume. (Recall, for example, the spectrum of the electromagnetic field in a box of linear size L: $E_n \sim \frac{n}{L}$.) The crucial thing here (in contrast to the case of photons) is that this energy stays finite even as $L \to \infty$.

   The excitations of such a system are massive particles[1].

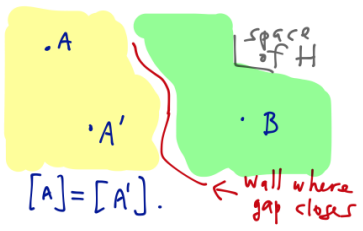
2. **Some:** An example of what I mean by 'some' is that the system can have excitations which are massless particles, like the photon.

   The lowest energy degrees of freedom occur at isolated points in momentum space: $\omega(k) = c\sqrt{\vec{k} \cdot \vec{k}}$ vanishes at $\vec{k} = 0$.

   In this category I also put the gapless fluctuations at a critical point; in that case, it's not necessarily true that $\omega \sim k^{\text{integer}}$ and those excitations are not necessarily *particles*. But they are still at $k = 0$[2].

3. **A lot:** What I mean by this is Fermi surfaces (but importantly, not just free fermions).

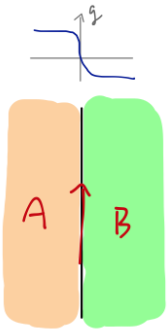
Let's go through that list one more time more slowly. Let's reconsider the case of gapped systems. Different gapped states are different if we can't deform the hamiltonian to get from one to the other without closing the gap. You might be bothered that it is hard to imagine checking that there is no way around the wall. It is therefore important to find sharp characterizations of such states, like integer labels, which cannot change smoothly.



[1] Verstraete et al claims to have proved a version of this statement.

[2] or some other isolated points in momentum space.

Even the lowest-energy (even below the gap) physics of gapped systems can be deeply fascinating. Such a thing is a topological field theory: it is a theory of groundstates, and it can provide a way to distinguish states of matter. For example, it may be that the *number* of groundstates depends on the topology of the space on which we put the system. This phenomenon is called topological order. Another possibility is that even if the system in infinite space has an energy gap, if we cut the space open, new stuff can happen; for example there may be gapless edge modes. Both of these phenomena happen in quantum Hall systems. A reason to think that an interface between the vacuum and a gapped state of matter which is distinct from the trivial one might carry gapless modes is that the couplings in the hamiltonian are forced to pass through the wall where the gap closes. (In fact there are important exceptions to this conclusion.)



An energy gap (and no topological order or special edge modes) should probably be the generic expectation for what happens if you pile together a bunch of degrees of freedom and couple them in some haphazard (but translation invariant!) way. At the very least this follows on general grounds of pessimism: if you generically got something interesting by doing this, physics would be a lot easier (or more likely: we wouldn't find it interesting anymore). **Gaplessness is something special that needs to be explained.** Here is a list of some possible reasons for gaplessness (if you find another, you should write it down):

1. broken continuous symmetry (Goldstone bosons)
2. tuning to a critical point – notice that this requires some agent to do the tuning, and will only occur on some subspace of the space of couplings of nonzero codimension.
3. continuous unbroken gauge invariance (photons)
4. Fermi surface (basically only in this case do we get gapless degrees of freedom at some locus of dimension greater than one in momentum space)
5. edge of topological phase: non-onsite realization of symmetry, anomaly inflow.
6. CFT with no relevant operators

Each entry in this list is something to be understood. We will get to many of them. I have promised various people that I would devote special attention to the second option, in the case of 1+1 dimensions; it is also relevant to the last two options.

Another important axis along which to organize states of matter is by symmetry. Specifically, we can label them according to the symmetry group G that acts on their Hilbert space and commutes with the Hamiltonian. You will notice that here I am speaking about what are called *global symmetries*, that is, symmetries (not redundancies of our labelling, like gauge transformations).

There are many refinements of this organization. We can ask how the symmetry G is *realized*, in at least three senses:

1. most simply, what representations of the group appear in the system?
2. is the symmetry preserved by the groundstate? If not, this is called 'spontaneous symmetry breaking'.
3. is it 'on-site'? Alternatively, is it 'anomalous'? What are its anomaly coefficients? I'll postpone the explanation of these terms.

Some useful lattice models (and their symmetry), which I hope to discuss this quarter:

- toric code (none), lattice gauge theory (none, necessarily)
- ising ($\mathbb{Z}_2$)
- potts, clock ($\mathsf{S}_N, \mathbb{Z}_N$)
- XY (spins, rotors) ($\mathsf{U}(1)$)
- Heisenberg ($\mathsf{SU}(2)$)
- $\mathsf{O}(n)$ ($\mathsf{O}(n)$)
- lattice (tight-binding) fermions
- lattice (tight-binding) bosons
- Kitaev honeycomb

Resources: The material which comes closest to my present goals is a course taught by Ashvin Vishwanath at Berkeley called 'Demystifying quantum field theory' and many insights can be found in his notes. Some material from the books by Sachdev, Wen, Fradkin, Polyakov will be useful.

1 QFT from springs

My purpose here is to remind us that QFT need not be regarded as a mystical beast that begins its life with infinitely many degrees of freedom near every point in space. In addition, the following discussion contains rudimentary versions of many of the ideas we'll need. Actually it contains even more of them than I thought when I decided to include it – some of them are simply missing from the textbook treatments of this system.

Consider a one-dimensional box of point masses, each connected by a spring to its neighbor. This is a nice simple model of a crystalline solid (we can speak about its shortcomings later).

When in equilibrium, the masses form a regular one-dimensional crystal lattice (equally spaced mass points). Now let q_n denote the displacement of the nth mass from its equilibrium position x_n and let p_n be the corresponding momentum. Assume there are N masses and impose periodic boundary conditions: $q_{n+N} = q_n$. The equilibrium positions themselves are

$$x_n = na, n = 1, 2...N$$

where a is the lattice spacing. The Hamiltonian for the collection of masses is:

$$\mathbf{H} = \sum_{n=1}^{N} \left(\frac{\mathbf{p}_n^2}{2m} + \frac{1}{2}\kappa \left(\mathbf{q}_n - \mathbf{q}_{n-1}\right)^2 \right) + \lambda \mathbf{q}^4. \tag{1}$$

I've include a token anharmonic term $\lambda \mathbf{q}^4$ to remind us that we are leaving stuff out; for example we might worry whether we could use this model to describe *melting*. Now set $\lambda = 0$ because we are going to study small deviations from $q = 0$.

This hamiltonian above describes a collection of coupled oscillators, with a matrix of spring constants $V = k_{ab}\mathbf{q}_a\mathbf{q}_b$. If we diagonalize the matrix of spring constants, we will have a description in terms of decoupled oscillators, called *normal modes*.

Since our system has (discrete) translation invariance, these modes are labelled by a wavenumber k[3]:

$$q_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{\mathbf{i}kx_n} q_n, \quad p_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{\mathbf{i}kx_n} p_n,$$

(Notice that in the previous expressions I didn't use boldface; that's because this step is really just classical physics. Note the awkward (but in field theory, inevitable) fact that we'll have (field) momentum operators $\mathbf{p}_k$ labelled by a wavenumber aka momentum.) The nice thing about the fourier kernel is that it diagonalizes the translation operator:

$$\mathbf{T}e^{\mathbf{i}kx} \equiv e^{\mathbf{i}k(x+a)} = e^{\mathbf{i}ka}e^{\mathbf{i}kx}.$$

Regulators: Because N is finite, k takes discrete values ($1 = e^{\mathbf{i}kNa}$); this is a long-wavelength "IR" property. Because of the lattice structure, k is periodic (only $e^{\mathbf{i}kan}, n \in \mathbb{Z}$ appears): $k \equiv k + 2\pi/a$; this is a short-distance "UV" property. The range of k can be taken to be

$$0 \leq k \leq \frac{2\pi(N-1)}{Na} \ .$$

[3]The inverse transformation is:

$$q_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-\mathbf{i}kx_n} q_k, \quad p_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-\mathbf{i}kx_n} p_k.$$

Because of the periodicity in k, we can equivalently label the set of wavenumbers by:

$$0 < k \leq \frac{2\pi}{a} \quad \text{or} \quad -\frac{\pi}{a} < k \leq \frac{\pi}{a}.$$

This range of independent values of the wavenumber in a lattice model is called the *Brillouin zone*. There is some convention for choosing a fundamental domain which prefers the last one but I haven't found a reason to care about this.

Summary: Because the system is in a box (periodic), k-space is discrete. Because the system is on a lattice, k-space is periodic. There are N oscillator modes altogether.

The whole hamiltonian is a bunch of decoupled oscillators, labelled by these funny wave numbers:

$$\mathbf{H} = \sum_k \left(\frac{\mathbf{p}_k \mathbf{p}_{-k}}{2m} + \frac{1}{2} m\omega_k^2 \mathbf{q}_k \mathbf{q}_{-k} \right)$$

where the frequency of the mode labelled k is

$$\omega_k \equiv 2\sqrt{\frac{\kappa}{m}} \sin \frac{|k|a}{2}. \tag{2}$$

Why might we care about this frequency? For one thing, consider the Heisenberg equation of motion for the deviation of one spring:

$$\mathbf{i}\partial_t \mathbf{q}_n = [\mathbf{q}_n, \mathbf{H}] = \frac{\mathbf{p}_n}{m}, \quad \mathbf{i}\partial_t \mathbf{p}_n = [\mathbf{p}_n, \mathbf{H}]$$

Combining these gives:

$$m\ddot{q}_n = -\kappa \left((q_n - q_{n-1}) - (q_n - q_{n+1}) \right) = -\kappa \left(2q_n - q_{n-1} - q_{n+1} \right).$$

In terms of the fourier-mode operators:

$$m\ddot{\mathbf{q}}_k = -\kappa \left(2 - 2\cos ka \right) \mathbf{q}_k \ .$$

Plugging in a fourier ansatz in time $q_k(t) = \sum_\omega e^{-\mathbf{i}\omega t} q_{k,\omega}$ turns this into an algebraic equation which says $\omega^2 = \omega_k^2 = \left(\frac{2\kappa}{m}\right) \sin^2 \frac{|k|a}{2}$ for the allowed modes. We see that (the classical version of) this system describes waves:

$$0 = \left(\omega^2 - \omega_k^2 \right) q_{k,\omega} \overset{k \ll 1/a}{\simeq} \left(\omega^2 - v_s^2 k^2 \right) q_{k,\omega}.$$

The result for small k is the fourier transform of the wave equation:

$$\left(\partial_t^2 - v_s^2 \partial_x^2 \right) q(x,t) = 0 \ . \tag{3}$$

v_s is the speed of propagation of the waves, in this case the speed of sound. Comparing to the dispersion relation (2), we have found

$$v_s = \frac{\partial \omega_k}{\partial k}|_{k \to 0} = a\sqrt{\frac{\kappa}{m}}.$$

The description we are about to give is a quantization of sound waves.

Notice that when $k = 0, \omega_k = 0$. We are going to have to treat this mode specially; there is a lot of physics in it.

So far the fact that quantumly $[\mathbf{q}_n, \mathbf{p}_{n'}] = \mathbf{i}\hbar\delta_{nn'}\mathbb{1}$ hasn't mattered in our analysis (go back and check). For the Fourier modes, this implies the commutator

$$[\mathbf{q}_k, \mathbf{p}_{k'}] = \sum_{n,n'} \mathbf{U}_{kn}\mathbf{U}_{k'n'}[\mathbf{q}_n, \mathbf{p}'_n] = \mathbf{i}\hbar\mathbb{1}\sum_n \mathbf{U}_{kn}\mathbf{U}_{k'n} = \mathbf{i}\hbar\delta_{k,-k'}\mathbb{1}.$$

(In the previous expression I called $\mathbf{U}_{kn} = \frac{1}{\sqrt{N}}e^{\mathbf{i}kx_n}$ the unitary matrix realizing the discrete Fourier kernel.)

To make the final step to decouple the modes with k and $-k$, introduce the annihilation and creation operators

$$\textbf{For } k \neq 0\textbf{:} \qquad \mathbf{q}_k = \sqrt{\frac{\hbar}{2m\omega_k}}\left(\mathbf{a}_k + \mathbf{a}^\dagger_{-k}\right), \quad \mathbf{p}_k = \frac{1}{\mathbf{i}}\sqrt{\frac{\hbar m\omega_k}{2}}\left(\mathbf{a}_k - \mathbf{a}^\dagger_{-k}\right).$$

They satisfy

$$[\mathbf{a}_k, \mathbf{a}^\dagger_{k'}] = \delta_{kk'}\mathbb{1}.$$

In terms of these, the hamiltonian is

$$\mathbf{H}_0 = \sum_k \hbar\omega_k\left(\mathbf{a}^\dagger_k\mathbf{a}_k + \frac{1}{2}\right) \quad + \frac{p_0^2}{2m}$$

– it is a sum of decoupled oscillators, and a free particle describing the center-of-mass. It is worth putting together the final relation between the 'position operator' and the phonon annihilation and creation operators:

$$\mathbf{q}_n = \sqrt{\frac{\hbar}{2m}}\sum_k \frac{1}{\sqrt{\omega_k}}\left(e^{\mathbf{i}kx}\mathbf{a}_k + e^{-\mathbf{i}kx}\mathbf{a}^\dagger_k\right) + \frac{1}{\sqrt{N}}\mathbf{q}_0 \tag{4}$$

and the corresponding relation for its canonical conjugate momentum

$$\mathbf{p}_n = \frac{1}{\mathbf{i}}\sqrt{\frac{\hbar m}{2}}\sum_k \sqrt{\omega_k}\left(e^{\mathbf{i}kx}\mathbf{a}_k - e^{-\mathbf{i}kx}\mathbf{a}^\dagger_k\right) + \frac{1}{\sqrt{N}}\mathbf{p}_0.$$

Notice that these expressions are formally identical to the formulae in a QFT textbook expressing a scalar field in terms of creation and annihilation operators.

[End of Lecture 1]

The groundstate is obtained from

$$|0\rangle \otimes |p_0 = 0\rangle, \qquad \text{where} \quad \mathbf{a}_k|0\rangle = 0, \ \forall\ k, \text{and } \mathbf{p}_0|p_0\rangle = p_0|p_0\rangle\ .$$

A set of excited states is

$$\mathbf{a}_k^\dagger|0\rangle$$

This state has energy $\hbar\omega_k$ above the groundstate. In a box of size $L = Na$, the smallest energy excitation has $k_1 = \frac{2\pi}{L}$ and energy

$$\Delta E \sim \frac{1}{L} \overset{L\to\infty}{\to} 0 \ . \tag{5}$$

(Note that here I am taking $L \to \infty$ to implement the thermodynamic limit of infinitely many degrees of freedom; the lattice spacing can remain finite for this purpose – it is not a continuum limit.) So according to our definition, this system is gapless. Why? Goldstone[4]: the system has a symmetry under $q_n \to q_n + \epsilon$ for all n. If everyone moves to the left three feet, none of the springs are stretched. This is the dance enacted by the $k = 0$ mode. If *nearly* everyone moves *nearly* three feet to the left, the springs will only be stretched a little; hence the modes with small k have small ω.

Tower of States: Now I will say a few words about the zeromode, which is horribly mistreated in all the textbook discussions of this system that I've seen (Le Bellac, Altland-Simons, and my own quantum mechanics lecture notes!). There is no potential at all for this mode – it drops out of the $(q_n - q_{n+1})^2$ terms. It just has a kinetic term, which we can think of as the center-of-mass energy of the system. How much energy does it cost to excite this mode? Notice that if everyone moves to the left by a, the system comes back to itself (I am assuming that the masses are indistinguishable particles): $|\{q_n\}\rangle \simeq |\{q_n + a\}\rangle$. In terms of the $k = 0$ mode, this is

$$q_0 = \frac{1}{\sqrt{N}}\sum_{n=1}^{N} q_n e^{-\mathbf{i}0x_n} \simeq \frac{1}{\sqrt{N}}\left(\sum_{n=1}^{N} q_n + Na\right), \quad i.e.\ q_0 \simeq q_0 + \sqrt{N}a.$$

This means that the wavefunction for the zeromode must satisfy

$$e^{\mathbf{i}p_0 q_0} = e^{\mathbf{i}p_0(q_0+\sqrt{N}a)} \implies p_0 \in \frac{2\pi\mathbb{Z}}{\sqrt{N}a}$$

and the first excited state has energy

$$\frac{p_0^2}{2m}|_{p_0=\frac{2\pi}{\sqrt{N}a}} = \frac{1}{2}\frac{1}{Nm}\left(\frac{2\pi}{a}\right)^2.$$

This is a victory at least in the sense that we expect the center of mass of the system to have an intertial mass Nm. Notice that the spacing of these states depends differently on the parameters than that of the ones from the nonzero-k phonon states.

But actually this phenomenon is ubiquitous: it happens whenever we take a system which breaks a continous symmetry (here: a solid breaks continuous translation invariance)[5] and

[4] I should probably give a review of Goldstone's theorem here. The relevant statement for our purposes is: if the groundstate is not invariant under a continuous symmetry of $\mathbf{H}$, the spectrum is gapless. A useful reference for a more specific statement is this paper.

[5] The fine print:

put it in finite volume, *i.e.* depart from the thermodynamic limit. In particular, in finite volume the zeromode associated with a conserved quantity (here the momentum) produces a *tower of states* with a different level-spacing (as a function of system size $L = Na$) than the usual modes (5). (It is sometimes called the *Anderson Tower of States* in the study of magnetism or the *Rotator spectrum* in lattice gauge theory). In this case, both towers go like $1/N$, but this is a coincidence. In other cases the tower from the zeromode is more closely spaced (it goes like $\frac{1}{\text{volume}} \sim \frac{1}{L^d} \sim \frac{1}{N}$) than the particle momentum tower (which always goes like $\frac{1}{L} \sim \frac{1}{N^{1/d}}$ (or maybe $\frac{1}{L^2}$)), so the tower of states from the zeromode is usually much closer together, and in the thermodynamic limit $L \to \infty$, they combine to form the degenerate vacua associated with spontaneous symmetry breaking. [6]

Phonons. Back to the phonon excitations. We can specify basis states for this Hilbert space

$$\left(\mathbf{a}_{k_1}^\dagger\right)^{n_{k_1}} \left(\mathbf{a}_{k_2}^\dagger\right)^{n_{k_2}} \cdots |0\rangle = |\{n_{k_1}, n_{k_2}, ...\}\rangle$$

by a collection of *occupation numbers* n_k, eigenvalues of the number operator for each normal mode.

We can make a state with two phonons:

$$|k, k'\rangle = \mathbf{a}_k^\dagger \mathbf{a}_{k'}^\dagger |0\rangle$$

and so on. As you can see, phonons are identical bosons: the state with several of them only says how many there are with each wavenumber.

So you see that we have constructed an approximation to the Fock space of a (massless) scalar field from a system with finitely many degrees of freedom per unit volume (here, length), and in fact finitely many degrees of freedom altogether, since we kept the IR regulator L finite. It is worth pausing to appreciate this: we've been forced to discover a framework for quantum systems in which particles can be created and annihilated, very different from the old-fashioned point of view where we have a fixed Hilbert space for each particle.

Many aspects of the above discussion are special to the fact that our hamiltonian was quadratic in the operators. Certainly our ability to completely solve the system is. Notice

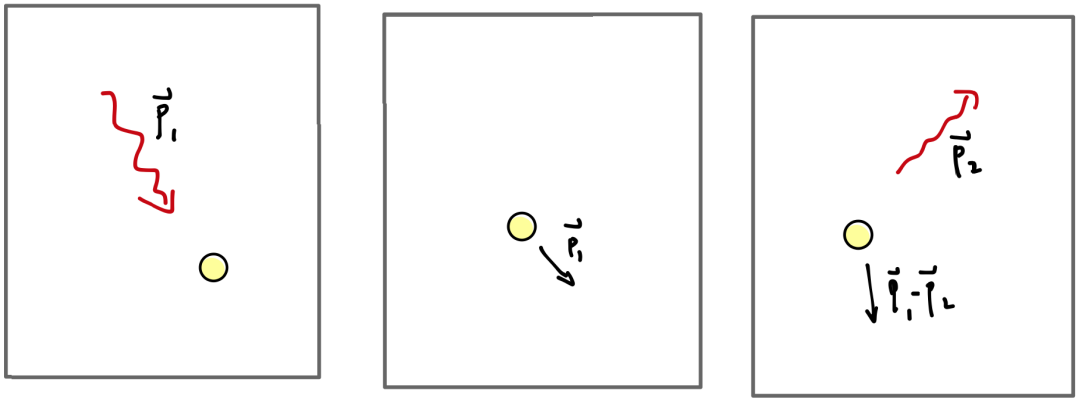
1. Actually it's important that the order parameter doesn't commute with the Hamiltonian; the exception is ferromagnets, where the order parameter is the total spin itself, which is a conserved quantity and therefore can be nonzero even in finite volume. So the tower is collapsed at zero in that case.

2. Actually, a one-dimensional mattress of oscillators will not spontaneously break continuous translation symmetry even in infinite volume. This is a consequence of the Coleman-Mermin-Wagner theorem: the positions of the atoms still fluctuate too much, even when there are infinitely many of them in a row; more than one dimension is required to have the crystal really sit still. You'll see the effects of these fluctuations on the problem set when you study the Debye-Waller factors. This does not vitiate our conclusions above at all.

[6] The definitive discussion of this subject can be found in the last few pages of P. Anderson, *Concepts in Solids*.

that the number of phonons *of each momentum* $\mathbf{n}_k \equiv \mathbf{a}_k^\dagger \mathbf{a}_k$ is conserved for each k. But if we add generic cubic and quartic terms in $\mathbf{q}$ (or if we couple our atoms to the photon field) even the number of phonons $\sum_k \mathbf{n}_k$ will no longer be a conserved quantity.[7] So a description of such particles which forced us to fix their number wouldn't be so great. More generally, I would like to emphasize that not every QFT can be usefully considered as nearly a bunch of harmonic oscillators. Finding ways to think about the ones which can't is an important job.

A nice example where we can see the importance of the tower of states and of the quantization of phonon number is the Mössbauer effect: when scattering high-energy photons off a solid, there is a finite amplitude for scattering events which produce *zero* phonons. This means that all the momentum transfer goes to the center of mass mode, which has negligible recoil as $N \to \infty$, since its inertial mass is Nm. This allows for very sharp absorption lines, which if the atom were in free space would be washed out (*i.e.* broadened to a width $E_{\text{recoil}} = \frac{(p_\gamma)^2}{2m}$.) by the random recoils of the individual atoms (as depicted in the comic strip below).



I've made a big deal about the regulators here. One reason we care about them is if we remove them ($N \to \infty, a \to 0$) and ask bad questions, we'll get infinity. For example, we could think about the vacuum energy $E_0 = \frac{1}{2}\sum_k \hbar\omega_k$. There is physics in there (*e.g.* Casimir forces), but we will not discuss that now.

[7] Note that it *is* possible to make a quadratic action for conserved particles, but this requires adding more degrees of freedom – the required $\mathsf{U}(1)$ symmetry must act something like

$$(q^1, q^2) \to (\cos\theta q^1, \sin\theta q^2).$$

We can reorganize this as a complex field $\Phi = q^1 + \mathbf{i}q^2$ on which the symmetry acts by $\Phi \to e^{\mathbf{i}\theta}\Phi$. This is illustrated on the problem set.

Continuum limit

The thing we have discovered is a *quantum field theory.*

At this point I will use the path-integral description of a 1d particle with $\mathbf{H} = \frac{\mathbf{p}^2}{2m} + V(\mathbf{q})$, a basic statement of which is the following formula for the propagator

$$\langle q|e^{-\mathbf{i}\mathbf{H}t}|q_0\rangle = \int_{q(0)=q_0}^{q(t)=q} [dq] e^{\mathbf{i}\int_0^t \mathrm{d}t\ \left(\frac{1}{2}\dot{q}^2 - V(q)\right)} \ .$$

Here $[dq] \equiv \mathcal{N}\prod_{l=1}^{M_\tau} dq(t_l)$ – the path integral measure is defined by a limiting procedure, and $\mathcal{N}$ is a normalization factor that always drops out of physical quantities.

Recall that the key step in the derivation of this statement is the evaluation of the propagator for an infinitesimal time:

$$\langle q_2|e^{-\mathbf{i}\mathbf{H}\Delta t}|q_1\rangle = \langle q_2|e^{-\mathbf{i}\Delta t\frac{\mathbf{p}^2}{2m}}e^{-\mathbf{i}V(\mathbf{q})}|q_2\rangle + \mathcal{O}(\Delta t^2) \ .$$

An integral expression for this can be obtained by inserting resolutions of the identity

$$\mathbb{1} = \mathbb{1}^2 = \left(\int \mathrm{d}p|p\rangle\langle p|\right)\left(\int \mathrm{d}q|q\rangle\langle q|\right)$$

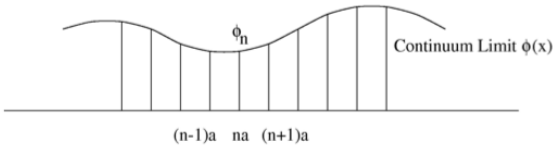
in between the two exponentials.

Scalar field theory in one dimension

If we use the path integral description, some of these things (in particular the continuum, sound-wave limit) are more obvious-seeming. The path integral for our collection of oscillators is

$$Z = \int [dq_1 \cdots dq_N] e^{\mathbf{i}S[q]}$$

with $S[q] = \int dt \left(\sum_n \frac{1}{2}m_n\dot{q}_n^2 - V(\{q\})\right)$. $V(\{q\}) = \sum_n \frac{1}{2}\kappa\left(q_{n+1} - q_n\right)^2$. Now let's try to take the continuum limit $a \to 0, N \to \infty$. Basically the only thing we need is to think of



$q_n = q(x = na)$ as defining a smooth function:
[Note that the continuum field is often called $\phi(x)$ instead of $q(x)$ for some reason. At least the letters $q(x)$ and $\phi(x)$ look similar.]

We now have

$$\left(q_n - q_{n-1}\right)^2 \simeq a^2\left(\partial_x q\right)^2|_{x=na}.$$

Now the path integral becomes:

$$Z = \int [Dq] e^{\mathbf{i}S[q]}$$

with Dq now representing an integral over all configurations $q(t,x)$ (*defined* by this limit) and

$$S[q] = \int dt \int dx \frac{1}{2}\left(\mu\left(\partial_t q\right)^2 - \mu v_s^2\left(\partial_x q\right)^2 - rq^2 - uq^4 - ...\right) \equiv \int dt \int dx \mathcal{L}$$

where I've introduced some parameters μ, v_s, r, u determined from $m, \kappa...$ in some ways that we needn't worry about right now (*e.g.* $\mu = m/a$, the mass per unit length of the chain). $\mathcal{L}$ is the Lagrangian *density* whose integral over space is the Lagrangian $L = \int dx \mathcal{L}$. The ellipses ... represent terms of higher order in the Taylor expansion of $q(x)$, which are suppressed by correspondingly more powers of a: $a^n \partial_x^n q$.

The equation of motion (stationary phase condition) is

$$0 = \frac{\delta S}{\delta q(x,t)} = -\mu \ddot{q} - \mu v_s^2 \partial_x^2 q - rq - 2uq^3 - ...$$

From the phonon problem, we automatically found $r = u = 0$, and the equation of motion is just the wave equation (3). This happened because of the symmetry $q_n \to q_n + \epsilon$. This is the operation that *translates* the whole crystal, It guarantees low-energy phonons near $k = 0$ because it means $q(x)$ can only appear in S via its derivatives.

The expressions for the field operators in terms of mode operators are now

$$\mathbf{q}(x) = \sqrt{\frac{\hbar}{2\mu}} \int dk \frac{1}{\sqrt{\omega_k}} \left(e^{\mathbf{i}kx}\mathbf{a}_k + e^{-\mathbf{i}kx}\mathbf{a}_k^\dagger\right)$$

and its canonical conjugate momentum

$$\mathbf{p}(x) = \frac{1}{\mathbf{i}} \sqrt{\frac{\hbar\mu}{2}} \int dk \sqrt{\omega_k} \left(e^{\mathbf{i}kx}\mathbf{a}_k - e^{-\mathbf{i}kx}\mathbf{a}_k^\dagger\right).$$

($\mathbf{p}(x)$ is the quantum operator associated with the field-momentum π above.) These equations are the same as (4), but with squinting. They are just the ones in introductory QFT textbooks; the stray factors of μ arise because we didn't 'canonically normalize' our fields and absorb the μs into the field, *e.g.* defining $\phi \equiv \sqrt{\mu} q$ would get rid of them. Notice that the position along the chain x here is just a *label* on the fields (it was n), not a quantum operator.

The field q is called a *scalar field* because it doesn't have any indices decorating it. This is to be distinguished from *e.g.* the Maxwell field, which is a vector field. (Note that vibrations of a crystal in three dimensions actually do involve vector indices. We will omit this complication from our discussion.)

2 Ising spins

The Ising model has many guises. There is this from statistical mechanics:

$$Z = \sum_{\{s_j\}} e^{-K\sum_{\langle jl\rangle} s_j s_l}.$$

There is this quantum spin system:

$$\mathbf{H}_{TFIM} = -J\sum_j \left(g_x \mathbf{X}_j + g_z \mathbf{Z}_j \mathbf{Z}_{j+1}\right) \ .$$

And there is this 2d conformal field theory:

$$S[\chi] = \int \mathrm{d}^2 z \left(\chi \overline{\partial}_z \chi\right) \tag{6}$$

which I first encountered on the worldsheet of a superstring. An important part of our job is to understand the connections between these things. One thing they have in common is a $\mathbb{Z}_2$ symmetry, $s_j \to -s_j$ or $\mathbf{Z}_j \to -\mathbf{Z}_j$ or $\chi \to -\chi$.

Let me say a few introductory words about quantum spin systems. This means that we have a collection of two-state systems (aka qbits) $\mathcal{H}_j = \mathrm{span}\{|\uparrow_j\rangle, |\downarrow_j\rangle\}$ distributed over space and coupled somehow:

$$\mathcal{H} = \bigotimes_j \mathcal{H}_j \ , \quad \dim\left(\mathcal{H}\right) = 2^N$$

where N is the number of sites.

One qbit: To begin, consider just one two-state system. There are four independent hermitian operators acting on this Hilbert space. Besides the identity, there are the three Paulis, which I will denote by $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ instead of $\boldsymbol{\sigma}^x, \boldsymbol{\sigma}^y, \boldsymbol{\sigma}^z$:

$$\mathbf{X} \equiv \boldsymbol{\sigma}^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{Y} \equiv \boldsymbol{\sigma}^y = \begin{pmatrix} 0 & -\mathbf{i} \\ \mathbf{i} & 0 \end{pmatrix}, \quad \mathbf{Z} \equiv \boldsymbol{\sigma}^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

This notation (which comes to us from the quantum information community) makes the important information larger and is therefore better, especially for those of us with limited eyesight. [End of Lecture 2]

They satisfy

$$\mathbf{XY} = \mathbf{iZ}, \quad \mathbf{XZ} = -\mathbf{ZX}, \quad \mathbf{X}^2 = \mathbb{1},$$

and all cyclic permutations $\mathbf{X} \to \mathbf{Y} \to \mathbf{Z} \to \mathbf{X}$ of these statements.

Multiple qbits: If we have more than one site, the paulis on different sites *commute*:

$$[\boldsymbol{\sigma}_j, \boldsymbol{\sigma}_l] = 0, \quad j \neq l \quad i.e. \quad \mathbf{X}_j \mathbf{Z}_l = (-1)^{\delta_{jl}} \mathbf{Z}_l \mathbf{X}_j,$$

where $\boldsymbol{\sigma}_j$ is any of the three paulis acting on $\mathcal{H}_j$.

2.1 Quantum-to-classical correspondence

[Kogut, Sachdev chapter 5, Goldenfeld §3.2]

Goals : Facilitate rapid passage between Lagrangian and Hamiltonian descriptions. Harness your knowledge of stat mech.

Important conclusions: correlation length $\sim \frac{1}{\text{gap}}$.

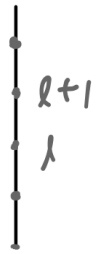
Let's begin with the classical ising model in a (longitudinal) magnetic field:

$$Z = \sum_{\{s_j\}} e^{-K \sum_{\langle jl \rangle} s_j s_l - h \sum_j s_j}. \tag{7}$$

Here I am imagining we have classical spins $s_j = \pm 1$ at each site of some graph, and $\langle jl \rangle$ denotes pairs of sites which share a link in the graph. You might be tempted to call K the inverse temperature, which is how we would interpret if we were doing classical stat mech; resist the temptation.

One qbit from ising chain

First, let's think about the case when the graph in (7) is just a chain:



$$Z_1 = \sum_{\{s_l=\pm 1\}} e^{-H_c}, \qquad H_c = -K\sum_{l=1}^{M_\tau} s_l s_{l+1} - h\sum_{l=1}^{M_\tau} s_l \tag{8}$$

These ss are now just M_τ numbers, each ± 1 – there are 2^{M_τ} terms in this sum. (Notice that the field h breaks the $s \to -s$ symmetry of the summand.) The parameter $K > 0$ is the 'inverse temperature' in the Boltzmann distribution; I put these words in quotes because I want you to think of it as merely a parameter in the classical hamiltonian.

For definiteness let's suppose the chain loops back on itself,

$$s_{l+M_\tau} = s_l \quad \text{(periodic boundary conditions)}.$$

Using the identity $e^{\sum_l (\dots)_l} = \prod_l e^{(\dots)_l}$,

$$Z_1 = \sum_{\{s_l\}} \prod_{l=1}^{M_\tau} T_1(s_l, s_{l+1}) T_2(s_l)$$

where

$$T_1(s_1, s_2) \equiv e^{K s_1 s_2}, \quad T_2(s) \equiv e^{hs} \ .$$

What are these objects? The conceptual leap is to think of $T_1(s_1, s_2)$ as a 2×2 matrix:

$$T_1(s_1, s_2) = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix}_{s_1 s_2} = \langle s_1 | \mathbf{T}_1 | s_2 \rangle,$$

which we can then regard as matrix elements of an operator $\mathbf{T}_1$ acting on a 2-state *quantum* system (hence the boldface). And we have to think of $T_2(s)$ as the *diagonal* elements of the same kind of matrix:

$$\delta_{s_1,s_2} T_2(s_1) = \begin{pmatrix} e^h & 0 \\ 0 & e^{-h} \end{pmatrix}_{s_1 s_2} = \langle s_1 | \mathbf{T}_2 | s_2 \rangle.$$

So we have

$$Z_1 = \mathrm{tr} \left(\underbrace{(\mathbf{T}_1 \mathbf{T}_2)(\mathbf{T}_1 \mathbf{T}_2) \cdots (\mathbf{T}_1 \mathbf{T}_2)}_{M_\tau \text{ times}} \right) = \mathrm{tr} \mathbf{T}^{M_\tau} \tag{9}$$

where I've written

$$\mathbf{T} \equiv \mathbf{T}_2^{\frac{1}{2}} \mathbf{T}_1 \mathbf{T}_2^{\frac{1}{2}} = \mathbf{T}^\dagger = \mathbf{T}^t$$

for convenience (so it's symmetric). This object is the *transfer matrix*. What's the trace over in (9)? It's a single two-state system – a single qbit (or quantum spin) that we've constructed from this chain of classical two-valued variables.

Even if we didn't care about quantum spins, this way of organizing the partition sum of the Ising chain does the sum for us (since the trace is basis-independent, and so we might as well evaluate it in the basis where $\mathbf{T}$ is diagonal):

$$Z_1 = \mathrm{tr}\mathbf{T}^{M_\tau} = \lambda_+^{M_\tau} + \lambda_-^{M_\tau}$$

where $\lambda_\pm$ are the two eigenvalues of the transfer matrix, $\lambda_+ \geq \lambda_-$:

$$\lambda_\pm = e^K \cosh h \pm \sqrt{e^{2K}\sinh^2 h + e^{-2K}} \overset{h\to 0}{\to} \begin{cases} 2\cosh K \\ 2\sinh K. \end{cases} \tag{10}$$

In the thermodynamic limit, $M_\tau \gg 1$, one of them dominates the free energy

$$e^{-F} = Z_1 = \lambda_+^{M_\tau}\left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^{M_\tau}\right) \sim \lambda_+^{M_\tau}.$$

Now I command you to think of the transfer matrix as

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

the propagator in euclidean time (by an amount $\Delta\tau$), where $\mathbf{H}$ is the *quantum* hamiltonian operator for a single qbit (note the boldface to denote quantum operators). So what's $\mathbf{H}$? To answer this, let's rewrite the parts of the transfer matrix in terms of paulis, thinking of $s = \pm$ as $\mathbf{Z}$-eigenstates. For $\mathbf{T}_2$, which is diagonal in the $\mathbf{Z}$ basis, this is easy:

$$\mathbf{T}_2 = e^{h\mathbf{Z}}.$$

To write $\mathbf{T}_1$ this way, stare at its matrix elements in the $\mathbf{Z}$ basis:

$$\langle s_1|\mathbf{T}_1|s_2\rangle = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix}_{s_1 s_2}$$

and compare them to those of

$$e^{a\mathbf{X}+b\mathbb{1}} = e^b e^{a\mathbf{X}} = e^b\left(\cosh a + \mathbf{X}\sinh a\right)$$

which are

$$\langle s_1|e^{a\mathbf{X}+b\mathbb{1}}|s_2\rangle = e^b \begin{pmatrix} \cosh a & \sinh a \\ \sinh a & \cosh a \end{pmatrix}_{s_1,s_2}$$

So we want $e^b \sinh a = e^{-K}, e^b \cosh a = e^K$ which is solved by

$$e^{-2K} = \tanh a\ . \tag{11}$$

So we want to identify

$$\mathbf{T}_1\mathbf{T}_2 = e^{b\mathbb{1}+a\mathbf{X}}e^{h\mathbf{Z}} \equiv e^{-\Delta\tau\mathbf{H}}$$

for small $\Delta\tau$. This requires that a, b, h scale like $\Delta\tau$, and so we can combine the exponents. Assuming that $\Delta\tau \ll 1$ (so that we can ignore the corrections to the leading term in the Baker-Campbell-Hausdorff formula), the result is

$$\mathbf{H} = E_0 - \frac{\Delta}{2}\mathbf{X} - \bar{h}\mathbf{Z} \ .$$

Here $E_0 = \frac{b}{\Delta\tau}, \bar{h} = \frac{h}{\Delta\tau}, \Delta = \frac{2a}{\Delta\tau}$. (Note that it's not surprising that the Hamiltonian for an isolated qbit is of the form $\mathbf{H} = d_0 \mathbb{1} + \vec{d} \cdot \vec{\boldsymbol{\sigma}}$, since these operators span the set of hermitian operators on a qbit; but the relation between the parameters that we've found will be important.)

To recap, let's go backwards: consider the quantum system consisting of a single spin with $\mathbf{H} = E_0 - \frac{\Delta}{2}\mathbf{X} + \bar{h}\mathbf{Z}$. Set $\bar{h} = 0$ for a moment. Then Δ is the energy gap between the groundstate and the first excited state (hence the name). The thermal partition function is

$$Z_Q(T) = \mathrm{tr} e^{-\mathbf{H}/T} = \sum_{s=\pm} \langle s|e^{-\beta\mathbf{H}}|s\rangle, \tag{12}$$

in the $\mathbf{Z}$ basis, $\mathbf{Z}|s\rangle = s|s\rangle$. I emphasize that T here is the temperature to which we are subjecting our quantum spin; $\beta = \frac{1}{T}$ is the length of the euclidean time circle. Break up the euclidean time circle into M_τ intervals of size $\Delta\tau = \beta/M_\tau$. Insert many resolutions of unity (this is called 'Trotter decomposition')

$$Z_Q = \sum_{s_1 \ldots s_{M_\tau}} \langle s_{M_\tau}|e^{-\Delta\tau\mathbf{H}}|s_{M_\tau - 1}\rangle \langle s_{M_\tau - 1}|e^{-\Delta\tau\mathbf{H}}|s_{M_\tau - 2}\rangle \cdots \langle s_1|e^{-\Delta\tau\mathbf{H}}|s_{M_\tau}\rangle \ .$$

The RHS[8] is the partition function of a classical Ising chain, Z_1 in (8), with $h = 0$ and K given by (11), which in the present variables is:

$$e^{-2K} = \tanh\left(\frac{\beta\Delta}{2M_\tau}\right) \ . \tag{13}$$

Notice that if our interest is in the quantum model with couplings E_0, Δ, we can use any M_τ we want – there are many classical models we could use[9]. For given M_τ, the couplings we should choose are related by (13).

A quantum system with just a single spin (for any $\mathbf{H}$ not proportional to $\mathbb{1}$) clearly has a unique groundstate; this statement means the absence of a phase transition in the 1d Ising chain.

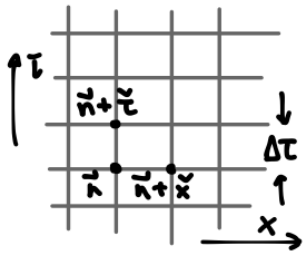
[8] RHS = right-hand side, LHS = left-hand side, BHS = both-hand side

[9] If we include the $\mathbf{Z}$ term, we need to take $\Delta\tau$ small enough so that we can write

$$e^{-\Delta\tau\mathbf{H}} = e^{\Delta\tau\frac{\Delta}{2}\mathbf{X}} e^{-\Delta\tau(E_0 - \bar{h}\mathbf{Z})} + \mathcal{O}(\Delta\tau^2)$$

More than one spin[10]

Let's do this procedure again, supposing the graph in question is a cubic lattice with more than one dimension, and let's think of one of the directions as euclidean time, τ. We'll end up with more than one spin.



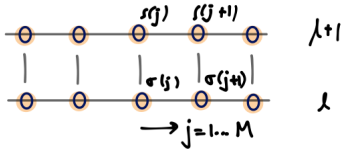
We're going to rewrite the sum in (7) as a sum of products of (transfer) matrices. I will draw the pictures associated to a square lattice, but this is not a crucial limitation. Label points on the lattice by a vector $\vec{n}$ of integers; a unit vector in the time direction is $\check{\tau}$. First rewrite the classical action H_c in $Z_c = \sum e^{-H_c}$ as[11]

$$\begin{aligned} H_c &= -\sum_{\vec{n}} \left(K s(\vec{n}+\check{\tau}) s(\vec{n}) + K_x s(\vec{n}+\check{x}) s(\vec{n}) \right) \\ &= K \sum_{\vec{n}} \left(\frac{1}{2} \left(s(\vec{n}+\check{\tau}) - s(\vec{n}) \right)^2 - 1 \right) - K_x \sum_{\vec{n}} s(\vec{n}+\check{x}) s(\vec{n}) \\ &= \text{const} + \sum_{\text{rows at fixed time, } l} L(l+1, l) \end{aligned} \tag{14}$$

with[12]

$$L(s,\sigma) = \frac{1}{2} K \sum_j \left(s(j) - \sigma(j) \right)^2 - \frac{1}{2} K_x \sum_j \left(s(j+1) s(j) + \sigma(j+1) \sigma(j) \right).$$

σ and s are the names for the spins on successive time slices, as in the figure at left.



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

$$\langle s | \mathbf{T} | \sigma \rangle = T_{s\sigma} = e^{-L(s,\sigma)},$$

in terms of which

$$Z = \sum_{\{s\}} e^{-H_c} = \sum_{\{s\}} \prod_{l=1}^{M_\tau} T_{s(l,j),s(l+1,j)} = \mathrm{tr}_{\mathcal{H}} \mathbf{T}^{M_\tau} \ .$$

This is just as in the one-site case; the difference is that now the hilbert space has a two-state system for every site on a fixed-l slice of the lattice. I will call this "space", and label these sites by an index j. (Note that nothing we say in this discussion requires space to be one-dimensional.) So $\mathcal{H} = \bigotimes_j \mathcal{H}_j$, where each $\mathcal{H}_j$ is a two-state system.

[10] This discussion comes from this paper of Fradkin and Susskind, and can be found in Kogut's review article.

[11] The minus signs in red were flubbed in lecture.

[12] Note that 'L' is for 'Lagrangian'. In retrospect, I should have called $H_c \equiv S = \int d\tau L$, as Kogut does. '$S$' is for 'action'.

The diagonal entries of $T_{s,\sigma}$ come from contributions where $s(l) = \sigma(l)$: they come with a factor of $\mathbf{T}_{s=\sigma} = e^{L(0 \text{ flips})}$ with

$$L(0 \text{ flips}) = -K_x \sum_j \sigma(j+1)\sigma(j).$$

The one-off-the-diagonal terms come from

$$\sigma(j) = s(j), \quad \text{except for one site where instead } \sigma(j) = -s(j).$$

This gives a contribution

$$L(1 \text{ flips}) = \underbrace{\frac{1}{2}K(1-(-1))^2}_{=2K} - \frac{1}{2}K_x \sum_j \left(\sigma(j+1)\sigma(j) + s(j+1)s(j)\right).$$

Similarly,

$$L(n \text{ flips}) = 2nK - \frac{1}{2}K_x \sum_j \left(\sigma(j+1)\sigma(j) + s(j+1)s(j)\right).$$

Now we need to figure out who is $\mathbf{H}$, as defined by

$$\mathbf{T} = e^{-\Delta\tau \mathbf{H}} \simeq 1 - \Delta\tau \mathbf{H} \ ;$$

we want to consider $\Delta\tau$ small and must choose K_x, K to make it so. We have to match the matrix elements $\langle s|\mathbf{T}|\sigma\rangle = T_{s\sigma}$:

$$\begin{aligned} T(0 \text{ flips})_{s\sigma} &= \delta_{s\sigma} e^{K_x \sum_j s(j)s(j+1)} &&\simeq \mathbb{1} - \Delta\tau \mathbf{H}|_{0 \text{ flips}} \\ T(1 \text{ flip})_{s\sigma} &= e^{-2K} e^{\frac{1}{2}K_x \sum_j (\sigma(j+1)\sigma(j) + s(j+1)s(j))} &&\simeq -\Delta\tau \mathbf{H}|_{1 \text{ flip}} \\ T(\text{n flips})_{s\sigma} &= e^{-2nK} e^{\frac{1}{2}K_x \sum_j (\sigma(j+1)\sigma(j) + s(j+1)s(j))} &&\simeq -\Delta\tau \mathbf{H}|_{\text{n flips}} \end{aligned} \tag{15}$$

From the first line, we learn that $K_x \sim \Delta\tau$; from the second we learn $e^{-2K} \sim \Delta\tau$; we'll call the ratio which we'll keep finite $g \equiv K_x^{-1} e^{-2K}$. To make τ continuous, we take $K \to \infty, K_x \to 0$, holding g fixed. Then we see that the n-flip matrix elements go like $e^{-nK} \sim (\Delta\tau)^n$ and can be ignored – the hamlitonian only has 0- and 1-flip terms.

To reproduce (15), we must take

$$\mathbf{H}_{\text{TFIM}} = -J \left(g \sum_j \mathbf{X}_j + \sum_j \mathbf{Z}_{j+1}\mathbf{Z}_j \right).$$

Here J is a constant with dimensions of energy that we pull out of $\Delta\tau$. The first term is the 'one-flip' term; the second is the 'zero-flips' term. The first term is a 'transverse magnetic field' in the sense that it is transverse to the axis along which the neighboring spins interact. So this is called the *transverse field ising model*. We're going to understand it completely below. As we'll see, it contains the *universal* physics of the 2d Ising model, including Onsager's solution. The word 'universal' requires some discussion.

[End of Lecture 3]

Symmetry of the transverse field quantum Ising model: $\mathbf{H}_{\rm TFIM}$ has a $\mathbb{Z}_2$ symmetry, generated by $\mathbf{S} = \prod_j \mathbf{X}_j$, which acts by

$$\mathbf{S}\mathbf{Z}_j = -\mathbf{Z}_j\mathbf{S},\ \mathbf{S}\mathbf{X}_j = +\mathbf{X}_j\mathbf{S},\quad \forall j;$$

On $\mathbf{Z}$ eigenstates it acts as:

$$\mathbf{S}|\{s_j\}_j\rangle = |\{-s_j\}_j\rangle \quad .$$

It is a symmetry in the sense that:

$$[\mathbf{H}_{\rm TFIM}, \mathbf{S}] = 0.$$

Notice that $\mathbf{S}^2 = \prod_j \mathbf{X}_j^2 = \mathbb{1},$, and $\mathbf{S} = \mathbf{S}^\dagger = \mathbf{S}^{-1}$.

By 'a $\mathbb{Z}_2$ symmetry,' I mean that the symmetry group consists of two elements $G = \{\mathbb{1}, \mathbf{S}\}$, and they satisfy $\mathbf{S}^2 = \mathbb{1}$, just like the group $\{1, -1\}$ under multiplication. This group is $G = \mathbb{Z}_2$. (For a bit of context, the group $\mathbb{Z}_N$ is realized by the Nth roots of unity, under multiplication.)

The existence of this symmetry of the quantum model is a direct consequence of the fact that the summand of the classical system was invariant under the operation $s_j \to -s_j, \forall j$. This meant that the matrix elements of the transfer matrix satisfy $T_{s,s'} = T_{-s,-s'}$ which implies the symmetry of $\mathbf{H}$. (Note that symmetries of the classical action do not so immediately imply symmetries of the associated quantum system if the system is not as well-regulated as ours is. This is the phenomenon called 'anomaly'.)

Quantum Ising in d space dimensions to classical ising in $d+1$ dims

[Sachdev, 2d ed p. 75] Just to make sure it's nailed down, let's go backwards again. The partition function of the quantum Ising model at temperature T is

$$Z_Q(T) = \mathrm{tr}_{\bigotimes_{j=1}^M \mathcal{H}_j} e^{-\frac{1}{T}\mathbf{H}_I} = \mathrm{tr}\left(e^{-\Delta\tau \mathbf{H}_I}\right)^{M_\tau}$$

The transfer matrix here $e^{-\Delta\tau\mathbf{H}_I}$ is a $2^M \times 2^M$ matrix. We're going to take $\Delta\tau \to 0, M_\tau \to \infty$, holding $\frac{1}{T} = \Delta M_\tau$ fixed. Let's use the usual[13] 'split-step' trick of breaking up the non-commuting parts of $\mathbf{H}$:

$$e^{-\Delta\tau\mathbf{H}_I} \equiv \mathbf{T}_x \mathbf{T}_z + \mathcal{O}(\Delta\tau^2).$$
$$\mathbf{T}_x \equiv e^{Jg\Delta\tau \sum_j \mathbf{X}_j}, \quad \mathbf{T}_z \equiv e^{J\Delta\tau \sum_j \mathbf{Z}_j \mathbf{Z}_{j+1}} \quad .$$

Now insert a resolution of the identity in the $\mathbf{Z}$-basis,

$$\mathbb{1} = \sum_{\{s_j\}_{j=1}^M} |\{s_j\}\rangle\langle\{s_j\}|, \quad \mathbf{Z}_j|\{s_j\}\rangle = s_j|\{s_j\}\rangle, \quad s_j = \pm 1.$$

many many times, one between each pair of transfer operators; this turns the transfer operators into transfer *matrices*. The $\mathbf{T}_z$ bit is diagonal, by design:

$$\mathbf{T}_z|\{s_j\}\rangle = e^{J\Delta\tau \sum_j s_j s_{j+1}}|\{s_j\}\rangle.$$

The $\mathbf{T}_x$ bit is off-diagonal, but only on a single spin at a time:

$$\langle\{s'_j\}|\mathbf{T}_x|\{s_j\}\rangle = \prod_j \underbrace{\langle s'_j|e^{Jg\Delta\tau\mathbf{X}_j}|s_j\rangle}_{2\times 2}$$

Acting on a single spin at site j, this 2×2 matrix is just the one from the previous discussion around (11):

$$\langle s'_j|e^{Jg\Delta\tau\mathbf{X}_j}|s_j\rangle = e^{-b}e^{Ks'_j s_j}, \quad e^{-2b} = \frac{1}{2}\sinh\left(2Jg\Delta\tau\right), \quad e^{-2K} = \tanh\left(Jg\Delta\tau\right) \quad .$$

Notice that it wasn't important to restrict to $1+1$ dimensions here. The only difference is in the $\mathbf{T}_z$ bit, which gets replaced by a product over all neighbors in higher dimensions:

$$\langle\{s'_j\}|\mathbf{T}_z|\{s_j\}\rangle = \delta_{s,s'} e^{J\Delta\tau \sum_{\langle jl\rangle} s_j s_l}$$

where $\langle jl\rangle$ denotes nearest neighbors, and the innocent-looking $\delta_{s,s'}$ sets the spins $s_j = s'_j$ equal for all sites.

[13] By 'usual' I mean that this is just like in the path integral of a 1d particle, when we write

$$e^{-\Delta\tau\mathbf{H}} = e^{-\frac{\Delta\tau}{2m}\mathbf{p}^2} e^{-\Delta\tau V(\mathbf{q})} + \mathcal{O}(\Delta\tau^2).$$

Label the time slices by a variable $l = 1...M_\tau$.

$$Z = \mathrm{tr} e^{-\frac{1}{T}\mathbf{H}_I} = \sum_{\{s_j(l)\}} \prod_{l=1}^{M_\tau} \langle \{s_j(l+1)\} | \mathbf{T}_z \mathbf{T}_x | \{s_j(l)\} \rangle$$

The sum on the RHS runs over the 2^{MM_τ} values of $s_j(l) = \pm 1$, which is the right set of things to sum over in the $d+1$-dimensional classical ising model. The weight in the partition sum is

$$\begin{aligned} Z &= \underbrace{e^{-bM_\tau}}_{\text{unimportant}} \sum_{\{s_j(l)\}_{j,l}} \exp\left(\sum_{j,l} \left(\underbrace{J\Delta\tau s_j(l) s_{j+1}(l)}_{\text{space deriv, from } \mathbf{T}_z} + \underbrace{K s_j(l) s_j(l+1)}_{\text{time deriv, from } \mathbf{T}_x} \right) \right) \\ &= \overbrace{\sum_{\text{spins}}}^{\text{constant}} e^{-H_{\text{classical ising}}} \end{aligned}$$

except that the the couplings are a bit anisotropic: the couplings in the 'space' direction $K_x = J\Delta\tau$ are not the same as the couplings in the 'time' direction, which satisfy $e^{-2K} = \tanh(Jg\Delta\tau)$. (At the critical point $K = K_c$, this can be absorbed in a rescaling of spatial directions, as we'll see later.)

Dictionary. So this establishes a mapping between classical systems in $d+1$ dimensions and quantum systems in d space dimensions. Here's the dictionary:

statistical mechanics in $d+1$ dimensions	quantum system in d space dimensions
transfer matrix	euclidean-time propagator, $e^{-\Delta\tau \mathbf{H}}$
statistical 'temperature'	(lattice-scale) coupling K
free energy in infinite volume	groundstate energy: $e^{-F} = Z = \mathrm{tr} e^{-\beta \mathbf{H}} \overset{\beta\to 0}{\to} e^{-\beta E_0}$
periodicity of euclidean time L_τ	temperature: $\beta = \frac{1}{T} = \Delta\tau M_\tau$
statistical averages	groundstate expectation values of time-ordered operators

Note that this correspondence between classical and quantum systems is not an isomorphism. For one thing, we've seen that *many* classical systems are related to the same quantum system, which does not care about the lattice spacing in time. There is a set of physical quantities which agree between these different classical systems, called *universal*, which is the information in the quantum system. More on this below.

Consequences for phase transitions and quantum phase transitions.

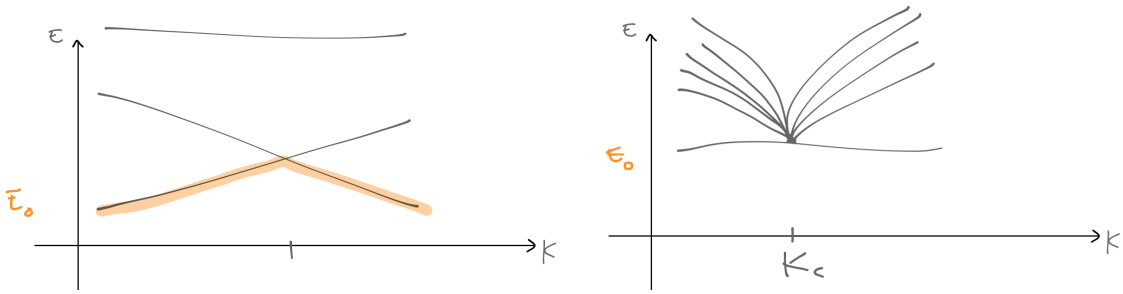
One immediate consequence is the following. Think about what happens at a *phase transition* of the classical problem. This means that the free energy $F(K, ...)$ has some kind of singularity at some value of the parameters, let's suppose it's the statistical temperature, *i.e.* the parameter we've been calling K. 'Singularity' means breakdown of the Taylor expansion, *i.e.* a disagreement between the actual behavior of the function and its Taylor series – a *non-analyticity*. First, this can *only* happen in the thermodynamic limit (at the very least $M_\tau \to \infty$), since otherwise there are only a finite number of terms in the partition sum and F is an analytic function of K (it's a polynomial in e^{-K}).

An important dichotomy is between *continuous* phase transitions (also called second order or higher) and first-order phase transitions; at the latter, $\partial_K F$ is discontinous at the transition, at the former it is not. This seems at first like an innocuous distinction, but think about it from the point of view of the transfer matrix for a moment. In the thermodynamic limit, $Z = \lambda_1(K)^{M_\tau}$, where $\lambda_1(K)$ is the largest eigenvalue of $\mathbf{T}(K)$. How can this have a singularity in K? There are two possibilities:

1. $\lambda_1(K)$ is itself a singular function of K. How can this happen? One way it can happen is if there is a level-crossing where two completely unrelated eigenvectors switch which is the smallest (while remaining separated from all the others).

This is a first-order transition. A distinctive feature of a first order transition is a *latent heat*: although the free energies of the two phases are equal at the transition (they have to be in order to exchange dominance there), their entropies (and hence energies) are not: $S \propto \partial_K F$ jumps across the transition.

2. The other possibility is that the eigenvalues of $\mathbf{T}$ have an *accumulation point* at $K = K_c$, so that we can no longer ignore the contributions from the other eigenvalues to $\mathrm{tr}\mathbf{T}^{M_\tau}$, even when $M_\tau = \infty$. This is the exciting case of a continuous phase transition. In this case the critical point K_c is really special.



Now translate those statements into statements about the corresponding quantum system. Recall that $\mathbf{T} = e^{-\Delta\tau \mathbf{H}}$ – eigenvectors of $\mathbf{T}$ are eigenvectors of $\mathbf{H}$! Their eigenvalues are related by

$$\lambda_a = e^{-\Delta\tau E_a} \ ,$$

so the largest eigenvalue of the transfer matrix corresponds to the smallest eigenvalue of $\mathbf{H}$: the groundstate. The two cases described above are:

1. As the parameter in $\mathbf{H}$ varies, two completely orthogonal states switch which one is the groundstate. This is a 'first-order quantum phase transition', but that name is a bit grandiose for this boring phenomenon, because the states on the two sides of the transition don't need to know anything about each other, and there is no interesting critical theory. For example, the *third* excited state need know nothing about the transition.

2. At a continuous transition in $F(K)$, the spectrum of $\mathbf{T}$ piles up at the top. This means that the spectrum of $\mathbf{H}$ is piling up at the bottom: the gap is closing. There is a gapless state which describes the physics in a whole neighborhood of the critical point.

Using the quantum-to-classical dictionary, the groundstate energy of the TFIM at the transition reproduces Onsager's tour-de-force free energy calculation.

Another failure mode of this correspondence: there are some quantum systems which when Trotterized produce a stat mech model with non-positive Boltzmann weights, *i.e.* $e^{-H_c} < 0$

for some configurations; this requires H_c to be complex. These models are less familiar! (An example where this happens is the spin-$\frac{1}{2}$ chain.) The quantum phase transitions of such quantum systems are not just ordinary finite-temperature transitions of familiar classical stat mech systems. So for the collector of QFTs, there is something to be gained by studying quantum phase transitions.

Correlation functions

[Sachdev, 2d ed p. 69] First let's construct correlation functions of spins in the classical Ising chain, (8), using the transfer matrix. (We'll study correlation functions in the TFIM in §2.2.3.) Let

$$C(l,l') \equiv \langle s_l s_{l'} \rangle = \frac{1}{Z_1} \sum_{\{s_l\}_l} e^{-H_c} s_l s_{l'}$$

By translation invariance, this is only a function of the difference $C(l,l') = C(l-l')$. For simplicity, set the external field $h = 0$. Also, assume that $l' > l$ (as we'll see, this is *time-ordering* of the correlation function). In terms of the transfer matrix, it is:

$$C(l-l') = \frac{1}{Z} \mathrm{tr} \left(\mathbf{T}^{M_\tau - l'} \mathbf{Z} \mathbf{T}^{l'-l} \mathbf{Z} \mathbf{T}^{l} \right). \tag{16}$$

Notice that there is only *one* operator $\mathbf{Z} = \boldsymbol{\sigma}^z$ here; it is the matrix

$$\mathbf{Z}_{ss'} = \delta_{ss'} s \ .$$

All the information about the index l, l' is encoded in the location in the trace.

Let's evaluate this trace in the basis of $\mathbf{T}$ eigenstates. When $h = 0$, we have $\mathbf{T} = e^K \mathbb{1} + e^{-K} \mathbf{X}$, so these are $\mathbf{X}$ eigenstates:

$$\mathbf{T}|\rightarrow\rangle = \lambda_+ |\rightarrow\rangle, \qquad \mathbf{T}|\leftarrow\rangle = \lambda_- |\rightarrow\rangle \ .$$

Here $|\rightarrow\rangle \equiv \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle)$.

In this basis

$$\langle \alpha | \mathbf{Z} | \beta \rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_{\alpha\beta}, \quad \alpha, \beta = \rightarrow \text{ or } \leftarrow .$$

So the trace (aka path integral) has two terms: one where the system spends $l' - l$ steps in the state $|\rightarrow\rangle$ (and the rest in $|\leftarrow\rangle$), and one where it spends $l' - l$ steps in the state $|\rightarrow\rangle$. The result (if we take $M_\tau \to \infty$ holding fixed $l' - l$) is

$$C(l'-l) = \frac{\lambda_+^{M_\tau - l' + l} \lambda_-^{l'-l} + \lambda_-^{M_\tau - l' + l} \lambda_+^{l'-l}}{\lambda_+^{M_\tau} + \lambda_-^{M_\tau}} \overset{M_\tau \to \infty}{\to} \tanh^{l'-l} K \ . \tag{17}$$

You should think of the insertions as

$$s_l = \mathbf{Z}(\tau), \ \tau = \Delta\tau l.$$

So what we've just computed is

$$C(\tau) = \langle \mathbf{Z}(\tau)\mathbf{Z}(0)\rangle = \tanh^l K = e^{-|\tau|/\xi} \tag{18}$$

where the *correlation time* ξ satisfies

$$\frac{1}{\xi} = \frac{1}{\Delta\tau} \ln\coth K \ . \tag{19}$$

Notice that this is the same as our formula for the gap, Δ, in (13).[14] This connection between the correlation length in euclidean time and the energy gap is general and important.

For large K, ξ is much bigger than the lattice spacing:

$$\frac{\xi}{\Delta\tau} \overset{K\gg 1}{\simeq} \frac{1}{2}e^{2K} \gg 1.$$

This is the limit we had to take to make the euclidean time continuous.

Notice that if we had taken $l < l'$ instead, we would have found the same answer with $l' - l$ replaced by $l - l'$.

[End of Lecture 4]

[14] Seeing this requires the following cool hyperbolic trig fact:

$$\text{If } \ e^{-2K} = \tanh X \ \text{ then } \ e^{-2X} = \tanh K \tag{20}$$

(*i.e.* this equation is 'self-dual') which follows from algebra. Here (13) says $X = \frac{\Delta}{TM_\tau} = \Delta\tau\Delta$ while (19) says $X = \Delta\tau/\xi$. Actually this relation (20) can be made manifestly symmetric by writing it as

$$1 = \sinh 2X \sinh 2K \ .$$

(You may notice that this is the same combination that appears in the Kramers-Wannier self-duality condition.) I don't know a slick way to show this, but if you just solve this quadratic equation for e^{-2K} and boil it enough, you'll find $\tanh X$.

Continuum scaling limit and universality

[Sachdev, 2d ed §5.5.1, 5.5.2] Now we are going to grapple with the term 'universal'. Let's think about the Ising chain some more. We'll regard $M_\tau \Delta\tau$ as a physical quantity, the proper length of the chain. We'd like to take a continuum limit, where $M_\tau \to \infty$ or $\Delta\tau \to 0$ or maybe both. Such a limit is useful if $\xi \gg \Delta\tau$. This decides how we should scale K, h in the limit. More explicitly, here is the prescription: Hold fixed physical quantities (*i.e.* eliminate the quantities on the RHS of these expressions in favor of those on the LHS):

$$\begin{aligned} \text{the correlation length,} \quad & \xi &&\simeq \Delta\tau \frac{1}{2} e^{2K}, \\ \text{the length of the chain,} \quad & L_\tau &&= \Delta\tau M_\tau, \\ \text{physical separations between operators,} \quad & \tau &&= (l - l')\Delta\tau, \\ \text{the applied field in the quantum system,} \quad & \bar{h} &&= h/\Delta\tau. \end{aligned} \tag{21}$$

while taking $\Delta\tau \to 0, K \to \infty, M_\tau \to \infty$.

What physics of the various chains will agree? Certainly only quantities that don't depend explicitly on the lattice spacing; such quantities are called *universal*.

Consider the thermal free energy of the single quantum spin (12)[15]: The energy spectrum of our spin is $E_\pm = E_0 \pm \sqrt{(\Delta/2)^2 + \bar{h}^2}$, which means

$$F = -T \log Z_Q = E_0 - T \ln \left(2 \cosh \left(\beta \sqrt{(\Delta/2)^2 + \bar{h}^2} \right) \right)$$

(just evaluate the trace in the energy eigenbasis). In fact, this is just the behavior of the ising chain partition function in the scaling limit (21), since, in the limit (10) becomes

$$\lambda_\pm \simeq \sqrt{\frac{2\xi}{\Delta\tau}} \left(1 \pm \frac{\Delta\tau}{2\xi} \sqrt{1 + 4\bar{h}^2 \xi^2} \right)$$

and so in the scaling limit (21)

$$F \simeq L_\tau \left(\underbrace{-\frac{K}{\Delta\tau}}_{\text{cutoff-dependent vac. energy}} \quad -\frac{1}{L_\tau} \ln \left(2 \cosh \frac{L_\tau}{2} \sqrt{\xi^{-2} + 4\bar{h}^2} \right) \right) ,$$

which is the same (up to an additive constant) as the quantum formula under the previously-made identifications $T = \frac{1}{L_\tau}$, $\xi^{-1} = \Delta$.

We can also use the quantum system to compute the correlation functions of the classical chain in the scaling limit (17). They are *time-ordered* correlation functions:

$$C(\tau_1 - \tau_2) = Z_Q^{-1} \mathrm{tr} e^{-\beta \mathbf{H}} \left(\theta(\tau_1 - \tau_2) \mathbf{Z}(\tau_1) \mathbf{Z}(\tau_2) + \theta(\tau_2 - \tau_1) \mathbf{Z}(\tau_2) \mathbf{Z}(\tau_1) \right)$$

[15] [Sachdev, 1st ed p. 19, 2d ed p. 73]

where

$$\mathbf{Z}(\tau) \equiv e^{\mathbf{H}\tau}\mathbf{Z}e^{-\mathbf{H}\tau}.$$

This time-ordering is just the fact that we had to decide whether l' or l was bigger in (16).

For example, consider what happens to this when $T \to 0$. Then (inserting $\mathbb{1} = \sum_n |n\rangle\langle n|$, in an energy eigenbasis $\mathbf{H}|n\rangle = E_n|n\rangle$),

$$C(\tau)|_{T=0} = \sum_n |\langle 0|\mathbf{Z}|n\rangle|^2 e^{-(E_n - E_0)|\tau|}$$

where the $|\tau|$ is taking care of the time-ordering. This is a *spectral representation* of the correlator. For large τ, the contribution of $|n\rangle$ is exponentially suppressed by its energy, so the sum is approximated well by the lowest energy state for which the matrix element is nonzero. Assuming this is the first excited state (which in our two-state system it has no choice!), we have

$$C(\tau)|_{T=0} \overset{\tau\to\infty}{\simeq} e^{-\tau/\xi}, \quad \xi = 1/\Delta,$$

where Δ is the energy gap.

In these senses, the quantum theory of a single qbit is *the* universal theory of the Ising chain. For example, if we began with a chain that had in addition next-nearest-neighbor interactions, $\Delta H_c = K' \sum_j s(j)s(j+2)$, we could redo the procedure above. The scaling limit would not be exactly the same; we would have to scale K' somehow (it would also have to grow in the limit). But we would find the same 2-state quantum system, and when expressed in terms of physical variables, the $\Delta\tau$-independent terms in F would be identical, as would the form of the correlation functions, which is

$$C(\tau) = \langle \mathbf{Z}(\tau)\mathbf{Z}(0)\rangle = \frac{e^{-|\tau|/\xi} + e^{-(L_\tau - |\tau|)/\xi}}{1 + e^{-L_\tau/\xi}}.$$

(Note that in this expression we did not assume $|\tau| \ll L_\tau$ as we did before in (18), to which this reduces in that limit.)

2.2 Transverse-Field Ising Model

Whether or not you liked the derivation above of its relation to the euclidean statistical mechanics Ising model, we are going to study the quantum system whose hamiltonian is

$$\mathbf{H}_{\rm TFIM} = -J\left(g\sum_j \mathbf{X}_j + \sum_{\langle jl\rangle}\mathbf{Z}_j\mathbf{Z}_l\right). \tag{22}$$

Some of the things we say next will be true in one or more spatial dimensions.

Notice that J has units of energy; we could choose units where it's 1. In 1d (or on bipartite lattices), the sign of J does not matter for determining what state of matter we realize: if $J<0$, we can relabel our operators: $\tilde{Z}_j = (-1)^j\mathbf{Z}_j$ and turn an antiferromagnetic interaction into a ferromagnetic one. So let's assume $g, J > 0$.

This model is interesting because of the competition between the two terms: the $\mathbf{X}_j$ term wants *each* spin (independently of any others) to be in the state $|\rightarrow\rangle_j$ which satisfies

$$\mathbf{X}_j|\rightarrow\rangle_j = |\rightarrow\rangle_j. \quad |\rightarrow\rangle_j = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_j + |\downarrow\rangle_j\right).$$

In conflict with this are the desires of $-\mathbf{Z}_j\mathbf{Z}_{j+1}$, which is made happy (*i.e.* smaller) by the more cooperative states $|\uparrow_j\uparrow_{j+1}\rangle$, or $|\downarrow_j\downarrow_{j+1}\rangle$. In fact, it would be just as happy about any linear combination of these $a|\uparrow_j\uparrow_{j+1}\rangle + b|\downarrow_j\downarrow_{j+1}\rangle$ and we'll come back to this point.

Another model which looks like it might have some form of competition is

$$\mathbf{H}_{\rm boring} = \cos\theta\sum_j \mathbf{Z}_j + \sin\theta\sum_j \mathbf{X}_j\ , \quad \theta\in[0,\frac{\pi}{2}]$$

Why is this one boring? Notice that we can continuously interpolate between the states enjoyed by these two terms: the groundstate of $\mathbf{H}_1 = \cos\theta\mathbf{Z} + \sin\theta\mathbf{X}$ is

$$|\theta\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + \sin\frac{\theta}{2}|\downarrow\rangle$$

– as we vary θ from 0 to $\pi/2$ we just smoothly rotate from $|\uparrow_z\rangle$ to $|\uparrow_x\rangle$.

How do we know the same thing can't happen in the transverse-field Ising chain? Symmetry. We've already seen that the Ising model has a $G=\mathbb{Z}_2$ symmetry which acts by $\mathbf{Z}_j\rightarrow\mathbf{S}\mathbf{Z}_j\mathbf{S}^\dagger = -\mathbf{Z}_j, \mathbf{X}_j\rightarrow\mathbf{S}\mathbf{X}_j\mathbf{S}^\dagger = +\mathbf{X}_j$, where the unitary $\mathbf{S}$ commutes with $\mathbf{H}_{\rm TFIM}$: $\mathbf{S}\mathbf{H}_{\rm TFIM}\mathbf{S}^\dagger = \mathbf{H}_{TFIM}$. The difference with $\mathbf{H}_{\rm boring}$ is that $\mathbf{H}_{\rm TFIM}$ has two phases in which G is realized differently on the groundstate.

$\boxed{g=\infty}$: First, let's take g so big that we may ignore the ZZ ferromagnetic term, so

$$\mathbf{H}_{g\rightarrow\infty} = -\sum_j \mathbf{X}_j\ .$$

(The basic idea of this discussion will apply in any dimension, on any lattice.) Since all terms commute, the groundstate is the simultaneous groundstate of each term:

$$\mathbf{X}_j|\text{gs}\rangle = +|\text{gs}\rangle, \quad \forall j, \qquad \Longrightarrow \ |\text{gs}\rangle = \otimes_j|\rightarrow\rangle_j.$$

Notice that this state *preserves the symmetry* in the sense that $\mathbf{S}|\text{gs}\rangle = |\text{gs}\rangle$. Such a symmetry-preserving groundstate is called a *paramagnet*.

$\boxed{g = 0}$: Begin with $g = 0$.

$$\mathbf{H}_0 = -J\sum_j \mathbf{Z}_j\mathbf{Z}_{j+1}$$

has groundstates

$$|+\rangle \equiv |\uparrow\uparrow\cdots\uparrow\rangle, \ |-\rangle \equiv |\downarrow\downarrow\cdots\downarrow\rangle,$$

or any linear combination. Note that the states $|\pm\rangle$ are not symmetric: $\mathbf{S}|\pm\rangle = |\mp\rangle$, and so we are tempted to declare that the symmetry is broken by the groundstate.

You will notice, however, that the states

$$|\ \text{  }_{\pm}\rangle \equiv \frac{1}{\sqrt{2}}\left(|+\rangle \pm |-\rangle\right)$$

are symmetric – they are $\mathbf{S}$ eigenstates, so $\mathbf{S}$ maps them to themselves up to a phase. It gets worse: In fact, in finite volume (finite number of sites of our chain), with $g \neq 0$, $|+\rangle$ and $|-\rangle$ are *not* eigenstates, and $|\ \text{  }_{+}\rangle$ is the groundstate. **BUT:**

1. The two states $|+\rangle$ and $|-\rangle$ only mix at order N in perturbation theory in g, since we have to flip all N spins using the perturbing hamiltonian $\Delta\mathbf{H} = -gJ\sum_j \mathbf{X}_j$ to get from one to the other. The tunneling amplitude is therefore

$$T \sim g^N\langle - |\mathbf{X}_1\mathbf{X}_2\cdots\mathbf{X}_N|+\rangle \overset{N\to\infty}{\to} 0.$$

2. There's a reason for the symbol I used to denote the symmetric states: at large N, these 'cat states' are superpositions of macroscopically distinct quantum states. Such things don't happen, because of decoherence: if even a single dust particle in the room measures the spin of a single one of the spins, it measures the value of the whole chain. In general, this happens very rapidly.

3. Imagine we add a small symmetry-breaking perturbation: $\Delta\mathbf{H} = -\sum_j h\mathbf{Z}_j$; this splits the degeneracy between $|+\rangle$ and $|-\rangle$. If $h > 0$, $|+\rangle$ is for sure the groundstate. Consider preparing the system with a tiny $h > 0$ and then setting $h = 0$ after it settles down. If we do this to a finite system, $N < \infty$, it will be in an excited state of the $h = 0$ Hamiltonian, since $|+\rangle$ will not be stationary (it will have a nonzero amplitude to

tunnel into $|-\rangle$). But if we take the thermodynamic limit *before* taking $h \to 0$, it will stay in the state we put it in with the 'training field' h. So beware that there is a singularity of our expressions (with physical significance) that means that the limits do not commute:

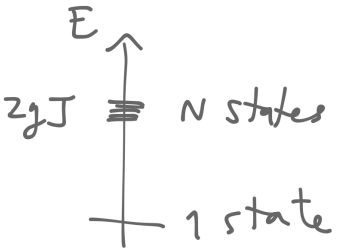
$$\lim_{N\to\infty}\lim_{h\to 0} Z \neq \lim_{h\to 0}\lim_{N\to\infty} Z.$$

The physical one is to take the thermodynamic limit first.

The conclusion of this brief discussion is that spontaneous symmetry breaking actually happens in the $N \to \infty$ limit. At finite N, $|+\rangle$ and $|-\rangle$ are approximate eigenstates which become a better approximation as $N \to \infty$.

This state of a $\mathbb{Z}_2$-symmetric system which spontaneously breaks the $\mathbb{Z}_2$ symmetry is called a ferromagnet.

So the crucial idea I want to convey here is that there must be a sharp phase transition at some finite g: the situation cannot *continuously* vary from one unique, symmetric groundstate $\mathbf{S}|\mathrm{gs}_{g\ll 1}\rangle = |\mathrm{gs}_{g\ll 1}\rangle$ to two symmetry-breaking groundstates: $\mathbf{S}|\mathrm{gs}_{\pm}\rangle = |\mathrm{gs}_{\mp}\rangle$. We'll make this statement more precise when we discuss the notion of long-range order. First, let's see what happens when we try to vary the coupling away from the extreme points.



$\boxed{g \gg 1}$ An excited state of the paramagnet, deep in the phase, is achieved by flipping one spin. With $\mathbf{H} = \mathbf{H}_\infty = -gJ\sum_j \mathbf{X}_j$, this costs energy $2gJ$ above the groundstate. There are N such states, labelled by which spin we flipped:

$$|n\rangle \equiv |\to\cdots\to \underbrace{\leftarrow}_{nth\ \text{site}} \to\cdots\rangle, \quad (\mathbf{H}_\infty - E_0)\,|n\rangle = 2gJ|n\rangle,\ \ \forall n$$

When g is not infinite, we can learn a lot from (1st order) degenerate perturbation theory in the ferrmomagnetic term. The key information is the matrix elements of the perturbing hamiltonian between the degenerate manifold of states. Using the fact that $\mathbf{Z}_j|\to\rangle = |\leftarrow\rangle$, so,

$$\mathbf{Z}_j\mathbf{Z}_{j+1}|\to_j\leftarrow_{j+1}\rangle = |\leftarrow_j\to_{j+1}\rangle$$

$$\langle n\pm 1|\sum_j \mathbf{Z}_j\mathbf{Z}_{j+1}|n\rangle = 1,$$

the ferromagnetic term hops the spin flip by one site. Within the degenerate subspace, it acts as

$$\mathbf{H}_{\text{eff}}|n\rangle = -J\left(|n+1\rangle + |n-1\rangle\right) + (E_0 + 2gJ)\,|n\rangle.$$

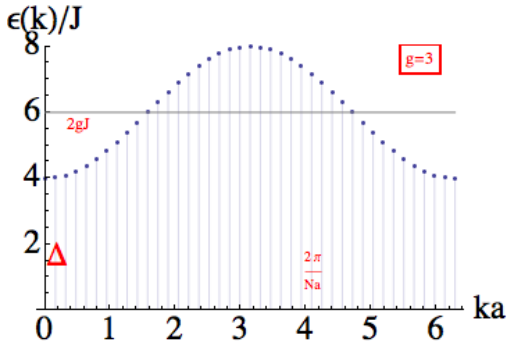
It is a kinetic, or 'hopping' term for the spin flip.

Let's see what this does to the spectrum. Assume periodic boundary conditions and N sites total. Again this is a translation invariant problem (in fact the same one, basically), which we solve by Fourer transform:

$$|n\rangle \equiv \frac{1}{\sqrt{N}} \sum_j e^{-\mathbf{i}kx_j} |k\rangle, \quad \begin{cases} x_j \equiv ja, \\ k = \frac{2\pi m}{Na}, \quad m = 1..N \end{cases}$$

On the momentum states, we have

$$(H - E_0)\,|k\rangle = (-2J\cos ka + 2gJ)\,|k\rangle.$$



The dispersion of these spinon particles is

$$\epsilon(k) = 2J(g - \cos ka) \overset{k\to 0}{\sim} \Delta + J(ka)^2 \qquad (23)$$

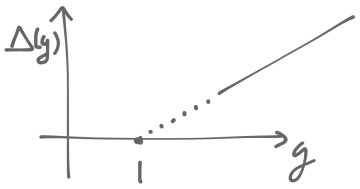
with $\Delta = 2J(g-1)$ – there is an energy gap (notice that Δ does not depend on system size). So these are *massive* particles, with dispersion $\epsilon = \Delta + \frac{k^2}{2M} + ...$ where Δ is the energy to create one at rest (notice that the rest energy is not related to its inertial mass $M^{-1} = 2Ja^2$).

A particle at j is created by the creation operator $\mathbf{Z}_j$:

$$|n\rangle = \mathbf{Z}_n |\mathrm{gs}_\infty\rangle.$$

And it is annihilated by the annihilation operator $\mathbf{Z}_j$ – you can't have two spin flips at the same location! These particles are their own antiparticles.

The *number* of such particles is counted by the operator $\sum_j (-\mathbf{X}_j)$. The number of particles is only conserved modulo two, however.



What happens as g gets smaller? The gap to creating a spin flip at large g looks like $2J(g-1)$. If we take this formula seriously, we predict that at $g = 1$ it costs zero energy to create spin flips: they should condense in the vacuum. Condensing spin flips means that the spins point in all directions, and the state is paramagnetic. (We shouldn't take it seriously because it's just first order in perturbation theory, but it turns out to be exactly right.)

It's possible to develop some more evidence for this picture and understanding of the physics of the paramagnetic phase in the Ising chain by doing more perturbation theory, and including states with two spin flips. Notice that for a state with two spin-flip particles, the total momentum k no longer uniquely determines the energy, since the two spin-flips can have a relative momentum; this means that there is a *two-particle continuum* of states, once we have enough energy to make two spin flips. For more on this, see *e.g.* Sachdev (2d ed) §5.2.2. In particular the two spin-flips can form boundstates, which means the two-particle continuum is actually slightly below 2Δ. [End of Lecture 5]

$\boxed{g \ll 1}$ Now let's consider excitations of the ferromagnet, about the state $|+\rangle = |\uparrow\uparrow\cdots\uparrow\rangle$. It is an eigenstate of $\mathbf{H}_0 = -J\sum_j \mathbf{Z}_j\mathbf{Z}_{j+1}$ and its (groundstate) energy is $E_0 = -JN$. We can make an excitation by flipping one spin:

$$|\cdots\uparrow\ \uparrow\ \uparrow\cdot\downarrow\cdot\uparrow\ \uparrow\ \uparrow\cdots\rangle$$

This makes two bonds unhappy, and costs $2J + 2J = 4J$. But once we make it there are many such states: the hamiltonian is the same amount of unhappy if we also flip the next one.

$$|\cdots\uparrow\ \uparrow\ \uparrow\cdot\downarrow\ \downarrow\cdot\uparrow\ \uparrow\cdots\rangle$$

The actual elementary excitation is a *domain wall* (or *kink*), which only costs $2J$. The domain wall should be regarded as living between the sites. It is not entirely a local object, since with periodic boundary conditions, we must make two, which can then move independently. To create two of them far apart, we must change the state of many spins.

At $g = 0$ the domain walls are *localized* in the sense that a domain wall at a fixed position is an energy eigenstate (just like the spinons at $g = \infty$), with the same energy for any position. But now the paramagnetic term $-\sum_j g\mathbf{X}_j$ is a kinetic term for the domain walls:

$$\mathbf{X}_{j+1}\underbrace{|\cdots\uparrow\uparrow\uparrow_j\cdot\downarrow_{j+1}\downarrow\downarrow\cdots\rangle}_{\bar{j}} = \underbrace{|\cdots\uparrow\uparrow\uparrow_j\uparrow_{j+1}\cdot\downarrow_{j+2}\downarrow\cdots\rangle}_{=|\bar{j}+1\rangle}.$$

Just like in our $g \gg 1$ discussion, acting on a state with an even number of well-separated domain walls

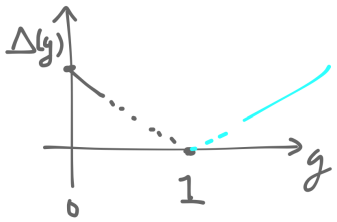
$$(H_{\text{eff}} - E_0)\,|\bar{j}\rangle = -gJ\,(|\bar{j}+1\rangle + |\bar{j}-1\rangle) + 2J|\bar{j}\rangle$$

where the diagonal term is the energy cost of one domain wall at rest. Again this is diagonalized in k-space with energy

$$\epsilon_{\text{one dwall}}(k) = 2J(1 - g\cos ka)$$

Again, this calculation is almost ridiculously successful at predicting the location of the phase transition:

$$\Delta_{DW} = 2J(1-g) \overset{g\to 1}{\to} 0.$$



Notice that although our discussion of the paramagnetic state $g \gg 1$ can be applied in any $d \geq 1$, the physics of domain walls is very dimension-dependent.

Interpretation of the stability of the SSB state in terms of domain walls:

If at finite N, with periodic boundary conditions, we prepare the system in the state $|+\rangle$, tunneling to $|-\rangle$ requires creation of a pair of domain walls $\Delta E = 4J$, which then *move* all the way around the circle, giving the tunneling rate

$$\prod_{j=1}^{N} \left(\frac{\langle \bar{j}+1 | \mathbf{H}_{\text{eff}} | \bar{j} \rangle}{\Delta E} \right) \sim \frac{(gJ)^N}{J^N} \sim g^N \sim e^{-N \log \frac{1}{g}}.$$

(For $g < 1$, $\log \frac{1}{g} > 0$.) The tunneling rate goes like e^{-N} – it is exponentially small in the system size.

2.2.1 Duality between spin flips and domain walls

The discussion we've just made of the small-g physics has a lot in common with the large-g physics. More quantitatively, the dispersion relation $\epsilon_{\text{one dwall}}(k)$ for a single domain wall looks nearly the same as that of one spin flip (23). In fact they are mapped to each other by the replacement

$$g \to \frac{1}{g}, \quad J \to Jg. \tag{24}$$

Notice that this takes small g (weak coupling of domain walls, strong coupling of spin flips) to large g (strong coupling of domain walls, weak coupling of spin flips).

In fact, there is a change of variables that (nearly) interchanges the two sides of the phase diagram. Suppose the system is on an interval – open boundaries – the chain just stops at $j = 1$ and $j = N$. (We do this to avoid the constraint of an even number of domain walls.) We can specify a basis state in the Z-basis by the direction (up or down along Z) of the first spin and the locations of domain walls.

Consider the operator, diagonal in this basis, which measures whether there is a domain wall between j and $j + 1$:

$$\boldsymbol{\tau}^x_{\bar{j}} \equiv \mathbf{Z}_{\bar{j}-\frac{1}{2}}\mathbf{Z}_{\bar{j}+\frac{1}{2}} = \begin{cases} +1, & \text{if } z_{\bar{j}-\frac{1}{2}} = z_{\bar{j}+\frac{1}{2}} \\ -1, & \text{if } z_{\bar{j}-\frac{1}{2}} = -z_{\bar{j}+\frac{1}{2}} \end{cases} = (-1)^{\text{disagreement}}.$$

Notice that $\boldsymbol{\tau}^2_{\bar{j}} = 1, \boldsymbol{\tau}^\dagger_{\bar{j}} = \boldsymbol{\tau}_{\bar{j}}$. Similarly, consider the operator that *creates* a domain wall at $\bar{j}$:

$$\boldsymbol{\tau}^z_{\bar{j}} \equiv \mathbf{X}_{\bar{j}+\frac{1}{2}}\mathbf{X}_{\bar{j}+\frac{3}{2}} \cdots = \prod_{j>\bar{j}} \mathbf{X}_j.$$

This operator flips all the spins to the right of the link in question (and fixes our reference first spin). It, too, is hermitian and squares to one. Finally, notice that

$$\boldsymbol{\tau}^z_{\bar{j}}\boldsymbol{\tau}^x_{\bar{j}'} = (-1)^{\delta_{\bar{j}\bar{j}'}}\boldsymbol{\tau}^x_{\bar{j}'}\boldsymbol{\tau}^z_{\bar{j}}$$

just like $\mathbf{Z}$ and $\mathbf{X}$ (since when $\bar{j} = \bar{j}'$, they contain a single $\mathbf{Z}$ and $\mathbf{X}$ at the same site). The domain walls can be represented in terms of two-state systems living on the links.[16]

Notice that the inverse of the map from $\mathbf{X}, \mathbf{Z}$ to $\boldsymbol{\tau}^x, \boldsymbol{\tau}^z$ is

$$\mathbf{X}_j = \boldsymbol{\tau}^z_{j-\frac{1}{2}}\boldsymbol{\tau}^z_{j+\frac{1}{2}} \ .$$

(The right hand side is an inefficient way to flip a single spin at j: namely, flip all the spins right of $j - 1$, and then flip back all the spins to the right of j.)

[16] Note that in lecture I reversed the names of τ^x and τ^z; I think this way is a little better.

So the 1d TFIM hamiltonian in bulk is

$$\begin{aligned}\mathbf{H}_{\text{TFIM}} &= -J\sum_j \left(g\mathbf{X}_j + \mathbf{Z}_j\mathbf{Z}_{j+1}\right) \\ &= -J\sum_{\bar{j}} \left(g\boldsymbol{\tau}^z_{\bar{j}}\boldsymbol{\tau}^z_{\bar{j}+1} + \boldsymbol{\tau}^x_{\bar{j}}\right). \end{aligned} \tag{25}$$

This is the TFIM hamiltonian again with $\mathbf{Z} \to \boldsymbol{\tau}^z$ and $\mathbf{X} \to \boldsymbol{\tau}^x$ and the couplings mapped by (24).

This is in fact the same map as Kramers-Wannier duality (or rather it is mapped to it by the quantum-to-classical map). As K&W argued, if there is a single phase transition it must occur at the self-dual point $g = 1$.

Notice that the paramagnetic (disordered) groundstate of the original system is a *condensate* of domain walls, in the following sense. The operator that creates a domain wall has an expectation value:

$$\langle \boldsymbol{\tau}^z_{\bar{j}} \rangle = \langle \text{gs}_{g=\infty} | \boldsymbol{\tau}^z_{\bar{j}} | \text{gs}_{g=\infty} \rangle = \langle \text{gs}_{g=\infty} | \prod_{j>\bar{j}} \mathbf{X}_j | \text{gs}_{g=\infty} \rangle = 1 \quad \forall \bar{j} \ .$$

(For $g \in (1, \infty)$, this expectation value is less than one but nonzero, just like how $|\langle \mathbf{Z} \rangle|$ decreases from 1 as g grows from zero.) Although there is a condensate, there is no order, in the sense that an expectation value of $\mathbf{X}$ does not break any symmetry of $\mathbf{H}_{\text{TFIM}}$. (There is another state where $\langle \boldsymbol{\tau}^x \rangle = -1$, namely the one where all the spins are pointing to the *left*. But (at large g) it's a very high-energy state.)

An important point (and the reason 'duality' is a dangerous word): the two sides of the phase diagram are *not the same*. On one side there are two groundstates related by the broken symmetry, on the other side there is a unique symmetric groundstate. That's how we knew there had to be a phase transition! I will say more about this mismatch.

Open boundaries. Let us make sure we can reproduce the correct number of groundstates in the two phases. To get this right, we have to be precise about the endpoint conditions. Let's study the case where we have N sites in a row; the first and last sites have only one neighbor. The Hamiltonian is

$$\mathbf{H}_{\text{TFIM}} = -J\left(\sum_{j=1}^{N-1} \left(g\mathbf{X}_j + \mathbf{Z}_j\mathbf{Z}_{j+1}\right) + g\mathbf{X}_N\right) \ .$$

The duality map is

$$\mathbf{Z}_j\mathbf{Z}_{j+1} = \boldsymbol{\tau}^x_{j+\frac{1}{2}}, j = 1, 2 \ldots N-1, \quad \mathbf{X}_j = \boldsymbol{\tau}^z_{j-\frac{1}{2}}\boldsymbol{\tau}^z_{j+\frac{1}{2}}, \quad j = 1 \ldots N \quad .$$

In terms of the domain-wall variables, the hamiltonian is

$$\mathbf{H}_{\text{TFIM}} = -Jg\left(\sum_{j=1}^{N-1}\left(\boldsymbol{\tau}^z_{j-\frac{1}{2}}\boldsymbol{\tau}^z_{j+\frac{1}{2}} + \frac{1}{g}\boldsymbol{\tau}^x_{j+\frac{1}{2}}\right) + \boldsymbol{\tau}^z_{N-\frac{1}{2}}\boldsymbol{\tau}^z_{N+\frac{1}{2}}\right) \ .$$

But now there are two special cases:
• $\boldsymbol{\tau}^z_{N+\frac{1}{2}} = 1$: this operator flips all the spins with $j > N$; but there are no spins with $j > N$. So it is the identity operator.
• $\boldsymbol{\tau}^x_{\frac{1}{2}}$: this operator measures whether or not there is a domain wall between $j = 1$ and $j = 0$, $\boldsymbol{\tau}^x_{\frac{1}{2}} = \mathbf{Z}_0\mathbf{Z}_1$. But there is no spin at $j = 0$. One way to think about this is to put a "ghost spin" at $j = 0$ which is always in the state $\mathbf{Z}_0 = 1$. So $\boldsymbol{\tau}^x_{\frac{1}{2}} = \mathbf{Z}_1$: it measures the value of our reference spin.

At $g = 0$: $\mathbf{H}_{g=0} = -J\sum_{j=2}^{N} \boldsymbol{\tau}^x_{j-\frac{1}{2}}$ and the groundstate is $\boldsymbol{\tau}^x_{j-\frac{1}{2}} = 1$ for $j = 2...N$. But $\boldsymbol{\tau}^x_{\frac{1}{2}}$ does not appear, so there are two degenerate groundstates, eigenstates of $\boldsymbol{\tau}^x_{\frac{1}{2}}$ with eigenvalue $\pm$, which are just $|\pm\rangle$, the states with no domain walls: all the other spins agree with the first one in a state where $\boldsymbol{\tau}^x_{j>1} = 1$.

At $g = \infty$, $\mathbf{H}_{g=\infty} = -Jg\left(\sum_{j=1}^{N-1} \boldsymbol{\tau}^z_{j-\frac{1}{2}}\boldsymbol{\tau}^z_{j+\frac{1}{2}} + \boldsymbol{\tau}^z_{N-\frac{1}{2}}\right)$. The first term requires agreement between $\boldsymbol{\tau}^z_{j-\frac{1}{2}} = \boldsymbol{\tau}^z_{j+\frac{1}{2}}$ for $j = 1...N-1$, but the second term requires $\boldsymbol{\tau}^z_{N-\frac{1}{2}} = 1$. So all the others must be $+1$, too. This is the unique groundstate of the paramagnet.

Comment: In the notation of our lattice model (14), the critical curve between the two phases is at

$$\sinh(2K_x)\sinh(2K) = 1 \ .$$

This relationship follows by *Kramers-Wannier duality*. In the scaling limit, $K_x \sim \Delta\tau \ll 1, K \gg 1$. So the critical condition is

$$1 = \sinh(2K_x)\sinh(2K) \simeq 2K_x\frac{e^{2K}}{2} = J\Delta\tau\frac{1}{Jg\Delta\tau} = \frac{1}{g}.$$

This determines the location of the quantum phase transition, in agreement with all the other stuff we've said about it.

2.2.2 Mean field theory

So far we have used perturbation theory about $g = 0$ and $g = \infty$ to delicately inch our way toward the interior of the phase diagram of the TFIM in one dimension (22). Here we take a plunge and try to guess the groundstate for all g. The nice thing about trying to guess the groundstate is the Rayleigh-Schrödinger hedge: the energy expectation in any state is an upper bound for the groundstate energy; minimizing the energy within a class of guesses is called a 'variational approach'.

The name for the particular guess we'll make is 'mean field theory', which means that we completely ignore entanglement between different sites, and suppose that the state is a product state

$$|MFT\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \cdots |\psi_j\rangle \cdots ...$$

If we further assume translational invariance then the state at every site is the same and we have one bloch sphere to minimize over for each g:

$$|\check{n}\rangle = \otimes_j |\uparrow_{\check{n}}\rangle_j = \otimes_j \left(\cos\frac{\theta}{2} e^{\mathbf{i}\varphi/2} |\rightarrow\rangle + \sin\frac{\theta}{2} e^{-\mathbf{i}\varphi/2} |\leftarrow\rangle \right)_j .$$

(Here θ is the angle $\check{n}$ makes with the x axis, and φ is the azimuthal angle in the yz plane, from the z-axis.) To evaluate the energy expectation in this state, we only need to know single-qbit expectations:

$$\langle \uparrow_{\check{n}} |\mathbf{X}| \uparrow_{\check{n}} \rangle = \cos\theta, \quad \langle \uparrow_{\check{n}} |\mathbf{Z}| \uparrow_{\check{n}} \rangle = \sin\theta\cos\varphi.$$

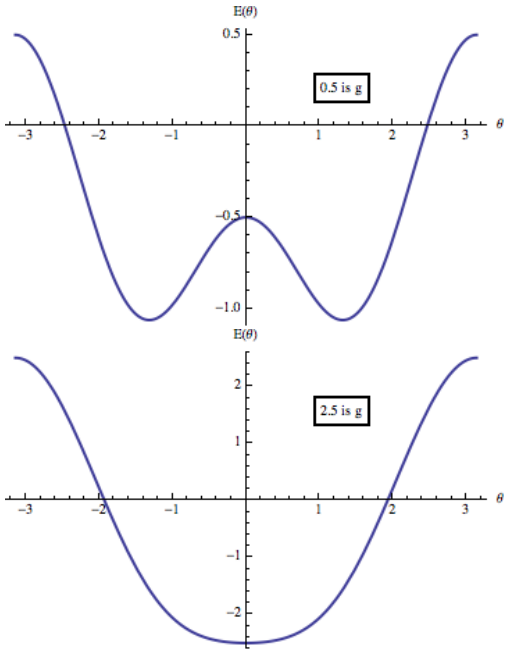
So the energy expectation is

$$E(\theta, \varphi) = -NJ \left(\sin^2\theta \cos^2\varphi + g\cos\theta \right).$$

This is extremized when $\varphi = 0, \pi$ and when

$$0 = \partial_\theta E = NJ \sin\theta \left(2\cos\theta - g \right).$$

Notice that when $\theta = 0$, the two solutions of φ are the same, since the φ coordinate degenerates at the pole. The solutions at $\cos\theta = g/2$ only exist when $g/2 < 1$. In that case they are minima (see the figure) since $\partial_\theta^2 E|_{\cos\theta = g/2} > 0$, while $\partial_\theta^2 E|_{\theta=0} = NJ(g-2)$ is negative for $g < 2$. (Notice that $\varphi = \pi$ can be included by allowing $\theta \in (-\pi, \pi]$, as in the figure.)



So mean field theory predicts a phase transition at $g = 2$, from a state where $\langle \mathbf{Z}_j \rangle = \sin\theta$ to one where $\langle \mathbf{Z} \rangle = 0$. It overestimates the range of the ordered phase because it leaves out fluctuations which tend to destroy the order.

Let's study the behavior near the transition, where θ is small. Then the energy can be approximated by its Taylor expansion

$$E(\theta) \simeq NJ\left(-2 + \frac{g-2}{2}\theta^2 + \frac{1}{4}\theta^4\right)$$

(where I have set $g = g_c = 2$ except in the crucial quadratic term). This has minima at

$$\langle \mathbf{Z}_j \rangle = \sin\theta \simeq \theta = \pm\sqrt{g_c - g}\ . \tag{26}$$

The energy behaves like

$$E_{MFT}(g) = \begin{cases} \frac{3}{4}(g_c - g)^2\ , & g < g_c \\ \\ 0\ , & g \geq g_c \end{cases}$$

Notice that $\partial_g E$ is continuous at the transition. (Recall that the groundstate energy of the quantum system is equal to the free energy of the corresponding stat mech system, so $\partial_g E \propto \partial_T F$ continuous is the same criterion for a continuous transition.) So mean field theory (correctly) predicts a continuous quantum phase transition between the ordered phase and the disordered phase. The location of the transition is wrong (mean field theory overestimates the size of the ordered region because it leaves out lots of order-destroying fluctuations), and so are other properties, such as the exponent in (26), which should be 1/8 instead of 1/2.

[End of Lecture 6]

2.2.3 Correlation functions and long-range order

A useful set of physical (observable! as we'll discuss in a moment) quantities are the correlation functions of the spins:

$$C(r,r') \equiv \langle 0|\mathbf{Z}_r\mathbf{Z}_{r'}|0\rangle \quad .$$

Notice that this is an *equal-time* correlator. If the groundstate is translation invariant, this is only a function of the separation between the spins $C(r,r') = C(r-r')$. (Notice that I am using rs and js to label the positions interchangably, but $r = ja$, where a is the lattice spacing) We can take further advantage of translation symmetry by thinking about the Fourier transform of this quantity, which is called the *static structure factor*:

$$S(q) \equiv \sum_r e^{-\mathbf{i}qr} C(r) \ ;$$

the sum over positions is over $r \in a\{1...N\}$, and (assuming periodic boundary conditions again) the argument $q \in \frac{2\pi}{Na}\{1...N\}$.

Correlations in the quantum Ising chain

At $g = \infty$, all the way in the paramagnetic phase,

$$\langle 0|\mathbf{Z}_j\mathbf{Z}_l|0\rangle|_{g=\infty} = \delta_{jl}$$

– different sites are totally uncorrelated. At finite but large $g \gg 1$ we have instead:

$$\langle 0|\mathbf{Z}_j\mathbf{Z}_l|0\rangle|_{g\gg 1} \overset{|x_j-x_l|\gg a}{\simeq} e^{-|x_j-x_l|/\xi} \tag{27}$$

– the fluctuations of the spin flips communicate between the sites, but only over a short range, because the spin flips are massive.

In contrast, for $g \ll 1$, in the ferromagnetic phase,

$$\langle 0|\mathbf{Z}_j\mathbf{Z}_l|0\rangle|_{g\ll 1} \overset{|x_j-x_l|\gg a}{\simeq} N_0^2(g) \tag{28}$$

where $N_0^2(g) = 1$ for $g = 0$. $N_0 < 1$ for $g > 1$.

More simply, if we train the system appropriately (as we discussed in point 3), then we'll have

$$\langle 0|\mathbf{Z}_j|0\rangle = \pm N_0(g).$$

N_0 is the order parameter, the magnetization. Another way to state the issue about the order of limits of $h \to 0, N \to \infty$: in a finite system the magnetization is zero:

$$\lim_{N\to\infty}\lim_{h\to 0}\langle \mathbf{Z}_i\rangle = 0, \quad \text{but} \quad \lim_{h\to 0}\lim_{N\to\infty}\langle \mathbf{Z}_i\rangle = N_0 \quad .$$

In this sense, the ferromagnet is said to have *long range order*.

OK but here's the simple important point: There's no smooth function of g that can interpolate between (27) and (28). A virtue of the notion of long-range order is it gives sharp distinctions between phases.

2.2.4 Physics of structure factors

Consider the limit where $N \to \infty$ so that the allowed momenta fill in a continuous Brillouin zone $q \in \frac{2\pi}{N}\{1..N\} \to (0, \pi] \simeq (-\pi, \pi]$. So the inverse Fourier transform is

$$C(R) = \int_{-\pi}^{\pi} \frac{dq}{2\pi} S(q) e^{\mathbf{i}qR} \ .$$

When there is long range order,

$$C(R) \overset{R\to\infty}{\sim} N_0^2$$

what does $S(q)$ do? It means a singularity at $q = 0$:

$$S(q) = 2\pi\delta(q)N_0^2 + \text{regular} \tag{29}$$

where 'regular' means terms which are smooth at $q = 0$. This singularity is a sharp signature of the broken symmetry.

We can learn something by considering a spectral representation of $S(q)$, obtained by resolving the identity by energy eigenstates $|n\rangle$:

$$S(q) = \langle 0|\mathbf{Z}_q \mathbf{Z}_{-q}|0\rangle = \sum_n |\langle n|\mathbf{Z}_q|0\rangle|^2 \ . \tag{30}$$

Here I used $\mathbf{Z}_q \equiv \frac{1}{\sqrt{N}} \sum_r e^{-\mathbf{i}qr} \mathbf{Z}_r = \mathbf{Z}_{-q}^\dagger$.

A sum rule: $C(R = 0) = \langle 0|\mathbf{Z}_r \mathbf{Z}_r|0\rangle = \langle 0|0\rangle = 1$. But the (inverse) Fourier transform is

$$1 = C(R = 0) = \int_{-\pi}^{\pi} \frac{dq}{2\pi} S(q) \ .$$

This is some short-distance information which remembers that the objects we're looking at square to one. Given the spectral interpretation, this sum rule means that as we vary parameters, the spectral weight can move around, but it can't go away. In particular, when the system orders as in (29), the $2\pi N_0^2$ contribution at $q = 0$ has to come from somewhere.

Dynamical structure factor

$$C(r, t) \equiv \langle 0|\mathbf{Z}_r(t)\mathbf{Z}_0(0)|0\rangle.$$

This reduces to $C(r)$ if we set $t = 0$. It is useful to think of $\mathbf{Z}_r(t)$ in Heisenberg representation: $\mathbf{Z}_r(t) = e^{+\mathbf{i}\mathbf{H}t}\mathbf{Z}_r e^{-\mathbf{i}\mathbf{H}t}$.

$$S(q,\omega) = \sum_r e^{-\mathbf{i}qr} \int_{-\infty}^{\infty} dt e^{+\mathbf{i}\omega t} C(r,t).$$

Notice that $S(q) = \int \mathrm{d}\omega S(q,\omega)$ (not $S(\omega = 0)$). The spectral representation of the dynamical structure factor is:

$$S(q,\omega) = \sum_n |\langle n|\mathbf{Z}_q|0\rangle|^2 2\pi\delta(\omega + E_0 - E_n) \ . \tag{31}$$

So (30) is obtained by integrating over ω[17] which simply removes the ω dependence.

In words: $S(q,\omega)$ is the number of states of momentum q and energy ω, weighted by (the square of) their overlap with the state engendered upon the groundstate by the spin operator. This formula is crying out to be interpreted in terms of Fermi's Golden Rule.

Before we do that let's include finite temperature, too. Let

$$C(x,t;x't') = \frac{1}{Z}\mathrm{tr} e^{-\mathbf{H}/T}\mathbf{Z}(x,t)\mathbf{Z}(x't')$$

where again $\mathbf{Z}$ are Heisenberg picture operators. Its fourier transform is

$$S(q,\omega) = \int d^d x \int_{-\infty}^{\infty} dt \ e^{\mathbf{i}\omega(t-t') - \mathbf{i}\vec{q}\cdot(\vec{x}-\vec{x}')} C(x,t;x't').$$

Notice that the time integral here is over real time, not euclidean time. To find the spectral representation, again, insert an energy eigenbasis $\mathbb{1} = \sum_n |n\rangle\langle n|$ in between each time evolution operator and each field, and do the space and time integrals. The result is:

$$S(q,\omega) = \frac{2\pi}{ZV} \sum_{n,n'} e^{-E_{n'}/T} \left|\langle n'|\mathbf{Z}(q)|n\rangle\right|^2 \delta\left(\omega - E_n - E_{n'}\right).$$

V is the volume of space.

Now let's interpret this in terms of a transition rate in a scattering experiment, via Fermi's Golden Rule. The initial state is $|n'\rangle$ with probability $p_{n'} = e^{-E_{n'}/T}/Z$ given by the Boltzmann distribution. (In the $T \to 0$ limit, we return to (31), where the initial state is the groundstate with probability 1.) The perturbing hamiltonian is linear in the spin operator $\mathbf{H}_{\text{pert}} \propto \mathbf{Z}(q)$. With these assumptions, the transition rate per unit time to the state $|m\rangle$ is exactly $S(q,\omega)$. This explains the connection between the dynamical structure factor and experiments involving scattering of neutrons: a neutron couples to the spins via its magnetic dipole moment, which couples linearly to the spin operator; ω, q are the change in energy and

[17] with the usual $\frac{1}{2\pi}$ convention: $\int \mathrm{d}\omega S(q,\omega) = S(q)$. Thanks to Sid Parameswaran for alerting me to my 2π problem here. His notes for a related course can be found here. You should trust his signs over mine!

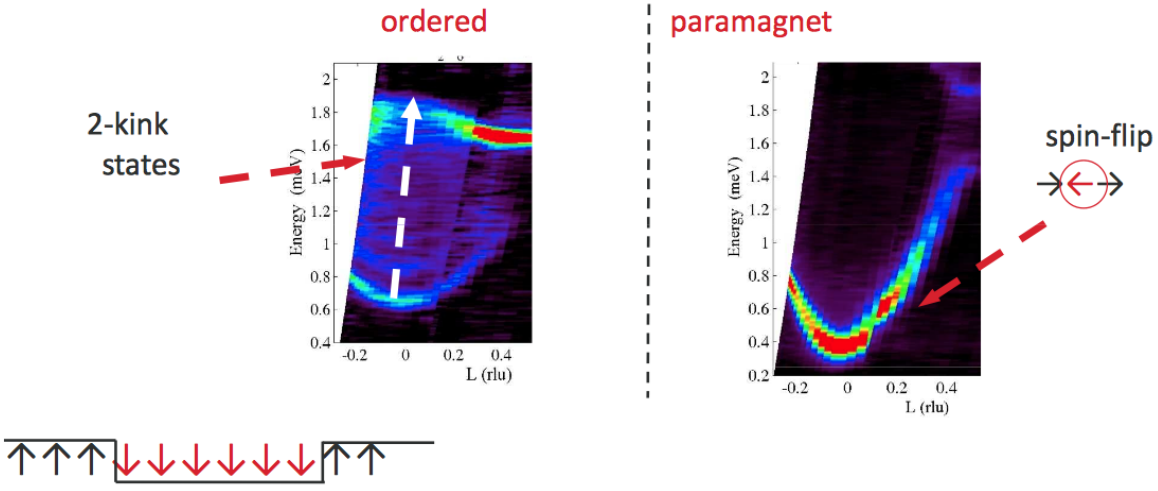
momentum of the system. The transition probability (at zero temperature) is proportional to something like:

$$|\langle n| \otimes \langle \text{neutron, final}|\vec{\mathbf{S}}_{\text{neutron}} \cdot \vec{\boldsymbol{\sigma}}|0\rangle \otimes |\text{neutron, initial}\rangle|^2 \ .$$

The (known) neutron states allow us to determine q, ω. You could ask me how we do an experiment which couples just to $\mathbf{Z}$, and not all of $\vec{\boldsymbol{\sigma}}$.

Given our understanding of the spectrum at large and small g, we have a prediction for the behavior of $S(q,\omega)$: at large g, the lowest-energy excitation is a single spin flip, which conveniently is created by $\mathbf{Z}$, and so contributes to $S(q,\omega)$. But it only contributes if q and ω are related in the right way, namely $\omega = \epsilon_q$. This means that there will be a sharp peak along this curve: for ω in the right range, $S(q,\omega) = Z\delta(\omega - \epsilon_q)$. For larger ω, we access the multi-particle continuum. In fact, $\mathbf{Z}|0\rangle$ only includes states with an odd number of particles, so the continuum in $S(q,\omega)$ starts near 3Δ. What happens at small g? At small g, the spin flip is not stable – it can decay into a pair of domain walls, which are free to move apart. Although these domain walls behave in many ways like particles, the operator $\mathbf{Z}_j$ cannot create a single domain wall. (Recall that the domain wall creation operator is $\boldsymbol{\tau}^x_{\tilde{j}}$ which is made of a string of $\mathbf{X}$s.) So the spectral function of $\mathbf{Z}$ has only a continuous mess at small g – a multiparticle continuum of domain walls, no sharp delta-function peak.

People[18] have done this experiment on a system which is well-described by the TFIM ($CoNb_2O_6$). It can be approximated as a collection of decoupled chains of spins with nearest-neighbor ferromagnetic Ising interactions. The transverse field term can be produced by applying a (surprise) transverse magnetic field, and hence g can be varied. Indeed at large g one sees a sharp single-particle peak, and at small g there is a featureless mess. Here is a summary of the results of the experiment (from Ashvin Vishwanath):



A further feature of these experiments which I must mention is: if we apply a *longitudinal* field, $\Delta\mathbf{H} = \sum_j h_z \mathbf{Z}_j$, we *confine* the domain walls, in the following sense. This term

[18] Coldea et al, *Science* **327** (2010) 177.

means that a region of down-spins costs an energy proportional to its length. This is an interaction energy between the domain walls at the edges of the region which is linear in their separation, a constant attractive force between them. Their spectrum of boundstates is extremely interesting and represents a physical appearance of E_8 symmetry.

2.2.5 Solution of Ising chain in terms of Majorana fermions

Where does the Ising model come from? Just as the classical ising model arises in many ways (lattice gas, double well potential....), quantum spin systems arise in many ways. for example:

1. spinful fermions at half-filling, with charge dofs quenched.
2. Similarly, spinless fermions hopping around on a lattice also realize a spin system: there are two states per site. How to relate fermion operators and spin operators? In 1+1 dimensions, we can answer this in complete explicitness, below.
3. If we view down-spin as vacuum and up-spin as the presence of a boson, we can view a hard-core boson system as a spin-1/2 system. here it is natural to preserve a $\mathsf{U}(1) \subset \mathsf{SU}(2)$ symmetry which rotates the phase of $\mathbf{X} + \mathbf{iY}$. the conserved quantity is $\mathbf{Z} = (2\mathbf{n}_B - 1)$ where $\mathbf{n}_B$ is the boson number. [End of Lecture 7]

Jordan-Wigner in 0+1 dimensions

Consider a single fermionic operator $\mathbf{c}$ satisfying

$$\mathbf{c}^2 = 0, \quad \{\mathbf{c}, \mathbf{c}^\dagger\} = 1 \ .$$

This constructs the Hilbert space of a qbit as follows:

$$\mathbf{c}|\rightarrow\rangle = 0, \quad \mathbf{c}^\dagger|\rightarrow\rangle = |\leftarrow\rangle, \mathbf{c}^\dagger|\leftarrow\rangle = 0, \mathbf{c}|\leftarrow\rangle = |\rightarrow\rangle \ .$$

The two states of the qbit just describe the presence or absence of the fermion in this state. We can rewrite the Pauli operators as

$$\mathbf{Z} = \mathbf{c} + \mathbf{c}^\dagger,$$

which flips the spin,

$$\mathbf{Y} = \frac{1}{\mathbf{i}} \left(\mathbf{c} - \mathbf{c}^\dagger \right),$$

$$\mathbf{X} = \mathbb{1} - 2\mathbf{c}^\dagger \mathbf{c}.$$

Here $\mathbf{c}^\dagger \mathbf{c}$ counts the number of fermions. Also note that the raising and lowering operators are $\boldsymbol{\sigma}^+ \equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = c^\dagger, \boldsymbol{\sigma}^- = c$ and the number operator $\mathbf{c}^\dagger \mathbf{c}$ (whose eigenvalues are 0 and 1) is $\mathbf{c}^\dagger \mathbf{c} = \frac{1}{2}\mathbf{X} + \mathbb{1}$.

The story is less simple with more than one fermion mode operator, because they must satisfy $\{\mathbf{c}_\alpha, \mathbf{c}_\beta\} = 0$, rather than just commuting, as the would if we just used the above map for each mode.

[0904.2771, Fradkin] Let's look at the TFIM again:

$$\mathbf{H}_{\text{TFIM}} = -J \sum_j (g\mathbf{X}_j + \mathbf{Z}_j\mathbf{Z}_{j+1})$$

has a phase transition between large-g and small-g phases.

(Disordered) large g: excitations are created by $\mathbf{Z}_j$ – they are spin flips. The groundstate is a condensate of domain walls: $\langle \boldsymbol{\tau}^z \rangle \neq 0$.

(Ordered) small g: excitations are created by the 'disorder' operator $\boldsymbol{\tau}^z_{\bar{j}} = \prod_{j>\bar{j}} \mathbf{X}_j$ – they are domain walls. The groundstate is a condensate of spins $\langle \mathbf{Z}_j \rangle \neq 0$, *i.e.* a ferromagnet.

So we understand what are the 'correct variables' (in the sense of the introductory discussion to this course) at large and small g. I claim that the Correct Variables *everywhere* in the phase diagram are obtained by "attaching a spin to a domain wall". These words mean the following: let

$$\begin{aligned} \boldsymbol{\chi}_j &\equiv \mathbf{Z}_j \boldsymbol{\tau}^z_{j+\frac{1}{2}} = \mathbf{Z}_j \prod_{j'>j} \mathbf{X}_{j'} \\ \tilde{\boldsymbol{\chi}}_j &\equiv \mathbf{Y}_j \boldsymbol{\tau}^z_{j+\frac{1}{2}} = -\mathbf{i}\mathbf{Z}_j \prod_{j'\geq j} \mathbf{X}_{j'} \end{aligned} \tag{32}$$

The first great virtue of this definition is that these operators agree with the creators of the elementary excitations in both regimes we've studied: When $g \ll 1$, $\langle \mathbf{Z}_j \rangle \simeq 1$ and more strongly, $\mathbf{Z}_j = \langle \mathbf{Z}_j \rangle + \text{small}$, so $\boldsymbol{\chi}_j \simeq \langle \mathbf{Z}_j \rangle \boldsymbol{\tau}^z_{j+\frac{1}{2}} \simeq \boldsymbol{\tau}^z_{j+\frac{1}{2}}$, the domain wall creation operator. Similarly, when $g \gg 1$, $\boldsymbol{\tau}^z_j \simeq 1 + \text{small}$, so $\boldsymbol{\chi}_j \simeq \mathbf{Z}_j \langle \boldsymbol{\tau}^z_{j+\frac{1}{2}} \rangle \simeq \mathbf{Z}_j$, which is the spin flipper on the paramagnetic vacuum.

Now let us consider the algebra of these $\boldsymbol{\chi}$s:

- They are *real*: $\boldsymbol{\chi}_j^\dagger = \boldsymbol{\chi}_j, \tilde{\boldsymbol{\chi}}_j^\dagger = \tilde{\boldsymbol{\chi}}_j$.

and

- They are *fermions*:

$$\text{if } i \neq j,\ \boldsymbol{\chi}_j\boldsymbol{\chi}_i + \boldsymbol{\chi}_i\boldsymbol{\chi}_j \equiv \{\boldsymbol{\chi}_j, \boldsymbol{\chi}_i\} = 0, \quad \{\tilde{\boldsymbol{\chi}}_j, \tilde{\boldsymbol{\chi}}_i\} = 0, \quad \{\boldsymbol{\chi}_j, \tilde{\boldsymbol{\chi}}_i\} = 0. \tag{33}$$

This is because the spin flip $\mathbf{Z}_j$ in $\boldsymbol{\chi}_j$ changes sign when it moves through the domain wall created by $\boldsymbol{\chi}_i$. When they are at the same site:

$$\boldsymbol{\chi}_j^2 = 1 = \tilde{\boldsymbol{\chi}}_j^2. \quad \text{In summary:} \quad \boxed{\{\boldsymbol{\chi}_i, \boldsymbol{\chi}_j\} = 2\delta_{ij},\ \{\tilde{\boldsymbol{\chi}}_i, \tilde{\boldsymbol{\chi}}_j\} = 2\delta_{ij},}.$$

Notice that (33) means that $\boldsymbol{\chi}_i$ cares about $\boldsymbol{\chi}_j$ even if $|i-j| \gg 1$. Fermions are weird and non-local!

Notice that these fermions don't have a spin index. I should mention that this change of variables is called *Jordan-Wigner*.

Real fermion operators like this are called *Majorana* fermion operators. We can make more familiar-looking objects by making complex combinations:

$$\mathbf{c}_j \equiv \frac{1}{2}\left(\boldsymbol{\chi}_j - \mathbf{i}\tilde{\boldsymbol{\chi}}_j\right) \qquad \Longrightarrow \qquad \mathbf{c}_j^\dagger = \frac{1}{2}\left(\boldsymbol{\chi}_j + \mathbf{i}\tilde{\boldsymbol{\chi}}_j\right)$$

These satisfy the more familiar anticommutation relations:

$$\{\mathbf{c}_i, \mathbf{c}_j^\dagger\} = \delta_{ij}, \quad \{\mathbf{c}_i, \mathbf{c}_j\} = 0, \quad \{\mathbf{c}_i^\dagger, \mathbf{c}_j^\dagger\} = 0,$$

and in particular, $\left(\mathbf{c}_i^\dagger\right)^2 = 0$, like a good fermion creation operator should.

We can write $\mathbf{H}_{\text{TFIM}}$ in terms of the fermion operators. We need to know how to write $\mathbf{X}_j$ and $\mathbf{Z}_j\mathbf{Z}_{j+1}$.

The operator which *counts* spin flips in the paramagnetic phase is

$$\mathbf{X}_j = -\mathbf{i}\tilde{\boldsymbol{\chi}}_j\boldsymbol{\chi}_j = -2\mathbf{c}_j^\dagger\mathbf{c}_j + 1 = (-1)^{\mathbf{c}_j^\dagger\mathbf{c}_j}.$$

(To get this we can use (32) and $\mathbf{YZ} = \mathbf{iX}$ and $(\boldsymbol{\tau}^z)^2 = 1$.) Here $\mathbf{c}_j^\dagger\mathbf{c}_j = \mathbf{n}_j$ measures the number of fermions at the site j and is either 0 or 1, since they are fermions. At each site

$$|\rightarrow_j\rangle = |n_j = 0\rangle, \qquad |\leftarrow_j\rangle = |n_j = 1\rangle, \qquad .$$

The number of spin flips is the number of fermions.

The operator which counts domain walls is

$$\mathbf{Z}_j\mathbf{Z}_{j+1} = \mathbf{i}\tilde{\boldsymbol{\chi}}_{j+1}\boldsymbol{\chi}_j.$$

(Check: $\mathbf{i}\tilde{\boldsymbol{\chi}}_{j+1}\boldsymbol{\chi}_j = \mathbf{i}\mathbf{Y}_{j+1}\prod_{k\geq j+2}\mathbf{X}_k\mathbf{Z}_j\prod_{l\geq j+1}\mathbf{X}_l = \left(\underbrace{\mathbf{i}\mathbf{Y}_{j+1}\mathbf{X}_{j+1}}_{=\mathbf{Z}_{j+1}}\right)\mathbf{Z}_j$.)

$$\mathbf{H}_{\text{TFIM}} = -J\sum_j\left(\mathbf{i}\tilde{\boldsymbol{\chi}}_{j+1}\boldsymbol{\chi}_j + g\mathbf{i}\boldsymbol{\chi}_j\tilde{\boldsymbol{\chi}}_j\right)$$

is quadratic in these variables, for any g! Free at last! (It is quadratic in the $\mathbf{c}$s, too, since they are linear in the $\boldsymbol{\chi}$s.)

Comments:

- Notice that the relation
$$\mathbf{X}_j = 1 - 2\mathbf{c}_j^\dagger \mathbf{c}_j$$
is exactly implementing the simple idea that *spinless* fermions on a lattice produce two-states per site which we can regard as spin up or spin down (in this case it's up or down along x): The states $\mathbf{X} = \pm 1$ correspond to $\mathbf{c}^\dagger\mathbf{c} = 0$ and 1 respectively.

- Notice that the description in terms of majoranas is preferred over the complex fermions because the phase rotation symmetry generated by the fermion number $\mathbf{c}^\dagger\mathbf{c}$ is not a symmetry of $\mathbf{H}_{\text{TFIM}}$ – in terms of the $\mathbf{c}$s, it contains terms of the form $\mathbf{c}_j\mathbf{c}_{j+1}$ which change the total number of $\mathbf{c}$ fermions (by ± 2). It is the hamiltonian for a *superconductor*, in which the continuous fermion number symmetry is broken down to a $\mathbb{Z}_2$ subgroup. Fermion number is still conserved mod two, and this is the $\mathbb{Z}_2$ symmetry of the Ising model, which acts by $\mathbf{Z} \to -\mathbf{Z}$.

- A useful thing to remember about majorana operators $\{\chi, \gamma\} = 0$ is that $(\mathbf{i}\chi\gamma)^\dagger = -\mathbf{i}\gamma\chi = +\mathbf{i}\chi\gamma$ is hermitian.

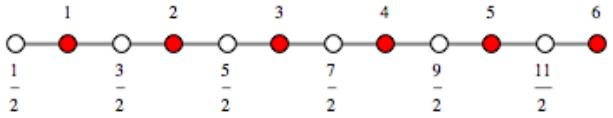
- Another useful fact:
$$\mathbf{c}(-1)^{\mathbf{c}^\dagger\mathbf{c}} = -\mathbf{c} \tag{34}$$
which is true because the BHS only nonzero if the number is nonzero *before* the annihilation operator acts, in which case we get $(-1)^1$. Similarly (the conjugate equation), $(-1)^{\mathbf{c}^\dagger\mathbf{c}}\mathbf{c}^\dagger = -\mathbf{c}^\dagger$, and $(-1)^{\mathbf{c}^\dagger\mathbf{c}}\mathbf{c} = \mathbf{c}$ and so on.

- This procedure of "attaching spin to a domain wall" led to fermions. This maybe isn't so surprising in one dimension. But there are analogs of this procedure in higher dimensions. In 2+1 dimensions, the analog is to attach charge to a vortex (or to attach magnetic flux to charge). This leads to transmutation of statistics from bosons to fermions and more generally to anyons and the fractional quantized Hall effect. In 3+1 dimensions, the analog seems to be attaching charge to a magnetic monopole to produce a 'dyon'.

Dual fermions

Let

$$\boldsymbol{\gamma}_{j+\frac{1}{2}} = -\tilde{\boldsymbol{\chi}}_{j+1}, \quad \tilde{\boldsymbol{\gamma}}_{j+\frac{1}{2}} = \boldsymbol{\chi}_j.$$



We are dividing up the pair of majorana modes that we had previously associated with the site j between the two neighboring links.

In terms of these $\mathbf{H}$ is

$$\mathbf{H}_{\text{TFIM}} = -J \sum_{\bar{j}=j+\frac{1}{2}} \left(\mathbf{i}\tilde{\boldsymbol{\gamma}}_{\bar{j}}\boldsymbol{\gamma}_{\bar{j}} + g\mathbf{i}\tilde{\boldsymbol{\gamma}}_{\bar{j}+1}\boldsymbol{\gamma}_{\bar{j}} \right).$$

We have interchanged the form of the two terms. The two expressions are related by $\boldsymbol{\chi} \leftrightarrow \boldsymbol{\gamma}, J \leftrightarrow Jg, g \leftrightarrow 1/g$.

Now let's think about the two phases in terms of the fermion fields.

$\boxed{g \gg 1}$ The disordered state, with a unique groundstate is governed by

$$\mathbf{H}_{g\to\infty} = -Jg \sum_j \mathbf{i}\boldsymbol{\chi}_j\tilde{\boldsymbol{\chi}}_j = -Jg \sum_j (-1)^{\mathbf{c}_j^\dagger \mathbf{c}_j}.$$

The groundstate of this is just the state $|0\rangle$ with no fermions, $\mathbf{c}_j|0\rangle = 0$ for all j, on which $(-1)^{\mathbf{c}_j^\dagger\mathbf{c}_j} = 1$.

$\boxed{g \ll 1}$

$$\mathbf{H}_{g\to 0} = -J \sum_j \mathbf{i}\tilde{\boldsymbol{\chi}}_{j+1}\boldsymbol{\chi}_j.$$

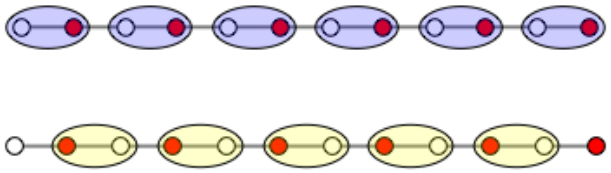
This is a hopping term for the χ fermions, which will be diagonalized by some bandstructure, which will then be partially filled by fermions. It's easier just to use the dual fermion variables, in terms of which

$$\mathbf{H}_{g\to 0} = -J \sum_{\bar{j}=j+\frac{1}{2}} \mathbf{i}\tilde{\boldsymbol{\gamma}}_{\bar{j}}\boldsymbol{\gamma}_{\bar{j}} = -J \sum_{\bar{j}=j+\frac{1}{2}} (-1)^{\check{\mathbf{c}}_{\bar{j}}^\dagger \check{\mathbf{c}}_{\bar{j}}}$$

where $\check{\mathbf{c}}_{\bar{j}} \equiv \frac{1}{2}\left(\boldsymbol{\gamma}_{\bar{j}} - \mathbf{i}\tilde{\boldsymbol{\gamma}}_{\bar{j}}\right)$ are the dual complex fermions. The groundstate of this is just the state with no dual fermions: $\check{\mathbf{c}}_{\bar{j}}|\check{0}\rangle = 0, \forall \bar{j}$.

Again we should think about the endpoint conditions. Consider what happens for an open chain. It's useful to account for the states as follows: think of dividing up each site into a pair of sites (white and black, as in the figure, one from the lattice, one from the dual lattice) with one majorana mode living at each. To help visualize, let's say $\tilde{\boldsymbol{\chi}}_j = -\boldsymbol{\gamma}_{j-\frac{1}{2}}$ lives at the white site (at $j-\frac{1}{2}$) and $\boldsymbol{\chi}_j = \tilde{\boldsymbol{\gamma}}_{j+\frac{1}{2}}$ lives at the red site to its right (at j). Note that $\boldsymbol{\gamma}_{\frac{1}{2}} = -\tilde{\boldsymbol{\chi}}_1, \tilde{\boldsymbol{\gamma}}_{\frac{1}{2}} = \boldsymbol{\chi}_0$. The g (para) term is a coupling between the white and red sites at fixed j. The ferro term couples a red site at $j-\frac{1}{2}$ to the next white site at j. This means that at small g, in an open chain, one site at each end will be *left out* – will not appear in the hamiltonian at all.

This results in *unpaired* majorana zeromodes at the ends of the chain. Let $\mathbf{a}^\dagger \equiv \frac{1}{2}(\mathbf{i}\tilde{\boldsymbol{\chi}}_1 + \boldsymbol{\chi}_N)$. The algebra $\{\mathbf{a}, \mathbf{a}^\dagger\} = 1$ must be represented on the groundstates. This requires a pair of states



$(g \gg 1)$

$(g \ll 1)$

$$|0\rangle = |+\rangle, \ \mathbf{a}^\dagger|0\rangle = |-\rangle.$$

The SSB degeneracy of the ferromagnet is mapped by this fermionization map to a *topological* degeneracy in terms of the fermions.

Splitting of the energy these two states is small because the modes are separated in space and the bulk is gapped. The modes are separated by a distance much bigger than the correlation length, $\xi \sim \frac{1}{\Delta}$. The splitting comes from a term

$$\Delta\mathbf{H} = \epsilon \mathbf{a}^\dagger \mathbf{a} = \epsilon \mathbf{i}\tilde{\boldsymbol{\chi}}_1 \boldsymbol{\chi}_N$$

and we (again) estimate that $\epsilon \sim e^{-N/\xi}$.

[End of Lecture 8]

2.2.6 Continuum limit

[Sachdev, p. 139 of 2nd ed, p. 48 of 1st ed]

We found in 2.1 that a quantum lattice model (with continuous time and discrete space) can be Trotterized into a lattice statistical mechanics model (with discrete (euclidean) time and discrete space). We also found a reverse map, via the continuum scaling limit. In this subsection we discuss the analogous scaling limit that takes a quantum lattice model to a continuum quantum field theory (with continuous time *and* continuous space).

Label site j of the lattice by its position $x_j = ja$, where a is the lattice spacing. $L = Na$ is the length of the chain. Wavevectors lie in the interval $k \in (-\frac{\pi}{a}, \frac{\pi}{a}]$.

In terms of complex fermions

$$\mathbf{X}_j = 1 - 2\mathbf{c}_j^\dagger \mathbf{c}_j, \quad \mathbf{Z}_j = -\prod_{i>j}(1 - 2\mathbf{c}_i^\dagger \mathbf{c}_i)\left(\mathbf{c}_j + \mathbf{c}_j^\dagger\right) = -\prod_{i>j}(-1)^{\mathbf{c}_i^\dagger \mathbf{c}_i}\left(\mathbf{c}_j + \mathbf{c}_j^\dagger\right).$$

and their Fourier modes $\mathbf{c}_k \equiv \frac{1}{\sqrt{N}}\sum_j \mathbf{c}_j e^{-\mathbf{i}kx_j}$, the TFIM hamiltonian is

$$\mathbf{H}_{\text{TFIM}} = J\sum_k \left(2(g - \cos ka)\mathbf{c}_k^\dagger \mathbf{c}_k - \mathbf{i}\sin ka\left(\mathbf{c}_{-k}^\dagger \mathbf{c}_k^\dagger + \mathbf{c}_{-k}\mathbf{c}_k\right) - g\right)$$

Notice that the second set of terms violates fermion number conservation in units of two; this is the same statement that the ferromagnetic term creates spin flips in pairs. The constant at the end is irrelevant unless we are after the Onsager free energy.

This Hamiltonian is quadratic in $\mathbf{c}_k$s, but not quite diagonal. The final solution for the spectrum involves one more operation the fancy name for which is 'Bogoliubov transformation', which is the introduction of new (complex) mode operators which mix particles and holes:

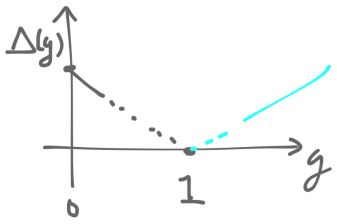
$$\boldsymbol{\gamma}_k = u_k \mathbf{c}_k - \mathbf{i}v_k \mathbf{c}_{-k}^\dagger$$

Demanding that the new variables satisfy canonical commutators $\{\boldsymbol{\gamma}_k, \boldsymbol{\gamma}_{k'}^\dagger\} = \delta_{k,k'}$ requires $u_k = \cos(\phi_k/2), v_k = \sin(\phi_k/2)$. We fix the angles ϕ_k by demanding that the hamiltonian in terms of $\boldsymbol{\gamma}_k$ be diagonal – no $\boldsymbol{\gamma}_k\boldsymbol{\gamma}_{-k}$ terms. The resulting condition is $\tan\phi_k = \frac{\epsilon_2(k)}{\epsilon_1(k)}$ with $\epsilon_1(k) = 2J(g - \cos ka), \epsilon_2(k) = -J\sin ka$, and $\mathbf{H} = \sum_k \epsilon_k\left(\boldsymbol{\gamma}_k^\dagger \boldsymbol{\gamma}_k - \frac{1}{2}\right)$, with $\epsilon_k = \sqrt{\epsilon_1^2 + \epsilon_2^2}$.

The end result is that the exact single-particle (single $\boldsymbol{\gamma}$) dispersion is

$$\epsilon_k = 2J\sqrt{1 + g^2 - 2g\cos ka}\ .$$

The argument of the sqrt is positive for $g \geq 0$. This is minimized at $k = 0$, which tells us the exact gap at all g:



$$\epsilon_k \geq \epsilon_0 = 2J|1 - g| = \Delta(g)$$

which, ridiculously, is just what we got from 1st order perturbation theory on each side of the transition.[19]

Comment on relation to Onsager. Notice that by the quantum-to-classical mapping, this solution has all the universal information of Onsager's solution. In particular, his exact free energy can be obtained by computing the groundstate energy of the fermions. (See the homework.)

The critical theory is scale invariant. At $g \to 1$, the fermions at $k = 0$ become gapless:

$$\epsilon_k \overset{g=g_c=1}{=} c|k|,$$

where the speed of propagation is $c = 2Ja$. There is an emergent Lorentz symmetry, at least in the spectrum, at the critical point. And the fact that the long-wavelength ($k \sim 0$) have the lowest energy allows a continuum description that forgets the lattice details.

Near $g \to g_c = 1$ (*i.e.* $|g - g_c| \ll g_c$),

$$\epsilon_k \overset{g\sim g_c=1}{=} c\sqrt{k^2 + \underbrace{\left(\frac{g-g_c}{a}\right)^2}_{m^2\to 0}}$$

There is a diverging length scale, which is $\xi = \frac{1}{m} = \frac{a}{|g-g_c|}$.

Some useful language:

- The *correlation length critical exponent* is ν in $\xi \sim |g - g_c|^{-\nu}$, the rate at which the correlation length diverges near the critical point. (Here $\nu = 1$.)
- The *dynamical critical exponent* is the power law in the dispersion *at* the critical point:

  $$\epsilon(k) \propto k^z. \tag{35}$$

  Notice that if we rescale space and time like

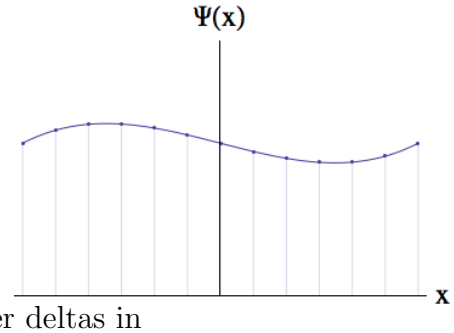
  $$x \to \lambda x, \qquad t \to \lambda^z t \tag{36}$$

  it will preserve the dispersion (35). (Here $z = 1$, as in a relativistic theory where time and space can be boosted into each other.)

We can expect the continuum description (at least the long-wavelength information) to be invariant under the scale transformation (36).

[19] Notice by the way that this 'single-particle excitation' in terms of the spins is a state with a single domain wall, and therefore only makes sense in the thermodynamic limit; in finite volume, the spin hilbert space only contains states with an even number of fermions. See §2.2.7 below for more on this point.

When $\xi \gg a$, define the continuum fermion field



$$\mathbf{\Psi}(x_j) \equiv \frac{1}{\sqrt{a}}\mathbf{c}_j$$

where the factor of $\sqrt{a}$ is designed to convert the kronecker deltas in

$$\{\mathbf{c}_j, \mathbf{c}^\dagger_{j'}\} = \delta_{jj'} \quad \text{into} \quad \{\mathbf{\Psi}(x), \mathbf{\Psi}^\dagger(x')\} = \delta(x-x').$$

The fourier transform is

$$\mathbf{c}_k = \int dx \frac{e^{-\mathbf{i}kx}}{\sqrt{L}} \mathbf{\Psi}(x) \ .$$

Let's plug this into $\mathbf{H}_{\rm TFIM}$, focussing on (the lightest) modes with $ka \ll 1$.

$$J\sum_k (g - \cos ka)\mathbf{c}^\dagger_k \mathbf{c}_k \rightsquigarrow (g-g_c)\int dx\ \mathbf{\Psi}^\dagger(x)\mathbf{\Psi}(x)$$

Here the $\rightsquigarrow$ means we are dropping terms of higher order in the expansion in powers of $a\partial_x$.

$$-\mathbf{i}J\sum_k \sin ka \mathbf{c}_{-k}\mathbf{c}_k \rightsquigarrow \frac{c}{2}\int dx\ \mathbf{\Psi}(x)^\dagger \partial_x \mathbf{\Psi}(x)^\dagger$$

So

$$\mathbf{H} \rightsquigarrow \frac{c}{2}\int dx\ \left(\mathbf{\Psi}(x)^\dagger\partial_x\mathbf{\Psi}(x)^\dagger - \mathbf{\Psi}(x)\partial_x\mathbf{\Psi}(x)\right) + \Delta\int dx \mathbf{\Psi}^\dagger\mathbf{\Psi}$$

with $\Delta = 2J|g-1|$.

[Fradkin, 2d ed p. 119] To understand the significance of this result it will be best to use the (fermion coherent state) path integral, which we'll do below in §2.2.8. But we can learn something just by thinking about the Heisenberg equations of motion.

$$\begin{aligned}\mathbf{H}_{\rm TFIM} &= -J\sum_j (g\mathbf{X}_j + \mathbf{Z}_j\mathbf{Z}_{j+1}) \\ &= -\mathbf{i}J\sum_j (g\boldsymbol{\chi}(j)\tilde{\boldsymbol{\chi}}(j) - \boldsymbol{\chi}(j)\tilde{\boldsymbol{\chi}}(j+1)) \ . \end{aligned} \tag{37}$$

Notice that the $\mathbb{Z}$ symmetry action on the fermions is

$$\mathbf{S} = \prod_j \mathbf{X}_j = \mathbf{i}^N \prod_j \boldsymbol{\chi}(j)\tilde{\boldsymbol{\chi}}(j).$$

The Heisenberg eom are $\mathbf{i}\partial_t\mathcal{O} = [\mathbf{H}, \mathcal{O}]$:

$$\mathbf{i}\partial_t\boldsymbol{\chi}(j) \ = \ \mathbf{i}J\left(g\tilde{\boldsymbol{\chi}}(j) - \tilde{\boldsymbol{\chi}}(j+1)\right)$$

$$\mathbf{i}\partial_t\tilde{\boldsymbol{\chi}}(j) \;=\; \mathbf{i}J\left(-g\boldsymbol{\chi}(j)+\boldsymbol{\chi}(j-1)\right) \tag{38}$$

(Free fields means linear equations of motion.) In the continuum,

$$\boldsymbol{\chi}(j+1)\simeq\boldsymbol{\chi}(x_j)+a\partial_x\boldsymbol{\chi}(x_j)+\mathcal{O}(a^2\partial_x^2)$$

so

$$\begin{aligned}\frac{1}{aJ}\partial_t\boldsymbol{\chi}(x) &\simeq -\left(\frac{1-g}{aJ}\right)\tilde{\boldsymbol{\chi}}(x)-\partial_x\tilde{\boldsymbol{\chi}}(x)\\ \frac{1}{aJ}\partial_t\tilde{\boldsymbol{\chi}}(x) &\simeq \left(\frac{1-g}{aJ}\right)\boldsymbol{\chi}(x)-\partial_x\boldsymbol{\chi}(x)\end{aligned} \tag{39}$$

If we let $\boldsymbol{\chi}_\pm=\frac{1}{2}\left(\boldsymbol{\chi}\mp\tilde{\boldsymbol{\chi}}\right)$ and rescale the speed of propagation into the time variable, $t\equiv aJx^0$,

$$\begin{aligned}\partial_0\boldsymbol{\chi}_+ &= \partial_x\boldsymbol{\chi}_+ + m\boldsymbol{\chi}_-\\ \partial_0\boldsymbol{\chi}_- &= -\partial_x\boldsymbol{\chi}_- - m\boldsymbol{\chi}_+\end{aligned} \tag{40}$$

with $m=\frac{1-g}{a}$. When $m\to 0$, at the critical point, these are *chiral* fermions:

$$\left(\partial_0\mp\partial_x\right)\boldsymbol{\chi}_\pm=0.$$

The $\boldsymbol{\chi}$s are still real, so these are majorana fermions. In a language we will introduce soon, this is an example of a 1+1d CFT with central charge $c=(\frac{1}{2},\frac{1}{2})$.

Even a bit away from the critical point, we can reconstruct the relativistic symmetry, and organize (40) into following Dirac equation

$$0=\mathbf{i}\gamma^\mu\partial_\mu\boldsymbol{\chi}+\mathbf{i}m\boldsymbol{\chi},$$

Notice that a spin index has emerged to save the spin-statistics theorem. Here we have chosen a basis of gamma matrices where

$$\gamma^0=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\qquad \gamma^1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

(in the $\pm$ basis) which satisfy $\{\gamma^\mu,\gamma^\nu\}=2\eta^{\mu\nu}$ as they should. So the Dirac equation above is

$$0=\mathbf{i}\begin{pmatrix}m & \partial_0+\partial_x\\ -\partial_0+\partial_x & m\end{pmatrix}\begin{pmatrix}\boldsymbol{\chi}_+\\ \boldsymbol{\chi}_-\end{pmatrix}\ .$$

(The Majorana condition I am imposing is $\bar{\boldsymbol{\chi}}=\boldsymbol{\chi}^T\gamma^0$, aka $\boldsymbol{\chi}^\dagger=\boldsymbol{\chi}^T$, so $\boldsymbol{\chi}_\pm$ are real.)

2.2.7 Duality and Jordan-Wigner with periodic boundary conditions

---

How do I define the creation operator for a single domain wall with periodic boundary conditions on the spins? I can't. The fact that I can only make an even number means that the *sign* of the operator is not well-defined – in fact you can't measure the single-domain wall operator. And since the fermions $\boldsymbol{\chi}$ are made from this operator, you can't measure a single fermion operator either.

Although there is a sign ambiguity, we can still define the domain wall operators:

$$\boldsymbol{\tau}^x_{\bar{j}} = \mathbf{Z}_{\bar{j}-\frac{1}{2}}\mathbf{Z}_{\bar{j}+\frac{1}{2}}, \quad \boldsymbol{\tau}^z_{\bar{j}} = \prod_{N\geq j>\bar{j}} \mathbf{X}_j$$

– in the second definition, we have made a choice: even though $\mathbf{Z}_{N+1} = \mathbf{Z}_1$, we've decided to pick out the link where $\bar{j} = N + \frac{1}{2}$ to stop the string of $\mathbf{X}$s. As we saw in the case of open boundary conditions, this means that $\boldsymbol{\tau}_{\frac{1}{2}}$ and $\boldsymbol{\tau}_{N+\frac{1}{2}}$ are special:

$$\boldsymbol{\tau}^z_{N+\frac{1}{2}} = 1$$

since the product is over the empty set.

$$\boldsymbol{\tau}^z_{\frac{1}{2}} = \prod_{j=1}^{N} \mathbf{X}_j = \mathbf{S}$$

is the symmetry generator. This is significant because this operator commutes with $\mathbf{H}_{TFIM}$, it means that eigenstates of $\boldsymbol{\tau}^z_{\frac{1}{2}}$ are energy eigenstates. And notice that it determines the boundary conditions on the $\boldsymbol{\tau}$s in following sense: combining the previous two equations,

$$1 = \boldsymbol{\tau}^z_{N+\frac{1}{2}} = \begin{cases} \boldsymbol{\tau}^z_{\frac{1}{2}} & \text{if } \prod_j \mathbf{X}_j = \mathbf{S} = 1 \\ -\boldsymbol{\tau}^z_{\frac{1}{2}} & \text{if } \prod_j \mathbf{X}_j = \mathbf{S} = -1 \end{cases}$$

This operator $\boldsymbol{\tau}^z_{\frac{1}{2}}$ which says whether the $\boldsymbol{\tau}$s are periodic or antiperiodic is a dynamical variable.

Gauge theory formulation of the 1d duality

There is a way to get around the obstacle of no-lone-domain wall on a circle, which generalizes to higher dimensions (but it is kind of overkill in this case). It is a reformulation of the TFIM as a lattice gauge theory. This is worth doing because in duality transformations in higher dimensions, the gauge theory formulation cannot be avoided. So this is a context where we can introduce the concepts of gauge theory without lots of other complications.

We begin by enlarging our system, by adding, in addition to the real qbits on the sites of the chain another set of fake qbits on the links:

$$\text{enlarge } \mathcal{H} = \otimes_j \mathcal{H}_j \quad \subset \quad \mathcal{H}_{\text{big}} \equiv (\otimes_j \mathcal{H}_j) \otimes \left(\otimes_{\bar{j}} \check{\mathcal{H}}_{\bar{j}}\right) \ .$$

Anticipating the answer, I will label the two states on a link $\bar{j}$ as $|\tau^z_{\bar{j}}\rangle = |\pm 1\rangle$. So as for any two-state system, associated with each link there are pauli operators

$$\boldsymbol{\tau}^z_{\bar{j}} |\tau^z_{\bar{j}}\rangle = \tau^z_{\bar{j}} |\tau^z_{\bar{j}}\rangle, \quad \boldsymbol{\tau}^x_{\bar{j}} |\tau^z_{\bar{j}}\rangle = |-\tau^z_{\bar{j}}\rangle \ , \ \boldsymbol{\tau}^z_{\bar{j}} \boldsymbol{\tau}^x_{\bar{j}'} = (-1)^{\delta_{\bar{j}\bar{j}'}} \boldsymbol{\tau}^x_{\bar{j}} \boldsymbol{\tau}^z_{\bar{j}'} \ .$$

We want to pick out a subspace of this larger fake Hilbert space which is the real one. One way to think about this is to declare that the following 'gauge transformation' is an equivalence:

$$\mathcal{G} : \tau^z_{\bar{j}} \to s_{\bar{j}} \tau^z_{\bar{j}}, \quad s_{\bar{j}} = \pm 1 \ . \tag{41}$$

Notice that this group $\mathcal{G}$ is big – we pick an element of $\mathbb{Z}_2$ (*i.e.* a sign) $s_{\bar{j}} = \pm 1$ for each link of the lattice, its order is $|\mathbb{Z}_2|^N = 2^N$. If a configuration of τ^zs is equivalent to its image under this map, we can just pick $s_{\bar{j}} = \tau^z_{\bar{j}}$ and this means that any configuration of τ^zs is equivalent to 1: we can just get rid of the τs and we are back at our original Hilbert space. This is called *unitary gauge*. Below we will implement this more concretely. But first we must decide how the gauge redundancy $\mathcal{G}$ acts on our original variables. It will be useful to choose

$$\mathcal{G} : \mathbf{X}_j \to s_{j-\frac{1}{2}} \mathbf{X}_j s_{j+\frac{1}{2}}, \quad \mathbf{Z}_j \to \mathbf{Z}_j. \tag{42}$$

This is the transformation law of a *link variable* in lattice gauge theory. The site of the original lattice is a link of the dual lattice. Let me use notation which emphasizes this:

$$\mathbf{X}_j \equiv \boldsymbol{\sigma}^x_{j-\frac{1}{2}, j+\frac{1}{2}}, \quad \mathbf{Z}_j \equiv \boldsymbol{\sigma}^z_{j-\frac{1}{2}, j+\frac{1}{2}} \ .$$

The *generator* of the gauge transformation is

$$\mathbf{G}_{\bar{j}} = \boldsymbol{\sigma}^z_{\bar{j}-1,\bar{j}} \boldsymbol{\tau}^x_{\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1}.$$

What I mean by 'generator' is that the action of the symmetry on states is $|\psi\rangle \to \mathbf{G}|\psi\rangle$ and the action on operators is

$$\mathcal{O} \to \mathbf{G} \mathcal{O} \mathbf{G}^\dagger.$$

The operator $\mathbf{G}_{\bar{j}}$ generates the transformation where $s_{\bar{j}} = -1$ and all the other $s_{\bar{j}'} = +1$. Check that this reproduces (41) and (42).

The physical hilbert space is cut out of $\mathcal{H}_{\text{big}}$ by projecting to eigenstates of these operators with eigenvalue 1:

$$\mathcal{H}_{\text{phys}} = \{|\psi\rangle \in \mathcal{H}_{\text{big}} \text{ such that } \mathbf{G}_{\bar{j}}|\psi\rangle = |\psi\rangle, \forall j\} \subset \mathcal{H}_{\text{big}}.$$

This is just the statement that physical states don't change under our made-up redundancy[20].

In order for this construction to make sense, our hamiltonian acting on the big hilbert space must be *gauge invariant*. That is, we must have

$$[\mathbf{H}, \mathbf{G}_{\bar{j}}] = 0, \quad \forall \bar{j}.$$

This way, we can simultaneously diagonalize $\mathbf{H}$ and $\mathbf{G}_{\bar{j}}$. A gauge invariant hamiltonian acting on the big hilbert space is:

$$\mathbf{H}_{\text{big}} = -J \sum_{\bar{j}} \left(\boldsymbol{\sigma}^z_{\bar{j}-1,\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1} + g \boldsymbol{\tau}^z_{\bar{j}} \boldsymbol{\sigma}^x_{\bar{j},\bar{j}+1} \boldsymbol{\tau}^z_{\bar{j}+1} \right) \tag{43}$$

In unitary gauge, $\boldsymbol{\tau}^z = 1$, we can erase the $\boldsymbol{\tau}^z$s and this reduces to $\mathbf{H}_{\text{TFIM}}$. This is why I chose it.

On the other hand, using the local constraints

$$1 = \mathbf{G}_{\bar{j}} \;\Leftrightarrow\; \boldsymbol{\sigma}^z_{\bar{j}-1,\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1} = \boldsymbol{\tau}^x_{\bar{j}} \quad \text{on} \quad \mathcal{H}_{\text{phys}},$$

we have

$$\mathbf{H}_{\text{big}} = -J \sum_{\bar{j}} \left(\boldsymbol{\tau}^x_{\bar{j}} + g \boldsymbol{\tau}^z_{\bar{j}} \boldsymbol{\sigma}^x_{\bar{j},\bar{j}+1} \boldsymbol{\tau}^z_{\bar{j}+1} \right) \ . \tag{44}$$

In this last expression, we still have the gauge freedom to play with – too many variables. Let's use it to get rid of the original $\boldsymbol{\sigma}$s. On an open chain, we can use the transformation

$$\boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1} \to s_{\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1} s_{\bar{j}+1} \tag{45}$$

to set $\boldsymbol{\sigma}^z = 1$ for each $\bar{j}$. In that case, the expression (44) precisely reduces to our dual description of the TFIM in terms of domain wall operators, (25).

But with periodic boundary conditions, the object

$$\mathbf{W} \equiv \prod_{\bar{j}=1}^{N} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1}$$

is a sign which does not change under the transformation (45). It is called a *Wilson line* operator. We can't get rid of it. In fact it is the familiar object

$$\mathbf{W} = \prod_j \mathbf{X}_j = \mathbf{S},$$

[20] In lecture I called this constraint the 'Gauss' Law constraint'. This is a correct and useful terminology which I have not explained yet. Sorry for the extra jargon. The reason it is correct is that if you *really* squint at this expression you can recognize it as a lattice version of the divergence of the electric flux coming out of the site $\bar{j}$. This will become clear in §5.2.

the $\mathbb{Z}_2$ symmetry charge.

You can see from (43) that if $\mathbf{W} = -1$, we have APBCs on $\boldsymbol{\tau}^z$. We can choose a gauge where we get rid of all the $\boldsymbol{\sigma}$s except the first, $\boldsymbol{\sigma}^z_{\frac{1}{2},\frac{3}{2}} = \mathbf{W}$. (This is a manifestation of the fact that we had to label a configuration of the spins by the location of the domain walls plus the value of one spin, which we are again choosing to be the first one.) Then the last link in the chain has an extra factor:

$$\mathbf{H}_{\text{dual}} = -J \sum_{\bar{j}=\frac{1}{2}}^{N-\frac{3}{2}} \left(\boldsymbol{\tau}^x_{\bar{j}} + g \boldsymbol{\tau}^z_{\bar{j}} \boldsymbol{\tau}^z_{\bar{j}+1} \right) + \boldsymbol{\tau}^x_{N-\frac{1}{2}} + \mathbf{W} g \boldsymbol{\tau}_{N-\frac{1}{2}} \boldsymbol{\tau}_{\frac{1}{2}}$$

[End of Lecture 9]

Gauge theory without metaphysics: Kitaev's toric code in 1d

Since there seemed to be some discontent about the sudden appearance of gauge theory in the previous discussion, I will try to make amends here by giving an example of whence this construction can come. It is a 1d version of Kitaev's toric code. The analogous construction in higher dimensions is extremely important as a demonstration of the emergence of gauge theory from a local lattice model. This silly reduction of it also does that, but gauge theory in 1d is not quite as exciting as it is in higher dimensions.

Suppose I tell you that the following is a high-energy description of a lattice quantum system: The Hilbert space is

$$\mathcal{H}_{\rm big} \equiv (\otimes_j \mathcal{H}_j) \otimes \left(\otimes_{\bar{j}} \check{\mathcal{H}}_{\bar{j}}\right)$$

(nothing is fake, there is no redundancy, no nonsense, this is really the Hilbert space), and the Hamiltonian is

$$\mathbf{H}_{\rm TC} = \mathbf{H}_{\mathbf{G}} + \mathbf{H}_{\rm big}$$

where $\mathbf{H}_{\rm big} = -J \sum_{\bar{j}} \left(\boldsymbol{\sigma}^z_{\bar{j}-1,\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1} + g \boldsymbol{\tau}^z_{\bar{j}} \boldsymbol{\sigma}^x_{\bar{j},\bar{j}+1} \boldsymbol{\tau}^z_{\bar{j}+1} \right)$ as before, and the new term relative to the previous discussion is

$$\mathbf{H}_{\mathbf{G}} \equiv -E_{\rm giant} \sum_{\bar{j}} \mathbf{G}_{\bar{j}}$$

with $\mathbf{G}_{\bar{j}} = \boldsymbol{\sigma}^z_{\bar{j}-1,\bar{j}} \boldsymbol{\tau}^x_{\bar{j}} \boldsymbol{\sigma}^z_{\bar{j},\bar{j}+1}$ as before. Here $E_{\rm giant}$ is an energy scale which is a gajillion times bigger than any other energy in the problem ($E_{\rm giant} \gg J \gg gJ$).

A few relevant facts:

$$[\mathbf{H}_{\rm big}, \mathbf{G}_{\bar{j}}] = 0, \qquad [\mathbf{G}_{\bar{j}}, \mathbf{G}_{\bar{j}'}] = 0, \quad \forall \bar{j}, \bar{j}'.$$

This means that we can diagonalize $\mathbf{H}_{\rm big}$ by first diagonalizing $\mathbf{H}_{\mathbf{G}}$ (which is a good idea, since its coefficient is so giant), and further, we can do *that* by diagonalizing each $\mathbf{G}_{\bar{j}}$ at the same time.

So who are these $\mathbf{G}_{\bar{j}}$s? $\mathbf{G}_{\bar{j}}$ is unitary (so should be thought of as enacting a transformation on $\mathcal{H}_{\rm big}$), and, in this example, $\mathbf{G}^2_{\bar{j}} = 1$, it is an element of $\mathbb{Z}_2$, so its eigenvalues are ± 1. For each $\mathbf{G}_{\bar{j}}$, the eigenstate with eigenvalue $+1$ is the low-energy one.

Let us pause to note that this model actually has a *local symmetry*: the symmetry generated by $\mathbf{G}_{\bar{j}}$ acts only on the dofs at or next to $\bar{j}$, and this is a symmetry for any $\bar{j}$. *However*, with the sign of $E_{\rm giant}$ as indicated, all states on which this symmetry actually acts nontrivially have an enormous energy $\sim E_{\rm giant}$. This means that we may forget about them and focus on the low-energy subspace of the Hilbert space

$$\mathcal{H}_{\rm phys} \equiv \{|\psi\rangle \in \mathcal{H}_{\rm big} \text{ such that } \mathbf{G}_{\bar{j}}|\psi\rangle = |\psi\rangle, \forall j\} \subset \mathcal{H}_{\rm big}.$$

You have seen this before, on the previous page: it is the Hilbert space of the $\mathbb{Z}_2$ gauge theory. We have emerged a gauge theory.

Now let me explain better what I was saying about "choosing a gauge". What I meant is the following. States of $\mathcal{H}_{\rm phys}$ correspond to *orbits* of the action of $\{\mathbf{G}_{\bar{j}}\}$ on $\mathcal{H}_{\rm big}$.[21] We can think of the projection to the $\mathbf{G}=1$ subspace as *choosing a representative of each orbit* of the action of (each) $\mathbf{G}$.

Consider an (oversimplified) example of a single two-state system with $\mathbf{G}=\boldsymbol{\sigma}^z$. In the $\mathbf{X}$-basis, this symmetry action has two orbits:

$$\frac{1}{\sqrt{2}}\left(|\rightarrow\rangle+|\leftarrow\rangle\right)=|\uparrow\rangle \quad \text{and} \quad \frac{1}{\sqrt{2}}\left(|\rightarrow\rangle-|\leftarrow\rangle\right)=|\downarrow\rangle$$

which are determined by the eigenvalue $\sigma^z=\uparrow,\downarrow$. If we *know* that we only want to keep the state with eigenvalue $+1$ then we can label an orbit by the first term. In this case there is only one orbit, and the projection is one-dimensional. We can describe this by saying that we have used the gauge redundancy to set $\sigma^x=1$, and simply set $\boldsymbol{\sigma}^x=1$ wherever it appears in our Hamiltonian.

More generally, we will have some action of a (local in space) group $g\in G$ of the form $\mathbf{U}_g$ on our $\mathcal{H}_{\rm big}$. This means that $\mathcal{H}_{\rm big}$ forms a (in general highly reducible) representation of G. The orbits of G are

$$\frac{1}{\sqrt{|G|}}\sum_{g\in G}\mathbf{U}_g|g_0\rangle$$

where $|G|$ is the order of the group G, and $|g_0\rangle$ is a reference state (like a highest-weight state in the theory of Lie groups). We may label the orbits by the reference state. And more to the point, states that lie in the *same* orbit correspond to the *same* unit eigenvector, and can be regarded as equivalent. This is the origin of the usual way of thinking about gauge redundancy.

We'll have more to say about the real toric code in $d>1$ later (in §5.2). We have shown that this model reduces at energies $E\ll E_{\rm giant}$ to the gauge theory description of the TFIM, which we have solved by the Jordan-Wigner transformation. The situation is a little bit different in 2d in that the model is not solvable away from the zero-correlation length limit.

[21] Given a unitary operator on a Hilbert space and a state $|\psi_0\rangle$ in that space, what I mean by 'the orbit of $|\psi_0\rangle$ under the action of $\mathbf{U}$' is the set of images under repeated action of $\mathbf{U}$:

$$\text{orbit of } |\psi_0\rangle \text{ under action of } \mathbf{U} \equiv \{|\psi_0\rangle, \mathbf{U}|\psi_0\rangle, \mathbf{U}^2|\psi_0\rangle ...\}$$

(if $\mathbf{U}$ represents a finite group, this contains a finite number of elements; call this order($\mathbf{U}$)). The associated eigenvector of $\mathbf{U}$ with eigenvalue 1 is

$$\frac{1}{\sqrt{\text{order}(\mathbf{U})}}\sum_{l=0}^{\text{order}(\mathbf{U})}\mathbf{U}^l|\psi_0\rangle.$$

It is just the average over the group. (For non-abelian groups, we must speak about the orbit of $|\psi_0\rangle$ under the action of the whole group.)

Boundary conditions on fermions

How does this arise in terms of the fermions? Just as for domain walls, the sign of the fermion operator is also not well-defined. With periodic boundary conditions on the spins ($\mathbf{Z}_{j+N} = \mathbf{Z}_j$), we can also define the fermions by putting a *branch cut* at $j = N$, that is, we'll write the Jordan-Wigner formula as $\boldsymbol{\chi}_j = \mathbf{Z}_j \prod_{N \geq i > j} \mathbf{X}_i$, where the string of $\mathbf{X}_i$s stops at N. This is a choice. The inverse map is then

$$\mathbf{Z}_j = \left(\mathbf{c}_j^\dagger + \mathbf{c}_j\right) \prod_{N \geq i > j} e^{\mathbf{i}\pi\mathbf{n}_i} \ . \tag{46}$$

The ferromagnetic term in the hamiltonian, for $j \neq N$, is

$$\mathbf{Z}_j\mathbf{Z}_{j+1} = \left(\mathbf{c}_j^\dagger + \mathbf{c}_j\right) e^{\mathbf{i}\pi\mathbf{n}_{j+1}} \left(\mathbf{c}_{j+1}^\dagger + \mathbf{c}_{j+1}\right) = \left(\mathbf{c}_j^\dagger + \mathbf{c}_j\right) \left(-\mathbf{c}_{j+1}^\dagger + \mathbf{c}_{j+1}\right)$$

where I used the identity (34). (This is as in our previous discussion.) But with periodic boundary conditions on $\mathbf{Z}$, the term at the branch cut is:

$$\mathbf{Z}_N\mathbf{Z}_1 = \left(\mathbf{c}_N^\dagger + \mathbf{c}_N\right) \left(\mathbf{c}_1^\dagger + \mathbf{c}_1\right) \prod_{N \geq j > 1} e^{\mathbf{i}\pi\mathbf{n}_j} \ .$$

This last red factor includes *all* of the sites except the first:

$$\prod_{N \geq j > 1} e^{\mathbf{i}\pi\mathbf{n}_j} = \prod_{j=1}^{N} e^{\mathbf{i}\pi\mathbf{n}_j} e^{\mathbf{i}\pi\mathbf{n}_1} = e^{\mathbf{i}\pi\mathbf{N}} e^{\mathbf{i}\pi\mathbf{n}_1} \equiv (-1)^{\mathbf{N}} e^{\mathbf{i}\pi\mathbf{n}_1} \ .$$

Here $\mathbf{N} \equiv \sum_j \mathbf{c}_j^\dagger \mathbf{c}_j$ operator which counts the total number of fermions, and $(-1)^{\mathbf{N}}$ is called the 'fermion parity'[22]. This is a symmetry generator: $[\mathbf{H}, (-1)^{\mathbf{N}}] = 0$, since $\mathbf{H}$ only contains terms with an even number of fermions. Returning to the outlying ferromagnetic coupling,

$$\begin{aligned} \mathbf{Z}_N\mathbf{Z}_1 &= \left(\mathbf{c}_N^\dagger + \mathbf{c}_N\right) \underbrace{\left(\mathbf{c}_1^\dagger + \mathbf{c}_1\right) e^{\mathbf{i}\pi\mathbf{n}_1}}_{=\mathbf{c}_1^\dagger - \mathbf{c}_1} (-1)^{\mathbf{N}} \\ &= \left(\mathbf{c}_N^\dagger + \mathbf{c}_N\right) \left(\mathbf{c}_1^\dagger - \mathbf{c}_1\right) (-1)^{\mathbf{N}}. \end{aligned} \tag{47}$$

This is the same as the other terms, but for a factor of $-(-1)^{\mathbf{N}}$. This means that we get periodic boundary conditions on $\mathbf{c}$, *i.e.* $\mathbf{c}_{N+1} = +\mathbf{c}_1$ if the total number of fermions is *odd* and antiperiodic boundary conditions, $\mathbf{c}_{N+1} = -\mathbf{c}_1$, in sectors where $\mathbf{N}$ is even. In summary[23] :

$$\boxed{\mathbf{c}_{N+1} = -\mathbf{c}_1(-1)^{\mathbf{N}}} \quad .$$

[22] please don't confuse the number operator $\mathbf{N}$ with the number of sites N – sorry about that

[23] This fact is perhaps simpler in terms of the majoranas:

$$\text{for } j < N, \ -\mathbf{i}\tilde{\boldsymbol{\chi}}_{j+1}\boldsymbol{\chi}_j = \mathbf{Z}_j\mathbf{Z}_{j+1}, \quad \text{but} \quad -\mathbf{i}\tilde{\boldsymbol{\chi}}_N\boldsymbol{\chi}_1 = \mathbf{Z}_N\mathbf{Z}_1\,(-\mathbf{S})\,.$$

So the full Hamiltonian is

$$\frac{1}{J}\mathbf{H}_{\text{TFIM}}^{\text{PBC on }\mathbf{z}} = -\sum_{j=1}^{N-1}\left(\mathbf{c}_j^\dagger + \mathbf{c}_j\right)\left(\mathbf{c}_{j+1}^\dagger - \mathbf{c}_{j+1}\right) - \left(\mathbf{c}_N^\dagger + \mathbf{c}_N\right)\left(\mathbf{c}_1^\dagger - \mathbf{c}_1\right)(-1)^{\mathbf{N}} - 2g\underbrace{\sum_j \mathbf{c}_j^\dagger \mathbf{c}_j}_{=\mathbf{N}}.$$

We can diagonalize this by fourier transformation

$$\mathbf{c}_j = \frac{1}{\sqrt{N}}\sum_k e^{\mathbf{i}kja}\mathbf{c}_k$$

where the allowed values of k depend on the boundary conditions on $\mathbf{c}$:

$$\begin{aligned} \text{PBC}, \quad (-1)^{\mathbf{N}} = -1: \qquad & k \in \frac{2\pi}{Na}\{1...N\} \\ \text{APBC}, \quad (-1)^{\mathbf{N}} = +1: \qquad & k \in \frac{2\pi}{Na}\left(\frac{1}{2} + \{1...N\}\right) \end{aligned} \tag{48}$$

And the Hamiltonian in each case is:

$$\mathbf{H}/J = \sum_k \left(\mathbf{c}_k^\dagger \mathbf{c}_k\left(2g - 2\cos ka\right) + \left(\mathbf{c}_k \mathbf{c}_{-k}\mathbf{i}\sin ka + h.c.\right)\right) \equiv \sum_k \mathbf{h}_k$$

The dynamics determine the boundary conditions: is it less energy to have an odd number of fermions or an even number? First notice that the dispersion has an inversion symmetry $k \to -k$, so each k value is degenerate in energy with $-k$, unless it is fixed by this map. The fixed points are $k = 0, \pi/a$ (recall that $ka = -\pi \simeq -\pi + 2\pi = \pi$). These only occur for PBC.

The twofold (in particular even) degeneracy for most k means that the occupation $\mathbf{N}_{k\neq -k}$ only changes by two, and we don't care about the associated $\mathbf{h}_k$. More algebraically, for PBC,

$$\mathbf{N} = \sum_j \mathbf{c}_j^\dagger \mathbf{c}_j = 2\sum_{k\neq 0,\pi} \mathbf{c}_k^\dagger \mathbf{c}_k + \mathbf{c}_0^\dagger \mathbf{c}_0 + \mathbf{c}_\pi^\dagger \mathbf{c}_\pi$$

(for APBC we can simply omit the last term and conclude that $\mathbf{N}$ is even, so $(-1)^{\mathbf{N}} = 1$.) So for determining $(-1)^{\mathbf{N}}$, we only care about $\mathbf{h}_{ka=0}$ and $\mathbf{h}_{ka=\pi}$.

$$\mathbf{h}_{ka=\pi} = \mathbf{c}_\pi^\dagger \mathbf{c}_\pi\left(2g + 2\right) > 0 \quad \implies \quad \text{empty in the groundstate}, \forall g \geq 0$$

As usual, the zero-momentum mode is the hero of the story:

$$\mathbf{h}_{ka=0} = \mathbf{c}_0^\dagger \mathbf{c}_0\left(2g - 2\right)$$

which *changes sign* at $g = g_c = 1$:

- For $g > g_c$,, $\mathbf{h}_0 > 0$ is a positive energy cost for filling the zeromode. So the groundstate for $g > g_c$ has this odd mode empty and hence $(-1)^{\mathbf{N}} = 1$, and APBCs.

- For $g < g_c$,, $\mathbf{h}_0 < 0$ is an incentive for filling the zeromode, of which there is odd number (namely, 1), so the system can lower its energy by having PBC and filling the $k = 0$ mode. In the thermodynamic limit $L = Na \to \infty$, these two groundstates become degenerate.

Comments on the splitting between the groundstates

Comments about the L-dependence of the splitting between these states:

(1) We would like to identify them with (linear combinations of) the two ferromagnetic groundstates, which become the fully-polarized (in $\mathbf{Z}$) states $|+\rangle$ and $|-\rangle$ as $g \to 0$. As we've argued from many other points of view, these states are split by an amount which is exponentially small in L: $\Delta E \sim e^{-L/\xi}$.

(2) *At the critical point*, where the fermions are massless, the energy splitting between the groundstates with PBC and APBC (in the scaling limit, where $a \to 0$), is an oft-used result in string theory, and is $\frac{1}{16}\frac{\hbar c}{L}$ (PBC is higher energy). The L-dependence is determined by dimensional analysis, since there is no other scale. Away from the critical point I don't know the answer at the moment.

(3) Note that these two sectors with different $(-1)^{\mathbf{N}}$ do not mix via $\mathbf{H}_{\text{TFIM}}$, since $(-1)^{\mathbf{N}}$ is a symmetry generator; you can't create a single fermion using $\mathbf{H}_{\text{TFIM}}$. There is *no* tunneling between these vacua. So these two states must in fact be the cat states $\frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)$ which are eigenvectors of the $\mathbb{Z}_2$ symmetry of the TFIM, which, after all, is

$$\mathbf{S} = \prod_{j=1}^{N} \mathbf{X}_j = \prod_{j=1}^{N} (-1)^{\mathbf{c}_j^\dagger \mathbf{c}_j} = (-1)^{\mathbf{N}}.$$

(Recall that the symmetric cat state $\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ is the ground state in finite volume, so this must be the fermion state with antiperiodic boundary conditions.) Being forced to superpose states with different boundary conditions may be discomfiting, but it is part of the resolution of the puzzle of an odd number of domain walls on a circle.

2.2.8 Coherent state path integrals for fermions

[Shankar, *Principles of QM*, path integrals revisited. In this chapter of his great QM textbook, Shankar sneaks in lots of insights useful for modern condensed matter physics]

Consider a single fermion mode operator (as in our discussion of Jordan-Wigner in 0+1 dimensions);

$$\{\mathbf{c}, \mathbf{c}\} = 0, \quad \{\mathbf{c}^\dagger, \mathbf{c}^\dagger\} = 0, \quad \{\mathbf{c}, \mathbf{c}^\dagger\} = 1 \ .$$

With a single mode, the general Hamiltonian is

$$\mathbf{H} = \mathbf{c}^\dagger \mathbf{c} \, (\omega_0 - \mu)$$

(ω_0 and μ are (redundant when there is only one mode) constants). We might be interested in its thermal partition function

$$Z = \text{tr } e^{-\frac{\mathbf{H}}{T}} \ .$$

(In this example, it happens to equal $Z = 1 + e^{-\frac{\omega_0 - \mu}{T}}$, as you can see by computing the trace in the eigenbasis of $\mathbf{n} = \mathbf{c}^\dagger \mathbf{c}$. But never mind that; the one mode is a proxy for many, where it's not quite so easy to sum.) How do we trotterize this? That is, what is the corresponding classical system? One answer is to use the 0d Jordan-Wigner map from §2.2.5. That route takes us back where we came from. Here's another answer. We can do the Trotterizing using any resolution of the identity on $\mathcal{H}$, so there can be many very-different-looking answers to this question.

Let's define coherent states for fermionic operators:

$$\mathbf{c}|\psi\rangle = \psi|\psi\rangle. \tag{49}$$

Here ψ is a c-number (not an operator), but acting twice with $\mathbf{c}$ we see that we must have $\psi^2 = 0$. ψ is a grassmann number. These satisfy

$$\psi_1 \psi_2 = -\psi_2 \psi_1, \psi \mathbf{c} = -\mathbf{c} \psi \tag{50}$$

– they anticommute with each other and with fermionic operators, and commute with ordinary numbers and bosons. They seem weird but they are easy. We'll need to consider multiple grassmann numbers when we have more than one fermion mode, where $\{\mathbf{c}_1, \mathbf{c}_2\} = 0$ will require that they anticommute $\{\psi_1, \psi_2\} = 0$ (as in the definition (50)); note that we will be simultaneously diagonalizing operators which *anti*commute.

It's very easy to solve the equation (49):

$$|\psi\rangle = |0\rangle - \psi|1\rangle$$

where as above $|0\rangle$ is the empty state ($\mathbf{c}|0\rangle = 0$) and $|1\rangle = \mathbf{c}^\dagger|0\rangle$ is the filled state. (Check: $\mathbf{c}|\psi\rangle = \mathbf{c}|0\rangle - \mathbf{c}\psi|1\rangle = +\psi\mathbf{c}|1\rangle = \psi|0\rangle = \psi|\psi\rangle$.)

Similarly, the left-eigenvector of the creation operator is

$$\langle\bar\psi|\mathbf{c}^\dagger = \langle\bar\psi|\bar\psi, \quad \langle\bar\psi| = \langle 0| - \langle 1|\bar\psi = \langle 0| + \bar\psi\langle 1|.$$

Notice that these states are weird in that they are elements of an enlarged hilbert space with grassmann coefficients (usually we just allow complex numbers). Also, $\bar\psi$ is *not* the complex conjugate of ψ and $\langle\bar\psi|$ is not the adjoint of $|\psi\rangle$. Rather, their overlap is

$$\langle\bar\psi|\psi\rangle = 1 + \bar\psi\psi = e^{\bar\psi\psi}.$$

Grassmann calculus summary. In the last expression we have seen an example of the amazing simplicity of Taylor's theorem for grassmann functions:

$$f(\psi) = f_0 + f_1\psi \ .$$

Integration is just as easy and its the same as taking derivatives:

$$\int \psi d\psi = 1, \qquad \int 1 d\psi = 0.$$

With more than one grassmann we have to worry about the order:

$$1 = \int \bar\psi\psi d\psi d\bar\psi = -\int \bar\psi\psi d\bar\psi d\psi.$$

The only integral, really, is the gaussian integral:

$$\int e^{-a\bar\psi\psi} d\bar\psi d\psi = a.$$

Many of these give

$$\int e^{-\bar\psi\cdot A\cdot\psi} d\bar\psi d\psi = \det A.$$

Here $\bar\psi\cdot A\cdot\psi \equiv \left(\bar\psi_1,\cdots,\bar\psi_M\right)\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & \ddots & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix}\psi_1 \\ \vdots \\ \psi_M\end{pmatrix}$. One way to get this expression is to change variables to diagonalize the matrix A.

$$\langle\bar\psi\psi\rangle \equiv \frac{\int \bar\psi\psi e^{-a\bar\psi\psi} d\bar\psi d\psi}{\int e^{-a\bar\psi\psi} d\bar\psi d\psi} = -\frac{1}{a} = -\langle\psi\bar\psi\rangle.$$

If for many grassman variables we use the action $S = \sum_i a_i\bar\psi_i\psi_i$ (diagonalize A above) then

$$\langle\bar\psi_i\psi_j\rangle = \frac{\delta_{ij}}{a_i} \equiv \langle\bar i j\rangle \tag{51}$$

and Wick's theorem here is

$$\langle\bar\psi_i\bar\psi_j\psi_k\psi_l\rangle = \langle\bar i l\rangle\langle\bar j k\rangle - \langle\bar i k\rangle\langle\bar j l\rangle.$$

Back to quantum mechanics: The resolution of $\mathbb{1}$ in this basis is

$$\mathbb{1} = \int d\bar{\psi} d\psi \; e^{-\bar{\psi}\psi} |\psi\rangle\langle\bar{\psi}| \tag{52}$$

And if $\mathbf{A}$ is a bosonic operator (made of an even number of grassmann operators),

$$\mathrm{tr}\mathbf{A} = \int d\bar{\psi} d\psi \; e^{-\bar{\psi}\psi} \langle -\bar{\psi}|\mathbf{A}|\psi\rangle \; .$$

(Note the minus sign; it will lead to a deep statement.) So the partition function is:

$$Z = \int d\bar{\psi}_0 d\psi_0 \; e^{-\bar{\psi}_0\psi_0} \langle -\bar{\psi}_0| \underbrace{e^{-\frac{\mathbf{H}}{T}}}_{=\underbrace{(1-\Delta\tau\mathbf{H})\cdots(1-\Delta\tau\mathbf{H})}_{M \text{ times}}} |\psi_0\rangle$$

Now insert (52) in between each pair of Trotter factors to get

$$Z = \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l e^{-\bar{\psi}_l\psi_l} \langle\bar{\psi}_{l+1}|(1-\Delta\tau\mathbf{H})|\psi_l\rangle \; .$$

Because of the $-\bar{\psi}$ in (52), to get this nice expression we had to define an extra letter

$$\bar{\psi}_M = -\bar{\psi}_0, \quad \psi_M = -\psi_0 \tag{53}$$

so we could replace $\langle -\bar{\psi}_0| = \langle\bar{\psi}_N|$. [End of Lecture 10]

Now we use the coherent state property to turn the matrix elements into grassmann-valued functions:

$$\langle\bar{\psi}_{l+1}| \left(1-\Delta\tau H(\mathbf{c}^\dagger,\mathbf{c})\right)|\psi_l\rangle = \langle\bar{\psi}_{l+1}| \left(1-\Delta\tau H(\bar{\psi}_{l+1},\psi_l)\right)|\psi_l\rangle \overset{\Delta\tau\to 0}{=} e^{\bar{\psi}_{l+1}\psi_l} e^{-\Delta\tau H(\bar{\psi}_{l+1},\psi_l)}.$$

It was important that in $\mathbf{H}$ all $\mathbf{c}$s were to the right of all $\mathbf{c}^\dagger$s, *i.e.* that $\mathbf{H}$ was normal ordered.)

So we have

$$\begin{aligned} Z &= \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l e^{-\bar{\psi}_l\psi_l} e^{\bar{\psi}_{l+1}\psi_l} e^{-\Delta\tau H(\bar{\psi}_{l+1},\psi_l)} \\ &= \int \prod_{l=0}^{M-1} d\bar{\psi}_l d\psi_l \exp\left(\Delta\tau \left(\underbrace{\frac{\bar{\psi}_{l+1}-\bar{\psi}_l}{\Delta\tau}}_{=\partial_\tau\bar{\psi}} \psi_l - H(\bar{\psi}_{l+1},\psi_l)\right)\right) \\ &\simeq \int [D\bar{\psi}D\psi] \exp\left(\int_0^{1/T} d\tau \; \bar{\psi}(\tau)\left(-\partial_\tau - \omega_0 + \mu\right)\psi(\tau)\right) = \int [D\bar{\psi}D\psi] e^{-S[\bar{\psi},\psi]}. \end{aligned} \tag{54}$$

Points to note:

- In the penultimate step we defined, as usual, continuum fields
$$\psi(\tau_l = \Delta\tau l) \equiv \psi_l, \qquad \bar{\psi}(\tau_l = \Delta\tau l) \equiv \bar{\psi}_l.$$
- We elided the difference $H(\bar{\psi}_{l+1}, \psi_l) = H(\bar{\psi}_l, \psi_l) + \mathcal{O}(\Delta\tau)$ in the last expression. This difference is usually negligible and sometimes helpful (an example where it's helpful is the discussion of the number density below).
- The APBCs (53) on $\psi(\tau + \frac{1}{T}) = -\psi(\tau)$ mean that in its fourier representation[24]
$$\psi(\tau) = T\sum_n \psi(\omega) e^{-\mathbf{i}\omega_n \tau}, \qquad \bar{\psi}(\tau) = T\sum_n \bar{\psi}(\omega) e^{\mathbf{i}\omega_n \tau} \tag{55}$$
the *Matsubara frequencies*
$$\omega_n = (2n+1)\pi T, \quad n \in \mathbb{Z}$$
are half-integer multiples of πT.
- The measure $[D\bar{\psi} D\psi]$ is defined by this equation, just as in the bosonic path integral.
- The derivative of a grassmann function is also defined by this equation; note that $\psi_{l+1} - \psi_l$ is not 'small' in any sense.
- In the last step we integrated by parts, *i.e.* relabeled terms in the sum, so
$$\sum_l \left(\bar{\psi}_{l+1} - \bar{\psi}_l\right)\psi_l = \sum_l \bar{\psi}_{l+1}\psi_l - \sum_l \bar{\psi}_l\psi_l = \sum_{l'=l-1} \bar{\psi}_{l'}\psi_{l-1} - \sum_l \bar{\psi}_l\psi_l = -\sum_l \bar{\psi}_l\left(\psi_l - \psi_{l-1}\right).$$
Note that no grassmanns were moved through each other in this process.

The punchline of this discussion for now (for purposes of our study of the scaling limit of the TFIM) is that the euclidean action is

$$\boxed{S[\bar{\psi}, \psi] = \int d\tau \left(\bar{\psi}\partial_\tau \psi + H(\bar{\psi}, \psi)\right)}\,.$$

The first-order kinetic term we've found $\bar{\psi}\partial_\tau\psi$ is sometimes called a 'Berry phase term'. Note the funny-looking sign.

Continuum limit warning (about the red $\simeq$ in (54)). The Berry phase term is *actually*
$$\sum_{l=0}^{N-1} \bar{\psi}_{l+1}\left(\psi_{l+1} - \psi_l\right) = T\sum_{\omega_n} \bar{\psi}(\omega_n)\left(1 - e^{\mathbf{i}\omega_n\tau}\right)\psi(\omega_n)$$

[24] $\bar{\psi}$ is still not the complex conjugate of ψ but the relative sign is convenient.

and in (54) we have kept only the leading nonzero term:

$$\left(1 - e^{\mathbf{i}\omega_n\tau}\right) \to \mathbf{i}\omega_n\tau.$$

Clearly this replacement is just fine if

$$\omega_n\tau \ll 1$$

for all ω_n which matter. Which ω_n contribute? I claim that if we use a reasonable $\mathbf{H} = \mathbf{H}_{\text{quadratic}} + \mathbf{H}_{\text{int}}$, reasonable quantities like Z, $\langle\mathcal{O}^\dagger\mathcal{O}\rangle$, are dominated by $\omega_n \ll \tau^{-1}$.

There's more we can learn from what we've done here that I don't want to pass up. Let's use this formalism to compute the fermion density at $T = 0$:

$$\langle\mathbf{N}\rangle = \frac{1}{Z}\mathrm{tr}e^{-\mathbf{H}/T}\mathbf{c}^\dagger\mathbf{c}.$$

This is an example where the annoying $\Delta\tau$s in the path integral not only matter, but are extremely friendly to us.

Frequency space, $T \to 0$.

Let's change variables to frequency-space fields, which diagonalize S. The Jacobian is 1 (since fourier transform is unitary):

$$D\bar\psi(\tau)D\psi(\tau) = \prod_n d\bar\psi(\omega_n)d\psi(\omega_n) \overset{T\to 0}{\to} D\bar\psi(\omega)D\psi(\omega).$$

The partition function is

$$Z = \int D\bar\psi(\omega)D\psi(\omega)\ \exp\left(T\sum_{\omega_n}\bar\psi(\omega_n)\left(\mathbf{i}\omega_n - \omega_0 + \mu\right)\psi(\omega_n)\right).$$

Notice that in the zero-temperature limit

$$T\sum_{\omega_n} \mapsto \int\frac{d\omega}{2\pi} \equiv \int \text{đ}\omega.$$

(This is the same fact as $V\sum_k \mapsto \int \text{đ}^d k$ in the thermodynamic limit.) So the zero-temperature partition function is

$$Z \overset{T\to 0}{=} \int D\bar\psi(\omega)D\psi(\omega)\ \exp\left(\int_{-\infty}^{\infty}\text{đ}\omega\bar\psi(\omega)\left(\mathbf{i}\omega - \omega_0 + \mu\right)\psi(\omega)\right).$$

Using the gaussian-integral formula (51) you can see that the *propagator* for ψ is

$$\langle\bar\psi(\omega_1)\psi(\omega_2)\rangle = \underbrace{\frac{\delta_{\omega_1,\omega_2}}{T}}_{\overset{T\to 0}{\to}\delta(\omega_1-\omega_2)}\ \frac{2\pi}{\mathbf{i}\omega_1 - \omega_0 + \mu}. \tag{56}$$

In particular $\langle\bar\psi(\omega)\psi(\omega)\rangle = \frac{2\pi/T}{\mathbf{i}\omega-\omega_0+\mu}$. $\delta(\omega=0) = 1/T$ is the 'volume' of the time direction.

Back to the number density. Using the same strategy as above, we have

$$\langle\mathbf{N}\rangle = \frac{1}{Z}\int \prod_{l=0}^{M-1+1}\left(d\bar\psi_l d\psi_l e^{-\bar\psi_l\psi_l}\right)\prod_{l=1}^{M-1}\langle\bar\psi_{l+1}|(1-\Delta\tau\mathbf{H}(\mathbf{c}^\dagger\mathbf{c}))|\psi_l\rangle \underbrace{\langle\bar\psi_{N+1}|\mathbf{c}^\dagger\mathbf{c}|\psi_N\rangle}_{=\bar\psi_{N+1}\psi_N=\bar\psi(\tau_N+\Delta\tau)\psi(\tau_N)}.$$

This formula has a built-in point-splitting of the operators!

$$\begin{aligned}\langle\mathbf{N}\rangle &= \frac{1}{Z}\int D\bar\psi D\psi\; e^{-S[\bar\psi,\psi]}\bar\psi(\tau_N+\Delta\tau)\psi(\tau_N)\\ &= \int_{-\infty}^{\infty}\text{đ}\omega\frac{e^{\mathbf{i}\omega\Delta\tau}}{\mathbf{i}\omega-\omega_0+\mu} = \theta(\mu-\omega_0).\end{aligned} \tag{57}$$

Which is the right answer: the mode is occupied in the groundstate only if $\omega_0 < \mu$. In the last step we used the fact that $\Delta\tau > 0$ to close the contour in the UHP; so we only pick up the pole if it is in the UHP. Notice that this quantity is very *UV sensitive*: if we put a frequency cutoff on the integral, $\int^\Lambda \frac{d\omega}{\omega} \sim \log\Lambda$, the integral diverges logarithmically. For most calculations the $\Delta\tau$ can be ignored, but here it told us the right way to treat the divergence. [25]

Bosons. This is a good place to mention quickly that we can do the same thing for bosons, using ordinary SHO (simple harmonic oscillator) coherent states. What I mean by 'bosons' is a many-body system whose Hilbert space can be written as $\mathcal{H} = \otimes_k\mathcal{H}_k$ where k is a label (could be real space, could be momentum space) and $\mathcal{H}_k = \text{span}\{|0\rangle_k, \mathbf{a}_k^\dagger|0\rangle_k, \frac{1}{2}\left(\mathbf{a}_k^\dagger\right)^2|0\rangle_k, ...\}$ is the SHO Hilbert space.

For each normal mode $\mathbf{a}$, coherent states are

$$\mathbf{a}|\phi\rangle = \phi|\phi\rangle; \qquad |\phi\rangle = \mathcal{N}e^{\phi\mathbf{a}^\dagger}|0\rangle.$$

The eigenbra of $\mathbf{a}$ is $\langle\phi|$, with

$$\langle\phi|\mathbf{a} = \langle\phi|\phi^\star, \quad \langle\phi| = \langle 0|e^{+\phi^\star\mathbf{a}}\mathcal{N}.$$

[25] The calculation between the first and second lines of (57) is familiar to field theorists – it can be described as a feynman diagram with one line between the two insertions. More prosaically, it is

$$\langle\bar\psi(\tau_N+\Delta\tau)\psi(\tau_N)\rangle \overset{(55)}{=} T^2\sum_{nm}e^{\mathbf{i}(\omega_n-\omega_m)\tau+\mathbf{i}\omega_n\Delta\tau}\langle\bar\psi(\omega_n)\psi(\omega_m)\rangle \overset{(56)}{=} T\sum_m\frac{e^{\mathbf{i}\omega_n\Delta\tau}}{\mathbf{i}\omega_n-\omega_0+\mu}\overset{T\to 0}{\to}\int\text{đ}\omega\frac{e^{\mathbf{i}\omega\Delta\tau}}{\mathbf{i}\omega-\omega_0+\mu}.$$

Their overlap is[26]:

$$\langle\phi_1|\phi_2\rangle = e^{\phi_1^\star\phi_2}.$$

If we choose $\mathcal{N} = e^{-|\phi|^2/2}$, they are normalized, but it is more convenient to set $\mathcal{N} = 1$. The overcompleteness relation on $\mathcal{H}_k$ is

$$\mathbb{1}_k = \int \frac{d\phi d\phi^\star}{\pi} e^{-|\phi|^2} |\phi\rangle\langle\phi|.$$

It will be convenient to arrange all our operators into sums of *normal-ordered* operators:

$$:\mathbf{a}_k\mathbf{a}_l^\dagger: =: \mathbf{a}_l^\dagger\mathbf{a}_k := \mathbf{a}_l^\dagger\mathbf{a}_k$$

with all annihilation operators to the right of all creation operators. Coherent state expectation values of such operators can be built from the monomials

$$\langle\phi| \prod_k \left(\mathbf{a}_k^\dagger\right)^{M_k} (\mathbf{a}_k)^{N_k} |\phi\rangle = \prod_k (\phi_k^\star)^{M_k} (\phi_k)^{N_k}.$$

Let the Hamiltonian be $\mathbf{H} = H(\{\mathbf{a}_k^\dagger\}, \{\mathbf{a}_k\}) =: \mathbf{H}:$, normal ordered. By now you know how to derive the path integral using this resolution of the identity:

$$\begin{aligned} Z &= \mathrm{tr} e^{-\mathbf{H}/T} \\ &= \int_{\phi_{N+1}=\phi_0} \prod_{l=0}^{N} d\phi_l \; e^{-\sum_{l=0}^{N}\left(\phi_{l+1}(\phi_{l+1}-\phi_l) - \Delta\tau H(\phi_{l+1}^\star,\phi_l)\right)} \\ &\simeq \int_{\phi(0)=\phi(1/T)} [D\phi] \; e^{-\int_0^{1/T} d\tau(\phi^\star\partial_\tau\phi + H(\phi^\star,\phi))}. \end{aligned} \tag{58}$$

Real time. If you are interested in real-time propagation, rather than euclidean time, just replace the euclidean propagator $e^{-\tau\mathbf{H}} \mapsto e^{-\mathbf{i}t\mathbf{H}}$. The result, for example, for the amplitude to propagate from one bose coherent state to another is

$$\langle\phi_f, t_f|e^{-\mathbf{i}t\mathbf{H}}|\psi_0, t_0\rangle = \int_{\phi(t_0)=\phi_0}^{\phi(t_f)=\phi_f} D\phi^\star D\phi \; e^{\frac{\mathbf{i}}{\hbar}\int_{t_0}^{t_f} dt(\mathbf{i}\hbar\phi^\star\partial_t\phi - H(\phi,\phi^\star))}.$$

Note the distinguishing feature of the Berry phase term that it produces a complex term in the real-time action.

---

[26] Check this by expanding the coherent states in the number basis and doing the integrals

$$\int \frac{d\phi d\phi^\star}{\pi} e^{-\phi\phi^\star} \phi^n (\phi^\star)^{n'} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{\mathbf{i}(n-n')\theta} \int_0^\infty du e^{-u} u^{\frac{n+n'}{2}}$$

to get $\mathbb{1} = \sum_n |n\rangle\langle n|$.

2.2.9 Scaling near the fixed-point theory of the Ising phase transition.

Armed with the fermion coherent state path integral, let's go back to the continuum hamiltonian describing the TFIM near $g = 1$:

$$\mathbf{H} \rightsquigarrow \frac{c}{2}\int dx\ \left(\mathbf{\Psi}(x)^\dagger \partial_x \mathbf{\Psi}(x)^\dagger - \mathbf{\Psi}(x)\partial_x \mathbf{\Psi}(x)\right) + \Delta \int dx \mathbf{\Psi}^\dagger \mathbf{\Psi} \equiv \int dx\ \mathfrak{h}$$

with $\Delta = 2J|g-1|$.

The euclidean-time action is $S[\Psi, \Psi^\dagger] = \int d\tau \int dx \mathcal{L}$, with

$$\begin{aligned} \mathcal{L} &= \bar{\Psi}\partial_\tau \Psi + \mathfrak{h} \\ &= \bar{\Psi}\partial_\tau \Psi + \frac{c}{2}\left(\bar{\Psi}\partial_x \bar{\Psi} - \Psi \partial_x \Psi\right) + \Delta \bar{\Psi}\Psi. \end{aligned} \tag{59}$$

Here I have simply replaced the field operators $\mathbf{\Psi}(x), \mathbf{\Psi}^\dagger(x)$ with their (right and left, respectively) grassmann eigenvalues $\Psi(x), \bar{\Psi}$. This lagrangian is more appealing in terms of the self-conjugate variables $\psi_\pm$:

$$\Psi = \psi_+ + \psi_- + \mathbf{i}\left(\psi_+ - \psi_-\right), \quad \bar{\Psi} = \psi_+ + \psi_- - \mathbf{i}\left(\psi_+ - \psi_-\right)$$

$$\mathcal{L} = \sum_\pm \psi_\pm \left(\partial_\tau \pm \mathbf{i}\partial_x\right)\psi_\pm + \Delta \psi_+ \psi_-.$$

If we let $z \equiv x + \mathbf{i}\tau$, and turn off the mass perturbation, this is the third avatar of the Ising model I mentioned at the beginning in (6).

Let's analyze its behavior under scale transformations. In order to make S invariant, we must scale time and space the same way ($z = 1$), and we must scale

$$x \to \lambda x, t \to \lambda t, \Psi \to \lambda^{-\frac{1}{2}}\Psi.$$

So just by scaling, we can see that its correlators behave as

$$\langle \mathbf{\Psi}(x)^\dagger \mathbf{\Psi}(0)\rangle \sim \frac{1}{x}.$$

(Since it's a gaussian theory, you can also calculate the coefficient.) More generally, the scaling dimension δ of an operator $\mathcal{O}(x)$ in $\mathcal{O} \to \lambda^{-\delta}\mathcal{O}$ in a scale-invariant QFT determines its vacuum autocorrelation functions to be

$$\langle \mathcal{O}(x)^\dagger \mathcal{O}(0)\rangle \sim \frac{1}{x^{2\delta}}.$$

(Later we will interpret δ as an eigenvalue of the dilatation operator.)

The mass perturbation Δ will violate the scale invariance, since

$$\int dx d\tau \bar{\Psi}\Psi \to \lambda^{+1} \int dx d\tau \bar{\Psi}\Psi\ .$$

As we make lengths and times bigger, it gets bigger – it's a *relevant* perturbation, which determines the behavior at long distances. The scaling dimension of this relevant perturbation determines the correlation length critical exponent: the correlation length is a length scale $\xi = a\lambda$ at which the relevant coupling $\Delta = \lambda^\nu \Delta_0 \sim 1$ has evolved to an order-one value (Δ_0 is its 'bare' value, at the lattice scale a, and ν is its scaling dimension). Eliminating $\lambda = \xi/a$ from these equations, we get

$$1 = (\xi/a)^\nu \Delta_0 \implies \xi \sim \frac{1}{\Delta_0^\nu},$$

so ν, the scaling dimension of the relevant operator, is indeed the correlation length critical exponent defined above.

Let's ask whether there are any other *relevant* operators, in the sense that they scale to larger values at larger λ (longer wavelengths). If there are more it means that our fixed point is *multicritical* – we would have to tune more than one parameter to reach it. We're going to demand that the Ising symmetry is preserved; this is the fermion number symmetry, $\Psi(x) \to -\Psi(x)$. So we should only consider operators made of an even number of fermions. And all the other operators we can make out of an even number of Ψ are *irrelevant*, in the sense that they scale to smaller values at larger λ. The marginal-looking operator $\bar\Psi\Psi\bar\Psi\Psi$ vanishes by Fermi statistics. Other operators have more derivatives have no chance:

$$\int dx d\tau \bar\Psi \partial_x^2 \Psi \sim \lambda^{-1}$$

$$\int dx d\tau \bar\Psi \partial_x \Psi \bar\Psi \partial_x \Psi \sim \lambda^{-2}$$

This means we have to tune only one parameter ($g - g_c \sim \Delta$) to get to the critical theory. (Since there are no classically-marginal operators, there is no need to do any perturbating to determine the effects of turning them on a little bit – small quantum corrections cannot overcome the order-1 amounts by which these operators are irrelevant.)

There *are* operators of dimension less than two, namely the operator $\mathbf{Z}$ itself; this is an operator which creates a *branch cut* in the fermions, which after all are made from the domain wall operators. To understand its dimension, it is useful to use a bosonized description, which we will do in the next section. Its scaling dimension turns out to be 1/8. This is important because it determines the rate at which the magnetization vanishes as $g \to g_c$ from below. It's important that the $\mathbb{Z}_2$ symmetry forbids us from adding this (relevant) operator to the action. So in the presence of the symmetry, the Ising fixed point has only one relevant perturbation (the fermion mass).

Notice how different this critical theory is from the guess would have made from the mean field theory above in §2.2.2: we would have guessed that the critical fluctuations should be those of the mean field θ, which when I set $g = g_c$ becomes a massless free boson, with many relevant operators which preserve the $\theta \to -\theta$ symmetry, since demanding scale invariance of

$$S_0[\theta] = \int dx d\tau \ \left((\partial_\tau \theta)^2 + (\partial_x \theta)^2\right)$$

implies that θ is dimensionlesss. So all even powers $\theta^2, \theta^4 ...$ (in fact, all even functions, like $\cos\theta$) have a chance to be relevant. We'll see that it's *not* actually quite so far off. (And in fact is the correct variables in $d > 1$, and the correct fixed point for $d \geq 3$ (the upper critical dimension).)

The nontrivial scaling of $\mathbf{Z}$ is an example of an *anomalous dimension*: the engineering dimension is different from the correct scaling behavior. Dimensional analysis is inviolable. What has happened is that powers of the short-distance scales are involved; that is: it is $J^{1/8}\mathbf{Z}$ (or $a^{-1/8}\mathbf{Z}$) that has finite correlation functions in the scaling limit. In relativistic QFT this phenomenon is called 'wavefunction renormalization' – we must multiplicatively renormalize our operators to get finite (not zero, not infinity) correlators in the long-wavelength limit.

2.2.10 Beyond the quantum Ising chain

The Jordan-Wigner transformation is useful in spin chains with many other choices of Hamiltonian and symmetry group, as you'll see on the problem set, and below.

Here we have a choice to make. Do we study systems which have more symmetry at each site or do we study ising spins in more dimensions? We'll get to both eventually.

[End of Lecture 11]

3 Bosons and Bosonization

I mentioned that another realization of two-state systems at each site comes from self-loathing bosons. Let's see how that comes about in a model with $\mathsf{U}(1)$ symmetry.

3.0.1 Spins, bosons, fermions

Consider the following spin system (qbits at the sites of a lattice)

$$\mathbf{H}_{XX} = -\frac{w}{2}\sum_{\langle ij\rangle}(\mathbf{X}_i\mathbf{X}_j + \mathbf{Y}_i\mathbf{Y}_j) + \frac{\mu}{2}\sum_j \mathbf{Z}_j.$$

On the homework, you Jordan-Wignered the 1d version of this[27] and found

$$\mathbf{H}_{XX} = -\sum_j \left(w\left(\mathbf{c}^\dagger_{j+1}\mathbf{c}_j + h.c.\right) + \mu\mathbf{c}^\dagger_j\mathbf{c}_j\right)$$

(Note that μ is playing the role of the transverse-field term.) The fact that we have equal 'ferromagnetic' couplings involving $\mathbf{X}$s and $\mathbf{Y}$s means that this system preserves a continuous $\mathsf{U}(1)$ symmetry, whose unitary implementation is

$$\mathbf{U}_\theta = \prod_j e^{\mathbf{i}\frac{\theta}{2}\mathbf{Z}_j}.$$

It acts on the spin operators by

$$\mathbf{X}_j \mapsto \mathbf{U}_\theta\mathbf{X}_j\mathbf{U}^\dagger_\theta = \cos\theta\mathbf{X}_j + \sin\theta\mathbf{Y}_j, \quad \mathbf{Y}_j \mapsto \mathbf{U}_\theta\mathbf{Y}_j\mathbf{U}^\dagger_\theta = \cos\theta\mathbf{Y}_j - \sin\theta\mathbf{X}_j, \quad \mathbf{Z}_j \mapsto \mathbf{Z}_j.$$

(Please beware my signs.) It therefore commutes with $\mathbf{H}_{XX}$. In terms of the fermions it acts by $\mathbf{c} \to e^{\mathbf{i}\theta}\mathbf{c}$: it is the total fermion particle number $\sum_j \mathbf{n}_j \equiv \sum_j \mathbf{c}^\dagger_j\mathbf{c}_j$ that is conserved as a result, $[\sum_j \mathbf{n}_j, \mathbf{H}] = 0$ – not just mod two like when we had only $\mathbb{Z}_2$ symmetry.

The fermion description is nice because it is quadratic; the fermions are free *in this subspace* of the phase diagram. If we add the nearest-neighbor $\mathbf{Z}_i\mathbf{Z}_j$ term, we add

$$\Delta\mathbf{H} = -v\sum_{\langle ij\rangle}\mathbf{Z}_i\mathbf{Z}_j = -v\sum_{\langle ij\rangle}\mathbf{n}_i\mathbf{n}_j,$$

which is quartic in the fermion operators, and therefore not simple. A description in terms of bosons will allow us to understand this direction in the phase diagram.

[27] using the (rotated $X \to Z$ relative to our previous one) map

$$\mathbf{Z}_j = 1 - 2\mathbf{c}^\dagger_j\mathbf{c}_j, \quad \mathbf{X}_j = -\prod_{i>j}(1-2\mathbf{c}^\dagger_i\mathbf{c}_i)\left(\mathbf{c}_j + \mathbf{c}^\dagger_j\right) = -\prod_{i>j}(-1)^{\mathbf{c}^\dagger_i\mathbf{c}_i}\left(\mathbf{c}_j + \mathbf{c}^\dagger_j\right), \quad \mathbf{Y}_j = -\mathbf{i}\mathbf{Z}_j\mathbf{X}_j$$

An explicit map from spin-$\frac{1}{2}$s to hardcore bosons is:

$$\mathbf{X}_j = \mathbf{b}_j + \mathbf{b}_j^\dagger, \quad \mathbf{Y}_j = -\mathbf{i}\left(\mathbf{b}_j - \mathbf{b}_j^\dagger\right), \quad \mathbf{Z}_j = 1 - 2\mathbf{b}_j^\dagger \mathbf{b}_j. \tag{60}$$

(The inverse is $\mathbf{b} = \boldsymbol{\sigma}^+$). Unlike Jordan-Wigner, we don't have to attach any kind of string to anybody to make this work: before the hardcore projection, the bosons *commute* at different sites

$$[\hat{\mathbf{b}}_i, \hat{\mathbf{b}}_j^\dagger] = \delta_{ij} \ .$$

The price of that niceness is that the map (60) relates the qbit hilbert space to the subspace of the (much bigger!) boson Hilbert space where the occupation numbers $\mathbf{b}_j^\dagger \mathbf{b}_j \equiv \mathbf{n}_j^b = 0$ or 1. And the objects in (60) only satisfy the Pauli algebra on this subspace (check it!). Worse, the boson operators, projected to the no-double-occupation ($\mathbf{b}_j^\dagger \mathbf{b}_j \leq 1 \forall j$) subspace are no longer canonically conjugate. The projected bosons

$$\mathbf{b} \equiv \mathbf{P}\hat{\mathbf{b}}\mathbf{P}, \mathbf{b}^\dagger \equiv \mathbf{P}\hat{\mathbf{b}}^\dagger\mathbf{P}, \qquad \mathbf{P} \equiv |0\rangle\langle 0| + |1\rangle\langle 1|$$

instead satisfy

$$[\mathbf{b}, \mathbf{b}^\dagger] = 1 - 2\mathbf{b}^\dagger\mathbf{b}. \tag{61}$$

Check it by acting with the both-hand side on $|n\rangle$ with $n = 0, 1$. (61) in fact says that the projected hardcore bosons *at a single site* anticommute to 1 . But for different sites they commute:

$$[\mathbf{b}_i, \mathbf{b}_j^\dagger] = \delta_{ij}\left(1 - 2\mathbf{b}_j^\dagger\mathbf{b}_j\right).$$

---

(60) is not the only way to map spins to bosons. The bosons above are closely related to *Holstein-Primakoff* bosons ($[\mathbf{h}, \mathbf{h}^\dagger] = 1$), which satisfy

$$\boldsymbol{\sigma}^+ \equiv \frac{1}{2}(\mathbf{X} + \mathbf{i}\mathbf{Y}) = \sqrt{2S - \mathbf{h}^\dagger\mathbf{h}}\,\mathbf{h} \quad , \boldsymbol{\sigma}^- = \left(\boldsymbol{\sigma}^+\right)^\dagger = \mathbf{h}^\dagger\sqrt{2S - \mathbf{h}^\dagger\mathbf{h}}, \quad \mathbf{Z} = 2(S - \mathbf{h}^\dagger\mathbf{h}), \tag{62}$$

with the (non-holonomic) constraint $0 \leq \mathbf{h}^\dagger\mathbf{h} \leq 2S$. S is the spin – $2S + 1$ is the number of states per site. These have the virtue that within the constrained subspace, the canonical boson commutators $[\mathbf{h}, \mathbf{h}^\dagger] = 1$ imply the Pauli spin algebra. The price is the sqrt factors.

A different map which has no sqrts or weird commutators, and makes the $\mathsf{SU}(2)$ action on the spins manifest, is *Schwinger bosons*:

$$\boldsymbol{\sigma}^+ = \mathbf{a}^\dagger\mathbf{b}, \quad \frac{1}{2}(\mathbf{X} - \mathbf{i}\mathbf{Y}) = \mathbf{b}^\dagger\mathbf{a}, \quad \mathbf{Z} = \frac{1}{2}\left(\mathbf{a}^\dagger - \mathbf{b}^\dagger\right).$$

We get spin S if we restrict to the sector with $\mathbf{a}^\dagger\mathbf{a} + \mathbf{b}^\dagger\mathbf{b} = 2S$ (which is a nice holonomic constraint that we can impose with a lagrange multiplier). If I don't satisfy your thirst for learning about this, see Dan Arovas' notes here.

Persevering nevertheless with our hardcore projected bosons, let's write the hamiltonian above in these variables:

$$\mathbf{H}_{XX} = -w \sum_{\langle ij \rangle} \left(\mathbf{b}_i^\dagger \mathbf{b}_j + \mathbf{b}_j^\dagger \mathbf{b}_i \right) - \mu \sum_j \mathbf{b}_j^\dagger \mathbf{b}_j \quad .$$

A lovely hopping term and a chemical potential for the boson number, $\sum_j \mathbf{n}_j^b = \sum_j \mathbf{b}_j^\dagger \mathbf{b}_j$. We can make the projection to no-double-occupancy more palatable by implementing it energetically (like we did the gauge constraint in the toric code construction above):

$$\mathbf{H}_B = -w \sum_{\langle ij \rangle} \left(\hat{\mathbf{b}}_i^\dagger \hat{\mathbf{b}}_j + \hat{\mathbf{b}}_j^\dagger \hat{\mathbf{b}}_i \right) - \mu \sum_j \mathbf{n}_j^b + U \sum_j \mathbf{n}_j^b \left(\mathbf{n}_j^b - 1 \right) . \quad .$$

where B is for 'big' or 'boson'. This last term is zero if $\mathbf{n}_j^b = 0, 1$, but exacts a huge energetic penalty $\Delta E = U$ if a site is occupied by two bosons. So the low-energy physics of $\mathbf{H}_B$ is that of $\mathbf{H}_{XX}$, with its slightly weird bosons.

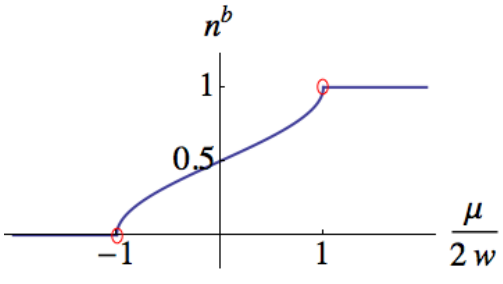
3.0.2 Using the JW solution to learn about the bosons

The free fermion solution is great when it is available, but it doesn't help much away from the solvable subspace of the parameters. Let's see what it says about the bosons and then we'll move away from the solvable subspace.

$$\mathbf{H}_{XX} = \sum_k \epsilon_k \mathbf{c}_k^\dagger \mathbf{c}_k, \quad \epsilon_k = -2w \cos ka - \mu.$$

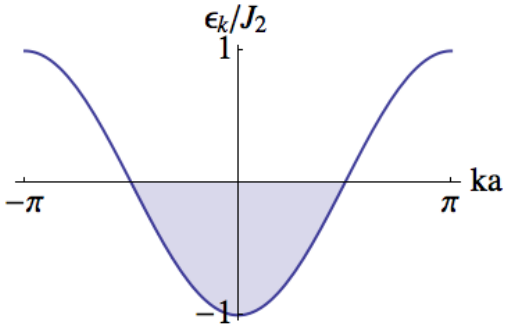
The boson number is the same as the fermion number, whose groundstate expectation value we can read off from the dispersion ϵ_k by counting negative-energy states:

$$\langle \mathbf{b}_j^\dagger \mathbf{b}_j \rangle = \frac{1}{2} \left(1 - \langle \mathbf{Z}_j \rangle \right) = \langle \mathbf{c}_j^\dagger \mathbf{c}_j \rangle = \begin{cases} 0, & \mu \leq -2w \quad \text{(I)} \\ 1 - \frac{1}{\pi} \cos^{-1} \left(\frac{\mu}{2w} \right), & |\mu| \leq 2w \quad \text{(II)} \\ 1, & \mu \geq 2w \quad \text{(III)} \end{cases}$$



The groundstates at $|\mu| > 2w$ are fully-polarized states of the spins – the spins all point up or all point down in the $\mathbf{Z}$-basis (like $g \to \infty$ of the TFIM). This is a paramagnetic state where the $\mathsf{U}(1)$ symmetry is preserved, since it rotates the spins about the $\mathbf{Z}$ axis. There is no long-range order, because there is no $\mathbf{Z} \to -\mathbf{Z}$ symmetry: the spins are told which way to point. In terms of the bosons, $\mu < -2w$ is a state where there are no bosons at all, a completely trivial insulator. The state at $\mu > 2w$ is more interesting. Here there is exactly one boson per site: this is a *Mott insulator*: the bosons can move in principle, but do not because they would have to go near their despised neighbor.

These states have a gap. In terms of the fermions, this is because the band is completely empty ($\mu < -2w$) or completely filled ($\mu > 2w$).



In the intermediate range of μ, we have a compressible superfluid state. 'Compressible' means that as we vary μ a little, the density changes a little, ($\frac{\partial n}{\partial \mu} \neq 0$) as you can see from the figure. 'Superfluid' is a little harder to explain. Note that although I say this is a superfluid, there is not in fact long-range order in this case, because of we have continuous symmetry in $d < 2$: the spins can't decide in which direction in the XY-plane to point, so in fact $\langle b \rangle = \frac{1}{2}\langle \mathbf{X} + \mathbf{iY} \rangle = 0$. This is a consequence of the Hohenberg-Mermin-Wagner-Coleman theorem which I hope to discuss more later. One thing you can see clearly, however, is that the state is gapless, since there is a Fermi surface. In fact, a finite compressibility requires gaplessness: since $[\mathbf{H}, \sum_j \mathbf{n}_j] = 0$, changing μ does not change the eigenvectors of $\mathbf{H}$, only the eigenvalues; so only if there is an unoccupied mode just above the chemical potential can we change the density by varying μ a little.

In summary, we have:

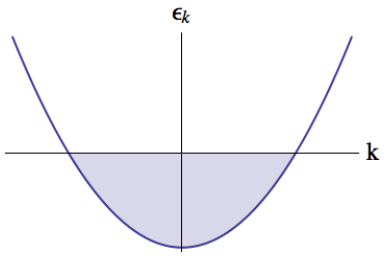
$$
\begin{aligned}
|\mathrm{gs, I}\rangle &= |\uparrow\uparrow\uparrow\cdots\uparrow\rangle = \Big|\ \Big\rangle = |\ \rangle \\
|\mathrm{gs, II}\rangle &= \Big|\ \Big\rangle \\
|\mathrm{gs, III}\rangle &= |\downarrow\downarrow\downarrow\cdots\downarrow\rangle = \Big|\ \Big\rangle = |\ \rangle
\end{aligned}
\tag{63}
$$

In region II we don't have a simple description of the states in terms of the spin or the bosonic variables so far.

The points $\mu/2w = \pm 1$ therefore describe Mott-superfluid and superfluid-Mott transitions (in terms of the fermions, they are band-insulator to metal transitions). The boson density demonstrates that there is critical behavior: near the $\mu/2w = -1$ transition, it is

$$
\langle \mathbf{n}_j^b \rangle \simeq \begin{cases} 0, & \mu < -2w \\ \left(1 + \frac{\mu}{2w}\right)^{1/2}, & \mu \gtrsim -2w \end{cases}
$$

3.0.3 Continuum description



The continuum description near the critical point $\mu + 2w \sim 0$ is

$$\mathbf{H} = \int dx \left(\mathbf{\Psi}^\dagger \frac{\nabla^2}{2m} \mathbf{\Psi} + \mu_F \mathbf{\Psi}^\dagger \mathbf{\Psi} \right)$$

where the fermion chemical potential is $\mu_F = \mu + 2w$. The $\mathsf{U}(1)$ symmetry acts as $\mathbf{\Psi} \to e^{\mathsf{i}\theta} \mathbf{\Psi}$. The inertial fermion mass here is $m_I = \frac{1}{2wa^2}$. On the other hand, the 'rest mass' (the energy to create a fermion at rest) is determined by μ_F (it's equal to $-\mu_F$ for $\mu_F < 0$).

As we did for the TFIM critical theory, let's study the scaling of possible perturbations of this critical theory. The euclidean action for the critical theory ($\mu_F = 0$) is

$$\begin{aligned} S[\bar{\Psi}, \Psi] &= \int d\tau dx \left(\bar{\Psi} \partial_\tau \Psi + \mathfrak{h} \right) \\ &= \int d\tau dx \left(\bar{\Psi} \left(\partial_\tau + \frac{\nabla^2}{2m} \right) \Psi \right). \end{aligned} \tag{64}$$

Notice that we have to scale time and space differently to make S invariant: $(x, \tau) \to (\lambda x, \lambda^2 \tau)$ so that both terms scale the same way. Thus, this fixed point has dynamical exponent $z = 2$. The fermion scaling is $\Psi \to \lambda^{-\frac{1}{2}} \Psi$ (again). Notice that $z \neq 1$ makes the measure $\int d\tau dx$ scale like $\lambda^{d+z} = \lambda^3$ – it's as if we were in higher dimensions. A marginal operator has dimension $d + z = 3$.

The operator $\bar{\Psi}\Psi$ changes μ_F and has dimension 1; this is the relevant operator we have to tune to get to the fixed point.

The operator $\bar{\Psi}\partial_x \Psi$ has dimension 2 and seems to be relevant. However: (1) it breaks the parity symmetry $x \to -x$ of the fixed point and so setting its coefficient to zero is technically natural. More importantly, (2) turning it on will merely move the minimum of the band away from $ka = \pi$ – it will not destroy the fact that the band has a minimum! So this just renormalizes the actual critical value of μ_F, which was non-universal anyway.

There are no other relevant operators. The four-fermion operator $\Psi^\dagger \Psi \Psi^\dagger \Psi = 0$ again. This means that the leading four-fermion operator coming from $\mathbf{Z}_j \mathbf{Z}_j$ is actually $\Psi^\dagger \partial_x \Psi \Psi^\dagger \partial_x \Psi$ which is irrelevant.

3.0.4 Fermi liquid in $d = 1 + 1$.

Now consider what happens for $\mu_F > 0$ ($|\mu| < 2w$). The groundstate in this regime (returning for a moment to the microscopic variables) is

$$\prod_{|k|\leq k_F} \mathbf{c}_k^\dagger |\Omega\rangle, \quad (\mathbf{c}_k|\Omega\rangle \equiv 0, \forall k)$$

With $\epsilon_k = -J\cos ka$, the Fermi surface is two points. For $\mu_F \ll w$, we can use the quadratic dispersion at the bottom of the band and the lowest-energy excitations above the groundstate occur at $k = \pm k_F = \pm\sqrt{2m\mu}$. Let's re-expand our continuum field to focus on these modes:

$$\mathbf{\Psi}(x,\tau) = e^{\mathbf{i}k_F x}\mathbf{R}(x,\tau) + e^{-\mathbf{i}k_F x}\mathbf{L}(x,\tau) \tag{65}$$

where the names arise from the direction of propagation of the modes.

$$\mathbf{H}_F = \int dx \left(\mathbf{R}^\dagger\left(-\mathbf{i}v_F\partial_x\right)\mathbf{R} + \mathbf{L}^\dagger\left(\mathbf{i}v_F\partial_x\right)\mathbf{L}\right) \equiv \int dx\ \mathfrak{h}_F$$

with $v_F = \partial_k\epsilon_k = 4Ja$ is the *Fermi velocity*. To do a scaling analysis it will help to use the action. The real-time Lagrangian is

$$\begin{aligned} L_F &= \int dx \left(\mathbf{i}L^\dagger\partial_t L + \mathbf{i}R^\dagger\partial_t R - \mathfrak{h}_F\right) \\ S_F &= \mathbf{i}\int dt \int dx \left(R^\dagger\left(\partial_t - v_F\partial_x\right)R + L^\dagger\left(\partial_t + v_F\partial_x\right)L\right) \end{aligned} \tag{66}$$

Now there are *no* symmetry-preserving ($\mathsf{U}(1)$ and translations) relevant operators that are not already present. Again the dimension of L, R is $\Delta_L = \Delta_R = \frac{1}{2}$ (Invariance of L_F under $(x,t) \to \lambda(x,t)$ requires $(L,R) \to \lambda^{-\frac{1}{2}}(L,R)$.)

- $\delta\mathcal{L} = \mu L^\dagger L$ (or $R^\dagger R$) is relevant ($\Delta = 1$), but merely shifts the location of the fermi surface (this is $\sum_j \mathbf{Z}_j$)
- $\delta\mathcal{L} = L^\dagger\partial_x L$ is marginal ($\Delta = 2$), but merely renormalizes the Fermi velocity.
- $\delta\mathcal{L} = L^\dagger\partial_x^2 L$ is the first correction that restores the microscopic bandstructure. It and all the others are irrelevant.
- $\delta\mathcal{L} = gL^\dagger R$ has nonzero momentum – it violates translation invariance. This is $\sum_j (-1)^j \mathbf{Z}_j$.
- $\delta\mathcal{L} = \Delta e^{\mathbf{i}\varphi} LR + h.c.$ *explicitly* breaks the $\mathsf{U}(1)$ symmetry and is therefore forbidden– this is a superconducting perturbation. This is $\sum_j e^{\mathbf{i}\varphi}\left(\mathbf{X}_j + \mathbf{i}\mathbf{Y}_y\right) + h.c.$.

- $\delta\mathcal{L} = vL^\dagger LR^\dagger R$ has dimension $\Delta = 2$ and so is marginal. So we'll have to think harder about this one. This comes from $v\sum_j \lambda \mathbf{Z}_j\mathbf{Z}_{j+1}$ in the spin hamiltonian. It turns out that this coupling is *exactly marginal* – *i.e.* there is a line of fixed points parametrized by v. The parameter is called the 'Luttinger liquid parameter' and we will understand it through bosonization.

- $\delta\mathcal{L} = L^\dagger\partial_x R\partial_x L^\dagger R$ has scaling dimension 3 and is therefore irrelevant. Anything with more derivatives or more fermions is even more irrelevant.

Many features of this discussion extend to Fermi surfaces in $d > 1+1$, in particular the fact that the relevant operators merely shift k_F and v_F. This is the origin of Fermi liquid theory. The analysis of the 4-fermion terms is even more interesting in that case.

[End of Lecture 12]

Chiral anomaly

Notice that it seems that we have enhanced the particle-number symmetry to separate conservation of $\mathbf{R}^\dagger\mathbf{R}$ and $\mathbf{L}^\dagger\mathbf{L}$.

But notice, too, that we would be overcounting degrees of freedom if we imagined that $\mathbf{R}, \mathbf{L}$ had all possible momenta: they are connected by the bottom of the band. This seemingly metaphysical consideration has an important consequence: if we apply an electric field, we convert $\mathbf{R}$ excitations into $\mathbf{L}$ excitations. This means that the *axial* charge $\mathbf{Q}_A \equiv \mathbf{R}^\dagger\mathbf{R} - \mathbf{L}^\dagger\mathbf{L}$ is not a conserved quantity. This is called the chiral anomaly.

See my penultimate lecture of Physics 215C for more on this. I reproduce part of the discussion here:

The physics of the chiral anomaly

[Polyakov, page 102; Kaplan 0912.2560 §2.1] Consider non-relativistic free (*i.e.* no 4-fermion interactions) fermions in 1+1 dimensions, *e.g.* with 1-particle dispersion $\omega_k = \frac{1}{2m}\vec{k}^2$. This is a convenient description of our lattice model if $\mu \ll t$, so we only care about the bottom of the band, but it is not crucial. The groundstate of N such fermions is described by filling the N lowest-energy single particle levels, up the Fermi momentum: $|k| \leq k_F$ are filled. We keep our infrared regulator so that the levels are discrete – put them in a box of length L, so that $k_n = \frac{2\pi n}{L}$. (In Figure 1, the red circles are possible 1-particle states, and the green ones are the occupied ones.) The lowest-energy excitations of this groundstate come from taking a fermion just below the Fermi level $|k_1| \lesssim k_F$ and putting it just above $|k_2| \gtrsim k_F$; the energy cost is

$$E_{k_1-k_2} = \frac{1}{2m}\left(k_F + k_1\right)^2 - \frac{1}{2m}\left(k_F - k_2\right)^2 \simeq \frac{k_F}{m}\left(k_1 - k_2\right)$$

– we get relativistic dispersion with velocity $v_F = \frac{k_F}{m}$. The fields near these Fermi points in k-space satisfy the Dirac equation

$$\left(\omega - \delta k\right)\psi_L = 0, \quad \left(\omega + \delta k\right)\psi_R = 0.$$

It would therefore seem to imply a conserved axial current – the number of left moving fermions minus the number of right moving fermions. But the fields ψ_L and ψ_R are not independent; with high-enough energy excitations, you reach the bottom of the band (near $k = 0$ here) and you can't tell the difference. This means that the numbers are *not* separately conserved.

We can do better in this 1+1d example and show that the amount by which the axial current is violated is given by the anomaly formula. Consider subjecting our poor 1+1d free fermions to an electric field $E_x(t)$ which is constant in space and slowly varies in time. Suppose we gradually turn it on and then turn it off; here gradually means slowly enough that the process is adiabatic. Then each particle experiences a force $\partial_t p = eE_x$ and its net change in momentum is

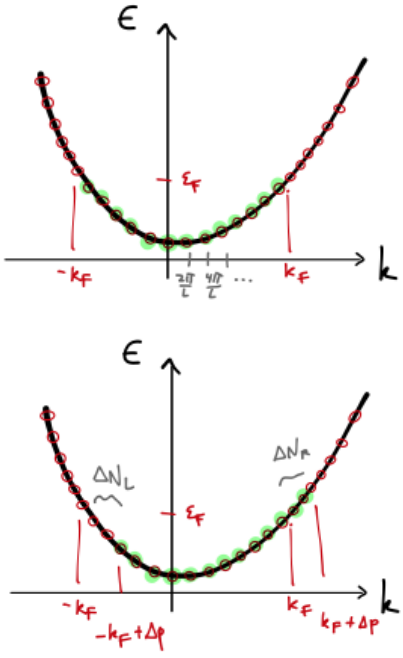


Figure 1: Green dots represent occupied 1-particle states. Top: In the groundstate. Bottom: After applying $E_x(t)$.

$$\Delta p = e \int \mathrm{d}t E_x(t).$$

This means that the electric field puts the fermions in a state where the Fermi surface $k = k_F$ has shifted to the right by Δp, as in the figure. Notice that the total number of fermions is of course the same – charge is conserved.

Now consider the point of view of the low-energy theory at the Fermi points. This theory has the action

$$S[\psi] = \int \mathrm{d}x\mathrm{d}t\bar{\psi}\left(\mathbf{i}\gamma^\mu\partial_\mu\right)\psi \ , \tag{67}$$

where γ^μ are 2×2 and the upper/lower component of ψ creates fermions near the left/right Fermi point. In the process above, we have added ΔN_R right-moving particles and taken away ΔN_L left-moving particles, that is *added* ΔN_L left-moving holes (aka anti-particles). The axial charge of the state has changed by

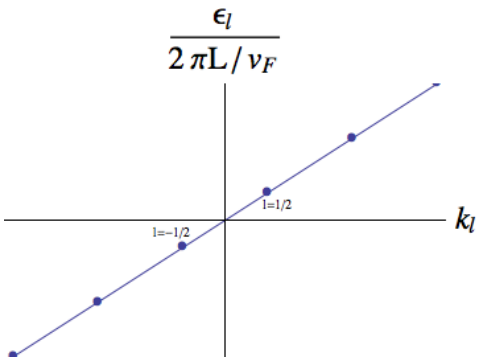
$$\Delta Q_A = \Delta(N_L - N_R) = 2\frac{\Delta p}{2\pi/L} = \frac{L}{\pi}\Delta p = \frac{L}{\pi}e\int \mathrm{d}t E_x(t) = \frac{e}{\pi}\int \mathrm{d}t\mathrm{d}x E_x = \frac{e}{2\pi}\int \epsilon_{\mu\nu}F^{\mu\nu}$$

On the other hand, the LHS is $\Delta Q_A = \int \partial^\mu J^A_\mu$. We can infer a local version of this equation by letting E vary slowly in space as well, and we conclude that

$$\partial_\mu J^\mu_A = \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu}.$$

This agrees exactly with the anomaly equation in $D = 1+1$ produced by the continuum QFT calculation of the anomaly using (67).

3.0.5 Bosonization, part 1: counting



[Sachdev, chapter 20] Let's quantize this system on a circle of length L with APBC: $\mathbf{L}(x+L) = -\mathbf{L}(x), \mathbf{R}(x+L) = -\mathbf{R}(x)$. The mode expansion is:

$$\mathbf{R}(x) = \frac{1}{\sqrt{L}} \sum_{l \in \mathbb{Z}+\frac{1}{2}} \mathbf{R}_l e^{\frac{2\pi l x \mathbf{i}}{L}}$$

and a similar expression for $\mathbf{L}$. We'll focus on $\mathbf{R}$ for a while. The modes satisfy

$$\{\mathbf{R}_l, \mathbf{R}_{l'}^\dagger\} = \delta_{ll'}.$$

and the hamiltonian (minus the vacuum energy) is

$$\mathbf{H}_R = \frac{2\pi v_F}{L} \sum_{l \in \mathbb{Z}+\frac{1}{2}} l \underbrace{\mathbf{R}_l^\dagger \mathbf{R}_l}_{\equiv \mathbf{n}_l^R} - E_0$$

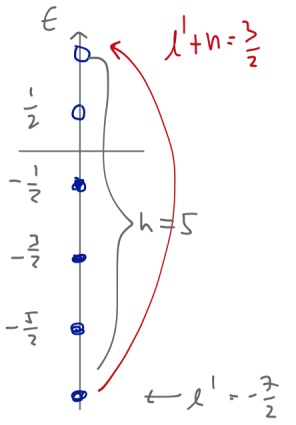
so $\mathbf{H}_R|\text{gs}\rangle = 0$. (The full $\mathbf{H}$ is $\mathbf{H} = \mathbf{H}_R + \mathbf{H}_L$.) The groundstate has all modes with $l > 0$ empty and all modes with $l < 0$ filled. Notice that the Hamiltonian is symmetric under the interchange $l \to -l$: *removing* a particle with negative momentum $-|l|$ (or energy) *adds* momentum $|l|$. The charge operator is

$$\mathbf{Q}_R = \sum_l : \mathbf{R}_l^\dagger \mathbf{R}_l : \in \mathbb{Z}\ , \quad \mathbf{Q}_R|gs\rangle = 0, \quad [\mathbf{Q}_R, \mathbf{H}] = 0.$$

Let's count states. The thermal partition sum for the rightmover is

$$\begin{aligned} Z_R(T) &= \text{tr}_{\mathcal{H}_R} e^{-\mathbf{H}_R/T} \\ &= \prod_{l \in \mathbb{Z}+\frac{1}{2}} \sum_{\mathbf{n}_l^R = 0,1} e^{-\frac{1}{T}\frac{2\pi v_F}{L}|l|\mathbf{n}_l^R} \equiv \prod_{l \in \mathbb{Z}+\frac{1}{2}} \sum_{\mathbf{n}_l^R = 0,1} q^{|l|\mathbf{n}_l^R} \qquad (q \equiv e^{-\frac{2\pi v_F}{TL}}) \\ &= \prod_{l \in \mathbb{Z}+\frac{1}{2}} \left(1 + q^{|l|}\right) \\ &= \prod_{n=1}^{\infty} \left(1 + q^{n-\frac{1}{2}}\right)^2 \end{aligned} \tag{68}$$

This is an infinite-product representation of an elliptic theta function.



Another way to arrive at this function is to keep track of the particle-hole excitations. If we take a particle out of the Fermi sea from a filled level l' and put it in an empty level $l'+n$, we create an excitation with momentum $l'+n-(l') = n$. n is an integer. The change in energy is

$$\Delta\epsilon = v_F \Delta k = \frac{2\pi v_F}{L} n.$$

Notice that it is independent of l'; this is because of the linear dispersion near the Fermi surface. The two fermions (the particle and the hole) move at the same speed, and in the same direction (because we are in 1d and they are both right-moving!) so it might be a good idea to think of them as a single particle. The operator which makes this happen is $\mathbf{R}^\dagger_{l'+n}\mathbf{R}_{l'}$. It is a bosonic operator. More precisely, the object $\boldsymbol{\rho}^\dagger_n \equiv \sum_{l'} \mathbf{R}^\dagger_{l'+n}\mathbf{R}_{l'}$ is a bosonic operator which accomplishes the above operation and can act many times without giving zero. It raises the momentum by n (in units of $\frac{2\pi}{L}$) and the energy by $\frac{2\pi n}{L} v_F$

An arbitrary fermion state is

$$|F\rangle = \prod_{l>0} \mathbf{R}^\dagger_l \prod_{l'<0} \mathbf{R}_{l'} |gs\rangle.$$

If it has charge Q_R, we can rearrange this as

$$|F\rangle = (\text{particle-hole excitations})\, |Q_R\rangle$$

where $|Q_R\rangle$ is the lowest-energy state with charge Q_R. This state has only ls of the same sign, and only the first Q_R of them. For $Q_R > 0$ it is

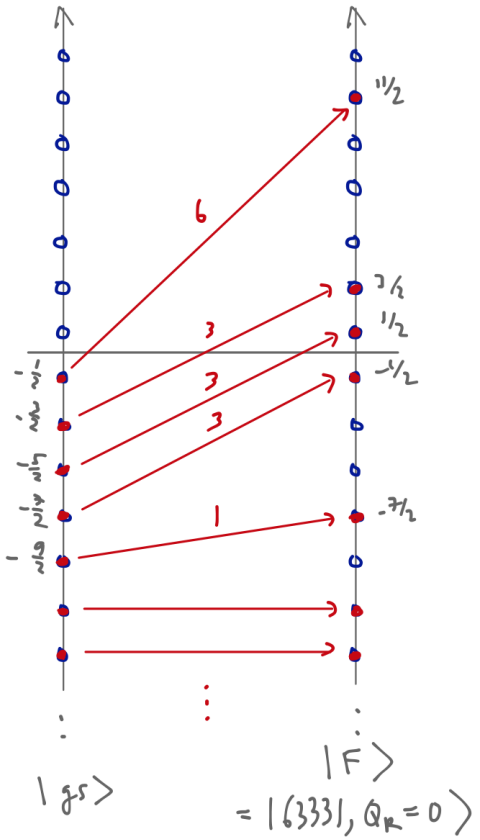
$$|Q_R\rangle = \mathbf{R}^\dagger_{Q_R+\frac{1}{2}} \cdots \mathbf{R}^\dagger_{\frac{3}{2}} \mathbf{R}^\dagger_{\frac{1}{2}} |\text{gs}\rangle.$$

(For $Q_R < 0$, remove the daggers and reverse the signs.) It has energy (for either sign of Q_R)

$$E_0(Q_R) = \frac{2\pi v_F}{L} \sum_{l=\frac{1}{2}}^{|Q_R|} l = \frac{\pi v_F}{L} Q_R^2 \ .$$

For simplicity, let's consider particle-hole excitations above the groundstate, $|\text{gs}\rangle = |Q_R = 0\rangle$, as in the figure at right. (To understand what happens for $Q_R \neq 0$, just shift all the labels by Q_R.) The state in the figure is

$$|63331\rangle = \left(\mathbf{R}^\dagger_{-\frac{7}{2}}\mathbf{R}_{-\frac{9}{2}}\right)\left(\mathbf{R}^\dagger_{-\frac{1}{2}}\mathbf{R}_{-\frac{7}{2}}\right)\left(\mathbf{R}^\dagger_{\frac{1}{2}}\mathbf{R}_{-\frac{5}{2}}\right)\left(\mathbf{R}^\dagger_{\frac{3}{2}}\mathbf{R}_{-\frac{3}{2}}\right)\left(\mathbf{R}^\dagger_{\frac{11}{2}}\mathbf{R}_{-\frac{1}{2}}\right)|\text{gs}\rangle$$



It is important that we act first (first means its at the right of the queue) with the boson operator making the largest jump – this guarantees that we never try to annihilate a fermion that is not there!

So we may label the excitation above $|Q_R\rangle$ by a set of numbers r_n indicating how many steps up the fermion in box $n = -l - \frac{1}{2}$ was pushed:

$$r_1 \geq r_2 \geq r_3 ... \geq r_{n_\star} = 0.$$

The corresponding state is:

$$\boldsymbol{\rho}^\dagger_{r_{n_\star}} \cdots \boldsymbol{\rho}^\dagger_{r_2} \boldsymbol{\rho}^\dagger_{r_1} |Q_R\rangle$$

This information can be represented by a *Young diagram*: For example, we represent the state in the figure, which has $\{r_n\} = \{6, 3, 3, 3, 1\}$, by:

$$\left| \text{  }, Q_R = 0 \right\rangle = \boldsymbol{\rho}^\dagger_1 \boldsymbol{\rho}^\dagger_3 \boldsymbol{\rho}^\dagger_3 \boldsymbol{\rho}^\dagger_3 \boldsymbol{\rho}^\dagger_6 |Q_R = 0\rangle$$

Notice that some of the r_n may be equal; the state still makes sense because these operators are bosons.

We can see how the Young diagram encodes the state in terms of the fermions, if we do anticommutator algebra to cancel $\mathbf{R}^\dagger_l$ and $\mathbf{R}_l$ whenever they both appear:

$$\left| \text{  } \right\rangle =$$

$$\mathbf{R}^\dagger_{-\frac{7}{2}} \mathbf{R}_{-\frac{9}{2}} \cdot \mathbf{R}^\dagger_{-\frac{1}{2}} \mathbf{R}_{-\frac{7}{2}} \cdot \mathbf{R}^\dagger_{\frac{1}{2}} \mathbf{R}_{-\frac{5}{2}} \cdot \mathbf{R}^\dagger_{\frac{3}{2}} \mathbf{R}_{-\frac{3}{2}} \cdot \mathbf{R}^\dagger_{\frac{11}{2}} \mathbf{R}_{-\frac{1}{2}} |\text{gs}\rangle$$

$$= \left(\mathbf{R}^\dagger_{\frac{11}{2}} \mathbf{R}_{\frac{3}{2}}{}^\dagger \mathbf{R}_{\frac{1}{2}}{}^\dagger \right) \mathbf{R}_{-\frac{3}{2}} \mathbf{R}_{-\frac{5}{2}} \mathbf{R}_{-\frac{9}{2}} |\text{gs}\rangle$$

Look at the momentum indices on the annihilation and creation operators. You can see these very numbers are the $(\pm)$ the numbers of boxes to the right and below the diagonal of the diagram respectively.

What is the hamiltonian in terms these boson operators?

$$\mathbf{H}^B = \frac{\pi v_F}{L} \mathbf{Q}^2_R + \frac{2\pi v_F}{L} \sum_{n=l+\frac{1}{2}=\frac{1}{2}}^{\infty} n \boldsymbol{\rho}^\dagger_n \boldsymbol{\rho}_n .$$

This is a *chiral* boson, since $n > 0$. Its partition sum is

$$\begin{aligned} Z^R_{\text{boson}}(T) &= \text{tr}_{\mathcal{H}_B} e^{-\mathbf{H}^B/T} \\ &= \left(\sum_{Q_R=-\infty}^{\infty} q^{Q^2_R/2} \right) \prod_{n=1}^{\infty} \underbrace{\sum_{m_n = \boldsymbol{\rho}^\dagger_n \boldsymbol{\rho}_n = 0,1,..\infty} q^{n m_n}}_{\frac{1}{1-q^n}} \end{aligned} \tag{69}$$

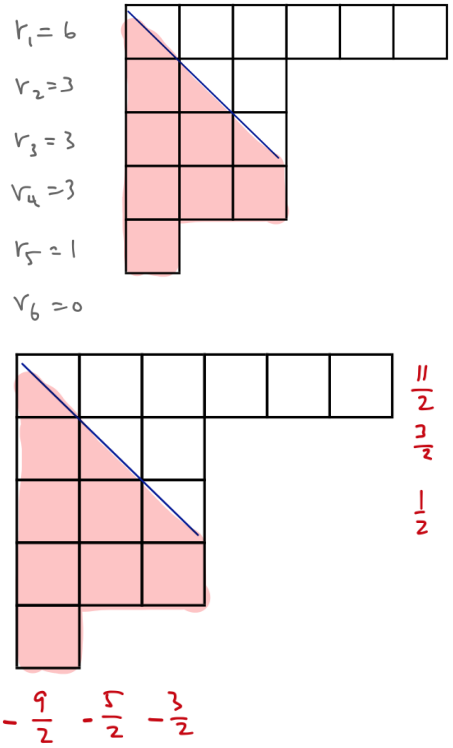


Figure 2: The two ways of looking at the Young diagram which determine the boson state and the fermion state, respectively.

This is an infinite-sum representation of the same elliptic theta function. (For more about the theta functions, if you must, see *e.g.* Polchinski vol. 1, page 215.)

3.1 Free boson CFT in 1+1d, aka Luttinger liquid theory

Now we take an apparent detour. One way to motivate what we'll do is to imagine that our system tries to produce by spontaneous symmetry breaking a Goldstone boson for our $\mathsf{U}(1)$ symmetry. We will study the physics of the resulting scalar field – whose field space is the broken group $\mathsf{U}(1)$, *i.e.* the circle – and see that the symmetry is in fact restored. We'll learn a lot of other things, too.

Consider the following (real-time) continuum action for a free massless scalar field in 1+1 dimensions:

$$S[\phi] = \frac{T}{2}\int dt \int_0^L dx \left((\partial_0\phi)^2 - (\partial_x\phi)^2\right) = 2T\int dxdt\partial_+\phi\partial_-\phi \ . \tag{70}$$

I have set the velocity of the bosons to $c = 1$ by rescaling t. Here $x^\pm \equiv t \pm x$ are lightcone coordinates; the derivatives are $\partial_\pm \equiv \frac{1}{2}(\partial_t \pm \partial_x)$. Space is a circle: the point labelled x is the same as the point labelled $x + L$.

We will assume that the field space of ϕ itself is periodic:

$$\phi(x,t) \equiv \phi(x,t) + 2\pi, \quad \forall x,t \ .$$

So the *field* space is a circle S^1 with (angular) coordinate ϕ. It can be useful to think of the action (70) as describing the propagation of a string, since a field configuration describes ana embedding of the real two dimensional space into the *target space*, which here is a circle. This is easier to picture if we consider more than one (say N) scalar fields, ϕ^a, which we can take to be coordinates on an N-dimensional space with metric $g_{ab}(\phi)$. Then we can consider the more general action

$$S[\phi^a] = T\int dxdt\partial_+\phi^a\partial_-\phi^b g_{ab}(\phi(x,t)) \ .$$

This is called a *non-linear sigma model*. Note that 'T' is for 'tension'[28].

[End of Lecture 13]

Let's retreat now to the case where the target space is a circle. To get back the action (70), we should set $g_{\phi\phi} = 2$. Notice that by changing coordinates on the field space, *e.g.* $\tilde{\phi} \equiv R\phi$ with α independent of x, t, we change the prefactor in the action:

$$2T\partial_+\phi\partial_-\phi = \frac{2T}{R^2}\partial_+\tilde{\phi}\partial_-\phi.$$

[28] As you might expect, there are many conventions in the literature for this quantity. The string theory literature is commonly written in terms of $\alpha' \equiv \frac{1}{2\pi T}$. The literature on Luttinger liquids (*e.g.* in Sachdev's book) uses $K \equiv 2\alpha' = \frac{1}{\pi T}$.

For example, by setting $R^2 = \sqrt{2T}$ it looks like we can make it disappear. But since ϕ is compact, it returns in the periodicity:

$$\tilde{\phi} \equiv \tilde{\phi} + 2\pi R.$$

So R is like the radius of the circle. This leads to a very annoying multiplicity of conventions. Notice that if ϕ is not compact, the path integral gets a very awkward infinite part from $\int_{-\infty}^{\infty} d\phi_0$ (with ϕ_0 constant). But these are just the same issues that arise in the path integral of a particle on the real line; this integral will produce delta functions which enforce conservation of momentum in the ϕ-space.

First a little bit of classical field theory. The equations of motion for ϕ are

$$0 = \frac{\delta S}{\delta \phi(x,t)} \propto \partial^\mu \partial_\mu \phi \propto \partial_+ \partial_- \phi$$

which is solved by

$$\phi(x,t) \equiv \phi_L(x^+) + \phi_R(x^-) \ .$$

In euclidean time, $\phi_{L,R}$ depend (anti-)holomorphically on the complex coordinate $z \equiv \frac{1}{2}(x + \mathbf{i}\tau)$ and the machinery of complex analysis becomes useful.

Symmetries: Since $S[\phi]$ only depends on ϕ through its derivatives, there is a simple symmetry $\phi \to \phi + \epsilon$. By the Nöther method[29] the associated current is

$$j_\mu = T \partial_\mu \phi \ . \tag{71}$$

This symmetry is translations in the target space, and I will sometimes call the associated conserved charge 'momentum'.

There is another symmetry which is less obvious. It comes about because of the topology of the target space. Since $\phi(x,t) \equiv \phi(x,t) + 2\pi m, m \in \mathbb{Z}$ describe the same point (it is a redundancy in our description, in fact a discrete gauge redundancy), we don't need $\phi(x+L,t) = \phi(x,t)$. It is enough to have

$$\phi(x+L,t) = \phi(x,t) + 2\pi m, \quad m \in \mathbb{Z}$$

The number m cannot change without the string breaking: it is a topological charge, a winding number:

$$m = \frac{1}{2\pi} \phi(x,t)|_{x=0}^{x=L} \overset{\text{FTC}}{=} \frac{1}{2\pi} \int_0^L dx \partial_x \phi \ . \tag{72}$$

[29] In case you have forgotten, here's the method: if the transformation with constant ϵ is a symmetry, then the variation of the action with $\epsilon = \epsilon(x,t)$ must be proportional to $\partial_\mu \epsilon$, and so that it vanishes $\forall \phi$ when ϵ is constant:

$$S[\phi + \epsilon(x)] - S[\phi] = \int dx dt \partial_\mu \epsilon(x) j^\mu \overset{\text{IBP}}{=} - \int \epsilon(x) \partial_\mu j^\mu \ .$$

But if the equations of motion are obeyed, then the action is invariant under *any* variation, including this one, for arbitrary $\epsilon(x)$. But this means that $\partial_\mu j^\mu = 0$, the current is conserved.

The associated current whose charge density is $\frac{1}{\pi}\partial_x\phi$ (which integrates over space to the topological charge) is

$$\tilde{j}_\mu = \frac{1}{2\pi}(\partial_x\phi, -\partial_0\phi)_\mu = \frac{1}{2\pi}\epsilon^{\mu\nu}\partial_\nu\phi.$$

This is conserved because of the equality of the mixed partials: $\epsilon^{\mu\nu}\partial_\mu\partial_\nu = 0$.

Let's expand in normal modes: $\phi = \phi_L + \phi_R$ with

$$\begin{aligned}
\phi_L(t+x) &= q_L + \underbrace{(p+w)}_{\equiv\frac{1}{2T}p_L}(t+x) - \mathbf{i}\sqrt{\frac{L}{4\pi T}}\sum_{n\neq 0}\frac{\rho_n}{n}e^{\mathbf{i}n(t+x)\frac{2\pi}{L}}, \\
\phi_R(t-x) &= q_R + \underbrace{(p-w)}_{\equiv\frac{1}{2T}p_R}(t-x) - \mathbf{i}\sqrt{\frac{L}{4\pi T}}\sum_{n\neq 0}\frac{\tilde{\rho}_n}{n}e^{\mathbf{i}n(t-x)\frac{2\pi}{L}},
\end{aligned} \tag{73}$$

The factor of $\frac{1}{n}$ is a convention whose origin you will appreciate below, as are the other normalization factors. Real ϕ means $\rho_n^\dagger = \rho_{-n}$ (If we didn't put the $\mathbf{i}$ it would have been $-\rho_{-n}$).

Here $q \equiv \frac{1}{L}\int_0^L dx\phi(x,t)$ is the center-of-mass position of the string. The canonical momentum for ϕ is $\pi(x,t) = T\partial_0\phi(x,t) = T\left(\partial_+\phi_L + \partial_-\phi_R\right)$.

QM. Now we'll do quantum mechanics. Recall that a quantum mechanical particle on a circle has momentum quantized in units of integers over the period. Since ϕ is periodic, the wavefunction(al)s must be periodic in the center-of-mass coordinate q with period 2π, and this means that the total (target-space) momentum must be an integer

$$\mathbb{Z} \ni j = \pi_0 \equiv \int_0^L dx\pi(x,t) = T\int_0^L dx\partial_t\phi \overset{(73)}{=} LT2p$$

So our conserved charges are quantized according to

$$p = \frac{j}{2LT}, \qquad w \overset{(73)(72)}{=} \frac{\pi m}{L}, \qquad j, m \in \mathbb{Z}\ .$$

(Don't confuse the target-space momentum j with the 'worldsheet momentum' n!)

(Note that this theory is scale-free. We could use this freedom to choose units where $L = 2\pi$.)

Now I put the mode coefficients in boldface:

$$\begin{aligned}
\boldsymbol{\phi}_L(x^+) &= \mathbf{q}_L + \frac{1}{2T}\mathbf{p}_L x^+ - \mathbf{i}\sqrt{\frac{L}{4\pi T}}\sum_{n\neq 0}\frac{\boldsymbol{\rho}_n}{n}e^{\mathbf{i}\frac{2\pi}{L}nx^+}, \\
\boldsymbol{\phi}_R(x^-) &= \mathbf{q}_R + \frac{1}{2T}\mathbf{p}_R x^- - \mathbf{i}\sqrt{\frac{L}{4\pi T}}\sum_{n\neq 0}\frac{\tilde{\boldsymbol{\rho}}_n}{n}e^{\mathbf{i}\frac{2\pi}{L}nx^-},
\end{aligned} \tag{74}$$

The nonzero canonical equal-time commutators are

$$[\boldsymbol{\phi}(x), \boldsymbol{\pi}(x')] = -\mathbf{i}\delta(x - x')$$

which determines the commutators of the modes (this was the motivation for the weird normalizations)

$$[\mathbf{q}_L, \mathbf{p}_L] = [\mathbf{q}_R, \mathbf{p}_R] = \mathbf{i}, \qquad [\boldsymbol{\rho}_n, \boldsymbol{\rho}_{n'}^\dagger] = n\delta_{n,n'}, \text{or} \quad [\boldsymbol{\rho}_n, \boldsymbol{\rho}_{n'}] = n\delta_{n+n'} \ ,$$

and the same for the rightmovers with twiddles. This is one simple harmonic oscillator for each $n \geq 1$ (and each chirality); the funny normalization is conventional.

$$\begin{aligned}
\mathbf{H} &= \int dx \left(\boldsymbol{\pi}(x)\dot{\boldsymbol{\phi}}(x) - \mathcal{L}\right) = \frac{1}{2}\int dx \left(\frac{\boldsymbol{\pi}^2}{T} + T\left(\partial_x \boldsymbol{\phi}\right)^2\right) \\
&= \underbrace{L\frac{1}{4T}\left(\mathbf{p}_L^2 + \mathbf{p}_R^2\right)}_{\frac{\boldsymbol{\pi}_0^2}{2T} + \frac{T}{2}\mathbf{w}^2} + \pi\sum_{n=1}^{\infty}\left(\boldsymbol{\rho}_{-n}\boldsymbol{\rho}_n + \tilde{\boldsymbol{\rho}}_{-n}\tilde{\boldsymbol{\rho}}_n\right) + \mathfrak{a} \\
&= \frac{1}{2L}\left(\frac{j^2}{T} + T\left(2\pi m\right)^2\right) + \pi\sum_{n=1}^{\infty} n\left(\mathbf{N}_n + \tilde{\mathbf{N}}_n\right) + \mathfrak{a}
\end{aligned} \tag{75}$$

Here $\mathfrak{a}$ is a (UV sensitive) constant which will not be important for us (it is very important in string theory), which is the price we pay for writing the hamiltonian as a sum of normal-ordered terms – the modes with negative indices are to the right and they annihilate the vacuum:

$$\boldsymbol{\rho}_n|0\rangle = 0, \quad \tilde{\boldsymbol{\rho}}_n|0\rangle = 0, \quad \text{for } n > 0 \ .$$

Energy eigenstates can be labelled by a target-momentum j and a winding m. Notice that there is an operator $\mathbf{w}$ whose eigenvalues are w. So when I write $|0\rangle$ above, I really should label a vacuum of the oscillator modes with p, w.

$\mathbf{N}_n \equiv \frac{1}{n}\boldsymbol{\rho}_{-n}\boldsymbol{\rho}_n$ is the number operator; if we redefine $\mathbf{a}_n \equiv \sqrt{n}^{-1}\boldsymbol{\rho}_n (n > 0)$, we have $[\mathbf{a}_n, \mathbf{a}_m^\dagger] = \delta_{nm}$ and $\mathbf{N}_n = \mathbf{a}_n^\dagger \mathbf{a}_n$ is the ordinary thing.

Notice that (71) means that there are separately-conserved left-moving and right-moving currents:

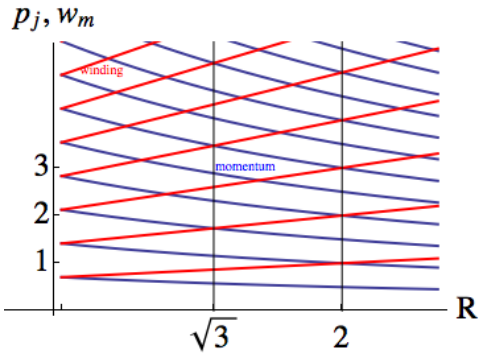
$$\begin{aligned}
(j_L)^\mu &= (j_L^z, j_L^{\bar{z}})^\mu \equiv (j_+, 0)^\mu \\
(j_R)^\mu &= (j_R^z, j_R^{\bar{z}})^\mu \equiv (0, j_-)^\mu
\end{aligned}$$

Here j_L only depends on the modes $\boldsymbol{\rho}_n$, and j_R only depends on the modes $\tilde{\boldsymbol{\rho}}_n$:

$$j_+ = \partial_+\phi = \partial_+\phi(x^+) = \mathbf{p} + \mathbf{w} + \sqrt{\frac{\pi}{LT}}\sum_{n\neq 0}\boldsymbol{\rho}_n e^{\mathbf{i}\frac{2\pi}{L}nx^+}$$

$$j_- = \partial_-\phi = \partial_-\phi(x^-) = \mathbf{p} - \mathbf{w} + \sqrt{\frac{\pi}{LT}}\sum_{n\neq 0}\tilde{\boldsymbol{\rho}}_n e^{\mathbf{i}\frac{2\pi}{L}nx^-}$$

Here's an **Observation (T-duality)**: At large T (think of this as a large radius of the target space), the momentum modes are closely-spaced in energy, and exciting the winding modes is costly, since the string has a tension, it costs energy-per-unit-length T to stretch it. But the spectrum (75) is invariant under the operation



$$m \leftrightarrow j, \quad T \leftrightarrow \frac{1}{(2\pi)^2 T}$$

which takes the radius of the circle to its inverse and exchanges the momentum and winding modes. This is called *T-duality*. The required duality map on the fields is

$$\phi_L + \phi_R \leftrightarrow \phi_L - \phi_R \ .$$

(The variable R in the plot is $R \equiv \sqrt{\pi T}$.)

You probably won't be surprised to learn that the value of T where the two spectra coincide – the self T-dual radius ($R = \sqrt{2}$) – is special. In fact there is an enhanced $\mathsf{SU}(2) \times \mathsf{SU}(2)$ symmetry at that point. We will understand this soon and there is a homework problem on it. There are other special points, such as $R = 2$ where a winding mode w_1 and a momentum mode p_2 become degenerate. This is the free-fermion radius. The next collision (between w_2 and p_3) occurs at $R = \sqrt{3}$ which is in fact a model with supersymmetry.

It will sometimes be useful to consider the euclidean-space formulation: $it \to \tau$. Then a function of the lightcone coordinates $(t \pm x)$ becomes a function of the complex coordinates

$$w = (x + \mathbf{i}\tau) \, , \quad \bar{w} = (x - \mathbf{i}\tau)$$

and the free scalar action is

$$S = T \int d^2w \ \partial_w \phi \bar{\partial}_w \phi .$$

(Note that $d^2w = 2dxd\tau$ and $\partial_w = \frac{1}{2}(\partial_x - \mathbf{i}\partial_\tau)$.) Note also that (following Polchinski's conventions) holomorphic = leftmoving.

Let's think about who are the good operators in this model. The first thing we might try is ϕ itself. Since the path integral is gaussian, we have

$$\langle \phi(z,\bar{z})\phi(0,0)\rangle = -\frac{1}{\pi T}\log\frac{z\bar{z}}{a^2} = -\frac{1}{\pi T}\log\frac{z}{a} - \frac{1}{\pi T}\log\frac{\bar{z}}{a} \tag{76}$$

where a is a UV cutoff.

Three ways to get this answer (not really independent): (1) Use the equations of motion (*i.e.* the Ward identity). What I mean is this: the path integral is independent of choice of

integration variable. Therefore:

$$0 = \int [d\phi] \frac{\partial}{\partial \phi(z,\bar{z})} \text{(anything)} \, .$$

In particular,

$$0 = \int [d\phi] \frac{\partial}{\partial \phi(z,\bar{z})} \left(e^{-S} \phi(0) \right) = \delta^2(z,\bar{z}) + 2T \partial\bar{\partial} \langle \phi(z,\bar{z}) \phi(0) \rangle \ . \tag{77}$$

Two comments: (a) in the Hamiltonian formalism, the delta function comes about when the derivatives hit the time-ordering theta-function. (b) The eom is in fact an *operator equation*, meaning that it ($\partial\bar{\partial}\phi(z) = 0$) is a true equation as long as there are no other operators nearby. In (77) we have one other operator nearby, which produces one contact term. We'll come back to this point.

We conclude that

$$\partial_z \bar{\partial}_z \langle \phi(z,\bar{z}) \phi(0,0) \rangle = -\frac{2}{T} \delta^2(z,\bar{z}) \ .$$

This equation is solved by the math fact[30]:

$$\partial\bar{\partial} \log |z|^2 = 2\pi \delta^2(z,\bar{z}) \ . \tag{78}$$

(Clearly it's true for $z \neq 0$. To see the delta function, integrate the both hand side.)

(2) Fourier: use the momentum space propagator (beware my factors here):

$$\langle \phi(z,\bar{z}) \phi(0,0) \rangle \propto - \int \mathrm{đ}^2 p \frac{e^{\mathbf{i} p \cdot x}}{p^2}$$

This is just the Fourier space solution of the PDE (78).

(3) We can also arrive at this answer using operator algebra by decomposing $\boldsymbol{\phi} = \boldsymbol{\phi}_> + \boldsymbol{\phi}_<$ into creation pieces and annihilation pieces.

Notice the virtue that this is a sum of holomorphic and antiholomorphic (actually meromorphic) pieces. Also notice that the cutoff dependence is just an additive constant. But this is a *weird* propagator: it *grows* in magnitude as the points get farther apart. Notice also that we got a log because a massless scalar in 1+1d has scaling dimension zero. This means that ϕ is not such a great operator – its effects are not local.

The Hohenberg-Mermin-Wagner-Coleman Fact. This (logarithmic) growth of the correlations of a massless scalar field is what prevents true spontaneous breakdown of continuous symmetries in two dimensions and below (notice from the fourier representation that the growth is only faster in $D < 2$).

[30] In case you are worried about the factors of two: I am using Polchinski's convention that $d^2z = 2dxdt$, so $\int d^2z \delta^2(z,\bar{z}) = 1$ means $\delta^2(z,\bar{z}) = \frac{1}{2}\delta(x)\delta(t)$.

If you look at Coleman's paper linked above it says some nonsense about 'there is no massless scalar field theory in $D = 2$'; this is false. The thing there isn't is Goldstone bosons, as his title correctly says. The field theory is fine, it's just that the ϕ operator itself is bad. And the fact that the fluctuations of ϕ (the supposed order parameter) are IR divergent means that there is no long-range order.

However, there are other operators which are good. To begin, the derivatives of ϕ are good operators. For example, the currents have

$$\langle j_+(z) j_+(0)\rangle = \frac{k}{z^2}.$$

The power here says that j_+ has scaling dimension 1. Actually, we can say something stronger: if we scale z and $\bar{z}$ independently, it has scaling dimension $(1,0)$.

$$\langle j_-(\bar{z}) j_-(0)\rangle = \frac{k}{\bar{z}^2}$$

has scaling dimension $(0,1)$.

(A current in a d-dimensional critical theory always has scaling dimension d: you have to integrate j_0 over space and get a pure number, the charge.)

Vertex operators. Since ϕ has logarithmic correlators, you might think that exponentiating it is a good idea. First let's take advantage of the fact that the correlator (76) splits into left and right to write $\phi(z,\bar{z}) = \phi_L(z) + \phi_R(\bar{z})$:

$$\langle \phi_L(z)\phi_L(0)\rangle = -\frac{1}{\pi T}\log\frac{z}{a}, \quad \langle \phi_R(\bar{z})\phi_R(0)\rangle = -\frac{1}{\pi T}\log\frac{\bar{z}}{a}\ , \quad \langle \phi_L(z)\phi_R(0)\rangle = 0\ . \tag{79}$$

Another set of good operators with definite scaling dimension arises from:

$$\mathcal{V}_{\alpha,\beta}(z,\bar{z}) =: e^{\mathbf{i}(\alpha\phi_L(z)+\beta\phi_R(\bar{z}))} :\ .$$

This is a composite operator which we have defined by normal-ordering. The normal ordering prescription is: $\boxed{q,p,-,+}$, that is: positive-momentum modes (lowering operators) go on the right, and p counts as a lowering operator, so in particular using the expansion (please beware my factors here): $\phi_L(z) = \mathbf{q}_L + \mathbf{p}_L z + \mathbf{i}\sum_{n\neq 0}\frac{\boldsymbol{\rho}_n}{n}w^n$, we have

$$: e^{\mathbf{i}\alpha\phi_L(z)} :\equiv e^{\mathbf{i}\alpha\mathbf{q}_L} e^{\mathbf{i}\alpha\mathbf{p}_L z} e^{\mathbf{i}\alpha\sum_{n<0}\frac{\boldsymbol{\rho}_n}{n}w^n} e^{\mathbf{i}\alpha\sum_{n>0}\frac{\boldsymbol{\rho}_n}{n}w^n}$$

(I used the definition $w \equiv e^{2\pi i z/L}$.)

How should we think about this operator ? In the QM of a free particle, the operator $e^{\mathbf{i}p\mathbf{x}}$ inserts momentum p – it takes a momentum-space wavefunction $\psi(p_0) = \langle p_0|\psi\rangle$ and gives

$$\langle p_0|e^{\mathbf{i}p\mathbf{x}}|\psi\rangle = \psi(p_0+p).$$

It's the same thing here, with one more twist.

In order for $\mathcal{V}_{\alpha,\beta}$ to be well-defined under $\phi \to \phi + 2\pi$, we'd better have $\alpha + \beta \in \mathbb{Z}$ – momentum is quantized, just like for the particle (the center of mass is just a particle). Let's consider what the operator $\mathcal{V}_{\alpha,\beta}$ does to a winding and momentum eigenstate $|w,p\rangle$ (with no oscillator excitations, $\boldsymbol{\rho}_n|p,w\rangle = 0, n < 0$):

$$\mathcal{V}_{\alpha\beta}(0)|w,p\rangle = e^{\mathbf{i}(\alpha+\beta)\mathbf{q}_0} e^{\mathbf{i}(\alpha-\beta)\tilde{\boldsymbol{\phi}}_0} e^{\mathbf{i}\alpha\sum_{n<0}\boldsymbol{\rho}_n} e^{\mathbf{i}\alpha\sum_{n>0}\boldsymbol{\rho}_n}|w,p\rangle = e^{\mathbf{i}\alpha\sum_{n<0}\boldsymbol{\rho}_n}|w+\alpha-\beta, p+\alpha+\beta\rangle \tag{80}$$

The monster in front here creates oscillator excitations; the important thing is that the winding number has been incremented by $\alpha - \beta$; this means that $\alpha - \beta$ must be an integer, too. We conclude that

$$\alpha + \beta \in \mathbb{Z}, \quad \alpha - \beta \in \mathbb{Z} \tag{81}$$

so they can both be half-integer, or they can both be integers.

Using the gaussian result from homework one (or moving the annihilation operators to the right) their correlators are

$$\langle \mathcal{V}_{\alpha,\beta}(z,\bar{z})\mathcal{V}_{\alpha',\beta'}(0,0)\rangle = \frac{D_0}{z^{\frac{\alpha^2}{\pi T}}\bar{z}^{\frac{\beta^2}{\pi T}}} \ . \tag{82}$$

The zeromode prefactor D_0 is:

$$D_0 = \langle e^{\mathbf{i}((\alpha+\alpha')\mathbf{q}_L + (\beta+\beta')\mathbf{q}_R)}\rangle_0 = \delta_{\alpha+\alpha'}\delta_{\beta+\beta'}.$$

This is charge conservation.

We conclude that the operator $\mathcal{V}_{\alpha,\beta}$ has scaling dimension

$$(h_L, h_R) = \frac{1}{2\pi T}(\alpha^2, \beta^2).$$

Notice the remarkable fact that the exponential of a dimension-zero operator manages to have nonzero scaling dimension. This requires that the multiplicative prefactor depend on the cutoff a to the appropriate power (and it is therefore nonuniversal). We could perform a multiplicative renormalization of our operators $\mathcal{V}$ to remove this cutoff dependence from the correlators.

The values of α, β allowed by single-valuedness of ϕ and its wavefunctional are integers. We see (at least) three special values of the parameter T:

- **The SU(2) radius**: When $2\pi T = 1$, the operators with $(n,m) = 1$ are *marginal*. Also, the operators with $(n,m) = (1,0)$ and $(n,m) = (0,1)$ have the scaling behavior of currents, and by holomorphicity are in fact conserved.
- **The free fermion radius**: when $2\pi T = 2$, $\mathcal{V}_{1,0}$ has dimension $(\frac{1}{2}, 0)$. which is the behavior of a left-moving free fermion.

- **The supersymmetric radius**: when $2\pi T = \frac{2}{3}$, $\mathcal{V}_{1,0}$ has dimension $(\frac{3}{2}, 0)$ and represents a supersymmetry current.

The expression for the correlator is easily generalized to multiple vertex operators:

$$\langle \prod_{j=1}^{N} \mathcal{V}_{\alpha_j}(z) \rangle = \frac{\delta_{\sum_j \alpha_j, 0}}{\prod_{i<j}^{N} (z_i - z_j)^{\frac{\alpha_i \cdot \alpha_j}{2\pi T}}}$$

Operator product expansion (OPE).

In QFT, if we act with a bunch of local operators close together (compared to their separation from other operators), their net effect is that of another local operator. If we have a *basis* of operators in our CFT, we may expand this net effect in terms of this basis. So far this is true in any QFT. In a scale-invariant theory, we can choose a basis of operators of definite scaling dimension[31],

$$\mathcal{O}_a \to \lambda^{-\Delta_a} \mathcal{O}_a, \qquad \langle \mathcal{O}_a^\dagger(z) \mathcal{O}_a(0) \rangle \sim \frac{1}{z^{2\Delta_a}}.$$

We can organize this expansion (the *operator product expansion*) in order of increasing dimension, $\{\mathcal{O}_a, \Delta(a) \leq \Delta(b) \text{ if } a < b\}$. Then the operators with the longest range of influence come first:

$$\begin{aligned} \mathcal{O}_A(z)\mathcal{O}_B(0) &\overset{z\to 0}{\sim} C_{AB1}(z)\mathcal{O}_1(0) + C_{AB2}(z)\mathcal{O}_2(0) + \cdots \\ &= \frac{c_{AB1}}{z^{\Delta_A+\Delta_B-\Delta_1}}\mathcal{O}_1(0) + \frac{c_{AB2}}{z^{\Delta_A+\Delta_B-\Delta_2}}\mathcal{O}_2(0) + \cdots \end{aligned} \tag{83}$$

In general, there will be power-law singularities in the coefficients, produced by the collisions of the operators; the powers are determined by scaling. But at some point in the expansion, the effects start to vanish as $z \to 0$, *i.e.* $\Delta_A + \Delta_B - \Delta_n < 0$ for Δ_n big enough. The terms in the OPE after this are *regular terms* and can often be ignored, for example because they don't effect correlation functions without more insertions nearby. In the future, we will be writing expressions like

$$\partial\phi(z)\partial\phi(0) \sim \frac{1}{z^2} + \text{regular terms}\ .$$

After this, "plus regular terms" will be implicit.

We saw earlier that free massless fermions in 1+1d have groundstate correlators that go like

$$\langle \mathbf{\Psi}^\dagger(x)\mathbf{\Psi}(0)\rangle \propto -\frac{1}{x}. \tag{84}$$

[31] Some fine print: actually, this requires that the dilatation operator is hermitian. This fails to be the case in some non-unitary theories, for example in *logarithmic CFTs* where the dilatation operator involves Jordan blocks, and where therefore the OPE involves also some logarithms (which shift under scale transformations).

In terms of the OPE, this can be expressed as $\psi_R^\dagger(x,t)\psi_R(0) \sim -\frac{1}{t-x}$ or in euclidean coords

$$\psi_R^\dagger(\bar{z})\psi_R(0) \sim \frac{1}{\bar{z}}.$$

The operators $e^{\mathbf{i}\phi(x)}$ satisfy

$$: e^{\mathbf{i}n\phi_R(z)} :: e^{\mathbf{i}m\phi_R(0)} :=: e^{\mathbf{i}(n+m)\phi_R(0)} : z^{\frac{2nm}{2\pi T}} + ... \tag{85}$$

Note that the power of z agrees with our general OPE formula since $n^2 + m^2 - (n+m)^2 = -2nm$. When we set $n = -m$ and $\pi T = 1$, this looks like the fermion expression (84).

In general the ... in (85) can contain further singular terms coming from the Taylor expansion of $\phi_L(z) = \phi_L(0) + z\partial\phi_L(0) + \cdots$. That is, going more slowly, we have

$$\begin{aligned} : e^{\mathbf{i}n\phi_R(z)} :: e^{\mathbf{i}m\phi_R(0)} : &= : e^{\mathbf{i}n\phi_R(z)+\mathbf{i}m\phi_R(0)} : z^{\frac{2nm}{2\pi T}} \\ &= : e^{\mathbf{i}(n+m)\phi_R(0)} : z^{\frac{2nm}{2\pi T}} + \mathbf{i}n\partial\phi_R(0) : e^{\mathbf{i}(n+m)\phi_R(0)} : z^{\frac{2nm}{2\pi T}+1} + ... \end{aligned} \tag{86}$$

Note that the bosonization identities which we will discuss further below are statements about the singular parts of the OPE of these operators – the finite, less-universal pieces are not part of the conversation.

[End of Lecture 14]

3.1.1 Application 1: Bosonization dictionary

[Fradkin 2d ed, page 133; Sachdev 2d ed, page 420] Let's set $T = \frac{1}{\pi}$ for a bit. The left and right $\mathsf{U}(1)$ currents in terms of bosons and fermions are

$$\Psi_L^\dagger \Psi_L = \frac{1}{2\pi}\partial_z \phi, \quad \Psi_R^\dagger \Psi_R = \frac{1}{2\pi}\bar{\partial}_z \phi.$$

The fermions themselves are:

$$\Psi_L(z)^\dagger =: e^{\mathbf{i}\phi_L(z)} :, \quad \Psi_R(\bar{z})^\dagger =: e^{\mathbf{i}\phi_R(\bar{z})} : \ .$$

How might a human have figured this out? (One human who did was Mandelstam.)

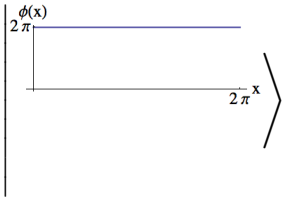
The relation for the charges says that the total fermion charge is the winding number:

$$\Psi_L^\dagger \Psi_L + \Psi_R^\dagger \Psi_R = \frac{1}{2\pi}\partial_x \phi. \tag{87}$$

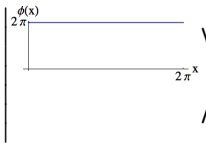
Let's consider the refined statement:

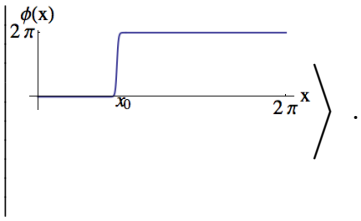
$$\Psi_L^\dagger \Psi_L = \frac{1}{2\pi}\partial_x \phi_L = j_L. \tag{88}$$

This means that if we create a left-moving fermion at $x = x_0$, the value of $\phi_L(x)$ should *jump* for all $x < x_0$, by a fixed amount so that if we integrate the BHS of (87) we get 1. (That amount is 2π.) For example, according to (88) a state with no fermions looks like:

$|\text{no fermions}\rangle =$ 

Acting on this state with one fermion at $x = x_0$ should do the following:

$\Psi_L^\dagger(x_0)$ 

$=$ 

.

Such a jump is produced by the conjugate field momentum (recall the translation operator in QM is $e^{-\mathbf{i}\mathbf{p}a}\psi(x) = \psi(x+a)$):

$$\Psi_L(x_0) \sim e^{\mathbf{i}2\pi \int_{-\infty}^{x_0} dx \boldsymbol{\pi}_L(x)} \ .$$

Now who is the canonical momentum $\boldsymbol{\pi}_L(x)$ in terms of ϕ_L, ϕ_R? It's the thing with a canonical commutator with ϕ_L. One way to figure this out is to go back to the mode expansions and show that

$$[\partial_x \boldsymbol{\phi}_L(x), \boldsymbol{\phi}_L(y)] = \frac{2\mathbf{i}}{T}\delta(x-y) \tag{89}$$

which shows that $\boldsymbol{\pi}_L(x) = \frac{T}{2}\partial_x \boldsymbol{\phi}_L(x)$ (and similarly $\boldsymbol{\pi}_R = -\frac{T}{2}\partial_x \boldsymbol{\phi}_R(x)$). (Notice that this equation is roughly the imaginary part of the x derivative of (79).) So (when $T = \frac{1}{\pi}$, which is when this operator is single-valued!) the fermion operator is

$$\Psi_L(x_0) \sim e^{\mathbf{i}\int_{-\infty}^{x_0} dx\ \partial_x \boldsymbol{\phi}_L} = e^{\mathbf{i}\boldsymbol{\phi}_L}\ .$$

We already saw that this kind of object creates winding in (80). So again the fermion is a domain-wall creation operator; notice that because the field ϕ is periodic, it can be possible to have a single such domain wall on a circle. We'll say a bit more about the periodicity of ϕ_L below.

In order for this to be a fermion operator, we require $\{\Psi_L(x), \Psi_L(x')\} = 0, x \neq x'$. Is it true? Yes in fact, as we can see by BCH, in the form $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{B}}e^{\mathbf{A}}e^{-[\mathbf{A},\mathbf{B}]}$ if $[\mathbf{A}, \mathbf{B}]$ is a c-number. A useful intermediate fact comes from the x-integral of (89) (from $-\infty$ to x) which says

$$[\boldsymbol{\phi}_L(x), \boldsymbol{\phi}_L(y)] = \mathbf{i}\pi \mathsf{sign}(x-y)$$

(the additive constant of integration on the RHS is fixed by demanding antisymmetry in $x \to y$ which is manifest on the LHS). Therefore:

$$:e^{\mathbf{i}\phi_L(x)}::e^{\mathbf{i}\phi_L(0)}:= e^{-\mathbf{i}\pi\mathsf{sign}(x)}:e^{\mathbf{i}\phi_L(0)}::e^{\mathbf{i}\phi_L(x)}:= -:e^{\mathbf{i}\phi_L(0)}::e^{\mathbf{i}\phi_L(x)}:.$$

More generally, the same calculation shows that

$$:e^{\mathbf{i}n\phi_L(x)}::e^{\mathbf{i}m\phi_L(0)}:= (-1)^{nm}:e^{\mathbf{i}m\phi_L(0)}::e^{\mathbf{i}n\phi_L(x)}:$$

So $e^{\mathbf{i}n\phi_L}$ is bosonic for even n and fermionic for odd n.

You can also see the fermi statistics from the OPEs:

$$:e^{\mathbf{i}\phi_L(z_1)}::e^{-\mathbf{i}\phi_L(z_2)}:\sim \frac{1}{z_1-z_2} = -\frac{1}{z_2-z_1}.$$

$$:e^{\mathbf{i}\phi_L(z_1)}::e^{\mathbf{i}\phi_L(z_2)}:\sim (z_1-z_2):e^{\mathbf{i}\phi_L(z_1)+\mathbf{i}\phi_L(z_2)}:$$

Notice that this object

$$e^{\mathbf{i}\frac{1}{\pi}\int_{-\infty}^{x_0} dx \partial_x \boldsymbol{\phi}_L} = e^{\mathbf{i}\int_{-\infty}^{x_0} dx j_L(x)}$$

has a lot in common with the Jordan-Wigner string, which counts the number of particles to the left.

Periodicity of chiral bosons. Since $\phi \simeq \phi + 2\pi m, m \in \mathbb{Z}$ and $\phi = \phi_L + \phi_R$ it is tempting to suggest that ϕ_L has period π. Such a period would mean that $e^{-\mathbf{i}\phi_L} \sim \Psi_L$ is not single-valued. But it is not quite correct. To see the correct statement, notice that the T-dual

field is $\tilde{\phi} = \phi_L - \phi_R$. With our normalization, this field also has periodicity 2π. The winding modes of the dual fields are momentum modes of the original one, so

$$\begin{matrix}\phi_L + \phi_R \simeq \phi_L + \phi_R + 2\pi m \\ \phi_L - \phi_R \simeq \phi_L + \phi_R + 2\pi j\end{matrix} \implies \begin{pmatrix}\phi_L \\ \phi_R\end{pmatrix} \simeq \begin{pmatrix}\phi_L \\ \phi_R\end{pmatrix} + \pi \begin{pmatrix}m+j \\ m-j\end{pmatrix} ,$$

where m and j are independent integers. But then $m \pm j$ are *not* independent integers: when $m + j$ is odd or even, so is $m - j$. This is the same fact as my earlier statement about the labels α, β on $\mathcal{V}_{\alpha,\beta}$ at Eq. (81).

Note that the T-dual field is often called $\vartheta \equiv \phi_L - \phi_R$ in discussions of Luttinger liquids (*e.g.* in Fradkin's and Sachdev's books).

Possibly Depressing comment. So now you are starting to see that this duality business is actually often a sad story: we thought we could solve *two* systems (free bosons and free fermions) but since they are really the same system in disguise, it turns out we can only solve one!

Multiple fermions. Notice that if we had N complex fermions ψ^a (for example a could be a spin index), we could bosonize them by N scalar fields at the free fermion radius:

$$\psi^a(z) \overset{?}{\sim} e^{\mathbf{i}\phi^a(z)}, \quad a = 1..N \ . \tag{90}$$

This formula *almost* works, but it has the problem that the operators on the RHS with different values of a will still commute (since the associated scalars don't have any singularities in their OPE, *i.e.* they commute). This must be fixed by so-called *Klein factors* (or 2-cocycles) which introduce the extra necessary signs. The necessary object is constructed by ordering the species of fermions $a = 1..N$ and defining:

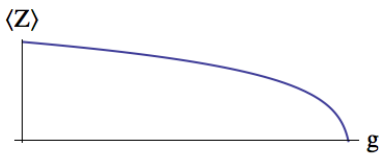
$$\mathfrak{c}_a \equiv (-1)^{\sum_{b<a} \mathbf{N}_b}$$

where $\mathbf{N}_b$ is the total fermion number of species b. The correct version of (90)

$$\psi^a(z) = \mathfrak{c}_a e^{\mathbf{i}\phi^a(z)}, \quad a = 1..N \qquad \text{(no sum on } a\text{)}. \tag{91}$$

Spin fields in (two copies of) the Ising model.

Now I will partially fulfil my promise to use bosonization to determine the scaling dimension of the spin operator at the Ising critical point. Recall that I claimed without any justification that in the TFIM, the behavior of the magnetization $\langle Z \rangle$ for $g \lesssim g_c$ was



$$\langle Z \rangle \sim (g_c - g)^{1/8}$$

(as in the figure). This 1/8 is the scaling dimension of the order parameter operator in the critical Ising CFT, the operator onto which the spin Z_j matches.

To begin, let's consider just the holomorphic part of a complex fermion $\psi(z) = \frac{1}{\sqrt{2}}(\chi_1(z) + \mathbf{i}\chi_2(z))$. This is not exactly the Ising critical theory, which is instead a non-chiral majorana mode, $\chi_L(z), \chi_R(\bar{z})$, but this is what we know how to bosonize. The bosonization map is

$$\psi(z) \sim e^{\mathbf{i}\phi(z)}.$$

Recall that the fermions in our spin chain arose as domain wall operators. They satisfy

$$\mathbf{Z}_l \mathbf{c}_j = -\mathbf{c}_j \mathbf{Z}_l, \quad l > j.$$

We can interpret this formula to say that the spin $\mathbf{Z}_j$ creates a branch cut for the fermion field. The CFT version of this statement is:

$$\psi(z)\sigma(0) \sim z^{-\frac{1}{2}}\mu(0) \quad . \tag{92}$$

Here σ, μ are *spin fields* (or more generally, *twist fields*) whose presence at $z = 0$ creates a branch cut. The actual location of the branch cut is not physical, but the location of the branch *points* (here $0, \infty$) is meaningful. Can we find operators which accomplish this?

Yes: they are

$$\sigma(w) = e^{-\frac{\mathbf{i}}{2}\phi(w)}, \quad \mu(w) = e^{+\frac{\mathbf{i}}{2}\phi(w)}. \tag{93}$$

To check this claim, their OPEs with the fermion operator are:

$$\begin{aligned} \psi(z)\sigma(0) &=: e^{\mathbf{i}\phi(z)} :: e^{-\frac{\mathbf{i}}{2}\phi(0)} :\sim e^{+\frac{\mathbf{i}}{2}\phi(0)} z^{-\frac{1}{2}} + \text{regular} \\ \psi(z)\mu(0) &=: e^{\mathbf{i}\phi(z)} :: e^{+\frac{\mathbf{i}}{2}\phi(0)} :\sim e^{+3\frac{\mathbf{i}}{2}\phi(0)} z^{+\frac{1}{2}} + \text{regular} \end{aligned} \tag{94}$$

As a check on my ability to do wick contractions, consider the OPE between (and hence the two-point function of) two of the spin fields, which is a special case of (82):

$$e^{\alpha \mathbf{i}\phi(z)} e^{-\alpha \mathbf{i}\phi(0)} \sim \frac{1}{z^{\alpha^2}} \ .$$

For $\alpha = \pm\frac{1}{2}$, this gives

$$\Delta_\sigma = \Delta_\mu = \frac{1}{2}\alpha^2|_{\alpha=\frac{1}{2}} = \frac{1}{8}.$$

The dimension of $e^{\mathbf{i}\frac{3}{2}\phi}$ is $\frac{1}{2}\left(\frac{3}{2}\right)^2 = \frac{9}{8}$ so the dimensional analysis in (94) works.

The answer $\Delta_\sigma = \frac{1}{8}$ is the dimension of the spin field in the theory of a *complex* chiral fermion. This is two copies of the left-moving part of the Ising CFT. The spin field in the Ising model which makes a branch cut for *both* χ_L and χ_R has the same dimension (the one that makes a branch cut in just χ_L has dimension $\frac{1}{16}$). The basic fact I am using here is that the spin-field dimension is additive in the number of fermion fields which feel the branch cut, that is, in the number of fermion fields whose boundary conditions are twisted.

To see this, let's make spin fields for N complex fermions ψ^a, each bosonized as we just did by N scalar fields

$$\psi^a(z) \sim e^{\mathbf{i}\phi^a(z)}, \quad a = 1..N$$

The operators

$$\sigma_s(z) \equiv e^{\mathbf{i}s^a\phi^a(z)}, \quad \{s^a = \pm\frac{1}{2}\}$$

create branch cuts for all N fermions at once:

$$\psi^a(z)\sigma_s(0) \sim z^{s^a}\sigma_{s'}(0) + \text{regular}$$

where $(s')^a = s^a + 1, (s')^{b\neq a} = s^b$. The dimension of these operators is the sum of the dimensions:

$$\Delta_s = \frac{1}{2}\left(s_1^2 + s_2^2 + \cdots + s_N^2\right) = \frac{N}{8}.$$

Now take $N = \frac{1}{2}$ to get the answer for a single majorana fermion. If you are not satisfied by this argument, there is a more direct way to compute the dimension of the Ising model spin field, using the defining equation (92) which requires a bit more CFT technology (it's explained on p. 74 of Ginsparg's CFT notes). We'll come to that soon.

Derivation of spin field.

Notice that the object (93) would not be considered single-valued on the original boson hilbert space! However, the operator which creates a branch cut for *both* ψ_L and ψ_R at the same time is $e^{\frac{\mathbf{i}}{2}(\phi_L(z)\pm\phi_R(\bar{z}))}$ and this *is* an allowed operator. And in fact, here is a *derivation* of its relation to the spin operator. Recall the JW formula

$$\boldsymbol{\sigma}_j^+ = (-1)^{\sum_{i<j} n_i}\mathbf{c}_j^\dagger.$$

(Previously we put the JW string to the right; here I defer to the bosonization convention we used above.) Here (in the basis we used earlier, where $\mathbf{Z}$ is the Ising order parameter), $\boldsymbol{\sigma}^+ \equiv \mathbf{Z} + \mathbf{i}\mathbf{Y}$, so $\boldsymbol{\sigma}^+ + \boldsymbol{\sigma}^- = 2\mathbf{Z}$. In the continuum, in te rms of the fermions, this is

$$\boldsymbol{\sigma}^+(x) = e^{\mathbf{i}\pi\int_{-\infty}^x dx' j_0(x')}\Psi(x)^\dagger \ .$$

Now, we must use (65) to expand our fermion operator near the fermi surface. Let's do just the R piece first (then we'll put it all together):

$$\begin{array}{rl} & e^{\mathbf{i}\pi\int_{-\infty}^x dx' j_0(x')}\Psi_R(x)^\dagger \\ \overset{(87)}{=} & e^{-\mathbf{i}\pi\int_{-\infty}^x dy\ \frac{\partial_y\phi_L+\partial_y\phi_R}{2\pi}}e^{-\mathbf{i}\phi_R(x)} \\ = & e^{\mathbf{i}\frac{\phi_L(x)+\phi_R(x)}{2}}e^{-\mathbf{i}\phi_R(x)} \\ \sim & e^{\mathbf{i}\frac{\phi_L(x)-\phi_R(x)}{2}} \end{array} \tag{95}$$

OK, that was practice; the full expression is:

$$\boldsymbol{\sigma}^+(x) \quad = \quad e^{\mathbf{i}\pi\int_{-\infty}^x dx' j_0(x')}\Psi(x)^\dagger$$

$$\overset{(87),(65)}{=} e^{-\mathbf{i}\pi\int_{-\infty}^{x}dy\ \frac{\partial_y\phi_L+\partial_y\phi_R}{2\pi}}\left(e^{-\mathbf{i}k_Fx}e^{-\mathbf{i}\phi_R(x)}+e^{+\mathbf{i}k_Fx}e^{-\mathbf{i}\phi_L(x)}\right)$$

$$\sim e^{\mathbf{i}\frac{\phi_L(x)-\phi_R(x)}{2}}e^{-\mathbf{i}k_Fx}+e^{-\mathbf{i}\frac{\phi_L(x)-\phi_R(x)}{2}}e^{+\mathbf{i}k_Fx} \tag{96}$$

[End of Lecture 15]

The role of the anomaly in bosonization. Recall that I promised that the chiral anomaly would play a role in bosonization. Here's one way in which it does. Consider the canonical commutator of the bose mode operators ($\partial_z\phi \propto \sum_n \boldsymbol{\rho}_n z^{-n-1}$):

$$[\boldsymbol{\rho}_n, \boldsymbol{\rho}_{-n'}] = \delta_{n,n'}$$

in terms of the fermions, $\boldsymbol{\rho}_n = \sum_l \mathbf{L}_l^\dagger \mathbf{L}_{l+n}$, we have:

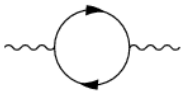
$$\begin{aligned}[\boldsymbol{\rho}_n, \boldsymbol{\rho}_{-n'}] &= \sum_{l,'l}[\mathbf{L}_l^\dagger\mathbf{L}_{l+n}, \mathbf{L}_{l'}^\dagger\mathbf{L}_{l'-n'}] \\ &= \sum_{l'}\left(\mathbf{L}_{l'-n}^\dagger\mathbf{L}_{l'-n'} - \mathbf{L}_{l'}^\dagger\mathbf{L}_{l'+n-n'}\right) \overset{?}{=} 0\end{aligned} \tag{97}$$

This expression has the classic form of an anomaly: it looks like this vanishes since we can redefine the dummy index in the second term by $n' \to n'+n$ and make the two terms look the same with opposite sign. *However*, in the presence of a UV cutoff, *i.e.* in the regulated theory, that shift will shift the cutoff and will not be innocuous since the summand is not zero for arbitrarily high energies.[32] If, for example, we regulate the short distance behavior using normal-ordered operators, we have

$$\begin{aligned}[\boldsymbol{\rho}_n, \boldsymbol{\rho}_{-n'}] &= \sum_{l,'l}[:\mathbf{L}_l^\dagger\mathbf{L}_{l+n}:, :\mathbf{L}_{l'}^\dagger\mathbf{L}_{l'-n'}:] \\ &= \delta_{nn'}\sum_l\left(\langle\mathbf{L}_{l-n}^\dagger\mathbf{L}_{l-n}\rangle - \langle\mathbf{L}_n^\dagger\mathbf{L}_n\rangle\right) = \delta_{nn'}n.\end{aligned} \tag{98}$$

This calculation is related to the anomaly because it says that the OPE of the chiral current j_R with itself has a singularity with a definite coefficient. Specifically (as you can check on HW 6),

$$j(z)j(0) \sim \frac{k}{z^2}$$

with $k = 1$ is called the *level* of the (here $\mathsf{U}(1)$) current algebra. This singularity implies the anomalous nonconservation when coupling to a background gauge field. In terms of the fermions, both of these effects come from the 'diangle' diagram: 

[32] To see how the high-energy stuff comes in consider the analogous infinite-volume expression

$$0 = \int_{-\infty}^{\infty}dp(f(p)-f(p)) \overset{?}{=} \int_{-\infty}^{\infty}dp(f(p+s)-f(p)) = \int dp(sf'(p)+\mathcal{O}(s^2)) = s(f(\infty)-f(-\infty)).$$

Twist fields and boundary conditions. Above I used the phrase "fermion fields whose boundary conditions are twisted" by the spin fields. With this phrase I betrayed that I am secretly using the following picture, which is called *radial quantization* (more on this at §4.2). Imagine that the radial coordinate of the plane is the time coordinate (it is euclidean so this interpretation is not problematic). Then equal-time surfaces are circles centered at the origin. (We'll see below that in a CFT this is equivalent to the ordinary notion of time evolution by a conformal transformation.) In this case, the boundary conditions on the spatial coordinate are determined by going around the origin; if there's a branch point at the origin, we must cross the branch cut in doing so, and the field comes back to itself up to the discontinuity across the branch cut, which here is a minus sign. So the spin fields can be regarded as "boundary-condition-changing operators": they take the fermions from the sector with PBC (and hence a fourier expansion with integer powers of z: $\psi_{PBC}(z) = \sum_{n\in\mathbb{Z}} z^n \psi_n$; this is called the *Ramond sector*) to APBC (half-integer mode expansion $\psi_{APBC}(z) = \sum_{n\in\mathbb{Z}} z^{n+\frac{1}{2}}\psi_{n+\frac{1}{2}}$; this is called the *Neveu-Schwarz sector*) and vice versa.

Understanding in terms of group theory. I must comment on the very essential group theory underlying the previous discussion of spin fields for multiple fermions. The theory of N complex free chiral fermions has a manifest $\mathsf{U}(N)$ symmetry rotating the complex fermions into each other: $\psi^a \to \mathsf{U}^a_b\psi^b$, $\mathsf{U}\mathsf{U}^\dagger = 1$. The free fermion action actually enjoys a larger symmetry which ignores the grouping into complex fermions: $\psi^a = \frac{1}{\sqrt{2}}(\chi^{2a-1} + \mathbf{i}\chi^{2a})$; since $S \propto \int \delta_{AB}\chi^A\bar\partial\chi^B$, these $2N$ chiral majorana fermions transform in the fundamental of an $\mathsf{SO}(2N)$ symmetry, $\chi^A \to \mathsf{O}^A_B\chi^B, \mathsf{O}^\mathsf{T}\mathsf{O} = 1$.

This $\mathsf{SO}(2N)$ acts on the hilbert space. In terms of the fermions, the generators are just fermion bilinears $j^{AB} = \chi^A\chi^B$, the currents (recall that the adjoint (the rep that the generators transform in) of $\mathsf{SO}(2N)$ is made from antisymmetrizing two fundamentals). There are $2N(2N-1)/2$ of these.

In terms of the bosons, the generators come in two types (in the theory of Lie groups this is called Cartan-Weyl basis). N of them are just the ordinary boson currents $\partial\phi^a$. This is called the *Cartan subalgebra* (the number of them is $N = \text{rank}(\mathsf{SO}(2N))$ – they all commute with each other. The non-commuting (charged) generators are vertex operators $e^{\mathbf{i}\phi^a w_a}$ where in order for this to be a conserved current, it must have dimension 1, which requires $\sum_{a=1}^N w_a w_a = 2$. (in Lie group terminology, w_a is a *root vector*). Recall that the entries of w must be integers. This condition is solved by vectors of the form $\{w_a\} = \{(0...0, \pm 1, 0\cdots, 0, \pm 1, ...)\}$, of which there are $N(N-1)/2\times 4$. Altogether we have recovered the same number of generators $N + 2N(N-1) = N(2N-1), \checkmark$.

The states in the Neveu-Schwarz (APBC) sector are of the form

$$\psi^A_{-\frac{1}{2}}\psi^B_{-\frac{3}{2}}\cdots|0\rangle, \qquad \psi^A_{r>0}|0\rangle = 0, \forall r$$

and therefore transform in various antisymmetric tensor representations of this $\mathsf{SO}(2N)$.

What is going on with these spin field objects $e^{\mathbf{i}s^a\phi^a} = e^{\mathbf{i}\left(\pm\frac{1}{2}\phi^1\pm\frac{1}{2}\phi^2\cdots\pm\frac{1}{2}\phi^N\right)}$? These operators

map the Neveu-Schwarz sector to the Ramond sector. For a hint about what they are doing, notice that the fermion zeromodes in the Ramond (PBC) sector satisfy

$$\{\psi_0^a, \psi_0^b\} = 2\delta^{ab} \implies \{\chi_0^A, \chi_0^B\} = \delta^{AB}$$

which is the algebra of Dirac gamma matrices for $\mathsf{SO}(2N)$, *i.e.* Clifford algebra. This algebra must be represented on the Ramond groundstates, which are therefore not unique, *i.e.* they are degenerate. The point in life of Clifford algebra is to construct spinor representations. So the answer is that these vectors $\{s^a\} = \{\pm\frac{1}{2}, \pm\frac{1}{2}, \cdots, \pm\frac{1}{2}\}$ are *weight vectors* for the spinor representation(s) of $\mathsf{SO}(2N)$, in which the Ramond groundstates transform. The spin fields create the Ramond groundstates from the NS groundstate:

$$\sigma_s(0)|0\rangle_{NS} = |s\rangle.$$

A very interesting application of this group theory arises in the demonstration (by Fidkowski and Kitaev) that eight free majorana chains in the topological phase (with the dangling majoranas at the ends) can be adiabatically connected to the trivial phase (only) via interacting Hamiltonians.

3.1.2 Application 2: Briefly, what is a Luttinger liquid?

[Fradkin p. 152] Let us consider a model which is actually made of fermions (which may have spin) at finite chemical potential. The actual density of the microscopic fermions Ψ is

$$\boldsymbol{\rho}_{\text{actual}}(x) = \sum_\sigma \Psi_\sigma^\dagger(x)\Psi_\sigma(x) = \underbrace{\rho_0}_{=N_{\text{spin}}\frac{k_F}{\pi}} + \underbrace{\boldsymbol{\rho}_L}_{=\psi_L^\dagger\psi_L} + \boldsymbol{\rho}_R + \sum_\sigma \left(e^{2k_F\mathbf{i}x}\psi_R^\dagger\psi_L + h.c.\right) ...$$

where I am reminding you that $\Psi(x) = e^{\mathbf{i}k_F x}\psi_R(x) + e^{-\mathbf{i}k_F x}\psi_L(x)$. The .... are even-higher momentum contributions which we will ignore. The quantity $2k_F$ is important in any system with a fermi surface as it is the *diameter* of the fermi surface, and (bosonic) particle-hole excitations near (just below) this momentum can be cheap despite its large value.

Consider the hamiltonian $\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_{\text{pot}}$ where the free Galilean-invariant hamiltonian is

$$\mathbf{H}_0 = \sum_\sigma \int đp\ pv_F \left(\psi_{\sigma R}^\dagger\psi_{\sigma R} - \psi_{\sigma L}^\dagger\psi_{\sigma L}\right) .$$

Subject this system to an (at-the-moment) external potential

$$\mathbf{H}_{\text{pot}} = \int dx\ eV(x)\boldsymbol{\rho}(x) = -e\int đp \int đq\tilde{V}(q)\Psi^\dagger(p+q)\Psi(p) .$$

Let us consider the special wavenumber $V(x) = V_0\cos(2k_F x)$, so

$$\mathbf{H}_{\text{pot}} = -eV_0\int dp\psi_R^\dagger(p)\psi_L(p) + h.c. = -eV_0\int dx\psi_R^\dagger(x)\psi_L(x) + h.c.$$

In the relativistic language, this is a mass term:

$$\mathbf{H} = \int dx \psi^\dagger \left(\sigma^3 v_F \mathbf{i} \partial_x + e V_0 \sigma^1 \right) \psi = \int dx \left(\bar{\psi} v_F \gamma^x \partial_x \psi + e V_0 \bar{\psi} \psi \right)$$

(here I used $\gamma^0 = \sigma^1, \gamma^x = \sigma^1 \sigma^3, \bar{\psi} \equiv \psi^\dagger \gamma^0$).

Where does this potential come from? It can be generated by (*e.g.* Coulomb) interactions between the fermions themselves

$$\mathbf{H}_{\text{int}} = \int dx e V \boldsymbol{\rho} + \frac{e}{2} \int \int V K V \ .$$

If we treat this interaction in mean field theory, we get back (for some choice of K more on which more below) the previous background potential, $\mathbf{H}_{\text{int}} \sim \mathbf{H}_{\text{pot}}$. On the other hand, if we integrate out V, we find a density-density interaction

$$\mathbf{H}_{\text{int}} = e \int \int \boldsymbol{\rho} K^{-1} \boldsymbol{\rho}$$

which for the right kernel, K, might be the Coulomb interaction. In a metal, the coulomb interaction is screened and therefore short-ranged. Expanding in this range, we find

$$\mathbf{H}_{\text{int}} = \int dx \left(\underbrace{2 g_2 \boldsymbol{\rho}_R(x) \boldsymbol{\rho}_L(x)}_{\text{back-scattering, 1d}} + \underbrace{g_4 \left(\boldsymbol{\rho}_R^2 + \boldsymbol{\rho}_L^2 \right)}_{\text{forward scattering, ubiquitous}} + g_u \lim_{y \to x} \left(\psi_R^\dagger(x) \psi_R^\dagger(y) \psi_L(x) \psi_L(y) + h.c. \right) \right)$$

The gs can be written in terms of various (fourier) components of the interaction potential K^{-1}. In the last term, you should think of $y = x + a$, one lattice spacing. The last term describes a process whereby two L/R fermions turn into two R/L fermions, which clearly violates momentum conservation by $\pm 2k_F$; this can happen in a lattice model if the two sides of the fermi surface differ by a reciprocal lattice vector:

$$k_F \equiv -k_F + \frac{2\pi}{a}.$$

This is called *umklapp* scattering. When does it happen? k_F is determined by the density of electrons. The maximum density is when all states are filled and is $N_s \frac{2\pi}{a}$, and then there is no fermi surface. Umklapp happens at *half-filling*, when $k_F = \frac{\pi}{a}$. Otherwise g_u can be ignored. Notice that this effect explicitly breaks the chiral symmetry $\psi_{L/R} \to e^{\mathbf{i}\alpha_{L/R}} \psi_{L/R}$ which independently rotates the phases of the left-movers and right-movers down to a $\mathbb{Z}_2$ symmetry, which is just fermion number (which is hard to violate).

Bosonize. Now we bosonize. Using $\phi = \phi_L + \phi_R, \Theta = \phi_L - \phi_R$ and our dictionary above, we find $\mathbf{H}_0 \equiv \int dx \mathbf{h}_0(x)$, with

$$\mathbf{h}_0(x) = \frac{v_F}{2} \left(\Pi^2 + (\partial_x \phi)^2 \right) = \frac{v_F}{2} \left((\partial_x \Theta)^2 + (\partial_x \phi)^2 \right)$$

where in the second step we used the canonical commutators we found previously. Using $\boldsymbol{\rho}_L = \frac{1}{2\pi}\partial_x\phi_L$ etc this is

$$\mathbf{h}_0(x) = \pi v_F \left((\boldsymbol{\rho}_L + \boldsymbol{\rho}_R)^2 + (\boldsymbol{\rho}_L - \boldsymbol{\rho}_R)^2 \right) = 2\pi v_F \left(\boldsymbol{\rho}_L^2 + \boldsymbol{\rho}_R^2 \right).$$

The forward scattering hamiltonian density is

$$\mathbf{h}_{\text{forward}}(x) = g_4 \left(\boldsymbol{\rho}_R^2 + \boldsymbol{\rho}_L^2 \right)$$

of exactly the same form as the free term (!) and therefore just renormalizes the velocity, the coefficient of the spatial gradient term. The back-scattering term is less innocuous

$$\mathbf{h}_{\text{back}}(x) = 2g_2 \boldsymbol{\rho}_R \boldsymbol{\rho}_L = \frac{g_2}{\#} \left((\partial_x \phi)^2 - \Pi^2 \right)$$

Umklapp is a little trickier: to define the bosonized form we must absorb the OPE singularity

$$: e^{-\mathbf{i}\phi_R(x)} :: e^{-\mathbf{i}\phi_R(y)} :\sim (x-y)^{\frac{1}{2\pi T}} : e^{-2\mathbf{i}\phi_R(x)} : + \ldots$$

into the coupling constant:

$$\mathbf{h}_u = \tilde{g}_u \cos 2\phi.$$

(It is useful to set $\tilde{g}_u = g_u a^{-\frac{1}{2\pi T}}$ where a is a short-distance cutoff.) With $\tilde{g}_u \neq 0$, the remaining chiral symmetry acts by $\phi \to \phi + \pi$.

The full hamiltonian density (without umklapp, so away from half-filling) is

$$\mathbf{h} = \frac{v}{2} \left(\frac{1}{K} (\partial_x \Theta)^2 + K (\partial_x \phi)^2 \right)$$

where v and K can be written in terms of the gs and v_F[33]. Notice that T-duality is clear in this expression as $\mathsf{T} : K \leftrightarrow \frac{1}{K}, \phi \leftrightarrow \Theta$.

A few comments:

- The backscattering term g_2 affects both the velocity and the 'stiffness' K (this is the parameter we called the tension T earlier). The signs are such that a repulsive interaction $g_2 > 0$ increases K.
- In the absence of umklapp, the chiral symmetry acts on the boson as $\phi \to \phi + \alpha$, $\alpha \in [0, 2\pi)$. Think about how this acts on the fermions: since $\boldsymbol{\rho} = \ldots + \frac{1}{2\pi}\partial\phi$, the density transforms as

  $$\boldsymbol{\rho}(x) \to \rho_0 + \boldsymbol{\rho}_R + \boldsymbol{\rho}_L + \left(e^{\mathbf{i}2(k_F x - \alpha)} \psi_R^\dagger(x) \psi_L(x) + h.c. \right) = \boldsymbol{\rho}\left(x - \frac{\alpha}{k_F} \right)$$

  – this is a rigid translation of the density profile! It makes sense then that lattice effects would break this symmetry.

[33] The relation is

$$v = \sqrt{(v_F + g_4/\pi)^2 - (g_2/\pi)^2}, \quad K = \sqrt{\frac{v_F + g_4/\pi + g_2/\pi}{v_F + g_4/\pi - g_2/\pi}}.$$

- We have ignored the fact that the dispersion will not be exactly linear arbitrarily far from the Fermi surface. Such terms can also be bosonized and become things like $(\partial_x \phi)^3$ which change the story dramatically; in particular they will lead to damping of the 'sound mode' described by ϕ even away from half-filling.

- If we study spinful fermions, there are even more 4-fermion terms to consider. A payoff of studying the case with spin is that it provides an example of *spin-charge separation*: by bosonizing like $\phi_c \equiv \phi_\downarrow + \phi_\uparrow, \phi_s \equiv \phi_\downarrow - \phi_\uparrow$ we can see that the charge sector (ϕ_c) and the spin sector (ϕ_s) decouple and can have independent vs and Ks – the spin excitations of the electron travel at a different speed than its charge! There is a lot more to say about this and I recommend Fradkin's treatment in chapter 6.

[End of Lecture 16]

How is a Luttinger liquid different from a Fermi liquid?

Continuing with the model with actual microscopic fermions, we have encountered a gapless phase in $D = 1 + 1$ described by free fermions, which we also know how to build in $D > 1+1$. The excitations are characterized by long-lived fermionic quasiparticles with the same $\mathsf{U}(1)$ charge as the constituent fermions $\mathbf{c}_j$ – they are Landau quasiparticles. This has the consequence that there are sharp peaks in the fermion green's functions,

$$\int e^{\mathbf{i}\omega t - \mathbf{i}kja} \langle \mathbf{c}_j^\dagger(t)\mathbf{c}_0(0)\rangle = G(\omega, k) \sim \frac{Z}{\omega - v_F(k - k_F)}$$

(the RHS is the contribution from the right-mover $\Psi_R(t - x)$). There is (therefore) a *jump* in the fermion occupation number at $k = k_F$:

$$n(k) \equiv \mathbf{c}^\dagger(k)\mathbf{c}(k) = \int_{-\infty}^{\infty} \frac{d\omega}{\pi} \underbrace{\mathrm{Im}G(\omega, k)}_{\sim Z\pi\delta(\omega - v_F(k-k_F))} = Z\theta(k - k_F).$$

[Fradkin §6.8] These last two characteristic features of Fermi liquid theory can be destroyed by interactions in $1 + 1$ dimensions, as we can see using the boson description. As a simple first step, notice that as we change T away from $\frac{1}{\pi}$, the dimension of the operator which at $T = \frac{1}{\pi}$ creates the right-moving fermion $e^{\mathbf{i}\phi_R}$ changes away from $\frac{1}{2}$:

$$\langle e^{\mathbf{i}\phi_R(x,t)} e^{-\mathbf{i}\phi_R(0,0)}\rangle \sim \frac{1}{|t - x|^{\frac{1}{\pi T}}} \ .$$

It is tempting to conclude that the fermion correlation function $\Psi(x) = e^{\mathbf{i}k_F x}\Psi_R(x) + ...$ is

$$\langle \Psi(x, t)^\dagger \Psi(0, 0)\rangle \overset{?}{\sim} \frac{\sin(k_F x)}{|t - x|^{\frac{1}{\pi T}}}$$

The fourier transform of this has a broad power-law peak (a branch point) instead of a pole; the spectral density has no delta function, but rather goes like

$$\mathrm{Im}\, G(\omega, k) \overset{?}{\sim} |\omega - v_F(k - k_F)|^{\frac{1}{\pi T}+1}\theta(\omega - v_F(k - k_F)),$$

and hence there is no jump at $k = k_F$.

This is not quite the correct story however, since the operator $e^{\mathbf{i}\phi_R}$ is not quite the operator which creates an electron, except at $T = \frac{1}{\pi}$ (for example, it is not a fermionic object!). Instead what happens is that the interactions mix the R and L operators; the operator which creates the rightmoving electron is actually the linear(-in-the-exponent) combination:

$$\Psi_R = e^{\mathbf{i}\alpha_+\phi_R + \mathbf{i}\alpha_-\phi_L}, \quad \text{with } \alpha_\pm = \frac{(K \pm 1)^2}{4K}$$

(recall that $K = \frac{1}{\pi T}$); it has total dimension $\Delta = \frac{1}{4}\left(K + \frac{1}{K}\right)$. The statement about the pole being replaced by a branch point is correct.

(Pieces of) the fermion Green's function are (relatively) directly measured in angle-resolved photoemission spectroscopy (ARPES) experiments. Other combinations of them can be measured by scanning-tunneling microscopy (STM).

[Fradkin, 2d ed. p. 170] There is a lot more to say about this subject. In particular, one can understand in great detail various ordering instabilities of interacting fermions in one dimension using the technology we have developed.

3.1.3 Application 3: sine-Gordon and Thirring

Let us reconsider the scale invariant theory described by the free scalar again.

$$\mathcal{L} = \frac{T}{2}\left(\partial_\mu \phi\right)^2 + g_n : \cos n\phi :$$

I allow the possibility of a cosine term. For the case of the Luttinger liquid we found that umklapp produced $n = 2$. There are several important points to make here:

(-1) In the absence of lattice effects or a background gauge field, there is a $\mathsf{U}(1) \times \mathsf{U}(1)$ symmetry acting on ϕ_L and ϕ_R independently.

(0) The theory is scale invariant for any T: there is a *line of fixed points* parametrized by T. That is, T is the coefficient of an *exactly marginal operator*. This is weird and the only examples in $D > 1 + 1$ (that I know) require supersymmetry.

(1) by changing T we can study interacting fermion systems.

(2) the fermion mass corresponds to a $\cos\phi$ term.

(3) These $\cos n\phi$ terms have a dimension that varies with T. This means that as we vary T they can go from irrelevant to relevant. If they are relevant they produce a mass for everybody (You can see this by Taylor expanding about the minimum of the potential). This happens when

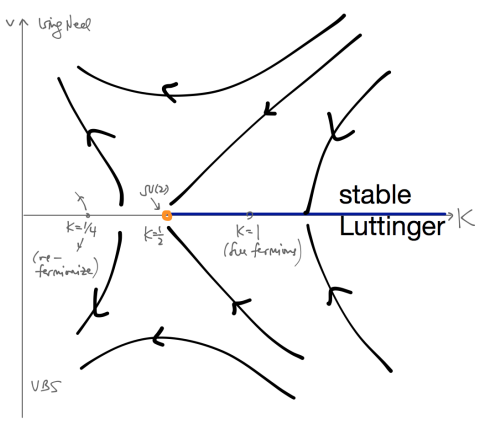
$$\Delta_n = \frac{n^2}{2\pi T} = K\frac{n^2}{2} \leq 2$$

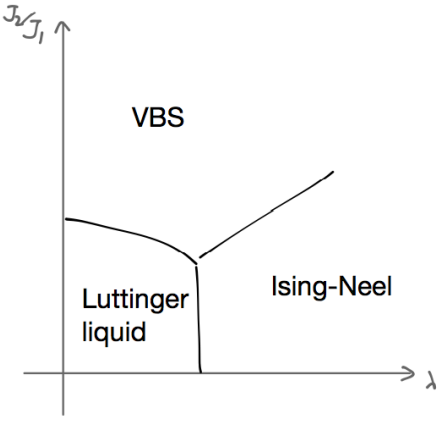
that is, for $T > T_c = \frac{n^2}{4\pi}$ (or $K < K_c = \frac{n^2}{2}$)

Brief comments on phase diagram. [Sachdev Ch. 20 – beware that $\phi_{\text{Sachdev}} = 2\phi_{\text{here}}$] The RG equations (for $n = 1$) are

$$\frac{dv}{d\ell} = (2 - 4K)v, \quad \frac{dK}{d\ell} = -\delta v^2$$

where $\delta > 0$ is some non-universal quantity. This leads to the following RG flow diagram in the space of these couplings (from Sachdev chapter 20):

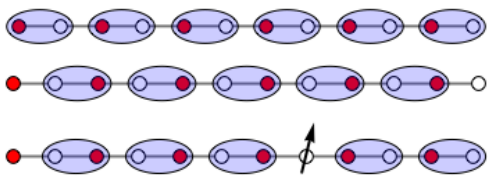




The figure at right is the phase diagram we can then infer for the model

$$\mathbf{H} = -J_1 \sum_{\langle ij \rangle} (\mathbf{X}_i \mathbf{X}_j + \mathbf{Y}_i \mathbf{Y}_j + \lambda \mathbf{Z}_i \mathbf{Z}_j) - J_2 \sum_{\langle\langle ij \rangle\rangle} (\mathbf{X}_i \mathbf{X}_j + \mathbf{Y}_i \mathbf{Y}_j + \mathbf{Z}_i \mathbf{Z}_j) \ .$$

where $\langle\langle ij \rangle\rangle$ denotes next-nearest neighbors. The simpler model we've been discussing moves along the λ axis.



• When $v \to -\infty$, we can find the groundstates by making the $-|v| \cos 2\phi$ happy first. The minima occur at $\phi = (n + \frac{1}{2})\pi$, $n \in \mathbb{Z}$. They can be distinguished by the value of $\sin \phi_i = (-1)^n = \langle (-1)^i \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} \rangle$. This is described by singlets on alternating bonds. This is called a *valence bond solid* state. The low energy excitations in this state are interesting – they are not just the gapped ϕ excitations rolling in the minima of the cosine potential. Rather, the domain walls between the two groundstates (created in the CFT by $e^{\mathbf{i}\frac{1}{2}\phi}$!) have lower energy. We can see this explicitly at the special point $K = \frac{1}{4}$ (which is $< K_c$) by 'refermionizing'. The previous original fermions where created by $\psi_L^\dagger = e^{\mathbf{i}\phi_R} = e^{\mathbf{i}\frac{1}{2}(\phi+\Theta)}$. Introduce new fermions

$$\tilde{\psi}_L^\dagger \equiv e^{\mathbf{i}\frac{1}{2}\left(2\phi + \frac{\Theta}{2}\right)}, \tilde{\psi}_R^\dagger \equiv e^{\mathbf{i}\frac{1}{2}\left(2\phi - \frac{\Theta}{2}\right)}.$$

At $K = \frac{1}{4}$ these are free fermions, and the $v \cos 2\phi$ term is just a mass term $v\tilde{\psi}_L^\dagger \tilde{\psi}_R + h.c.$. Interestingly, there is a regime where the quantum of the boson can *decay* into a pair of these (fermionic) domain walls.

• The phase at $v \to +\infty$ is similar. Minimizing $+|v| \cos 2\phi$ gives $\phi = n\pi$ which are distinguished by $\cos \phi = (-1)^n = \langle (-1)^i \mathbf{Z}_i \rangle$. This is 'Ising Neel' order, that is: staggered Ising

spins. (Neel order refers to staggered magnetization in an arbitrary direction, spontaneously breaking $\mathsf{SU}(2)$ spin rotation symmetry.)

The two phases at large $|v|$ are related by the relabelling $v \to -v, \phi \to \phi + \pi/2$. In terms of the bosons or fermions both of these large $|v|$ states are charge density waves, where the density of particles breaks the translation symmetry down to a smaller discrete subgroup.

Most continuous phase transitions occur by tuning the coefficient of a relevant operator to zero (recall the TFIM example where we had to tune $g \to g_c$ to get massless fermions). This is *not* what happens in the XY model; rather, we are varying a *marginal* parameter and the dimensions of other operators depend on it and become relevant at some critical value of that marginal parameter. This leads to very weird scaling near the transition, of the form $e^{-\frac{a}{\sqrt{K-K_c}}}$ (for example, in the correlation length, the exponential arises from inverting expressions involving $G_R(z) = -\frac{1}{4\pi K}\log z$) – it is sometimes called an 'infinite order' phase transition, because all derivatives of such a function are continuous. I don't want to say more about how this comes about because I know Congjun Wu explained it in the RG class.

In the high-energy literature, this is called duality between the Thirring model and the sine-Gordon model.

3.1.4 XY transition from SF to Mott insulator

[This discussion is from Ashvin Vishwanath's lecture notes.] So far in this chapter, we've applied bosonization to understand systems made of spins, fermions or bosons. It will be useful to make the application to bosons more direct and explicit. Let's return for a moment to the Bose-Hubbard model (in any dimension)

$$\mathbf{H}_{BH} = -\tilde{J}\sum_{\langle ij\rangle}\left(\mathbf{b}_i^\dagger \mathbf{b}_j + h.c.\right) + \frac{U}{2}\sum_i \mathbf{n}_i\left(\mathbf{n}_i - 1\right) - \mu\sum_i \mathbf{n}_i$$

with $\mathbf{n}_i \equiv \mathbf{b}_i^\dagger \mathbf{b}_i$, and $[\mathbf{b}_i, \mathbf{b}_j^\dagger] = \delta_{ij}$. This Hilbert space has a useful *number-phase* representation in terms of

$$[\mathbf{n}_i, \phi_j] = -\mathbf{i}\delta_{ij}, \quad \phi_i \equiv \phi_i + 2\pi, \quad \mathbf{n}_i \in \mathbb{Z}$$

(where the last statement pertains to the eigenvalues of the operator). The bosons are

$$\mathbf{b}_i = e^{-\mathbf{i}\phi_i}\sqrt{\mathbf{n}_i}, \quad \mathbf{b}_i^\dagger = \sqrt{\mathbf{n}_i}e^{+\mathbf{i}\phi_i} \ ;$$

these expressions have the same algebra as the original $\mathbf{b}$s. In terms of these operators, the hamiltonian is

$$\mathbf{H}_{BH} = -\tilde{J}\sum_{\langle ij\rangle}\left(\sqrt{\mathbf{n}_i}e^{\mathbf{i}(\phi_i-\phi_j)}\sqrt{\mathbf{n}_j} + h.c.\right) + \frac{U}{2}\sum_i \mathbf{n}_i\left(\mathbf{n}_i - 1\right) - \mu\sum_i \mathbf{n}_i.$$

If $\langle \mathbf{n}_i \rangle = n_0 \gg 1$, so that $\mathbf{n}_i = n_0 + \Delta\mathbf{n}_i, \Delta\mathbf{n}_i \ll n_0$ then $\mathbf{b}_i = e^{-\mathbf{i}\phi}\sqrt{\mathbf{n}_i} \simeq e^{-\mathbf{i}\phi_i}\sqrt{n_0}$ and

$$\mathbf{H}_{BH} \simeq -\underbrace{2\tilde{J}n_0}_{\equiv J}\sum_{\langle ij\rangle}\cos\left(\phi_i - \phi_j\right) + \frac{U}{2}\sum_i \left(\Delta\mathbf{n}_i\right)^2 \equiv \mathbf{H}_{\text{rotors}}$$

where we set $n_0 \equiv \mu/U \gg 1$. This is a *rotor model*.

This model has two phases:
$\boxed{U \gg J}$: then we must satisfy the U term first and the number is locked, $\Delta\mathbf{n} = 0$ in the groundstate. This is a Mott insulator, with a gap of order U. Since $\mathbf{n}$ and ϕ are conjugate variables, definite number means wildly fluctuating phase.

$\boxed{U \ll J}$: then we must satisfy the J term first and the phase is locked, $\phi = 0$ in the groundstate, or at least it will try. This is the superfluid (SF). That is, we can try to expand the cosine potential[34]

$$\mathbf{H}_{\text{rotors}} = U\sum_i \mathbf{n}_i^2 - J\sum_{\langle ij\rangle}\cos\left(\phi_i - \phi_j\right) \simeq U\sum_i \mathbf{n}_i^2 - J\sum_{\langle ij\rangle}\left(1 - \frac{1}{2}\left(\phi_i - \phi_j\right)^2 + ...\right)$$

which is a bunch of harmonic oscillators and can be solved by Fourier: $\phi_i = \frac{1}{\sqrt{N^d}}\sum_i e^{-\mathbf{i}k\cdot x_i}\phi_k$, so

$$\mathbf{H} \simeq \sum_k \left(U\pi_k\pi_{-k} + J\left(1 - \cos ka\right)\phi_k\phi_{-k}\right)$$

This has gapless *phonon* modes at $k = 0$. I have written the hamiltonian in 1d notation but nothing has required it so far. The low energy physics is described by the continuum lagrangian density

$$L_{\text{eff}} = \frac{\rho_s}{2}\left(\frac{\left(\partial_\tau\phi\right)^2}{c} + c\left(\vec{\nabla}\phi\right)^2\right)$$

with $\rho_s = \sqrt{J/U}, c = \sqrt{JU}$. ρ_s is called the superfluid stiffness. This is the free scalar theory above, with $T = \rho_s$.

Now 1d comes in: In $d > 1$, there is long range order – the bosons condense and spontaneously break the phase rotation symmetry $\phi \to \phi + \alpha$; the variable ϕ is a Goldstone boson. In 1d there is no long-range order. The two phases are still distinct however, since one has a gap and the other does not. The correlators of the boson operator $b_i \sim e^{\mathbf{i}\phi_i}$ diagnose the difference. In the Mott phase they have exponential decay. In the "SF" they have

$$\langle e^{\mathbf{i}\phi(x)}e^{-\mathbf{i}\phi(y)}\rangle = \frac{c_0}{r^\eta}, \quad \eta = \frac{1}{2\pi\rho_s} = \frac{1}{2K}.$$

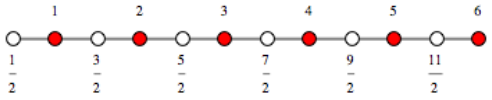
This is *algebraic long range order*.

However: what happens as U grows and makes the phase fluctuate more? Our continuum description in terms of harmonic oscillators hides (but does not ignore) the fact that $\phi \simeq \phi + 2\pi$. The system admits vortices, aka winding modes.

[34] From now on the background density n_0 will not play a role and I will write $\mathbf{n}_i$ for $\Delta\mathbf{n}_i$.

Lattice T-duality. To see their effects let us do T-duality on the lattice. As usual, the dual variables live on the bonds, labelled by $\bar{i} = \frac{1}{2}, \frac{3}{2}....$ Introduce

$$\mathbf{m}_{\bar{i}} \equiv \frac{\phi_{i+1} - \phi_i}{2\pi}, \quad \Theta_{\bar{i}} \equiv \sum_{j<\bar{i}} 2\pi \mathbf{n}_j \tag{99}$$



which together imply

$$[\mathbf{m}_{\bar{i}}, \Theta_{\bar{j}}] = -\mathbf{i}\delta_{\bar{i}\bar{j}}.$$

To understand where these expressions come from, notice that the operator

$$e^{\mathbf{i}\Theta_{\bar{i}}} = e^{\mathbf{i}\sum_{j<\bar{i}} 2\pi \mathbf{n}_j}$$

rotates the phase of the boson on all sites to the left of $\bar{i}$ (by 2π). It inserts a vortex in between the sites i and $i+1$. The rotor hamiltonian is

$$\begin{aligned} \mathbf{H}_{\text{rotors}} &= \frac{U}{2} \sum_{\bar{i}} \left(\frac{\Theta_{\bar{i}+1} - \Theta_{\bar{i}}}{2\pi} \right)^2 - J \sum_{\bar{i}} \cos 2\pi \mathbf{m}_{\bar{i}} \\ &\simeq \sum_{\bar{i}} \left(\frac{U}{2} \left(\frac{\Delta\Theta}{2\pi} \right)^2 + \frac{J}{2} (2\pi \mathbf{m}_{\bar{i}})^2 \right) \end{aligned} \tag{100}$$

where the second term should be regarded as a π^2 kinetic energy term. This looks like simple harmonic oscillators again. BUT: we must not forget that $\Theta \in 2\pi\mathbb{Z}$! It's oscillators with discretized positions. We can rewrite it in terms of continuous Θ at the expense of imposing the condition $\Theta \in 2\pi\mathbb{Z}$ energetically by adding a term $-\lambda \cos\Theta$[35]. The resulting model has the action

$$L_{\text{eff}} = \frac{1}{2(2\pi)^2 \rho_s} (\partial_\mu \Theta)^2 - \lambda \cos\Theta.$$

This $\Theta = \phi_L - \phi_R$ is the T-dual variable of our previous discussion; recall the ETCRs

$$[\phi(x), \Theta(y)] = 2\pi \mathbf{i} \mathsf{sign}(x-y). \tag{101}$$

This follows directly from the definition of Θ (99). (101) means that the operator $\cos\Theta(x)$ jumps the SF phase variable θ by 2π – it inserts a 2π vortex, as we designed it to do.

This system has two regimes, depending on the scaling dimension of the vortex insertion operator:

[35] This step seems scary at first sight, since we're adding degrees of freedom to our system, albeit gapped ones. $\Theta_{\bar{i}}$ is the number of bosons to the left of $\bar{i}$ (times 2π). An analogy that I find useful is to the fact that the number of atoms of air in the room is an integer. This constraint can have some important consequences, for example, were they to solidify. But in our coarse-grained description of the fluid phase, we use variables (the continuum number density) where the number of atoms (implicitly) varies continuously. The nice thing about this story (both for vortices and for air) is that the system tells us when we can't ignore this quantization constraint.

• If λ is an irrelevant coupling, we can ignore it in the IR and we get a superfluid, with algebraic LRO.

• If the vortices are relevant, $\lambda \to \infty$ in the IR, and we pin the dual phase, $\Theta_{\bar{i}} = 0, \forall \bar{i}$. This is the Mott insulator, since $\Theta_{\bar{i}} = 0$ means $\mathbf{n}_i = 0$ – the number fluctuations are frozen.

When is λ relevant? Expanding around the free theory,

$$\langle e^{\mathbf{i}\Theta(x)} e^{-\mathbf{i}\Theta(0)} \rangle = \frac{c}{x^{2\pi\rho_s}}$$

this has scaling dimension $\Delta = \pi\rho_s$ which is relevant if $2 > \Delta = \pi\rho_s$. Since the bose correlators behave as $\langle b^\dagger b \rangle \sim x^{-\eta}$ with $\eta = \frac{1}{2\pi\rho_s}$, we see that only if $\eta < \frac{1}{4}$ do we have a stable SF phase. If $\eta > \frac{1}{4}$, the SF is unstable to proliferation of vortices and we end up in the Mott insulator.

[End of Lecture 17]

CFT at the edge of Quantum Hall states. I must mention another important situation where the previous analysis is useful, namely at the interface between a quantum Hall insulator and vacuum. Such an edge hosts a $D = 1 + 1$ CFT which is determined by the (gapped) bulk state. In simple examples (such as integer QHE and Laughlin states) it is a chiral Luttinger liquid.

With hindsight, the nature of this CFT can be determined using very simple reasoning which can be found, *e.g.* in the QFT textbook by Tony Zee. The basic idea is first to argue that the effective field theory of the bulk is a Chern-Simons gauge theory; the edge states of such a theory are then determinable, just as in the final section of Witten's Jones Polynomial paper.

The preceding results can then be used to understand, for example, tunneling of electrons into a quantum Hall sample.

This connection is also the place where the most interesting CFTs may be realized in nature.

4 CFT

I haven't been very explicit about it so far, but we've been studying CFT for a while now. I decided that it would be useful to understand the important examples well before saying anything about the more formal aspects of the structure of CFT and constraints on physics from conformal invariance.

So far we have been studying examples of CFT which have a free-field description (maybe not always an obvious one). When this crutch is absent, what do we do? Remarkably, the answer is not 'nothing'. In particular we can define what we mean by CFT without recourse to perturbation theory, and sometimes we can even *solve* such a thing (I will try to make clear what 'solve' means here).

[$D > 2$: Rychkov. $D = 2$: Ginsparg.] Here is the way to think about CFT non-perturbatively: A CFT is a list of operators with definite scaling dimensions $\{\mathcal{O}_a, \Delta_a\}$ and their OPE structure constants, c_{abc}, appearing in (83). From this information you can compute any correlation function of local operators by successive uses of the OPE. So the Ising model is[36]:

$$\{\underbrace{\mathbb{1}}_{(0,0)}, \underbrace{\chi_L}_{(\frac{1}{2},0)}, \underbrace{\chi_R}_{(0,\frac{1}{2})}, \underbrace{\varepsilon \equiv \mathbf{i}\chi_L\chi_R}_{(\frac{1}{2},\frac{1}{2})}, \underbrace{\sigma}_{(\frac{1}{16},\frac{1}{16})}, \underbrace{\mu}_{(\frac{1}{16},\frac{1}{16})}, \underbrace{T \equiv \chi_L\partial\chi_L}_{(2,0)}, \underbrace{\bar{T} \equiv \chi_R\bar{\partial}\chi_R}_{(0,2)}, \cdots\}$$

and $c_{\chi_L\chi_R\epsilon} = \mathbf{i}, c_{\sigma,\psi,\mu} = 1, ...,$ or more explicitly:

$$\sigma \times \sigma = \mathbb{1} + c_{\sigma\sigma\varepsilon}\varepsilon, \varepsilon \times \varepsilon = \mathbb{1} + \underbrace{c_{\varepsilon\varepsilon\varepsilon}}_{=0}\varepsilon, \sigma \times \varepsilon = c_{\sigma\sigma\varepsilon}\sigma.$$

Notice that I include the identity operator in this list; it has dimension zero for sure. In this way of enumerating the operators the list goes on forever: *e.g.* we can keep appending more derivatives. We'll learn to how do better soon by listing only the *primary operators* at the top of each representation of the conformal group. But in fact, as in effective field theory, one can often get far just knowing about the few lowest-dimension operators[37] A special role is played by the operator $T(z)$ in this list, the *stress tensor*.

4.1 The stress tensor and conformal invariance (abstract CFT)

Let's think a bit about relativistic field theory in the continuum. Any continuum QFT has a stress tensor[38]; there are two useful perspectives on this operator. The simpler, but worse,

[36] Actually, I should leave out the fermions. As this model arises from a spin system, a single fermion operator is not allowed. From the point of view advocated in §2.2.7, it is not gauge invariant.

[37] Naturally, this strategy is generally called 'effective conformal field theory'. A recent victory in this direction can be found in this paper, which solves QCD in $D = 2$ by diagonalizing a 5×5 matrix (!).

[38] Maybe this is not true. The following is a very technical comment which you should ignore if you want. It is certainly true if the field theory has a definition in terms of a lagrangian or a local hamiltonian. Possible

one is to regard it as the Nöther current for spacetime translations. The stress tensor $T_{\mu\nu}$ is conserved if the action is translation-invariant: that is, invariant under the replacement $\phi(x^\mu) \to \tilde{\phi}(\tilde{x}) \equiv \phi(x^\mu + a^\mu)$. This is $d+1$ symmetries, so we have $d+1$ conserved currents: $T^\mu_\nu, \partial_\mu T^\mu_\nu = 0$. The associated charges are the energy and momentum

$$\mathbf{H} = \int_{\text{space}} T_{00}, \quad \mathbf{P}_i = \int_{\text{space}} T_{0i} \ .$$

These generate translations in time and space by commutators (or classically by Poisson brackets):

$$\delta_a \phi(x) = \tilde{\phi}(\tilde{x}) - \phi(x) = a^\mu [\mathbf{P}_\mu, \phi(x)].$$

The finite solution of this equation is

$$\phi(x+a) = e^{\mathbf{i}\mathbf{P}_\mu a^\mu} \phi(x) e^{-\mathbf{i}\mathbf{P}_\mu a^\mu}.$$

The current we get from the Nöther method is not symmetric in its indices.

The better way to think about the stress tensor is as the (linear) response of the system to a local, small perturbation of the metric of spacetime:

$$\delta S = \int \sqrt{g} T^{\mu\nu} \delta g_{\mu\nu} \ , \quad \text{aka} \quad T^{\mu\nu}(x) = \frac{1}{\sqrt{g}} \frac{\delta S}{\delta g_{\mu\nu}(x)}. \tag{102}$$

Here $\sqrt{g} \equiv \sqrt{\det g}$ is the covariant measure.[39] Consider in particular making a *scale transformation*, which we can accomplish by changing the metric by:

$$\delta g_{\mu\nu} = 2\lambda g_{\mu\nu} \quad \to \quad \delta S = \int T^\mu_\mu 2\lambda \tag{103}$$

where λ is a constant. Therefore we see that if $T^\mu_\mu = 0$ then the theory is scale invariant. In the other direction, scale invariance (vanishing of (103)) actually only implies that $T^\mu_\mu = \partial^\mu K_\mu$ is a total derivative. But the object K is a vector operator whose scaling dimension must be $D-1$ (since that of the stress tensor is D). On general grounds, a vector operator of dimension $D-1$ is a conserved current (at least this is the only way I know to protect its dimension). And a conserved current has $\partial^\mu K_\mu = 0$, so that $T^\mu_\mu = 0$ anyway.

If $T^\mu_\mu = 0$ then nothing we said above depended on the fact that λ is a constant, and we should also consider the following transformation:

$$\delta g_{\mu\nu} = 2\delta\Omega(x) g_{\mu\nu} \quad \to \quad \delta S = \int T^\mu_\mu 2\delta\Omega(x) \tag{104}$$

exceptions come from: (1) field theory in a fixed AdS geometry; the absence of dynamical gravity means no stress tensor. (An attempt to take such a thing seriously as a CFT is in this paper by Heemskerk et al.) (2) In his CFT notes, Rychkov discusses a long-range Ising system with a fixed point with no stress tensor; he claims without further discussion that this is the same as the first example. (3) Weird things 'defined' by scaling limits of string theory. Does a 2d CFT without a stress tensor have a central charge?

[39] The $T^{\mu\nu}$ from the metric variation can be related to the Nöther current for translations by "improvement," which means adding boundary terms to the action; this modifies the Nöther current.

Such a transformation, which rescales the whole metric (and therefore preserves angles between vectors at the same point) in a position-dependent way, is a *Weyl transformation*, closely related to a conformal transformation. The difference is that for arbitrary $\Omega(x)$, the new metric will be curved ($\mathcal{R} \propto \nabla^2 \log \Omega$). If we are not interested in QFT in curved spacetime, we should restrict ourselves to choices of Ω which preserve the initial choice of curvature; this means that they can be undone by a coordinate transformation. Infinitesimally, such a transformation is $x^\mu \to (x')^\mu = x^\mu + \xi^\mu(x)$, and the metric change is

$$\delta g_{\mu\nu}(x) = \partial_\mu \xi_\nu(x) + \partial_\nu \xi_\mu(x) \stackrel{!}{=} 2\delta\Omega g_{\mu\nu}. \tag{105}$$

This a set of PDEs for ξ_μ and $\delta\Omega$.

For the case of $g_{\mu\nu} = \eta_{\mu\nu}$, Minkowski spacetime, the stage of special relativity, the constraint (105) is solved by the following formulae[40] (and by translations and rotations and boosts, which don't change the Minkowski metric at all). The conserved currents and charges of the transformations above (in flat spacetime) are:

$$S_\mu = x^\nu T_{\mu\nu} \qquad \to \; D \equiv \int S_0 d^d x = \int x^\mu p_\mu \tag{106}$$

$$C_{\mu\nu} = (2x_\mu x_\lambda - x^2 g_{\mu\lambda}) T^\lambda_\nu \qquad \to \; C_\mu \equiv \int C_{0\mu} d^d x \tag{107}$$

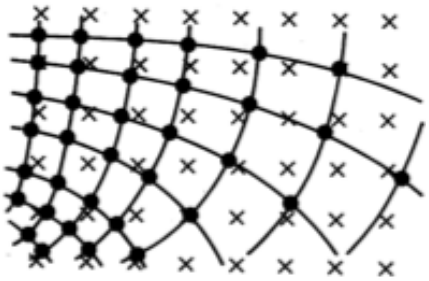
since both $\partial^\mu S_\mu$ and $\partial^\mu C_{\mu\nu}$ are proportional to T^μ_μ. In the last equation of the first line, p_μ is the momentum *density*; the action of the integrand on fields is via $x^\mu \partial_\mu$, a rescaling.

Here is the right way to think about this condition on finite transformations. We are demanding a change of coordinates which accomplishes the following:

$$\eta_{\mu\nu}(dx')^\mu (dx')^\nu \stackrel{!}{=} \Omega^2(x) dx^\mu dx^\nu \eta_{\mu\nu}.$$

The jacobian for this change of coordinates must therefore satisfy

$$J^\mu_\nu \equiv \frac{\partial (x')^\mu}{\partial x^\nu} = \Omega(x) \mathsf{O}^\mu_\nu(x)$$



[Fig credit: Rychkov]

where $\mathsf{O}^\mu_\nu(x)$ is a spacetime-dependent Lorentz transformation ($\mathsf{O}\eta\mathsf{O} = \eta$). This is a spacetime-dependent rescaling and rotation; you should think of it as a position-dependent RG transformation.

We conclude from the above discussion that, at least classically, if $T^\mu_\mu = 0$ (and $T^{\mu\nu}$ is symmetric) then the theory has both scale invariance and conformal invariance. The precise logical relation between scale invariance and conformal invariance is a subject of a lot of discussion. There is no example of a reasonable relativistic fixed point without conformal invariance.

[40] The details are at the beginning of the Ginsparg notes, or on page 3 here.

The resulting set of transformations forms an extension of the Poincare group. There are $d+2$ extra generators $\mathbf{C}_\mu$ and $\mathbf{D}$. In fact the algebra is $\mathsf{so}(d+1,2)$.

Conformal algebra. Here is how to think about this algebra. Most of the commutators just say that vectors (like $\mathbf{C}_\mu$ and $\mathbf{P}_\mu$) transform as vectors and D is a scalar. The important ones are:

$$[\mathbf{D},\mathbf{P}_\mu] = \mathbf{i}\mathbf{P}_\mu, \quad \mathbf{P}_\mu \text{ is a raising operator for } \mathbf{D}$$
$$[\mathbf{D},\mathbf{C}_\mu] = -\mathbf{i}\mathbf{C}_\mu, \quad \mathbf{C}_\mu \text{ is a lowering operator for } \mathbf{D}$$
$$[\mathbf{C}_\mu,\mathbf{P}_\nu] = 2\mathbf{i}\left(\eta_{\mu\nu}\mathbf{D} - \mathbf{M}_{\mu\nu}\right)$$

which says you can recover $\mathbf{D}$ and the spin from $\mathbf{C}_\mu$.

The right way to think about the transformation $\mathbf{C}$ generates is:

$$\text{inversion} \circ \text{translation} \circ \text{inversion:} \qquad \frac{x'^\mu}{\vec{x}'.\vec{x}'} = \frac{x^\mu}{\vec{x}.\vec{x}} + b^\mu \quad . \tag{108}$$

Inversion is $x^\mu \to -\frac{x^\mu}{x^\nu x_\nu}$; inversion plus poincaré implies conformal.

So if we want to study representations of this algebra, we can diagonalize $\mathbf{D}$ and some of the rotation generators. (Notice that the familiar Poincaré casimir $\mathbf{P}^\mu\mathbf{P}_\mu$ is *not* a casimir of the conformal group.) Then we can build representation starting with states that have been lowered as much as possible, so $\mathbf{C}|\text{primary}\rangle = 0$; the rest of the representation is obtained by acting with derivatives, *i.e.* $\mathbf{P}_\mu$ (and $\mathsf{SO}(d)$ rotations).

Fields also form representations of the spacetime symmetry group. A primary field is one which is local enough that its transformation rule under a conformal transformation is the same as if it were just a scale transformation:

$$\Phi(x) \to \Phi'(x') = \Omega(x)^{-\Delta}\Phi(x)$$

for a scalar field, or

$$\Phi(x) \to \Phi'(x') = \Omega(x)^{-\Delta}\mathcal{R}[\mathsf{O}(x)^\mu_\nu]\Phi(x')$$

for a field with spin. $\mathcal{R}$ is the representation matrix for the rotation O in the Φ representation; for a vector field, this is just $\mathsf{O}(x)^\mu_\nu$.

Fields which are not primaries (for example, the derivative of a primary) are called descendants and have more complicated transformation rules. Fields can be organized into eigenstates of $\mathbf{D}$, of definite scaling; every such field is a primary or a descendant. The algebra determines the behavior of descendants from that of primaries.

4.1.1 Geometric interpretation of conformal group

Here is a geometric interpretation of the conformal group, called the *projective null cone* construction (due to Dirac (!)). The conformal group in $\mathbb{R}^{d,1}$ is isomorphic to $\mathsf{SO}(d+1,2)$,

the Lorentz group of a space with two extra dimensions. This space $\mathbb{R}^{d+1,2}$ has metric

$$\eta_{ab} = \text{diag}(-++...+-)_{ab} \tag{109}$$

where the last two dimensions are the 'extra' ones. So we can find linear representations of the conformal group by temporarily adding some extra coordinates. A light ray in this space can be parameterized by $d+1$ dimensional coordinates x^μ in the following way:

$$\zeta^a = \kappa(x^\mu, \frac{1}{2}(1-x^2), \frac{1}{2}(1+x^2))^a \tag{110}$$

where κ is some arbitrary constant. The group $\mathsf{SO}(d+1,2)$ moves these light rays around. We can interpet these transformations as maps on the x^μ. In fact these transformations (combined by rescalings to get back to the original slice) are precisely the conformal transformations.

Regard the expression (110) as an embedding of $\mathbb{R}^{d,1} \subset \mathbb{R}^{d+1,2}$; this is an *isometric embedding*, *i.e.* the induced metric is again $ds^2 = -dt^2 + d\vec{x}^2$. To see that Lorentz boosts in the embedding space are conformal transformations on the null slice, note that on the slice, $\xi^a\xi_a = 0 \implies \xi^a d\xi_a = 0$, and so the transformation $\xi^a \to \lambda(x)\xi^a$ (which preserves a null subspace) takes $d\xi^a d\xi_a \to \lambda(x)^2 d\xi^a d\xi_a$, a conformal transformation.

Invariants in $\mathbb{R}^{d+1,2}$ should therefore be conformal invariants. Consider the object:

$$\zeta_1 \cdot \zeta_2 = \eta_{ab}\zeta_1^a\zeta_1^b = \frac{1}{2}\kappa_1\kappa_2(x_1-x_2)^2. \tag{111}$$

ζ^a and $\lambda\zeta^a$ are identified with the same x^μ, so κ is a redundant variable. So conformal invariants actually are cross ratios of invariants in $\mathbb{R}^{d+1,2}$, for example

$$\frac{\zeta_1 \cdot \zeta_2 \zeta_3 \cdot \zeta_4}{\zeta_1 \cdot \zeta_3 \zeta_2 \cdot \zeta_4}. \tag{112}$$

An extremely useful consequence of this is the statement that $r_{12}^2 \equiv (x_1-x_2)^\mu(x_1-x_2)_\mu$ transforms under a conformal transformation by a rescaling

$$x_{12}^2 \to \frac{x_{12}^2}{\Omega(x_1)\Omega(x_2)}$$

where $\Omega(x_i) \equiv 1+2b\cdot x_i + b^2 x_i^2$ is the rescaling of the metric at x_i under the special conformal transformation $x^\mu \to \frac{x^\mu + b^\mu}{1+2b\cdot x + b^2x^2}$.

4.1.2 Infinite conformal algebra in $D = 2$.

Here there is an important division between $D = 1+1$ and $D > 1+1$. In $D = 2$, in lightcone or holomorphic coordinates, tracelessness of T says

$$T^\mu_\mu \propto T_{z\bar{z}} = 0\ .$$

Conservation of the stress tensor $0 = \partial^\mu T_\mu^\nu$ then says

$$\bar{\partial}_z T_{zz} = 0, \quad \partial_z T_{\bar{z}\bar{z}} = 0.$$

That is $T(z) \equiv T_{zz}$ is holomorphic (and $\tilde{T}(\bar{z}) \equiv T_{\bar{z}\bar{z}}$ is antiholomorphic). In the quantum theory, these statements are true as operator equations; that is: they are exactly true away from other operator insertions in the path integral (the lingo for this is 'up to contact terms'). This holomorphic factorization has the following dramatic consequence: Given an arbitrary holomorphic function[41] $\xi(z)$, the current

$$j_\mu^{(\xi)} = (j_z, j_{\bar{z}})_\mu \equiv (\xi(z)T(z), 0)_\mu$$

is also conserved (!):

$$\partial^\mu j_\mu^{(\xi)} = \bar{\partial}_z j_z - \partial j_{\bar{z}} = 0.$$

This is infinitely many conserved currents! Basically, just from scale invariance.

What are these transformations? Recall that the current $T_{\mu\nu}$ generates translations, which by abuse of notation[42] we can write as $x^\mu \to x^\mu + a^\mu$. Accordingly, the current $T(z)$ generates 'holomorphic translations': $z \to z + \xi$ with ξ constant. So it's not too shocking that $j_\mu^{(\xi)}$ generates the *local* transformation $z \to z + \xi(z)$. The finite version of this transformation is just an arbitrary holomorphic map:

$$z \to z'(z), \quad \bar{z} \to \bar{z}'(\bar{z}).$$

(The important thing here is that z' does not depend on $\bar{z}$!) What does this transformation do to the metric? The flat metric in holomorphic coordinates is

$$ds^2 = d\tau^2 + dx^2 = dz d\bar{z} \to \frac{\partial z}{\partial z'} dz \frac{\partial \bar{z}}{\partial \bar{z}'} d\bar{z} = f(z, \bar{z}) dz d\bar{z}.$$

The metric has only changed by an overall function. This means that the *angle* between any two vectors has not changed. This is the definition of a *conformal transformation*. The conformal group is infinite dimensional in $D = 2$.

For the example of the free massless scalar in $D = 2$ (with curved-worldsheet action $S[\phi] = \frac{1}{2\pi K} \int dx dt \sqrt{g} g^{\mu\nu}(x) \partial_\mu \phi \partial_\nu \phi$) the stress tensor is

$$T_{\mu\nu} = \frac{1}{2\pi K} \left(\partial_\mu \phi \partial_\nu \phi - \frac{1}{2} g_{\mu\nu} (\partial \phi)^2 \right).$$

[41] More precisely, since there is one for z and one for $\bar{z}$, $\xi^z, \xi^{\bar{z}}$ are components of a holomorphic vector field.

[42] This abuse of notation is both very tempting and very confusing. We are not merely relabeling our coordinates; that doesn't do anything – physics is coordinate-independent. We are transforming our fields by

$$\phi \to \phi'(x') = \phi(x) \ .$$

If you promise to keep this in mind, then we can use the less cumbersome expressions below.

Notice that it is traceless: $g^{\mu\nu}T_{\mu\nu} \equiv T^\mu_\mu = 0$. In holomorphic coordinates, $ds^2 = dzd\bar{z}$, the nonzero components are

$$\begin{aligned} T_{zz} &\equiv T(z) = \frac{1}{2\pi K}\partial_z\phi\partial_z\phi = \frac{1}{2\pi K} : \partial_z\phi_L\partial_z\phi_L :, \\ \bar{T}_{\bar{z}\bar{z}} &\equiv T(\bar{z}) = \frac{1}{2\pi K}\bar{\partial}_z\phi\bar{\partial}_z\phi = \frac{1}{2\pi K} : \bar{\partial}_z\phi_R\bar{\partial}_z\phi_R : \end{aligned} \tag{113}$$

In the last step I've emphasized the factorization into L and R parts, and the fact that quantumly we must define this composite operator somehow, and we are doing it by normal ordering. This involves a choice of additive constant, about which there is a bit more to say.

Conformal invariance constrains the operator algebra of a CFT, and since (various moments of) the stress tensor components are generators of conformal transformations, their OPEs are highly constrained.

4.2 Radial quantization

Here is an important example of a conformal transformation: Our coordinate on the *cylinder* was $w = x + \mathbf{i}\tau$; I call it a cylinder because $x \simeq x + L$. Let's set $L = 2\pi$ for convenience. Consider the map

$$w \mapsto z = e^{-\mathbf{i}w} = e^{\tau - \mathbf{i}x}$$

Since it's holomorphic ($z(w, \bar{w})$ doesn't depend on $\bar{w}$), this is a conformal transformation. The image is the complex z-plane. Equal-τ surfaces are circles. The spatial momentum operator was $\partial_\sigma = z\partial_z - \bar{z}\bar{\partial}_z$, which acts by $z \to e^{\mathbf{i}b}z$ – rotations about the origin. The hamiltonian on the cylinder was $\partial_\tau = z\partial_z + \bar{z}\bar{\partial}_z$. So time translations become rescaling about the origin of the z-plane: the dilatation operator (the operator $\mathbf{S}$ that generates scale transformations) is the hamiltonian in radial quantization. Time-ordered correlators on the cylinder are *radially-ordered*. One reason this is nice is that it provides an IR cutoff even when studying infinite-volume physics. (Notice that the operators $z\partial_z \pm \bar{z}\bar{\partial}_z$ commute and so we can label states by their scaling dimension Δ and *spin*.

So far, this all goes through for CFT in any D. In $D > 2$, the equal-radial-time-slices are $D-1$-spheres, and therefore the single quantum number for spin would be replaced by a representation of $SO(D)$.

Example. The mode expansion of the free boson (about the origin of the z plane) is now[43]

$$\partial_z\phi(z) = \sum_{n\in\mathbb{Z}} \boldsymbol{\rho}_n z^{-n-1}$$

(with $\boldsymbol{\rho}_0 \equiv p$), and we can extract them by a contour integral:

$$\boldsymbol{\rho}_n = \oint_{C_0} \frac{dz}{2\pi\mathbf{i}} z^n \partial\phi(z)$$

[43] Allow me to use $\phi_{\text{here}} = \sqrt{\pi T}\phi_{\text{before}}$, so that $\phi_{\text{here}} \simeq \phi_{\text{here}} + 2\pi R$ and the action is $S = -\frac{1}{4\pi}\int d^2z\partial\phi\bar{\partial}\phi$.

where C_0 is a contour encircling $z = 0$ (and no other operators!).

Primaries and quasiprimaries.

Def: A primary operator (or field) $\Phi_{h,\bar{h}}(z,\bar{z})$ of weight $(h,\bar{h})$ transforms under the conformal transformation[44]

$$(z,\bar{z}) \to (f(z),\bar{f}(\bar{z})) \qquad \text{by} \qquad \Phi_{h,\bar{h}}(z,\bar{z}) \to (\partial_z f)^h \left(\bar{\partial}_z \bar{f}\right)^{\bar{h}} \Phi(f(z),\bar{f}(\bar{z})).$$

(The way to remember this is that $\Phi(z)dz^h d\bar{z}^{\bar{h}}$ is a scalar.) For example,

$$\begin{array}{rlll} \text{scaling:} & z \to e^{\lambda} z \implies \Phi \to e^{\lambda\Delta}\Phi, & \Delta = h+\bar{h} & \text{(scaling dimension)} \\ \text{rotations:} & z \to e^{\mathbf{i}\theta} z \implies \Phi \to e^{\mathbf{i}s\theta}\Phi, & s = h-\bar{h} & \text{(spin)} \end{array} \tag{114}$$

The infinitesimal transformation $f(z) = z + \xi(z)$ results in

$$\delta_\xi \Phi(z) = (\xi\partial_z + h\partial_z\xi)\,\Phi + \text{antiholomorphic bits.}$$

This transformation is *generated* by $\int \frac{dz}{2\pi\mathbf{i}}\xi(z)T(z) \equiv \mathbf{L}[\xi]$ (as in 4.1.2) in the sense that $\delta_\xi\Phi = \mathbf{i}[\mathbf{L}[\xi],\Phi]$.

Consider for a while a holomorphic operator, with $\bar{h} = 0$. It has a mode expansion

$$\Phi_h(z) = \sum_{n\in\mathbb{Z}} \Phi_n z^{-n-h}.$$

The shift by h in the moding on the plane comes from the conformal transformation from the cylinder:

$$\Phi_h(w = -\mathbf{i}\ln z) = \sum_n \Phi_n e^{-\mathbf{i}wn}$$

where this is just fourier expansion, and the conformal factor is $\left(\frac{\partial z}{\partial w}\right)^h = z^h$. Note that $n < 0$ is positive energy.

The modes of the stress tensor are called Virasoro operators

$$T(z) = \sum_n \mathbf{L}_n z^{-n-2}, \quad \mathbf{L}_n = \oint_{C_0} \frac{dz}{2\pi\mathbf{i}} z^{n+1} T(z).$$

The definition of primary implies that

$$[\mathbf{L}_0, \Phi_{h,\bar{h}}(0)] = h\Phi_{h,\bar{h}}(0), \qquad [\mathbf{L}_n, \Phi_{h,\bar{h}}(0)] = 0, \ \forall n > 0.$$

[44]Recall that in general dimension $D \geq 2$, the transformation of a primary operator of dimension Δ is

$$\Phi_\Delta(x) \to |\frac{\partial x'}{\partial x}|^\Delta \Phi'_\Delta(x')$$

where $|\frac{\partial x'}{\partial x}|$ is the Jacobian of the conformal map.

This in turn implies that the state

$$|\Phi_{h,\bar{h}}\rangle \equiv \Phi_{h,\bar{h}}(0)|0\rangle$$

is a *highest weight state of the Virasoro operators*, in the sense that

$$\mathbf{L}_0|\Phi_{h,\bar{h}}\rangle = h|\Phi_{h,\bar{h}}\rangle, \quad \mathbf{L}_n|\Phi_{h,\bar{h}}\rangle = 0, \ \forall n > 0.$$

Note that the modes with $n > 0$ *raise* the value of $\mathbf{L}_0$. They include the ordinary special conformal generators $C_\mu = (\mathbf{L}_1, \bar{\mathbf{L}}_1)_\mu$.

In $D = 1 + 1$ it is important to distinguish between Vir primary and ordinary conformal primary, which is just killed by $\mathbf{L}_1$ and not $\mathbf{L}_{n\geq 2}$.

Contours and commutators. You may be bothered by the connection between the algebra in terms of OPEs on the complex plane

$$(\text{like} \quad \partial\phi(z)\partial\phi(w) \sim \frac{1}{(z-w)^2} + \cdots \quad)$$

and the perhaps-more-familiar algebra of mode operators.

$$(\text{like} \quad [\boldsymbol{\rho}_n, \boldsymbol{\rho}_m] = n\delta_{n+m} \quad).$$

The very direct connection between the two comes from radial quantization. Recall that the path integral on the plane produces *radially ordered* correlators:

$$\int [D\phi] \qquad \underbrace{e^{-S[\phi]}...A(z)B(w).....}_{\text{these are numbers, order doesn't matter}} \qquad = \langle T(....\mathbf{A}(z)\mathbf{B}(w)...)\rangle$$

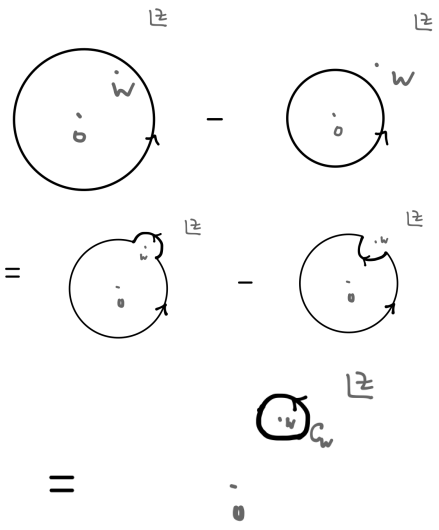
with

$$T(\mathbf{A}(z)\mathbf{B}(w)) = \begin{cases} \mathbf{A}(z)\mathbf{B}(w), & |z| > |w| \\ \mathbf{B}(w)\mathbf{A}(z), & |z| < |w| \end{cases}$$

So consider, for example, the commutator

$$[L[\xi], \Phi_h(w)] = \left(\oint_{|z|>|w|} - \oint_{|z|<|w|}\right) \frac{dz}{2\pi\mathbf{i}}\xi(z)T(z)\Phi_h(w)$$

$$= \oint_{C_w} \frac{dz}{2\pi\mathbf{i}}\xi(z)T(z)\Phi_h(w).$$



This is a general rule: the commutator of modes of two fields is given by the contour integral of one about the other.

The previous expression is of interest since it describes the conformal transformation of the primary Φ_h parameterized by the holomorphic vector field ξ:

$$\mathbf{i}[\mathbf{L}[\xi], \Phi_h(w)] = \delta_\xi \Phi_h(w) = (\xi \partial_z + h \partial_z \xi)\Phi.$$

Compare this expression to the general OPE

$$T(z)\Phi_h(0) = \sum_n \frac{1}{z^{n+1}} \mathcal{O}^{(n)}(0)$$

and we determine

$$\begin{cases} \mathcal{O}^{(1)} = & h\Phi_h \\ \mathcal{O}^{(0)} = & \partial\Phi_h \\ \mathcal{O}^{(\text{else})} = & 0 \end{cases}.$$

That is, we have shown that the OPE of the stress tensor with a *primary operator* of dimension Δ is

$$T(z)\mathcal{O}(0) \sim \frac{\Delta\mathcal{O}}{z^2} + \frac{\partial\mathcal{O}}{z}. \tag{115}$$

Notice that L_0 is scaling, L_{-1} is translations, and L_1 is special conformal.

Some examples to check: (1) $T(z)\partial\phi(0)$ for the free boson theory. (2) $T(z)T(0)$ for the free boson theory (this one is a trick question as we'll see in 4.2.1).

Not all operators of definite scaling dimension behave this way, and this can be taken as a definition of a primary operator. It implies that under a conformal transformation $(z, \bar{z}) \to (w(z), \bar{w}(\bar{z}))$,

$$\mathcal{O}_{h,\bar{h}}(z,\bar{z}) \to \left(\frac{\partial w}{\partial z}\right)^h \left(\frac{\partial \bar{w}}{\partial \bar{z}}\right)^{\bar{h}} \mathcal{O}_{h,\bar{h}}(w(z), \bar{w}(\bar{z}))$$

State-operator correspondence. (any D)

To get a CFT state $|\Phi\rangle$ from a local operator $\Phi(z)$ just insert that operator at the origin acting on the CFT vacuum in radial quantization about the origin. It's easier in equations:

$$\Phi(0)|0\rangle \equiv |\Phi\rangle.$$

The state $|0\rangle$ is then the image of the identity operator under this map.

To get an operator from a state, let's think about the path integral on the ball Σ. Think of it as a functional of the boundary conditions on the fields (which I'll call ϕ; θ is a coordinate on $\partial\Sigma$):

$$\int_{\phi|_{\partial\Sigma}(\theta)=\phi_0(\theta)} [D\phi] e^{-S} = \Psi_0[\phi_0] = \langle\phi_0|0\rangle$$

This is an integral representation of the groundstate wavefunctional. If instead we consider the path integral with a local operator insertion, we get a wavefunctional for a different state:

$$\int_{\phi|_{\partial\Sigma}(\theta)\phi_0(\theta)} [D\phi]e^{-S}\Phi(0) = \Psi_\Phi[\phi_0] = \langle\phi_0|\Phi\rangle \ .$$

In a CFT this is related by a conformal transformation $z = e^{-\mathbf{i}w}$ to the path integral on the cylinder with the state $|\Phi\rangle$ inserted in the far past. So to get the operator corresponding to an arbitrary state just glue this path integral around where you want to put the operator. To move it to a different place, just act with translation generators: $\Phi(x) = e^{-\mathbf{i}\mathbf{P}^\mu x_\mu}\Phi(0)e^{\mathbf{i}\mathbf{P}^\mu x_\mu}$.

For more on the beauty of the state-operator correspondence in D-dimensional CFT see the notes here Rychkov.

Convergence of OPE. With this realization in mind, it is clear that the OPE in CFT is a *convergent* expansion: we are simply inserting a resolution of the identity on the hilbert space in a particular basis of eigenstates of the dilatation operator.

Adjoint in radial quantization. The adjoint is a bit weird: it is just the adjoint on the cylinder, but two things. First, the state in the far future on the cylinder gets mapped to $z = \infty$. Second, we must remember that the map from the cylinder to the sphere produces a Jacobian. Since the adjoint operation involves an *inversion* (a particular element of the conformal group), the primary Φ acquires an extra factor from the Jacobian. This produces an extra factor in the state:

$$|\Phi\rangle^\dagger = \lim_{z\to\infty}\langle\Phi|z^{-h} \ .$$

This is sometimes called the BPZ adjoint after its discoverers.

4.2.1 The Virasoro central charge in 2d CFT

The OPE of the stress tensor with itself is

$$T(z)T(0) \sim \frac{c_L/2}{z^4} + \frac{2T}{z^2} + \frac{\partial T}{z} \tag{116}$$

The quantity c appearing here is called the *Virasoro central charge*. The word 'central' is because it is a c-number, not an operator. It is a crucial piece of data about the CFT. It can be extracted from

$$\langle T(z)T(0)\rangle = \frac{c_L/2}{z^4} \ .$$

There is also a right-moving central charge which appears in the OPE $\bar{T}\bar{T}$.

For free bosons, with $T(z) = \frac{1}{4\pi K}(\partial\phi)^2$ it is equal to $c_L = 1$ (note that the Ks all cancel out). Notice that it is additive: with N free bosons, the answer is $c_L = N$.

For a majorana fermion, with $T(z) = \chi_L \partial \chi_L$, it is equal to $c_L = \frac{1}{2}$. A basic check of bosonization is that a complex fermion has central charge $\frac{1}{2} + \frac{1}{2} = 1$.

It is monotonic under RG flows between CFTs: $c_{UV} > c_{IR}$. Therefore it is a useful measure of the number of degrees of freedom. In fact, using the definition

$$c = \lim_{z \to 0} 2z^4 \langle T(z) T(0) \rangle$$

it can be extended away from RG fixed points; Zamolodchikov proved that this quantity is monotonic along any flow.

Comparing (115) and (116) we see that nonzero c means that the stress tensor itself is not a primary. It transforms weirdly under conformal transformations, in a very definite way. This means, for example, that under the map from the cylinder to the plane, the additive normalization of the stress tensor changes. For an application of this to Casimir energy of CFT on a sphere S^d, see HW 6.

For this reason c can be interpreted as a *conformal anomaly* or *Weyl anomaly*.

$$T^\mu_\mu = \frac{c_L}{4\pi} \mathcal{R}.$$

The relation between this statement and the stress tensor OPE is the same as that between the chiral anomaly in the form $\partial^\mu j^A_\mu = k \frac{F}{2\pi}$ and the anomalous commutator in (97), or better the current-current OPE

$$j(z) j(0) \sim \frac{k}{z^2} \ .$$

In the presence of a source coupling to the current (A for j, or curvature for $T_{\mu\nu}$), the statement about the two-point function implies a statement about the one-point function.

4.3 Back to general dimensions.

The conformal group in $D > 2$ is finite dimensional, but it still produces powerful constraints.

So we can make our list of operators specifying the CFT much shorter by simply enumerating the primaries (and their dimensions, spins, and structure constants). All the structure constants for the descendants are determined from this data by conformal invariance.

4.3.1 Constraints on correlation functions from CFT

(any D) We consider correlators of primaries; to get correlators of descendants just take derivatives of those of their primaries. We'll also focus for simplicity on scalar operators. A

correlation function of N primaries transforms under a conformal transformation as

$$\langle \prod_i \Phi_i'(x_i') \rangle = \prod_i \Omega(x_i)^{-\Delta_i} \langle \prod_i \Phi_i(x_i) \rangle.$$

You should regard the operators Φ' on the LHS as the images of Φ under a spacetime-dependent RG transformation.

Suppose we have in our possession a conformal *invariant* I which depends on N spacetime positions.

- Translations imply that I depends only on differences $x_i - x_j$ (of which there are $D(N-1)$).
- Rotations imply that I depends only on distances $r_{ij} \equiv |x_i - x_j|$ of which there are $\frac{N(N-1)}{2}$.
- Scale invariance implies that I depends only on ratios of differences $\frac{r_{ij}}{r_{kl}}$.
- Special conformal transformations act on these distances by

  $$(r_{12}')^2 = \frac{r_{12}^2}{\Omega_1 \Omega_2}$$

  where

  $$\Omega_i \equiv 1 + 2b \cdot x_i + b^2 x_i^2$$

  is the factor by which the metric at x_i rescales under the associated transformation. (The easiest way to see this is by doing it for an inversion first, and then using the construction of special conformal transformations as $I \circ T \circ I$. Alternatively, this is where the projective null cone comes into its own.) Only *cross-ratios*

  $$\frac{r_{ij} r_{kl}}{r_{ik} r_{jl}}$$

  are invariant under this. There are $\frac{N(N-3)}{2}$ of these (and none at all for $N < 4$).

This discussion of invariants has the following implications.

1. One-point functions of primaries of nonzero dimension vanish if there is a conformally invariant vacuum, $\mathbf{D}|0\rangle = 0$. That this is so can be seen as follows. The infinitesimal scale transformation of a primary is

   $$\delta\Phi(x) = \left(\Omega^{-\Delta}\Phi(x') - \Phi(x)\right)|_{\Omega = e^{-\lambda}, x' = e^{-\lambda}x} = \lambda \left(\Delta + x^\mu \partial_\mu\right) \Phi(x) \ = \mathbf{i}\lambda[\mathbf{D}, \Phi(x)].$$

   The last step is the statement that the dilatation operator $\mathbf{D}$ generates scale transformations. Therefore:

   $$\langle 0|\Phi_\Delta(0)|0\rangle = \frac{\mathbf{i}}{\Delta}\langle 0|[\mathbf{D}, \Phi_\Delta(0)]|0\rangle = 0.$$

2. In CFT, two point functions of primaries are only nonzero if the primaries have the same dimension. This requires conformal invariance; the rest is determined by scale, translation and rotation:
$$\langle 0|\phi_1(z)\phi_2(w)|0\rangle = \frac{\delta_{h_1,h_2}}{r_{12}^{2h_1}}.$$

3. $$\langle \prod_{i=1}^{3} \phi_i(z_i)\rangle = C_{123} r_{12}^{2(-h_1-h_2+h_3)} r_{23}^{2(-h_2-h_3+h_1)} r_{31}^{2(-h_3-h_1+h_2)}$$
C_{123} is the OPE coefficient between the three operators. Notice that this is more than scaling, which just says that the sum of the powers should be $h_1 + h_2 + h_3$.

4. $$\langle \prod_{i=1}^{4} \phi_i(z_i)\rangle = f(x_1, x_2) \prod_{i<j} r_{ij}^{-2(h_i+h_j+\sum_k h_k/3)}$$
where x_{12} are the two cross-ratios. In two dimensions, these are
$$x = \frac{z_{12}z_{34}}{z_{13}z_{24}}$$
and $\bar{x}$.

It seems like $N \geq 4$-point functions are underconstrained. However, we have see that the successive use of the OPE relates them (via *known* functions) to sums of 3-point functions. The unknown functions of cross-ratios are actually determined by this process! The general name for such objects is *conformal blocks*.

4.3.2 Thermodynamics of a CFT

Thermodynamics of scale-invariant theories is very constrained. The partition function is defined as
$$Z_{CFT} = \mathrm{Tr}_{CFT}\left(\exp(-H/T)\right).$$
In the thermodynamic limit, $\ln Z$ is *extensive*, *i.e.* proportional to the volume of the space. But $\ln Z$ is a dimensionless quantity. Hence, we must have $\ln Z \sim VT^d$ (d is the number of spatial dimensions) in the absence of any other energy scales (such as a chemical potential for some conserved charge). The free energy then will be
$$F = -T \ln Z = cVT^{d+1}.$$
where this c should be regarded as a rough measure of the number of degrees of freedom of the CFT. In $d = 2$ it is proportional to the Virasoro central charge. In other cases, this statement is less sharp.

As with any conservation law, $T^\mu_\mu = 0$ is an operator equation in the full quantum theory. What happens if we put it inside $\mathrm{Tr}(e^{-H/T})$, with no other insertions? The operator equation then translates into the following equation

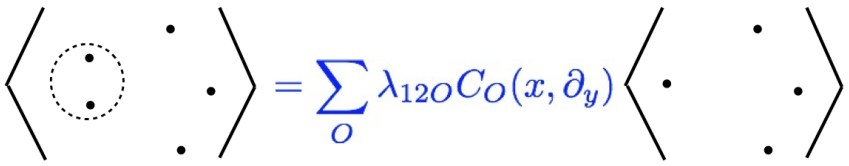
$$0 = \mathrm{Tr}(T^\mu_\mu e^{-H/T}) = \langle T_{00}\rangle - \langle T_{ii}\rangle = \mathcal{E} - dP.$$

This last relation gives the speed of the sound

$$c_s = \sqrt{\left(\frac{\partial P}{\partial \mathcal{E}}\right)_S} = \sqrt{\frac{1}{d}} \tag{117}$$

4.3.3 Too few words about the Conformal Bootstrap

Earlier I said that given a 'solution' to a CFT in the form of a list of primaries and their dimensions and OPE structure constants, you could compute any n-point correlator by repeated use of the OPE to reduce it to a two-point function, by the following kind of operation:



[from Rychkov's CFT notes]

Some interesting questions and previews of the answers:

(1) Which sets of scaling dimensions arise from CFTs which actually exist?

The general answer is not known. One set of constraints follows from unitarity in the form of positivity of the inner product on the CFT Hilbert space. For example, the dimension of a scalar operator (other than $\mathbb{1}$) must be larger than the free field dimension $(\frac{D-2}{2})$[45].

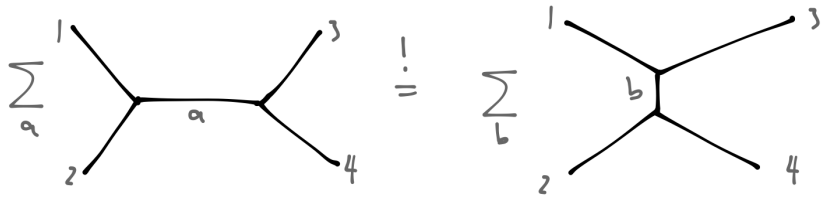
(2) In implementing the reduction to three-point functions and structure constants, you have a choice about the order in which you group the operators. Do you get the same answer independent of the order?

[45] This follows from the fact that $\mathbf{D}$ appears on the RHS of $[\mathbf{C}_\mu, \mathbf{P}_\nu]$. A crucial ingredient comes from the fact that the dagger of a lowering operator is a raising operator (obvious, right?). In radial quantization this implies the weird-looking formula:

$$\mathbf{P}^\dagger_\mu = \mathbf{C}_\mu \ .$$

For more details, I recommend the discussion leading up to eqn (5.57) of Jared Kaplan's AdS/CFT notes.

The condition that you do – the associativity of the OPE – imposes significant constraints on the structure constants and dimensions.



(3) Are there more constraints from higher-order diagrams?

Claim: no. The analog of the 'pentagon identity' is automatic.

To make use of the (many!) associativity constraints, a further ingredient is required, which is called *conformal blocks*. For more in this direction, a good starting point is [Rychkov] and the CFT lectures cited above.

In $D = 2$ it really works [BPZ]. This approach has led to complete solutions of all unitary CFTs with $c < 1$ (this is explained in Ginsparg's notes), and a number of interesting examples (*e.g.* Liouville theory) with $c > 1$. More recently, this program has had success in $2 + 1$ dimensions: the 3d Ising model has been cornered.

[End of Lecture 18]

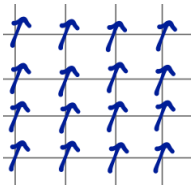
5 $D = 2 + 1$

5.1 Duality for the transverse-field Ising model in $D = 2 + 1$

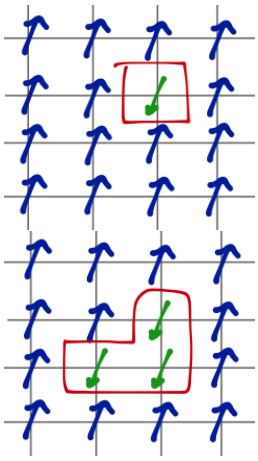
On a lattice of any dimension, with qbits at the sites r ($\mathcal{H} = \otimes_r \mathcal{H}_{\frac{1}{2}}$), the tranvserse-field Ising model Hamiltonian is:

$$\mathbf{H}_{\text{TFIM}} = -J \left(\sum_{\langle rr' \rangle} \mathbf{Z}_r \mathbf{Z}_{r'} + g \sum_r \mathbf{X}_r \right) .$$

The $\mathbb{Z}_2$ symmetry operator is $\mathbf{S} = \prod_r \mathbf{X}_r$. $J > 0$ is a ferromagnetic coupling.



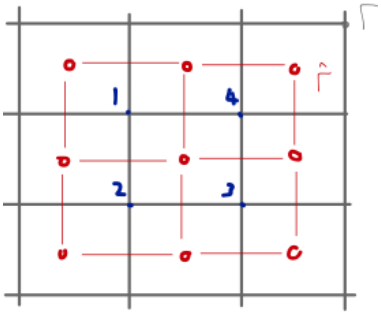
At small g, the groundstates (into which the cat states decohere) spontaneously break the $\mathbb{Z}_2$ symmetry; at $g = 0$, they are simply $|\uparrow\uparrow\uparrow\cdots\rangle$ and $|\downarrow\downarrow\downarrow\cdots\rangle$. There is a gap of order $2J$ to flipping a spin. Small g mixes in states where some spins are flipped.



So far, everything is as in $D = 1 + 1$. The first difference is that in $D > 1$, a domain wall is not a pointlike object. A domain wall has codimension one – its location is specified by one real condition. So in $D = 2 + 1$, it is a string. More precisely, since it is the boundary of a region (of down spins in a sea of up spins, say), it is a closed string: the boundary of a region has no boundary. This is a similar topological statement to the fact that they come in pairs in one dimension.

Clearly at $g = \infty$, the groundstate is a paramagnet, $\prod_r |\rightarrow\rangle$. What happens in between? Can we describe the paramagnet in terms of the proliferation of domain walls as we did in $D = 1 + 1$?

Wegner duality. The theory can be reformulated in terms of domain walls as follows. Domain walls cover links of the dual lattice, $\hat{\Gamma}$, whose sites correspond to faces of the direct lattice Γ. Each links of Γ crosses a link of $\hat{\Gamma}$.



We can rewrite the ferromagnetic interaction term by

$$\boldsymbol{\tau}^x_{\bar{l}} \equiv \mathbf{Z}_r \mathbf{Z}_{r'}$$

where $\bar{l}$ is the link of $\hat{\Gamma}$ crossed by the link rr'. But on a 2d lattice with N sites, there are $2N$ links, and hence $2N$ such $\boldsymbol{\tau}$s – way too many variables! The domain wall variables satisfy the following *local constraints*. For each plaquette (little loop) in the direct lattice, we have:

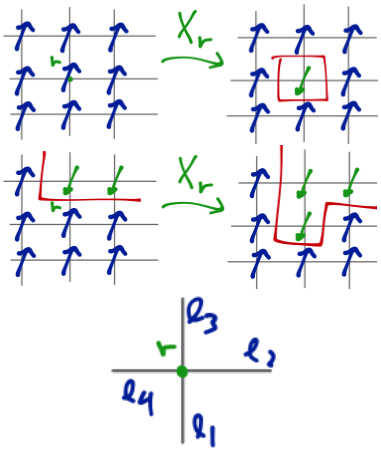
$$1 = (\mathbf{Z}_1 \mathbf{Z}_2)(\mathbf{Z}_2 \mathbf{Z}_3)(\mathbf{Z}_3 \mathbf{Z}_4)(\mathbf{Z}_4 \mathbf{Z}_1)$$

But in terms of the domain wall operators, this plaquette is associated with a site $+$, and we have

$$1 = (\mathbf{Z}_1\mathbf{Z}_2)(\mathbf{Z}_2\mathbf{Z}_3)(\mathbf{Z}_3\mathbf{Z}_4)(\mathbf{Z}_4\mathbf{Z}_1) = \prod_{\bar{l}\in +} \boldsymbol{\tau}^x_{\bar{l}} \equiv \mathbf{A}_+, \forall \text{ sites } + \in \hat{\Gamma} \ .$$

In 1d, there were no loops (except with periodic boundary conditions on the spins).

What about the transverse field term in $\mathbf{H}_{\text{TFIM}}$? $\mathbf{X}_r$ flips a spin at r; on a sea of up spins, this creates a domain wall surrounding the single resulting down spin. If there was already a domain wall there, it annihilates it ($\mathbf{X}$ is its own inverse). More generally it moves the boundaries of the domain walls adjacent to r. It flips the value of the domain wall operators on all bonds surrounding r:



$$\mathbf{X}_r = \prod_{\bar{\ell}\in \boxed{r}} \boldsymbol{\tau}^z_{\bar{l}} \equiv \mathbf{B}_{\boxed{r}}$$

where $\boxed{r}$ denotes the set of links of the dual lattice surrounding r.

An important observation: each term in $\mathbf{H}_{\text{TFIM}}$ commutes with the $\mathbf{A}_+ = 1$ constraint:

•

$$[\mathbf{B}_{\boxed{r}}, \mathbf{A}_+] = 0, \qquad \forall r, +$$

since each $+$ and $\square$ share *two* links ($\mathbf{X}_1\mathbf{X}_2\mathbf{Z}_1\mathbf{Z}_2 = (-1)^2\mathbf{Z}_1\mathbf{Z}_2\mathbf{X}_1\mathbf{X}_2$).

• since $\mathbf{ZZ} = \boldsymbol{\tau}^x$ is made from $\boldsymbol{\tau}^x$s, just like $\mathbf{A}_+$, the ferromagnetic term is unproblematic.

In the domain wall variables, we have

$$\mathbf{H}_{\text{TFIM}} = -J\left(\sum_{\bar{l}\in\hat{\Gamma}} \boldsymbol{\tau}^x_{\bar{l}} + g \sum_{\square \text{ of } \hat{\Gamma}} \prod_{\bar{l}\in\square} \boldsymbol{\tau}^z_{\bar{l}}\right)$$

with the *constraint* that $1 = \mathbf{A}_+ = \prod_{l\in+} \boldsymbol{\tau}^x_l, \forall +$. This is Ising gauge theory. The gauge transformation acts as

$$\boldsymbol{\tau}^z_{\bar{i}\bar{j}} \mapsto s_{\bar{i}}\boldsymbol{\tau}^z_{\bar{i}\bar{j}}s_{\bar{j}}, \quad \boldsymbol{\tau}^x_{\bar{i}\bar{j}} \mapsto \boldsymbol{\tau}^x_{\bar{i}\bar{j}}, \quad s_{\bar{i}} = \pm 1.$$

The transformation with $s_{\bar{i}} = -1$ and all others $+1$ is *generated* by $\mathbf{A}_+(\bar{i})$, in the sense that

$$\mathcal{O} \mapsto \mathbf{A}^\dagger_+(\bar{i})\mathcal{O}\mathbf{A}_+(\bar{i}) \ .$$

And since $\mathbf{A}_+ = 1$ on the Hilbert space, this is a (huge!) *redundancy* of the description. One more time: it does not act on the physical Hilbert space, which is

$$\mathcal{H}_{\text{phys}} = \{|\psi\rangle \mid \mathbf{A}_+(\bar{i})|\psi\rangle = |\psi\rangle, \forall \bar{i}\} \subset \otimes_{\text{links of } \hat{\Gamma}} \mathcal{H}_{\frac{1}{2}} = \mathcal{H}_{\text{big}} \ .$$

Global issues. As in the 1+1d case, there are some global subtleties with the quantum duality transformation: In particular, we must specify the value of a single spin in order for the locations of domain walls to specify the spin configuration. The gauge theory variables do not produce the symmetry-breaking degeneracy.

Notice by our counting argument ($2N$ links between N sites) that this means that some of the constraints $\mathbf{A}_+ = 1$ must also be redundant. An example is $\prod_{\text{all sites of } \hat{\Gamma}} \mathbf{A}_+ = 1$, automatically, since each link is included twice.

This mismatch is a crucial thing: the spin system can have symmetry-breaking order and this gauge theory cannot; as we will see below, the gauge theory can have *topological order* and this spin system cannot.

From now on in this discussion, we will take the gauge theory as 'fundamental' and ignore these issues. I will therefore stop writing the bars on the coordinates of $\hat{\Gamma}$, the lattice which the gauge theory inhabits.

5.2 $\mathbb{Z}_2$ lattice gauge theory

To make it look more like gauge theory familiar from high energy physics, regard the non-gauge-invariant variable $\boldsymbol{\tau}^z$ as

$$\text{“} \qquad \boldsymbol{\tau}^z_{\bar{i}\bar{j}} = e^{\mathbf{i}\int_{\bar{i}}^{\bar{j}} \vec{a}\cdot d\vec{s}} \qquad \text{”}$$

the holonomy of some fictitious continuum gauge field integrated along the link. More precisely, let

$$\boldsymbol{\tau}^z_{\bar{i}\bar{j}} \equiv e^{\mathbf{i}\pi \mathbf{a}_{\bar{i}\bar{j}}}, \quad \mathbf{a}_{\bar{i},\bar{j}} = 0, 1.$$

Then the plaquette operator is

$$\mathbf{B}_\square = \prod_{l\in\square} \boldsymbol{\tau}_l^z \text{“} \;=\; e^{\mathbf{i}\oint_{\partial\square}\vec{a}\cdot d\vec{l}} \text{”} = e^{\mathbf{i}\pi \sum\!\!\!\!\!\!\diagup_{\partial\square}\mathbf{a}} \overset{\text{Stokes}}{=} e^{\mathbf{i}\pi \mathbf{b}_\square} \,,$$

where $\mathbf{b}_\square$ is the (discrete) magnetic flux through the plaquette $\square$. In the penultimate expression, the symbol $\sum\!\!\!\!\!\!\diagup$ is intended to emphasize that we are summing the $\mathbf{a}$s around the closed loop.

In the Hamiltonian description of gauge theory, the field momentum for $\vec{\mathbf{a}}$ is the electric field $\vec{\mathbf{e}}$. So, we call

$$\boldsymbol{\tau}^x_l \equiv e^{\mathbf{i}\pi \mathbf{e}_l}$$

The star operator is

$$\mathbf{A}_+ = \prod_{l\in+} \boldsymbol{\tau}^x_l = e^{\mathbf{i}\pi \sum_{l\in+}\mathbf{e}_{ij}} \equiv e^{\mathbf{i}\pi \Delta\cdot\mathbf{e}}$$

which is a lattice divergence operator. The constraint is

$$1 = \prod_{l \in +} \boldsymbol{\tau}_l^x \quad \leftrightarrow \quad \Delta \cdot \mathbf{e} = 0 \bmod 2.$$

This is *binary electrodynamics*, electrodynamics mod two. Electric charges are violations of the Gauss' Law constraint: if

$$(\Delta \cdot \mathbf{e})\,(i) = 1 \bmod 2$$

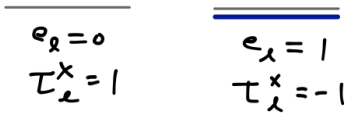
at some site i, we say there is a $\mathbb{Z}_2$ charge at site i. Notice that this is *not* something we can do in the spin system: such a site is the *end* of a domain wall.

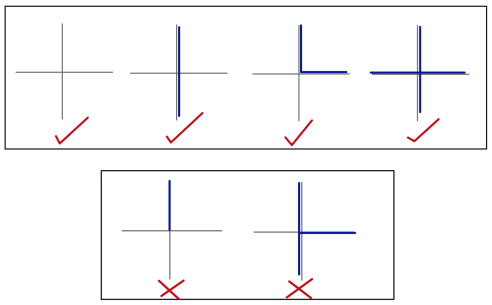
Emergence of gauge theory from a spin system. We can perform the same toric code trick as we did in $D = 1+1$ (in §2.2.7) and realize the Gauss' Law constraint energetically, on a larger Hilbert space, on which the local gauge transformations actually do act. That is: The physics of the gauge theory is the same as the low energy physics ($E \ll E_{\text{giant}}$) of

$$\mathbf{H}_{\text{TC}} = -J \sum_{\tilde{l} \in \hat{\Gamma}} \boldsymbol{\tau}_{\tilde{l}}^x - Jg \sum_{\square \text{ of } \hat{\Gamma}} \mathbf{B}_\square - E_{\text{giant}} \sum_{+ \in \hat{\Gamma}} \mathbf{A}_+ \ .$$

This maneuver (due to Kitaev and Wen) is conceptually quite important, since it means that the gauge theory can emerge from an ordinary spin system, with hilbert space $\mathcal{H}_{\text{big}} = \otimes_{\text{links of } \hat{\Gamma}} \mathcal{H}_{\frac{1}{2}}$.

Which states satisfy the star condition? In the $\boldsymbol{\tau}^x$ basis there is an extremely useful visualization: we say a link l of $\hat{\Gamma}$ is covered with a segment of string (an electric flux line) if $\mathbf{e}_l = 1$ (so $\boldsymbol{\tau}_l^x = -1$) and is not covered if $\mathbf{e}_l = 0$ (so $\boldsymbol{\tau}_l^x = +1$):



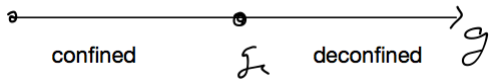


These strings are just the domain walls across which the spins disagree. The star condition is precisely the condition that the strings don't end.

5.3 Phase diagram of $\mathbb{Z}_2$ gauge theory and topological order

Confinement. $\boxed{g \ll 1}$ At small g we must minimize $-\boldsymbol{\tau}_l^x$ for each l, subject to the constraint $\prod_{l \in +} \boldsymbol{\tau}_l^x = 1, \forall +$. This is easy, just set $\boldsymbol{\tau}_l^x = 1$ for every link. In the dual spin variables, this is one of the states with all the spins the same, but we can't specify which one, so there is a unique state of the gauge theory. In terms of the binary electromagnetic fields, this is a

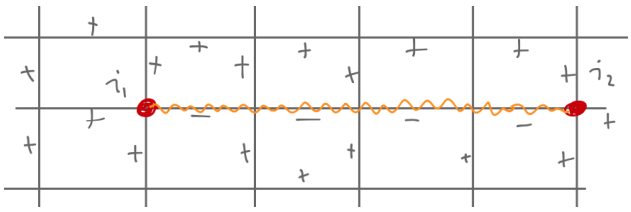
state where $\langle \mathbf{e}_l \rangle = 0, \forall l$: electric flux is costly in energy. (This means that the magnetic field is wildly fluctuating.)



This is the *confined* phase. The origin of this term comes from asking what happens if we insert external test charges, separated by a distance x, and ask what force one exerts on the other via the gauge interactions. Choose two sites $i_{1,2}$ separated by x lattice sites, and demand that $A_+(i_{1,2}) = \prod_{l \in +(i_{1,2})} \boldsymbol{\tau}_l^x = -1$, while $A_+(i) = 1$ for all other sites i. The latter condition implies that the links covered with electric flux ($\mathbf{e}_l = 1, \boldsymbol{\tau}_l^x = -1$) form a *string*, which only terminates at the external charges. This is the familiar statement that lines of electric flux terminate and originate at electric charges.

But now to find the potential between static charges $V(x)$, we need to minimize

$$\mathbf{H}(g=0) = -J \sum_l \boldsymbol{\tau}_l^x = E_0 + 2J \mathsf{L}(\text{string}) \ .$$



Here $E_0 = -J2N$ is the energy of the state with no electric flux and no external charges, where $2N$ is the number of sites. $\mathsf{L}(\text{string})$ is the length of the electric flux string: the string can be said to have a nonzero tension (energy per unit length), $2J$. Clearly this minimization is accomplished by a straight line, and the potential between the charges is

$$V(x) = +2Jx$$

which is linearly rising with separation, and implies a constant attractive force

$$F = -\partial_x V = 2J.$$

Deconfinement. $\boxed{g \gg 1}$ At large g, we minimize the magnetic flux term

$$-gJ\mathbf{B}_\square = -gJ \prod_{l \in \square} \boldsymbol{\tau}_l^z$$

subject to the constraint $\mathbf{A}_+ = 1, \forall +$. The groundstate has $\mathbf{B}_\square = 1, \forall \square$: the magnetic flux is zero (mod 2). (Note that this satisfies the constraint, since $\mathbf{B}$ and $\mathbf{A}$ commute and can therefore be simulataneously diagonalized.)

The excitations arise by flipping $\mathbf{B}_\square = -1$ for some $\square$; this costs energy $2Jg$.

This is definitely a different phase from the $g \ll 1$ phase above, as we will see now, but there is not *local* order parameter (note that here the global disagreement in the duality transformation is important) which distinguishes them (this was Wegner's original motivation for studying this system).

Definite $\mathbf{b}$ means wildly fluctuating $\mathbf{e}$. This is one reason to call this state deconfined: electric flux is cheap. It will be useful to write the groundstate explicitly in the electric

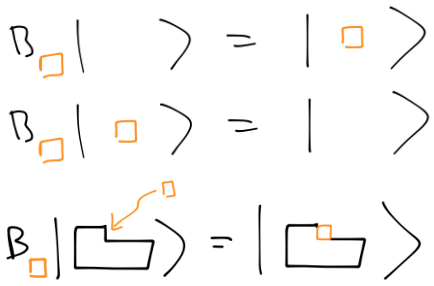
variables. The states of the electric variables satisfying the $\mathbf{A}_+ = 1$ constraint are described by configurations of closed unoriented strings C on the lattice, just as before. Explicitly: a link with $\mathbf{e} = 1$ is covered with a string segment.

We can write the groundstate in this basis:

$$|\text{gs}\rangle = \sum_C \Psi(C)|C\rangle$$

and we will now determine the groundstate wavefunction $\Psi(C)$. The eigenvalue equation for $\mathbf{B}_\square$ is then

$$|\text{gs}\rangle \stackrel{!}{=} \mathbf{B}_\square|\text{gs}\rangle = \sum_C \Psi(C)\mathbf{B}_\square|C\rangle \ .$$



So: how does $\mathbf{B}_\square$ act on a closed string configuration? This is just the discussion we had above when we determined (in §5.1) that $\mathbf{X}_r = \mathbf{B}_{\boxed{r}}$ – the domain walls *are* the loops of electric flux, and they are moved around by the magnetic field operator:

$$\mathbf{B}_\square|C\rangle = |\mathcal{C} + \square\rangle$$

$\mathbf{B}_\square$ *creates* a loop of string around the plaquette $\square$. If one is there already it annihilates it. If a string has a segment covering $\square$, $\mathbf{B}_\square$ moves the string around $\square$.

So the eigenvalue equation $\mathbf{B}_\square = 1$ says $\Psi(C) = \Psi(C')$ if C' and C can be continuously deformed into each other by attaching or removing plaquettes. If the lattice is simply connected – if all curves are the boundary of some region contained in the lattice – then this means

$$|\text{gs}\rangle = \sum_C |C\rangle$$

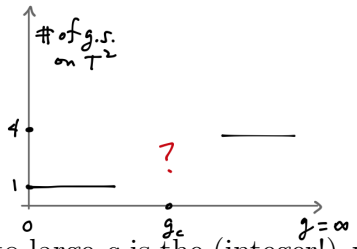
is a uniform superposition of all loops.

String condensation. Notice that the deconfined phase of the gauge theory involves the condensation of the electric flux strings, in the sense that the operators $\mathbf{B}_\square$ which create these strings have a nonzero groundstate expectation value:

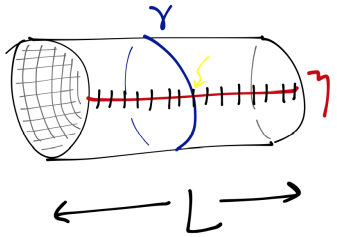
$$\langle\text{gs}|\mathbf{B}_\square|\text{gs}\rangle \stackrel{g=\infty}{=} 1.$$

As with an ordinary condensate of bosons, away from the zero-correlation-length limit ($g = \infty$), the condensate will not be exactly 1, since finite g suppresses configurations with electric flux. But within the deconfined phase it will be nonzero.

Topological order. If the space has non-contractible loops, then the eigenvalue equation does not determine the relative coefficients of loops of different topology! On a space with $2g$ independent non-contractible loops, there are 2^{2g} independent groundstates.



This is (a paradigmatic example of) *topological order*. No local operator can tell the different groundstates apart! A (nonlocal!) order parameter for the phase transition from small g to large g is the (integer!) number of groundstates on a space with nontrivial topology. Two gapped states where this number is different certainly cannot be adiabatically connected, since an integer cannot vary continuously.



Consider the cylinder. There is one nontrivial class of loops, call a representative γ. Let η be a line running along the cylinder. The two groundstates are generated by the action of the 'Wilson loop operator'

$$\mathbf{W}_x(\eta) \equiv \prod_{l \text{ crossed by } \eta} \boldsymbol{\tau}_l^x$$

in the sense that [46]

$$|\mathrm{gs}_2\rangle = \mathbf{W}_x(\eta)|\mathrm{gs}_1\rangle \ .$$

This is also a groundstate (at $g=\infty$) since there is no plaquette which violates $\mathbf{B}_p$ (more simply: $[\mathbf{H}_{g=\infty}, \mathbf{W}_x(\eta)] = 0$). They are *distinguished* by $\mathbf{W}_z(\gamma) \equiv \prod_{l\in\gamma} \boldsymbol{\tau}_l^z$ in the sense that

$$\mathbf{W}_z(\gamma)|\mathrm{gs}_\alpha\rangle = (-1)^\alpha |\mathrm{gs}_\alpha\rangle, \quad \alpha = 1,2.$$

This follows since $\{\mathbf{W}_x(\eta), \mathbf{W}_z(\gamma)\} = 0$ – they share a single link (the one pointed to by the yellow arrow in the figure).

At finite g (and in finite volume), there is tunneling between the topologically degenerate groundstates, since then

$$[\mathbf{H}, \prod_{l\in\gamma} \boldsymbol{\tau}_l^z] \neq 0.$$

This means

$$\langle \mathrm{gs}_2|\mathbf{H}|\mathrm{gs}_1\rangle \equiv \Gamma \neq 0.$$

However, the amplitude Γ requires the creation of magnetic flux on some plaquette (*i.e.* a plaquette P with $B_P = -1$, which costs energy $2Jg$), which then must hop (using the $-J\boldsymbol{\tau}^x$ term in $\mathbf{H}$) all the way along the path η, of length L, to cancel the action of $\mathbf{W}_x(\eta)$. The amplitude for this process goes like

$$\Gamma \sim \frac{\langle \mathrm{gs}_2| \left(-J\boldsymbol{\tau}_1^x\right)\left(-J\boldsymbol{\tau}_2^x\right)\cdots\left(-J\boldsymbol{\tau}_L^x\right)|\mathrm{gs}_1\rangle}{2gJ\cdot 2gJ\cdot \ldots 2gJ} \sim \left(\frac{J}{2gJ}\right)^L = e^{-L\log 2g}$$

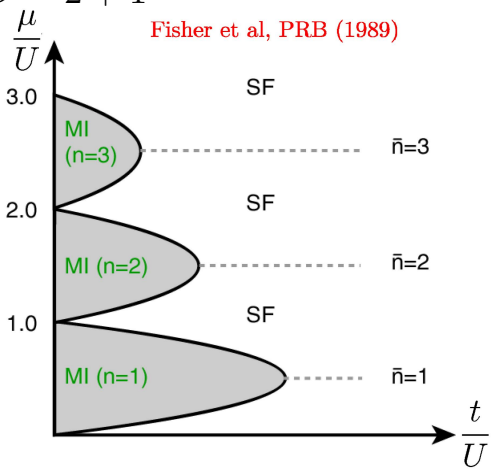
[46] Notice that in the literal dual of the spin system,

$$\prod_{l \text{ crossed by } \eta} \boldsymbol{\tau}_l^x \overset{\text{spins}}{=} \left(\mathbf{Z}_1\mathbf{Z}_2\right)\left(\mathbf{Z}_2\mathbf{Z}_3\right)\cdots\left(\mathbf{Z}_{N-1}\mathbf{Z}_N\right) = \mathbf{Z}_1\mathbf{Z}_N$$

this operator is trivial. (If we identify the two ends of the cylinder to make a torus, $\mathbf{Z}_1 = \mathbf{Z}_N$, it is just $\mathbb{1}$.) Here is where the global mismatch between the spin system and the 'dual' gauge theory is crucial. The TFIM does not have topological order. However, the toric code construction is one demonstration that it is possible for an ordinary spin system to exhibit this phenomenon.

which is *extremely tiny* in the thermodynamic limit. The way to think about this is that the Hamiltonian is itself a local operator, and cannot distinguish the groundstates from each other. It takes a non-perturbative process, exponentially suppressed in system size, to create the splitting.

5.4 Superfluid and Mott Insulator in $D = 2 + 1$



(Fig credit: Roman Lutchyn)

Mean field theory gives the famous picture at right, with lobes of different Mott insulator states with different (integer!) numbers of bosons per site. We have seen above (in §3.1.4) that the prevention of vortices is essential to superfluidity, which is the condensation of bosons. In $D = 1 + 1$, vortices are events in spacetime. In $D = 2+1$, vortices are actual particles, *i.e.* localizable objects, around which the superfluid phase variable winds by 2π (times an integer).

More explicitly, if the boson field which condenses is $b(x) = ve^{\mathbf{i}\phi}$, and we choose polar coordinates in space $x + \mathbf{i}y \equiv Re^{\mathbf{i}\varphi}$, then a vortex is a configuration of the order parameter field of the form $b(x) = f(R)e^{\mathbf{i}\varphi}$, where $f(R) \overset{R\to\infty}{\to} v$ far away: the phase of the order parameter winds around. Notice that the phase is ill-defined in the core of the vortex where $f(R) \overset{R\to 0}{\to} 0$.

To see the role of vortices in destroying superfluidity more clearly, consider superfluid flow in a 2d annulus geometry, with the same polar coordinates $x + \mathbf{i}y = Re^{\mathbf{i}\varphi}$. If the superfluid phase variable is in the configuration $\phi(R, \varphi) = n\varphi$, then the current is

$$\vec{J}(R, \varphi) = \rho_s \vec{\nabla}\phi = \check{\varphi}\rho_s \frac{n}{2\pi R}.$$

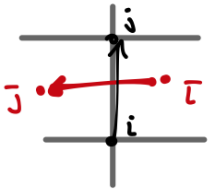
The current only changes if the integer n changes. This happens if vortices enter from the outside; removing the current (changing n to zero) requires n vortices to tunnel all the way through the sample, which if they are gapped and the sample is macroscopic can take a cosmologically long time.

There is a dual statement to the preceding two paragraphs: a state where the bosons themselves are gapped and localized – that is, a mott insulator – can be described by the condensation of vortices. To see this, let us introduce dual variables to the $2 + 1$d rotor model

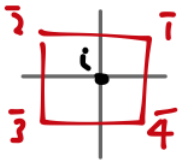
$$\mathbf{H}_{\text{rotors}} = U \sum_i \mathbf{n}_i^2 - J \sum_{\langle ij \rangle} \cos(\phi_i - \phi_j) \ .$$

Introduce a dual lattice whose sites are (centered in) the faces of the original (direct) lattice; each link of the dual lattice crosses one link of the direct lattice.

• First let $e_{\bar{i}\bar{j}} \equiv \frac{\phi_i - \phi_j}{2\pi}$. Here we define $\bar{i}\bar{j}$ by the right hand rule: $ij \times \bar{i}\bar{j} = +\check{z}$ (ij denotes the unit vector pointing from i to j). This is a lattice version of $\vec{e} = \check{z} \times \vec{\nabla}\phi \frac{1}{2\pi}$. Defining lattice derivatives $\Delta_x \phi_i \equiv \phi_i - \phi_{i+\check{x}}$, the definition is $e_x = \frac{\Delta_x \phi}{2\pi}, e_y = \frac{\Delta_y \phi}{2\pi}$. It is like an electric field vector.



• The conjugate variable to the electric field is $a_{\bar{i}\bar{j}}$, which must therefore be made from the conjugate variable of ϕ_i, namely $\mathbf{n}_i$: $[\mathbf{n}_i, \phi_j] = -\mathbf{i}\delta_{ij}$. Acting with $\mathbf{n}_i$ translates ϕ_i, which means that it shifts all the $e_{\bar{i}\bar{j}}$ from the surrounding plaquettes. More precisely:



$$2\pi \mathbf{n}_i = a_{\bar{1}\bar{2}} + a_{\bar{2}\bar{3}} + a_{\bar{3}\bar{4}} + a_{\bar{4}\bar{1}}.$$

This is a lattice, integer version of $n \sim \frac{1}{2\pi}\vec{\nabla} \times a \cdot \check{z}$. In terms of these variables,

$$\mathbf{H}_{\text{rotors}} = \frac{U}{2}\sum_i \left(\frac{\Delta \times a}{2\pi}\right)^2 - J\sum_{\langle \bar{i}\bar{j}\rangle} \cos\left(2\pi e_{\bar{i}\bar{j}}\right)$$

with the following *constraint*. If it were really true that $\vec{e} = \frac{1}{2\pi}\check{z} \times \vec{\nabla}\phi$, with single-valued ϕ, then $\vec{\nabla} \cdot \vec{e} = \vec{\nabla} \cdot \left(\check{z} \times \vec{\nabla}\phi\right) = 0$. But there are vortices in the world, where ϕ is not single valued. The number of vortices $n_v(R)$ in some region R with $\partial R = C$ is determined by the winding number of the phase around C:

$$2\pi n_v(R) = \oint_C d\vec{\ell} \cdot \vec{\nabla}\phi \overset{\text{Stokes}}{=} 2\pi \int_R d^2x \vec{\nabla} \cdot \vec{e}$$

(More explicitly, $2\pi\vec{\nabla} \cdot \vec{e} = \epsilon_{zij}\partial_i\partial_j\phi = [\partial_x, \partial_y]\phi$ clearly vanishes if ϕ is single-valued.) Since this is true for any region R, we have

$$\vec{\nabla} \cdot \vec{e} = 2\pi\delta^2(\text{vortices}).$$

Actually, the lattice version of the equation has more information (and is true) because it keeps track of the fact that the number of vortices is an integer:

$$\Delta_x e_x + \Delta_y e_y \equiv \vec{\Delta} \cdot \vec{e}(\bar{i}) = 2\pi n_v(\bar{i}), \qquad n_v(\bar{i}) \in \mathbb{Z}.$$

It will not escape your notice that this is Gauss' law.

Phases of the 2d rotors. Since $\vec{e} \sim \vec{\nabla}\phi$ varies continuously, *i.e.* electric flux is not quantized, this is called *noncompact* electrodynamics. Again we will impose the integer constraint $a \in 2\pi\mathbb{Z}$ energetically, *i.e.* let $a \in \mathbb{R}$ and add (something like) $\Delta\mathbf{H} \overset{?}{=} -t\cos a$ and see what happens when we make t finite. The expression in the previous sentence is not quite right, yet, however: This operator does not commute with our constraint $\vec{\Delta} \cdot \vec{e} - 2\pi n_v = 0$ – it jumps $\vec{e}$ but not n_v[47]

[47] A set of words which has the same meaning as the above: $\cos a$ is not gauge invariant. Understanding these words requires us to think of the operator $G(\bar{i}) \equiv \vec{\Delta} \cdot \vec{e} - 2\pi n_v$ as the generator of a transformation,

$$\delta\mathcal{O} = \sum_{\bar{i}} s(\bar{i})[G(\bar{i}), \mathcal{O}].$$

It can be a useful picture.

We can fix this by introducing explicitly the variable which creates vortices:

$$[n_v(\bar{i}), \chi(\bar{j})] = -\mathbf{i}\delta_{\bar{i}\bar{j}} \ .$$

Since $n_v \in \mathbb{Z}$, $\chi \simeq \chi + 2\pi$ lives on a circle. So:

$$\mathbf{H} \sim \sum_{\bar{i}} \left(\frac{U}{2} \left(\frac{\Delta \times a}{2\pi} \right)^2 + \frac{J}{2} (2\pi e)^2 - t \cos(\Delta\chi - a) \right)$$

still subject to the constraint $\vec{\Delta} \cdot \vec{e} = 2\pi n_v$.

Two regimes:
$\boxed{J \gg U, t}$: This suppresses e and its fluctuations, which means a fluctuates. The fluctuating a is governed by the gaussian hamiltonian

$$\mathbf{H} \sim \sum \left(\vec{e}^2 + \vec{b}^2 \right)$$

with $b \equiv \frac{\Delta \times a}{2\pi}$, which should look familiar. This *deconfined* phase has a gapless photon; a $2+1$d photon has a single polarization state. This is the goldstone mode, and this regime describes the superfluid phase (note that the parameters work out right in the original variables). The relation between the photon a and the original phase variable, in the continuum is

$$\epsilon_{\mu\nu\rho} \partial_\nu a_\rho = \partial_\mu \phi.$$

$\boxed{t \gg U, J}$: In this regime we must satisfy the cosine first. Like in $D = 1+1$, this can be described as the statement that vortices condense. Expanding around its minimum, the cosine term is

$$\mathbf{h} \ni t (a - \partial\chi)^2$$

– the photon gets a mass by eating the phase variable χ. There is an energy gap. This is the Mott phase.

If the vortices carry other quantum numbers, the Mott phase can be more interesting.

Compact electrodynamics in $D = 2+1$. Note that this free photon phase of $D = 2+1$ electrodynamics is not accessible if e is quantized (so-called compact electrodynamics) where monopole instantons proliferate and gap out the photon. This is the subject of §5.5.

5.5 Compact electrodynamics in $D = 2 + 1$

Consider a quantum system on a two-dimensional lattice (say, square) with rotors $\Theta_l \equiv \Theta_l + 2\pi m$ on the links l. (Think of this as the phase of a boson or the direction of an

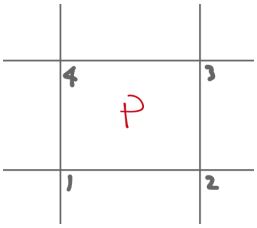
easy-plane spin.) The conjugate variable $\mathbf{n}_l$ is an integer

$$[\mathbf{n}_l, \Theta_{l'}] = -\mathbf{i}\delta_{l,l'}.$$

Here $\mathbf{n}_{ij} = \mathbf{n}_{ji}, \Theta_{ij} = \Theta_{ji}$ – we have not oriented our links (yet). We also impose the *Gauss' law constraint*

$$\mathbf{G}_s \equiv \sum_{l \in v(s)} \mathbf{n}_l = 0 \quad \forall \text{ sites } s,$$

where the notation $v(s)$ means the set of links incident upon the site s ('v' is for 'vicinity'). We'll demand that the Hamiltonian is 'gauge invariant', that is, that $[\mathbf{H}, \mathbf{G}_s] = 0 \forall s$. Any terms which depend only on $\mathbf{n}$ are OK. The natural single-valued object made from Θ is $e^{\mathbf{i}\Theta_l}$, but this is not gauge invariant. A combination which is gauge invariant is the *plaquette operator*, associated to a face p of the lattice:

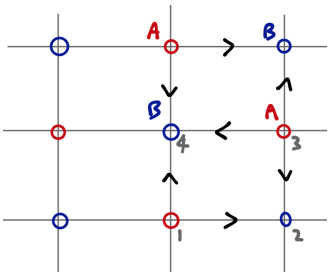


$$\prod_{l \in \partial p} e^{(-1)^y \mathbf{i}\Theta_l} \equiv e^{\mathbf{i}(\Theta_{12} - \Theta_{23} + \Theta_{34} - \Theta_{41})}$$

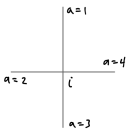
– we put a minus sign on the horizontal links. ∂p denotes the links running around the boundary of p. So a good hamiltonian is

$$\mathbf{H} = \frac{U}{2} \sum_l \mathbf{n}_l^2 - K \sum_\square \cos \left(\sum_{l \in \partial \square} (-1)^y \Theta_l \right).$$

A useful change of variables gets rid of these annoying signs. Assume the lattice is bipartite: made of two sublattices A, B each of which only touches the other. Then draw arrows from A sites to B sites, and let



$$\begin{array}{rcl} \mathbf{e}_{ij} & \equiv & \eta_i \mathbf{n}_{ij} \\ \mathbf{a}_{ij} & \equiv & \eta_i \Theta_{ij}, \end{array} \quad \eta_i \equiv \begin{cases} +1, & i \in A \\ -1, & i \in B \end{cases}.$$



Then the Gauss constraint now reads

$$0 = \mathbf{e}_{\bar{i}\bar{1}} + \mathbf{e}_{\bar{i}\bar{2}} + \mathbf{e}_{\bar{i}\bar{3}} + \mathbf{e}_{\bar{i}\bar{4}} \equiv \Delta \cdot \mathbf{e}(\bar{i}).$$

This is the lattice divergence operation. The plaquette term reads

$$\cos\left(\Theta_{12} - \Theta_{23} + \Theta_{34} - \Theta_{41}\right) = \cos\left(\mathbf{a}_{12} + \mathbf{a}_{23} + \mathbf{a}_{34} + \mathbf{a}_{41}\right) \equiv \cos\left(\Delta \times \mathbf{a}\right)$$

– the lattice curl (more precisely, it is $(\Delta \times \mathbf{a}) \cdot \check{z}$). In these variables,

$$\mathbf{H} = \frac{U}{2} \sum_l \mathbf{e}_l^2 - K \sum_\square \cos\left((\Delta \times \mathbf{a}) \cdot \check{n}_\square\right)$$

(in the last term we emphasize that this works in $D \geq 2 + 1$ if we remember to take the component of the curl normal to the face in question). This is (compact) lattice $\mathsf{U}(1)$ gauge

theory, with no charges. The word 'compact' refers to the fact that the charge is quantized; the way we would add charge is by modifying the Gauss' law to

$$\underbrace{\Delta \cdot \mathbf{e}(\bar{i})}_{\in \mathbb{Z}} = \underbrace{\text{charge at } \bar{i}}_{\Longrightarrow \in \mathbb{Z}}$$

where the charge must be quantized because the LHS is an integer. (In the noncompact electrodynamics we found dual to the superfluid, it was the continuous angle variable which participated in the Gauss' law, and the discrete variable which was gauge variant.)

What is it that's compact in compact QED?

The operator appearing in Gauss' law

$$\mathbf{G}(x) \equiv \left(\vec{\nabla} \cdot \vec{\mathbf{e}}(x) - 4\pi \mathbf{n}(x) \right)$$

(here $\mathbf{n}(x)$ is the density of charge) is the generator of gauge transformations, in the sense that a gauge transformation acts on any operator $\mathcal{O}$ by

$$\mathcal{O} \mapsto e^{-\mathbf{i}\sum_x \alpha(x)\mathbf{G}(x)} \mathcal{O} e^{\mathbf{i}\sum_x \alpha(x)\mathbf{G}(x)} \tag{118}$$

This is a fact we've seen repeatedly above, and it is familiar from ordinary QED, where using the canonical commutation relations

$$[\mathbf{a}^i(x), \mathbf{e}^j(y)] = -\mathbf{i}\delta^{ij}\delta(x-y), \quad [\phi(x), \mathbf{n}(y)] = -\mathbf{i}\delta(x-y)$$

(ϕ is the phase of a charged field, $\Phi = \rho e^{\mathbf{i}\phi}$) in (118) reproduce the familiar gauge transformations

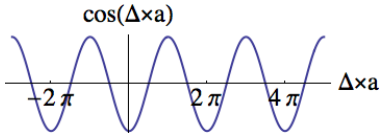
$$\vec{\mathbf{a}} \to \vec{\mathbf{a}} + \vec{\nabla}\alpha, \qquad \phi \to \phi + \alpha \quad .$$

SO: if all the objects appearing in Gauss' law are integers (which is the case if charge is quantized and electric flux is quantized), it means that the gauge parameter α itself only enters mod 2π, which means the gauge transformations live in $\mathsf{U}(1)$, as opposed to $\mathbb{R}$. So it's the gauge group that's compact.

This distinction is very important, because (in the absence of matter) this model does not have a deconfined phase! To see this result (due to Polyakov), first consider strong coupling:

$\boxed{U \gg K}$: The groundstate has $\mathbf{e}_{\bar{l}} = 0$, $\forall \bar{l}$. (Notice that this configuration satisfies the constraint.) There is a gap to excitations where some link has an *integer* $\mathbf{e} \neq 0$, of order U. (If $\mathbf{e}$ were continuous, there would not be a gap!) In this phase, electric flux is confined, *i.e.* costs energy and is generally unwanted.

$\boxed{U \gg K}$: The surprising thing is what happens when we make the gauge coupling weak.



Then we should first minimize the magnetic flux term: minimizing $-\cos(\Delta \times \mathbf{a})$ means $\Delta \times \mathbf{a} \in 2\pi\mathbb{Z}$. Near each minimum, the physics looks like Maxwell, $\mathbf{h} \sim \mathbf{e}^2 + \mathbf{b}^2 + \cdots$. BUT: it turns out to be a colossally bad idea to ignore the tunnelling *between* the minima. To see this, begin by solving the Gauss law constraint $\Delta \cdot \mathbf{e} = 0$ by introducing

$$\mathbf{e}_{\bar{1}\bar{2}} \equiv \frac{1}{2\pi}\left(\phi_2 - \phi_1\right)$$

(*i.e.* $\vec{\mathbf{e}} = \check{z} \cdot \Delta\phi \frac{1}{2\pi}$.) ϕ is a (discrete!) 'height variable'. Then the operator

$$e^{\mathbf{i}(\Delta\times\mathbf{a})(\bar{i})}$$

increases the value of $\mathbf{e}_{\bar{i}a}$ for all neighbors a, which means it jumps $\phi \to \phi + 2\pi$. So we should regard

$$(\Delta \times \mathbf{a})\,(\bar{i}) = \Pi_\phi(\bar{i})$$

in the sense that

$$[\Pi_\phi(r), \phi(r')] = -\mathbf{i}\delta_{rr'}.$$

The hamiltonian is now

$$\mathbf{H} = \frac{U}{2}\sum_l \left(\Delta\phi\right)^2 - K\sum_r \cos\Pi_\phi(r)$$

with no constraint. Deconfinement limit should be $K \gg U$, in which case it looks like we can Taylor expand the cosine $\cos\Pi_\phi \sim 1 - \frac{1}{2}\Pi_\phi^2$ and get harmonic oscillators. But: tunneling between the neighboring vacua of $\Delta \times \mathbf{a}$ is accomplished by the *flux-insertion operator* (or *monopole operator*)

$$e^{\mathbf{i}\phi}, \quad [e^{\mathbf{i}\phi(r)}, (\Delta \times \mathbf{a})\,(r')] = e^{\mathbf{i}\phi(r)}\delta_{rr'}$$

–that is, $e^{\mathbf{i}\phi}$ is a raising operator for $\Delta \times \mathbf{a}$. To analyze whether the Maxwell limit survives this, let's go to the continuum and study perturbations of the free hamiltonian

$$\mathbf{H}_0 = \int \left(\frac{U}{2}\left(\vec{\nabla}\phi\right)^2 + \frac{K}{2}\Pi_\phi^2\right)$$

by

$$\mathbf{H}_1 = -\int V_0 \cos\phi\ .$$

Notice that again we can think of the addition of this term as energetically imposing the condition that $\phi \in 2\pi\mathbb{Z}$.

So: is V_0 irrelevant? Very much no. In fact

$$\langle \cos\phi(r)\cos\phi(0)\rangle_0 \sim \text{const} \tag{119}$$

has constant amplitude at large r! That means that the operator has dimension zero, and the perturbation in the action has $[S_1 = -\int V_0 \cos\phi d^2x d\tau] \sim L^3$, very relevant. The result

is that it pins the ϕ field (the would-be photon mode) to an integer, from which it can't escape.

To see (119) begin with the gaussian identity

$$\langle e^{\mathbf{i}s\phi(x)} e^{\mathbf{i}s'\phi(0)} \rangle = e^{-\frac{ss'}{2}\langle\phi(x)\phi(0)\rangle} ,$$

with $s, s' = \pm$. The required object is

$$\begin{aligned}
\langle\phi(x)\phi(0)\rangle &= \frac{\mathbf{i}}{T}\int \text{đ}p \frac{e^{\mathbf{i}\vec{p}\cdot\vec{x}}}{p^2} = \mathbf{i}\frac{2\pi}{(2\pi)^3 T}\int_0^\infty dp \underbrace{\int_{-1}^{1} d\cos\theta e^{\mathbf{i}px\cos\theta}}_{=\frac{2\sin px}{px}} \\
&= \mathbf{i}\frac{2}{(2\pi)^2 T}\int_0^\infty dp \frac{\sin px}{px} \\
&= \mathbf{i}\frac{2}{2\pi T}\frac{1}{x}\frac{1}{2}\underbrace{\int_{-\infty}^{\infty} d\bar{p}\frac{\sin\bar{p}}{\bar{p}}}_{=\pi} \\
&= \frac{\mathbf{i}}{2Tx}.
\end{aligned} \tag{120}$$

(I have set the velocity of propagation to 1, and $T = U/K$.) So

$$\langle e^{\mathbf{i}s\phi(x)} e^{\mathbf{i}s'\phi(0)} \rangle = e^{-\mathbf{i}\frac{ss'}{4xT}} .$$

And

$$\langle \cos\phi(x) \cos\phi(0)\rangle = \cos\frac{1}{4Tx}$$

which does not decay at long distance, and in fact approaches a constant.

Adding matter helps to produce a deconfined phase! In particular, the presence of enough massless charged fermions can render the monopole operator irrelevant. I recommend this paper by Tarun Grover for more on this.

[End of Lecture 19]

Next steps:

- phase diagram of $\mathbb{Z}_2$ gauge theory.
- entanglement in many body quantum systems.
