Notes for Physics 217A

Leon Balents

*Leon Balents**

UCSB

September 24th, 2015 – December 3rd, 2015

Contents

*balents@kitp.ucsb.edu

1 What this course is about

Lecture 1 (1.5h)
September 24th, 2015

One of the advantages of teaching an advanced graduate course is the flexibility to interpret the subject. In the past this class has really been mostly about techniques - we teach coherent states, path integrals, perturbation theory, etc., a set of many body tools used often by theorists. These tools have a lot in common with those in quantum field theory (QFT). Indeed, another way to teach this class is as "QFT for condensed matter physics". This is a popular sort of approach. It somehow embodies the view that QFT is an ideal of beauty in physics, and that it is the ultimate that we could hope to achieve in condensed matter physics (or elsewhere). A lot of theorists really believe this - maybe they are right. There are a couple of nice courses online that I found which seem to take this view. They make good reading:

- There is this class "Demystifying quantum field theory" by Ashvin Vishwanath at Berkeley.
- Here is another class, "Whence quantum field theory" by John McGreevy at UCSD.

There's a lot to like about these classes. They approach QFT from a decidedly CM perspective, and hit quite a few modern topics. Moreover, they focus a fair amount on physics rather than technique.

I want to take this a bit further, really putting the physics front and center, and so this course is going to be about:

"Universal and non-universal properties of matter"

In my view QFT is just one technique to understand this subject. Actually the more one studies QFT the less one is sure of where it is even well-defined. The real subject of this class - properties of matter - is always defined. It includes but transcends QFT. (where the subject gets hazy is when gravity and quantum mechanics come together, and there, well, nobody really knows what is going on and we will not enter there). If you love QFT, well great, good for you. You are not alone.

So here are the general topics:

- **The thermodynamic limit:** We will for the most part be interested in matter in the thermodynamic limit, i.e. in systems with at least one spatial dimension very large, with a finite number of degrees of freedom per unit volume. This limit gives rise to emergent phenomena and universal behaviors.
- **Phases of matter:** Solid, liquid, gas - you've heard of this. We will talk about what constitutes a phase, what the properties of phases are, and examples of important phases.
- **Entanglement:** How quantum are different states of matter? We'll look a bit into how quantum information theory informs the study of matter.
- **Quasiparticles:** What are the excitations of quantum ground states? Often they behave like particles.
- **Correlation and response:** What do we actually measure in condensed matter? How is it related to what we can calculate?
- **Functional integral techniques:** You gotta know this stuff. But I hope we don't need to spend too much time on it.
- **Transfer matrix:** How to turn quantum problems into classical stat mech ones and vice versa
- **Symmetry breaking:** One of the most basic and pervasive phenomena in nature
- **Physical examples:** Quantum Ising model, bosonic superfluids and Mott insulators, AKLT chain, toric code

2 Quantum Ising model

First we are going to look at an iconic simple model of a quantum many body system: the transverse field Ising model. It is very simple. We take a set of sites i, and on each site we put a spin-1/2 system, i.e. a two-dimensional Hilbert space spanned by $|\uparrow\rangle = |+1\rangle$ and $|\downarrow\rangle = |-1\rangle$. The full space is a direct product of these, i.e. by states of the form $|s_1\rangle \otimes |s_2\rangle \cdots ...$, where each $s_1 = \pm 1$-. We define Pauli matrices in the usual way σ_i^μ acting on site i, that is really $\sigma_i^\mu = 1 \otimes 1 \otimes \cdots \otimes \sigma^\mu \otimes 1 \cdots$. Then the Hamiltonian is written as

$$H = -J \sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z - h \sum_i \sigma_i^x, \tag{1}$$

Here $\langle ij \rangle$ means the first sum is over nearest-neighbor pairs of sites on a lattice. For our purpose you can think of a hypercubic lattice, where sites are specified by integer coordinates $i = (n_1, n_2, \ldots, n_d)$ in d dimensions. We'll take $J \geq 0$ (a ferromagnetic Ising coupling).

Why is this a good model? It is simple, but it also satisfies a lot of basic physical requirements:

- It has a finite number of degrees of freedom (read finite-dimensional Hilbert space) locally
- It is local: H is a sum of terms, each of which acts only on sites near one another. Implicitly to define this we needed the direct product structure of the Hilbert space, so that a local operator (like σ_i^μ) acts only within a few members of the direct product.
- It has an Ising symmetry: H is invariant under the unitary transformation which takes $\sigma_i^\mu \to -\sigma_i^\mu$ for $\mu = y, z$. This is generated by the unitary (and Hermitian) operator $U = \prod_i \sigma_i^x$, which clearly commutes with H.
- It has translational symmetry: H is unchanged by taking $\sigma_i^\mu \to \sigma_{i'}^\mu$, where i'(i) is just the site i translated by some lattice vector (i.e. vector with integer components). So every site "looks" the same.

The first, second, and last properties are things we are going to nearly always assume in this course. The third is a particular example of a symmetry, the simplest we can imagine.

2.1 *Limits*

Weak exchange

Now let us consider the two extreme limits. Suppose $J = 0$. Then σ_i^x commutes with H and is a good quantum number $= \pm 1$ for each site. The ground state (we take $h > 0$ without loss of generality) is the state with $\sigma_i^x = +1$ on all sites. Schematically,

$$|0\rangle = |\to\to\cdots\to\rangle. \tag{2}$$

This ground state is unique, and is invariant under all the symmetries of H. Consequently, it is sometimes called a "quantum paramagnet". Importantly, it is a product state – there is no quantum entanglement. What are the excitations? The simplest ones, which we would call elementary excitations, are states in which a single spin is flipped. These have energy $+2h$ compared to the ground state,

$$\begin{aligned} |i\rangle &= |\to\cdots\to\leftarrow_i\to\cdots\to\rangle, \\ &= \sigma_i^z|0\rangle. \end{aligned} \tag{3}$$

These states are odd under the Ising symmetry. They are not translationally invariant, but we can form plane wave combinations, which we will do momentarily. The next higher excitations are two-particle ones, which we could write as, for example,

$$|i,j\rangle = \sigma_i^z\sigma_j^z|0\rangle, \tag{4}$$

which has energy $+4h$. Note that $|i,j\rangle = |j,i\rangle$ which means that the spin flips are identical particles, in the quantum sense: if we exchange two particles, it

is not a linearly independent state. Indeed the lack of a sign would indicate that these particles are bosons.

Now let us consider the effect of J in perturbation theory. To obtain the first order correction to the energies, we just need to consider matrix elements of the J term within each degenerate manifold at $J = 0$. As the $J = 0$ states are eigenstates of σ_i^x, each term in J flips two spins from $+$ to $-$ or vice versa. Hence in the ground state manifold there is no energy shift,

$$\Delta E_0 = O(J^2), \tag{5}$$

Now let us look at the one particle states. Here there are non-zero matrix elements. Specifically

$$\langle j|\sigma_k^z\sigma_l^z|i\rangle = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{kl}\delta_{ij}, \tag{6}$$

which means that the flipped spin can be "moved" from one site to its neighbor (the last negative term is just there to cancel double counting if $k = l$ which is trivial). Writing

$$|\psi\rangle = \sum_i \psi_i|i\rangle, \tag{7}$$

we obtain

$$-J \sum_{j\in NN(i)} \psi_j = \Delta E\psi_i. \tag{8}$$

This is like a one-particle Schrödinger equation. It is solved by defining plane waves,

$$\psi_j = e^{ik\cdot r_j}, \tag{9}$$

i.e.

$$|\mathbf{k}\rangle = \sum_j e^{i\mathbf{k}\cdot\mathbf{r}_j}|j\rangle, \tag{10}$$

which gives

$$\Delta E(k) = -2J \sum_{\mu=1}^{d} \cos k_\mu. \tag{11}$$

This we see describes a "band" for a particle which has a minimum energy at $k = 0$, and bandwidth of $4Jd$. The gap is slightly reduced, by $-2Jd$, from the value of $+2h$ at $J = 0$. The plane wave nature means that these particles propagate, and the energy shift is the kinetic energy of the particle. The minimum energy excitation occurs at $k = 0$, and has the energy (note that there is no shift in the ground state energy to linear order in J)

$$\epsilon_{qp}(k = 0) = 2h - 2dJ. \tag{12}$$

So we see that the gap between the group state and the minimum energy quasiparticle decreases with J, at least when J is small.

Suppose we look not at the energy but at the wavefunction of the ground state. Then there is a first order effect. As usual, this is obtained as

$$|0\rangle' = |0\rangle + (E_0 - H_0)^{-1} H' |0\rangle + O(J^2), \tag{13}$$

where H_0 is the transverse field term and H' is the J term. This gives

$$|0\rangle' = |0\rangle + (J/4h) \sum_{\langle ij \rangle} |i,j\rangle + O(J^2). \tag{14}$$

We see that pairs of nearby particles are mixed into the modified ground state wavefunction at this order. We can say that these are "virtual pairs" of particles. Note that there are no actual excitations in this state. It is still the ground state, it is a unique state, and the energy is not even shifted at this order.

Weak field

Now consider the opposite limit, $h = 0$. Now σ_i^z is a good quantum number instead of σ_i^x. To minimize the energy, all spins should be aligned in a ground state: the energy is just $-J$ times the number of aligned bonds. There are now *two* ground states,

$$|\uparrow\uparrow \ldots \uparrow\rangle \qquad \text{and} \qquad |\downarrow\downarrow \cdots \downarrow\rangle. \tag{15}$$

These are the ferromagnetic states with positive and negative magnetization. Each of these is *not* invariant under the Ising symmetry: this is the symptom of *spontaneous symmetry breaking*. We could if we wanted to choose different linear combinations,

$$|\text{cat}\pm\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\uparrow \ldots \uparrow\rangle \pm |\downarrow\downarrow \cdots \downarrow\rangle \right). \tag{16}$$

These "cat" states are eigenstates of the symmetry, that is $U|\text{cat}\pm\rangle = \pm|\text{cat}\pm\rangle$. Back in quantum mechanics class, we were taught that we can always choose Hamiltonian eigenstates to also be eigenstates of a complete set of commuting symmetry generators. So in this sense the cat states are "better" than the simpler-looking ferromagnetic states. However, in physical terms they are usually worse. The problem is that the cat states consist of superpositions of macroscopically different states. There are local operators (e.g. σ_i^z) which distinguish the two ferromagnetic components. Consequently any local measurement can "collapse" a cat state into one of the two ferromagnetic states. Since these measurements can occur anywhere in the system, even if there is a very low probability to make this measurement locally, the probability adds up over the whole system and when there are many sites, collapse becomes almost certain. So the cat states are very fragile.

Another important feature is that the two ferromagnetic components are

"super-orthogonal" (I made up that term!). Not only are they orthogonal, but if we act with any local operator on one ferromagnetic state the result remains orthogonal to the opposite magnetization state. This means that in any superposition of the two ferromagnetic states, expectation values of local operators behavior probabilistically: they involve only diagonal matrix elements of the two components. So even if somehow a cat state is realized, we can understand its properties just by classically weighting properties of the two ferromagnetic states.

The low energy excited states are those in which a small number of bonds are misaligned, and the excitation energy of such a state is $+2J$ times the number of misaligned bonds. What is the smallest number of such bonds? Above one dimension, and forgetting about boundaries, the minimum number is $2d$, and correponds to just flipping one spin,

$$|i\rangle_\uparrow = \sigma_i^x |\uparrow\cdots\uparrow\rangle, \tag{17}$$

and similarly we can construct the state $|i\rangle_\downarrow$ with all the spins flipped. These states correspond to the smallest "domain" of flipped spins. Their excitation energy is $+4Jd$. The state $|i\rangle_\uparrow$ is a "quasiparticle" for the ferromagnetic state, similar to $|i\rangle$ in Eq. (3) for the quantum paramagnet.

Higher excited states consist of flipping larger domains. For example, if we flip two neighboring spins, it results in $4d-2$ misaligned bonds (the minus two is because the bond between the two flipped spins is not misaligned), hence excitation energy $(8d-4)J$. If instead we flip two non-adjacent spins, then the energy is $2\times 4Jd = 8Jd$. Note that the excitation energy for two adjacent spin flips is *less* than that of the two separated spin flips in both cases. We can therefore consider the adjacent spin flip state as a *bound state* of two elementary quasiparticles.

One dimension is special. If we consider two adjacent flipped spins we see that the exciation energy above the ground state is *identical* to that of a single flipped spin. It costs no energy to flip the second spin! In fact, we can continue to flip more spins, growing the domain of flipped spins as much as we want, with no energy cost. What this means is that in one dimension, we got the minimum energy excited state wrong. In fact, we can (in an infinite system or one with open boundaries) arrange for only *one* not two misaligned bonds, which has an energy cost of $2J$ instead of $4J$ like the single spin flip state. To do this, we flip a *semi-infinite* string of spins:

$$|i+1/2\rangle_{\uparrow\downarrow} = |\uparrow\cdots\uparrow_i\downarrow_{i+1}\cdots\downarrow\rangle. \tag{18}$$

This state is a *domain wall* or *soliton*, which connects the two different symmetry broken states. We can imagine a domain wall in any ferromagnet, but only in one dimension does this state have finite energy – in general it has an energy propertional to the cross-sectional area of the domain wall. In 1d, it is actually the lowest energy excited state.

Lecture 2 (1.5h)
September 29th, 2015

We'll come back to 1d later; for now let us go back to $d>1$. Here, the situation seems fairly similar to that in the small J limit, except for the existence of two degenerate ground states in this case. We can carry through similar

manipulations to see the perturbative effects of the transverse field, just as we looked at the perturbative effects of J in Sec. 2.1. That is, we take H_0 to be the exchange term and H' the transverse field term, opposite to the previous case. There is a small twist compared to the small J case, which is that the first order energy shift of *all* eigenstates is zero. This is because H' flips only one spin which always changes the zeroth order energy (In $d > 1$). Thus energy shifts occur first at second order in h and we need to be a little more sophisticated. We will now indulge in this exercise, which is a bit painful. The only real point here is to show that the gap for the quasiparticle state *decreases* with h. If you don't care how this is shown, then you can skip ahead to Eq. (28). If you do care, maybe you can check my algebra – there is a high probability of some sloppy mistake.

In general, we can use the following procedure. Consider the subspace of all unperturbed states of a given energy E_0, and define the projector $\mathcal{P}$ onto this subspace. Then consider an exact eigenstate $|\Psi\rangle$ which connects to this subspace as H' is taken to zero. Then we can define $|\Psi_0\rangle = \mathcal{P}|\Psi\rangle$ as the component of the exact wavefunction in the "bare" ground state subspace. It satisfies the equation

$$\left[E_0 + \mathcal{P}H' \sum_{n=0}^{\infty} \mathcal{R}^n \mathcal{P}\right] |\Psi_0\rangle = E|\Psi_0\rangle, \tag{19}$$

where

$$\mathcal{R} = (E - H_0)^{-1}(1 - \mathcal{P})H' \tag{20}$$

is called the "resolvent". Eq. (19) appears to give an effective Hamiltonian for the problem. It is not quite a Hamiltonian since the object in the square brackets depends upon E through the resolvent in Eq. (20). Hence it is really a non-linear eigenvalue problem. However, to any given order in perturbation theory, E may be expanded in a series as well to obtain a true Hamiltonian form. The leading order energy shift of any given state is obtained by replacing $E \to E_0$ inside $\mathcal{R}$ and taking the first non-zero term in the sum over n.

Now we can apply this to the ground state first. Here the space spanned by $\mathcal{P}$ is one dimensional so the effective Hamiltonian is a number. The first non-zero term is just $n = 1$:

$$H_{\text{eff}}^{(GS)} = E_0 + \mathcal{P}H'(E_0 - H_0)^{-1}H'\mathcal{P}. \tag{21}$$

Now H' just flips a spin so the eigenvalue of H_0 is increased from E_0 by $4dJ$, and to get back to the original space (so as not to be annihilated by the left-most $\mathcal{P}$ operator, the same spin must be flipped back. Hence we have

$$E_{GS} = E_0 - N\frac{h^2}{4dJ}, \tag{22}$$

where the factor of N comes from the sum over possible sites of the flipped "virtual" spin.

Now consider the energy shift of the quasiparticle states. This time $\mathcal{P}$ should be taken to be the projector onto the N-dimensional subspace of states $|i\rangle_\uparrow$. We have

$$H_{\text{eff}}^{(qp)} = E_0^{qp} + \mathcal{P}H'(E_0^{qp} - H_0)^{-1}H'\mathcal{P}. \tag{23}$$

Here again the right-most H' flips one spin (we will call this the "first" spin flipped) and the left-most one flips another. Consider the action of this Hamiltonian upon a quasiparticle state $|i\rangle_\uparrow$. There are several possibilities:

1. The first spin flipped (call it j) is neither i nor one of its neighbors, and the second spin flipped is again j. Then the denominator $E_0^{qp} - H_0 = -4dJ$.

2. The first spin flipped is neither i nor one of its neighbors, but the second spin flipped is spin i. Again$E_0^{qp} - H_0 = -4dJ$.

3. The spin j is a neighbor of i, and the second spin flipped is also j. The energy denominator is $4dJ - (8d-4)J = -4(d-1)J$.

4. The spin j is a neighbor of i, and the second spin flipped is i. The energy denominator is $4dJ - (8d-4)J = -4(d-1)J$.

5. The spin j is i, and the second spin flipped is arbitrary (call it k). The energy denominator is $+4dJ$.

So we obtain

$$\begin{aligned} H_{\text{eff}}^{(qp)}|i\rangle_\uparrow &= E_0^{(qp)}|i\rangle_\uparrow - \frac{h^2}{4dJ}(N-(2d+1))|i\rangle_\uparrow \\ &\quad - \frac{h^2}{4dJ}\sum_{|j-i|>1}|j\rangle_\uparrow \\ &\quad - \frac{h^2}{4(d-1)J}(2d)|i\rangle_\uparrow \\ &\quad - \frac{h^2}{4(d-1)J}\sum_{|j-i|=1}|j\rangle_\uparrow \\ &\quad + \frac{h^2}{4dJ}\sum_k |k\rangle_\uparrow. \end{aligned} \tag{24}$$

Each line corresponds to a line in the list above. We can simplify by rewriting the term in the second line as $\sum_{|j-i|>1}|j\rangle_\uparrow = \sum_k |k\rangle_\uparrow - \sum_{|j-i|=1}|j\rangle_\uparrow - |i\rangle_\uparrow$. This allows us to cancel the final term and write (using Eq. (22) and $E_0^{(qp)} = E_{GS} +$

$4dJ$)

$$H_{\text{eff}}^{(qp)}|i\rangle_\uparrow = \left[E_{GS} + 4dJ + \frac{h^2}{4dJ}(2d+1) - \frac{h^2}{4(d-1)J}(2d) + \frac{h^2}{4dJ}\right]|i\rangle_\uparrow \tag{25}$$

$$- \left[\frac{h^2}{4(d-1)J} - \frac{h^2}{4dJ}\right] \sum_{|j-i|=1} |j\rangle_\uparrow$$

$$= \left[E_{GS} + 4dJ - \frac{h^2}{2d(d-1)J}\right]|i\rangle_\uparrow - \frac{h^2}{4d(d-1)J} \sum_{|j-i|=1} |j\rangle_\uparrow. \tag{26}$$

This now has the form of a hopping Hamiltonian for the quasiparticle. By going to momentum space, as in Eq. (10), we obtain

$$\epsilon_{\text{qp}}(k) = E_{\text{qp}} - E_{GS} = 4dJ - \frac{h^2}{2d(d-1)J} - \frac{h^2}{4d(d-1)J} \sum_\mu 2\cos k_\mu. \tag{27}$$

The minimum energy quasiparticle has $k_\mu = 0$, with the energy

$$\epsilon_{\text{qp}}(0) = 4dJ - \frac{1}{d(d-1)}\frac{h^2}{J}. \tag{28}$$

We see that the gap to the lowest energy quasiparticle above the ferromagnetic state decreases with increasing h, just as the corresponding gap did in the paramagnetic state with increasing J (albeit quadratically in the present case).

Putting this together with the results of the previous section, we observe that both the ferromagnetic and paramagnetic phases exhibit a gap, at least in their extreme limits, but that this gap decreases moving toward the intermediate regime. In fact, we will argue later that the gap is always non-zero except at a special value of J/h, which defines the point at which the paramagnetic and ferromagnetic phases meet. This is a continuous quantum phase transition, or quantum critical point.

There is one new feature we can think about, which is the possible splitting of the ground state degeneracy between the up and down polarized states. We do not expect this degeneracy for $h \neq 0$ in a finite system. In general we only expect degeneracies in a finite system if symmetry demands it. It is not the case here. Quantum mechanics tells us we can choose eigenstates with definite values of $U = \pm 1$, which are the cat states in Eq. (16). But there is no reason they need to have the same energy. You can see that the ground state energy does not exist in general by an argument of *reductio ad absurdem*. Take just two sites, with Hamiltonian $H = -J\sigma_1^z\sigma_2^z - h\sigma_1^x - h\sigma_2^x$. It is straightforward to find the ground state, which has the form

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}\left[\cos\theta\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right) + \sin\theta\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right)\right], \tag{29}$$

where $\tan\theta = (\sqrt{J^2+4h^2} - J)/2h$. This is smoothly connected to the $|\text{cat}+\rangle$ state, which it becomes as $h \to 0$. The first excited state is exactly the $|\text{cat}-\rangle$

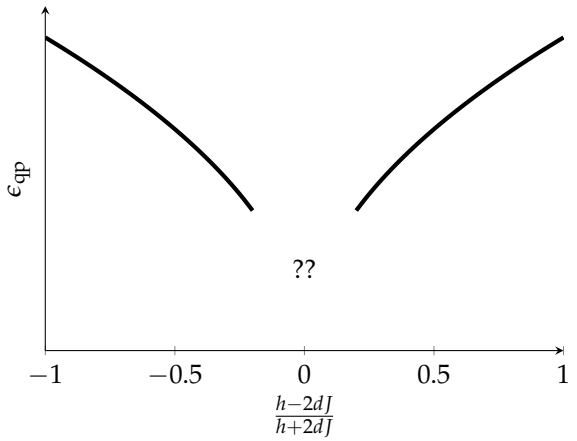


Figure 1: Energy gap in the Ising model (schematic).

state. The energy gap is $E_1 - E_0 = \sqrt{J^2 + 4h^2} - J \approx 2h^2/J$ for $h \ll J$. We can understand this result from degenerate perturbation theory. The ground state is split by at $\mathcal{O}(h^2)$ as two spin flips are required to connect the up and down polarized ferromagnetic states. While an exact solution is more complicted for a larger system, we may anticipate from the perturbation theory argument that the splitting between ground states in general is of order $J(h/J)^N$, for a lattice with N sites (we are using the logic of Eq. (19)). This vanishes exponentially with the volume of the system. So the cat states are exponentially close in energy, and both are well separated from the quasiparticle states. We see that the phenomena of spontaneous symmetry breaking is robust in the large system size limit.

What did we learn?

By looking at the two limits, we have made a few observations. Let me summarize the key points:

- The quantum Ising model has two phases, paramagnetic and ferromagnetic. Each phase has a limit in which the ground state becomes a product state, with zero entanglement.
- The lowest energy excitations near these limits are quasiparticles, which behave like particles hopping on the lattice. They have a definite energy-momentum relation.
- A quasiparticle costs a non-zero minimum energy to produce: we say there is an energy gap. The gap is reduced upon perturbing from either zero entanglement limit. In fact, though we did not show it, the gap is non-zero everywhere except at the phase transition point.
- In the ferromagnetic phase, the Ising symmetry is spontaneously broken, and the low ground state and low energy states are separated into two sectors (opposite magnetization) which are disconnected by any local

operator. In a finite system, the sectors mix to form cat states, but the splitting between these states is exponentially small in volume, and they are very fragile with respect to perturbations.

2.2 *Phase diagram and mean field theory*

Above we found the behavior in the two extreme limits. From these limits, we notice a clear qualitative difference. In the small J limit, the ground state is invariant or even under the Ising symmetry and unique. Moreover, all the low energy states can be assigned definite Ising parity: the single quasiparticle states are odd under U, two quasiparticle states are even, etc. There is no degeneracy in the thermodynamic limit of states of different parity. In the small h limit, the ground state is two-fold degenerate, and the "natural" eigenstates have indefinite parity. Even if we choose to form parity eigenstates, we find that states with different parity are degenerate in the thermodynamic limit.

The difference between these two situations is qualitative, provided we maintain the Ising symmetry of H. Indeed, it turns out that the two limits are separated by a *quantum phase transition*. Let us look at this in a simple approximation. We saw that in both limits, the ground is becomes a product state. So we can hope that this provides a reasonable description of the physics in between. Let's assume this is the case, and solve the problem variationally. We take a trial state

$$|\psi\rangle = \otimes_i |\phi\rangle_i, \tag{30}$$

where $|\phi\rangle$ is some arbitrary, normalized, single spin state. Then the variational energy is

$$E_{\text{var}} = \langle\psi|H|\psi\rangle = -J\sum_{\langle ij\rangle}(m^z)^2 - h\sum_i m^x = N\left[-Jd(m^z)^2 - hm^x\right], \tag{31}$$

where N is the number of sites, and $m^\mu = \langle\phi|\sigma^\mu|\phi\rangle$. One can also interpret this approximation in terms of a "mean field" Hamiltonian which just represents decoupled spins:

$$H_{\text{mf}} = -\sum_i \mathbf{h}_{\text{eff}}\cdot\boldsymbol{\sigma}. \tag{32}$$

such that $|\psi\rangle$ is the ground state of H_{var}. One has simply $\mathbf{m} = \mathbf{h}_{\text{eff}}/|\mathbf{h}_{\text{eff}}|$. As the overall magnitude of $\mathbf{h}_{\text{eff}}$ does not enter, we can choose to write $\mathbf{h}_{\text{eff}} = h\hat{\mathbf{x}} + h_{\text{ex}}\hat{\mathbf{z}}$, where h_{ex} is called the exchange field. It coincides with the Curie-Weiss average (mean) field seen by a single spin due to the exchange interaction J. Consequently we can refer to this approximation as Mean Field Theory (MFT). There are many formulations of MFT and we will encounter some others later in the course.

It is easy to show that the magnetization obeys the constraint $|\mathbf{m}| = \sqrt{\sum_\mu (m^\mu)^2} = 1$. So we can choose $m^x = \cos\theta$ and $m^z = \sin\theta$ (there is no

advantage to taking non-zero m^y), and the energy becomes

$$E_{\text{var}}/N = -Jd\sin^2\theta - h\cos\theta = -Jd + Jd\cos^2\theta - h\cos\theta. \tag{33}$$

Note that the energy is even in θ: this is the Ising symmetry $m^z \to -m^z$. Minimizing with respect to θ we find

$$m^z = \sin\theta = \begin{cases} 0 & h > 2dJ \\ \pm\sqrt{1-\left(\frac{h}{2dJ}\right)^2} & h < 2dJ \end{cases} \tag{34}$$

One sees that MFT predicts a phase transition at $h = 2dJ$ between the paramagnetic and ferromagnetic phases. The magnetization m^z is non-analytic at the transition, exhibiting a square root singularity. Non-analyticities are characteristic of phase transitions (we have a whole course on this – Physics 220).

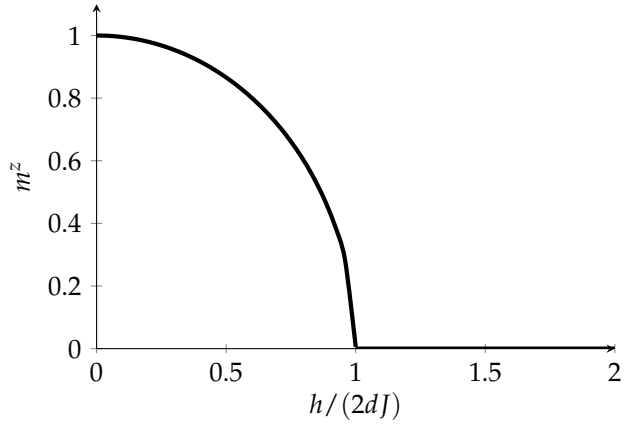


Figure 2: Magnetization in mean field theory

3 Entanglement

3.1 *Block picture*

We're going to take a bit of time now to try to generalize/abstract from our treatment of the quantum Ising model. We already saw that, at least in the two extreme limits, the ground state becomes a product state, its excitations are quasiparticles, and they have a gap. To what extent are these robust properties? Clearly the exact ground state away from the two extremes is not exactly a product state, but the latter two properties seem to persist, at least in perturbation theory. Are these observations connected?[1]

It is rather natural to think that they are. Why might the ground state be a product state? Note we mean not necessarily a product over single sites, but over blocks. This works if the connections between the blocks are not so

[1] A warning: we will discuss these questions in physical terms, without very much rigor – I am not capable of it.

important. You might even think this is always the case (warning: it is not), because the Hamiltonian is local. The terms in H connect only neighboring sites. In a variational sense, only the expectation value of single spins and nearest-neighbor pairs matters. So we can do a perfectly good job optimizing the energy of nearly all terms in the Hamiltonian with such a product state. We miss only terms that straddle the boundaries. So if we can find a decent way to smooth these boundaries, and connect the blocks just a little, we can expect to get a good energy.

The physical implication of a block product state is that a finite subsystem of just a few spins captures the essentials of the system. This then directly connects to the nature of the excitations. An excitation within any single block is like an excited state of a finite quantum system: its energy is always non-zero. This is why we expect that a product-like state should have an excitation gap. How is this "particle-like"? Well, the minimum energy state we can make in such a block-like system is just to excite a single block. This is a localized excitation: by measuring local operators we can tell which block it is in. Outside that block, the system remains in its ground state. This is physically like a particle in the vacuum. We are free to introduce superpositions of states with the excitation in different blocks, and thereby create plane wave states. In this way, we create states with a well-defined dispersion, i.e. a definite excitation energy for a given momentum.

With this picture, it is also natural that these properties are robust to small perturbations. A perturbation within a block obviously just makes small changes to the block ground and excited states, but this preserves the product nature. A small coupling between blocks can be treated perturbatively, and because of the gap between the ground and excited states, only small excitations of neighboring blocks are introduced into the ground state. We may need to expand the blocks slightly to write the new ground state in block form, but we can expect it still has one, approximately.

We can also think in terms of energies. We can expect that a finite (not infinitesimal) perturbation is needed to close a finite excitation gap. We indeed saw explicitly that the quasiparticle gaps in the Ising model suffered only infinitesimal decreases under infinitesimal introduction of perturbations around the product state limit. We emphasize that this is really only natural once you accept both the product state/finite size picture of the ground state and excitations *and* you assume locality of the Hamiltonian. This is because H is in fact a sum of a number of terms of order the system volume, which diverges in the thermodynamic limit. So the absolute change in the Hamiltonian is large even if you make a tiny change in a parameter such as J or h. Indeed, the absolute ground state energy is proportional to N, and so is the change in this quantity. To get the change in the *excitation* energy of a quasiparticle in Sec. 2.1, this large quantity needed to cancel in the difference of energies of ground and excited states. That it does so is due to locality.

So what do we conclude from these, admittedly vague, arguments? Local quantum systems can have ground states that are well-approximated by products of finite blocks. In this situation, we generically expect a gap to all excitations, and that the lowest excitations are quasiparticle-like. These prop-

erties should be robust: i.e. they should persist under arbitrary local perturbations that are not too large. The region of parameter space where this does persist defines a zero temperature phase of matter. The Ising paramagnet and ferromagnet are examples of two such phases.

Lecture 3 (1.5h)
October 1st, 2015

3.2 *Entanglement measures*

To begin to sharpen up the vagueness of the previous subsection, let us refine the notion of something that is "almost" a product state. A strict product state is one of the form

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle, \tag{35}$$

where A and B are disjoint subsystems and the Hilbert space has a direct product form. Here because of locality, we speak of subsystems that are spatially disjoint, i.e. A is a collection of some lattice sites, and B is the rest. In such a state, measurements of operators that reside entirely within A are completely independent of those in B. We say there is zero entanglement between A and B.

Any state that cannot be written as such a product state is said to be entangled. A measure of this entanglement is provided by entanglement entropy. For any state, we can define a density matrix,

$$\rho = |\psi\rangle\langle\psi|. \tag{36}$$

This is a pure state density matrix, i.e. a projection operator. It has a very simple eigenvalue spectrum: one eigenvalue is 1 and all the others are 0. Now we can construct a reduced density matrix,

$$\rho_A = \mathrm{Tr}_B\, \rho. \tag{37}$$

If $|\psi\rangle$ is a product state as in Eq. (35), then ρ_A is also a pure state density matrix, but otherwise it is not. So any deviation from the pure state eigenvalue spectrum can be used to measure the entanglement of A and B. The most common measure is called the von Neumann entanglement entropy

$$S = -\mathrm{Tr}\left[\rho_A \ln \rho_A\right]. \tag{38}$$

By construction 0 and 1 eigenvalues contribute zero to S. In general, $S \geq 0$ and is bounded by $\ln 2$ times the smaller of the two dimensions of subspaces A or B. An instructive example to get intuition for S is to consider a state $|\psi\rangle$ which is a product of n entangled pairs of spins (e.g. singlets $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$), with one spin of each pair living in A and the other in B, times a state which is a product of components in the remainder of A and B. For such a state, one finds $S = n \ln 2$.

With the tool of entanglement entropy, we can try to diagnose product-like states. If we had a perfect product state, then if we subdivide the system in such a way as to cleanly separate all blocks within A or B, then $S = 0$. However, we do not a priori know where the blocks are. Moreover, we can expect

that even an approximate product state has some mixing of states across the boundaries between blocks. What we do expect is that if we take regions A and B to be very large, they will each contain many blocks entirely within them (the idea is that the blocks have some fixed size, while the size of A or B can be made as big as we want). Then the contribution to S comes entirely from the blocks on or near the boundary between A and B. Hence we expect that S is proportional to the *area of the boundary between A and B*, i.e.

$$S \sim s_0 L^{d-1} + \cdots , \tag{39}$$

where s_0 is some constant and L is the linear size of region A. This is the *area law* of entanglement. We generically expect it for any product-like state. Unfortunately, the converse is not true: non-product states also can obey the area law. So the area law is *not* enough to discriminate, in general, between other types of states. In fact, we can exhibit many examples of non-product states, even in phases with zero gap, where the area law is obeyed. One way to distinguish these is to look at the subleading $(\cdots)$ terms in Eq. (39), but we will discuss others. If you want to read more about the area law, there is a review article here. Conversely, there are (gapless) ground states that violate the area law. We will come to that later.

More refined measures of entanglement are possible. Most generally, given a partition into subsystems A and B, we can write the full ground state in a *Schmidt decomposition*:

$$|\psi\rangle = \sum_i c_i |i_A\rangle |i_B\rangle, \tag{40}$$

where $|i_A\rangle$ is an orthonormal set in A and $|i_B\rangle$ is an orthonormal set in B. The c_i are "Schmidt coefficients" which can be chosen real (by choice of phase of the states) and $0 \leq c_i \leq 1$, and $\sum_i c_i^2 = 1$. The entanglement entropy is just $S = -\sum_i p_i \ln p_i$, where $p_i = c_i^2$ are the eigenvalues of the reduced density matrix. The set of these eigenvalues – called the entanglement spectrum – constitutes a characterization of entanglement which is independent of the basis choice in A and B. One can use the entanglement spectrum to more finely distinguish different types of states.

3.3 *Matrix product states*

In one dimension, the area law reduces to a simple constant behavior at large L. Unlike in higher dimensions, it is believed that in 1d, the area law does guarantee the product-like nature of a state. Specifically, any state obeying an area law can be smoothly deformed into a product state, provided no symmetry is imposed upon the state during the deformation (we will come back to this point). Moreover, it has been proven that any gapped system in one dimension obeys the area law – see this reference.

So it is somewhat natural to focus on wavefunctions which obey the area law in one dimension. This is actually a very restrictive assumption. It turns out that if we were to randomly guess a wavefunction of a large system, it would have zero probability of displaying an area law. Moreover, if we con-

sider eigenstates that are far from the ground state (i.e with an excitation energy of order ϵL^d, with $\epsilon > 0$), then these generically obey a volume law (i.e. $S \sim L^d$) rather than an area law – and these are by far the majority of eigenstates. Anyway, because imposing an area law on a state is so restrictive, there is an efficient "compression" of the representation of area law wavefunctions. This is given, in one dimension, by the Matrix Product State (MPS) construction. We can think of an MPS as a (major) improvement on the idea of a block product state – it realizes the area law while at the same time maintaining translational invariance, allowing quite general states, including ones that are not quite product-like. An MPS is written as follows. Consider a direct product Hilbert space of spins spanned by a basis of states $|s_1\rangle \otimes \cdots \otimes |s_N\rangle = |s_1 \cdots s_N\rangle$. A completely general state would be written as

$$|\psi\rangle = \sum_{s_1 \cdots s_N} \psi(s_1, \cdots, s_N) |s_1 \cdots s_N\rangle, \tag{41}$$

where $\psi(s_1, \cdots, s_N)$ is a set of q^N complex numbers, if the spins have q basis states per site (e.g. $q = 2$ for our Ising model). This is an exponentially large amount of information when N is large. In an MPS, this is reduced drastically, by taking a specific form for these coefficients,

$$\psi(s_1, \cdots, s_N) = \mathrm{Tr}\left[A_1(s_1) A_2(s_2) \cdots A_N(s_N)\right], \tag{42}$$

where $A_i(s)$ is an $m \times m$ dimensional matrix – m is called the inner dimension, or bond dimension, of the MPS. Note that we do not need to even assume translational invariance here. Even so, the number of parameters in Eq. (42) is *much* smaller than for a general state. Each matrix has m^2 elements, and there are q matrices for each site, and N such matrices, so the total number of such state is qm^2N. If one further assumes translational invariance, so that all the matrices are identical, then the factor of N is removed. According to rigorous results, one obtains an excellent approximation to the exact ground state wavefunction of a gapped, and hence area law, wavefunction by taking a *finite* m[2] So one is left with a finite number of parameters to describe such states!

It is nice to check that this wavefunction makes sense for some physical situation. Let's look at the quantum Ising model of Sec. 2.1. There we already worked out the perturbative form of the wavefunction in the weak exchange limit, $J \ll h$, in Eq. (14), which we transcribe here for the case of one dimension:

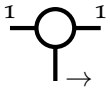
$$|0\rangle' = |0\rangle + (J/4h) \sum_i |i, i+1\rangle + O(J^2). \tag{43}$$

More explicitly in terms of spins, it is

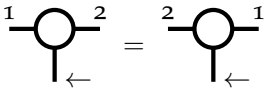
$$|\psi\rangle = |\rightarrow \cdots \rightarrow\rangle + (J/4h) \sum_i |\rightarrow \cdots \rightarrow_{i-1} \leftarrow_i \leftarrow_{i+1} \rightarrow \cdots \rightarrow\rangle + O(J^2). \tag{44}$$

[2] I am a little unclear on exactly what an "excellent approximation" means, but maybe you can figure out it from these papers: *MPS represent ground states* and *An area law for one-dimensional quantum systems*.

In fact, *to this order*, we can reproduce this wavefunction straightforwardly with an $m = 2$ MPS. We need just the elements:

(45)


$= 1,$

(46)


$= \sqrt{\frac{J}{4h}},$

and all other matrix elements zero. With this choice, it is easy to see that Eq. (44) is reproduced to $O[(J/h)]$. The MPS with these elements of course has higher order terms in J/h, which we did not check in perturbation theory, but any property we calculate to leading order will be properly reproduced. At the same time, the MPS so obtained has very nice properties, and can be thought of as a more physical "resummation" of terms in the perturbation theory to elevate Eq. (44) to a more appropriate wavefunction which has a good thermodynamic limit.

OK back to generalities. While proving that the MPS form is always good is not easy, it is reasonable, and the converse, that an MPS always obeys the area law, is simple. Suppose we subdivide the system into some interval A of L spins within the full system, so that say A contains sites 1 through L. Then we can write Eqs. (41,42) as

(47)
$$|\psi\rangle = \sum_{ab} \left(\sum_{s_1 \cdots s_L} [A(s_1) \cdots A(s_L)]_{ab} \, |s_1 \cdots s_L\rangle \right) \otimes \left(\sum_{s_{L+1} \cdots s_N} [A(s_{L+1}) \cdots A(s_N)]_{ba} \, |s_{L+1} \cdots s_N\rangle \right),$$

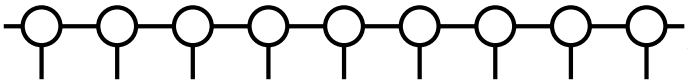
which we can immediately see has the Schmidt form of Eq. (40) with just m^2 Schmidt components. Thus the entanglement entropy is bounded above by $S \leq 2 \ln(m)$, and hence must obey the area (constant) law.

Clearly studying MPS's drastically simplifies the problem of studying quantum states. They are fun to work with too. An MPS has a nice graphically representation. It is most easily seen by writing out the matrix multiplication in Eq. (42) explicitly:

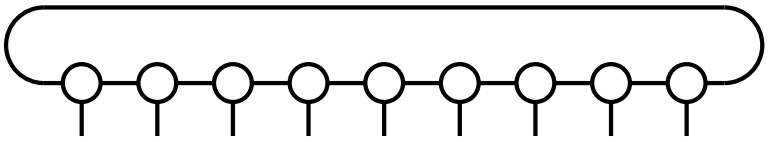
(48) $$|\psi\rangle = \sum_{s_1 \cdots s_N} \sum_{i_1 \cdots i_N} A_{i_1 i_2}(s_1) A_{i_2 i_3}(s_2) \cdots A_{i_N i_1}(s_N) |s_1 \cdots s_N\rangle.$$

We can draw each matrix as a "node" (I'll draw a circle) with three lines emerging from it: two internal legs representing the two matrix indices i, i' and one external leg representing the physical spin index s. The internal indices must be paired up with a neighboring matrix which we represent by connecting those lines. Indices on connected lines must match, and are summed. We

arrive at, for an infinite system

(49) $|\psi\rangle_\infty =$ 

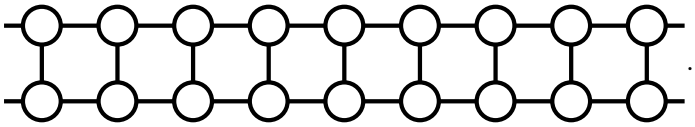
For a finite system with periodic boundary conditions, the picture is instead

(50) $|\psi\rangle_N =$ 

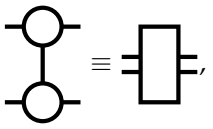
There are a lot of nice manipulations one can do with MPS's. For example, one can write down expectation values pretty easily. Let us first find the expectation value of the identity – i.e. the normalization of the state (this of course gives some normalization condition on $A(s)$). Formally, we have

$$\langle\psi|\psi\rangle = \sum_{s_1\cdots s_N} \sum_{i_1\cdots i_N} \sum_{j_1\cdots j_N} A^*_{j_1 j_2}(s_1) A_{i_1 i_2}(s_1) \cdots A^*_{j_N j_1}(s_N) A_{i_N i_1}(s_N), \tag{51}$$

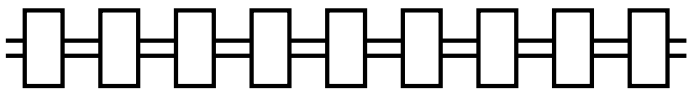
which is shown graphically as

(52) $\langle\psi|\psi\rangle =$ 

This object has the same structure as a one-dimensional partition function of a classical statistical mechanics model, where the i_a and j_a variables, i.e. the internal states of the MPS, are the "spins" of this fictitious classical system. Indeed one can define a *transfer matrix*

(53) $T_{ij;i'j'} = \sum_s A^*_{jj'}(s) A_{ii'}(s) =$  $\equiv$,

where now the double line means a composite index, or spin variable, composed of a pair of the original states (i, j). The norm, or partition function, is then a product over those transfer matrices,

$\langle\psi|\psi\rangle$ = 

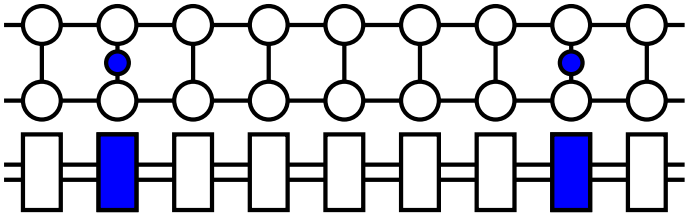
(54) $= \mathrm{Tr}\left(T^N\right),$

where I added a trace in the second line which really means I should have wrapped around the loose ends, but I was lazy in drawing. We see this is something relatively easy to calculate, since it depends only upon the eigenvalues of T, which is a finite $(m^2 \times m^2)$ matrix. For a large system it is com-

pletely dominated by the largest eigenvalue of T.

We can also calculate expectation values of operators easily. For example, if we have two local operators $\mathcal{O}_i$ and $\mathcal{O}_j$ which act only on the sites i and j, then their expectation value amounts to two insertions in the diagram:

$\langle\psi|\mathcal{O}_i\mathcal{O}_j|\psi\rangle$ =



=

$$= \mathrm{Tr}\left(T^{N-|i-j|}M[\mathcal{O}]T^{|i-j|}M[\mathcal{O}]\right), \tag{55}$$

where

$$M[\mathcal{O}]_{ij;i'j'} = \sum_{s_1 s_2} A^*_{jj'}(s')\langle s'|\mathcal{O}|s\rangle A_{ii'}(s) \tag{56}$$

is the modification of the transfer matrix by the operator $\mathcal{O}$. From this we can write the correlation function in terms of the eigenstates of T:

$$\langle\mathcal{O}_i\mathcal{O}_j\rangle = \frac{\mathrm{Tr}\left(T^{N-|i-j|}M[\mathcal{O}]T^{|i-j|}M[\mathcal{O}]\right)}{\mathrm{Tr}\left(T^N\right)}$$

$$\to_{N\to\infty} \sum_{n=1}^{m^2}\left(\frac{t_n}{t_0}\right)^{|i-j|}|\langle n|M[\mathcal{O}]|0\rangle|^2, \tag{57}$$

where $|n\rangle$ and t_n are the eigenstates and eigenvalues of T, ranked with $t_0 > t_1 > \cdots$. We see that correlations are the sum of exponentials. This is an expected property in a gapped state.

Lecture 4 (1.5h)
October 6th, 2015

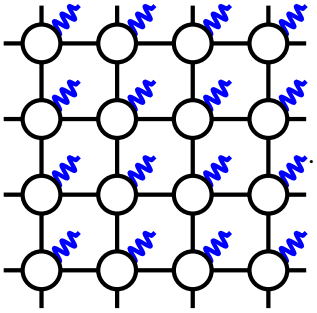
3.4 *Tensor Network States*

MPS are great in one dimension. There is a natural analog in higher dimension, which we may expect to describe gapped phases in that case. This is a tensor network state. We can write such a state as

$$|\psi\rangle = \sum_{\{s_i\}}\sum_{abc\cdots}\left[T_{abcd}(s_1)T_{aefg}(s_2)\cdots\right]|s_1\cdots s_N\rangle, \tag{58}$$

where now $T_{abcd}(s_i)$ are tensors defined on the physical sites i. Just as for MPS, there are internal "states" or indices, indicated here by roman letters $a, b, \cdots$, which are paired off and summed. This is an example of a Tensor Network State (TNS), specifically known as a Projected Entangled Pair State, or PEPS. It also has a graphical representation, which is indeed much more convenient

here than the explicit writing of tensors:

(59) $|\psi\rangle =$  .

Here we have drawn each "external" leg representing a physical spin index s_i as a wiggly line, to differentiate them from the internal ones. This helps for clarity in this two-dimensional drawing. In general, evaluating matrix elements between TNS reduces to some d-dimensional statistical mechanics-like problem, i.e. a sum over discrete variables on the lattice with weights that are a product of local terms. However, there is no guarantee that the weights are positive definite, which allows for more richness.

TNS have some properties that are similar to MPS. They build in the area law. However, there is no guarantee that TNS have exponential correlations in $d > 1$. This is because a discrete stat mech model in $d \geq 2$ can have ordered or quasi-long-range ordered phases, and sustain classical critical points. There is also no proof that a TNS can represent an arbitrary gapped state. In fact it seems difficult or impossible to properly represent so-called *chiral* phases, such as integer quantum Hall states.

3.5 *More on product-like states*

While the TNS of the last subsection are attractive from many perspectives, they are not necessarily product-state-like. Actually quite interesting states including states with *intrinsic topological order*, which we'll talk about later, can be represented as TNS. So if we are interested in defining product-like states, we need to do something else.

I'm not aware of any explicit general way to describe such a state. But we can define it implicitly. Pretty much the idea is just a fancier version of the perturbation theory we employed earlier in Sec. 2.1: we say that a ground state is product-like if it can be obtained by perturbation theory from some product state. Perturbation theory is just some smooth interpolation between the initial product state and the final ground state, where all the intermediate states are also "good" ground states. The formalization of this is that the ground state is obtained by a *local unitary* U acting on a product state,

(60) $|\Psi\rangle = U(t)|0\rangle,$

where $|0\rangle$ is a strict product state, and U is a unitary operator. But not just

any unitary operator, it should be of the form

$$U(t) = T_\tau e^{-i\int_0^t d\tau\, \mathcal{H}(\tau)}, \tag{61}$$

where $\mathcal{H}(\tau)$ is a sum of finite local hermitian operators (which might depend upon τ), and t is finite (e.g. $t = 1$ by choice of normalization of $\mathcal{H}$). The locality is important: it is what guarantees that the state obtained in this way is not highly entangled. We can think of U as enacting a finite time quantum evolution under the "Hamiltonian" $\mathcal{H}$ from an initial product state . We expect that under evolution by a local Hamiltonian, any information can only propagate over a finite length in a finite time. This includes entanglement. This is why we expect that $|\Psi\rangle$ defined this way remains "product-like".

Actually this statement has been proven. It goes by the name of a *Lieb-Robinson bound*, which says that in fact information propagates no faster than a finite "velocity" per unit time evolution by a local Hamiltonian. We can think of this in many ways (take a look at the wiki page for example to see the actual statement of the bound). But one which is intuitive is to think about the evolution of some local operator by U:

$$\mathcal{O}(t) = U(t)^\dagger \mathcal{O} U(t). \tag{62}$$

What is actually proven is this:

$$\| \left[\mathcal{O}(t), \mathcal{O}'\right] \| \leq c e^{-a(|X-Y| - v_{\mathrm{LR}}|t|)}, \tag{63}$$

where c, a and v_{LR} are constants, and $|X - Y|$ is the distance between the support of the operator $\mathcal{O}$ (at $t = 0$) and that of the operator $\mathcal{O}'$. The double lines denote the operator norm.

We can interpret this more explicitly as follows. Suppose $\mathcal{O}$ is a product of spin operators over some local region, for example a single spin at site i, $\mathcal{O} = \sigma_i^x$. Under evolution by $U(t)$, it will become a sum of terms which can be expressed as products of operators involving multiple sites,

$$\sigma_i^x(t) = \sum_j c_{ij\mu}(t)\sigma_j^\mu + \sum_{jk} c_{ijk\mu\nu}(t)\sigma_j^\mu \sigma_k^\nu + \sum_{jkl} c_{ijkl\mu\nu\lambda}(t)\sigma_j^\mu \sigma_k^\nu \sigma_l^\lambda + \cdots \tag{64}$$

What the Lieb-Robinson bound tells us is that terms in this sum (i.e. the coefficients $c_\cdot$) which involve *any* spin j far from i decay exponentially with the distance $|i - j|$ when this distance is larger than some "Lieb-Robinson velocity" v_{LR}, i.e. when $|i - j| > v_{\mathrm{LR}} t$ (and I think in fact even the sum of all such terms). In general, the "non-locality" of an operator generated by such local Unitary time evolution is restricted to distances less than $v_{\mathrm{LR}} t$. The Lieb-Robinson velocity itself is a non-universal property reflecting the Hamiltonian $\mathcal{H}$ defining the time evolution, but it is finite. The Lieb-Robinson result is pretty cool, as it produces something like the "light cone" you would have in relativistic systems, even though the problem is entirely non-relativistic. I would say it is extremely intuitive, and in no way surprising. But it is still remarkable to me that it actually proven rigorously.

What does this mean for product-like states? Well, if Eq. (60) holds, then we know that in $|0\rangle$, which is a strict product state, all connected correlation functions (which means correlation functions of operators with zero average) are zero. Correspondingly, correlations in $|\Psi\rangle$ can be regarded as correlations of dressed operators evolved by U in $|0\rangle$. The Lieb-Robinson result implies that these decay at least exponentially to zero at long distance. So the corrections to quantities like this, relative to a strict product state, are very weak.

We can also turn this around. We can always write a strictly single-site Hamiltonian which has $|0\rangle$ as its ground state, which is just the sum of projection operators onto the ground state of each site. Now we know that $|\Psi\rangle$ is the ground state of this single-site Hamilton evolved by U. According to the Lieb-Robinson result, this Hamiltonian for which $|\Psi\rangle$ is the ground state is a sum of terms which are quasi-local, i.e. whose terms involving many spins decay exponentially with the number of spins. So in this way one can find a quasi-local "parent" Hamiltonian for any finite time local unitary evolution of a product state.

OK, this has been a bit formal, and we will move away from formality now. What is the upshot? Well, the main thing to understand is that product-like states are rather stable, and, up to exponential corrections, are what we may expect for many Hamiltonians. Correlations in these states decay exponentially. Soon we will encounter states that do not have purely exponential correlations. This requires an explanation!

4 From Ising model to scalar field theory

So far we have tackled the quantum mechanics of the transverse field Ising model head-on. In general this is a hard model to solve except in some special limits, where we could do perturbation theory. Based on those limits we made some guesses as to general properties, and discussed how that is captured in some of the modern language of quantum information. We could go a lot further with this, but we will turn now to a different language, that of the effective action.

You might well have been a little dissatisfied with the MPS and TNS treatment, since it focuses so much on ground states. We gave hand-waving arguments that connect the ground state to excited states, but this was not very explicit. Can we describe all of this together in one framework? And maybe get a better argument that the simple properties of the FM and PM states of the Ising model hold everywhere except an isolated critical point?

A good step in this direction is to think about a formulation of the full statistical mechanics of the problem in terms of an "effective action". This is a very general way to study quantum systems. It makes an extremely powerful connection between the quantum physics of a d-dimensional model and a classical statistical mechanics problem in $d+1$ dimensions. Let's construct one such relation for our quantum Ising model. I should tell you that there are *many* ways to do this. I am going to choose a somewhat unconventional route, in order to get quickly to a "scalar field theory", which recovers the mean field approximation in a simple way. I'll probably comment on other

ways later.

Consider the quantum partition function,

$$Z = \mathrm{Tr}\left[e^{-\beta H}\right], \tag{65}$$

where H is the quantum Ising Hamiltonian in Eq. (1). From this we can obtain the full thermodynamics of H, and by making small modifications, calculate just about anything else we want. Now the difficulty with this is that we do not know how to diagonalize H, so it is hard to use the eigenbasis for the trace. It is also hard to evaluate the trace in a simple basis like σ^z or σ^z eigenstates, because H contains both types of operators. So there is no really easy way to do the trace. The basic strategy to deal with this is to try to separate the σ^z and σ^z terms, which then each can be dealt with easily. The problem is that, in general, the exponential of a sum of non-commuting operators is not the product of exponentials of the terms of the sum. However, if the terms in the sum are *small*, then this is true to leading order. So what we can do is to break up the exponential into a product of small terms,

$$Z = \mathrm{Tr}\left[\underbrace{e^{-\Delta\tau H}e^{-\Delta\tau H}\cdots e^{-\Delta\tau H}}_{\beta/\Delta\tau \text{ terms}}\right], \tag{66}$$

where we introduce the seemingly odd notation $\Delta\tau$ for a small number. Here we are allowed to break up the exponential since H commutes with itself. Now in the limit $\Delta\tau \to 0$, we can split each exponential into its constituent parts,

$$Z = \lim_{\Delta\tau\to 0} \mathrm{Tr}\left[e^{-\Delta\tau H_X}e^{-\Delta\tau H_Z}e^{-\Delta\tau H_X}e^{-\Delta\tau H_Z}\cdots e^{-\Delta\tau H_X}e^{-\Delta\tau H_Z}\right], \tag{67}$$

with $H_X = -h\sum_i \sigma_i^x$ and $H_Z = -J\sum_{\langle ij\rangle}\sigma_i^z\sigma_j^z$. It is helpful to introduce the exchange matrix J_{ij} which is equal to J when i, j are nearest neighbors and 0 otherwise. Then

$$H_Z = -\frac{1}{2}\sum_{ij} J_{ij}\sigma_i^z\sigma_j^z = -\frac{1}{2}\underline{\sigma^z}\cdot\underline{\underline{J}}\cdot\underline{\sigma^z}, \tag{68}$$

if we want to use matrix notation.

An expression like Eq. (67) is the standard route to constructing a path integral. It is called "Trotterizing" the exponential. Now we have a couple of choices. The most common one is at this point to insert resolutions of the identity between each exponential in the σ_i^z basis. This expresses Z in terms of a large sum over discrete classical Ising spins which are indexed by both i and a new index τ that labels the insertion. We call τ the "imaginary time". The result is an anisotropic classical Ising model in $d+1$ dimensions. This is very direct, but the downside of this approach is that it is not so easy to take the $\Delta\tau \to 0$ limit: it is hard to take a continuum limit of discrete variables.

Lecture 5 (1.5h)
October 8th, 2015

So instead we will make a different choice. We will not insert a resolution

of the identity at all. Instead let's use an identity:

$$e^{\frac{1}{2}\underline{x}\cdot\underline{\underline{M}}\cdot\underline{x}} = \frac{1}{\sqrt{\det\left(2\pi\underline{\underline{M}}\right)}}\int d\underline{y}\, e^{-\frac{1}{2}\underline{y}\cdot\underline{\underline{M}}^{-1}\cdot\underline{y}}\, e^{\underline{x}\cdot\underline{y}}. \tag{69}$$

Here $\underline{x}, \underline{y}$ are n-component vectors, and $\underline{\underline{M}}$ is an $n \times n$ symmetric matrix, and the integral is over each component of y, taken along the entire real axis. Actually we can relax the latter requirement and can take it along any infinite open contour such that the integral converges. In actual applications, we often have to do this implicitly. Don't worry too much about this.[3] This identity is sometimes called a *Hubbard-Stratonovich transformation*. I drew a box around this because it is so extremely useful and important you should memorize it. We can apply it here

$$\begin{aligned} e^{-\Delta\tau H_Z} &= e^{\frac{1}{2}(\Delta\tau\underline{\sigma}^z)\cdot\left(\frac{1}{\Delta\tau}\underline{\underline{J}}\right)\cdot(\Delta\tau\underline{\sigma}^z)} \\ &= \left(\det\left(\frac{2\pi}{\Delta\tau}\underline{\underline{J}}\right)\right)^{-1/2}\int d\underline{\phi}\, e^{-\frac{1}{2}\underline{\phi}\cdot\left(\Delta\tau\underline{\underline{J}}^{-1}\right)\cdot\underline{\phi}}\, e^{\underline{\phi}\cdot\underline{\sigma}^z\Delta\tau} \\ &= \text{Const.}\times\int[d\phi_i]\, e^{-\frac{1}{2}\sum_{ij}J_{ij}^{-1}\phi_i\phi_j\Delta\tau+\sum_i\phi_i\sigma_i^z\Delta\tau}. \end{aligned} \tag{70}$$

Here we used the square bracket symbol to indicate integration over the full set of ϕ_i. Now we use this identity to rewrite each factor of $e^{-\Delta\tau H_Z}$ in Eq. (67). In doing so, we introduce $\beta/\Delta\tau$ separate integrations over a set of ϕ_i. So to distinguish these dummy integration variables, we label them by their position in the product, introducing a τ label: $\phi_i \to \phi_{i\tau}$. Then we have

$$\begin{aligned} Z &= \lim_{\Delta\tau\to 0}\text{Const.}\times\int[d\phi_{i\tau}]\prod_\tau\left(e^{-\frac{1}{2}\Delta\tau\sum_{ij}J_{ij}^{-1}\phi_{i\tau}\phi_{j\tau}}\right) \\ &\quad\times\text{Tr}\left[e^{-\Delta\tau H_X}e^{\Delta\tau\sum_i\phi_{i\tau_n}\sigma_i^z}e^{-\Delta\tau H_X}e^{\Delta\tau\sum_i\phi_{i\tau_{n-1}}\sigma_i^z}\cdots e^{-\Delta\tau H_X}e^{\Delta\tau\sum_i\phi_{i\tau_1}\sigma_i^z}\right], \end{aligned} \tag{71}$$

where $\tau_a = a\Delta\tau$, and $a = 1\cdots n$ with $n = \beta/\Delta\tau$. Now we notice that the sites i have been decoupled, so we can take the trace separately for each i, and taking the $\Delta\tau \to 0$ limit gives

$$Z = Z_0\int[d\phi_i(\tau)]e^{-\int_0^\beta d\tau\,\frac{1}{2}J_{ij}^{-1}\phi_i\phi_j}\prod_i\text{Tr}_i\left[T_\tau e^{-\int_0^\beta d\tau\,\mathcal{H}_i(\tau)}\right], \tag{72}$$

where we defined $\phi_i(\tau) = \phi_{i\tau}$, Z_0 is a constant and

$$\mathcal{H}_i(\tau) = -h\sigma_i^x - \phi_i(\tau)\sigma_i^z. \tag{73}$$

By moving the trace term into the exponential, we can formally write this as

[3] If you do worry, it will probably confuse you a lot, but in the end it can be resolved. It used to bother me a lot when I first encountered this. To really do it right in the physical applications requires a rather clever choice of contours, but in fact it does not matter.

a statistical mechanical-like problem with an "effective action",

$$Z = Z_0 \int [d\phi_{i\tau}] e^{-S[\{\phi_i(\tau)\}]}, \tag{74}$$

with

$$S[\phi] = \int_0^\beta d\tau \, \frac{1}{2} J_{ij}^{-1} \phi_i \phi_j - \sum_i \ln z[\phi_i], \tag{75}$$

with

$$z[\phi] = \mathrm{Tr}\left[T_\tau e^{-\int_0^\beta d\tau \, (-h\sigma^x - \phi(\tau)\sigma^z)} \right]. \tag{76}$$

Formally, at this point, we can declare at least a partial victory. In Eqs. (74-76), we have reformulated the quantum Ising partition function as a functional integral over a scalar variable $\phi_i(\tau)$ on the sites of a discrete lattice and continuous time, weighted by a factor e^{-S}. This is a partially discretized scalar field theory. The action S, however, is a somewhat overly complicated functional of the scalar field. Can we understand it?

We can recognize $z[\phi]$ as a single-spin partition function. This hints at a relation to mean field theory. To pursue this, let us consider the saddle point approximation. The quantity $S[\phi]$ is a functional of the full set of configurations $\phi_i(\tau)$. From Eq. (74) we can see that the configuration with maximal S gives the largest contribution to Z. The saddle point approximation basically consists of approximating the full integral by this leading term. So what is the saddle point? Well it should be found by differentiating, $\delta S / \delta \phi_i(\tau) = 0$. This is a functional derivative. We can obtain it using the chain rule and $\delta \phi_j(\tau') / \delta \phi_i(\tau) = \delta_{ij} \delta(\tau - \tau')$. We find

$$\sum_j J_{ij}^{-1} \phi_j(\tau) - \langle \sigma_i^z(\tau) \rangle = 0, \tag{77}$$

where

$$\langle \sigma_i^z(\tau) \rangle = \frac{\mathrm{Tr}\left[T_\tau \, \sigma_i^z(\tau) e^{-\int_0^\beta d\tau \, \mathcal{H}_i(\tau)} \right]}{z[\phi_i]}. \tag{78}$$

It is really natural to expect that at the saddle point ϕ_i is independent of τ. Then this all simplifies since $\mathcal{H}_i$ becomes τ-independent. Then we see that the expectation value in Eq. (78) is just the standard expectation value of a spin in the single-site Hamiltonian $\mathcal{H}_i$. And moreover from Eq. (77) we get simply

$$\phi_i = \sum_j J_{ij} \langle \sigma_j^z \rangle. \tag{79}$$

So the saddle point ϕ_i has the interpretation of the mean exchange field on site i due to the Ising coupling to other spins. Hence "mean field theory". At

zero temperature, $\beta \to \infty$, the expectation value becomes just the quantum one in the ground state of $\mathcal{H}_i$, so that

$$\langle \sigma_i^z \rangle =_{T=0} \frac{\phi_i}{\sqrt{h^2 + \phi_i^2}}. \tag{80}$$

Using this in Eq. (79) gives a set of self-consistent equations for ϕ_i, whose solution is *identical* to the one we gave in Sec. 2.2 by assuming a product form for the ground state. They are related by the condition $\phi_i = \phi = 2dJm^z$, and $\langle \sigma_i^z \rangle = m^z$ – *Exercise*: check that using this in Eq. (80) one recovers Eq. (34).

This might seem like a really complicated way to redo MFT. But it has the advantage that it is pretty easy to look at the corrections to MFT, by considering fluctuations. From the point of view of Eq. (74), this is just the proper way to approximate the integrals. What we are assuming is that the integral is dominated by the region near the saddle point, i.e. near the mean field solution. So we should write $\phi_i(\tau) = \phi_i^{(0)} + \varphi_i(\tau)$, and expand the action to leading order in φ, and then carry out the integral. So formally we have

$$S[\{\phi_i(\tau)\}] \approx S[\{\phi^{(0)}] + \delta S, \tag{81}$$

with δS of order φ^2 since the first derivative vanishes at the saddle point. To obtain δS we expand Eq. (75) to second order. With a little work, we see that

$$\delta S = \int_0^\beta d\tau \, \frac{1}{2} J_{ij}^{-1} \varphi_i \varphi_j - \frac{1}{2} \sum_i \int d\tau d\tau' \, \left[\langle T_\tau \, \sigma^z(\tau) \sigma^z(\tau') \rangle - \langle \sigma^z \rangle^2 \right] \varphi_i(\tau) \varphi_i(\tau'). \tag{82}$$

Let us for example consider this in the PM phase, where $\phi_i^{(0)} = 0$. Then $\langle \sigma^z \rangle = 0$, and the mean field Hamiltonian is just $\mathcal{H}_i = -h\sigma_i^x$. We have, for $\tau > \tau'$,

$$\begin{aligned} \langle T_\tau \, \sigma^z(\tau) \sigma^z(\tau') \rangle &= \langle \rightarrow | \sigma^z e^{-\mathcal{H}(\tau - \tau')} \sigma^z | \rightarrow \rangle \\ &= e^{-2h(\tau - \tau')}. \end{aligned} \tag{83}$$

So by symmetry the argument of the exponential above is generally $-2h|\tau - \tau'|$. So we find

$$\delta S = \int_0^\beta d\tau \, \frac{1}{2} J_{ij}^{-1} \varphi_i \varphi_j - \frac{1}{2} \sum_i \int d\tau d\tau' \, e^{-2h|\tau - \tau'|} \varphi_i(\tau) \varphi_i(\tau'). \tag{84}$$

Lecture 6 (1.5h)
October 13th, 2015

To go further we need to be a bit more explicit about the matrix J^{-1}. As usual it is defined by the condition that $J^{-1}J = 1$, which in components means

$$\sum_j J_{ij}^{-1} J_{jk} = \delta_{ik}. \tag{85}$$

Now to solve this let us make the ansatz that

$$J_{ij}^{-1} = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)}}{\tilde{J}(\mathbf{k})}. \tag{86}$$

Inserting this in Eq. (85) we have

$$\int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{1}{\tilde{J}(\mathbf{k})} \sum_j J_{jk} e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)} = \delta_{ik}. \tag{87}$$

We carry out the sum over j using the explicit form of J_{jk} which is equal to J when j, k are nearest neighbors, i.e.

$$\int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{1}{\tilde{J}(\mathbf{k})} \sum_{\mu=1}^{d} J \left(e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_k-\hat{x}_\mu)} + e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_k+\hat{x}_\mu)} \right) = \delta_{ik}, \tag{88}$$

which simplifies to

$$\int \frac{d^d\mathbf{k}}{(2\pi)^d} \left[\frac{\sum_{\mu=1}^{d} 2J\cos k_\mu}{\tilde{J}(\mathbf{k})} \right] e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_k)} = \delta_{ik}. \tag{89}$$

We recognize that if the quantity in the square brackets is chosen to be unity, then the left hand side produces the desired Kronecker delta. Hence we conclude that

$$\tilde{J}(\mathbf{k}) = \sum_{\mu=1}^{d} 2J\cos k_\mu. \tag{90}$$

Now it is convenient to Fourier transform the field,

$$\varphi_i(\tau) = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{d\omega}{2\pi} e^{i\mathbf{k}\cdot\mathbf{r}_i + i\omega\tau} \varphi(\mathbf{k},\omega). \tag{91}$$

The factors of 2π in the measure are a nice convention. This is just a change of variables in the functional integral defining the partition function. Plugging this into Eq. (84) we obtain

$$\delta S = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{d\omega}{2\pi} \frac{1}{2} \left[\frac{1}{\tilde{J}(\mathbf{k})} - \frac{4h}{\omega^2 + 4h^2} \right] \varphi(\mathbf{k},\omega)\varphi(-\mathbf{k},-\omega). \tag{92}$$

We are mostly interested in the long-time and long-distance behavior (this is a recurring theme), so it is sensible to expand the quantity in the square brackets in small momentum and frequency. This is essentially the same as

taking a "continuum limit". We obtain

$$\delta S = \int \frac{d^d\mathbf{k}}{(2\pi)^d}\frac{d\omega}{2\pi}\frac{\mathcal{Z}}{2}\left[\omega^2 + v^2k^2 + m^2\right]\varphi(\mathbf{k},\omega)\varphi(-\mathbf{k},-\omega), \tag{93}$$

where

$$m^2 = 4h^2\left(\frac{h}{2dJ} - 1\right), \qquad v^2 = \frac{h^3}{d^2J}, \qquad \mathcal{Z} = \frac{1}{4h^3}. \tag{94}$$

This is the action of a very famous quantum field theory: the relativistic free scalar field. The overall coefficient $\mathcal{Z}$ can be absorbed in a normalization of φ, while v and m have interpretations of the "speed of light" and "mass", respectively, in that context. Here v indeed represents a characteristic velocity, while m represents a *energy gap*, which we can at least see dimensionally since it has dimensions of frequency which is the same (with $\hbar = 1$) as energy in quantum mechanics.

From the point of view of the functional integral, we are to use Eq. (93) for example to calculate the partition function via

$$Z = e^{-S_0}\int [d\varphi(\mathbf{k},\omega)]e^{-\delta S[\varphi]}. \tag{95}$$

Since the quantity in the square brackets in Eq. (93) governs how fast the exponential weight $e^{-\delta S}$ decays with φ, we can say fluctuations around the saddle point are largest when this quantity is smallest. In the long-wavelength low-frequency limit, this is controlled by m. Hence when m vanishes, fluctuations become very large. Inspecting Eq. (94) we see that m vanishes only when $h = 2dJ$ (note: we must take $h \geq 2dJ$ since we assumed that $\langle\sigma_i^z\rangle = 0$). That is, the mass remains non-zero throughout the PM phase and vanishes just at the phase transition to the FM state. The upshot is that we expect that the saddle point approximation and the quadratic expansion is qualitatively good everywhere in the PM phase, and it is only the phase transition point and its features that may require more careful treatment.

On the other hand it is pretty clear that if we do go exactly to the transition point, $m = 0$, then Eq. (93) is inadequate. This is because a constant shift of all φ costs zero action, so integrating over this "center of mass" coordinate will lead to a divergence. It is not physical though, and arises only because we truncated the effective action to quadratic order. To cure it, and write a theory that is valid even up to the critical point, we should include higher order terms in φ (which is equal to ϕ for $h > 2dJ$). We can do this in a couple of ways. The most obvious is just to continue the perturbation theory expansion of $z[\phi]$ to higher order. We'll take a slightly less obvious path. Since the problem occurs for constant φ, we can just assume φ is independent of τ in evaluating $z[\phi = \varphi]$. In that case, Eq. (76) is greatly simplified. We have

$$z[\varphi] = \mathrm{Tr}\left[e^{-\beta(-h\sigma^x - \varphi\sigma^z)}\right]. \tag{96}$$

Since the operator in the exponential is time-independent, we can just diago-

nalize it once, and evaluate the trace in this basis:

$$z[\varphi] = e^{\beta\sqrt{h^2+\varphi^2}} + e^{-\beta\sqrt{h^2+\varphi^2}} \to_{\beta\to\infty} e^{\beta\sqrt{h^2+\varphi^2}}, \tag{97}$$

taking the zero temperature limit. Hence in this limit we obtain, for constant φ,

$$S[\varphi] = \int_0^\beta d\tau \left\{ \frac{1}{2} J_{ij}^{-1} \varphi_i \varphi_j - \sum_i \sqrt{h^2 + \varphi_i^2} \right\}, \qquad \text{for } \partial_\tau \varphi_i = 0. \tag{98}$$

If we perform the same manipulations as before, taking the continuum limit, we obtain to quadratic order *every term* exactly as in Eqs. (93-94), except the ω^2 one, which is of course absent by assumption for constant φ. But we can straightforwardly expand this further in φ, to obtain the quartic correction

$$S_4 = \sum_i \frac{\varphi_i^4}{8h^3} \to \int d^d\mathbf{x}\, \frac{\varphi^4}{8h^3}, \tag{99}$$

with of course still higher order terms of order φ^6 etc. Transforming Eq. (93) back to real space and imaginary time τ, we obtain

$$S = \int d^d\mathbf{x} d\tau \left\{ \frac{\mathcal{Z}}{2} \left[(\partial_\tau \varphi)^2 + v^2 (\nabla\varphi)^2 + m^2 \varphi^2 \right] + u\varphi^4 + \cdots \right\}, \tag{100}$$

where we defined $u = \frac{1}{8h^3}$. This is the famous φ^4 or scalar field theory. It is believed to describe properly all the long-distance and long-time properties of the quantum Ising model, both away from and arbitrarily close to its transition point. It also describes the long-distance properties of the classical Ising model in $d+1$ dimensions, in a similar fashion. You will study this thing, perhaps *ad nauseum*, in Physics 220.

5 Stuff that can be measured

5.1 *Types of measurements*

So we talked about both fundamental physics and formalism. What do we *do* with any of this? Ultimately we would like to address experiments. One of the unique aspects of condensed matter physics is the amazing variety of experimental probes, which is both revealing – because you can learn so very much about a system – and bewildering – because it is hard to know what all these different measurements really tell you. While this is not supposed to a course about condensed matter phenomena (we have 223ab for that), you should understand a bit about how many body theory connects to experiment. I'll try to divide the types of experiments into different categories:

1. **Thermodynamics:** There are measurements that study extensive, equilibrium properties. This includes heat capacity, magnetization, and com-

pressibility (or pressure versus volume). All these quantities can be calculated from the free energy or partition function. Consequently they are very fundamental but also have very little specific information.

2. **Correlations:** A different sort of measurement is to just "watch" a system somewhere. We can study different local operators, like our spins σ_i^μ in the Ising model. We might measure the spin at some site at some time, and then another one at another position at the same time or a later time. The outcome of any one such measurement, according to quantum mechanics, is actually stochastic, and we can only predict the probabilities of possible outcomes. However, we often do many such measurements in parallel, and really look at the average. This is a correlation function. These are very natural objects theoretically.

3. **Spectra:** Some measurements specifically identify energies of excited states, or rather the energies of transitions between states. Obviously the Lyman spectra of atoms is a canonical example. In many-body systems, there is an awfully large set of levels, and the existence of thermodynamics tells us that in principle there are energy levels spread over a range of energies of order volume. But it is hard to impart that much energy, and you cannot measure it precisely. So we reserve spectroscopic measurements to discuss energy changes that are order one, i.e. finite in the thermodynamic limit, or in other words, transitions which do not change the energy density. Such measurements can be made by "scattering" type measurements, or by adding or removing electrons from a system. Or they might be made by turning on a classical field which is resonant with some excitation. A priori a measurement of spectra is very clean to think about theoretically. We just need to find the low-lying energy levels.

4. **Linear Response:** Instead of just watching a system, we can perturb it just a bit. If our perturbation is weak, then we can use time-dependent perturbation theory to see how it responds. Usually the response to a weak perturbation is linear. That might fail if we take an extreme limit: $T \to 0$, $\omega \to 0$ *and* $\mathbf{k} \to 0$ simultaneously – this is because in that limit you probe macroscopic response of an infinite number of coherent degrees of freedom, and that infinity can induce a divergence, i.e. a non-linear response. But usually linear response is good. Some scattering measurements can be understood as linear response, if the thing which we scatter off the solid interacts weakly with it. In addition, measurements like conductivity or susceptibility are linear response probes. There is a general theoretical framework for relating linear responses to correlation functions, which we can try to calculate theoretically.

5. **Real time dynamics:** We can also try to monitor some property of a system continuously in time, usually in combination with driving the system somehow either continuously or with some pulse. This for example is what one does in nuclear magnetic resonance (NMR) or muon spin resonance (μSR). Lots of measurements in ultra-cold atoms are like this.

These are some of the most complex experiments to try to think about. The full quantum dynamics of a many-body system is hard!!! Unless there is some simplification (which there usually is in NMR for example), we will not be able to say much that is general, and it is difficult to do real calculations.

Somewhat surprisingly, the second, third, and fourth types of measurements are all related.

Lecture 7 (1.5h)
October 15th, 2015

5.2 *Correlation functions*

We can consider a variety of correlation functions. The different definitions can be a bit bewildering sometimes. Don't worry too much about this: all of them are pretty similar, and it is not so important, at least initially, to know all the differences. So many different ones have been defined just because they are used so often, it can be convenient. Unfortunately, conventions are not entirely uniform, so you should be a bit careful what people are talking about, when it comes down to the details.

By principles of statistical mechanics, the expected value of any quantity is

$$\langle \mathcal{Q} \rangle = \frac{1}{Z} \mathrm{Tr} \left(e^{-\beta H} \mathcal{Q} \right). \tag{101}$$

A correlation function is just the case where $\mathcal{Q}$ is a product of local operators, which may be measured at different times

$$C(\mathbf{x} - \mathbf{x}', t - t') = \langle \mathcal{O}(\mathbf{x}, t) \mathcal{O}(\mathbf{x}', t') \rangle, \tag{102}$$

where the spatial argument indicates that the operator has support only in the region around that coordinate, and, as usual, time evolution proceeds according to

$$\mathcal{O}(x, t) = e^{iHt} \mathcal{O}(\mathbf{x}) e^{-iHt}. \tag{103}$$

For a translationally invariant system, the correlation function depends only on the space and time differences, as indicated in Eq. (102). So Eq. (102) defines a simple "real time" correlation function, i.e. it reflects an actual time evolution of a quantum system.

Another popular real time correlation function is defined in terms of the commutator

$$D_{\mathrm{ret}}(\mathbf{x} - \mathbf{x}', t - t') = -i \langle \left[\mathcal{O}(\mathbf{x}, t), \mathcal{O}(\mathbf{x}', t') \right] \rangle \Theta(t - t'). \tag{104}$$

Here I included several conventional factors: the $-i$ which helps to make it real, since the commutator of Hermitian operators changes sign under complex conjugation, and multiplied it by the Θ function to make it non-zero only for $t > t'$. We define it this way because in this form it arises in linear response theory. It is called the retarded correlation function.

A third useful correlation function which we will often encounter is the imaginary time one. What occurs most often is the *imaginary time ordered* correlation function,

$$C(\mathbf{x}-\mathbf{x}',\tau-\tau') = \langle T_\tau\, \mathcal{O}(\mathbf{x},\tau)\mathcal{O}(\mathbf{x}',\tau')\rangle, \tag{105}$$

where

$$\mathcal{O}(\mathbf{x},\tau) = e^{H\tau}\mathcal{O}(\mathbf{x})e^{-H\tau}. \tag{106}$$

and T_τ is the time-ordering symbol, so that, explicitly,

$$C(\mathbf{x}-\mathbf{x}',\tau-\tau') = \begin{cases} \langle \mathcal{O}(\mathbf{x},\tau)\mathcal{O}(\mathbf{x}',\tau')\rangle & \tau > \tau' \\ \langle \mathcal{O}(\mathbf{x}',\tau')\mathcal{O}(\mathbf{x},\tau)\rangle & \tau < \tau' \end{cases} \tag{107}$$

The imaginary time-ordering is not especially physical, but it is what is "spit out" by the path integral prescription. Indeed this is so much the case that we will often omit the T_τ in writing Eq. (105).

5.3 *Spectral representations*

Real time correlation function

Quantum mechanics allows us to write correlation functions in terms of eigenstates, at least as a matter of principle. For example, consider the real time correlation function in Eq. (102). Making the time evolution explicit, we have

$$C(\mathbf{x},t) = \langle e^{iHt}\mathcal{O}(\mathbf{x})e^{-iHt}\mathcal{O}(0)\rangle. \tag{108}$$

Then we can evaluate the trace in the eigenbasis, and use a resolution of the identity in the middle of the expression in the same basis to get

$$C(\mathbf{x},t) = \frac{1}{Z}\sum_{m,n}\langle m|\mathcal{O}(\mathbf{x})|n\rangle\langle n|\mathcal{O}(0)|m\rangle e^{-i(E_n-E_m)t}e^{-\beta E_m}. \tag{109}$$

You can see that the energy level differences appear as complex exponentials. By Fourier transforming we can single out specific frequencies,

$$C(\mathbf{x},\omega) = \int dt\, e^{i\omega t}C(\mathbf{x},t) = \frac{1}{Z}\sum_{m,n}\langle m|\mathcal{O}(\mathbf{x})|n\rangle\langle n|\mathcal{O}(0)|m\rangle e^{-\beta E_m}\delta(\omega-(E_n-E_m)). \tag{110}$$

So if we can calculate the Fourier transform of the correlation function, we can in principle extract at least energy level differences. In translationally invariant systems, we can make this a bit nicer. Let us define a translation operator T_x, so that

$$\mathcal{O}(\mathbf{x}) = \mathsf{T}_x^{-1}\mathcal{O}(0)\mathsf{T}_x. \tag{111}$$

Provided $\mathsf{T}_x^{-1} H \mathsf{T}_x = H$ (which defines translational invariance), we can choose the energy levels to be eigenstates of T_x, and introduce this label explicitly, $|n\rangle \to |\mathbf{k}, n\rangle$, where

$$\mathsf{T}_x |\mathbf{k}, n\rangle = e^{i\mathbf{k}\cdot\mathbf{x}} |\mathbf{k}, n\rangle. \tag{112}$$

Making these substitutions in Eq. (110), we obtain

$$C(\mathbf{x}, \omega) = \frac{1}{Z} \sum_{m,n} \sum_{\mathbf{q},\mathbf{p}} \langle \mathbf{p}, m | \mathcal{O}(0) | \mathbf{q}, n \rangle \langle \mathbf{q}, n | \mathcal{O}(0) | \mathbf{p}, m \rangle e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}} e^{-\beta E_m} \delta(\omega - (E_n - E_m)). \tag{113}$$

Then, Fourier transforming in space as well,

$$\begin{aligned} C(\mathbf{k}, \omega) &= \int d^d\mathbf{x}\, C(\mathbf{x}, \omega) e^{-i\mathbf{k}\cdot\mathbf{x}} \\ &= \frac{1}{Z} \sum_{m,n} \sum_{\mathbf{q},\mathbf{p}} |\langle \mathbf{q}, n | \mathcal{O}(0) | \mathbf{p}, m \rangle|^2 \, e^{-\beta E_m} (2\pi)^d \delta(\mathbf{k} - (\mathbf{q} - \mathbf{p})) \delta(\omega - (E_n - E_m)). \end{aligned} \tag{114}$$

Things are particularly nice at zero temperature, where the sum over m reduces to just the ground state. Typically we can assume it has $\mathbf{p} = 0$. Then

$$C(\mathbf{k}, \omega) = \sum_n \sum_{\mathbf{q}} |\langle \mathbf{q}, n | \mathcal{O}(0) | 0 \rangle|^2 \, (2\pi)^d \delta(\mathbf{k} - \mathbf{q}) \delta(\omega - \epsilon_n), \tag{115}$$

where $\epsilon_n = E_n - E_0$ is the excitation energy of the state n. Here we can see that the result is just the sum of delta functions that select states with a given excitation energy and a given momentum, weighted by a *positive* function which is just the square of a matrix element. This is called a *spectral representation* and the weight is a *spectral weight*.

For future use, we can obtain a spectral representation slightly differently, by Fourier transforming Eq. (102) directly in both coordinates. One obtains

$$C(\mathbf{k}, \omega) = \frac{1}{VZ} \sum_{m,n} |\langle m | \mathcal{O}_k | n \rangle|^2 \, e^{-\beta E_m} \delta(\omega - E_n + E_m), \tag{116}$$

where V is the volume. This is of course equivalent to Eq. (114), which one can see by expressing $\mathcal{O}(\mathbf{x} = 0) = \int d^d\mathbf{k} / (2\pi)^d \mathcal{O}_\mathbf{k}$ and using a little algebra.

Retarded correlation function

Let's write a spectral representation for the retarded correlation function in Eq. (104). We have

$$D(\mathbf{x} - \mathbf{x}', t - t') = \frac{-i}{Z} \sum_{m,n} \langle m | \mathcal{O}(x) | n \rangle \langle n | \mathcal{O}(x') | n \rangle \left(e^{-\beta E_m} - e^{-\beta E_n} \right) e^{-i(E_n - E_m)(t-t')} \Theta(t - t'). \tag{117}$$

We get this by writing out the commutator and inserting complete sets of states in each of the two terms. Now let's Fourier transform with respect to both $\mathbf{x}$ and $\mathbf{x}'$ (i.e. multiply by $e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}$ and integrate over both coordinates. We obtain

(118)
$$VD(\mathbf{k}, t-t') = \frac{-i}{Z}\sum_{m,n} |\langle m|\mathcal{O}_{\mathbf{k}}|n\rangle|^2 \left(e^{-\beta E_m} - e^{-\beta E_n}\right) e^{-i(E_n-E_m)(t-t')}\Theta(t-t'),$$

where V is the volume. Fourier transforming now with respect to time, we add an infinitesimal factor $e^{-\delta(t-t')}$ with $\delta = 0^+$ to guarantee convergence. This gives

$$D(\mathbf{k},\omega) = \frac{1}{VZ}\sum_{m,n} |\langle m|\mathcal{O}_{\mathbf{k}}|n\rangle|^2 \frac{\left(e^{-\beta E_m} - e^{-\beta E_n}\right)}{\omega - (E_n - E_m) + i\delta}. \tag{119}$$

Now we get a spectral representation by taking the imaginary part,

$$-\frac{1}{\pi}\mathrm{Im}\, D(\mathbf{k},\omega) = \frac{1}{VZ}\sum_{m,n} |\langle m|\mathcal{O}_{\mathbf{k}}|n\rangle|^2 \left(e^{-\beta E_m} - e^{-\beta E_n}\right)\delta(\omega - E_n + E_m). \tag{120}$$

So the imaginary part of the retarded correlation function is directly related to the spectrum of the system, i.e. it is non-zero only a frequencies that correspond to possible transitions between levels. We'll see why this is the case soon. In the zero temperature limit this simplifies to

$$-\frac{1}{\pi}\mathrm{Im}\, D(\mathbf{k},\omega) =_{T=0} \frac{1}{V}\sum_{n} |\langle 0|\mathcal{O}_{\mathbf{k}}|n\rangle|^2 \,\mathrm{sign}(\omega)\delta(|\omega| - \epsilon_n). \tag{121}$$

By comparing Eq. (120) to Eq. (116), we can relate the ordinary and retarded correlation functions. Specifically by using the delta function in Eq. (120) to write $E_n = E_m + \omega$, we obtain

$$\boxed{-\frac{1}{\pi}\mathrm{Im}\, D(\mathbf{k},\omega) = \left(1 - e^{-\beta\omega}\right) C(\mathbf{k},\omega).} \tag{122}$$

We will see in the next section that D describes the linear response of the system to a perturbation of a force, and hence the work done by the perturbation, i.e. dissipation. Hence this relation is considered a *fluctuation-dissipation* relation – it relates a linear response quantity describing dissipation (D) to one which describes fluctuations (C).

Imaginary time correlation function

So finally, we can look at the Imaginary time correlation function in Eq. (105). First let us be explicit and take $\tau' = 0$, $\tau > 0$. Then

$$C(\mathbf{x}, \tau) = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H} e^{\tau H}\mathcal{O}(\mathbf{x}) e^{-\tau H}\mathcal{O}(\mathbf{0})\right) \tag{123}$$

From this, note that if we take $\tau = \beta$,

$$C(\mathbf{x}, \beta) = \frac{1}{Z}\mathrm{Tr}\left(\mathcal{O}(\mathbf{x}) e^{-\beta H}\mathcal{O}(\mathbf{0})\right) = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}\mathcal{O}(\mathbf{0})\mathcal{O}(\mathbf{x})\right) = C(\mathbf{x}, \tau = 0^-), \tag{124}$$

using cyclicity of the trace. So the imaginary time correlation function is a β-periodic function of τ. This means that it has discrete Fourier components $\omega_n = \frac{2\pi}{\beta} n$, with integer n. These are called "Matsubara" frequencies. With that knowledge we can proceed with the spectral representation. Getting tired of it yet??

We multiply by $e^{i\omega_n \tau}$ and integrate τ from 0 to β:

$$C(\mathbf{x}, \omega_n) = \int_0^\beta e^{i\omega_n \tau} C(\mathbf{x}, \tau) = \frac{1}{Z}\sum_{m,n} \langle m|\mathcal{O}(\mathbf{x})|n\rangle\langle n|\mathcal{O}(0)|m\rangle e^{-\beta E_m} \frac{1 - e^{-\beta(E_n - E_m - i\omega_n)}}{E_n - E_m - i\omega_n}. \tag{125}$$

I apologize for the two different uses of n – in the subscript of the Matsubara frequency and to label states. It should be clear which is which from context. In momentum space, this becomes

$$C(\mathbf{k}, \omega_n) = \frac{1}{VZ}\sum_{m,n} |\langle m|\mathcal{O}_{\mathbf{k}}|n\rangle|^2 \frac{e^{-\beta E_m} - e^{-\beta E_n}}{E_n - E_m - i\omega_n}, \tag{126}$$

where I used $e^{i\beta\omega_n} = 1$ which is true for discrete Matsubara frequencies, and multiplied through the $e^{-\beta E_m}$ factor. Comparing Eq. (126) to Eq. (119), we see that there is a very simple relation:

$$C(\mathbf{k}, \omega_n) \to_{i\omega_n \to \omega + i\delta} -D(\mathbf{k}, \omega). \tag{127}$$

This is a useful result because it means that we can calculate the Matsubara correlation function (which is what we naturally obtain using path integrals, for example) and then by carrying out the "analytic continuation" in Eq. (127), we obtain the retarded correlation function. And the retarded one is quite physical, as we will soon see. Unfortunately, if one does not have an explicit form for the Matsubara correlation function, it can be difficult to carry out this analytic continuation (think about it!!).

Summary

We can obtain spectral representations of any correlation function we want, which basically give them as a weighted sum of matrix elements of exact eigenstates, organized by momentum and energy. The spectral weights, written this way, have definite signs, which reflect the positivity of squared matrix elements. This means that such correlation functions are not arbitrary functions. They are special because they arise from unitary time evolution. Using the spectral representation we can see directly how to relate various different correlation functions.

5.4 *Linear response*

Let us think about the response of our system to some perturbation. This consists of starting with some initial Hamiltonian H (which we assume is time-independent) and adding a perturbation $\delta H(t)$, which can be time dependent. We suppose $\delta H(t) = 0$ for $t < t_0$. Then we can consider the time evolution operator from time t_0 to some later time t, which is a time-ordered exponential:

$$U(t, t_0) = T e^{-i \int_{t_0}^{t} dt' \, (H + \delta H(t'))} \tag{128}$$

Since the perturbation is small, we can try to expand it in δH. This is a little tricky because of the time ordering, but we can do it by breaking the time-ordered exponential into a Trotter product, expanding each piece, and recollecting them. We obtain

$$U(t, t_0) = e^{-iH(t-t_0)} - i \int_{t_0}^{t} dt' \, e^{-iH(t-t')} \delta H(t') e^{-iH(t'-t_0)} + O(\delta H^2). \tag{129}$$

Now we can consider the time-dependent expectation value of some operator

$$\langle \mathcal{O}(t) \rangle = \langle U(t_0, t) \mathcal{O} U(t, t_0) \rangle. \tag{130}$$

Inserting Eq. (129), we obtain, to linear order

$$\begin{aligned} \mathcal{O}(t) \quad = \quad & e^{iH(t-t_0)} \mathcal{O} e^{-iH(t-t_0)} \\ & -i \int_{t_0}^{t} dt' \left[e^{iH(t-t_0)} \mathcal{O} e^{-iH(t-t')} \delta H(t') e^{-iH(t'-t_0)} - e^{iH(t'-t_0)} \delta H(t') e^{iH(t-t')} \mathcal{O} e^{-iH(t-t_0)} \right] + O(\delta H^2), \end{aligned} \tag{131}$$

which can be recollected by defining operators whose the time evolution is with respect to H not including δH, as in Eq. (103) – this is called the "interaction picture" in textbooks:

$$\mathcal{O}^{(0)}(t) = e^{iH(t-t_0)} \mathcal{O} e^{-iH(t-t_0)}, \tag{132}$$

etc. Then we obtain from Eq. (131) and Eq. (132), taking the expectation value,

$$\langle \mathcal{O}(t) \rangle = \langle \mathcal{O} \rangle - i \int_{t_0}^{t} dt' \, \langle \Big[\mathcal{O}^{(0)}(t), \delta H^{(0)}(t') \Big] \rangle + O(\delta H^2). \tag{133}$$

The right hand side expresses the result of the perturbation explicitly in terms of time-dependent correlation functions of the unperturbed system.

Lecture 8 (1.5h)
October 20th, 2015

Now let us take a specific form of the perturbation. Let's assume

$$\delta H(t) = \int d^d\mathbf{x}' \, h(\mathbf{x}', t) \mathcal{O}(\mathbf{x}'). \tag{134}$$

Here I indicate only the explicit time-dependence, not any Heisenberg dependence. Then we find

$$\delta \langle \mathcal{O}(\mathbf{x}, t) \rangle = -i \int_{t_0}^{t} dt' \int d^d\mathbf{x}' \, \langle \Big[\mathcal{O}^{(0)}(\mathbf{x}, t), \mathcal{O}^{(0)}(\mathbf{x}', t') \Big] \rangle h(\mathbf{x}', t'). \tag{135}$$

If we assume $h(\mathbf{x}', t') = 0$ for $t' < t_0$, we can extend the lower limit fo the integration to minus infinity, and extend the upper limit to plus infinity by adding a Theta function to the integrand. The result is

$$\delta \langle \mathcal{O}(\mathbf{x}, t) \rangle = \int d^d\mathbf{x}' dt \, D(\mathbf{x} - \mathbf{x}', t - t') h(\mathbf{x}', t'). \tag{136}$$

This is the central result of linear response. It shows that the linear change of an expectation value with respect to a perturbation is given by the retarded correlation function of the operator which is observed and the one which constitutes the perturbation. In Fourier space, it is simply

$$\boxed{\delta \langle \mathcal{O}(\mathbf{k}, \omega) \rangle = D(\mathbf{k}, \omega) h(\mathbf{k}, \omega).} \tag{137}$$

5.5 *Example application to Ising model and ϕ^4 theory*

OK, that's a lot of technical stuff. Let's go back to our field theory of the quantum Ising model. We saw in Sec. 2 that the low energy excitations of the two phases, away from the QCP, are quasiparticles. I claim this is captured by the scalar field theory. Now we have learned how to extract excitations from correlation functions, so we are ready to see what excitations are actually encoded in the field theory.

Relation to neutron scattering

We are going to examine the correlation functions of the field φ, in the scalar field theory, which we can interpret using the formalism of the previous subsection. This is a nice exercise. But it is not *just* an exercise. It essentially gives the correlation function of the spin operator σ^z of the original Ising model.

That, in turn, is actually just what is measured in one of the classic experimental probes of magnetism: neutron scattering. The neutron interacts with electrons through its magnetic moment, which is coupled linearly to any magnetic moments in the material. A Fermi's Golden rule treatment of this interaction gives the rate to scatter a neutron with a change of momentm $\mathbf{k}$ and change of energy E as

$$\frac{d^2\sigma}{dEd\Omega} \propto \sum_{\alpha\beta} \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2}\right) S_{\alpha\beta}(\mathbf{k}, \omega = E), \tag{138}$$

where

$$S_{\alpha\beta}(\mathbf{k}, \omega) = \int d^d\mathbf{x} dt\, e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} \langle M_\alpha(\mathbf{x}, t) M_\beta(\mathbf{0}, 0)\rangle, \tag{139}$$

which is nothing but a real-time correlation of the form in Eq. (108), with M_α the magnetization operator. One can rationalize this by remembering that Fermi's golden rule give a rate proportional to the square of a matrix element, which is precisely the form that the spectral representation, Eq. (115) gives. The prefactor is the result of the dipole-dipole interaction. We see that $S(k, \omega)$, which is called the inelastic *structure factor*, gives a direct measurement of the spectrum of excitation created by the magnetization operator. In a simple spin system, M_α is just proportional to the spin operator, so that the neutron structure factor is determined from the $\sigma^z - \sigma^z$ correlation function.

Gaussian integrals and Wick's theorem

Let's start with the free field theory, $u = 0$, in Eq. (93). The action is quadratic at this level, so we can calculate with it. To do so, we need to quickly talk again about Gaussian integrals. Recall Eq. (69). Write two equalities by taking $\underline{\underline{M}} = \underline{\underline{K}}^{-1}$, once with $\underline{x} = 0$ and once with $\underline{x}$ unchanged, and take the ratio. One obtains

$$\left\langle e^{\underline{x}\cdot\underline{y}} \right\rangle_y \equiv \frac{\int d\underline{y} e^{-\frac{1}{2}\underline{y}\cdot\underline{\underline{K}}\cdot\underline{y}} e^{\underline{x}\cdot\underline{y}}}{\int d\underline{y} e^{-\frac{1}{2}\underline{y}\cdot\underline{\underline{K}}\cdot\underline{y}}} = e^{\frac{1}{2}\underline{x}\cdot\underline{\underline{K}}^{-1}\cdot\underline{x}}. \tag{140}$$

This gives the expectation value for an arbitrary exponential of a linear function of the field (the integration variable, in this case $\underline{y}$). We can get from it any polynomial expectation value, by just Taylor expanding both sides, and equating coefficients of x_i. You can think of the expectation value of the exponential thus as a *generating function* for the correlators of $\underline{y}$. Most fundamentally,

$$\langle y_i y_j \rangle = (K^{-1})_{ij} \equiv G_{ij}. \tag{141}$$

Higher orders are also straightforward, for example

$$\langle y_i y_j y_k y_l \rangle = G_{ij}G_{kl} + G_{ik}G_{jl} + G_{il}G_{jk}. \tag{142}$$

One can see that this result can be obtained by taking the sum of all possible partitions of the y variables into pairs, and for each partition replacing the pair by its fundamental correlation G via Eq. (141). This rule is called "Wick's theorem".

Free field correlations

OK, that's all we need. Now let us consider the basic correlation function in the scalar field theory:

$$C_{\varphi\varphi}(\mathbf{x}-\mathbf{x}',\tau-\tau') = \langle\varphi(\mathbf{x},\tau)\varphi(\mathbf{x}',\tau')\rangle \equiv \frac{\int[d\varphi]e^{-S[\varphi]}\varphi(\mathbf{x},\tau)\varphi(\mathbf{x}',\tau')}{\int[d\varphi]e^{-S[\varphi]}}. \tag{143}$$

We can use Eq. (141) to calculate this. Here K is the kernel of the quadratic form in the scalar theory. It is simplest in Fourier space, Eq. (93). From this, we can read off the Fourier transform

$$C_{\varphi\varphi}(\mathbf{k},\omega_n) = \frac{1}{\mathcal{Z}}\frac{1}{\omega_n^2+v^2k^2+m^2}. \tag{144}$$

Now in the quantum theory, Eq. (143) should be interpreted as the imaginary time Green's function, Eq. (105), whose Fourier transform is the Matsubara Green's function, Eq. (126). So we can obtain the retarded Green's function using Eq. (127), which gives

$$D_{\varphi\varphi}(\mathbf{k},\omega) = \frac{1}{\mathcal{Z}}\frac{1}{(\omega+i\delta)^2-v^2k^2-m^2} = \frac{1}{\mathcal{Z}}\frac{1}{\omega^2-v^2k^2-m^2+i\mathrm{sign}(\omega)\delta}, \tag{145}$$

where I use δ to always indicate a positive infinitesimal, but not any specific value. The spectral function is obtained from the Imaginary part, Eq. (121), which gives

$$-\frac{1}{\pi}\mathrm{Im}D(\mathbf{k},\omega) = \frac{\mathrm{sign}(\omega)}{\mathcal{Z}}\delta(\omega^2-v^2k^2-m^2) = \frac{\mathrm{sign}(\omega)}{\mathcal{Z}}\frac{\delta(|\omega|-\sqrt{m^2+v^2k^2})}{2\sqrt{m^2+v^2k^2}}. \tag{146}$$

We see that there is an excitation with a sharply defined energy $\epsilon(k) = \sqrt{m^2+v^2k^2}$. The coefficient of the delta-functions gives the matrix element in Eq. (121), i.e. the overlap of the state $\varphi|0\rangle$ with the exact excitation ($|n\rangle$ in the formal expression). The coefficient is called the "quasiparticle weight".

Perturbation expansion

Now it is interesting to compute some corrections when the $u\varphi^4$ term is included. The idea is simple: start with Eq. (143) and expand both numerator and denominator in a series in u. Let's start with the denominator, which is

the partition function itself:

(147)
$$Z = \int [d\varphi] e^{-S[\varphi]} = \int [d\varphi] e^{-S_0[\varphi]} \left(1 - u \int d^d\mathbf{x} d\tau\, \varphi^4 + \frac{u^2}{2} \int d^d\mathbf{x} d^d\mathbf{x}' d\tau d\tau'\, \varphi^4_{x,\tau} \varphi^4_{x',\tau'} + \cdots \right),$$

where $S_0 = S(u = 0)$ is the quadratic part of the action. By dividing by the first term, it becomes

(148)
$$Z = Z_0 \left(1 - u \int d^d\mathbf{x} d\tau \left\langle \varphi^4_{x,\tau} \right\rangle_0 + \frac{u^2}{2} \int d^d\mathbf{x} d^d\mathbf{x}' d\tau d\tau' \left\langle \varphi^4_{x,\tau} \varphi^4_{x',\tau'} \right\rangle_0 + \cdots \right),$$

where the average is with respect to S_0. We can evaluate this if we want, but it will actually not be necessary. Let us instead look at the expansion of the numerator. In a similar way, we have

$$N = \int [d\varphi] e^{-S[\varphi]} (\varphi_{x,\tau} \varphi_{x',\tau'}) = Z_0 \Big(\langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0$$
$$- u \int d^d\mathbf{x}_1 d\tau_1 \left\langle \varphi_{x,\tau} \varphi_{x',\tau'} \varphi^4_{x_1,\tau_1} \right\rangle_0 + \frac{u^2}{2} \int d^d\mathbf{x}_1 d^d\mathbf{x}'_2 d\tau_1 d\tau'_2 \left\langle \varphi_{x,\tau} \varphi_{x',\tau'} \varphi^4_{x_1,\tau_1} \varphi^4_{x_2,\tau_2} \right\rangle_0 \Big)$$
(149)

Now we can take the ratio, and expand expand the full expression consistently to second order

(150)
$$\begin{aligned} \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle &= \frac{N}{Z} \\ &= \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 - u \int_1 \left(\left\langle \varphi_{x,\tau} \varphi_{x',\tau'} \varphi^4_1 \right\rangle_0 - \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 \left\langle \varphi^4_1 \right\rangle_0 \right) \\ &+ \frac{u^2}{2} \int_{1,2} \Big(\left\langle \varphi_{x,\tau} \varphi_{x',\tau'} \varphi^4_1 \varphi^4_2 \right\rangle_0 + 2 \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle \left\langle \varphi^4_1 \right\rangle \left\langle \varphi^4_2 \right\rangle_0 \\ &- \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle \left\langle \varphi^4_1 \varphi^4_2 \right\rangle_0 - 2 \left\langle \varphi_{x,\tau} \varphi_{x',\tau'} \varphi^4_1 \right\rangle \left\langle \varphi^4_2 \right\rangle_0 \Big), \end{aligned}$$

where we have abbreviated x_1, τ_1 by 1 and similarly abbreviated the integral measures. Let us look just at the term linear in u. Using Wick's theorem,

$$\begin{aligned} &\langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle = \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 \\ &- u \int_1 \left(\langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 \left\langle \varphi^4_1 \right\rangle_0 + 12 \langle \varphi_{x,\tau} \varphi_1 \rangle_0 \langle \varphi_{x',\tau'} \varphi_1 \rangle_0 \left\langle \varphi^2_1 \right\rangle_0 - \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 \left\langle \varphi^4_1 \right\rangle_0 \right) + O(u^2) \\ &= \langle \varphi_{x,\tau} \varphi_{x',\tau'} \rangle_0 - 12u \int_1 \langle \varphi_{x,\tau} \varphi_1 \rangle_0 \langle \varphi_{x',\tau'} \varphi_1 \rangle_0 \left\langle \varphi^2_1 \right\rangle_0 + O(u^2). \end{aligned}$$
(151)

Note that after we applied Wick's theorem, the expression simplified quite a bit, due to cancellations between the numerator and denominator. This type of cancellation is general and has a graphical interpretation, which we will come to. We can rewrite this result in terms of the "bare" correlation function in Eq. (143), which we will now call C_0:

(152)
$$C(x-x',\tau-\tau') = C_0(x-x',\tau-\tau') - 12uC(0,0)^2 \int_1 C_0(x-x_1,\tau-\tau_1)C_0(x_1-x',\tau_1-\tau') + O(u^2).$$

One can recognize the second term as a convolution, which tells us that Fourier transform will simplify it: the Fourier transform of a convolution is the product of Fourier transforms. Hence

(153) $$C(\mathbf{k},\omega_n) = C_0(\mathbf{k},\omega_n) - 12uC(0,0)^2 C_0(\mathbf{k},\omega_n)^2 + O(u^2).$$

This can be rearranged in an a priori funny but standard way:

$$\begin{aligned} C(\mathbf{k},\omega_n) &= C_0(\mathbf{k},\omega_n)\left[1 - 12uC(0,0)^2 C_0(\mathbf{k},\omega_n)\right] \\ &= C_0(\mathbf{k},\omega_n)\frac{1}{1+12uC(0,0)^2 C_0(\mathbf{k},\omega_n)} \\ &= \frac{1}{C_0^{-1}(\mathbf{k},\omega_n) + 12uC(0,0)^2} \equiv \frac{1}{C_0^{-1}(\mathbf{k},\omega_n)+\Sigma}, \end{aligned} \tag{154}$$

where we have ensured each line is equivalent to previous one only to order u, and defined the "self energy"

(155) $$\Sigma = 12uC(0,0)^2.$$

Formally, we can always *define* Σ so that the left hand side and final form on the right hand side are *exactly* equal.

(156) $$\boxed{C(\mathbf{k},\omega_n) \equiv \frac{1}{C_0^{-1}(\mathbf{k},\omega_n) + \Sigma(\mathbf{k},\omega_n)}}.$$

So it is perfectly fine to trade the exact correlation function for an exact self-energy. It turns out that the perturbation series for the self-energy is simpler than that for the correlation function itself. This is another simplification on top of the cancellations already observed between the numerator and denominator.

To linear order in u, we obtained a rather trivial result: the self-energy is a constant, independent of frequency and momentum. It can therefore be interpreted as a *renormalization* of the mass m^2, i.e. to this order

(157) $$C(\mathbf{k},\omega_n) = \frac{1}{\mathcal{Z}}\frac{1}{\omega^2 + v^2k^2 + m^2 + \Delta m^2},$$

with $\Delta m^2 = 12uC(0,0)^2/\mathcal{Z} > 0$. This means that the energy gap of the quasi-particles is increased by the correction. We can associate the point at which the energy gap vanishes with the QCP. This indicates that the location of the QCP is shifted, in such a way that the domain of the PM phase is increased relative to the free theory, i.e. relative to MFT. This makes physical sense: quantum fluctuations suppress the tendency to order.

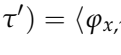
Lecture 9 (1.5h)
October 22nd, 2015

Second order terms and a quick primer on diagrammatics

A non-trivial self-energy occurs at $O(u^2)$. It becomes helpful to introduce the notion of diagrams. The idea is as follows. Every polynomial in φ can be thought of as a "vertex" out of which several "lines" emerge, each line indicating a factor of φ,

(158) $\varphi^2_{x,\tau} =$  , $\varphi^4_{x,\tau} =$  , etc.

We can use the vertex to indicate fields that share a point in space-time. Then diagramatically, the Green's function is

(159) $C_0(x-x',\tau-\tau') = \langle\varphi_{x,\tau}\varphi_{x',\tau'}\rangle_0 =$  .

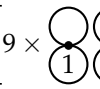
For an expectation value in a free (quadratic) theory, the rules of Wick's theorem apply. This is where the diagrams become useful. Wick's theorem basically says that the result of an expectation value is given by the sum of all ways of connecting the lines in pairs. For two fields this is trivial,

(160) $C_0(x-x',\tau-\tau') =$,

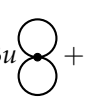
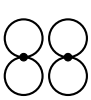
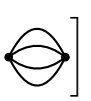
so we can recognize every line connecting a pair of vertices as a correlation function or "propagator". Now we can do perturbation theory graphically, for example for the partition function:

(161) $Z = Z_0 \left\langle 1 - u \int_1$  $+ \frac{u^2}{2} \int_{1,2}$  $+ \cdots \right\rangle$

Using Wick's theorem, this becomes

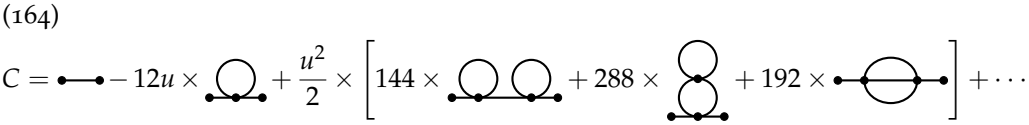
(162) $Z = Z_0 \left(1 - u \int_1 3 \times$  $+ \frac{u^2}{2} \int_{1,2} \left[9 \times$  $+ 4! \times$  $\right] + \cdots \right).$

Here I included factors which count how many different pairings lead to the same term. Often we will just skip writing the integrals and labels, and leave those implied, and so write just

(163) $Z = Z_0 \left(1 - 3u$  $+ \frac{u^2}{2} \left[9 \times$  $+ 4! \times$  $\right] + \cdots \right).$

Notice that the first term in the square brackets is "disconnected", and is just the square (up to the prefactor) of the term at previous order. This causes it to cancel if we look at a more physical object, like the "free energy", or the logarithm of Z. This is an example of a "linked cluster" theorem. The expansion of $\ln Z$ contains only connected diagrams. A similar result holds for the expansion of the correlation function itself, which as we saw, requires expanding both the numerator and denominator in Eq. (143). Upon combining those series, a cancellation of disconnected terms occurs. One finds

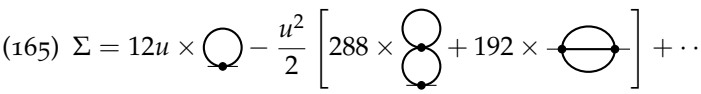
(164)



Note that I have drawn only the connected terms, which means that the "external" vertices (from the two φ fields defining C) and all the internal vertices are connected.

The self-energy has an additional simplification. Here we take all the terms in the correlation function, and first remove the two external legs (see how this worked in Eq. (154)), and *then* drop all those which are "one particle reducible", which means that they can be separated into two parts by removing a single line. This causes us to drop both the first term (the free correlation function) and the third term (first one inside the square brackets) in Eq. (164):

(165)



where I've drawn thin short lines to indicate where the legs were that were removed. The diagrammatic rules for the self-energy simplify things a lot. There would have been a lot more terms had we looked directly at the correlation function, for example. Anyway, we can see from Eq. (165) that there are two terms in the self-energy at $O(u^2)$. The first one is, like the $O(u)$ term, a constant, and so just another trivial shift of m^2. The second $O(u^2)$ term is not a constant, and so is more interesting.

Second order self-energy and three-particle contributions

Let's sketch what this thing actually looks like, to get a bit of physics out of it. In real space/time, it is simple because it is just the product of three correlation functions:

(166)

But the correlation function is much easier to understand in momentum and freqency space. So we need to Fourier tranform this:

(167) $\Sigma^{(2)}(\mathbf{k},\omega_n) =$  $= -96u^2 \int d^d\mathbf{x} d\tau\, C_0(\mathbf{x},\tau)^3 e^{i\mathbf{k}\cdot\mathbf{x}+i\omega_n\tau}$

Now we can write

(168) $C_0(\mathbf{x},\tau) = \int \frac{d^d\mathbf{k}_a}{(2\pi)^d}\frac{d\omega_a}{2\pi} C_0(\mathbf{k}_a,\omega_a) e^{-i\mathbf{k}_a\cdot\mathbf{x}-i\omega_a\tau},$

using subscripts $a = 1,2,3$ for each factor of C_0. Then doing the integral over space, we obtain

(169) $\Sigma^{(2)}(\mathbf{k},\omega_n) = -\frac{96u^2}{\mathcal{Z}^3}\int\limits_{1,2,3}\int d\tau \prod_{a=1}^{3}\frac{e^{-i\omega_a\tau}}{\omega_a^2+v^2k_a^2+m^2} \times e^{i\omega_n\tau}\delta(\mathbf{k}-\sum_a \mathbf{k}_a).$

Now we can do all the frequency integrals which gives

(170) $\Sigma^{(2)}(\mathbf{k},\omega_n) = -\frac{12u^2}{\mathcal{Z}^3}\int\limits_{\mathbf{k}_a}\int d\tau \frac{1}{\mathcal{Z}^3} e^{i\tau\omega_n}\frac{e^{-|\tau|\sum_a\epsilon_a}}{\epsilon_1\epsilon_2\epsilon_3}\delta(\mathbf{k}-\sum_a\mathbf{k}_a),$

with $\epsilon_a = \sqrt{m^2+v^2k_a^2}$ is the (bare) quasiparticle energy. Now one can do the τ integral to get

(171)
$$\Sigma^{(2)}(\mathbf{k},\omega_n) = -\frac{12u^2}{\mathcal{Z}^3}\int\limits_{\mathbf{k}_a}\frac{1}{\epsilon_1\epsilon_2\epsilon_3}\left[\frac{1}{i\omega_n+\sum_a\epsilon_a}+\frac{1}{-i\omega_n+\sum_a\epsilon_a}\right]\delta(\mathbf{k}-\sum_a\mathbf{k}_a).$$

This expression is nice because we can see that the integrand is very well behaved when $Re[\omega_n]$ is smaller than $3m$ in absolute value. When it is larger, the integrand sums over poles, leading to a branch cut structure on the real frequency axis. This is very physical. It is particularly nice to analytically continue, which is straightforward here because the frequency is explicit, and to look at the Imaginary part:

(172) $\frac{1}{\pi}\mathrm{Im}\Sigma^{(2)}(\mathbf{k},\omega>0) = -\frac{12u^2}{\mathcal{Z}^3}\int\limits_{\mathbf{k}_a}\frac{1}{\epsilon_1\epsilon_2\epsilon_3}\delta(\omega-\sum_a\epsilon_a)\delta(\mathbf{k}-\sum_a\mathbf{k}_a).$

This tells us that there are contributions arising from states of three quasiparticles, when ω and $\mathbf{k}$ are equal to the sum of the energy and momentum, respectively, of the three quasi-particles. The imaginary part in Σ translates to an imaginary part in C, through the relation between them. It indicates that interactions produce additional spectral weight in the correlation function corresponding to three particle states – see Fig. 3. The interpretation is that, in the interacting theory, the exact quasiparticle states are "dressed", so that the "bare" quasiparticle state created by the operator φ has overlap not only with

an exact one-quasiparticle state but also with exact 3-quasiparticle states.[4]

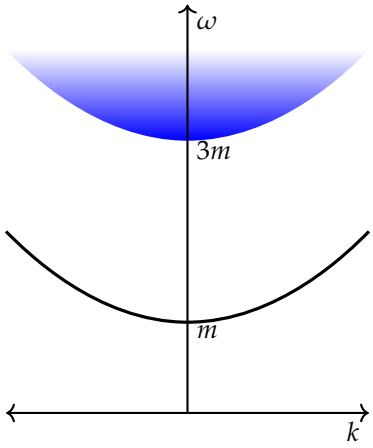


Figure 3: Schematic intensity plot of spectral function for the scalar field theory in the paramagnetic phase. At zero momentum, a delta-function quasiparticle peak appears at $\omega = m$, and the three-quasiparticle continuum appears for $\omega \geq 3m$.

6 Bosons, superfluidity, etc.

Lecture 10 (1.5h)
October 27th, 2015

Up to now we focused on the situation of a discrete (Ising) symmetry, in which, except if one tunes to a QCP, the system is gapped and the ground state is "strictly" product-like: we expect that the area law of entanglement entropy is obeyed with corrections that become exponentially small in the size of the subregion defining the partition, when this is taken large. A gap to all excitations is rather generic in the situation of discrete symmetry (though we will eventually encounter exceptions). However, in the presence of continuous symmetry, excitations with arbitrarily low energy can emerge, as we will see. This goes along with the phenomena of spontaneously broken continuous symmetry. This abstract concept underlies the physical phenomena of Bose-Einstein condensation and superfluidity.

6.1 *Second quantization, coherent states, and coherent state path integral*

Second quantization

We need to set up a bit of formalism. Hopefully you are already familiar with some of it. The most basic is second quantization. I quickly review it here, without proving the statements. You can find them all over. For a system of particles, we learn in elementary quantum mechanics how to write a many-electron Schrödinger equation. We introduce N coordinates $\mathbf{x}_i$ and momenta

[4] It has zero overlap with exact *two*-quasiparticle states because quasiparticles carry the Ising parity, and so only odd parity states mix with odd parity states.

$\mathbf{p}_i$, and write a general Hamiltonioan as

$$H_N = \sum_{i=1}^{N} h^{(1)}(\mathbf{x}_i, \mathbf{p}_i) + \frac{1}{2}\sum_{i\neq j} h^{(2)}(\mathbf{x}_i, \mathbf{p}_i; \mathbf{x}_j, \mathbf{p}_j), \tag{173}$$

where the superscripts indicate one-body and two-body terms, and as usual, in the coordinate representation, $\mathbf{p}_i = -i\boldsymbol{\nabla}_i$. The Schrödinger equation is

$$H_N \Psi(\mathbf{x}_1, \cdots, \mathbf{x}_N) = E\Psi(\mathbf{x}_i, \mathbf{p}_i). \tag{174}$$

For identical bosons, for which we consider only symmetric wavefunctions, we can employ second quantization (we can do it for Fermions too, but Andreas says not to tell you about that!). Now we introduce the many-particle Fock space, in which the number of particles is arbitrary ($\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 + \mathcal{H}_2 + \cdots$, where $\mathcal{H}_N$ is the Hilbert space of N particles, including the space spanned by the vacuum state, $\mathcal{H}_0 = \{c|0\rangle\}$). The second quantized Hamiltonian $H = H_0 \oplus H_1 \oplus H_2 \oplus \cdots$ is written

$$H = \sum_{ij}\langle j|h^{(1)}|i\rangle a_j^\dagger a_i + \frac{1}{2}\sum_{ijkl}\langle kl|h^{(2)}|ji\rangle a_k^\dagger a_l^\dagger a_i a_j, \tag{175}$$

where the states $|i\rangle$ form an arbitrary orthonormal basis of single particle states, and we introduced creation/annihilation operators satisfying

$$[a_j, a_k^\dagger] = \delta_{jk}. \tag{176}$$

The interpretation is that $a_j^\dagger$ creates a particle in state j, and a_j annihilates one. The number of particles in state j is $n_j = a_j^\dagger a_j$. The vacuum state $|0\rangle$ is annihilated by all a_j:

$$a_j|0\rangle = 0. \tag{177}$$

For ordinary bosons in free space, for example cold atoms in a trap, we could take

$$h^{(1)}(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2m} + U(\mathbf{x}), \qquad h^{(2)}(\mathbf{x}, \mathbf{x}') = V(\mathbf{x}, \mathbf{x}'), \tag{178}$$

where $U(\mathbf{x})$ is some one-particle (e.g. trapping) potential, and $V(\mathbf{x}, \mathbf{x}')$ is an interaction. In this case, it is convenient to use a real-space basis, i.e. $|i\rangle = |\mathbf{x}\rangle$ of states localized at a point in space (i.e. delta-function wavefunctions), and $a_i \to \psi(\mathbf{x})$ – the label i becomes the continuous coordinate $\mathbf{x}$, and we have

$$H = \int d^d\mathbf{x}\psi^\dagger(\mathbf{x})\left(-\frac{\nabla^2}{2m} + U(\mathbf{x})\right)\psi(\mathbf{x}) + \frac{1}{2}\int d^d\mathbf{x}d^d\mathbf{x}'\psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{x}')V(\mathbf{x}, \mathbf{x}')\psi(\mathbf{x}')\psi(\mathbf{x}), \tag{179}$$

where

(180) $[\psi(\mathbf{x}), \psi^\dagger(\mathbf{x}')] = \delta^{(d)}(\mathbf{x} - \mathbf{x}')$.

In general, one can go back from Eq. (179) to Eq. (175) by the change of basis formula,

(181) $\psi(\mathbf{x}) = \sum_i \phi_i(\mathbf{x}) a_i$,

where $\phi_i(\mathbf{x}) = \langle \mathbf{x}|i\rangle$ is the wavefunction of state i.

Coherent states

Coherent states are useful in various applications. They can serve as a way to understand the classical limit of harmonic oscillators. They are also useful to construct path integrals. We'll use this. A coherent state is an eigenstate of an annihilator operator:

(182) $a|\phi\rangle = \phi|\phi\rangle$,

where we consider just a single mode, and suppress the label (i etc.) – with many single particle states we just use a direct product of these things. The parameter ϕ can be complex, since a is not hermitian. One can explicitly find such a state as follows:

(183) $|\phi\rangle = e^{\phi a^\dagger}|0\rangle$,

where $|0\rangle$ is the vacuum state. If you expand the exponential you will see that the coherent state is a superposition of all particle numbers. Of course, by complex conjugation,

(184) $\langle\phi| = \langle 0|e^{\phi^* a}$.

And of course

(185) $\langle\phi|a^\dagger = \phi^*\langle\phi|$.

How does the raising operator act on a ket? We can differentiate Eq. (183) to obtain

(186) $a^\dagger|\phi\rangle = \dfrac{\partial}{\partial\phi}|\phi\rangle$.

Note that $a^\dagger$ behaves similarly to the usual momentum operator in this basis. It is important to note that since a is not hermitian, different coherent states are not orthogonal. In fact, one can show by Taylor expanding or other means that

(187) $\langle\phi|\phi'\rangle = e^{\phi^*\phi'}$.

We will need the resolution of the identity,

$$\mathbb{1} = \int \frac{d\phi^* d\phi}{2\pi i} e^{-\phi^*\phi} |\phi\rangle\langle\phi|. \tag{188}$$

Here we have written the measure in a standard way, which is slightly formal. It is defined as though ϕ^* and ϕ are independent variables. Equivalently, we can write $\phi = \mathrm{Re}\phi + i\mathrm{Im}\phi$, and $\phi^* = \mathrm{Re}\phi - i\mathrm{Im}\phi$, and then use the change of variables formula

$$d\phi^* d\phi = d\mathrm{Re}\phi d\mathrm{Im}\phi \left| \frac{\partial(\phi^*, \phi)}{\partial(\mathrm{Re}\phi, \mathrm{Im}\phi)} \right| = 2i\, d\mathrm{Re}\phi\, d\mathrm{Im}\phi. \tag{189}$$

The resolution of the identity differs from the usual one because the states are not orthonormal. You can prove it by using the definitions, Eq. (183) and Eq. (184), writing out the exponentials explicitly as sums over number eigenstates, and then carrying out the integral. Using Eq. (188), we can easily find an expression for the trace,

$$\mathrm{Tr}\mathcal{O} = \int \frac{d\phi^* d\phi}{2\pi i} e^{-\phi^*\phi} \langle\phi|\mathcal{O}|\phi\rangle. \tag{190}$$

All these formula are readily elevated to the full space of many single-particle states. We replace $|\phi\rangle \to |\phi_1, \phi_2, \cdots\rangle = |\{\phi\}\rangle$. Usually, though it may be confusing, we will just denote this by the same symbol $|\phi\rangle$. Then the key identities are

$$\langle\phi|\phi'\rangle = e^{\sum_i \phi_i^* \phi_i'}, \tag{191}$$

$$\mathbb{1} = \int \prod_i \frac{d\phi_i^* d\phi_i}{2\pi i} e^{-\sum_i \phi_i^*\phi_i} |\phi\rangle\langle\phi|, \tag{192}$$

and

$$\mathrm{Tr}\mathcal{O} = \int \prod_i \frac{d\phi_i^* d\phi_i}{2\pi i} e^{-\sum_i \phi_i^*\phi_i} \langle\phi|\mathcal{O}|\phi\rangle. \tag{193}$$

Coherent state path integral

Now we can use the coherent states to construct a path integral for the partition function. With bosons, it is usual to consider the grand canonical partition function $Z = \mathrm{Tr} e^{-\beta(H-\mu N)}$, where N is the number of boson and μ is the chemical potential. Since we can write $N = \sum_i n_i = \sum_i a_i^\dagger a_i$, we can easily absorb the chemical potential into $h^{(1)}$. So henceforth we will not write the $-\mu N$ explicitly: it is included in H which is then a "grand canonical Hamiltonian".

We follow the usual strategy to construct the path integral. First write the

trace in the desired basis, which here is the coherent states:

$$Z = \int \prod_i \frac{d\phi_i^* d\phi_i}{2\pi i} e^{-\sum_i \phi_i^* \phi_i} \langle \phi | e^{-\beta H} | \phi \rangle. \tag{194}$$

Now we split up the exponential into a Trotter product, and insert a resolution of the identity in between each term. We thereby get

$$Z = \int [d\phi^*][d\phi] \prod_{\tau=0}^{\beta-\Delta\tau} e^{-\sum_i \sum_{\tau=0}^{\beta-\Delta\tau} \phi_i^*(\tau)\phi_i(\tau)} \prod_{\tau=0}^{\beta-\Delta\tau} \langle \phi(\tau+\Delta\tau) | e^{-\Delta\tau H} | \phi(\tau) \rangle, \tag{195}$$

where we introduced the imaginary time label as we did earlier in Sec. 4, and we define $\phi(\beta) = \phi(\tau)$ since the outer bra and ket are the same in the trace. The measure is now being left implicit, but there is in fact an integral over complex ϕ for each i and each τ. Now for infinitesimal $\Delta\tau$, we can expand the exponential to linear order and it is simple to evaluate provided we have, as in Eq. (175), all creation operators on the left, and annihilation operators in the right, in each term of H. This is called *normal ordering*, and we will assume that H is normal ordered. If not, one can reorder the terms in H to normal order them. Normal ordering is good because the a_i operators on the right act on a ket, and each gives a factor of $\phi_i(\tau)$, while the $a_i^\dagger$ acts on the left to produce a factor of $\phi_i^*(\tau+\Delta\tau)$, hence

$$\langle \phi(\tau+\Delta\tau) | e^{-\Delta\tau H} | \phi(\tau) \rangle = e^{-\Delta\tau H(\phi^*(\tau+\Delta\tau),\phi(\tau))} e^{\sum_i \phi_i^*(\tau+\Delta\tau)\phi_i(\tau)}. \tag{196}$$

When Eq. (196) is inserted in Eq. (195), the second exponential in Eq. (196) will combine wiht the first exponential in Eq. (195), since neither involves H. One obtains

$$Z = \int [d\bar{\phi}][d\phi] e^{-\sum_{i\tau} (\bar{\phi}_i(\tau) - \bar{\phi}_i(\tau+\Delta\tau))\phi_i(\tau)} e^{-\Delta\tau \sum_{i\tau} H(\bar{\phi},\phi)}, \tag{197}$$

where we are suppressing the detailed labels of ϕ inside the Hamiltonian. We also defined $\bar{\phi} = \phi^*$ which looks nicer. At this point, we can take the $\Delta\tau \to 0$ limit to obtain

$$Z = \int [d\bar{\phi}][d\phi] e^{-S}, \tag{198}$$

where the boundary condition $\phi_i(\beta) = \phi_i(0)$ is assumed, and the action is

$$S = \int_0^\beta d\tau \left[\sum_i \bar{\phi}_i \partial_\tau \phi_i + H(\bar{\phi}, \phi) \right]. \tag{199}$$

In taking $\Delta\tau \to 0$, we can a priori neglect the difference between the imaginary time of the ϕ^* argument and the ϕ argument of H. This is sometimes important if some operator ordering issue arises, but usually it is ok.

6.2 *The superfluid state*

Consider a collection of bosons, for simplicity with a fixed density in the continuum. This is described by the Hamiltonian of Eq. (179). This time around we will discuss the physics in the opposite order that we did in the Ising model – starting with the path integral formulation. Later we will go back to think about the actual wavefunction.

Effective action

Applying the formalism of the previous subsection, we see that the partition function for the system of bosons can be written as a path integral,

$$Z = \int [d\overline{\psi}][d\psi] e^{-S[\overline{\psi},\psi]}, \tag{200}$$

with the action

$$S = \int_0^\beta d\tau \left\{ \int d^d\mathbf{x}\, \overline{\psi} \left(\partial_\tau - \frac{\nabla^2}{2m} - \mu \right) \psi + \frac{1}{2} \int d^d\mathbf{x} d^d\mathbf{x}'\, |\psi(x)|^2 V(\boldsymbol{x}-\boldsymbol{x}') |\psi(x')|^2 \right\}. \tag{201}$$

Here we consider a translationally-invariant system with no one-body potential, for simplicity. Now let us look for a saddle point. It is extremely natural to expect the dominant saddle point is constant in time and space, the former because we expect a stationary state, and the latter because we seek to describe a uniform gas or Bose fluid. For such a configuration the action simplies to

$$S = \beta V \left[-\mu|\psi|^2 + \frac{1}{2}\tilde{V}(\mathbf{0})|\psi|^4 \right], \tag{202}$$

where $\tilde{V}(\boldsymbol{q}) = \int d^d\mathbf{x}\, e^{i\boldsymbol{q}\cdot\boldsymbol{x}} V(\boldsymbol{x})$. It is obviously minimized by $|\psi_0|^2 = \mu/\tilde{V}(\mathbf{0})$ for $\mu > 0$ and $\psi = 0$ otherwise. In general, we should tune the chemical potential so that the expectation value of the density operator gives the physical boson density,

$$\langle \psi^\dagger(\boldsymbol{x})\psi(\boldsymbol{x})\rangle = N/V \equiv \overline{n}. \tag{203}$$

If we evaluate the left-hand side at the saddle point level, then obviously this requires $\mu > 0$, but in fact fluctuations around the saddle point give a non-zero left-hand side even for $\mu < 0$ at $T > 0$, which describes a free of quasi-free Bose gas above its Bose-Einstein condensation temperature. Here we will focus on $T = 0$, where the $\mu > 0$ saddle point is the correct one.

Lecture 11 (1.5h)
October 29th, 2015

An interesting feature of this case is the *degeneracy* of the saddle point. Minimizing the action fixes the magnitude of ψ but not its phase:

$$\psi_0 = \sqrt{\frac{\mu}{\tilde{V}(\mathbf{0})}} e^{i\theta}, \tag{204}$$

where θ is arbitrary. This reminds one of the two-fold degeneracy of the ground state of the Ising model in the ferromagnetic phase, and indeed a corresponding two-fold degeneracy of saddle points in the path integral formulation of that problem. Now we are taught that the overall phase of the wavefunction is not physical, so the meaning of this degeneracy is not so clear. This is not exactly the phase of the many-body wavefunction, but it seems similar. Formally, the boson Hamiltonian has a conserved total number, $N = \int d^d\mathbf{x}\, \psi^\dagger(\boldsymbol{x})\psi(\boldsymbol{x})$, which generates the $U(1)$ phase-rotation symmetry via $U(\chi) = e^{i\chi N}$:

$$U(1): \qquad \psi(\boldsymbol{x}) \to e^{i\chi}\psi(\boldsymbol{x}). \tag{205}$$

This is why the value of the action is independent of θ. It is pretty obvious too that $\langle\psi\rangle = 0$ in any state with fixed particle number. So the meaning of the phase in Eq. (204) is certainly tricky.

Number-phase representation and sound mode

One can recognize that, although the overall phase is not physical, a phase difference is. This means that a phase *gradient* has physical meaning. We can expect that if a configuration with proper magnitude and constant phase minimizes the action, then a configuration with a small phase gradient will have a small action, and hence be important in the path integral. Motivated by this, we can try the change of variables

$$\psi(\boldsymbol{x}) = \sqrt{n(\boldsymbol{x})}e^{i\theta(\boldsymbol{x})}. \tag{206}$$

This transformation has a unit Jacobean in the path integral. So

$$Z = \int [dn][d\theta] e^{-S[n,\theta]}, \tag{207}$$

and by some simple algebra, one obtains

$$\begin{aligned} S[n,\theta] \;=\; & \int_0^\beta d\tau \Big\{ \int d^d\mathbf{x} \left[in\partial_\tau\theta + \frac{n}{2m}|\boldsymbol{\nabla}\theta|^2 + \frac{|\boldsymbol{\nabla}n|^2}{8mn} - \mu n \right] \\ & + \frac{1}{2}\int d^d\mathbf{x} d^d\mathbf{x}'\, n(\boldsymbol{x})V(\boldsymbol{x}-\boldsymbol{x}')n(\boldsymbol{x}') \Big\}. \end{aligned} \tag{208}$$

Here we used spatial integration by parts, neglecting boundary terms, to simplify the parts arising from the boson's kinetic energy. The saddle point in these variables corresponds to $n = n_0 = \mu/\tilde{V}(\mathbf{0})$ and θ constant. Let's expand around the saddle point, by writing $n = n_0 + \delta n$, and keeping terms of

$O(\delta n^2, \theta^2, \delta n\theta)$. We have $S = S_0 + \delta S$, with

$$\begin{aligned} \delta S &= \int_0^\beta d\tau \Big\{ \int d^d\mathbf{x} \left[i n_0 \partial_\tau \theta + i\delta n\, \partial_\tau \theta + \frac{n_0}{2m} |\boldsymbol{\nabla}\theta|^2 + \frac{|\boldsymbol{\nabla}\delta n|^2}{8mn_0} \right] \\ &\quad + \frac{1}{2} \int d^d\mathbf{x} d^d\mathbf{x}'\, \delta n(\boldsymbol{x}) V(\boldsymbol{x}-\boldsymbol{x}') \delta n(\boldsymbol{x}') \Big\}. \end{aligned} \tag{209}$$

The $n_0\partial_\tau\theta$ term is a total derivative and can be neglected here. The remainder is a quadratic form and is simplified by Fourier tranformation

$$\begin{aligned} \delta S &= \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{d\omega_n}{2\pi} \Big\{ \frac{1}{2} (\tilde{V}(\boldsymbol{k}) + \frac{k^2}{4mn_0}) \delta n_{\mathbf{k},\omega_n} \delta n_{-k,-\omega_n} \\ &\quad + \frac{n_0 k^2}{2m} \theta_{\mathbf{k},\omega_n} \theta_{-k,-\omega_n} + \omega_n \delta n_{\mathbf{k},\omega_n} \theta_{-k,-\omega_n} \Big\}. \end{aligned} \tag{210}$$

We can write this in a matrix form

$$\delta S = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{d\omega_n}{2\pi} \frac{1}{2} \begin{pmatrix} \delta n & \theta \end{pmatrix}_{\mathbf{k},\omega_n} \begin{pmatrix} \tilde{V}(\boldsymbol{k}) + \frac{k^2}{4mn_0} & \omega_n \\ -\omega_n & \frac{n_0 k^2}{m} \end{pmatrix} \begin{pmatrix} \delta n \\ \theta \end{pmatrix}_{-k,-\omega_n}. \tag{211}$$

Now it is clear that the correlation functions of δn and θ are just given by the inverse of this matrix. Since the excitations are related to poles of the correlation functions, these occur when the determinant of this matrix vanishes, i.e.

$$\left(\tilde{V}(\boldsymbol{k}) + \frac{k^2}{4mn_0} \right) \frac{n_0 k^2}{m} + \omega_n^2 = 0, \tag{212}$$

and using $i\omega_n \to \omega$ we see that excitations occur at

$$\omega = \sqrt{\left(\tilde{V}(\boldsymbol{k}) + \frac{k^2}{4mn_0} \right) \frac{n_0 k^2}{m}}. \tag{213}$$

This is kind of a nice formula. If we consider large momentum $k \gg 4mn_0\tilde{V}$, we obtain simply $\omega \sim k^2/(2m)$, the dispersion of a free boson. At small k, however, $\tilde{V}(\boldsymbol{k}) \approx \tilde{V}(\mathbf{0}) \gg k^2/(4mn_0)$ and we have

$$\omega \sim vk, \tag{214}$$

with $v = \sqrt{n_0\tilde{V}(\mathbf{0})/m}$. Eq. (214) describes an *acoustic*, or sound-like mode, with linear dispersion and obviously no excitation gap. The absence of an excitation gap is a direct consequence of the degeneracy of the saddle point. Because the action of a uniform phase is independent of this phase, the action of a slowly varying gradient is very small, since locally it is approximately the same as a constant phase, and the interactions are local (we assumed some degree of locality when we declared at $\tilde{V}(\mathbf{0})$ is finite – for example the Fourier

transform of a Coulomb interaction diverges at small momentum like $1/k^2$ in three dimensions). The sound mode corresponds to an excitation with such a small gradient.

We can see this more explicitly by going back to Eq. (209), and making approximations equivalent to the small k limit. At long wavelengths, we can approximate $V(\boldsymbol{x}-\boldsymbol{x}') \approx \tilde{V}(\mathbf{0})\delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}')$, and neglect the $|\boldsymbol{\nabla}\delta n|^2$ term, as well as the total derivative term we neglected earlier. Then we have

$$\delta S \approx \int d^d\mathbf{x}\, d\tau \, \left\{ \frac{\tilde{V}(\mathbf{0})}{2}(\delta n)^2 + i\delta n\, \partial_\tau\theta + \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2 \right\}. \tag{215}$$

Here we can see that a density fluctuation is penalized by a constant "mass", i.e. a quadratic term which does not vanish at small momenta and frequencies, while the action for small gradients of θ is small. To this end, we can integrate out δn:

$$Z = Z_0 \int [d\delta n][d\theta] e^{-\delta S[\delta n,\theta]} = \tilde{Z}_0 \int [d\theta] e^{-S_\theta[\theta]}, \tag{216}$$

which is easy to do because it is just a Gaussian integral. We obtain then an action for θ alone:

$$S_\theta = \int d^d\mathbf{x}\, d\tau \, \left[\frac{1}{2\tilde{V}(\mathbf{0})}(\partial_\tau\theta)^2 + \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2 \right]. \tag{217}$$

This is just the action of a free massless scalar boson. We obtained something similar for the Ising model at its quantum critical point (with φ instead of θ), but here we found this for an arbitrary parameter in the ordered phase. Moreover, in the Ising model, even at the QCP, there was a correction in the form of a φ^4 term, which is important because it lacks any gradient. No such term is possible here, because of the $U(1)$ symmetry in Eq. (205): the action must be invariant under $\theta \to \theta + \chi$. Thus correction terms must involve powers of space-time gradients of θ, rather than θ itself. This suppresses their effects at low energy and in fact makes the predictions of the simple free action robust.

This means that the linear acoustic mode is a generic and universal feature of the superfluid. It is an example of a *Goldstone mode*. Goldstone modes occur whenever continuous symmetries are spontaneously broken. In the boson context, even though the global phase is not meaningful, the presence of the Goldstone "sound" mode gives a physical significance to the $U(1)$ symmetry breaking.

We can also think about the sound mode as a quasiparticle of the superfluid. To relate to the spectrum, we can consider the Green's function of θ. This is mathematically identical to the Green's function C_0 of φ we obtained in Sec. 5.5 for the Ising scalar field, except with $m = 0$. This indeed shows that there is a mode with $\omega = v|k|$, and moreover that this mode contributes a delta-function to the spectral weight of the $\theta - \theta$ correlator. Thus we can indeed regard the sound mode as a branch of quasiparticles, and, in the current approximation, these quasiparticles are infinitely sharp, i.e. they have a delta-function spectral weight, just as the quasiparticles did in the Ising case.

In general, the width of the peak representing a quasiparticle excitation in a spectral function is related to the *lifetime* of the excitation. If it can decay, then the peak will not be infinitely sharp, but have a width which is given by the decay rate, i.e. the inverse lifetime. For the Ising quasiparticle, the lifetime is truly infinite, because there are no lower-energy excitaitons into which to decay. However, here, because the quasiparticle states extend down to zero energy, decay is *a priori* possible. So we should not really expect an infinite lifetime. How large a lifetime is long enough? Typically, we regard a quasiparticle as well-defined if its width in energy is narrower than its mean energy. This means that it survives long enough before decaying for its quantum wavefunction to undergo more than one phase oscillation.

Lecture 12 (1.5h)
November 3rd, 2015

Lifetime of phonon

Let's see if we can actually get the lifetime of the sound mode or phonon. To do so, we will need to add a few terms to the minimal action, in order to allow the mode to decay. Symmetry allows the following modifications, which have the minimal number of powers of θ and derivatives:

$$S_\theta = \int d^d\mathbf{x}\, d\tau \ \left[\frac{1}{2}(\partial_\tau\theta)^2 + \frac{1}{2}|\boldsymbol{\nabla}\theta|^2 + \frac{\alpha}{2}|\nabla^2\theta|^2 + i\lambda\partial_\tau\theta|\boldsymbol{\nabla}\theta|^2\right]. \tag{218}$$

Here I rescaled coordinates so that the coefficients of the two leading terms are just $1/2$, which simplifies later calculations. The α term is still quadratic in θ, and so can be included exactly. It modifies the zeroth order correlation function to

$$C_0(\boldsymbol{k},\omega_n) = \frac{1}{\omega_n^2 + \epsilon_{\boldsymbol{k}}^2}, \tag{219}$$

with

$$\epsilon_{\boldsymbol{k}} = \sqrt{k^2 + \alpha k^4} \approx |k| + \frac{\alpha}{2}|k|^3. \tag{220}$$

The λ term is cubic in θ, and so can only be handled perturbatively. To get a decay, we need to consider the self-energy. The leading diagram is $O(\lambda^2)$,

$\Sigma =$ (221)

I guess I need to explain in more detail how to calculate these things. There are certain "rules" for translating such diagrams into integrals directly in momentum space. I will not go into it for the moment, and just write down a result

$$\Sigma(k,\omega_n) \propto \lambda^2 \int\limits_{\boldsymbol{q},\Omega_n} \frac{\left[\omega_n(q^2 + 2\boldsymbol{k}\cdot\boldsymbol{q}) + \Omega_n(k^2 + 2\boldsymbol{k}\cdot\boldsymbol{q})\right]^2}{(\Omega_n^2 + \epsilon_{\boldsymbol{q}}^2)((\omega_n + \Omega_n)^2 + \epsilon_{\boldsymbol{k}+\boldsymbol{q}}^2)}. \tag{222}$$

The two denominators are simply the Fourier transform of the two Green's functions represented by the lines. The numerator is a result of all the derivatives in the λ term – you can at least count them and see that there are six powers total of frequency and momentum, which corresponds to the two λ terms that went into this diagram. We can do the Ω_n integral by completing the contour in the upper half plane, and evaluating the residue due to two poles. This gives

(223)
$$\Sigma(\boldsymbol{k},\omega_n) \propto \lambda^2 \int_{\boldsymbol{q}} \frac{(k^2+2\boldsymbol{k}\cdot\boldsymbol{q})^2\epsilon_q^2\epsilon_{k+q} + (q^2+2\boldsymbol{k}\cdot\boldsymbol{q})^2\epsilon_{k+q}\omega_n^2 + \epsilon_q[(k^2+2\boldsymbol{k}\cdot\boldsymbol{q})^2\epsilon_{k+q}^2 + (k^2-q^2)^2\omega_n^2]}{2\epsilon_q\epsilon_{k+q}(\epsilon_q+\epsilon_{k+q}-i\omega_n)(\epsilon_q+\epsilon_{k+q}+i\omega_n)}$$

Now we can analytically continue and extract the imaginary part. If we consider $i\omega_n \to \omega + i\delta$ and $\omega > 0$, we obtain a delta-function constraining $\omega = \epsilon_q + \epsilon_{k+q}$ (another delta-function contributes only for negative frequency):

(224)
$$\mathrm{Im}\Sigma(\boldsymbol{k},\omega>0) \propto \lambda^2 \int_{\boldsymbol{q}} \frac{[(q^2-k^2)\epsilon_q + (q^2+2\boldsymbol{k}\cdot\boldsymbol{q})\epsilon_{k+q}]^2}{\epsilon_q\epsilon_{k+q}}\delta(\omega-\epsilon_q-\epsilon_{k+q}).$$

We can understand this physically as corresponding to the creation of two quasiparticles which split the total momentum $\boldsymbol{k}$ into parts $\boldsymbol{q}$ and $\boldsymbol{k}+\boldsymbol{q}$. To interpret the imaginary part, let us write the correlation function including it:

(225)
$$\begin{aligned} -\mathrm{Im}D(\boldsymbol{k},\omega) &= \mathrm{Im}\left[\frac{1}{-\omega^2+\epsilon_k^2+i\Sigma'}\right] \\ &= \frac{-\Sigma'}{(\omega^2-\epsilon_k^2)^2+(\Sigma')^2}. \end{aligned}$$

To examine the peak, let $\omega = \epsilon_k + \delta\omega$, then for $\delta\omega \ll \epsilon_k$, we obtain

(226)
$$-\mathrm{Im}D(\boldsymbol{k},\omega) \approx \frac{-\Sigma'}{4\epsilon_k^2\,[(\delta\omega)^2+(\Sigma'/\epsilon_k)^2]},$$

so we see that the inverse lifetime, or decay rate is

(227)
$$\frac{1}{\tau} = \frac{|\Sigma'(\mathbf{k},\omega=\epsilon_k)|}{\epsilon_k}.$$

Therefore, evaluating Eq. (224) and dividing by the energy,

(228)
$$\frac{1}{\tau} \propto \frac{\lambda^2}{\epsilon_k}\int_{\boldsymbol{q}} \frac{[(q^2-k^2)\epsilon_q + (q^2+2\boldsymbol{k}\cdot\boldsymbol{q})\epsilon_{k+q}]^2}{\epsilon_q\epsilon_{k+q}}\delta(\epsilon_k-\epsilon_q-\epsilon_{k+q}).$$

Now the delta-function constraint is going to pick out some subspace of the full $\boldsymbol{q}$ integration. We are interested in the small k limit, for which the αk^3 term

in Eq. (220) is a small perturbation. So let's first consider what would happen if we set $\alpha = 0$ entirely. Then the constraint becomes $|k| = |q| + |\boldsymbol{k} + \boldsymbol{q}|$. Squaring this out, we find that this is satisfied only when $\boldsymbol{k} \cdot \boldsymbol{q} = -kq$, which requires that $\boldsymbol{q}$ is antiparallel to $\boldsymbol{k}$. This is kind of a singular limit: we would expect generally that the single delta function would reduce the d-dimensional q integration to a $d-1$-dimensional one. But it seems that for $\alpha = 0$ in three dimensions it reduces the 3-dimensional integral to a one-dimensional one (over the modulus q). In 1d it seems to reduce the integral from the full space to a half-space. So in general we need to include α to get sensible results.

We will ignore $d = 1$ for now. In $d > 1$, for small k we expect that the solution of the constraint will be given by wavevectors $\boldsymbol{q}$ which are close to but not exactly antiparallel with $\boldsymbol{k}$. Without loss of generality, take $\boldsymbol{k} = k\hat{\boldsymbol{z}}$ and $\boldsymbol{q} = q_z\hat{\boldsymbol{z}} + \boldsymbol{q}_\perp$, with $\boldsymbol{q}_\perp \cdot \hat{\boldsymbol{z}} = 0$. If we assume $k \ll 1$ and $q_z = O(k)$ and $|\boldsymbol{q}_\perp| \ll k, q_z$, we obtain, for $q_z < 0$,

$$\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{q}} - \epsilon_{\boldsymbol{k}+\boldsymbol{q}} \approx -3\alpha k q_z(k + q_z) + \frac{k q_\perp^2}{2q_z(k+q_z)}. \tag{229}$$

For $q_z > 0$ the quantity is always large and position, so the delta-function cannot be satisfied. Define a dimensionless variable by letting $q_z = -kr$. This gives

$$\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{q}} - \epsilon_{\boldsymbol{k}+\boldsymbol{q}} \approx k^3 \left[3\alpha r(1-r) - \frac{q_\perp^2}{2k^4 r(1-r)} \right] \tag{230}$$

We see that the delta-function is satisfied when

$$r(1-r) = \sqrt{\frac{1}{6\alpha}} \frac{q_\perp}{k^2}. \tag{231}$$

Since the left-hand side is bounded above by $1/4$ (when $r = 1/2$), we require

$$q_\perp < \frac{\sqrt{6\alpha} k^2}{4}, \tag{232}$$

which shows that we were indeed consistent assuming $q_\perp \ll k, |q_z|$. Note that we get a non-zero result only when $\alpha > 0$, which is a recognized feature of such problems (see *Chernyshev+Zhitomirsky* for much more than you really want to know). To proceed, let us change variables to

$$x = -\frac{q_z}{k}\left(1 + \frac{q_z}{k}\right) = r(1-r), \qquad \boldsymbol{w} = \frac{\boldsymbol{q}_\perp}{\sqrt{6\alpha} k^2}. \tag{233}$$

Making these substitutions in Eq. (228), and replacing $\epsilon_{\boldsymbol{k}} \to k$, we obtain, up

to a proportionality constant

$$\begin{aligned}\frac{1}{\tau} &\sim \lambda^2 k^{2d-2} \alpha^{(d-1)/2} \int_0^{1/4} \frac{dx}{\sqrt{1-4x}} \int_0^{\infty} d^{d-1}\boldsymbol{w}\, k^4 x\, \delta[3\alpha k^3 (x - \boldsymbol{w}^2/x)] \\ &\sim \lambda^2 k^{2d-1} \alpha^{(d-3)/2}. \qquad (234)\end{aligned}$$

This means that the quasiparticle peak has a width in energy of order k^3 or k^5 in two or three dimensions, respectively, which is much smaller than the mean energy k. This is a typical behavior of a quasiparticle excitation in a gapless system: the peak has an intrinsic width, but it narrows as the energy is lowered toward zero, so that in this limit it becomes sharp compared to its energy. At order one energy, there is *a priori* no reason for the quasiparticle to be sharp, and it will not be, unless interactions are for some reason weak on microscopic scales.

The one dimensional case is much more subtle and has actually been a subject of recent research.

Persistent flow

The sound mode is an elementary *excitation* of the uniform state. But phase gradients play another role: they can comprise non-trivial *topological* textures in the superfluid. Indeed these textures are why the superfluid is "super". Now we consider time-independent configurations of θ, which describe distinct states of the system. One can understand this by considering the partition function at "large" temperature, i.e. small β. When β is sufficiently small, then gradients of θ as a function of imaginary time become too costly in action, and the τ-independent configurations dominate. Hence

$$\begin{aligned} Z &= \int [d\theta(\boldsymbol{x},\tau)] e^{-\int d^d\mathbf{x}\, d\tau \left[\frac{1}{2\tilde{V}(\mathbf{0})}(\partial_\tau\theta)^2 + \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2\right]} \\ &\approx \int [d\theta(\boldsymbol{x})] e^{-\beta\int d^d\mathbf{x}\, \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2}. \qquad (235)\end{aligned}$$

This has the form of a classical partition function. We should expect that at non-zero temperature, there is some "decoherence", which reduces the system to an effective classical one on long scales. We see that in this classical limit, however, the phase itself survives as a new variable describing the superfluid. We recognize from Eq. (235) the "free energy"

$$F = \int d^d\mathbf{x}\, \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2. \qquad (236)$$

So we can think of equilibrium configurations of the superfluid as described by different functions $\theta(\boldsymbol{x})$. The coefficient determining the energy of a phase gradient, n_0, is known as the *superfluid density*.

Now consider an annular container, i.e. a superfluid confined within two concentric cylinders. The phase θ is defined within the superfluid. Now consider a loop $\mathcal{C}$ that lies within the superfluid and encircles the inner cylinder

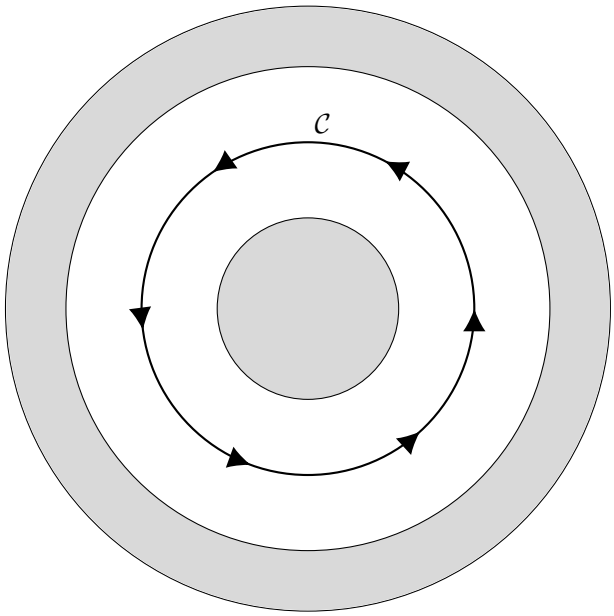


Figure 4: Superfluid confined to an annulus. The circulation around the countour $\mathcal{C}$ is quantized.

clockwise, returning to its original point. This is a non-trivial loop insofar as it cannot be smoothly deformed to a small loop or point without leaving the space. The line integral of the phase gradient must be an integer multiple of 2π, since the phase itself is defined only modulo 2π:

$$\frac{1}{2\pi} \oint_{\mathcal{C}} \boldsymbol{\nabla}\theta \cdot d\boldsymbol{x} \in \mathbb{Z}. \tag{237}$$

The integer defined by this integral is called the *winding number*. The different winding number states are really distinct physical configurations of the system. Due to the quantization of this line integral, a "trivial" configuration in which the integral vanishes is disconnected from a non-trivial one it which it is non-zero. This affords "topological" stability to the latter configurations. Note that constant θ, i.e. zero phase gradient, minimizes the free energy in Eq. (236), so configurations with non-zero winding number have higher energy than the ground state, and hence are only metastable. But they have extremely long lifetime. This is because any smooth time evolution of the system cannot change this winding number, since it is a topological invariant. Instead, some non-smooth evolution is required – we will see how this works later.

Now this is a bit abstract, but it is time to associate some physics to the phase gradient. In fact, the phase gradient has a natural interpretation in terms of the local *superfluid velocity*. To see this, consider a uniform phase gradient, $\theta = \boldsymbol{k} \cdot \boldsymbol{x}$. This corresponds to a plane wave form of ψ, suggesting non-zero momentum. More formally, we can write the boson number current operator,

$$\boldsymbol{J} = \frac{-i}{2m}\left(\psi^\dagger \boldsymbol{\nabla}\psi - \boldsymbol{\nabla}\psi^\dagger\,\psi\right). \tag{238}$$

This satisfies the continuity equation for particle conservation

$$\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{J} = 0. \tag{239}$$

Now inserting Eq. (206) in Eq. (238), we obtain

$$\boldsymbol{J} = \frac{n}{m}\boldsymbol{\nabla}\theta. \tag{240}$$

It is conventional to define $\boldsymbol{J} = n\boldsymbol{v}_s$, with the superfluid velocity

$$\boldsymbol{v}_s = \frac{\hbar}{m}\boldsymbol{\nabla}\theta. \tag{241}$$

Here for fun I restored the $\hbar$ (= 1) on the right hand side. Using Eq. (241), the integer winding number in Eq. (237) can be rewritten

$$\frac{\oint \boldsymbol{v}_s\cdot d\boldsymbol{x}}{h/m} \in \mathbb{Z}, \tag{242}$$

which expresses the *quantization of circulation* in units of h/m.

Vortices

Now we can understand that for an annular container, a persistent flow can exist, and it is very difficult for it to decay because this decay must happen somehow through abrupt steps, due to the quantization of circulation. But what about an ordinary cylindrical container, like a bucket? There is a nice experiment you can think of. Take the bucket in the normal state, i.e. at higher temperature, and spin up the fluid so that it is flowing around the center. If you wait it will gradually slow and stop. But suppose before that happens you cool it quickly to low temperature, within the superfluid phase. Now the circulation around any loop is supposed to be quantized. And if the fluid is still flowing it obviously it is non-zero if you take a circle around the center of the bucket. The thing is, that it also should be independent of the exact shape of the loop, so you can smoothly shrink the loop until it is almost a point. For such a small loop, the integral defining the circulation must vanish. Which says that the fluid cannot be rotating!

Something is fishy. If we carry out the cooling carefully the total angular momentum of the system remains conserved, so it *must* remain rotating. So what's the problem? Well, we have assumed that the superfluid phase θ is defined everywhere, which is the case if $|\psi|$ is non-zero everywhere. What actually happens is that in the rotating fluid the superfluidity is destroyed in some locations, i.e. $|\psi| = 0$ at some locations. These locations define *defects* in the superfluid, and around these defects there can be circulation. In two dimensions, such a defect is a point *vortex*, around which the phase θ winds

by a multiple of 2π. In three dimensions, the defect is a curved *vortex line*, such that the phase winds by a multiple of 2π on any curve encircling the defect. A rotating superfluid in a bucket is supported by an array of vertical vortex lines, whose net effect mimics rigid-body-like rotation of the superfluid.

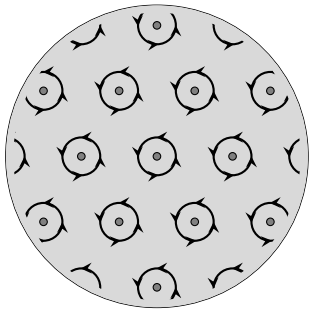


Figure 5: Vortices form a lattice in a rotating superfluid.

Lecture 13 (1.5h)
November 5th, 2015

Lifetime of persistent flow

Once we understand vortices, we can appreciate the stability of persistent flow in the annulus. Basically if a flow is set up with N quanta of circulation in the annulus, it is as though we have a large "supervortex" of strength N, or equivalently, N elementary vortices, inside the hole (central cylinder) of the annulus. The circulation decays by these vortices escaping from the hole. That is, a vortex line must be slowly exit the central cylinder, move through the bulk of the superfluid, and then excape out the outer wall. After this process, the circulation will be reduced by one quantum. The process is energetically favorable: the reduction of circulation reduces the angular momentum of the fluid, and hence its kinetic energy. However, there is an energy barrier which must be overcome. This is basically because the core of a vortex costs energy: the superfluid at the center of the vortex must be converted to normal fluid, even though the equilibrium state is superfluid, which means that the normal fluid is higher in energy. There is also a hydrodynamic contribution due to the profile of velocity around the vortex.

So what is the energy of a vortex? Well we can use the free energy of Eq. (236). Consider an ideal straight vortex line along the z axis. The phase θ is constant as a function of z and winds proportionally to the angle ϕ in the plane normal to z. The magnitude of the phase gradient is constant around a circle of constant radius r (distance from the vortex core), and since it winds by 2π around the circle, we have $|\boldsymbol{\nabla}\theta| = 1/r$. Then the free energy is

$$F_{\text{vortex}} = \int_0^L dz \int_a^R 2\pi r dr \frac{\hbar^2 n_0}{2m} \frac{1}{r^2} = \frac{\pi\hbar^2 n_0 L}{m} \ln(R/a). \tag{243}$$

Here we restored the needed factor $\hbar^2$, included a short-distance cutoff a representing the radius of the core, and approximated the geometry of the system

as a cylinder of height L and radius R around the vortex. Roughly we should take L the depth of the annular container and R of order the width of the channel. Putting in numbers for He, and approximating the logarithm by one, we obtain

$$F^{\rm He}_{\rm vortex} \gtrsim 0.1K \times \left(\frac{L}{1\mathring{A}}\right). \tag{244}$$

For even a not very deep container where $L \sim 1\mu m$, this is already prohibitively large compared to temperature, which must be below the superfluid transition temperature of helium is 2.2K. This means that thermal activated escape of vortices, which is proportional to the Arrhenius factor e^{-F/k_BT}, will be extremely slow.

6.3 *Superfluid wavefunctions*

We've used the path integral formulation to discuss the phenomena of superfluidity, the gapless Goldstone phonon mode, and other consequences of the emergent macroscopic phase. Let's now go back and understand some of these things in the quantum formulation in terms of wavefunctions. We understood the superfluid phase in the path integral by starting with a very simple saddle point $\psi(\boldsymbol{x},\tau) \to \psi_0 = \sqrt{n_0}$. What is the analog wavefunction? Remembering that the path integral was constructed in the coherent state basis, this corresponds to a product of coherent states at each site,

$$\begin{aligned} |\psi_0\rangle &= \prod_{\boldsymbol{x}} |\phi(\boldsymbol{x}) = \psi_0\rangle = \prod_{\boldsymbol{x}} e^{\psi_0 \psi^\dagger(\boldsymbol{x})}|0\rangle \\ &= e^{\int d^d\mathbf{x}\, \psi_0 \psi^\dagger(\boldsymbol{x})}|0\rangle, \end{aligned} \tag{245}$$

where we have not normalized the state. It is instructive to carry out a change of basis using Eq. (181) to momentum eigenstates. This gives

$$\int d^d\mathbf{x}\, \psi^\dagger(\boldsymbol{x}) = \sqrt{V} a^\dagger_{k=0}, \tag{246}$$

where V is the volume of the system. Then

$$|\psi_0\rangle = e^{\sqrt{N} a_{\mathbf{0}}^\dagger}|0\rangle, \tag{247}$$

where $N = n_0 V$ is the number of bosons. We can also write this out in Fock space

$$|\psi_0\rangle = \sum_{n=0}^{\infty} \frac{N^{n/2}}{n!} \left(a_{\mathbf{0}}^\dagger\right)^n |0\rangle = \sum_{n=0}^{\infty} \sqrt{\frac{N^n}{n!}} |n\rangle_{\boldsymbol{k}=0}, \tag{248}$$

where the kets are states with n bosons in the $\boldsymbol{k} = 0$ state. Obviously this is a state of uncertain boson number. For a closed system obviously we can choose exact eigenstates to have definite boson number. We can ask how bad the problem is – i.e. how peaks is the distribution of boson numbers? To do

so, let us examine the weight using Stirling's formula:

$$\sqrt{\frac{N^n}{n!}} = e^{1/2[n\ln N - \ln(n!)]} \sim e^{1/2[n\ln N - n\ln n + n]}, \tag{249}$$

which decays at large enough n since the $n\ln n$ term dominates, but grows at small n. So it is peaked at some n, which by differentiation we find is just $n = N$. Expanding around this we obtain, near the peak,

$$\sqrt{\frac{N^n}{n!}} \sim e^{\frac{N}{2} - \frac{(n-N)^2}{4N}}, \tag{250}$$

which indicates that the number fluctuations are of order $\sqrt{N}$, and so large, but still small compared to N itself. If we take just the largest component, or equivalently project onto the component with exactly N particles, we have simply

$$P_N|\psi_0\rangle \propto \left(a_{\mathbf{0}}^\dagger\right)^N |0\rangle. \tag{251}$$

This is just the ground state of non-interacting bosons, i.e. a Bose-Einstein Condensate (BEC). It is about as featureless as you could imagine. Indeed, one can ask about the first-quantized wavefunction of the BEC. It is

$$\begin{aligned} \Psi(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) &= \langle \boldsymbol{x}_1 \cdots \boldsymbol{x}_N | \psi_0 \rangle \\ &= \langle 0 | \psi(\boldsymbol{x}_1) \cdots \psi(\boldsymbol{x}_N) | \psi_0 \rangle = \text{Const.} \end{aligned} \tag{252}$$

Can't be more simple than that! It means there are absolutely no correlations between the positions of particles. This is an expression of the uncertainty principle: the momentum of the bosons is completely certain, so their conjugate positions must be completely uncertain.

Actually the non-interacting limit is so simple that it is a little pathological. It does not have a linearly dispersing Goldstone mode, and the chemical potential exactly vanishes. Basically there is an infinite compressibility, i.e. there is no energy cost to add particles. If you look at Eq. (213) where we described the sounds mode, you see that it becomes quadratic in the limit $V(\boldsymbol{k}) \to 0$, which is the non-interacting one. So we can expect that the BEC wavefunction is a little bit off from a generic superfluid one.

Obviously we need to build in *some* correlations between the boson positions to correct Eq. (252). To get an idea of how to do this, we can go back to our path integral theory and take a look at the fluctuations of the boson density. We can from Eq. (215), and since we are interested in the density, insteading of proceeding to Eq. (217) by integrating out δn, we do the opposite and integrate out θ. We obtain

$$Z \propto \int [d\delta n] e^{-S_n}, \tag{253}$$

with

$$S_n = \int \frac{d^d\mathbf{k}}{(2\pi)^d}\frac{d\omega_n}{2\pi}\frac{1}{2}\left[\tilde{V}(\mathbf{0}) + \frac{m}{n_0}\frac{\omega_n^2}{k^2}\right]\delta n_{\boldsymbol{k},\omega_n}\delta n_{-\boldsymbol{k},-\omega_n}. \tag{254}$$

This means that the density-density correlations have the form

$$\langle \delta n_{\boldsymbol{k},\omega_n}\delta n_{\boldsymbol{k}',\omega_n'}\rangle = \frac{(2\pi)^{d+1}\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}')\delta(\omega_n+\omega_n')}{\tilde{V}(\mathbf{0}) + \frac{m}{n_0}\frac{\omega_n^2}{k^2}}. \tag{255}$$

Now suppose we measure the correlations of the density in the ground state $|\psi\rangle$. This gives the equal time correlations:

$$\langle\psi|\delta n_{\boldsymbol{k}}\delta n_{\boldsymbol{k}'}|\psi\rangle = \langle \delta n_{\boldsymbol{k}}(\tau)\delta n_{\boldsymbol{k}'}(\tau)\rangle, \tag{256}$$

where $\delta n(\boldsymbol{x}) = \psi^\dagger(\boldsymbol{x})\psi(\boldsymbol{x}) - n_0$ and $\delta n_{\boldsymbol{k}}$ is its Fourier transform. The equal time correlation function is the Fourier transform of the frequency domain one, hence

$$\begin{aligned}\langle\psi|\delta n_{\boldsymbol{k}}\delta n_{\boldsymbol{k}'}|\psi\rangle &= \int \frac{d\omega_n}{2\pi}\frac{(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}')}{\tilde{V}(\mathbf{0}) + \frac{m}{n_0}\frac{\omega_n^2}{k^2}} \\ &= \sqrt{\frac{n_0k^2}{4m\tilde{V}(\mathbf{0})}}(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}') \end{aligned} \tag{257}$$

Note this has a non-analytic dependence upon k – it is proportional to $|k|$. Fourier transforming to real space implies that there is a power-law dependence,

$$\langle\psi|\delta n_{\boldsymbol{x}}\delta n_{\boldsymbol{x}'}|\psi\rangle \sim \frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|^{d+1}}, \tag{258}$$

at large separations. This is of course very different from the BEC wavefunction. To remedy it, we can multiply the BEC wavefunction by what is called a "Jastro factor" to create the desired correlations:

$$|\psi\rangle_{FB} = \exp\left[-\frac{1}{4}\int d^d\mathbf{x}d^d\mathbf{x}'v(\boldsymbol{x}-\boldsymbol{x}')\delta n_{\boldsymbol{x}}\delta n_{\boldsymbol{x}'}\right]|\psi_0\rangle. \tag{259}$$

Alternatively we can write in first quantized form

$$\Psi(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N) \propto \exp\left[-\frac{1}{2}\sum_{i<j}v(\boldsymbol{x}_i-\boldsymbol{x}_j)\right]. \tag{260}$$

Honestly I do not know the exact history but this type of wavefunction is associated with Feynmann and Bijl. We are to adjust $v(\boldsymbol{x})$ to get the desired density-density correlations to match Eqs. (257,258). This is pretty easy since actually $|\psi_0\rangle$ is totally uniform. It means that the expectation value of density operators is just a classical gas, with an effective Boltzmann weight which is

that of a gas with interaction potential $v(\boldsymbol{x}-\boldsymbol{x}')$. Since the weight is quadratic in δn, this is just a Gaussian integral, and so it can be done easily. This implies that the density-density correlation function is just

$$\langle\psi|\delta n_{\boldsymbol{k}}\delta n_{\boldsymbol{k}'}|\psi\rangle = \frac{(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}')}{\tilde{v}(\boldsymbol{k})}. \tag{261}$$

Then by comparison to Eq. (257) we have $\tilde{v}(\boldsymbol{k}) = \sqrt{\frac{4m\tilde{V}(\boldsymbol{0})}{n_0k^2}}$, and hence

$$\tilde{v}(\boldsymbol{x}) \sim \sqrt{\frac{4m\tilde{V}(\boldsymbol{0})}{n_0}}\frac{1}{|x|^{d-1}}. \tag{262}$$

We have established that a power-law form of the Jastrow factor is necessary to reproduce the density correlations of the superfluid. This shows explicitly that the superfluid state is not quite product-like.

Entanglement entropy

This will also relate to our earlier discussion of entanglement and ground states of quantum systems. We argued in Sec. 3 that for gapped systems, the ground state has the form of a product-like state, and conversely, product-like states imply the existence of a gap. This means that the ground state of a superfluid is somewhat different from a product state. We can expect it has some "extra" entanglement that a gapped phase does not have.

To get an idea of this, let us calculate the entanglement entropy for the BEC state. This is not the same as for the true superfluid state, but it gives an idea, and it is a bit of an easier exercise. We use the fixed particle number state of Eq. (251):

$$|\psi\rangle = \frac{1}{\sqrt{N!}}\left(a_0^\dagger\right)^N|0\rangle. \tag{263}$$

Now we can consider the partition of space into a region A with volume V_A and the rest, called B, with volume $V_B = V - V_A$. Let's define $y = V_A/V$ the volume fraction, $0 < y < 1$. Then we can define two new operators, $a_A = 1/\sqrt{V_A}\int_A d^d\mathbf{x}\psi(\boldsymbol{x})$ and $a_B = 1/\sqrt{V_B}\int_B d^d\mathbf{x}\psi(\boldsymbol{x})$. Then $a_0 = \sqrt{y}a_A + \sqrt{1-y}a_B$. So

$$|\psi\rangle = \frac{1}{\sqrt{N!}}\left(\sqrt{y}a_A^\dagger + \sqrt{1-y}a_B^\dagger\right)^N|0\rangle. \tag{264}$$

Expanding this out, we obtain

$$\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{N!}}\sum_{n=0}^{N}\binom{N}{n}y^{n/2}(1-y)^{(N-n)/2}\left(a_A^\dagger\right)^n\left(a_B^\dagger\right)^{N-n}|0\rangle \\
&= \frac{1}{\sqrt{N!}}\sum_{n=0}^{N}\binom{N}{n}\sqrt{n!}\sqrt{(N-n)!}\,y^{n/2}(1-y)^{(N-n)/2}|n\rangle_A\otimes|N-n\rangle_B \\
&= \sum_{n=0}^{N}\sqrt{\binom{N}{n}}y^{n/2}(1-y)^{(N-n)/2}|n\rangle_A\otimes|N-n\rangle_B \\
&\equiv \sum_{n=0}^{N}c_n|n\rangle_A\otimes|N-n\rangle_B, \qquad (265)
\end{aligned}$$

with

$$p_n = c_n^2 = \binom{N}{n}y^n(1-y)^{N-n}. \qquad (266)$$

Eq. (265) is just a Schmidt decomposition with the probabilities p_n. Hence we can obtain the entanglement entropy by just calculating $S = -\sum_n p_n \ln p_n$. Physically, this is just a classical expectation value, $S = -\langle \ln p_n\rangle$, with respect to the probability measure p_n. Now let us consider the large volume limit, i.e. both V_A and V large but y fixed, and likewise fixed density N/V, and hence the fraction of particles in region A, $x = n/N$ becomes almost equal to the volume fraction y, so x is also fixed. We write $n = xN$. Then

$$p_n \to p(x) = \binom{N}{xN}y^{Nx}(1-y)^{N(1-x)} \to \frac{1}{\sqrt{2\pi N y(1-y)}}\exp\left[-N\frac{(x-y)^2}{2y(1-y)}\right], \qquad (267)$$

where we used the leading approximation to the binomial distribution for large numbers, which is of course Gaussian. This says that x is sharply peaked around y with a variance $\langle (x-y)^2\rangle = y(1-y)/N$. Now we have

$$\begin{aligned}
S &= -\langle \ln p_n\rangle \to \ln\sqrt{2\pi N y(1-y)} + N\left\langle N\frac{(x-y)^2}{2y(1-y)}\right\rangle \\
&= 1 + \frac{1}{2}\ln(2\pi n_0 V_A(1-V_A/V))) \qquad (268) \\
&\sim \frac{d}{2}\ln(n_0 L_A) + \text{const.}, \qquad (269)
\end{aligned}$$

where the last line holds for a region of linear size L_A embedded in a much larger region ($V_A/V \to 0$).

We observe two things. First, there is no area law term, which would be proportional to L_A^{d-1}. This is missing due to the triviality of the BEC state. We should expect that even including interactions perturbatively we would recover an area law contribution. More importantly, the correction to the (zero) area law term is *non-analytic* and logarithmic in length L_A. In general, any con-

tribution to the entanglement entropy which is neither of area law form nor exponentially suppressed in size (i.e. decaying faster than any power law) indicates a non-product-like state. The additive logarithmic contribution above is such an example. It arises, clearly, due to relatively large number fluctuations in the BEC.

This result is very close to but not identical to that of the superfluid. In the latter, there is also an additive logarithmic term in the entanglement entropy, but with a different coefficient: $(d-1)/2$ instead of $d/2$ (see arXiv:1112.5166v2, where this is obtained with a lot more work than for the BEC). The coefficient is a little smaller than in the BEC, which makes sense since the number fluctuations are suppressed somewhat by interactions.

Lecture 14 (1.5h)
November 10th, 2015

Anderson tower of states

So far we've skirted around an issue: how does the superfluid manage to break the $U(1)$ particle conservation symmetry anyway? We know that for any finite closed system, the number of bosons must be conserved. So clearly the ground state must be a particle number eigenstate. This implies that under the $U(1)$ symmetry, the state transforms at most by a phase. This property in turn implies that any observable which is not particle number conserving must have a zero expectation value. Thus $\langle GS|\psi|GS\rangle$ is certainly zero. How are these features compatible with broken symmetry in the thermodynamic (infinite volume) limit?

The way it works is that, for very large but not infinite systems, levels with *almost* the same particle number become *almost* degenerate. In the infinite volume limit, the separation of these levels becomes negligibly small, and we can form superpositions of finite volume eigenstates with indefinite particle number. These states correspond to the symmetry broken ones. Let's look at this in detail. Consider the action of Eq. (215). We can "undo" the path integral construction to write the Hamiltonian this represents. To do so, we recognize the term $i\delta n\partial_\tau\theta$ as the "Berry phase" term arising from the overlap of momentum and coordinate-like basis states. It implies that as operators δn and θ are conjugate variables,

$$\left[\delta n(\boldsymbol{x}), \theta(\boldsymbol{x}')\right] = i\delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{270}$$

The rest of the action gives the time integral of the Hamiltonian,

$$H = \int d^d\mathbf{x}\left[\frac{\tilde{V}(\mathbf{0})}{2}(\delta n)^2 + \frac{n_0}{2m}|\boldsymbol{\nabla}\theta|^2\right]. \tag{271}$$

Now we can make the following canonical transformation (assuming periodic

boundary conditions in finite volume)

$$\delta n(\boldsymbol{x}) = \sum_{k\neq 0} \sqrt{\frac{\omega_k}{2V\tilde{V}(\mathbf{0})}} \left[a_k^\dagger e^{ik\cdot x} + a_k e^{-ik\cdot x}\right] + \frac{\delta N}{V}, \tag{272}$$

$$\theta(\boldsymbol{x}) = -i \sum_{k\neq 0} \sqrt{\frac{\tilde{V}(\mathbf{0})}{2V\omega_k}} \left[a_k^\dagger e^{ik\cdot x} - a_k e^{-ik\cdot x}\right] + \Theta, \tag{273}$$

where V is the system volume, and a_k and $a_k^\dagger$ are ladder operators,

$$[a_k, a_{k'}^\dagger] = \delta_{k,k'}, \tag{274}$$

and δN is an integer-valued operator,

$$\delta N = \int d^d\mathbf{x}\, \delta n(\boldsymbol{x}), \tag{275}$$

which just counts the total number of bosons in the system. It is conjugate to the zero mode operator Θ, i.e.

$$[\delta N, \Theta] = i. \tag{276}$$

Also $\omega_k = v|k|$ is the phonon energy. Plugging this into the Hamiltonian, we obtain

$$H = \frac{\tilde{V}(\mathbf{0})}{2V}(\delta N)^2 + \sum_{k\neq 0} \omega_k \left(a_k^\dagger a_k + \frac{1}{2}\right). \tag{277}$$

We see that the spectrum consists not only of the phonon modes, which begin at an excitation energy $2\pi v/L$ arising from the smallest quantized momentum $|k| = 2\pi/L$, but also a set of "zero mode" states due to non-zero integer values of δN, which begin at an energy $\tilde{V}(\mathbf{0})/V = \tilde{V}(\mathbf{0})/L^d$. We see that the zero mode states have much lower energy than the phonons for large L. This shows the very near degeneracy of states with different particle number, which go together to produce the symmetry broken states. One often makes a plot of the energy levels, with δN as the x axis, and energy on the y axis. Then the lowest levels lie on a rather flat parabola with very small curvature, and for each x value, there is a column of higher energy levels corresponding to the excitation of the phonon modes. The collection of levels is called the "Anderson tower" of states.

Off-diagonal long range order etc.

From the above discussion, we can see that it is the "zero mode", Θ, with $\boldsymbol{k} = \mathbf{0}$ which restores the quantization of total boson number in a finite system. We can ask, for a large but finite system, how do we characterize then the broken symmetry? One way is to observe the behavior of the Anderson tower of states. But we can also try to look more directly for the symmetry breaking. In a finite system with definite number, we cannot ever get a non-zero expec-

tation value of the boson field, i.e. clearly $\langle\psi(\boldsymbol{x})\rangle = \langle\psi^\dagger(\boldsymbol{x})\rangle = 0$. What is the alternative? The answer is to look at the two-point correlation function,

$$G(\boldsymbol{x}-\boldsymbol{x}') = \langle\psi(\boldsymbol{x}')\psi^\dagger(\boldsymbol{x})\rangle. \tag{278}$$

This can be considered the two-particle density matrix. It is also the equal-time boson Green's function. Since the product of the two operators conserves the total number of bosons in the system, it is non-zero. We can ask about the long-distance behavior of this function. The idea is that when $\boldsymbol{x}$ and $\boldsymbol{x}'$ are far apart (further than some correlation length), the two points are basically independent, and we *define* this to give the product of two expectation values:

$$\lim_{|\boldsymbol{x}-\boldsymbol{x}'|\to\infty} G(\boldsymbol{x}-\boldsymbol{x}') \equiv |\langle\psi\rangle|^2. \tag{279}$$

Of course for a finite system we cannot take the infinite separation limit, but we are ok so long as the system size is large compare to some relevant correlation length ξ, i.e. we want $\xi \ll |\boldsymbol{x}-\boldsymbol{x}'| \ll L$. If we like we can take successive limits, with first $L\to\infty$ and then $|\boldsymbol{x}-\boldsymbol{x}'|\to\infty$.

If the limit in Eq. (279) is non-zero, we say that there is *Off-Diagonal Long Range Order* (ODLRO). It is "off diagonal" because G so defined gives the off-diagonal term in the density matrix. The term long range order should be self-explanatory. It means that the correlations between $\psi(\boldsymbol{x}')$ and $\psi^\dagger(\boldsymbol{x})$ persist when the points are infinitely far apart – hence "long range".

How does this connect to the discussion of the Anderson tower? Well, we saw that for a large but finite system the number quantization involved only the zero mode $\theta(\boldsymbol{k}=\boldsymbol{0})$ and the conjugate N variable. This represents quantum fluctuations of the phase which are constant in space. Such fluctuations drop out completely from $G(\boldsymbol{x}-\boldsymbol{x}')$, so we can see that ODLRO is perfectly consistent with the Anderson tower.

Note that we actually could apply the same type of thinking, of symmetry breaking as a long-distance limit of a correlation function, to the Ising model. Everything is analogous. Due to the Ising symmetry, in a finite system we must have $\langle\sigma_i^z\rangle = 0$. However, we can define the magnetization by the limit,

$$\lim_{|i-j|\to\infty} \langle\sigma_i^z\sigma_j^z\rangle \equiv m^2. \tag{280}$$

When this is non-zero we say there is Long Range Order (LRO) in the Ising model. We saw in Sec. 2 and subsequent discussions that the quantum Ising model has LRO in one dimension and higher. Obviously in zero dimensions, i.e. for a finite system, there is no real symmetry breaking in this case. At non-zero temperature, the Ising model has symmetry breaking only in $d \geq 2$ (I think I did not actually discuss this yet, but I will do so in the next Section). So we can see the trend that in lower dimensions, spontaneous symmetry breaking is less stable.

Spontaneous breaking of continuous symmetry, like in the bose fluid, is less stable than the discrete symmetry breaking of the Ising model. The reason

is that fluctuations are possible into states of arbitrarily low energy, since there is no gap in the superfluid phase, in contrast to the Ising model, where there is a gap everywhere except at the QCP. This is true whenever a continuous symmetry is spontaneously broken, because there is a Goldstone mode. We can expect these fluctuation effects are stronger in lower dimensions. So we can ask, how low in dimension can we go and still spontaneously break the U(1) symmetry? In the language of this section, in how low a dimension does ODLRO persist?

We can address this by just calculating $G(\boldsymbol{x})$. The answer, as in the quantum Ising model, depends on whether we are at zero or non-zero temperature. To do the calculation, we go back to the number-phase representation, and let $\psi(\boldsymbol{x}) = \sqrt{n_0}e^{i\theta(\boldsymbol{x})}$, keeping fluctuations of θ but not of n, following the logic that led to the phase action, Eq. (217). Then we have

$$G(\boldsymbol{x}) = n_0 \left\langle e^{i(\theta(\boldsymbol{x})-\theta(\boldsymbol{0}))} \right\rangle . \tag{281}$$

We should evaluate this expectation value using the quadratic action S_θ in Eq. (217). A basic rule, which follows from the formulae we already derived for Gaussian integrals, e.g. Eq. (140), is that, if $\mathcal{O}$ is a linear function of the fields, then

$$\left\langle e^{\mathcal{O}} \right\rangle = \exp\left[\langle \mathcal{O} \rangle + \frac{1}{2}\left(\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 \right) \right]. \tag{282}$$

Applying this here, we obtain

$$G(\boldsymbol{x}) = n_0 \exp\left[-\frac{1}{2} \left\langle (\theta(\boldsymbol{x}) - \theta(\boldsymbol{0}))^2 \right\rangle \right]. \tag{283}$$

Using the Matsurbara method, we obtain

$$\frac{1}{2} \left\langle (\theta(\boldsymbol{x}) - \theta(\boldsymbol{0}))^2 \right\rangle = \frac{1}{\beta} \sum_{\omega_n} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{1-\cos \boldsymbol{k}\cdot\boldsymbol{x}}{\frac{\omega_n^2}{\tilde{V}(\boldsymbol{0})} + \frac{n_0 k^2}{m}}. \tag{284}$$

Now we can consider zero and non-zero temperature. At $T = 0$, the sum over frequencies is converted to an integral, which is easily done, to obtain

$$\frac{1}{2} \left\langle (\theta(\boldsymbol{x}) - \theta(\boldsymbol{0}))^2 \right\rangle_{T=0} = \sqrt{\frac{m\tilde{V}(\boldsymbol{0})}{4n_0}} \int_0^\Lambda \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{1-\cos \boldsymbol{k}\cdot\boldsymbol{x}}{|k|}. \tag{285}$$

It is always good to check that you get a finite integral after doing such a calculation! The integral should be cut off at large k by some microscopic scale $\Lambda \sim 1/a_0$, related to the distance between bosons etc. At small k the integral is generally well-behaved because the cosine factor goes to unity for small k. So for fixed $\boldsymbol{x}$, everything is fine. We are interested in large x, to see if there is ODLRO. At large x, the cosine term oscillates rapidly and will be negligible *except* at very small k. Since this is a small region of the integration

domain, we can try the naive limit $|x| \to \infty$ by just dropping the cosine. By power counting, we can see that the integral remains finite provided $d > 1$. This means that this correlation function has a finite limit, and hence $G(\boldsymbol{x})$ remains non-zero at infinite distance. So ODLRO persists at $T = 0$ for $d > 1$.

For $d = 1$, there is a problem. The naïve limit of dropping the cosine does not work! If we do it, we obtain a logarithmically divergent integral. In fact, in one dimension, we have

$$\int_{-\Lambda}^{\Lambda} \frac{dk}{2\pi} \frac{1-\cos kx}{|k|} \sim \frac{1}{\pi} \ln(\Lambda|x|), \qquad \text{for } \Lambda|x| \gg 1. \tag{286}$$

This implies that the boson Green's function decays to zero at infinity, but it does so as a power law:

$$G(\boldsymbol{x}) \sim \frac{1}{(\Lambda|x|)^{\eta}}, \tag{287}$$

with $\eta = \sqrt{\frac{m\tilde{V}(\mathbf{0})}{4\pi^2 n_0}}$. This power law behavior is known as (Off-Diagonal) Quasi-Long Range Order, or ODQLRO. It is intermediate between what would be expected in a truly disordered phase, such as in a bose gas at high temperature, where $G(\boldsymbol{x})$ decays exponentially at long distance, and the case of ODLRO, where it decays only to a finite value. This behavior indicates that one dimension is a special dimension for zero temperature superfluids, what is known as the (zero temperature) lower critical dimension, which separates true LRO from the complete absence of symmetry breaking in lower dimensions.

Now consider $T > 0$. Then the sum in Eq. (284) remains discrete. We should analyze the long-distance behavior of this expression. Similarly to before, at large $\boldsymbol{x}$ the cosine oscillates rapidly and we can naïvely neglect it. Then the integrand is singular only for one term in the sum, $\omega_n = 0$. So we expect that the long-distance properties are dominated by this zero Matsubara frequency mode, and write

$$\frac{1}{2}\left\langle (\theta(\boldsymbol{x}) - \theta(\mathbf{0}))^2 \right\rangle \sim_{|\Lambda x| \gg 1} \frac{k_B T m}{n_0} \int_0^{\Lambda} \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{1-\cos \boldsymbol{k}\cdot\boldsymbol{x}}{k^2}. \tag{288}$$

The divergence at $k = 0$ is worse now than at $T = 0$. We see that the integral grows logarithmically with $|x|$ in $d = 2$ instead of $d = 1$:

$$\frac{1}{2}\left\langle (\theta(\boldsymbol{x}) - \theta(\mathbf{0}))^2 \right\rangle_{T>0,d=2} \sim \frac{k_B T m}{2\pi n_0} \ln(\Lambda|x|), \tag{289}$$

leading to ODQLRO in $d = 2$ at $T > 0$ (i.e. the same behavior as Eq. (287), but with

$$\eta_{2d}(T) = \frac{k_B T m}{2\pi n_0}. \tag{290}$$

In one dimension at $T > 0$ the correlation function in Eq. (288) grows linearly with $|x|$, and so $G(\boldsymbol{x})$ decays exponentially. There is no superfluidity in any sense in one dimension at non-zero temperature.

The behavior of the boson Green's function is nice theoretically, but it is a bit formal. It does not tell us whether there is superfluidity in 1d at $T = 0$ or in 2d at $T > 0$ in the sense of superflow, e.g. persistent currents, zero viscosity, etc. This is a subtle question and the answer is, not really, but the behavior may be almost superfluid like, depending upon what exactly you measure. To explain further is beyond the level of discussion I'd like to give here in this class.

Lecture 15 (1.5h)
November 12th, 2015

7 Domain walls, the 1+1d Ising model, and duality

So far we have discussed the quantum Ising model and the bose gas, and looked at their ground state and elementary excitations. We were able to show in both cases that, away from any phase transition, the low energy excitations are quasiparticle like. This means they behave like particles, for example one can form a wave packet out of them, and one can assign a definite energy-momentum relation to them. In the Ising model's disordered phase, the quasiparticle is a spin flip relative to the field, while in its ordered phase, at least in $d > 1$ it is a spin flip relative to the ferromagnetic background. Both of these are not only quasiparticle-like but also *local*, insofar as a single excitation can be created by a local operator, i.e. a single σ_i^z or σ_i^x operator in these two cases. In the superfluid, the phonon is also local, which we know because the density-density correlation function contains a delta-function peak, which implies that a single phonon can be create by the density operator. So in all cases we have local quasiparticle excitations.

This is *not* necessarily the case in all phases of matter, even in gapped ones. Indeed we already had a hint of this in Sec. 2.1, where we saw that in a *one-dimensional* Ising model in its ordered phase, the minimum energy excitation is actually a domain wall or soliton. It is (as we will see) a quasiparticle, and in most respects behaves just the other quasiparticles we discussed, except for one important difference. It is not local. This means that there is no local operator which can create a single domain wall. This is simply because an infinite number of spins must be flipped to create a single domain wall, and so the operator which does it must be (semi)-infinite.

7.1 *Ising duality*

Actually we can make this all very explicit through what is called a *duality* transformation. We begin with the 1d quantum Ising Hamiltonian,

$$H = \sum_i \left[-J\sigma_i^z \sigma_{i+1}^z - h\sigma_i^x \right]. \tag{291}$$

Let us try to change variables from the σ basis which describes the spins to ones which describe domain walls. A complete basis of states is the set of eigenstates of σ_i^z for all i. Each element of the basis is specified by a string of N

bits, with N sites, with each bit given by state of σ_i^z. Now for each such string, i.e. classical arrangement of Ising spins, we can determine whether each bond, connecting neighboring spins, is satisfied or not. If it is satisfied, i.e. the two spins are parallel, we say there is no domain wall there. If it is not, i.e. the two spins are anti-parallel, we say there is a domain wall there. Mathematically, we can assign a "dual" spin $\tau^x_{i+\frac{1}{2}} = \pm 1$ to each bond according to

$$\tau^x_{i+\frac{1}{2}} = \sigma^z_i \sigma^z_{i+1}, \tag{292}$$

where $\tau^x = +1$ if there is no domain wall, and $\tau^x = -1$ is there is a domain wall. Let's assume for the moment periodic boundary conditions. Then given σ_i^z values we uniquely determine the τ_a^x values. The converse is almost true. Suppose we are given a set of τ_a^x values for all bonds. Then we can assign the first spin σ_1^z arbitrarily, and then by knowing $\tau^x_{1/2}$ we obtain from Eq. (292) σ_2^z, and by induction, the remaining spins. But with periodic boundary conditions, we must get back to the same original spin. This means there must be a constraint,

$$\prod_a \tau_a^x = \prod_{i=1}^N \tau^x_{i+\frac{1}{2}} = 1. \tag{293}$$

This just means physically that in a periodic system, there is always an even number of domain walls: each domain has two ends. If this constraint is satisfied, then there are *two* linearly independent physical spin states for each set of τ_a^x variables.

To fix this, we can take advantage of the symmetry of the problem. The Ising symmetry means that H commutes with the generator of the symmetry,

$$U = \prod_i \sigma_i^x = \pm 1. \tag{294}$$

So we can work with states of fixed eigenvalue of $U = \pm 1$. This choice of U is the additional quantum number we need to make the set of domain wall variables complete. In other words we can trade

$$\{\sigma_i^z\} \leftrightarrow \{U, \tau_a^x\}, \tag{295}$$

with the constraint of Eq. (293). It is interesting to note the similarity of Eq. (293) to Eq. (294). The difference is that only one sign, the positive one, of the product over τ_a^x is physical, while both signs of the product over σ_i^x are physical.

Now let us look at the conjugate variables. The operator σ_i^x flips the i^{th} spin, which therefore creates or annihilates domain walls on two bonds which share site i. Thus we expect that

$$\sigma_i^x = \mu_i \tau^z_{i-\frac{1}{2}} \tau^z_{i+\frac{1}{2}}, \tag{296}$$

where $\mu_i = \pm 1$ is a c-number, not an operator. Independent of the choice of μ_i, this achieves the desired sign flip since τ_a^z anticommutes with τ_a^x. The parameters μ_i are almost arbitrary and just represent a convention. There is a constraint, however, when we recognize that we work in a subspace of fixed U. By explicitly multiplying, we have

$$U = \prod_i \mu_i, \tag{297}$$

so we must choose the μ_i to obtain the desired value (± 1) of U. For the even parity sector, $U = +1$, we can take simply $\mu_i = 1$. For the odd parity sector, $U = -1$, we can choose for example $\mu_1 = -1$ and $\mu_i = +1$ for $i > 1$.

One can check that Eq. (296) and Eq. (292) are consistent. That is, if we assume that σ_i^z and σ_i^x anticommute on the same site and commute on different ones, then the same holds for τ_a^z and τ_z^x. It is almost obvious due to the symmetry of these two equations. You might be dis-satisfied with Eq. (296), since it expresses σ_i^x in terms of τ_a^z and not the other way around. What is a formula for τ_a^z in terms of σ_i^x? Well, strictly speaking there is none! The reason is that the physical space obeys the constraint of Eq. (293). Only operators which preserve this constraint are physical. A single τ_a^z operator anticommutes with $\prod_a \tau_a^x$ and so violates the contraint: there is no way to create a single domain wall. Instead, a pair of domain walls can be created, and there is a good formula for this:

$$\tau_a^z \tau_b^z = \prod_{Min(a,b)<i<Max(a,b)} \mu_i \sigma_i^x. \tag{298}$$

If we are a little bit cavalier – not very by physicists' standards – we can imagine "pushing" one domain wall off to $-\infty$ and writing a formula for a single domain wall operator in an infinite system,

$$\tau_a^z \text{ ''} = \text{''} \prod_{i<a} \mu_i \sigma_i^x. \tag{299}$$

This formula is correct in the sense that if we apply it to any product of an even number of τ_a^z operators it gives the right answer. It is rather physical: to create a single domain wall we need to flip a semi-infinite string of spins to the left of the wall. The nice thing about Eq. (299) is that it makes it explicit that τ_a^z is a non-local operator.

We expect the ground state will be in the even parity sector. So let's assume $U = +1$ and proceed – this only captures half the states. Then we can take $\mu_i = +1$ and simplify the calculations. It is straightforward to use Eq. (292) and Eq. (296) to rewrite the Hamiltonian, $H_\pm = P_\pm H P_\pm$, with $P_\pm = (1 \pm U)/2$ the projection operator onto the $U = \pm 1$ sector. We obtain

$$H_+ = \sum_a \left[-J\tau_a^x - h\tau_a^z \tau_{a+1}^z \right]. \tag{300}$$

Remarkably, this has an *identical* form to H written in terms of the original spins, but with $h \leftrightarrow J$. We say that the transformation from σ to τ variables

is a *duality* transformation, and the one dimensional quantum Ising model is *self-dual*.

The duality implies that if the model with parameters (J, h) is critical, then so is the model with parameters (h, J). Now from Sec. 2.2 we expect that there is single quantum critical point. If this is the case, then we can determine its location exactly: the only possible unique point is $h = J$. So we conclude that $h_c = J$ for the 1d quantum Ising model! Recall that in MFT we found $h_c = 2dJ = 2J$ in one dimension. So we see that the mean field result is off by a factor of two. It makes sense that the true h_c is smaller than the mean field one, since MFT underestimates fluctuations.

7.2 *Two-soliton continuum*

What else can we do with the duality? Well, because it interchanges h and J, it interchanges the ferromagnetic and paramagnetic phases. That is, the physical ferromagnetic phase in which $\langle \sigma_i^z \rangle \neq 0$ corresponds to the apparently paramagnetic phase of the dual theory, where $\langle \tau_a^z \rangle = 0$.[5] This means that to study the domain wall excitations of the original Ising model, which exist in the FM phase, we can study the paramagnetic phase of the dual Hamiltonian, Eq. (300). This is a slight simplification since deep in the FM limit, $h \ll J$, the dual Hamiltonian in Eq. (300) is to leading order just a decoupled sum over "fields" τ_a^x on individual dual sites. A single domain wall of the original model corresponds in the $h = 0$ limit to a configuration with $\tau_b^x = +1$ everywhere except for a single site with $\tau_a^x = -1$, where the domain wall resides. This is mathematically identical to the single spin-flip excitation of the original problem.

Now we can use the duality mapping to investigate the spin-spin correlation function, to look to see what happens to the quasiparticle peak in the FM phase. Recall that in the PM phase, we calculated the spin-spin correlation function using field theory methods in Sec. 5.5. As a consequence of local spin-flip quasiparticles, the corresponding spectral function has a delta-function pole, c.f. Fig. 3. Now let us consider the spin-spin correlation function in the PM phase. Specifically, consider the σ^x correlation function,

$$C(x,t) = \langle \sigma_{i+x}^x(t) \sigma_i^x(0) \rangle, \tag{301}$$

and the corresponding spectral function obtained by Fourier transformation (we use real-time dynamics here just to keep it simple). According to the duality transformation, we have

$$C(x,t) = \langle \tau^z_{i+x+\frac{1}{2}}(t) \tau^z_{i+x-\frac{1}{2}}(t) \tau^z_{i+\frac{1}{2}}(0) \tau^z_{i-\frac{1}{2}}(0) \rangle. \tag{302}$$

This formula is correct and involves no μ_i factors at zero temperature because the ground state has $U = +1$ and also σ_i^x commutes with U and so the intermediate states are also in this sector. If we imagine inserting a complete set of intermediate states, we see that contributions occur from states which have

[5] Strictly speaking, we should use correlation functions instead of expectation values, to avoid using a single unphysical τ_a^z operator.

overlap with two dual "spin flips" acting on the ground state.

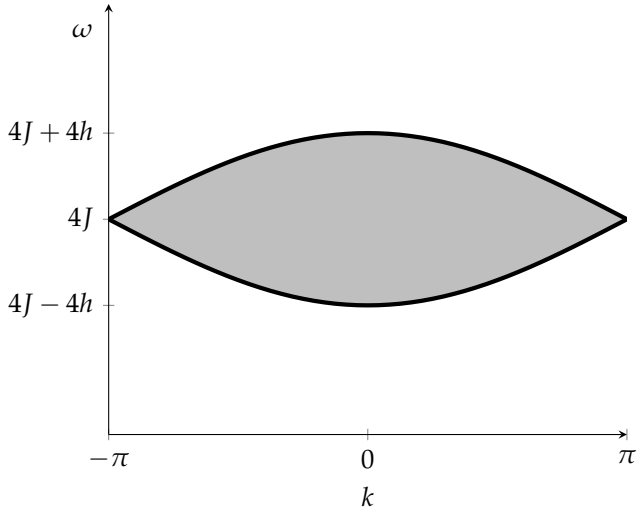


Figure 6: Domain of spectral weight from two-soliton excitations

Let us evaluate this a little schematically. In the dual Hamiltonian, the small h limit corresponds to the small exchange one, and we obtained the wavefunction and energy of a single dual spin flip excitation in Eq. (10) and Eq. (11). In the current notation, specializing to $d = 1$, we have:

$$|k\rangle \propto \sum_a e^{ikx_a} \tau_a^z |0\rangle, \tag{303}$$

which has an energy, using the duality dictionary, of $\epsilon(k) = 2J - 2h\cos k$. Now we expect that the Fourier transform $C(k,\omega)$ obtains contributions from states with two dual spin flips, with total momentum k and total energy ω. The weight of each such contribution is the square of a matrix element, $|M| \sim |\langle k_1 k_2 | \tau_{i+\frac{1}{2}} \tau_{i-\frac{1}{2}} |0\rangle|^2$, which is something like the probability to find the two particles one site apart. Let us assume this is momentum independent. Then we will have, approximately

$$\begin{aligned} S(k,\omega) &\sim \int dq\, \delta(\omega - \epsilon(q) - \epsilon(k-q)), \\ &\sim \int dq\, \delta(\omega - 4J - 4h\cos(k/2)\cos(q)), \\ &\sim \frac{\Theta[4h|\cos\frac{k}{2}| - |\omega - 4J|]}{\sqrt{(4h\cos\frac{k}{2})^2 - (\omega - 4J)^2}}. \end{aligned} \tag{304}$$

We see that there is a continuum of weight for $4J - |4h\cos\frac{k}{2}| < \omega < 4J + |4h\cos\frac{k}{2}|$. There is no quasiparticle pole in sight! This is a sign of non-local quasiparticles. Because the elementary domain walls are non-local, they cannot be created singly by *any* local operator, so a dominant continuum should be expected in any correlation function, not only that of the σ_i^x operator, which

we studied here.

We won't do it now, but I should point out that the one dimension quantum Ising model is actually exactly soluble by what is called a Jordan-Wigner transformation, which is a non-local change of variables which converts the problem to a free fermion chain. Using this method one can calculate the ground state energy and σ_i^x correlation functions exactly. In that way one can obtain the two-soliton continuum without any approximations, and confirm the above result. We might return to this solution when we talk about intrinsic topological order.

Lecture 16 (1.h)
November 17th, 2015

Confinement

It is interesting to ask what happens in the presence of a small applied longitudinal field, by adding to the Hamiltonian the term

$$H' = -\sum_i h_\parallel \sigma_i^z. \tag{305}$$

We suppose $h_\parallel \ll J, h$, so it is really a weak perturbation. Yet it is important because it explicitly breaks the Ising symmetry. Consequently, it splits the degeneracy of the two ferromagnetic ground states, even at the classical level. The energy difference between the two states is a small number, $2h_\parallel$ *per site*. The total energy difference between the two ground states is $2h_\parallel N$, which becomes arbitrarily large for a large system. So the "wrong" ground state is pushed up to infinite energy in the thermodynamic limit.

What about the excitations? Well, because of the field, the classical energy of a domain wall is no longer independent of its length. So an isolated domain wall in an infinite system also has infinite energy. This means that completely non-local excitations are no longer possible: the solitons are *confined*. Let us look at how this works in a little more detail.

Start by thinking classically, i.e. $h = 0$, and we have a single domain of length ℓ. It has an energy cost of $4J$ for the two endpoints plus a contribution of $2h_\parallel \ell$ from the domain. So there is a linear "confining potential" between the two domain walls. The lowest energy configuration is just when $\ell = 1$, i.e. a single flipped spin ($\ell = 0$ is just the ground state). Now when $h > 0$, this is not quite correct, because the walls have kinetic energy, and so their zero point energy competes with the confining potential. It is possible to write down a lattice Schrödinger equation for the two domain wall states, just by projecting the full Hamiltonian into the manifold of these states. But we will do something simpler. Let's assume $h_\parallel \ll h$, so that the kinetic energy actually dominates. Then the solitons are "almost" free, although when they get really far apart they will eventually feel the confining potential. Let us focus on the lowest energy two soliton states, with energy close to $4J - 4h$. This originates from two domain walls whose momenta are both close to $k = 0$. So we can expand the dispersion relation around $k = 0$ and approximate it by an effective mass. The energy of two domain walls with momenta k_1 and k_2 is then

$$\epsilon(k_1, k_2) \approx 4J - 4h + \frac{k_1^2}{2m} + \frac{k_2^2}{2m}, \tag{306}$$

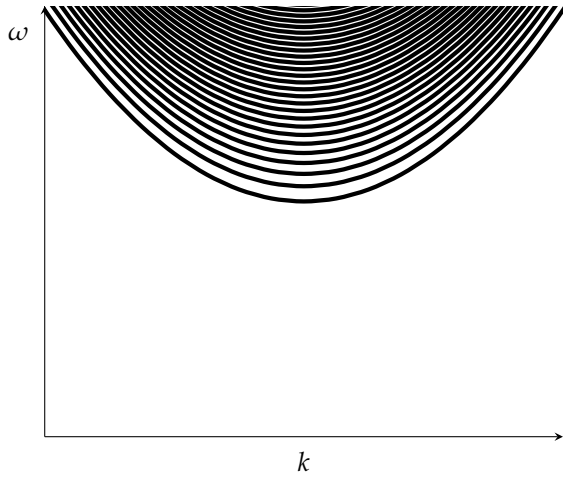


Figure 7: Breakup of bottom of the two-soliton continuum into magnon-like bound states.

with $m = 1/(2h)$. Now we can write a continuum Schrödinger equation including the effect of the parallel field:

$$\left[4J - 4h - \frac{1}{2m}\frac{\partial^2}{\partial x_1^2} - \frac{1}{2m}\frac{\partial^2}{\partial x_2^2} + h_\parallel |x_1 - x_2|\right] \psi(x_1, x_2) = \epsilon \psi(x_1, x_2). \tag{307}$$

In terms of the relative wavefunction, we have $\psi(x_1, x_2) = e^{ik(x_1+x_2)/2}\psi(x_1 - x_2)$, with

$$\left[-\frac{1}{2m^*}\frac{\partial^2}{\partial x^2} + h_\parallel |x|\right] \psi(x) = \Delta\epsilon \psi(x), \tag{308}$$

where $m^* = m/2$ is the reduced mass, and $\Delta\epsilon = \epsilon - (4J - 4h + \frac{k^2}{4m})$. We note that we should restrict to $x > 0$, because since the FM ground state is definite (all spins up), the domain wall which flips up to down with increasing x must always be to the left of the one which flips down to up with increasing x. So we have $\psi(x \leq 0) = 0$, and this is the Schrödinger equation for a particle in a linear (triangular) potential well and a hard wall, whose solutions are Airy functions. The spectrum is

$$\Delta\epsilon_n = \left(\frac{h_\parallel^2}{2m^*}\right)^{1/3} |z_n|, \tag{309}$$

where the z_n are the zeros of the Airy function. We can understand the behavior by dimensional analysis. A bound state confined to a length ℓ has a kinetic energy of order $1/m^*\ell^2$ which should be balanced by the potential energy $h_\parallel \ell$. Equating the two we obtain $\ell \sim 1/(m^* h_\parallel)^{1/3}$, and then forming the energy $h_\parallel \ell$ we obtain the prefactor in Eq. (309).

This means that the continuum in Eq. (304) breaks up into a series of sharp peaks, representing a whole set of bound states formed by the linear confining potential (see Fig. 7). Each bound state is quasiparticle-like, but the narrow spacing of the peaks, and consequent small weight in each peak, is a remnant of the non-locality of the excitation in zero applied parallel field.

Amazingly, this sequence of Airy-function states has actually been seen in experiment. Please look at the beautiful experimental paper (and especially Figure 3B) by R. Coldea *et al*.

8 Symmetry protected topological phases

Thus far, everything we discussed seemed to be fully determined by symmetry – of the Hamiltonian and of the ground state. In our Ising model, the Hamiltonian had Z_2 symmetry and the ground state might or might not preserve this symmetry, which thereby determined two phases, and only two phases, of the model. Similarly, the Bose gas at low temperature breaks the $U(1)$ symmetry of the Hamiltonian, while at high temperature it does not, and these are the only two phases of that problem.

However, this is not the only way to distinguish phases. There can be multiple phases with exactly the same symmetry. This might already seem familiar from experience with H_2O: water can be liquid or gas which both have the same symmetry but you can observe a phase transition when water boils. However, classical liquids and gases are really the same phase insofar as one can, by applying both pressure and temperature, connect the two continuously. A stronger distinction is possible: there can be two phases of exactly the same symmetry which can never be connected continuously. These can fully legitimately be called different phases.

This distinction is by construction "topological". In general we consider some space of allowed (with some specified symmetries, plus locality) Hamiltonians (e.g. coupling constants) which have non-critical ground states. For example, we could require that they are gapped (which is a bit of a too strong requirement in general but is the easiest to define). The existence of multiple phases in the sense that there are states that cannot be deformed into one another continuously means that the space of Hamiltonians with the gap requirement contains disconnected components.

This is rather abstract, and formal since the space of Hamiltonians is obvious very large. But topology enters into the discussion in other, less abstract, ways. Actually we already encountered topology a little bit when we talked about vortices in the Bose superfluid. These are topological defects in an $U(1)$ order parameter. So they describe excitations or metastable configurations of the ordered state. The equilibrium or ground state itself does not contain any vortices, or any topological aspects. Here we discuss topological characterizations of the ground state itself. One of the consequences of the existence of multiple distinct phases with the same symmetry is that it is difficult to bring them together spatially. That is, a boundary between two of these topologically distinct phases with the same symmetry often has some kind of "breakdown" phenomena associated with it. Something unexpected happens there, most of

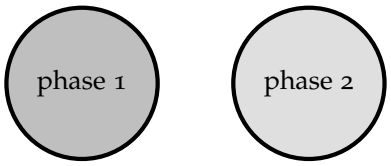


Figure 8: Space of non-critical ground states of local Hamiltonians with some given symmetry, in the case of a Z_2 topological classification. You can think of the space of Hamiltonians itself, or of for example tensor network states. It is important to impose that the states are potential ground states of some physical, i.e. local, Hamiltonian, and that they are not at some phase transition. A strong requirement might be that there is a gap above the ground state.

the time.

It has become extremely popular theoretically to classify phases, i.e. to give an *exhaustive* accounting for the different phases that can occur in some given symmetry class, and usually with some additional constraints. This is also exhausting, and we do not have time for it, nor do I have the inclination. We will instead proceed by example.

Lecture 17 (1.5h)
November 19th, 2015

8.1 *One dimensional topological insulator*

One of the first places that topology of ground states entered condensed matter was in the Su-Schrieffer-Heeger model which was supposed to describe electronic states of some polymers. It makes a nice illustration of topology, though it is really physics of non-interacting particles, not many-body physics as we are supposed to be studying in this course. Still it is a simple place to start, and it connects pretty nicely to our next example.

Basic model: one dimensional topological insulator with chiral symmetry

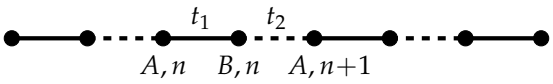


Figure 9: 1d lattice with 2 site basis.

So let us consider a particle, which could be an electron, hopping in a one dimensional lattice. Our lattice has a two site unit cell, so we define two sublattices, type A and B, which alternate. We label them A, n and B, n, as shown in Fig. 9. The state with the particle on site n of type A is $|A, n\rangle$ and similarly for the B sublattice. We consider the Schrödinger equation:

$$\begin{aligned} H|A,n\rangle &= v|A,n\rangle - t_1|B,n\rangle - t_2|B,n-1\rangle, \\ H|B,n\rangle &= -v|B,n\rangle - t_2|A,n+1\rangle - t_1|A,n\rangle. \end{aligned} \tag{310}$$

where v is a staggered on-site potential and t_1 and t_2 are real hopping amplitudes. The eigenstates of this model are easily found by using Bloch's theorem.

We write

$$|\psi_k\rangle = \sum_n \sum_{a=A,B} u_a e^{ikn} |a,n\rangle, \tag{311}$$

and observe that this is an eigenstate provided

$$\mathsf{H}(k)u = \epsilon u, \tag{312}$$

with

$$\mathsf{H}(k) = \begin{pmatrix} v & -t_1 - t_2 e^{-ik} \\ -t_1 - t_2 e^{ik} & -v \end{pmatrix}, \tag{313}$$

and

$$u = \begin{pmatrix} u_A \\ u_B \end{pmatrix}. \tag{314}$$

If we impose some extra symmetry, we can require $v = 0$. The simplest thing to do is to impose *chiral* symmetry, which is requires

$$\sigma^z \mathsf{H}(k) \sigma^z = -H(k), \qquad \text{(chiral symmetry)}. \tag{315}$$

In real space, this corresponds to the condition that hopping is only between the two sublattices, never within one sublattice.

It is straightforward to diagonalize $\mathsf{H}(k)$. The eigenvalues are simply

$$\epsilon_\pm(k) = \pm|t_1 + t_2 e^{ik}| \equiv \pm|f(k)|, \tag{316}$$

with $f(k) = t_1 + t_2 e^{ik}$. There are two bands which are separated at each k by a energy difference $2|f(k)|$ unless $f(k) = 0$, where touching occurs. Such vanishing really requires two conditions, since $f(k)$ is a complex function, so that both its real and imaginary parts must vanish. The vanishing occurs only when $k = 0, \pi$ and requires $|t_2| = |t_1|$.

Winding number

Now with the condition $v = 0$, we can see that there are actually distinct classes of Hamiltonians, which differ in topology. In particular, provided the gap never closes, then $|f(k)| > 0$ for all k, and hence we can define the the phase $\theta(k)$ via $f(k) = |f(k)|e^{i\theta(k)}$. For a smooth f, $\theta(k)$ is defined modulo an overall constant integer multiple of 2π. Then we can consider a winding number,

$$w = \frac{1}{2\pi} \int_0^{2\pi} dk \, \frac{d\theta}{dk}, \tag{317}$$

which must be an integer. Equivalently, we can define $\hat{f}(k) = f(k)/|f(k)|$, which has unit modulus. Then

$$w = \frac{-i}{2\pi} \int_0^{2\pi} dk\, \hat{f}^* \frac{d}{dk} \hat{f}. \tag{318}$$

By inspection, for $|t_2| < |t_1|$, $w = 0$. However, if $|t_2| > |t_1|$, then $w = 1$. In principle, we could obtain higher winding numbers, if we included longer-range hopping (between sites of opposite sublattices).

Note that if we take $t_1 = t_2$, then the system is at a quantum phase transition between topological phases. At that point $f(k)$ vanishes at $k = \pi$ ($k = -\pi$ is equivalent), so there are gapless excitations whose energy vanishes linearly with $q = k - \pi$: this is a one-dimensional Dirac fermion.

Boundary states and Dirac equation

What is the significance of the winding number? By construction it separates band structures into different classes. It turns out that the boundary between different classes of insulators has a protected state there. Let's look at this explicitly. Let us consider a Hamiltonian of the form in Eq. (310), but with spatially-dependent hopping, $t_1 \to t_1(n)$ and $t_2 \to t_2(n)$, and let us furthermore assume chiral symmetry $v = 0$ and moreover (which is a useful simplification) that $|t_1| - |t_2| \ll |t_1|, |t_2|$. This means that the system is everywhere close to the phase transition between the two topological phases. In this case, we can focus on the states near $k = \pi$. So we modify Eq. (310) to make t_1 and t_2 functions of n, and use the ansatz

$$|\psi\rangle = \sum_n \sum_{a=A,B} u_a(x = n)(-1)^n |a, n\rangle, \tag{319}$$

where the small wave-vector q is accounted for by variations of $u_a(x) \sim e^{iqx}$. Now we apply the Hamiltonian to this state:

$$\begin{aligned}
H|\psi\rangle &= \sum_n \sum_a u_a(x = n)(-1)^n H|a, n\rangle \\
&= \sum_n \sum_a u_a(x = n)(-1)^n \left(-t_1|-a, n\rangle - t_2|-a, n-a\rangle\right), \\
&= -t_1 \sum_n \sum_a u_{-a}(x = n)(-1)^n |a, n\rangle + t_2 \sum_n \sum_a u_{-a}(x = n-a)(-1)^n |a, n\rangle \\
&= \sum_n \sum_a \left(-t_1 u_{-a}(x = n) + t_2 u_{-a}(x = n-a)\right)(-1)^n |a, n\rangle, \\
&\approx \sum_n \sum_a \left(-(t_1 - t_2) u_{-a}(x = n) - t_2 a \frac{d}{dx} u_{-a}(x = -n)\right)(-1)^n |a, n\rangle,
\end{aligned}$$

where in the last line we assumed that $u_a(x)$ is a slowly-varying function of x, which is equivalent to assuming q is small. We also used the notation $a = A = 1$ and $a = B = -1$ to keep the algebra compact. Imposing the

eigenstate condition, $H|\psi\rangle = \epsilon|\psi\rangle$, we find

$$-(t_1 - t_2)u_{-a}(x = n) - t_2 a\frac{d}{dx}u_{-a}(x = -n) = \epsilon u_a(x). \tag{320}$$

Collecting this into the spinor u in Eq. (314), we have

$$\mathcal{H}u \equiv \left[-(t_1 - t_2)\sigma^x - it_2\sigma^y\frac{d}{dx}\right]u = \epsilon u, \tag{321}$$

which defines the continuum Hamiltonian

$$\mathcal{H} = -iv\sigma^y\frac{d}{dx} - m\sigma^x, \tag{322}$$

with

$$v \approx t_1 \approx t_2, \qquad m = t_1 - t_2 \ll t. \tag{323}$$

Here we will consider v constant, but $m = m(x)$ varying from $m = -M$ for $x \to -\infty$ and $m = +M$ for $x \to +\infty$, which describes an interface between the two topological phases. Note that if we consider m constant, then the eigenstates of Eq. (322) are just plane waves with energy $\pm\sqrt{m^2 + v^2q^2}$, which is consistent with the lattice solution. What happens if m changes sign? Then it seems possible there is some special state localized near where the sign change happens. Let us look for an eigenstate of $\mathcal{H}$ with zero energy $\epsilon = 0$. Then $\mathcal{H}u = 0$ implies

$$\left(-m(x) - v\frac{d}{dx}\right)u_B = 0, \tag{324}$$

$$\left(-m(x) + v\frac{d}{dx}\right)u_A = 0. \tag{325}$$

The solutions are

$$u_B(x) = u_{B0}e^{-\int_0^x dx'\frac{m(x')}{v}}, \tag{326}$$

$$u_A(x) = u_{A0}e^{+\int_0^x dx'\frac{m(x')}{v}}. \tag{327}$$

Now we see that for $m(x) \to \pm M$ for $x \to \pm\infty$, only the u_B solution is normalizable if $M > 0$, and conversely only the u_A solution is normalizable if $M < 0$. If instead we had taken $m(x)$ *not* to change sign, neither solution would have been normalizable. We see that when the mass $m(x)$ changes sign, there is one and only one bound state at zero energy. This lies *inside* the gap for delocalized plane-wave states far from the boundary. This phenomena of a zero energy mode when the mass gap changes sign in the 1d Dirac equation is a well-known phenomena, discovered by Jackiw and Rebbi and others.

The existence of a zero energy bound state is a robust topological feature guaranteed at the interface between the two topological insulators when chiral symmetry is preserved. One way to understand the robustness beyond the

above continuum Dirac treatment is to think about chiral symmetry directly. It implies that, provided the chiral symmetry is preserved, even in the discrete lattice model with inhomogeneous hopping, states come in pairs with energy ϵ and $-\epsilon$. The zero energy state is special: the chiral symmetry does not map this to another state. Such a state is "locked" to zero energy because if it were to shift away from zero it would need to bifurcate into two states, and you cannot make two states out of one. The only way for the state to disappear is for the energy gap to close, at which point other zero energy states appear and can pair up with this one. Of course, a zero energy state, so long as there is a gap, must be a bound state, since it lies in the forbidden region for bulk propagating states. This argument does not tell you that there must be a zero energy state in a given model with chiral symmetry, but it does tell you that once it is present, it is robust to arbitrary perturbations.

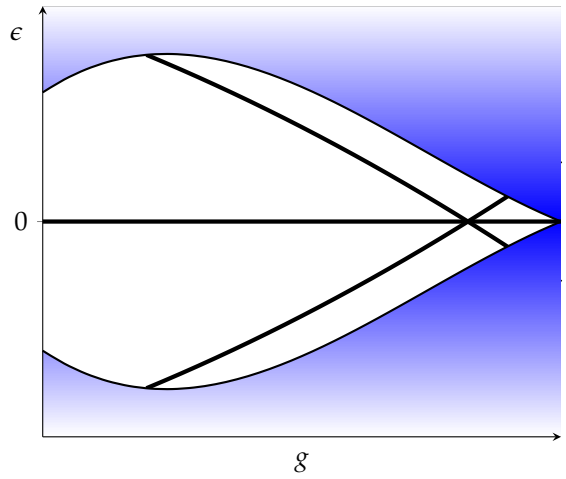


Figure 10: Schematic evolution of the spectrum of a boundary between two topological phases (or an endpoint of the non-trivial one) with chiral symmetry, as a function of some Hamiltonian g. A zero energy bound state is robust because it cannot split and any other bound states approach it in pairs. It can be removed only if the continuum (shaded region) of unbound states moves to zero energy, signaling a bulk phase transition, shown here at $g = 1$.

A physical picture for the two topological phases can be given by an illustration, Fig. 11, which accurately describes the two extreme limits $t_1 \gg t_2$ and $t_2 \gg t_1$. The two bands correspond to bonding and anti-bonding states on the strong bonds, which are shifted relative to one another by one site in the two phases. One can see that at the end a chain in which the strong bonds are broken, or at the interface between two chains, there is a site left out, which corresponds to an additional state for a particle, which is neither a bonding or anti-bonding state, and in fact has zero energy.

Lecture 18 (1.5h)
November 24th, 2015

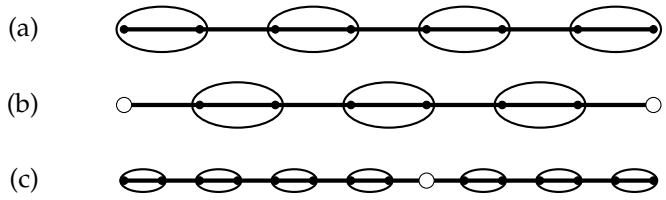


Figure 11: Topological phases in one dimension. In (a) and (b), two possible phases are shown, corresponding to different winding numbers in the chiral symmetry class. The states are visualized as bonding or antibonding states on the strong bonds, indicated by ellipses. For a properly terminated chain, which preserves chiral symmetry, one of the phases has zero energy bound states at the ends (shown as open circles), while the other does not. The interface between the two phases, shown in (c), also has a zero energy bound state (again if chiral symmetry is preserved at the interface).

Polarization and Berry phase

We can obtain the distinction between these two topological classes by looking at states instead of wavefunctions. We notice that for the lower energy band, with energy $\epsilon_-(k)$, the periodic part of the Bloch wavefunction is

$$u \equiv |u\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \hat{f} \end{pmatrix}. \tag{328}$$

Then we see that

$$w = -2\frac{\gamma}{2\pi}, \tag{329}$$

with

$$\gamma = \int_0^{2\pi} dk \, \langle u | i\partial_k | u \rangle. \tag{330}$$

The object γ is known as the *Zak phase*. According to the prior result, it is quantized to be either zero or π.

Let us further discuss the Zak phase. It actually is an example of a Berry phase. This is because the integral is over the full Brillouin zone, which makes it a periodic orbit: we can choose the states $|\psi_k\rangle$ to be smooth periodic functions of k. You can think of it as the geometrical phase accumulated by a particle as it moves in momentum space, for example if it were driven by a very weak electric field. The quantization can be understood in more general terms, provided we assume *inversion symmetry*. First of all, regardless of symmetry, there is an ambiguity in the definition, Eq. (330), of the Zak phase. This is because one can always choose different phase conventions for the Bloch wavefunctions, which may depend upon k. Two conventions differ by the "gauge transformation" $|u\rangle \to e^{i\chi(k)}|u\rangle$, with $\chi(k)$ a 2π-periodic function of k. This implies γ is defined only up to a multiple of 2π. Now consider the

implications of inversion symmetry. This maps $k \to -k$ and interchanges A and B sublattices. It implies that the state $\sigma^x|u,-k\rangle$ is a good eigenstate of $\mathsf{H}(-k)$ with the same energy as $|u,k\rangle$. However, a priori the state we obtain this way may differ in phase convention from our starting one. Hence we have

$$\sigma^x|u,-k\rangle = e^{i\chi(k)}|u,k\rangle, \tag{331}$$

where again $\chi(k)$ should be a 2π-periodic function. Starting from the definition of the Zak phase, we change $k \to -k$ in the integral and apply Eq. (331)

$$\begin{aligned}
\gamma &= -\int_0^{2\pi} dk\, \langle u,-k|i\partial_k|u,-k\rangle \\
&= -\int_0^{2\pi} dk\, \langle u,-k|\sigma^x i\partial_k \sigma^x|u,-k\rangle \\
&= -\int_0^{2\pi} dk\, \langle u,k|e^{-i\chi(k)} i\partial_k \left(e^{i\chi(k)}|u,k\rangle\right) \\
&= -\gamma + \int_0^{2\pi} \partial_k \chi(k) = -\gamma + 2\pi q, \qquad (332)
\end{aligned}$$

with some integer q. This implies that $\gamma = 0, \pi$ modulo 2π.

Note that inversion symmetry leads to a $\mathbb{Z}_2$ classification, i.e. there are only two distinct topological phases, while for chiral symmetry in principle there is an integer winding number (though only two values of it are realized for this simple model).

A physical interpretation of the Zak phase is in terms of *electric polarization* of the electronic system, if all the states in the lower band are occupied. To understand this, you can consider a Wannier function,

$$|X\rangle = \int_0^{2\pi} \frac{dk}{2\pi}\, e^{ikX}|\psi_k\rangle, \tag{333}$$

As is well-known, Wannier states for bands separated by gaps are well-localized functions and form a complete and orthonormal basis (by including all values of X). Now one can measure the mean particle position in such a Wannier state, $\langle X|\hat{x}|X\rangle$. One obtains, by using Eq. (311),

$$\langle X|\hat{x}|X\rangle = \frac{\gamma}{2\pi}. \tag{334}$$

This is the Wannier center, and we see it can be 0 or 1/2. It is entirely consistent with the pictures of Fig. 11(a,b), and simply corresponds to the two choices of the center of the bonds. The Zak phase, or equivalently the Wannier center, is a property of the entire band, not of a single momentum state. For electrons, if

the Fermi energy is in the gap separating the two bands, then all these states are occupied, and this can be considered a ground state property. This is an example of a topological distinction between ground states.

These arguments tell us that there are two distinct topological classes of bands. This means that there should still be a gap closing (phase transition) if one evolves from one class to another. Let us check this. If our lattice has inversion symmetry, then we require in general

$$\sigma^x \mathsf{H}(k)\sigma^x = H(-k). \tag{335}$$

Such a Hamiltonian in general has the form

$$\mathsf{H}(k) = \begin{pmatrix} \epsilon_0(k) + v(k) & f(k) \\ f^*(k) & \epsilon_0(k) - v(k) \end{pmatrix}, \tag{336}$$

with $f^*(k) = f(-k)$, $\epsilon_0(k) = \epsilon_0(-k)$ and $v(k) = -v(-k)$. The band gap is $\Delta = \sqrt{v^2 + |f|^2}$, so that now f can vanish without the gap closing. This means the previous argument about integer quantization of w is no longer valid. However, so long as the gap does not close, the Zak phase in Eq. (330) is still non-zero and it is quantized to either 0 or π. If we try to vary from one class of states to the other, we will need to pass $f(\pi)$ through zero (it is real at $k = \pi$ since π and $-\pi$ are equivalent, so it can be varied through zero here by changing a single parameter). However, by the same reasoning, $v(\pi) = 0$, so a gap will close if we do this.

Now we can ask if there are robust bound states at the ends of inversion symmetry protected TIs. The answer is – I think – that this is not in general guaranteed, because inversion symmetry can be broken at the boundary. For an end of a 1d chain this is obvious, and inversion symmetry is *always* broken. For an interface between the two types of TIs, it is possible to preserve inversion symmetry around a site at the boundary (see for example Fig. 7(c)). In that case there may be a protected bound state, but I am not sure.

8.2 *AKLT/Haldane chain*

In the topological insulator example, the many body ground state is a Slater determinant, i.e. it is described by a set of filled single particle levels. Consequently, the topology of the ground state with a gap is equivalent to the topology of the filled bands. However, this is a crutch, and the notion of topologically distinct ground states extends also to systems which do not have a non-interacting description. So we can ask for examples in which we do not have a free particle description. The Haldane or AKLT chain is a famous example.

Historically, there was uncertainty about the ground state of the simple Heisenberg antiferromagnetic quantum spin chain,

$$H_{\text{Heisenberg}} = J \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1}, \tag{337}$$

where $\boldsymbol{S}_i$ is a spin-s operator, i.e. $\boldsymbol{S} \cdot \boldsymbol{S} = s(s+1)$, with possible values $s =$

$1/2, 1, 3/2, \cdots$. It is obvious that in the limit $s \to \infty$, the ground state is classical, and consists of spins which become rigid vectors of length s pointing alternately in the $+\hat{n}$ and $-\hat{n}$ directions, where the direction $\hat{n}$ is arbitrary. But for quantum spins this is not possible, because of the famous Mermin-Wagner theorem, which requires that there are no broken continuous symmetries even at $T = 0$ in one dimension, unless the order parameter itself commutes with the Hamiltonian. We already saw this in the context of the Bose superfluid, where we found there is no long range order in one dimension at $T = 0$ (the Bose gas has QLRO at $T = 0$ in this case). One piece of information came from Bethe's exact solution of the $s = 1/2$ case, which he worked out in 1935 – way ahead of its time! Theorists would spend many decades puzzling out what it really meant. Bethe found that the $s = 1/2$ chain was indeed not magnetically ordered, but was gapless with some mysterious excitations that seemed quite a lot like fermions. This is not consistent with a trivial (i.e. unentangled) disordered phase, which we would expect to have a gap. So one could think that this is like the one dimensional superfluid, with QLRO. It is true that there is QLRO, but a full description is even richer. We do not have time to discuss it here, but the Bethe chain is an example of an interesting conformal field theory.

In any case, since the $s = 1/2$ case was found to have QLRO, and we expect that larger s is more classical, it would be very natural to think that all s would have at least QLRO. Amazingly, this is not the case. It was Duncan Haldane who first suggested (based on a field theory argument which I will not try to reproduce here) that there might be a qualitative difference between integer and half-integer s, and that integer spins would have true disordered ground states with a spin gap. He turned out to be right. So the $s = 1$ Heisenberg antiferromagnetic chain is often called the Haldane chain. However, what he did not anticipate was that the ground state is not a trivial gapped one, but an SPT state (that is modern nomenclature). The topological nature of the $s = 1$ chain was clarified by Affleck, Kennedy, Tasaki, and Lieb, which introduced a slightly modified version of the Heisenberg Hamiltonian which is exactly soluble, at least for the ground state:

$$H_{AKLT} = J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + \frac{1}{3} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 \right]. \tag{338}$$

The AKLT Hamiltonian differs from the Heisenberg model by the addition "biquadratic" term with a specified coefficient. It is solvable because this particular combination is proportional (for $s = 1$ spins) to a spin *two* projection operator,

$$P_2(i, i+1) = \frac{1}{2} \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + \frac{1}{6} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 + \frac{1}{3}. \tag{339}$$

This operator annihilates total spin zero or one states, and gives unity on spin two states, which can be seen by writing $P_2 = \frac{1}{24} S^2_{i,i+1} (S^2_{i,i+1} - 2)$, with $\boldsymbol{S}_{i,i+1} = \boldsymbol{S}_i + \boldsymbol{S}_{i+1}$. If we introduce the basis of total spin states $|s_t m_t\rangle$, where s_t

is the total spin quantum number of $\boldsymbol{S}_i + \boldsymbol{S}_{i+1}$ and $m_t = -s_t, \ldots, s_t$, then

$$P_2(i,i+1) = \sum_{m_t=-2}^{2} |2,m_t\rangle\langle 2,m_t|. \tag{340}$$

We see that

$$H_{AKLT} = J\sum_i \left(2P_2(i,i+1) - \frac{2}{3}\right). \tag{341}$$

Like any projection operator, $P_2(i,i+1)$ has eigenvalues $= 1, 0$ only. So an eigenstate which saturates a lower bound on the energy – i.e. a ground state – is achieved by any state which is annihilated by *all* $P_2(i,i+1)$. Can we find such a state?

AKLT showed how to construct it. The trick is to represent each spin-1 spin as a pair of spin-1/2 spins. Of course, two spin-1/2 spins can be combined into a spin-0 state and a spin-1 state, so we can always represent a spin-1 this way. More formally, given a state of two s=1/2 spins, we can project it to a physical spin-1 state by applying the operator

$$P_{\text{phy}} = |1\rangle\langle\uparrow\uparrow| + |0\rangle\frac{1}{\sqrt{2}}\left(\langle\uparrow\downarrow| + \langle\downarrow\uparrow|\right) + |-1\rangle\langle\downarrow\downarrow|. \tag{342}$$

We can introduce two spin-1/2 auxiliary spins for each site i, and such an operator for each site i, and then define the full projection $P_{\text{phys}} = \prod_{i=1}^{N} P_{\text{phys}}(i)$. This operator takes a state in the Hilbert space of $2N$ $s = 1/2$ spins and maps it to a state of N $s = 1$ spins. This allows us to write *any* state of the $s = 1/2$ auxiliary spins and from it obtain an $s = 1$ state. AKLT had the inspired idea to use this trick for their model. They wrote the auxiliary state,

$$|\psi\rangle_{\text{aux}} = \prod_{i=1}^{N}\left(|\uparrow\rangle_{i,R}|\downarrow\rangle_{i+1,L} - |\downarrow\rangle_{i,R}|\uparrow\rangle_{i+1,L}\right), \tag{343}$$

where we introduced the labels L, R to denote the two $s = 1/2$ spins associated with each site, and took periodic boundary conditions $N + 1 = 1$. This state is, up to normalization, a product of spin-0 singlet states combining pairs of auxiliary spins between neighboring sites. Because of this, the auxiliary state itself is already manifestly a total spin zero state. This product of singlets can also be drawn graphically,

(344) $|\psi\rangle_{\text{aux}} =$ 

The L, R auxiliary spins of one site are marked.

By applying a projection operator to it, we obtain the physical AKLT state:

$$|\Psi\rangle_{AKLT} = P_{\text{phys}}|\psi\rangle_{\text{aux}}. \tag{345}$$

This is also manifestly a total spin zero state, since P_{phys} commutes with the total spin. We can also draw it graphically, by indicating how the auxiliary spins are grouped into physical ones,

(346) $|\Psi\rangle_{AKLT} =$ 

Now let us think about the action of the AKLT Hamiltonian on this state. It is a ground state of H_{AKLT} provide the total *physical* spin on each bond is less than two. Now we can write the total spin on bond $i, i+1$ in terms of the auxiliary spins:

(347)
$$\boldsymbol{S}_{i,i+1} = \boldsymbol{S}_i + \boldsymbol{S}_{i+1} = \boldsymbol{S}_{i,L} + \boldsymbol{S}_{i,R} + \boldsymbol{S}_{i+1,L} + \boldsymbol{S}_{i+1,R} = (\boldsymbol{S}_{i,R} + \boldsymbol{S}_{i+1,L}) + (\boldsymbol{S}_{i,L} + \boldsymbol{S}_{i+1,R})\,.$$

The last form shows that it is the sum of two different total spin operators, which are the combinations grouped in parenthesis. Each of these can be decomposed into spin-0 and spin-1 channels. In the AKLT state of Eq. (326), the first of these is, by design, spin-0. The second can be at most spin one, and so the sum can only have total spin zero or one, and never spin two. This proves that the state constructed this way indeed has physical spin less than two on every bond, i.e. it is a ground state of H_{AKLT}! It turns out that it is also unique – i.e. it is the only ground state – at least for a chain with periodic boundary conditions, which is what we have written.

If you look at the picture in Eq. (346), you immediate see a similarity to the diagrammatic representation of MPS's. Indeed the AKLT *is* an MPS, i.e. we can write it as

(348) $$|\psi\rangle_{AKLT} = \sum_{\sigma_1 \cdots \sigma_N = -1,0,1} \text{Tr}\,[A(\sigma_1)\cdots A(\sigma_N)]\,|\sigma_1 \cdots \sigma_N\rangle.$$

To see this explicitly, we map $|\downarrow\rangle_{i,L} \to |\uparrow\rangle_{i,L}$ and $|\uparrow\rangle_{i,L} \to -|\downarrow\rangle_{i,L}$, which maps the singlet

(349) $$|\uparrow\rangle_{i,R}|\downarrow\rangle_{i+1,L} - |\downarrow\rangle_{i,R}|\uparrow\rangle_{i+1,L} \to \sum_{\sigma=\uparrow,\downarrow} |\sigma\rangle_{i,R}|\sigma\rangle_{i+1,L},$$

which corresponds to the index contraction on bonds of an MPS. Then with the same transformation, the projector becomes

(350) $$P_{\text{phys}}(i) = |1\rangle(-\langle\downarrow\uparrow|) + |0\rangle\frac{1}{\sqrt{2}}(\langle\uparrow\uparrow| - \langle\downarrow\downarrow|) + |-1\rangle\langle\uparrow\downarrow|.$$

This operator represents the set of three matrices $A(\sigma)$ with $\sigma = -1, 0, 1$ the three $s = 1$ states of site i. We read of the matrix $A(\sigma)$ as the coefficient of the $|\sigma\rangle$ ket, and just conjugate the first auxiliary bra (corresponding to i, L) to a ket, i.e.

(351)

$$A(1) = -|\downarrow\rangle\langle\uparrow|, \qquad A(0) = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right), \qquad A(-1) = |\uparrow\rangle\langle\downarrow|.$$

Up to normalization and overall phase of the wavefunction, this agrees with "standard" MPS notations for the AKLT state, e.g. on wikipedia.

With periodic boundary conditions the AKLT state is unique. Using the MPS formulation it is easy to compute correlation functions in the state. One can see then directly that they are all short-range. The spin-spin correlation function, and all other connected correlation functions, decay exponentially. Thus we can view the AKLT state as a realization of Haldane's proposal that a spin-1 antiferromagnetic chain is disordered. It also has a spin gap, which physically means that to excite away from the ground state, some singlets of the AKLT state must be broken. However, while the excitations have been computed variationally and numerically, they are not known exactly for the AKLT state.

What was not appreciated even by AKLT so far as I know was the topological character of the state. This becomes apparent only when it is written for a finite open chain rather than a periodic one. Then we can still follow the auxiliary spin construction, and succeed in find ground states by forcing total spin < 2 on all the bonds by forming singlets between auxiliary spins. However, in this case, there are two auxiliary spins left out at the ends of the chains:

(352) $|\Psi\rangle^{\text{open}}_{AKLT} =$ 

The outermost (first and last) auxiliary spins (solid dots in the diagram) are *completely free*. So one can construct 4 linearly independent auxiliary states, and project each of them to obtain AKLT ground states. These states clearly have a net spin-1/2 at each end. This is something amazing! We start with a collection of $s = 1$ spins, which by the rules of angular momentum addition can only produce total spin with integer values, and we end up with effectively 2 $s = 1/2$ spins at the ends! These days it is fashionable to call this "symmetry fractionalization". Regardless of the name, it is very cool.

The presence of the $s = 1/2$ end spins is a robust feature of a *phase*. It persists as long as the spin-rotation symmetry of the Hamiltonian is preserved (in fact other symmetries can protect the boundary spins as well), and the bulk gap of the spectrum is maintained, i.e. until a quantum phase transition occurs. One can see this by continuity. It is best to think of a semi-infinite chain, which has then only one $s = 1/2$ spin, i.e. a two-fold ground state degeneracy. To remove the $s = 1/2$ spin degeneracy of the ground state, some other level would need to cross the ground state. If this happens in the bulk, it is a phase transition. Can a level crossing occur at the edge? For some models, additional bound states can appear at the edge, but these must be built out of the excitations in the bulk, which have integer spin. So any such state will have an integer spin *plus* the original half-integer boundary spin, and so has a total half-integer spin, and in particular a total spin zero is impossible. Thus any levels that are below the bulk gap must have a half-integer spin, and consequently a minimal two-fold degeneracy. Crossing of such a level with

the original one preserves the boundary spin. We see that the boundary spins are protected properties of the phase. A full discussion of the symmetries that can protect the topological phase has been given by several authors – for example it is apparent from the above discussion that time-reversal symmetry is sufficient, since it allows distinguishing half-integer and integer spin by Kramer's degeneracy and $\Theta^2 = -1$ versus $\Theta^2 = +1$.

So we expect a range of Hamiltonians "near" the AKLT one are in the same topological phase. Happily, this includes the simple Heisenberg Hamiltonian, Eq. (319), for the Haldane chain. On moving away from the AKLT point, the boundary spin spreads out a bit, and no longer resides exactly on the end site of the lattice, but it is still localized to the end. This means that, if we study the ground state with $S^z = +1/2$, the $\langle S_i^z \rangle$ will be non-zero near the end of the chain, summing to $1/2$, but decaying exponentially away from the end of the chain (in fact it also oscillates, reflecting the underlying antiferromagnetic interactions, but sums to a positive answer).

One can ask how to detect the emergent $s = 1/2$ spins. For a finite system, since there are two of these, if we add the full spin of the whole system it will be integer. This just means we can recombine the two $s = 1/2$ spins into a total spin $s = 0$ or $s = 1$. So if we do a numerical experiment and just segregate states by their total spin, we will find that there are four low energy states, a singlet and a triplet, separated by a gap from the excited ones. The fact that the singlet and triplet are degenerate is the indication of the fractionalization into two spin-1/2 spins. Also, note that separation into the total spin basis is a highly non-local operation. The minimally entangled basis is that of the individual $s = 1/2$ spins.

One can also think experimentally. The individual end spins are readily detectable! They give the Curie susceptibility of $s = 1/2$ spins, and a magnetic entropy of $k_B \ln 2$ per end. These are sharp characteristics that distinguish $s = 1/2$ spins. Many other experiments can be imagined. A very nice one is the NMR measurement of Tedoldi *et al*, who studied Haldane chains of $s = 1$ Ni spins doped by a low density of Mg substituting for Ni. Each site with a substitute Mg is missing its $s = 1$ spin, which cuts the chain into two. Consequently, two $s = 1/2$ impurity spins appear near the impurity, at the ends of the two chains so created. In an NMR experiment, a magnetic field is applied which polarizes the end spins. Then nuclei at different positions along the chain experience different fields due to the spread of the end spin, which is ~ 6 lattice sites for the Heisenberg model. This leads to a characteristic family of NMR resonant frequencies, which were measured by Tedoldi *et al*, in excellent agreement with theory.

Finally, let us note that there is a remarkable similarity to the existence of boundary states in the chiral hopping model we discussed in the previous subsection. The AKLT chain is part of a family of systems enjoying protected boundary excitations, and embodying some topological structure. This includes not only the one dimensional topological insulators, but also two and three dimensional time-reversal invariant topological insulators, quantum Hall states, and various topological superconductors. The latter are understood theoretically but a definitive experimental identification is very much

an active subject today.

9 Toric code

Lecture 19 (1.5h)
December 1st, 2015

In the previous section, we saw that states with the same symmetry may still be distinguished by topology. The non-trivial phases of this type can have interesting boundary degrees of freedom which are anomalous, like the $s = 1/2$ boundary spins of the AKLT chain. These SPT phases are, however, fragile, because they rely on a symmetry to distinguish them from the trivial phase. If we, for example, apply a magnetic field to the AKLT chain, it immediately becomes indistinguishable from a trivial state. The two states of the end spin are no longer degenerate.

There are, however, even more exotic things. It is possible for a system to enter a non-trivial phase which is stable to *all* perturbations. Even with no symmetry at all, it remains non-product within some finite domain of stability. The simplest example of this is the two-dimensional *toric code* model of Kitaev, which is actually largely a reinterpretation of the Ising lattice gauge theory introduced much earlier by Wegner. There is a very nice review article by Kitaev and Laumann which presents this, and a number of other topics, in quite readable form. Consider a set of spin-1/2 "spins" on the middle of the links of a square lattice, with the Hamiltonian

$$H_{\text{tc}} = -K \sum_p P_p - K' \sum_s S_s, \tag{353}$$

where the sums are over plaquettes p and sites s, the plaquette operator $P_p = \prod_{i \in p} \sigma_i^z$ is a product over the spins on the bonds surrounding the plaquette p, and the "star" operator $S_s = \prod_{i \in s} \sigma_i^x$ is a product over the spins on bonds neighboring the site s.

9.1 *Massive superposition in the ground state*

Obviously all the star operators commute as do the plaquette operators, and one can easily verify that the stars and plaquettes do as well, $[S_s, P_p] = 0$ for all p, s. This makes the toric code model especially easy to solve: ground states are simply those states for which $S_s = P_p = +1$ for every star and plaquette.

While this looks fairly trivial in these variables, the state itself in any *local* basis is highly entangled. Consider for instance the σ_i^x basis, which is complete and local. States with $S_s = 1$ are those in which an even number of spins composing each star have eigenvalues $\sigma_i^x = -1$. One may represent these states by coloring the links with negative spins, and according to $S_s = 1$ the links form closed loops. Since the operator P_p is off-diagonal in this basis, it *requires* superposition of the loop states. One can construct a state with $P_p = +1$ everywhere by taking a representative as a "base" state a direct product state (eigenstate of σ_i^x),

$$|\psi_0\rangle = \otimes_i |\sigma_i^x = s_i\rangle, \tag{354}$$

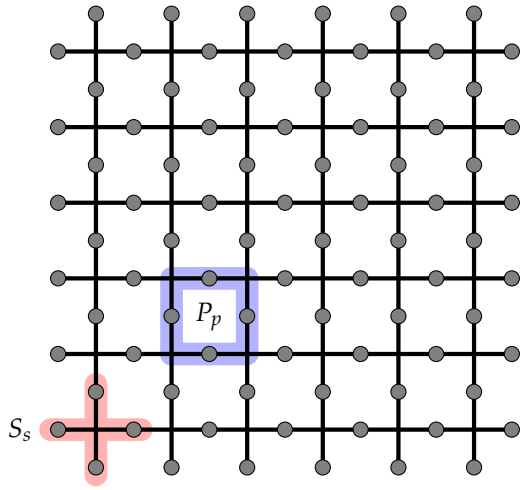


Figure 12: Toric code. Spins reside on the sites of the gray dots. The blue shaded region shows the spins involved in a plaquette operator, while the red shaded regions shows a star operator.

where the $s_i = \pm 1$ are chosen to satisfy the star rule $S_s = 1$. For example, we can take $s_i = 1$ everywhere. Then, we can act on it with the projectors

$$Q_p = \frac{1 + P_p}{2} = \frac{1}{2} \sum_{q_p = 0,1} P_p^{q_p}, \tag{355}$$

which projects onto states with $P_p = 1$. Using this,

$$|0\rangle = \prod_p Q_p |\psi_0\rangle. \tag{356}$$

By writing out the product explicitly,

$$|0\rangle = 2^{-N} \sum_{q_1 \cdots q_N = 0,1} \prod_p P_p^{q_p} |\psi_0\rangle, \tag{357}$$

we see that it is the sum of all possible products of plaquette operators acting on the base state. Each plaquette operator flips the s_i values on the sites contained in that plaquette. If we choose $s_i = +1$, then $|\psi_0\rangle$ is the "zero loop" state, and each product of plaquette operators creates loops on the plaquettes with non-zero q_p. Hence $|0\rangle$ appears to be a massive superposition of loop states, in which two loop configurations related to another by a flip of a minimal square plaquette appear with equal weight. This looks highly entangled indeed.

9.2 *Ground state degeneracy on the torus*

It might seem there are many such states, depending upon the choice of s_i. However, using $Q_p = Q_p P_p$, we can see that states in which the s_i are re-

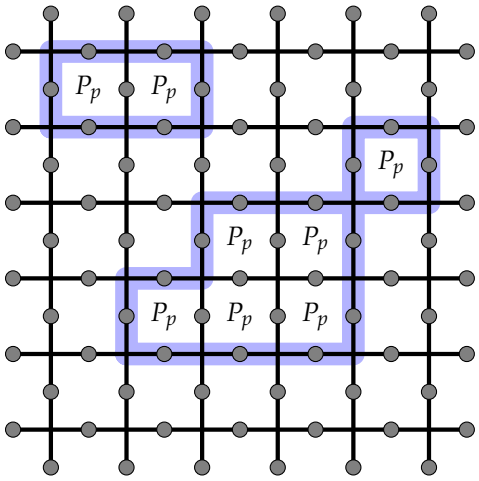


Figure 13: A loop configuration in the toric code. One loop configuration corresponds to a product state in the σ_i^x basis, with $\sigma_i^x = -1$ for the spins contained in the loops. The loop configuration shown is given by the product of P_p operators on the plaquettes indicated, acting upon $|\psi_0\rangle$ in Eq. (354) defined with $s_i = 1$ for all i.

lated by flipping spins around a minimal square plaquette are the same (after projection). That is, $|0\rangle$ is invariant under the transformation

$$|\psi_0\rangle \to P_p|\psi_0\rangle. \tag{358}$$

This means that any two choices of s_i related by Eq. (358) will give the same state $|0\rangle$. This leaves very few distinct states.

To see how many, let us do some counting. For concreteness, let us do this with periodic boundary conditions, i.e. on the torus, with L sites in each direction and $N = l^2$ total sites. We can choose *a priori* 2^{2N} states $|\psi_0\rangle$ by arbitrary choice of s_i on every link. However, we require the condition $S_s = 1$ on every site. One such condition cuts the number of states by 2. Each condition is independent *except* the last one. This is because the product over all sites,

$$\prod_s S_s = 1, \tag{359}$$

since every spin is included in two stars, and $(\sigma_i^x)^2 = 1$. So there are $N - 1$ constraints, which reduces the number of base states by $1/2^{N-1}$. This leaves 2^{N+1} base states. This is just the number of loop coverings, i.e. graphs like in Fig. 13, where every site is visiting by an even number of colored bonds. Now any two such states which differ by the transformation in Eq. (358) yield the same state $|0\rangle$. Graphically, this transformation changes the color of the links on the plaquette p, i.e. turns uncolored to colored and vice versa. So most loop coverings yield the same state $|0\rangle$. Let us try to count further. We can start with the trivial $|\psi_0\rangle$, which is the "no loop" state. Then acting on it with P_p on each plaquette gives another loop state. In principle by choosing

whether or not to apply each P_p we can double the number of states. However, if we apply *all* the P_p, then every bond gets its color switched twice, and one recovers the same loop state. Actually there is an operator identity:

$$\prod_p P_p = 1. \tag{360}$$

Thus by starting with the trivial base state, we can generate 2^{N-1} loop states by acting with all but one of the P_p operators. And all of these loop states obtained in this way lead to the same $|0\rangle$. So we need to divide the 2^{N+1} possible base states by the number, 2^{N-1}, which lead to the same physical state after projection. This means that the true number of distinct ground states is $2^{N+1}/2^{N-1} = 4$. Note that in this argument we used periodic boundary conditions several times, in assuming all sites and plaquettes were equivalent, for example, and in the identities of Eqs. (359,360).

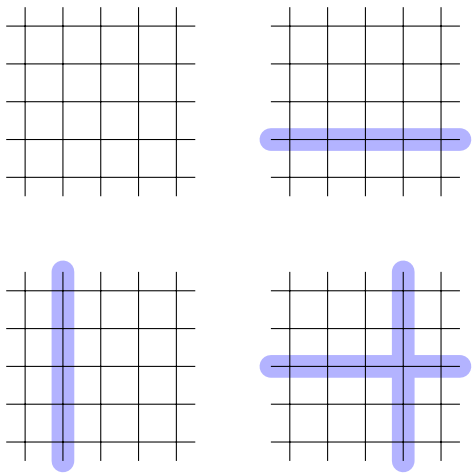


Figure 14: Four base states which differ by non-trivial loops (representing spins $\sigma_i^x = -1$ along the loop). Each of these base states generates, via Eq. (356), a distinct ground state on the torus.

Let us think explicitly about the 4 ground states on the torus. Each of them can be obtained by starting from a single base state, such that the four base states *cannot* be obtained from one another by the action of P_p operators. The four types of base states (see Fig. 14) can be explicitly constructed using loop configurations in which there are *large* loops that wind all the way around either the x or y axis. The *parity* of the number of such loops is conserved by the action of P_p. To see this formally, we can construct an operator to measure this parity. Consider a non-trivial closed loop $\tilde{\ell}$ defined on the dual lattice (see Fig. 15):

$$X_{\tilde{\ell}} = \prod_{i \in \tilde{\ell}} \sigma_i^x. \tag{361}$$

The operator $X_{\tilde{\ell}}$ obviously commutes with the star operators but slightly less

trivially commutes with P_p for any p, since the plaquette operator changes the sign of an even number of spins on the loop. It is equal to the parity of the number of loops crossing the dual loop $\tilde{\ell}$. For convenience we can specify two particular loops $\tilde{\ell}_1$ and $\tilde{\ell}_2$ as shown in Fig. 15, and denote the corresponding loop operators by X_1, X_2. The states $|0\rangle_{1,2,3,4}$ constructed from the base states pictured in Fig. 14 are eigenstates of the loop parity operators with $(X_1, X_2) = (1,1), (1,-1), (-1,1), (-1,-1)$. In this way the loop parities parametrize the degenerate ground state subspace.

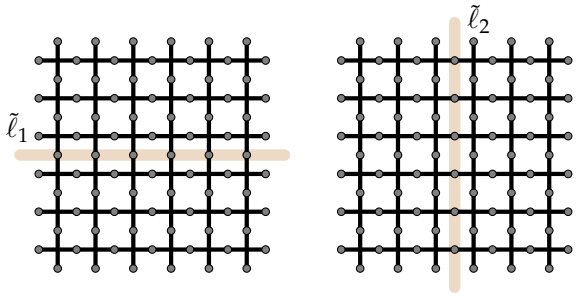


Figure 15: Non-trivial loops $\tilde{\ell}_1$ and $\tilde{\ell}_2$ on the dual lattice.

These four ground states constitute the famous *topological degeneracy* of the toric code on the torus. The interesting thing about this degeneracy is that the different degenerate states cannot be distinguished locally. More precisely, if we take any local operator $\mathcal{O}$ (i.e. a product over σ_i^μ operators over some set of nearby sites which does not span the system), then

(362) $\langle m|\mathcal{O}|n\rangle = \overline{\mathcal{O}}\delta_{mn},$

if $|m\rangle$ with $m = 1,2,3,4$ are the orthonormal degenerate ground states. This can be shown formally, but the physics is simply that the ground states differ only by the global parity of the loops, and this does not affect the local configuration space. Essentially because of the above property, one can show that the the topological degeneracy is very robust: it is maintained even when the toric code Hamiltonian is perturbed by arbitrary perturbations, provided these are below some threshold, which defines the stability region for the toric code phase.

Lecture 20 (1.5h)
December 3rd, 2015

9.3 *Anyons*

Actually we can find not only the ground state but all the levels in the toric code. Let us again assume a torus geometry. We can refine the thinking of the previous subsection into a more formal change of variables. The Hilbert space is that of $2N$ spins, as there are twice as many links as sites, and so is 2^{2N}-dimensional. Instead of a direct product basis in real space, we can trade this for a different basis of commuting observables. The natural ones are the star and plaquette operators, which, like for example σ_i^z, take two values. There are N star and N plaquette operators. Due however to the constraints,

Eq. (359) and Eq. (360), this describes just $2N-2$ independent spins, i.e. a 2^{2N-2}-dimensional Hilbert space. This is almost but not quite enough to fully specify a new basis. The additional operators we need are the loop operators X_1 and X_2 defined in Eq. (361) for the two non-trivial cycles of the torus. The entire new basis is specified by states $|X_1, X_2; \{S_s, P_p\}\rangle$, with each argument of the state taking values ± 1, subject to the constraints of Eqs. (359,360). In this basis, the Hamiltonian is diagonal! So we have found all the energy levels.

e and m particles

The X_1, X_2 variables do not enter the energy. Hence for this model every energy level has this fourfold degeneracy. These are global degeneracies. *Excitations* from the ground state correspond to terms in the Hamiltonian which are not minimized. A priori, the lowest energy excitations consist of changing one star or one plaquette from $+1$ to -1. This increases the energy by $2K'$ or $2K$, respectively. We can consider such excitations the quasiparticles of the toric code. There is a concern, however, which is that changing a single star or plaquette is not consistent with the constraints of Eqs. (359,360). Physically, this does not mean that these are the wrong quasiparticles, but rather that they are non-local objects. Like a domain wall in the 1d Ising ferromagnet, each such quasiparticle always occurs with a counterpart in a closed periodic system. The latter requirement is, however, a global constraint, and for a large system the two quasiparticles may be far away from one another so that neither affects its "partner" significantly. Also like the 1d Ising soliton, a single quasiparticle in the toric code cannot be created by any local operator. Indeed, for example acting with σ_i^z on a ground state, which has $S_s = +1$ everywhere, leads to a state which has $S_s = -1$ on the two stars sharing the spin i. Similarly, acting with σ_i^x creates a state with $P_p = -1$ on two plaquettes.

For convenience, we will call the quasiparticles consisting of stars with $S_s = -1$ "electric" or e particles, and those consisting of plaquettes with $P_p = -1$ "magnetic" or m particles. A pair of electric particles is created by acting with a product, or "string", or σ_i^z operators connecting the two sites,

$$|\mathsf{e}_s; \mathsf{e}_{s'}\rangle = \prod_{i \in \ell(s,s')} \sigma_i^z |0\rangle, \tag{363}$$

where $\ell(s, s')$ is a chain of bonds on the lattice connecting sites s and s' (see Fig. 16). Similarly, a pair of m anyons is created by a string running on the dual lattice.

We can now explain the names of the e and m anyons. The *definition* of the e anyon is that it is a defect where $S_s = -1$. Away from the anyon, the lines of negative σ_i^x form closed loops. The smallest such loop is created by the plaquette operator P_p. We can think of these loops as field lines, and they are "divergenceless". Everywhere, that is, away from the e anyon. A the anyon's location, there is a divergence: that is, the e particle is a source of "flux" of these field lines. The string of σ_i^x operators in Eq. (363) creates this flux. We refer to the lines of $\sigma_i^x = -1$ as "electric" field lines, and the e particle carries "electric" charge. The difference from the usual charge in electromagnetism is

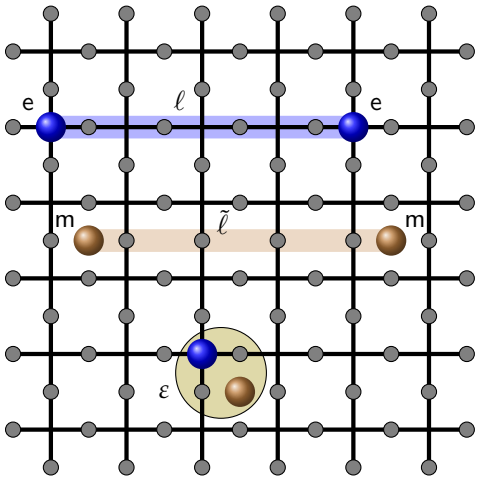


Figure 16: Anyons. Two e particles, shown as blue balls, are created by acting with σ_i^z operators on the string ℓ. Similarly, two m particles, are created at the ends of a string of σ_i^x operators on the dual path $\tilde{\ell}$. A composite ε particle consists of a neighboring e and m particle.

that this charge is $\mathbb{Z}_2$-valued.

Similarly, the m anyon can be considered to carry a "magnetic" charge, more conventionally called a flux. So sometimes the m particle is called a $\mathbb{Z}_2$ vortex. We can define dual field lines which lie on the dual lattice, and consider of lines of sites on which $\sigma_i^z = -1$. Again these form loops, away from any m anyon. The star operator S_p creates the smallest such loop. The m particle is a source for these dual field lines.

The very existence of the e and m excitations as quasi-particles can be thought of in terms of these field lines. The defining feature of the particles is that they are sources of flux. For them to be finite energy excitations, it must be possible for this flux to "spread" somehow, so that its energy cost remains finite (this happens in ordinary electromagnetism, where the electric field of a point charge spreads out like $1/r^2$, and so the energy density is $\sim 1/r^4$, whose integral converges at long distances r from the charge in three dimensions). In the toric code, the analogous spreading occurs due to massive superposition of different configurations of the field lines in the wavefunction. It is remarkable that this $\mathbb{Z}_2$ flux can extend infinitely far away from the quasiparticles, yet they remain of finite energy. The existence of the flux lines implies that there is some way to "sense" the presence of a quasiparticle far away, without directly measuring in the immediate vicinity of the quasiparticle. We can see this directly below.

Statistics of e and m

Let's focus on the two e anyons. Now we can imagine moving the two anyons "smoothly", i.e. one site at a time, so that they are interchanged. This process consists of changing the endpoints of the string $\ell(s, s')$ to move s or s', which

just multiplies the wavefunction in Eq. (363) by a factor of σ_i^z, which either extends the product in the string or cancels an existing factor in the product, since $(\sigma_i^z)^2 = 1$. The net process of interchanging the two e anyons results in multiplying the initial wavefunction by a factor,

$$|\psi_{\text{final}}\rangle = \prod_{i \in \mathcal{L}} \sigma_i^z |\psi_{\text{initial}}\rangle, \tag{364}$$

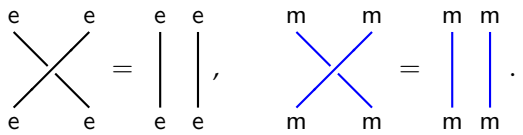
where $\mathcal{L}$ is the closed loop describing the path of the two e particles during the interchange. Now we can use the discrete equivalent of Stoke's theorem, to rewrite this loop product as a product over plaquettes,

$$\prod_{i \in \mathcal{L}} \sigma_i^z = \prod_{p \in \mathcal{A}} P_p, \tag{365}$$

where $\mathcal{A}$ is the area spanned by the loop $\mathcal{L}$. Acting on the state with two e particles, which has no m particles, $P_p = 1$ and so we obtain $|\psi_{\text{final}}\rangle = |\psi_{\text{initial}}\rangle$. The act of interchanging two e particles returns one to the same state, with no phase change. We conclude that the e particles are bosons. The same argument applies to the m particles.

We can represent the bosonic exchange by diagrams:

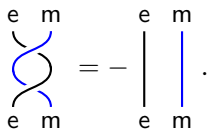
(366)



You should interpret these diagrams as a view of the "world-lines" of the particles, viewing the 2d plane side-on, so the x axis is the physical x axis, the physical y axis is into the paper, and the y axis is time. You need to be careful about when a line goes behind or in front of another. This represents winding.

However, we can consider an interesting variation: what happens if we move an e particle in a closed loop, containing an m particle? This again must return to the same state, up to a phase. Since we only move the e particle, Eq. (364) continues to hold, as does Eq. (365), which is just a mathematical identity. But the product of plaquette operators is -1 if the area $\mathcal{A}$ contains an m particle. We find the remarkable result that upon moving an e particle around an m particle, the wavefunction changing sign (sometimes we say it accumulates a π phase). This type of phase is called *mutual statistics*. The effect is independent of the distance between the e and m particles, and so it is a sign of their non-locality. Very crudely, if you think of how the two particles must have been created, they each leave behind a "string", and the e particle must cross the string of the m particle. In any case, we say that the e and m particles are "mutual anyons". The effect is similar to an Aharonov-Bohn phase in quantum mechanics. It is as if the m particle appears as a magnetic flux of half a flux quantum to the e particle, and vice-versa.

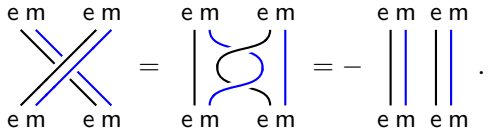
Here, the diagram is different

(367) 

The ε particle (dyon) is a fermion

A consequence of this statistical interaction between e and m particles is that a composite – a pair of nearby e and m particles, which is sometimes called a "dyon" or an ε particle – behaves as a *fermion*! Basically if we make an interchange of two ε's, we effectively wind a constituent e particle of one ε half-way around the m of the other ε particle, and vice-versa, and the net effect is to change the overall sign of the wavefunction.

There is a formal demonstration using these diagrams. It goes like this:

(368) 

Seems like magic. Personally I like to see what this means explicitly. Consider two ϵ particles, described by some initial state $|\psi_{\text{init}}\rangle$, such that, acting on $|\psi_{\text{init}}\rangle$, the star operators S_s give $+1$ except for two sites s_1 and s_2, where $S_s = -1$, and likewise the plaquette operators P_p give $+1$ except for two plaquettes p_1 and p_2, where $P_p = -1$. We choose s_1 and p_1 to be adjacent, comprising one ε "particle", and s_2 and p_2 to be adjacent, making up the other ε particle, but the s_1, p_1 will be far from s_2, p_2. For concreteness we take the configuration shown in Fig. 17. Note that some delicacy is required here because in the toric code model, there is no real interaction (only the statistical one) at all between the e and m particles – so there is no actual bound state of the two. This means that different configurations of dyons are actually degenerate (for example, we can place the m particle on any of the neighboring plaquettes of the e site s). Due to degeneracy of levels, it is not so easy to define the *adiabatic* phase. To do it, we must follow the path of a dyon which we hold together "by hand" in the process of evolution.

Now we wish to exchange the two ε particles. We will do this by a sequence of unitary transformations, so that

$$|\psi_{\text{final}}\rangle = U|\psi_{\text{init}}\rangle, \tag{369}$$

where the total unitary transformation is made in n steps:

$$U = U_{t_n} \cdots U_{t_2} U_{t_1}, \tag{370}$$

and U_t gives a "small" transformation which is local and moves anyons by a short distance. At each step, we can act with a single σ_i^z and σ_j^x operator to

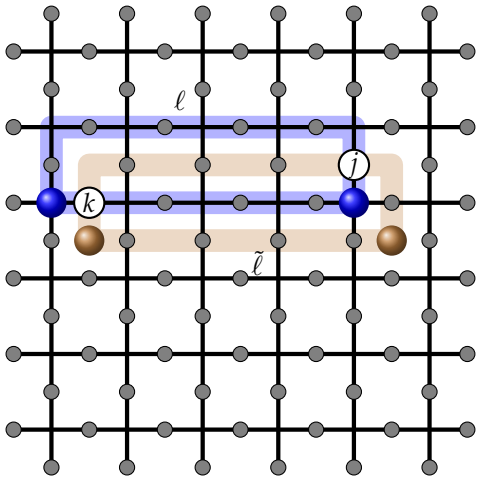


Figure 17: Exchange of two ε "particles". The anyons are moved along the indicated paths in the clockwise direction. The two paths intersect at the sites j and k, marked by white circles. Hence sites j and site k occur in both paths ℓ and $\tilde{\ell}$, but they occur *first* in ℓ and then in $\tilde{\ell}$, i.e. $t_j(\ell) < t_j(\tilde{\ell})$ and likewise $t_k(\ell) < t_k(\tilde{\ell})$.

move an e and m excitation, so

$$U_t = \sigma^x_{i_{\tilde{\ell}}(t)} \sigma^z_{i_\ell(t)}, \tag{371}$$

where ℓ and $\tilde{\ell}$ are the paths which the e and m particles sweep out, respectively, and $i_{\tilde{\ell}}(t)$ and $i_\ell(t)$ are specific spins (links) along these paths, at the "time" t of the evolution. We choose the steps such that the e and m constituents of each ε composite stay together during the evolution, and preserve their orientation (see also the discussion below). We choose to make the exchange by a clockwise rotation – see Fig. 17. Then we see that in general the paths ℓ and $\tilde{\ell}$ intersect twice, at sites i and j. All other sites involved in the paths are distinct, and so the operators involved all commute. We need to be careful about the spin operators at sites j and k. By inspection of Fig. 17, we observe that both of these sites are visited first by path ℓ and later by path $\tilde{\ell}$. This implies that for both j and k, the σ^x operator occurs to the left (i.e. "later") of the σ^z operator. Since all the other spin operators commute, we can therefore move *all* the σ^x_i operators to the left of the σ^z_i operators, for every site i. This gives

$$U = \prod_{i \in \tilde{\ell}} \sigma^x_i \times \prod_{i \in \ell} \sigma^z_i. \tag{372}$$

Now we can rewrite this as

$$U = \prod_{s \in \mathcal{A}(\tilde{\ell})} S_s \times \prod_{p \in \mathcal{A}(\ell)} P_p. \tag{373}$$

Here $\mathcal{A}(\tilde{\ell})$ and $\mathcal{A}(\ell)$ are the areas enclosed by $\tilde{\ell}$ and ℓ, respectively. This makes it easy to find the action of U on $|\psi_{\text{init}}\rangle$, Eq. (369), since these operators just measure the presence of anyons in the initial state. We see that the path ℓ contains no anyons, while the path $\tilde{\ell}$ encircles a single e anyon. So we obtain $U|\psi_{\text{init}}\rangle = -|\psi_{\text{init}}\rangle$, hence

$$|\psi_{\text{final}}\rangle = -|\psi_{\text{init}}\rangle. \tag{374}$$

We have successfully shown that the ε particles are fermions! This is pretty remarkable as our original model contains only bosonic spins, i.e. all the microscopic operators commute at different sites.

Relationship of ground state degeneracy and anyons

We found a fourfold ground state degeneracy of the toric code earlier by direct construction of these states. It turns out that this degeneracy follows from the existence of the e and m anyons. Because of this, it is most natural to regard the anyons as the fundamental characteristic of the phase of matter represented by the toric code. Once we know that they exist, we can understand the ground state degeneracy of the model with *any* boundary conditions, i.e. on different genus surfaces.

The basic idea is to relate the anyonic statistics to an algebra of loop operators. These operators describe the adiabatic motions of anyons. So imagine a time evolution as follows: Start from a ground state with no excitations. Create a pair of e anyons nearby, and then slowly move them apart, taking them in a non-trivial path around the torus, and then annihilating them (alternatively, you can keep one fixed and move the second one around the torus, and then return to annihilate it on the other side). In the toric code, we would do this explicitly by first acting with σ_i^z somewhere, which creates the anyon pair, and then acting with additions σ_i^z operators along a loop to move one of the e particles. In the end we arrive at an product of σ_i^z operators along some loop ℓ which spans one of the periodic directions of the torus. Let us define these operators as

$$Z_a = \prod_{i \in \ell_a} \sigma_i^z, \tag{375}$$

where ℓ_1 and ℓ_2 are loops traversing the periodic x and y directions, on the direct lattice (we use the capital Z symbol because these involve σ_i^z matrices).

Let's think a bit more physically about this Z_a operator. What we are doing is pulling an e anyon around the torus. As it does so, it pulls along with it an electric field line. Once the e anyon makes the full circuit around the torus and annihilates with the other e anyon, it leaves behind a pure loop of electric field. This is exactly one of the non-trivial loops shown in Fig. 14. So the act of creating a pair of e anyons, winding one around the cylinder, and annihilating again, takes the system from one ground state to another. The Z_a operator creates the electric field loop.

We can also of course consider the adiabatic evolution of the m anyons, similarly. Here we would act in the toric code with a series of σ_i^x operators

along some dual paths,

$$X_a = \prod_{i \in \tilde{\ell}_a} \sigma_i^x, \tag{376}$$

where $\tilde{\ell}_a$ give loops on the dual lattice winding in the x and y directions.

The loops winding in orthogonal directions cross, and so X and Z operators do not commute. We can find directly that

$$Z_x X_y = -X_y Z_x, \qquad Z_y X_x = -X_x Z_y. \tag{377}$$

These two equations actually imply the ground state degeneracy of 4. This is because, by construction, these operators start from a ground state and end in a ground state: any quasiparticles created are annihilated at the end of each operation. In the toric code model, this is obvious because the operators all commute with H. Once we have Eq. (377), then we know that the ground state(s) must form a representation of this algebra. And it is easy to show that the minimum dimension of the algebra is 4-dimensional (I leave it as an exercise to the reader!).

Note that, while we can obtain this algebra directly from the explicit operators in the toric code, it is expected to hold *anywhere* in the toric code phase, i.e. even once the exactly soluble model is perturbed. The idea is that under such a perturbation, the e and m anyons still exist, are gapped, and their mutual statistics is robust to perturbations. Then we can consider the adiabatic evolution, and define operators X_a an Z_a to carry this out. In a general model, these operators will not be given by Eqs. (375,376). However, the commutation relations, Eq. (377), actually directly follows from the mutual statistics, i.e. Eqs. (366,367). This requires some geometrical thinking, and unfortunately needs 3d pictures, which I find a bit hard to draw in latex. We anyway, here is how I understand it. We can rewrite for example the first condition in Eq. (377) as

$$Z_x^{-1} X_y^{-1} Z_x X_y = -1. \tag{378}$$

So the left hand side can be regarded as a sequence of operations to be performed, after which the wavefunction returns to itself, with an overall minus sign. It turns out that topologically, this sequence is equivalent to taking an e anyon around an m anyon, i.e. the non-trivial mutual statistics of Eq. (367). To see this, we need to visualize the operations of Eq. (378) in 2+1 space-time dimensions. We have four "world-lines" of particles which are taken around the torus. The "endpoints" (intersection with the boundary of the periodic cell) of the two e lines can be brought together and pinched off to form a single loop. The same can be done with the two m lines. Having done so, one finds that the two loops are linked! This linked configuration is, in turn, deformable to winding one anyone around the other. So, by the mutual statistics of e and m, it is equal to -1. Thus the non-commutativity of the loop operators follows from the mutual statistics, and in turn so does the topological degeneracy on the torus.

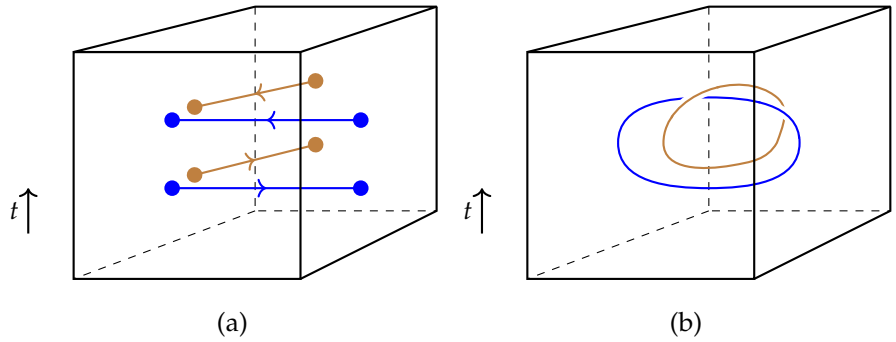


Figure 18: Equivalence of Eq. (378) to braiding an e and m anyon. In (a), one observes the sequence of e and m particles threading opposite cycles of the torus in alternating fashion. By contracting the endpoints on the periodic boundaries, this deforms to the diagram in (b), in which the world-lines of e and m particles are clearly linked.

9.4 *Stability of the toric code*

So far we mostly analyzed the exactly soluble toric code model. This would be a purely formal exercise if it were not for the fact that it represents a stable phase of ($T = 0$) matter. The universal properties of this phase – the ground state degeneracy, the existence and statistics of anyons, etc. – persist under *arbitrary* small perturbations of the model. This might be surprising as the toric code is certainly very finely tuned: it has an infinite number of local operators which commute with H, so in a sense it has an infinite number of local symmetries. Global symmetries lead to effects which are usually fragile to perturbations. For example, the gapless Goldstone mode of a system with U(1) symmetry becomes gapped upon breaking the U(1) symmetry. The SPT phases of the previous section become smoothly connected to trivial phases is the symmetry protecting them is broken. Here we have started with much more symmetry, which will always be broken explicitly in any realistic model. Yet amazingly, the key properties of the toric code are robust. Actually they are *more* robust than for ordinary phases: no matter what symmetries are broken, the phase is stable to arbitrary local perturbations. This is not true for any of the previous examples we considered.

Why is the toric code phase stable? Let us go back to the essential features of the phase, which is really that it supports gapped anyons. This is possible only because lines of flux are conserved (i.e. they form loops) in the ground state, and because they fluctuate wildly, i.e. the ground state is a superposition of loops up to arbitrarily large size. Let us think how these factors are affected by a perturbation. Consider the modification of the Hamiltonian by additional terms

$$H' = \sum_i h_i \sigma_i^x + \tilde{h}_i \sigma_i^z. \tag{379}$$

These terms break all the global and local symmetries of the toric code (except

time-reversal, which we can also break, but do not bother to do so for simplicity). Consider the action of H' on a ground state $|0\rangle$. Acting once with H' creates a sum of terms, each of which consists of a *pair* of anyons on nearby sites. Equivalently, we can think of these as short open "strings". So H' does induce violations of the flux conservation, but these violations are localized. In terms of the $\mathbb{Z}_2$ charges, the violations are neutral, viewed from outside. This was of course guaranteed by the non-locality of the anyons. Further action of H', at higher orders in perturbation theory, can induce more widely separated anyons, or ends of strings, but they are penalized by the denominator in perturbation theory, so one expects that the probability to find a pair of anyons separated by a distance ℓ will decay exponentially with ℓ.

This is a typical behavior for *virtual* particle/anti-particle excitations in a theory with a gap. Indeed just the same occurs for the spin flip excitations in the paramagnetic phase of the Ising model. Because the virtual excitations only occur locally, they do not cause long-distance effects, and most importantly, the charges of any real (not virtual) anyons remain well-defined.

Other properties can be analyzed similarly. We can consider for example the topological degeneracy on the torus. While it is exact for the soluble toric code mode, in general it is not for a finite lattice with nonzero H'. Clearly neither X_a nor Z_a commute with H'. So for a finite lattice the topological degeneracy will be split. However, it is restored in the thermodynamic limit. In particular, the four lowest states approach one another with an energy spacing which decreases exponentially with the linear size of the system, $\sim e^{-L/\xi}$, with some constant ξ. One can see for example that, perturbatively in H', the splitting scales as the L^{th} power of H', as that many actions are required to connect states which differ by a global winding. More physically, the action Z_a corresponds to creating a virtual pair of e anyons and winding them around the torus before annihilating them. This is a quantum tunneling process, passing through a high energy state (of twice the gap for the e anyon) for a tunneling time which is proportional to the length around the torus. So the amplitude of the process is proportional to the exponential of its action, which is the product of these two and hence exponential in the length.

Another way to think about the stability of the toric code phase is to ask, how is it eventually destabilized when the perturbation is not small? In particular, we should understand how the anyons are removed from the spectrum, as they must not be present in a topologically trivial, unentangled state. Obviously for large enough h_i and $\tilde{h}_i$ in Eq. (379), the system reaches such a state. A good picture for how this happens is that the gap for one of the anyons, either e or m closes, and since these are bosons, they "condense" (the one whose gap vanishes depends upon whether h or $\tilde{h}$ is larger). In this condensate, *all* the anyons cease to exist as well-defined quasiparticles. For example, suppose the m particle condenses. Then an individual m particle is no longer well-defined, because the number of m particles in the ground state is uncertain. In terms of flux, the magnetic field lines now can end in arbitrary places in the sample, already in the ground state. We say that the m particle is "Higgsed", like the Higgs boson in the standard model. What of the e particle? Well because the ground state is full of a superposition of different m states, the mutual statis-

tics of e and m particles means that the e particle can no longer propagate coherently. Another viewpoint is based on the uncertainty principle: the m condensation means that the magnetic operators σ_i^z, i.e. the lines of magnetic force on the dual lattice with $\sigma_i^z = -1$, become highly fluctuating. Because they anticommute with σ_i^x, the latter become very sharply defined and do not fluctuate. It means that the electric field lines become concentrated and cannot spread out. Consequently an e particle emits a concentrated field line which because it cannot spread, costs an energy proportional to its length. Thus an isolated e particle has an infinite energy. We say that the e particle is "confined". The ε particle is also confined because it also carries the electric charge.

We see that the combination of the Higgs phenomena and confinement removes all the anyons as quasiparticles. So the transition to a trivial phase occurs by driving the gap for one of the bosonic anyons to zero. This cannot happen with a weak perturbation of the toric code, because the gap is non-zero and large initially, and it is a smooth function of the perturbation for small perturbations. Thus the stability of the toric code phase is assured by the gap of the anyons.

If you don't like any of these physical arguments, there is a rigorous proof of stability.

9.5 *Recap*

The toric code provides a very simple and completely soluble model for a topological phase. We see that there is a robust phase of matter which is entirely characterized by means other than symmetry – topology and entanglement. Actually there are many such phases. In the lab, the fractional quantum Hall effect is a well-studied example. That is rather an old and mostly solved problem, but the perspective on it has changed somewhat as people have come to realize the role of entanglement and topology in it, and the generality of the phenomena (in theory anyway). The existence of novel non-local excitations is the key feature of this type of state.

Many other examples are of interest. There is a fairly complete understanding of topological phases in two dimensions, at an universal level anyway. But three dimensional topological phases are still a subject of much research. There are intriguing relations between phases with intrinsic topological order and SPT phases. Highly entangled phases which are not topological also exist, and are arguably more interesting. These can be gapless and there are many more unsolved mysteries there. The pursuit of highly entangled phases in experiment (for example in spin systems, or artificial structures) or realistic theories is another major current challenge. Quantum computing might be the "killer app" for these phenomena. There are many connections of these ideas to quantum field theory – topological field theories, gauge theory, etc. But we are out of time. I hope you found what we were able to cover interesting.