PHY392T NOTES: TOPOLOGICAL PHASES OF MATTER

ARUN DEBRAY
NOVEMBER 12, 2019

These notes were taken in UT Austin's PHY392T (Topological phases of matter) class in Fall 2019, taught by Andrew Potter. I live-TEXed them using `vim`, so there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu. Any mistakes in the notes are my own.

Contents

Lecture 1.

: 8/29/19

Lecture 2.

Second quantization: 9/3/19

Today we'll describe second quantization as a convenient way to describe many-particle quantum-mechanical systems.

In "first quantization" (only named because it came first) one considers a system of N identical particles, either bosons or fermions. The wavefunction $\psi(r_1, \dots, r_N)$ is redundant: if σ is a permutation of $\{1, \dots, N\}$, then

$$\psi(r_1, \dots, r_n) = (\pm 1)\psi(r_{\sigma(1)}, \dots, r_{\sigma(N)}, \tag{2.1}$$

where the sign depends on whether we have bosons or fermions, and on the parity of σ.

For fermionic systems specifically, $\psi(r_1, \dots, r_N)$ is the determinant of an $N \times N$ matrix, which leads to an exponential amount of information in N. It would be nice to have a more efficient way of understanding many-particle systems which takes advantage of the redundancy (2.1) somehow; this is what second quantization does.

Another advantage of second quantization is that it allows for systems in which the total particle number can change, as in some relativistic systems.

The idea of second quantization is to view every degree of freedom as a quantum harmonic oscillator

$$H \coloneqq \frac{1}{2}\omega^2(p^2 + x^2). \tag{2.2}$$

We set the lowest eigenvalue to zero for convenience. If $a \coloneqq (x+ip)/\sqrt{2}$ and $a^\dagger \coloneqq (x-ip)/\sqrt{2}$, then $\hat{n} \coloneqq a^\dagger a$ computes the eigenvalue of an eigenstate.

Now let's assume our particles are all identical bosons. Then we introduce these operators $a_\sigma(\mathbf{r}), a_\sigma^\dagger(\mathbf{r})$ which behave as annihilation, respectively creation operators, in that they satisfy the commutation relations

$$\begin{aligned} [a_\sigma^\dagger(\mathbf{r}), a_{\sigma'}(\mathbf{r}')] &= -\delta_{\sigma\sigma'}\delta(\mathbf{r}-\mathbf{r}') \\ [a^\dagger, a^\dagger] &= 0. \end{aligned} \tag{2.3}$$

The Hamiltonian is generally of the form

$$H \coloneqq \sum_{\sigma,\sigma'} \int_{\mathbf{r},\mathbf{r}'} a_\sigma^\dagger(\mathbf{r}) h_{\sigma\sigma'}(\mathbf{r}-\mathbf{r}') a_{\sigma'}(\mathbf{r}) + V_{\alpha\beta\gamma\delta} a_\alpha^\dagger a_\beta^\dagger a_\gamma a_\delta, \tag{2.4}$$

where the first term is the free part and the second term determines a two-particle interaction.

Letting $n_\sigma(\mathbf{r}) \coloneqq a_\sigma^\dagger(\mathbf{r}) a_\sigma(\mathbf{r})$, which is called the *number operator* (since it counts the number of particles in state σ), there is a state $|\varnothing\rangle$ called the *vacuum* which satisfies $n_\sigma(\mathbf{r})|\varnothing\rangle = 0$ and $a_\sigma(\mathbf{r})|\varnothing\rangle = 0$. Particle creation operators commute, in that

$$a^\dagger(\mathbf{r}_1)a^\dagger(\mathbf{r}_2)|\varnothing\rangle = a^\dagger(\mathbf{r}_2)a^\dagger(\mathbf{r}_1)|\varnothing\rangle. \tag{2.5}$$

This is encoding that the particles are bosons: we exchange them and nothing changes.

The fermionic story is similar, but things should anticommute rather than commute. Letting α be an index, let f_α, resp. $f_\alpha^\dagger$ be the annihilation, resp. creation operators for a fermion in state α. There's again a vacuum $|\varnothing\rangle$, with $f_\alpha|\varnothing\rangle = 0$ for all α. Now we impose the relation

$$f_\alpha^\dagger f_\beta^\dagger|\varnothing\rangle = -f_\beta^\dagger f_\alpha^\dagger|\varnothing\rangle. \tag{2.6}$$

That is, define the *anticommutator* by

$$\{f_\alpha^\dagger, f_\beta^\dagger\} \coloneqq f_\alpha^\dagger f_\beta^\dagger + f_\beta^\dagger f_\alpha^\dagger. \tag{2.7}$$

Then we ask that $\{f_\alpha^\dagger, f_\beta^\dagger\} = 0$, and $\{f_\alpha^\dagger, f_\beta\} = \delta_{\alpha\beta}$.

Again we have a number operator $n_\alpha \coloneqq f_\alpha^\dagger f_\alpha$; it satisfies $n_\alpha f_\alpha = f_\alpha(n_\alpha - 1)$, and measures the number of particles in the state α. Because

$$(f_\alpha^\dagger)^2 = f_\alpha^\dagger f_\alpha^\dagger = -f_\alpha^\dagger f_\alpha^\dagger = 0, \tag{2.8}$$

then n_α is a projector (i.e. $n_\alpha^2 = n_\alpha$), and therefore its eigenvalues can only be 0 or 1. This encodes the Pauli exclusion principle: there can be at most a single fermion in a given state.

We'd like to write our second-quantized systems with quadratic Hamiltonians, largely because these are tractable. Let $(h_{\alpha\beta})$ be a self-adjoint matrix and consider the Hamiltonian

$$H \coloneqq \sum_{\alpha,\beta} f_\alpha^+ h_{\alpha\beta} f_\beta. \tag{2.9}$$

The *number operator* $N \coloneqq \sum n_\alpha$ commutes with the Hamiltonian, which therefore defines a symmetry of the system. The associated conserved quantity is the particle number. Slightly more explicitly, we have a symmetry of the group U_1 (i.e. the unit complex numbers under multiplication): for $\theta \in [0, 2\pi)$, let

$$u_\theta \coloneqq \exp(i\theta N). \tag{2.10}$$

Then

$$u_\theta^\dagger H u_\theta = \sum_{\alpha,\beta} \underbrace{u_\theta^\dagger f_\alpha^\dagger u_\theta}_{=e^{-i\theta} f_\alpha^\dagger} h_{\alpha\beta} \underbrace{u_\theta^\dagger f_\beta u_\theta}_{=e^{i\theta} f_\beta} = H. \tag{2.11}$$

When you see a Hamiltonian, you should feel a deep-seated instinct to diagonalize it: we want to find $\lambda_n, v^{(n)}$ such that $h_{\alpha\beta} v_\beta^{(n)} = \lambda_n v_\alpha^{(n)}$ and $vv^\dagger = \mathrm{id}$. Let $v_{n\alpha} \coloneqq v_\alpha^{(n)}$ and

$$\psi_n \coloneqq \sum_\alpha v_{n\alpha} f_\alpha. \tag{2.12}$$

Then $\psi_n^\dagger$ and ψ_n satisfy the same creation and annihilation relations as $f_\alpha^\dagger$ and f_α:

$$\{\psi_n^\dagger, \psi_m\} = \{\sum_\alpha v_{n\alpha}^* f_\alpha^\dagger, \sum_\beta v_{m\beta} f_\beta\} \tag{2.13}$$

$$= \sum_{\alpha,\beta} v_{n\alpha}^* v_{m\beta} \underbrace{\{f_\alpha^\dagger, f_\beta\}}_{=\delta_{\alpha\beta}} \tag{2.14}$$

$$= \sum_\alpha v_{m\alpha} (v^\dagger)_{n\alpha} = \delta_{m,n}. \tag{2.15}$$

Let $\hat{n}_n \coloneqq \psi_n^\dagger \psi_n$. Now the Hamiltonian has the nice diagonal form

$$H = \sum_n \lambda_n \psi_n^\dagger \psi_n, \tag{2.16}$$

and we can explicitly calculate its action on a state:

$$H\psi_{n_1}^\dagger \psi_{n_2}^\dagger \cdots \psi_{n_N}^\dagger |\varnothing\rangle = \underbrace{\left(\sum_m \lambda_m \psi_m^\dagger \psi_m \psi_{n_1}^\dagger\right)}_{(*)} \psi_{n_2}^\dagger \cdots \psi_{n_N}^\dagger |\varnothing\rangle. \tag{2.17}$$

The term $(*)$ is equal to

$$\psi_m^\dagger(\delta_{mn} - \psi_{n_1}^\dagger \psi_m) = \delta_{mn_1} \psi_{n_1}^\dagger + \psi_{n_1}^\dagger \psi_m^\dagger \psi_m. \tag{2.18}$$

Then (2.17) is equal to

$$(2.17) = \lambda_{n_1} \psi_{n_1}^\dagger (\psi_{n_2}^\dagger \cdots \psi_{n_N}^\dagger)|\varnothing\rangle, \tag{2.19}$$

so we've split off a term and can induct. The final answer is

$$= \left(\sum_{i=1}^N \lambda_i\right) \psi_{n_1}^\dagger \cdots \psi_{n_N}^\dagger |\varnothing\rangle. \tag{2.20}$$

Example 2.21 (1d tight binding model)**.** Let's consider the system on a circle with L sites (you might also call this periodic boundary conditions). We have operators which create fermions at each state and also some sort of tunneling operators. The Hamiltonian is

$$H \coloneqq -t \sum_{j=1}^L (f_{j+1}^\dagger f_j + f_j^\dagger f_{j+1}) - \mu \sum_{j=1}^L f_j^\dagger f_j, \tag{2.22}$$

where $j+1$ is interpreted mod L as usual. One of t and N (TODO: which?) can be interpreted as the chemical potential. The eigenstates are the Fourier modes

$$\psi_k \coloneqq \frac{1}{\sqrt{L}} \sum_{j=1}^L e^{ikj} f_j, \tag{2.23}$$

where $k = 2\pi n/L$. Hence in particular $e^{ik(L+1)} = e^{ik}$. Now we can compute

$$\sum_{j=1}^{L} f_{j+1}^\dagger f_j = \frac{1}{L}\sum_{j,k,k'} e^{ik'(j+1)} e^{-ikj} \psi_{k'}^\dagger \psi_k \tag{2.24}$$

$$= \frac{1}{L}\sum_{k,k'} e^{ik'} \sum_j e^{ij(k-k')} \psi_{k'}^\dagger \psi_k \tag{2.25}$$

$$= \sum_k e^{ik} \psi_k^\dagger \psi_k. \tag{2.26}$$

That is, the diagonalized Hamiltonian is

$$H = \sum_{k=1}^{L} (-2t\cos k - \mu)\psi_k^\dagger \psi_k. \tag{2.27}$$

You can plot λ_k as a function of k, but really k is defined on the circle $\mathbb{R}/2\pi\mathbb{Z}$, which is referred to as the *Brillouin zone*. The ground state of the system is to fill all states with negative energy:

$$|\text{G.S.}\rangle = \left(\prod_{k:\lambda_k<0} \psi_k^\dagger\right)|\varnothing\rangle. \tag{2.28}$$

If L is fixed, k only takes on L different values, but implicitly we'd like to take some sort of thermodynamic limit $L \to \infty$, giving us the actually smooth function $\lambda_k = -2t\cos k - \mu$. ◂

We said that second quantization is useful when the particle number can change, so let's explore that now. This would involve a Hamiltonian that might look something like

$$H = f_\alpha^\dagger h_{\alpha\beta} f_\beta + \frac{1}{2}\Big(\Delta_{\alpha\beta} f_\alpha^\dagger f_\beta^\dagger + \Delta_{\alpha\beta}^\dagger f_\alpha f_\beta\Big). \tag{2.29}$$

These typically arise in mean-field descriptions of superconductors. This typically arises in situations where electrons are attracted to each other — this is a little bizarre, since electrons have the same charge, but you can imagine an electron moving in a crystalline solid with some positive ions. The electron attracts the ions, but they move more slowly, so the electron keeps moving and we get an accumulation of positive charge, and this can attract additional electrons.

This binds pairs of electrons together at a certain point, and this forms a *condensate*, i.e. a superposition of states with different particle numbers. (2.29) describes a superconducting condensate, in which $\Delta_{\alpha\beta}$ describes pairs of particles appearing or disappearing in the condensate. To learn more, take a solid-state physics class.

Remark 2.30. You have to have pairs of fermionic terms — if you try to include an odd number of fermions, or a single fermionic term, you'll get nonlocal interactions between the lone fermion and others. Thus, even though the particle number is not conserved, its value mod 2, which is called *fermion parity*, is conserved. ◂

If you try to directly diagonalize (2.29), some weird stuff happens, so we'll rewrite the Hamiltonian such that it looks like it's particle-conserving, and then apply our old trick. This approach is due to Nambu. Let

$$\Psi_{\alpha,\tau} \coloneqq \begin{pmatrix} f_\alpha \\ f_\alpha^\dagger \end{pmatrix}, \tag{2.31}$$

where τ denotes the vertical index. We can rewrite the Hamiltonian as

$$H = \frac{1}{2}\begin{pmatrix} f_\alpha^\dagger & f_\alpha \end{pmatrix}\begin{pmatrix} h_{\alpha\beta} & \Delta_{\alpha\beta} \\ \Delta_{\alpha\beta}^\dagger & -h_{\alpha\beta}^\dagger \end{pmatrix}\begin{pmatrix} f_\beta \\ f_\beta^\dagger \end{pmatrix} + (\text{constant}) = \frac{1}{2}\Psi_{\alpha\tau}^\dagger \mathcal{H}_{\alpha\beta\tau\tau'}\Psi_{\beta\tau'}. \tag{2.32}$$

However, Ψ and $\Psi^\dagger$ have some redundancy: if σ^x denotes the Pauli matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, then $\Psi_\tau^\dagger = \sigma^x_{\tau\tau'}\Psi_{\tau'}$.

This is telling us that Ψ and $\Psi^\dagger$ create particles with energies (say) e and $-e$, respectively. Now

$$\Psi^\dagger H \Psi = \Psi^{\mathrm{T}}\sigma^x H \sigma^x (\Psi^\dagger)^{\mathrm{T}} = -\Psi^\dagger \sigma^x H^{\mathrm{T}} \sigma^x \Psi, \tag{2.33}$$

and therefore $\mathcal{H} = -\sigma^x \mathcal{H}^* \sigma^x$. Using this, we can determine the eigenstates: $\mathcal{H}v = Ev$ iff $\mathcal{H}\sigma^x v^* = -E\sigma^x v^*$. Then

$$\mathcal{H}\sigma^x v^* = \sigma^x(\sigma^x \mathcal{H}\sigma^x)v^* = \sigma^x(\sigma^x \mathcal{H}^* \sigma^x v)^* = \sigma^x(-\mathcal{H}v)^* = -E\sigma^x v^* \tag{2.34}$$

and

$$\gamma_E \coloneqq \sum_{\alpha,\tau} v_{\alpha\tau}\Psi_{\alpha\tau} \tag{2.35}$$

satisfies $\gamma_{-E} = \gamma_E^\dagger$. TODO: what are we trying to show here?

This $E \leftrightarrow -E$ symmetry is an instance of what's traditionally called "particle-hole symmetry," but it's a little weird — we can't break this symmetry by introducing additional terms to the Hamiltonian. So it might be more accurate to call it *particle-hole structure*, which conveniently has the same acronym.

TODO: some other stuff I missed. I think $\{\Psi_{\alpha\tau}, \Psi^\dagger_{\beta\tau'}\} = \delta_{\alpha\beta}\delta_{\tau\tau'}$ and $\{\gamma_E, \gamma^\dagger_{E'}\} = \delta_{EE'}$, which tells us these (I think) behave like creation and annihilation operators.

At zero energy, $\gamma_0 = \gamma_0^\dagger$, so we have a fermion which is its own antiparticle. This is called a *Majorana fermion*. It will be helpful to have a slightly different normalization here, which we'll discuss more later.

Lecture 3.

The Majorana chain: 9/5/19

Today we will discuss a one-dimensional system studied by Kitaev [Kit01]. Introduce periodic boundary conditions, so that the sites live on a circle with length L. At each site i, we have a local Hilbert space $\mathcal{H}_i \coloneqq \mathbb{C}\cdot\{|0\rangle, |1\rangle\}$, and the total Hilbert space of states is the tensor product of these over all of the states.

Let c_j and $c_j^\dagger$ denote the annihilation, resp. creation operators at site j. Then the Hamiltonian is

$$H \coloneqq -\sum_{j=1}^{L} t(c_{j+1}^\dagger c_j + c_j^\dagger c_{j+1}) - \mu\sum_{j=1}^{L} c_j^\dagger c_j - \Delta(c_{j+1}^\dagger c_j^\dagger + c_j c_{j+1}). \tag{3.1}$$

Here t, Δ, and μ are parameters; μ is called the *chemical potential*.

To solve this Hamiltonian, we will introduce a different set of creation and annihilation operators: let

$$\tilde{c}_k \coloneqq \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{ikj} c_j \tag{3.2a}$$

$$\tilde{c}_k^\dagger \coloneqq \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{ikj} c_j^\dagger. \tag{3.2b}$$

Here $k \in 2\pi n/L$, like last time, and we only consider those k in $[0, 1)$.[1] Using these, we can rewrite (3.1) as

$$H = \sum_k (-2t\cos k - \mu)\tilde{c}_k^\dagger \tilde{c}_k - \Delta\sum_k \left(e^{ik}\tilde{c}_k^\dagger \tilde{c}_{-k}^\dagger + e^{ik}\tilde{c}_{-k}\tilde{c}_k\right). \tag{3.3}$$

To get the last term, use the fact that $\tilde{c}_k^\dagger \tilde{c}_{-k}^\dagger = -\tilde{c}_{-k}^\dagger \tilde{c}_k^\dagger$, so

$$\frac{1}{2}\sum_k \tilde{c}_k^\dagger \tilde{c}_{-k}^\dagger e^{ik} + \frac{1}{2}\sum_k \left(-e^{-ik}\tilde{c}_k^\dagger \tilde{c}_{-k}^\dagger\right). \tag{3.4}$$

Again introduce the Nambu spinor $\Psi \coloneqq \begin{pmatrix}\tilde{c}_k \\ \tilde{c}_k^\dagger\end{pmatrix}$; then we can rewrite (3.3) as

$$H = \frac{1}{2}\sum_k \Psi_k^\dagger \begin{pmatrix} -2t\cos k - \mu & 2i\Delta\sin k \\ -2i\Delta\sin k & 2t\cos k + \mu\end{pmatrix}\Psi_k. \tag{3.5}$$

So now all we have to do is diagonalize a 2×2 matrix, which isn't so hard. In particular, the eigenvalues (energy levels) are

$$E_k = \pm\frac{1}{2}\sqrt{(2t\cos k + \mu)^2 + (2\Delta\sin k)^2}. \tag{3.6}$$

In particular, we can plot these as k varies and see whether the system is gapped.

- Suppose $\Delta = \mu = 0$. Then there are values of k such that the spectrum isn't gapped, but as soon as you make $\Delta \neq 0$, there is a spectral gap.

[1] To me (Arun), this looks like k is in the Pontrjagin dual $(\mathbb{Z}/L)^\vee$, which would be appropriate if this is a Fourier transform.

- If $\mu = -2t$, we again close the gap at $\Delta = 0$ and $k = 0$, but in general there is a gap.

So the phase diagram in μ appears to have three phases and two transitions between them, and is symmetric about $\mu \mapsto -\mu$. For $\mu \to -\infty$, this is adiabatically connected to a trivial phase, and thus is itself trivial: there are no particles. For $\mu \to \infty$, it is also trivial: every site is occupied in the ground state. The third phase is a topological superconductor (though we have yet to show it).

So the two phase transitions happen at $\mu = \pm 2t$. Suppose $\mu = -2t + M$, where M is close to zero, so that we can study the phase transition. Since $-2t \cos k + 2t = O(k^2)$, we'll ignore it, and therefore replace M with $M - 2t \cos k + 2t$. Similarly,

$$2i\Delta \sin k = 2i\Delta k + O(k^3), \tag{3.7}$$

and we will drop the higher-order terms. Under these approximations, the Hamiltonian now is

$$\begin{aligned} H &\approx \frac{1}{2} \sum_k \Psi + k^\dagger \left(\begin{pmatrix} M & 0 \\ 0 & -M \end{pmatrix} + \begin{pmatrix} 0 & 2 - \Delta \\ -2i\Delta & 0 \end{pmatrix} k \right) \Psi_k \\ &= \frac{1}{2} \sum_k \Psi_k^\dagger (M\sigma^z - 2k\Delta \sigma^y) \Psi_k. \end{aligned} \tag{3.8}$$

Here σ^z and σ^y are the usual Pauli matrices. Now (3.8) looks like a Dirac equation with a mass term: letting $a \coloneqq -2\Delta$, we get

$$H \approx \frac{1}{2} \sum_k \Psi_k^\dagger (ak\sigma^y + M\sigma^z) \Psi_k \approx \frac{1}{2} \int \mathrm{d}x \, \psi^\dagger(x)(-iv\partial_x \sigma^y + M(x)\sigma^z)\psi(x). \tag{3.9}$$

Let's let M vary in space, so that we have a defect between the two phases at $x = 0$. We'll show that the defect has a bound state.

Consider the $E = 0$ solution to the continuum approximation in (3.9). Then

$$\tilde{\gamma} = \int \mathrm{d}x \, \phi_\alpha(x) \hat{\psi}_\alpha(x), \tag{3.10}$$

and we end up with an ordinary differential equation for the solution:

$$(-iv\partial_x \sigma^y + M(x)\sigma^z)\phi_\alpha(x) = 0, \tag{3.11}$$

and therefore $\frac{\partial}{\partial x}\phi_\pm = \pm(M/v)\phi$ (TODO: some details here are missing). Therefore

$$\phi_\pm(x) = \exp\left(\pm \int_0^x \frac{M(x)}{v} \, \mathrm{d}x \right) \phi_\pm(0). \tag{3.12}$$

One of these blows up at infinity and makes no physical sense, but there is a solution which is largest at $M = 0$ and decays to zero at infinity. This is the bound zero mode ϕ_+. Here are a few more facts about this zero mode.

- As $x \to \pm\infty$, $\phi_+(x) \to \exp(-|M_0||x|)$.
- We didn't use much about $M(x)$, only the fact that it switches sign at $M = 0$. This is the sense in which the zero mode is topological: we can deform $M(x)$ and obtain the same behavior.[2]
- $\gamma = \gamma^\dagger$: in a sense, this mode is both a creation and annihilation operator. For this reason, it's called a *Majorana zero mode*.
- The side of these bound states is determined by the *correlation length* $v/M = 1/3$.

This is not a critical system — besides this zero mode, all other phases are gapped. For $M > 0$, we get a trivial insulator, and for $M < 0$, we have the topological phase, a topological superconductor. The bound state at the defect is what implies that the $M < 0$ phase isn't trivial.

How realistic are the periodic boundary conditions? Well, we can't create an infinite wire in the lab, so maybe we should work on the unit interval of length L, which is large with respect to the correlation length. Then, you maybe can convince yourself that there are two Majorana modes, one at each boundary site, and they overlap a little bit in the bulk, approximately at order $e^{-L/3}$. Call these modes γ_L and γ_R. If you let $\psi = (\gamma_M + i\gamma_R)/2$ and $\psi^\dagger = (\gamma_L - i\gamma_R)/2$, then these satisfy the anticommutation relations of creation and annihilation operators of ordinary fermions: for example, $\{\psi, \psi^\dagger\} = 1$. This is a little bit weird.

[2] This is an instance of a very general theorem in mathematics called the Atiyah-Singer index theorem, which can be used to produce zero modes in fermionic systems.

Another weird aspect of this system is that if L is large enough, you can't couple to both γ_L and γ_R at the same time. If you tried to perturb the system, say by introducing a bosonic field with an electric potential $V = \phi\gamma_L$, well, that's not allowed, because you would get an odd number of fermions. So these Majorana modes are protected by small perturbations, and in that sense might be useful if you care about quantum memory. The drawback is that you can't put a state with an even number of fermions and a state with an odd number of fermions into superposition, which is unfortunate; the solution is to consider several separate copies of this system.

So let's work with N wires, meaning we have $2N$ Majorana zero modes $\gamma_1, \dots, \gamma_{2N}$, hence N ordinary fermion creation/annihilation operator pairs $\psi^\dagger, \psi$ as we discussed above. This system has a 2^N-dimensional space of ground states: for any subset $S \subset \{1, \dots, n\}$, we say that the fermion state ψ is occupied for $i \in S$ and unoccupied for $i \notin S$.

The fermion parity

$$P_f \coloneqq \prod_{i=1}^{N} i\gamma_{2i-1}\gamma_{2i} \in \{\pm 1\} \tag{3.13}$$

is a conserved quantum number of this system (intuitively, it tells us whether there are an even or odd number of fermions present). Therefore we have 2^{N-1} states available as quantum memory.

These give us different ways to label the ground states, but different labelings interact in complicated ways. For example, if $N = 2$,

$$|P_{14} = 1, P_{23} = 1\rangle = \frac{1}{\sqrt{2}}(|P_{12} = 1, P_{34} = 1\rangle - |P_{12} = -1, P_{34} = -1\rangle). \tag{3.14}$$

You can imagine this as follows: we begin with no particles, and create two fermions on each copy of the chain. This doesn't change the parity, because we created them from nothing. Now, we smush together sites 2 and 3 and measure there, and get zero. Then, this is telling us that the remaining states are maximally uncertain. This was an operator

$$|{+}{+}\rangle \longmapsto \frac{1}{\sqrt{2}}(|{+}{+}\rangle - |{-}{-}\rangle). \tag{3.15}$$

This is a topologically protected operation, which is exciting if you want to make quantum computers. But it's proven to not be universal, i.e. we can't get (or even well-approximate) all quantum gates in the Majorana chain. In fact, what we get can be efficiently approximated by a classical computer, and this isn't even universal for classical computer! But there are other examples of phases which are universal for quantum computing, and Microsoft is researching how these might be actually implemented.

These states can have (quasi)particle modes akin to the Majorana modes here. In general these are called *nonabelian anyons* or *nonabelian defects*. The defining property of these is that there is a topologically protected ground state degeneracy associated to the zero modes, and it grows exponentially in the number of particles present. The process of turning two particles into something else will be called *fusion*; for the Majorana chain we have the relation

$$\gamma \times \gamma = 0 + f, \tag{3.16}$$

as we either get nothing or a fermion. This is akin to the fact that if we collide two spin-1/2 particles, they could annihilate each other or produce a spin-1 particle. The fact that the Hilbert space grows exponentially is reminiscent of the fact that for a spin-s particle, the Hilbert space of states has dimension $2s + 1$ to the number of particles: the local dimension is the number of objects. Here, though, we will encounter examples of nonabelian anyons whose quantum dimensions are irrational.

Next time, we'll argue that the Majorana chain is the only nontrivial topological phase that can occur among 1D superconductors (unless we add additional symmetries to the Hamiltonian). We'll also discuss how to see that the phase is nontrivial in the bulk; after that, we'll discuss some possible physical realizations in real system, such a spin-orbit coupled semiconductor wire, put in contact with a normal superconductor.

Lecture 4.

The Majorana chain, II: 9/10/19

Today, we'll continue to discuss the Kitaev chain from last time. Recall that the phase diagram in μ has two trivial phases for $\mu \gg 0$ and $\mu \ll 0$, and the phase in between them is topological, specifically some kind of topological superconductor. The trivial phase corresponds to $\mu \neq 0$ and $t, \Delta = 0$, and the topological phase to $\mu = 0$ and $t = \Delta$.

In the topological phase formulated on an interval, there are protected zero modes at the boundary, corresponding to operators γ_L and γ_R (for the left-hand and right-hand sides of the interval, respectively). These are both self-adjoint. Writring $\psi \coloneqq (\gamma_L + i\gamma_R)/2$, $\psi + \psi^\dagger = \gamma_L$ and $\psi - \psi^\dagger = \gamma_R$.

In the topological phase, we can simplify the Hamiltonian slightly:

$$H = \Delta \sum_i (c_{i=1}^\dagger - c_{i+1})(c_i^\dagger + c_i) = i\Delta_j \overline{\gamma_{j+1}} \gamma_j, \tag{4.1}$$

where $\gamma_i = c_i + c_{i+1}^\dagger$ and $\overline{\gamma_i} = -i(c_i - c_{i+1}^\dagger)$. These look like particle creation and annihilation operators for a pair of Majorana fermions. These commute with the Hamiltonian.

This system admits the following interpretation. At each site i, we have two places where there can be a Majorana fermion, γ_i and $\overline{\gamma_i}$. However, the Majorana corresponding to γ_i and the one corresponding to $\overline{\gamma_{i+1}}$ are coupled. This explains what we see on the interval: at each boundary site, one of the two terms can't be paired up, and we obtain the boundary zero modes. With periodic boundary conditions, all Majoranas can be paired up, and we obtain a single ground state.

We can also consider anti-periodic boundary conditions, where we say that traversing once around the circle picks up a minus sign: specifically, throw a minus sign onto the interaction term between sites 1 and N.[3] This means that we again obtain a single ground state, but the number of fermions differs by 1 from the ground state of the periodic boundary conditions.[4] This was a little handwavy, but we'll go into more detail in a bit; in general, the difference in fermion parity in different boundary conditions is a useful invariant of these systems.

Remark 4.2. The trivial phase of the Kitaev chain admits a similar description: we have $\gamma_i, \overline{\gamma_i}$ as before, but now γ_i is coupled to $\overline{\gamma_i}$, and there are no interactions between different sites. Therefore on an interval or circle, we always pair up the Majorana fermions, and don't obtain any boundary zero modes. ◁

Another thing we can do is combine phases: formulate two copies of the Kitaev chain on the interval, but such that they interact very weakly, certainly not enough to close the gap. This operation, called *stacking*, is an algebraic operation on phases. In particular, given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ with Hamiltonians $H_i \colon \mathcal{H}_i \to \mathcal{H}_i$, the stacked phase has Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ and Hamiltonian $H_1 \otimes \mathrm{id}_{\mathcal{H}_2} + \mathrm{id}_{\mathcal{H}_1} \otimes H_2$, maybe plus some small interaction term. By commutativity of tensor product, stacking is an associative and commutative operation; moreover, the trivial insulator is an identity for this operation. All of the one-dimensional phases P we encounter will be *invertible*, i.e. there's some other phase P' such that $P \otimes P'$ is trivial. We consider two phases the same if they can be connected by a deformation which is local and never closes the gap.

So given a dimension, we obtain an abelian group, called the group of invertible phases in that dimension. This is not always finite, e.g. the phases delineated by the integer quantum Hall effect produce a countable subgroup in dimension 2. The group itself is expected to be finitely generated.

Determining these groups is an interesting theoretical question in condensed-matter physics. Let's see what it is in dimension 1. Stack together two copies of the Kitaev chain in its topological phase, and add a small interaction term between the two copies of γ_L, and between the two copies of γ_R. Then they can be paired up, and in the end there are no edge states, suggesting that this tensor product is trivial. There's additional evidence that it's trivial, e.g. switching from periodic to antiperiodic boundary conditions doesn't change fermion parity on this system, because the two changes cancel out. Hence we obtain a $\mathbb{Z}/2$ subgroup

[3] In the continuum perspective, this corresponds to choosing the nonbounding spin structure on S^1, rather than the bouning spin structure.

[4] The actual fermion parity of the ground state can be changed by a local potential: flip some signs in the Hamiltonian. But the difference between the two boundary conditions is an invariant.

of the group of phases in dimension 1. This may seem very abstract, but similar considerations apply to more general fermionic systems, and this has guided researchers on where to look for such systems in real materials.

In general, if ground states have even fermion parity on all closed boundary systems, and there are no edge modes, then the phase is trivial. The argument is that we can divide the interval[5] into a bunch of pieces, and then each piece can be adiabatically transformed into the trivial insulator.

$$\sim \cdot \sim$$

Now let's make this story into something more concrete. The *fermion parity* is defined to be

$$P_F = (-1)^{\sum_i \widehat{n}_i} = \prod_j i\gamma_j\overline{\gamma_j}, \tag{4.3}$$

so 1 if there is an even number of fermions and -1 if there's an odd number. This commutes with the Hamiltonian, and in fact commutes with any local Hamiltonian, because local Hamiltonians must have an even number of fermions in each term, as we saw before.

Fermionic parity looks like a discrete symmetry, but it's a bizarre one, in that we can't break it or couple it to a local background field. Otherwise it behaves mostly like any other discrete symmetry.

The change in boundary conditions between periodic and anti-periodic correspond to a symmetry flux for fermion parity: given a local operator $\mathcal{O}$, we can act on it (thought of, I think, as moving it past the antiperiodic boundary condition) and obtain $P_F\mathcal{O}P_F^\dagger$.

We will later discuss how to promote a global symmetry, such as fermion parity, to a local one, and then gauge it. This will be a powerful nonperturbative way to study phases. In the Kitaev chain, this concretely looks like considering fermions at different sites with bonds between sites i and $i+1$ with a σ^z on the bond. Then we rewrite the Hamiltonian by

$$i\gamma_i\overline{\gamma_{i+1}} \longmapsto i\gamma_i\sigma^z_{i,i+1}\overline{\gamma_{i+1}}. \tag{4.4}$$

Therefore given $s_i \in \mathbb{Z}/2$ for each i, we can act by fermion parity at site i, which sends $\gamma_i \mapsto s_i\gamma_i$, $\overline{\gamma_i} \mapsto s_i\overline{\gamma_i}$, and

$$\sigma^z_{i,i+1} \longmapsto s_{i+1}s_i\sigma^z_{i,i+1}. \tag{4.5}$$

The generator of this symmetry is the operator

$$\prod_i \left((\sigma^x_{i-1,i}\sigma^x_{i,i+1})(\gamma_i\overline{\gamma_i})\right)^{(1+s_i)/2}. \tag{4.6}$$

This may look a little bizarre, but is reminiscent of something more familiar in electromagnetism: given a function ϕ on spacetime, we act on the gauge field by

$$A_\mu \longmapsto A_\mu + \partial_\mu\phi \tag{4.7}$$

and $\psi(\mathbf{r},t) \mapsto \exp(i\phi(\mathbf{r},t))\psi(\mathbf{r},t)$. The generator of this symmetry is

$$\exp\left(i\int \phi(\mathbf{r})(n(\mathbf{r}) - \nabla\cdot\mathbf{E})\right). \tag{4.8}$$

This has promoted a global number symmetry to a gauge symmetry for the gauge group U_1. In our example, U_1 is replaced with $\{\pm 1\}$, $\sigma^z_{i,i+1}$ plays the role of $\exp(i\int_i^{i+1} A\cdot \mathrm{d}\ell)$, and s_i plays the role of $\exp(i\phi_r)$ restricted to site i — if this doesn't make perfect sense, that's OK, because this is an approximate analogy.

The physical states are obtained by projecting onto the gauge-invariant states, which you can think of as averaging over the elements of a larger group, the product of copies of $\{\pm 1\}$ indexed over the sites. Explicitly, if you call this group G,

$$|\psi_{\text{physical}}\rangle = \frac{1}{|G|}\sum_{g\in G}|\psi_{\text{gauge-dependent}}\rangle. \tag{4.9}$$

The flux $\prod_i \sigma^z_{i,i+1}$ is gauge-invariant, in that a gauge transformation $(s_i) \in G$ acts by

$$\prod_i \sigma^z_{i,i+1} \longmapsto \prod_i (s_i\sigma^z_{i,i+1}s_{i+1}) = \dots s_i\sigma^z_{i,i+1}s_{i+1}s_{i+1}\sigma^z_{i+1,i+2}\dots \tag{4.10}$$

[5]This argument applies in dimension 1, but the result is true more generally

and these two copies of s_{i+1} cancel. The flux squares to 1, so has eigenvalues $\{\pm 1\}$; the states with eigenvalue 1 are said to have no flux, and those with eigenvalue -1 are said to have flux. This will be our way of formalizing that we're in the antiperiodic boundary condition.

Recall that the Hamiltonian for the trivial phase is

$$H_{\text{triv}} - \mu \sum_i c_i^\dagger c_i. \tag{4.11}$$

If you add σ^z bonds, this Hamiltonian does not change (TODO: I think). For the topological phase, we had

$$H_{\text{top}} = -i\Delta \sum_j \gamma_j \overline{\gamma_{j+1}} \longmapsto i\Delta \sum_j \gamma_i \sigma^z_{j,j+1} \overline{\gamma_{j+1}}. \tag{4.12}$$

The total fermion parity of the ground state is

$$P_F^{GS} = \prod_j i\overline{\gamma_j}\gamma_j \tag{4.13}$$

$$= \pm \prod_j i\gamma_j \overline{\gamma_{j+1}}. \tag{4.14}$$

The sign is equal to $\prod_j \sigma^z_{j,j+1}$, which is -1 to the number of fluxes. Let F denote the operator that measures fluxes (TODO: didn't quite understand the definition), then

$$F P_F F^{-1} P_F^{-1} = -1 \tag{4.15}$$

in the topological phase, but not in the trivial phase, so (4.15) is the topological invariant of phases in this dimension.

Remark 4.16. A fact (which we won't prove) is that in general fermionic systems, one can use fluxes to remove fermions unless an unusual commutation relation such as (4.15) holds. ◂

Searching for a condition such as (4.15) is a powerful way to classify invertible fermionic phases: since it didn't make reference to band structures, it also applies to interacting systems. It can be souped up a bit and made even more powerful.

Remark 4.17. Sometimes this phase is called a symmetry-protected phase (SPT), where the symmetry is fermion parity, but as mentioned above this behaves differently than other symmetries, so calling this an SPT seems a little strange. In any case, much of what we've just discussed generalizes to other symmetry groups and hence to SPTs. ◂

Now let's turn to the classification question mentioned above. Band structures are our first tool. Consider a (Fourier-transformed) Hamiltonian

$$H = \sum_k \psi^\dagger_{k,\alpha} h^{(k)}_{\alpha\beta} \psi_{k,\beta}. \tag{4.18}$$

So if there are L lattice sites, $k \in (\mathbb{Z}/L)^\vee$, or, in physical language, k is a momentum. As k varies, the spectrum is a bunch of curves, called *bands*; suppose that N are above zero and M are below zero. We can use local terms to flatten the bands, preserving the gap to be at least Δ for all bands. Concretely, suppose

$$h(k) = \sum_n \varepsilon_n(k) \prod_n (k). \tag{4.19}$$

Here $\varepsilon_n(k)$ is the n^{th} eigenvalue when the momentum is k. Thus we want to diagonalize (TODO: was erased before I could write it down), and we're interested in the topology of such basis transformations: gap-preserving deformations of the Hamiltonian produce a homotopy of these unitary operators.

As we saw, taking $L \to \infty$ means we think of k as living on the circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ (the Brilloun zone); in general dimension d, it lives on the torus $T^d = (S^1)^{\times d}$.

So in general, we're interested in maps $T^d \to \mathrm{U}_{N+M}$, which following what we did above we call a *band structure*. However, unitary operators in U_N or U_M give us trivially systems, in that they shuffle the indices of the bands around, so what we really want to compute is the set of homotopy classes of maps

$$T^d \to \mathrm{U}_{N+M}/(\mathrm{U}_N \times \mathrm{U}_M). \tag{4.20}$$

We will delve more into this next time.

Lecture 5.

Classification of band structures: 9/12/19

As we mentioned last time, it's possible to *stack* two topological phases: put them both on the same medium, with very weak interactions between them. This defines an associative, commutative operation on phases, and the trivial phase is a unit; hence those phases which are invertible under this operation form an abelian group, and we're interested in studying this group.

For interacting systems in arbitrary dimensions, this is still a bit of an open question; certainly people know what the expected answer is, but we haven't figured out all the details.

Today, we'll restrict ourself to something less ambitious: translation-invariant, noninteracting fermion systems in dimension d.[6] In this setting the mathematical formalism is understood.

We impose translation-invariance so that we can work in momentum space (i.e. take the Fourier transform of the Hamiltonian); specifically, translation invariance imposes periodic boundary conditions, so momentum space, which we also call the Brillouin zone, is a d-dimensional torus. The Fourier-transformed Hamiltonian has the general form

$$H = \sum_{k\in T^d} \psi^\dagger_{k,\alpha} h_{\alpha,\beta}(k)\psi_{k,\beta}. \tag{5.1}$$

Since we work on a lattice and let the number of sites go to infinity, really k also lives on a lattice inside T^d, where we also refine this lattice. We want $\sigma^x h^*(-k)\tau^x = -h(k)$, where σ^x is $\left(\begin{smallmatrix} I_n & 0 \\ 0 & -I_m \end{smallmatrix}\right)$: replacing ± 1 with identity matrices.

Hence h is self-conjugate. We also want the gap in the spectrum of H to be preserved for all k. Therefore the eigenvalues $\varepsilon_1(k), \dots, \varepsilon_n(k)\colon T^d \to \mathbb{R}\setminus 0$, called *bands*, are continuous functions that do not cross the origin, so we obtain two invariants: let n denote the number of positive bands and m the number of negative bands.

We can diagonalize h by unitary matrices $U(k)$ that also vary continuously in k:

$$h = U(k)^\dagger \begin{pmatrix} \varepsilon_1(k) & & \\ & \ddots & \\ & & \varepsilon_n(k) \end{pmatrix} U(k). \tag{5.2}$$

In this case the Hamiltonian simplifies to

$$H = \Delta U(k)^\dagger \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix} U(k), \tag{5.3}$$

where I_ℓ is the identity matrix of size $\ell \times \ell$.

Let U_ℓ denote the Lie group of $\ell \times \ell$ unitary matrices. By the above discussion, a gapped phase defines a map $T^d \to \mathrm{U}_{n+m}$. However, this overdetermines the classification:

- You can permute the indices of the positive bands, and of the negative bands, so really the target space should be $\mathrm{U}_{n+m}/(\mathrm{U}_n \times \mathrm{U}_m)$.
- A smooth deformation of the Hamiltonian by local terms induces a homotopy between the maps, so the classification should deal with homotopy classes of maps, denoted $[T^d, \mathrm{U}_{n+m}/(\mathrm{U}_n \times \mathrm{U}_m)]$.
- Finally, one can add extra degrees of freedom, so the classification stabilizes in n and m. Therefore what we really need to consider is homotopy classes of maps

$$T^d \longrightarrow \bigcup_{n,m} \mathrm{U}_{n+m}/(\mathrm{U}_n \times \mathrm{U}_m). \tag{5.4}$$

This is a mathematical question, and while the tools used to solve it are slightly beyond the scope of this class, they're well-understood; this is the part of mathematics called *K-theory*. Specifically, one can compute that in $d = 1$, the group is trivial. For $d = 2$, there's a $\mathbb{Z}$, corresponding to some quantized quantity. In general, the group is 0 for d odd and $\mathbb{Z}$ for d even; there's always a periodicity in this classification.

Now we haven't used all of the structure yet (which I think means topological superconductors rather than topological insulators? TODO). The self-conjugacy condition on h means we can diagonalize the system using

[6] Noninteracting bosons are also simple: they just form a superfluid, which cannot form a gapped phase.

orthogonal matrices rather than just unitary ones, and forces $m = n$. Therefore the target space is instead

$$\bigcup_n \mathrm{O}_{2n}/\mathrm{U}_n. \tag{5.5}$$

The classification in $d = 1$ is now $\mathbb{Z}/2$, corresponding to the $\mathbb{Z}/2$ generated by the Kitaev chain that we studied before. In general, the group is always 0, $\mathbb{Z}/2$, or $\mathbb{Z}$, and has periodicity of order 8.

$$\sim \cdot \sim$$

Whether or not you like this level of mathematical formalism, you can use it to guide your considerations in experiments. For example, in a dimension where the classification is a $\mathbb{Z}/2$, you might guess it has something to do with unpaired fermions like in the Majorana chain.

If you try to do this naïvely, you end up asking for same-spin coupling, which is very unusual, and would be difficult to engineer. This led to a different, more relativistic answer invoking *spin-orbit coupling*. Specifically, if you have a one-dimensional wire with an inversion-breaking crystal structure (we can tell apart the diractions along the wire and in the wire, e.g. if there's more charge at the bottom than the top), there's an effect called *Rashba spin-orbit coupling*

$$H_R \coloneqq \alpha \mathbf{E} \cdot (\mathbf{P} \times \mathbf{S}) = \alpha p \sigma^y. \tag{5.6}$$

Adding this kind of term to the usual Hamiltonian

$$H = \left(\frac{p^2}{2m} - \mu\right) - B\sigma^z - \alpha p \sigma^y \tag{5.7}$$

shifts the bands: previously the bands were two upward-facing parabolas, one intersecting the x-axis and one above it. Now, both parabolas cross the x-axis, and there's a degenerate point where they intersect. Then, one adds a magnetic field, which pushes the bands away from that degenerate point, and it looks better.

So take your favorite semiconductor with spin-orbit coupling and put it on top of a superconductor, close enough that electrons (and specifically Cooper pairs) can tunnel through, causing a *proximity effect*. This adds a term

$$\Delta_{\text{eff}}(c_\uparrow^\dagger c_\downarrow^\dagger + \text{h.c.}) \tag{5.8}$$

to the Hamiltonian, where

$$\Delta_{\text{eff}} \approx \begin{cases} t^2/\Delta, & t \ll \Delta \\ \Delta, & t \gtrsim \Delta. \end{cases} \tag{5.9}$$

Then, all you have to do is turn on an external magnetic field.

Suppose you actually do this in the lab. How would you detect whether there are Majorana fermions at the end of the wire? One approach is that the presence of a Majorana fermion affects the qualitative properties of electron tunneling near the end of the wire. However, there are other ways to cause similar-looking effects: any real material has impurities, and this together with the external magnetic field can degrade the superconductor enough to produce low-energy states. These are bound pairs of Majoranas, and in particular they aren't topologically protected and are not what we're looking for. This approach has been investigated in experiments, e.g. Zhang et al (*Nature*, 2018) and Mourik et al (*Science*, 2012).

These experiments gave nice signatures of topological behavior, which is typical of topological phases. Then one has to argue further as to why these signatures don't come from impurities; for example, Majorana fermions produce signatures at very specific values.

In more detail, suppose we have a metallic lead and a tunneling link Γ to a wire hosting a Majorana zero mode. The experiment sends in an electron at very low energy from the lead to the wire. It will either bounce off the interface or transmit as a Cooper pair and emit backscattering. The Schrödinger equation allows us to compute the conductance:

$$H = \int_{-\infty}^{0} \mathrm{d}x\, \psi^\dagger(x)\left(\frac{p^2}{2m} - \mu\right)\psi(x) + i\Gamma\gamma(\psi(0) + \psi^\dagger(0)). \tag{5.10}$$

We're at energies too low for the electron to propagate into the superconductor, so what we have to compute is

$$\Psi_0 = \int_{x<0} e^{ik_F x}\psi(x) + ue^{-ik_F x}\psi(x) + ve^{ik_F x}\psi^\dagger(x). \tag{5.11}$$

The last term comes from a particle-hole symmetry consideration. We can simplify with $[H_{x<0}, \psi_0^\dagger] = 0$ and

$$i\Gamma\gamma\{(\psi(0) + \psi^\dagger(0)), \Psi_0\} = 0. \tag{5.12}$$

Additionally,

$$2eI = v_F \cdot \left(eV \frac{\partial n}{\partial \mu}\right)(2e), \tag{5.13}$$

where v_F is the *Fermi velocity* and

$$n = \int \frac{\mathrm{d}k}{2\pi} \Theta(-\varepsilon_k + \mu). \tag{5.14}$$

Differentiating gives us

$$\frac{\partial n}{\partial \mu} = \int \frac{\mathrm{d}k}{2\pi} \delta(\mu - \varepsilon_k) = \frac{1}{2\pi}\left(\left|\frac{\partial \varepsilon_k}{\partial k}\right|\right)^{-1} = \frac{1}{2\pi v_F}. \tag{5.15}$$

Hence the conductivity is $I = (2e^2/2\pi)(V)$ (if you pay attention to factors of $\hbar$, there's one in the denominator). One then can check that there are only two possible options, which correspond to the trivial and topological phases.

The first experiments didn't match this more detailed analysis, but by 2017 or so newer devices produced much better-looking curves.

One can see from these experiments that in the topological phase, $\frac{\mathrm{d}I}{\mathrm{d}V}$ does not depend very strongly on the strength of the tunneling barrier (there are a few wiggles, but not nearly as many as in the trivial phase). So while there's still a little bit of uncertainty, these figures look pretty good. The wiggles in the graph may have something to do with the fact that the Majoranas are actually fairly large relative to the length of the wire used in the experiment. There is still plenty of work before getting qubits out of this, and some more things to rule out (though, see https://arxiv.org/abs/1908.05549 for some progress).

Lecture 6.

Review of symmetries in quantum mechanics: 9/17/19

Andrew Potter isn't here today, so (TODO: didn't get his name since I was late) gave today's lecture. On Thursday, we'll begin discussing symmetry-protected topological phases, so today we'll review how symmetries manifest in quantum mechanics.

Suppose we have a phase diagram of some gapped physical system with two parameters, and suppose the system has some symmetry which is preserved by the first parameter λ_1 and is broken by the second parameter λ_2. Moving along λ_1 does not change the physical behavior of the system under any observations you can make which respect the symmetry — but for paths which vary λ_2, this is not true. But maybe the things you care about don't see that symmetry anyways, in which case this whole system is in the same phase. That is, the classification of phases of a system depends on what parameters and what symmetries you allow; more symmetries means "being in the same phase" is a stricter equivalence relation, which often means more phases. Studying this has led to a great deal of interesting research.

$\sim \cdot \sim$

Definition 6.1. A *symmetry* of a quantum system is a set G of operations on quantum states that leave physical measurements unchanged. We can combine any two operations g_1 and g_2 by first doing g_2, then g_1; call this new operation $g_1 g_2$. We ask for G to be a *group*, meaning it satisfies the following axioms.

(1) For any $g_1, g_2 \in G$, the product $g_1 g_2 \in G$.
(2) There is an element $e \in G$, called the *identity* or *unity*, such that $eg = ge = g$ for all $g \in G$. Physically, this is the operation that doesn't do anything.
(3) Every operation can be undone: for each $g \in G$ there's a $g^{-1} \in G$ such that $gg^{-1} = g^{-1}g = e$.

When one studies groups abstractly in mathematics, one must impose another axiom, called *associativity*, on triples of elements, but if your groups arise as sets of transformations on some system, this is already true.

Remark 6.2. It is *not* true that $g_1 g_2 = g_2 g_1$ in general! If that holds for all $g_1, g_2 \in G$, G is called *abelian*. ◂

Example 6.3. Let N be a positive integer. The group of *integers mod* N, denoted $\mathbb{Z}/N\mathbb{Z}$, $\mathbb{Z}/N$, or $\mathbb{Z}_N$,[7] is $\{0, 1, 2, \ldots, n-1\}$, where the group operation is defined to be addition mod n, i.e. given $p, q \in \mathbb{Z}/N$, take the remainder of dividing $p+q$ by N. For example, the hours of a clock form a $\mathbb{Z}/12$, unless you have a 24-hour clock, in which case it's $\mathbb{Z}/24$.

These groups are all abelian. ◂

Example 6.4. The *circle group* $\mathrm{U}(1)$[8] describes phases. Concretely, the underlying set is $[0, 1)$, and the group operation is addition mod 1, i.e. add two elements, then take the non-integer part. This group is also abelian. ◂

Example 6.5. The *unitary group* $\mathrm{U}(N)$ (also U_N) is the group of unitary $N \times N$ matrices, and the group operation is matrix multiplication. Similarly, the *special unitary group* $\mathrm{SU}(N)$ (or SU_N, etc.) is the subset of $\mathrm{U}(N)$ of matrices with determinant 1 — since this is preserved under matrix multiplication, this is indeed a group.

Similarly, the *orthogonal group* $\mathrm{O}(N)$ denotes the group of $N \times N$ orthogonal matrices (with real coefficients), and $\mathrm{SO}(N)$, the *special orthogonal group*, denotes the subgroup with determinant equal to 1. ◂

We now have two different things called U_1; they had better be the same. And indeed, a 1×1 unitary matrix is a unit complex number, and these all can be described as $\exp(2\pi i\theta)$ for exactly one $\theta \in [0, 1)$; furthermore, the identity, multiplication, and inverses match. So this is not a problem.

We care about groups because they do stuff. Wigner's theorem states that a symmetry operation on a quantum system can be represented by linear operations that are unitary or antiunitary. See Weinberg's book for a proof. Unitary means $U^\dagger U = 1$, and linear means that

$$U(\alpha\psi_1 + \beta\psi_2) = \alpha \cdot U(\psi_1) + \beta \cdot U(\psi_2). \tag{6.6}$$

So $\langle \phi \mid \psi \rangle = \langle U\phi \mid U\psi \rangle$. An *antiunitary* operator A satisfies $\langle \phi \mid \psi \rangle^* = \langle A\phi \mid A\psi \rangle$ and *antilinearity*:

$$A(\alpha\psi_1 + \beta\psi_2) = \alpha^* \cdot A(\psi_1) + \beta^* \cdot A(\psi_2). \tag{6.7}$$

Time-reversal symmetries are examples of antiunitary symmetries.

If your Hilbert space is N-dimensional, U_N therefore is important, since it contains the unitary symmetry operations for the system.

Definition 6.8. A *unitary representation* of a group G is a map $\rho\colon G \to \mathrm{U}_N$ such that $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$. In particular, $\rho(e) = e$.

We will also need a more general notion.

Definition 6.9. A *projective unitary representation* of a group G is data of two maps, $\rho\colon G \to \mathrm{U}_N$ and $\phi\colon G \times G \to \mathrm{U}_1$ such that $\rho(g_1)\rho(g_2) = \phi(g_1, g_2)\rho(g_1 g_2)$.[9]

The study of group representations is a huge subject of mathematics; in particular, for the groups we care about in physics, the classifications of their representations are all known.

One interesting example is that there are interesting projective representations of SO_N which do not come from actual representations. These correspond to the notion of "half-integer spin" in physics. TODO: didn't follow the notation. In any case, there's a group structure on the set of isomorphism classes of projective representations that aren't actual representations; for $N = 3$ that group is $\mathbb{Z}/2$.

Lecture 7.

Symmetry-protected topological phases: 9/19/19

There are some situations in condensed-matter physics where we have a two-parameter system, where varying one of the parameters preserves some symmetry and varying the other breaks it. For example, we could have a time-reversal-symmetric system where we're able to introduce a magnetic field, which breaks the symmetry.

[7] The names $\mathbb{Z}/N\mathbb{Z}$ and $\mathbb{Z}/N$ are more common in mathematics and $\mathbb{Z}_N$ is more common in physics. This is because there's a different object called $\mathbb{Z}_N$ in number theory.

[8] This group goes by several other names, including U_1, $\mathbb{T}$, and S^1, in different fields in mathematics.

[9] TODO: we may need to impose a condition on ϕ; I don't remember it offhand and missed it during the lecture.

Here something curious can happen: as long as we only make deformations which preserve the symmetry, we can find phases which are different from the trivial phase, in that if you deform it to the trivial phase, there's necessarily a phase transition closing the gap; but if you allow the symmetry to be broken, there is a gap-preserving deformation. Such phases are called *symmetry-protected topological (SPT)* phases.

Here are some examples of commonly-considered symmetries in the study of SPTs.

- Time-reversal symmetry is an antiunitary symmetry squaring to 1, and is customarily denoted $\mathbb{Z}_2^T$. Often this will act on the system by complex conjugation. The generator of this symmetry is denoted $\mathcal{T}$.
- Some systems preserve particle number. This is a U_1 symmetry, and the generator is $\exp(-i\theta\widehat{N})$.
- Some systems have a rotation symmetry acting on the spins. This is SO_3 for integer spins, but SO_3 only acts projectively for half-integer phases, so we have to pass to the double cover $\mathrm{Spin}_3 = \mathrm{SU}_2$. The

Example 7.1 (The AKLT phase)**.** Affleck-Kennedy-Lieb-Tasaki [AKLT88] wrote down the first example of an SPT, before the idea of an SPT was ever coined, for the symmetries $\mathbb{Z}_2^T$ and SO_3. Consider a wire with L vertices as we did for the Majorana chain, and suppose that at each site, there are three options for the spins. That is, the local Hilbert space is $\mathbb{C}^3$, and the total Hilbert space is the tensor product of the local Hilbert spaces over all sites. Introduce the Pauli-like operators

$$S^z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad S^x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad S^y = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{7.2}$$

and let $\mathbf{S} = (S^z, S^y, S^x)$. Haldane [Hal83a, Hal83b, Hal85] famously studied the Hamiltonian

$$H = J\sum_{i=1}^{L} \mathbf{S}_i \cdot \mathbf{S}_{i+1}, \tag{7.3}$$

and conjectured this is gapless for integer spins and gapped for half-integer spins. Affleck-Kennedy-Lieb-Tasaki perturbed this Hamiltonian to

$$H = J\sum_{i=1}^{L} \mathbf{S}_i \cdot \mathbf{S}_{i+1} + K\sum_{i=1}^{L} (\mathbf{S}_i \cdot \mathbf{S}_{i+1})^2. \tag{7.4}$$

The first term is generally called a *Heisenberg interaction*, and the second is called a *biquadratic interaction*. This system has an SO_3 rotation symmetry (if the spin is an integer) and a $\mathbb{Z}_2^T$ time-reversal symmetry, which is complex conjugation. The terms in (7.4) don't commute, but because this Hamiltonian satisfies a weaker condition of being *frustration-free*, we can write down a nice ground state, which for judiciously chosen J and K is an SPT.

Let's suppose $J = \cos\theta$ and $K = \sin\theta$; then we can plot the phase diagram in θ.

- For θ near $-\pi$, the system is a dimer: nearby neighbors pair up.
- For θ in the second quadrant and part of the third quadrant, the system is a ferromagnet: nearby spins tend to point in the same direction.
- For θ a bit less than π, the system is critical.
- For the remaining θ, in particular at the *AKLT point* $K = J/3$, the system is an SPT!

We restrict to the AKLT point because it's easier to analyze.

TODO: I missed something here.

In order to analyze the SO_3-symmetry, let's assume the system has spin 1, so that the $\mathbb{C}^3$ at each site is the defining representation $\mathbf{1}$ of SO_3. Therefore the interactions between sites i and $i+1$ are in the representation

$$\mathbf{1} \otimes \mathbf{1} \cong \mathbf{0} \oplus \mathbf{1} \oplus \mathbf{2}. \tag{7.5}$$

Let $\Pi_{j,j+1}^{(i)}$ denote the orthogonal projection $\mathbf{1} \otimes \mathbf{1} \twoheadrightarrow \mathbf{i}$, for $i \in \{0,1,2\}$; this map is SO_3-equivariant.

At the AKLT point, $K = J/3$, which means that in $\mathbf{0}$, the total energy is $-2J/3$, in $\mathbf{1}$, we get $-2J/3$, and in $\mathbf{2}$, we get $4J/3$. Therefore

$$(7.4) = 2J\sum_j \Pi_{j,j+1}^{(2)} + (\text{constants}). \tag{7.6}$$

There is a projection P_1 which sends a spin-1/2 state to a spin-1 state. (TODO: I don't understand this.) The ground state at the AKLT point is P_1 applied to the *singlet state*

$$|\psi_{\text{sing}}\rangle = \bigotimes_i \frac{1}{\sqrt{2}}(\uparrow_{B,i}\downarrow_{A,i+1} - \downarrow_{B,i}\uparrow_{A,i+1}). \tag{7.7}$$

The singlet state pairs up neighbors, and therefore there are unpaired edge states, though they're not as exotic as the ones we found in the Majorana chain.

We should check that $P_1|\psi_{\text{sing}}\rangle$ is actually a ground state. One can check that the eigenvalues of the Hamiltonian are nonnegative. Then, each term in (7.4), reformulated as in (7.6), kills $P_1|\psi_{\text{sing}}\rangle$:

$$\Pi^{(2)}_{j,j+1}P_1|\psi_{\text{sing}}\rangle = P_1\widetilde{\Pi}^{(2)}_{j,j+1}|\psi_{\text{sing}}\rangle = 0. \tag{7.8}$$

If you introduce a magnetic field term $-\mathbf{B}\cdot\sum \mathbf{S}_i$ to the Hamiltonian, you can kill these edge states, but this breaks both the SO_3 and the $\mathbb{Z}_2^T$ symmetries. Thus this system is an SPT for either SO_3 or $\mathbb{Z}_2^T$ separately; in fact, we can pick smaller subgroups of SO_3, including anything for which $-\mathbf{B}\cdot\boldsymbol{\sigma}$ isn't symmetric. For example, we could pick the $\mathbb{Z}/2\times\mathbb{Z}/2$ subgroup of rotations by π around two different axes.

If we consider the system on an interval, we have four ground states, corresponding to what the two spins on the edges are doing: each can be up or down. We therefore get spin projector operators on the edges, e.g. S^z_L, which is the identity in the bulk and on the right-hand side, but which acts by σ^x on the left-hand side; we can analogously see the other Pauli matrices and the operators on the right-hand side.

Here another strange thing happens — the macroscopic degrees of freedom in the bulk are all spin-1, but these edge operators obey the rules of a spin-1/2 system! This is an example of fractionalization. Symmetry is crucial: we can only see spin using the SO_3 (or SU_2) symmetry. ◂

In the rest of today's lecture, we'll discuss another system in the same phase whose analysis doesn't require introducing fictitious degrees of freedom. Next time, we'll discuss the entanglement properties of the AKLT phase: what happens if you cut the system somewhere? This will allow us to conclude that the system is in an SPT phase by just looking at the bulk, leading eventually to the matrix product state representation of SPTs and a general 1D classification.

Example 7.9 (Cluster state)**.** This state is easier to analyze than the AKLT system, though as $\mathbb{Z}/2\times\mathbb{Z}/2$-symmetry protected phases, they're in the same phase. It also appears in quantum information, where certain properties make it easier to reason about qubits.

The Hilbert space is again a tensor product of local Hilbert spaces, each of which is four-dimensional, with basis $|0\rangle,\dots,|3\rangle$. Let g_A generate the first $\mathbb{Z}/2$ and g_B generate the second $\mathbb{Z}/2$. We ask for $\mathbb{Z}/2\times\mathbb{Z}/2$ to act on the system as follows:

- g_A fixes $|0\rangle$ and $|1\rangle$ and flips the signs of $|2\rangle$ and $|3\rangle$.
- g_B fixes $|0\rangle$ and $|2\rangle$ and flips the signs of $|1\rangle$ and $|3\rangle$.[10]

Write $X_{a,i}$ for g_A acting at site i, and $X_{B,i}$ for g_B acting at site i. Let $Z_{A,i}$ and $Z_{B,i}$ switch the basis vectors at site i (TODO: I didn't get which ones specifically).[11] Then the Hamiltonian is

$$H = -J\sum_i (Z_{A,i}X_{B,i}Z_{A,i+1} + Z_{B,i}X_{A,i+1}Z_{B,i+1}). \tag{7.10}$$

So this looks like a ZXZ chain, though we've formulated it somewhat differently (e.g. this isn't exactly a spin system).

This is a *stabilizer Hamiltonian*, which means

- each ZXZ term in the Hamiltonian squares to 1, so its eigenvalues are in $\{\pm 1\}$;
- all pairs of these terms commute; and
- no term is a product of other terms.

We can call each such term S a *stabilizer*. If you specify an eigenvalue for each stabilizer separately, you've determined a unique state, and therefore we can analyze the Hamiltonian and its ground states without writing down the entire wavefunction.

[10] If σ denotes the sign representation of $\mathbb{Z}/2$, this tells us our local Hilbert space is $(\mathbf{1}\oplus\sigma_A)\otimes(\mathbf{1}\oplus\sigma_B)$.

[11] I think under the description of the local Hilbert space as a tensor product, X_A and Z_A are the usual Pauli matrices acting on the first tensor factor, and X_B and Z_B act on the second tensor factor.

Often people write down the projectors $\Pi_S \coloneqq (1+S)/2$, which projects into the subspace where S has eigenvalue 1.

By inspection, we can write down some operators which commute with the Hamiltonian.

- Let $\Sigma^1 \coloneqq X_{A,1}Z_{A,1}$; this clearly commutes with distant terms, and you can check that it commutes with the nearby ZXZ terms.
- In the same way, you can define $\Sigma^2 \coloneqq Y_{A,1}Z_{B,1}$ and $\Sigma^3 \coloneqq Z_{A,1}$ and observe that they commute with the Hamiltonian.

So, just like in the AKLT chain, these generate a spin-1/2 at the boundary: $[\Sigma^i, \Sigma^j] = 2i\epsilon^{ij}_k\Sigma^k$.

We should next check how the $\mathbb{Z}/2 \times \mathbb{Z}/2$ symmetry acts on $\mathbf{\Sigma}$.

$$g_A\mathbf{\Sigma}g_A = X_{A,1}\begin{pmatrix} X_AZ_B \\ Y_AZ_B \\ Z_A \end{pmatrix} X_{A,1} = \begin{pmatrix} \Sigma^1 \\ -\Sigma^2 \\ -\Sigma^3 \end{pmatrix} = \Sigma^1\mathbf{\Sigma}\Sigma^1. \tag{7.11a}$$

Similarly,

$$g_B\mathbf{\Sigma}g_B = \begin{pmatrix} -\Sigma^1 \\ -\Sigma^2 \\ \Sigma^3 \end{pmatrix} = \Sigma^3\mathbf{\Sigma}\Sigma^3. \tag{7.11b}$$

TODO: g_A and g_B commute, but Σ^1 and Σ^3 anticommute. What happened?

TODO: also there was some other stuff which I didn't follow. ◂

Lecture 8.

Entanglement: 9/24/19

Let's say we're studying some quantum system on a wire. We can imagine cutting the wire into two pieces A and A^c, as in Figure 1.

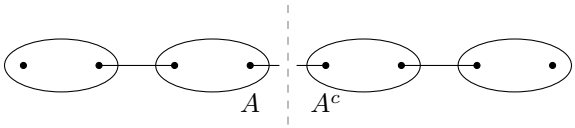


FIGURE 1. Splitting a spin chain into two subsystems which might be entangled.

Since the Hilbert space is a tensor product of local Hilbert spaces, it factorizes:[12]

$$\mathcal{H} = \bigotimes_{i=1}^{L}\mathcal{H}_i = \underbrace{\left(\bigotimes_{i\in A}\mathcal{H}_i\right)}_{=:\mathcal{H}_A} \otimes \underbrace{\left(\bigotimes_{i\in A^c}\mathcal{H}_i\right)}_{=:\mathcal{H}_{A^c}}. \tag{8.1}$$

Hence, if the system is in a state $|\psi\rangle$, we can ask whether $|\psi\rangle$ factors as a product state

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_{A^c}\rangle. \tag{8.2}$$

If so, we call $|\psi\rangle$ *entangled*; if not, it's *unentangled*.

If you choose bases $\{|i_A\rangle\}$ for $\mathcal{H}_A$ and $\{|j_{A^c}\rangle\}$ for $\mathcal{H}_{A^c}$, then you can check whether

$$|\psi\rangle = \sum_{i,j}\lambda_{ij}|i_A\rangle \otimes |j_{A^c}\rangle, \tag{8.3}$$

in which case $|\psi\rangle$ is entangled, but whether you can do this is basis-dependent, which is annoying; there's a luck component of choosing a good basis. Let's try something smarter.

[12] Well, this is true for bosonic systems. For fermions, this must be modified slightly to take the nonlocal effects of particle statistics into account, which amounts to using the $\mathbb{Z}/2$-graded tensor product.

Remark 8.4. TODO: Be especially aware of mistakes in the following, coming from my confusion over the notation: at each site, $\dim(\mathcal{H}_i)$ isn't always 1, so straightening out what exactly we're indexing over confused me thoroughly during class. It shouldn't be too hard, but I wasn't able to figure it out in time. Sorry about that.

I also missed some stuff (e.g. some stuff about the entanglement entropy). ◂

Let's choose a basis s_i for each site i in the chain. Then we can write $|\psi\rangle$ as

$$|\psi\rangle = \sum_{s_1,\dots,s_L} \psi^{s_1,\dots,s_L} \tag{8.5a}$$

$$= \sum_{\substack{s_i : i\in A \\ s_j : j\in A^c}} \psi^{(s_i)_{i\in A},(s_j)_{j\in A^c}} |(s_i)_{i\in A}\rangle \otimes |(s_j)_{j\in A^c}\rangle. \tag{8.5b}$$

Think of $\psi^{(s_i)_{i\in A},(s_j)_{j\in A^c}}$ as an $|A| \times |A^c|$ matrix; if $|A| = |A^c|$, we could diagonalize and use that to check whether $|\psi\rangle$ is a product state. But in general, this is not a square matrix, so we can't diagonalize it, but there is a "next best thing," called the *singular value decomposition*. In general, an $m \times n$ matrix M factors as $M = U\Lambda V$, where

- U is an $n \times n$ unitary matrix; specifically, it's the change-of-basis matrix which diagonalizes the $n \times n$ matrix $MM^\dagger$.
- Similarly, V is an $m \times M$ unitary matrix, which is the change-of-basis matrix diagonalizing the $m \times m$ matrix $M^\dagger M$.
- Λ is a "diagonal" $m \times n$ matrix, in that its only nonzero values are the diagonal entries $\lambda_1, \dots, \lambda_n$. These are called the *singular values* of M.

This is unique in the same way as diagonalization is, and in particular the set of singular values is an invariant of M. Applying this to (8.5b),

$$|\lambda\rangle = (8.5b) = \sum \lambda_\gamma U^{\alpha\gamma} |(s_i)_{i\in A}\rangle V^{\gamma\beta} |(s_j)_{j\in A^c}\rangle. \tag{8.6}$$

Here α indexes the basis elements $|(s_i)_{i\in A}\rangle$ and β indexes the analogue for A^c.

Example 8.7. Let's consider a chain of length 2, with a qubit (i.e. a two-dimensional Hilbert space) at each site, and we'll split in the only nontrivial way. Let

$$\psi := \frac{|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{2}, \tag{8.8}$$

which is a product state. For $\beta, \alpha \in \{\uparrow, \downarrow\}$, we have

$$(\psi^{\alpha,\beta}) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{8.9}$$

so computing the SVD is the same thing as diagonalizing $(\psi^{\alpha,\beta})$:

$$(\psi^{\alpha,\beta}) = \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\Lambda} \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{V}. \tag{8.10}$$

In general, we order the singular values by $|\lambda_1| \geq |\lambda_2| \geq \dots$, so $\lambda_1 = 1$ and $\lambda_2 = 0$. ◂

In general, $|\psi\rangle$ is in a product state iff $|\lambda_1| = 1$ and $|\lambda_i| = 0$ for $i > 1$. Taking the SVD in this context is called the *Schmidt decomposition*; the two states are called *Schmidt states* and the singular values are called *Schmidt weights*.

There's something (TODOnot sure what) probabilistic going on here, and when you have a random variable, it's often useful to study its entropy.

Definition 8.11. The *entanglement entropy* of $|\psi\rangle$, A, and A^c is

$$S_A(|\psi\rangle) = -\sum_i |\lambda_i|^2 \log|\lambda_i|^2 = -\operatorname{tr}(\rho_A \log \rho_A), \tag{8.12}$$

where ρ_A is the partial trace.

Hence, $S_A(|\psi\rangle) = 0$ iff ψ is unentangled. The entanglement entropy satisfies

$$S_A(|\psi_1\rangle \otimes |\psi_2\rangle) = S_A(|\psi_1\rangle) + S_A(|\psi_2\rangle) \tag{8.13a}$$

and is *subadditive*, in that

$$S_{A\cup B} \leq S_A + S_B \tag{8.13b}$$

for any choice of $|\psi\rangle$.

Entanglement entropy tells you a lot about your system. For example, the "typical state" (i.e. an expectation value under some sort of Haar random measure) has entanglement entropy

$$S_A(|\psi_{\text{typical}}\rangle) = |A| \log|\mathcal{H}_i| - (\text{constant}) + \dots, \tag{8.14}$$

where the unwritten terms tend to 0 as $|A| \to \infty$. However, the ground states of a gapped Hamiltonian behave differently:

$$S_A(|\psi_{\text{GS}}\rangle) = (\text{constant}) + \dots. \tag{8.15}$$

In higher dimensions, it's conjectured that this constant is proportional to $|\partial A|$.

Example 8.16. In a 1D conformal field theory, $S_A = (c/3)\log|A| + \dots$, where c is the central charge. For CFTs in higher dimensions, we have terms $S_A = c_1|\partial A| + c_2 \log|A| + \dots$.

In a Fermi liquid, $S_A = (k_F L_A)^{d-1} \log(L_A)$. ◂

As is so often true in physics, it's helpful to think of this as a statistical-mechanics system: the term $p_i := |\lambda_i|^2$ can be valued in $[0, 1]$, or alternatively we can look at $\varepsilon_i := -\log p_i$. The *entanglement spectrum* is the plot of ε_i against i; one of the hallmarks of an SPT, which we will discover using the matrix product states formalism, is a highly nontrivial entanglement spectrum.

That means it's about time to discuss matrix product states. First, let's cut just the first site from the remaining $L - 1$ sites, giving a matrix Λ_1 of singular values. Then iterate with the next site, giving another matrix Λ_2, and so on, leading to a factorization

$$\psi^{s_1 \cdots s_N} = \sum \lambda_{1,\alpha} \psi^{s_1}_{A,\alpha} \left(\psi^{s_2 \cdots s_N}_{A^c,\alpha} \right) \tag{8.17}$$

$$= \sum_{\alpha_1} \lambda_{1\alpha_1} \psi^{s_1}_{A\alpha_1} \sum_{\alpha_2} \lambda_{2\alpha_2} \psi^{s_2}_{A\alpha_1\alpha_2} \psi^{s_3 \cdots s_N}_{\alpha_1,\alpha_2}. \tag{8.18}$$

If we iterate and let $M^{s_i}_{\alpha_{i-1}\alpha_i} := \lambda_{i\alpha_i} \psi^{s_i}_{A\alpha_{i-1}\alpha_i}$, then this rewrites $|\psi\rangle$ as a product of these matrices:

$$= M^{s_1}_{\alpha_1} M^{s_2}_{\alpha_1\alpha_2} M^{s_3}_{\alpha_2\alpha_3} \cdots M^{s_{L-1}}_{\alpha_{L-2,L-1}} M^{s_L}_{\alpha_{L_1}}. \tag{8.19}$$

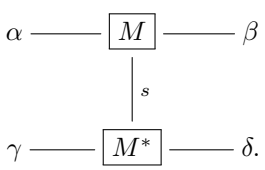
Hence the name "matrix product state."

Remark 8.20. If you want to actually simulate quantum systems, this looks computationally too big to really work with, but it's actually tractable, and much more efficient than *a priori* descriptions of ground states of gapped Hamiltonian systems, and nothing like it works for arbitrary states. ◂

You can think of this as a system of contractions across a trivalent tensor network. There are multiple ways to do this, and they're all necessarily equivalent.

For example, let's consider a simple contraction... TODO

(8.21)



This corresponds to $\sum_s M^s_{\alpha\beta} M^{*s}_{\gamma\delta} = \dots$

Lecture 9.

Matrix product states: 10/1/19

At the end of the last lecture, we described the ground state of a topological phase by cutting at every junction and describing the state as a bunch of contractions of matrices. This admits a diagrammatic representation as a *tensor network*, a graph where the vertices are tensors and the edges, labeled by indices, represent contracting the specified index. Thus a tensor network represents a computation of a number. You can contract the indices in different orders, and the answer doesn't change, but the computational complexity can change drastically. In fact, for networks representing d-dimensional systems for $d \geq 2$, it's NP-hard to determine the optimum order for evaluation!

In dimension 1, things are a little easier. In almost all experiments, what we compute is a correlation function $\langle \psi \mid \mathcal{O}_1 \mathcal{O}_2 \mid \psi \rangle$, and we can represent this as a "ladder," an $N \times 2$ tensor network with $\mathcal{O}_1$ and $\mathcal{O}_2$ inserted somewhere in the middle. Here, we know what the optimum evaluation order is: first contract each leg of the ladder, then go along the line. This is polynomial in the size of the indices and linear in N.

The key operation is the transfer matrix (8.21), representing

$$T_{\alpha\gamma,\beta\delta} = \sum_s M^s_{\alpha\beta} M^*_{\gamma\delta}. \tag{9.1}$$

Diagonalizing the transfer matrix tells us that

$$T^M = \sum_{j=1}^{\chi^2} \tau_j^M V_j^\dagger V_j, \tag{9.2}$$

where $\tau_1 = 1 > \tau_2 \geq \tau_3 \geq \dots$. The fact that there is a unique largest eigenvalue is a linear-algebraic consequence of the fact that there's a unique ground state.

Suppose the insertions of $\mathcal{O}_1$ and $\mathcal{O}_2$ are separated by R steps in the ladder, where R is large. Then most of the terms in (9.2) are very small, and we conclude that

$$\langle \psi \mid \mathcal{O}_1 \mathcal{O}_2 \mid \psi \rangle - \langle \mathcal{O}_1 \rangle \langle \mathcal{O}_2 \rangle \sim e^{-R/\xi}, \tag{9.3}$$

where $\xi \sim \tau_2$, the largest remaining component.

Thus matrix product states have exponentially decaying correlation functions and limited entanglement, which is exactly what we expect from the ground states of gapped Hamiltonians. In fact, in $d = 1$, there is a one-to-one correspondence: given a gapped Hamiltonian, there's a matrix product state representation of its ground state, and given a matrix product state one can reconstruct a gapped Hamiltonian. Since matrix product states are computationally efficient to work with, this is very fortunate.

Example 9.4. Consider a trivial state where the spins are alternating. We can represent this with a $\chi = 1$ matrix product state, where $M^s = \delta_{s,\uparrow}$ for i even and $M^s = \delta_{s,\downarrow}$ for i odd. ◂

Example 9.5. For a more interesting example, consider the GHZ state

$$\frac{|\uparrow\uparrow\uparrow\cdots\uparrow\rangle + |\downarrow\downarrow\downarrow\cdots\downarrow\rangle}{\sqrt{2}}. \tag{9.6}$$

This can be represented as a $\chi = 2$ matrix product state, where

$$M^\uparrow = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad M^\downarrow = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{9.7a}$$

and

$$V_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad V_R = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{9.7b}$$

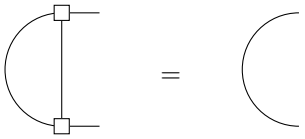
◂

In the exercises, you'll construct a matrix product state representation for the ground state of the AKLT chain (Example 7.1).

One can run a kind of renormalization on the tensor network where we take a few legs, say 3, and regard them as one larger node. Then we obtain a new transfer matrix $\widetilde{T} \coloneqq T^3$. If we keep doing this, in the limit the eigenvalues τ_i for $i > 1$ will die off, and we'll just get the projector $T = V_1^\dagger V_1$.

Remark 9.8. The matrix product state representation of a ground state is not unique: for example, we can insert W^{-1}, then W, into the tensor network. This is sometimes called a *gauge redundancy*. One can cut down on this redundancy by asking for the chain to be in *(left) canonical form*, meaning...

(9.9)



This guarantees that the ground state is normalized correctly, even in infinite chains, and simplifies some other computations (TODO: I missed some of what came afterwards.) ◂

Since matrix product states describe ground states of gapped systems, we can use them to study SPTs. First, though, we need to understand how symmetries of a system are implemented for matrix product states. This will lead us to a complete understanding of 1d SPTs.

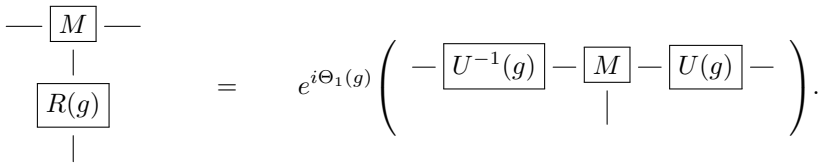
Let S be a symmetry group of a system, acting via a representation R on the Hilbert space. For now, we assume R is unitary (so no time-reversal symmetry), and that it represents an *internal symmetry*: R factors as a tensor product over all sites. Geometrically, the symmetry is just doing something independently to each site, rather than something like translation or reflection which exchanges sites. Spin rotation is a good example of such an R.

We want to describe the ground state $|\psi\rangle$ as a matrix product state, i.e. $R(g)|\psi\rangle = e^{i\Theta(g)}|\psi\rangle$. What does this mean for the matrices in the matrix product state description of $|\psi\rangle$? One clear sufficient condition is that

$$R_{ss'}(g)M^{s'}_{\alpha,\beta} = e^{i\Theta_1(g)}M^{s}_{\alpha,\beta}, \tag{9.10}$$

but this turns out to be too strong: one can write down symmetric ground states which do not have this property. This is related to the gauge redundancy that we discussed above: we can ask for

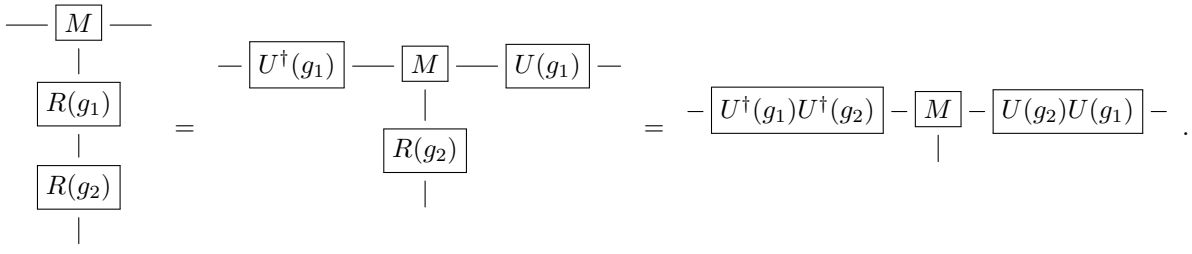
(9.11)



These compute the same ground state, as we described above, and one can show that this is both a necessary and a sufficient condition.

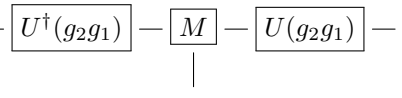
Now, what happens when we compose two elements of the symmetry group? We know $R(g_2)R(g_1) = R(g_2g_1)$, and can also compute what happens to the bond degrees of freedom, by first pulling $R(g_1)$ back into the bond space à la (9.11), then pulling back $R(g_2)$:

(9.12)



We can compare this to what happened if we combined $R(g_1)$ and $R(g_2)$ before pulling back into the bond space, in which case we obtain

(9.13)



It would be nice to say these are equal, but this is not always true: it could be the case that

$$U(g_2g_1) = \omega(g_1, g_2)U(g_2)U(g_1) \tag{9.14}$$

for some function $\omega\colon S \times S \to \mathrm{U}_1$.

Example 9.15. In the AKLT state, $S = \mathbb{Z}/2 \times \mathbb{Z}/2$. Let a denote the generator of the first copy of $\mathbb{Z}/2$ and b denote the generator of the second. Then $U(a)U(b) = -U(b)U(a)$, indicating that $\omega(a,b)\omega(b,a) = -1$, so ω is nontrivial. ◂

So we get a projective representation, which is reasonable. This can also explain something interesting: in a closed system, we get a unique ground state, but in a system with a boundary, there must be multiple ground states, which you can argue using the linear algebra of matrix product states. This is ultimately because we must choose boundary conditions, and in general have multiple choices. We say this ground state degeneracy in both the Kitaev chain and the AKLT chains: in both cases, there's a unique ground state on the circle, but multiple on the interval.

This tensor product decomposition assumes that we're in a bosonic system, because locality is a little different. But the upshot is that to understand and classify 1d bosonic SPT phases for an internal symmetry group G, we should classify projective representations of the symmetry group.[13]

We arrived at projective representations through $\omega\colon S \times S \to \mathrm{U}_1$, but asking for ω to describe a projective representation puts both some conditions on ω and implies some redundancy. First, the condition: using associativity of multiplication in g,

$$\omega(g_1,g_2)\omega(g_1g_2,g_3)U(g_1g_2g_3) = \omega(g_1,g_2g_3)\omega(g_2,g_3)U(g_1g_2g_3). \tag{9.16}$$

Since $U(g_1g_2g_3)$ is invertible, we conclude that

$$\frac{\omega(g_1,g_2)\omega(g_1g_2,g_3)}{\omega(g_1,g_2g_3)\omega(g_2,g_3)} = 1. \tag{9.17}$$

This is called the 2-*cocycle condition*.

The redundancy arises as follows: given a phase $\beta \in \mathrm{U}_1$, we can conjugate $\mathrm{U}(g)$ by β and obtain what should be the same projective representation. Under this transformation,

$$\omega(g_1,g_2) \longmapsto \frac{\beta(g_1)\beta(g_2)}{\beta(g_2g_1)}\omega(g_1,g_2). \tag{9.18}$$

Therefore the group of isomorphism classes of projective representations into U_1 is the group of maps $\omega\colon S \times S \to \mathrm{U}_1$ satisfying the 2-cocycle condition, modulo the relation (9.18). This group is denoted $H^2(S;\mathrm{U}_1)$, the *second group cohomology of S with coefficients in* U_1. Here, "second" has to do with the fact that ω takes in two copies of S.

The idea is that $H^2(S;\mathrm{U}_1)$ classifies 1d bosonic SPTs with internal symmetry S: such an SPT has edge modes transforming in the given projective representation.

Lecture 10.

The field-theoretic perspective: 10/3/19

In condensed-matter systems, we can compute partition functions by a sort of path integral:

$$U_{fi} = \langle f \mid e^{-iHt} \mid i\rangle = \sum_{i\to f} e^{iS_{cl}(\text{path})}, \tag{10.1}$$

where the sum is over all classical paths $i \to f$, and S_{cl} is the classical action, $\int_0^\beta \mathrm{d}\tau\, L$, where L is the Lagrangian.

We're not always so interested in time evolution in condensed-matter physics, but if we rotate to imaginary time, this tells you information about thermal behavior:

$$H \longrightarrow H + \frac{1}{\beta}\int \mathrm{d}x\, \lambda(x)\phi(x). \tag{10.2}$$

This is a sort of generating function: if

$$Z_\lambda \coloneqq \mathrm{tr}\Big(e^{-\beta H - \int \lambda\phi}\Big), \tag{10.3}$$

[13] This is really a 1d story. One can generalize this to higher dimensions, where there's a generalization of the notion of a projective representation, but even then this misses some SPT phases. Still, it's an interesting story.

then, e.g.,

$$\frac{\partial^2 \ln Z_\lambda}{\partial\lambda(x_1)\partial\lambda(x_2)}\bigg|_{\lambda\to 0} = \frac{\partial}{\partial\lambda(x_1)}\left(\frac{\operatorname{tr}\left(-\phi(x_2)\exp\left(-\beta H - \int\lambda\phi\right)\right)}{\operatorname{tr}\exp\left(-\beta H - \int\lambda\phi\right)}\right) \tag{10.4}$$

$$= \frac{\operatorname{tr}\phi(x_1)\phi(x_2)e^{-\beta H}}{\operatorname{tr}e^{-\beta H}} - \frac{\operatorname{tr}\phi(x_1)e^{-\beta H}\operatorname{tr}\phi(x_2)e^{-\beta H}}{(\operatorname{tr}e^{-\beta H})^2} \tag{10.5}$$

$$= \langle\phi(x_1)\phi(x_2)\rangle - \langle\phi(x_1)\rangle\langle\phi(x_2)\rangle, \tag{10.6}$$

which is telling you thermal information.

Today, we'll study a very simple example from this field-theoretic perspective, which will be surprisingly involved and technical. After that we'll apply this formalism to the AKLT spin chain.

We begin with the path integral for spins. There are various ways to derive the answer, but going into them would take us a little far off course, so we'll just write down the answer for now and see that it works.

The first question is, what do we do with open paths in the path integral? Defining the Berry phase doesn't work terribly well here, and we can't just write down a Lagrangian, because what Lagrangian defines a Berry phase?

Anyways, if we write $\mathbf{S} = s\mathbf{n}$, where s is the spin size and $|\mathbf{n}|^2 = 1$, we'll think of $|\mathbf{n}\rangle$, which is on the Bloch sphere, evolving by a phase, $|\mathbf{n}\rangle \mapsto e^{i\phi(\mathbf{n})}|\mathbf{n}\rangle$. Write $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Then the answer for the classical action is

$$S = s\int \mathrm{d}t\,(1-\cos\theta)\frac{\partial\phi}{\partial t}. \tag{10.7}$$

The derivative is a total derivative, so we don't have to worry about how fast a given path goes. A general path can be approximated by a bunch of small rectangles, so let's focus on a rectangular path given by four segments s_1, s_2, s_3, s_4. Then the contribution of, e.g., $s_1 + s_3$ is

$$s_1 + s_3 = s((1-\cos\theta_1)\,\Delta\Phi + (1-\cos\theta_3)(-\Delta\phi)) \tag{10.8}$$

$$= \Delta\left(-\frac{\partial}{\partial\theta}(1-\cos\theta)\,\Delta\theta\,\Delta\phi\right) \tag{10.9}$$

$$= s\sin\theta\,\Delta\theta\,\Delta\phi. \tag{10.10}$$

This works, but is awkward in that we had to choose coordinates for $\mathbf{n}$, and there's no global coordinate system that works. The solution, which manages to write this in a more invariant way, is due to Wess, Zumino, and Witten. We'll have to make a choice of a reference value $\mathbf{n}_0$, but any two choices will give the same answer.

Introduce a fictitious extra coordinate $u \in [0,1]$, which we'll think of as interpolating smoothly from some fixed configuration $\mathbf{n}(0,t) \coloneqq \mathbf{n}_0$ to $\mathbf{n}(1,t) \coloneqq \mathbf{n}(t)$, the configuration we're actually in. (So the two coordinates together form a ball instead of a sphere, with $(0,t)$ at the center.) Now the action looks like

$$S_{\mathrm{WZW}} = \frac{s}{2}\int_0^1 \mathrm{d}u\int_{t_i}^{t_f}\mathrm{d}t\,\mathbf{n}\cdot(\partial_\mu\mathbf{n}\times\partial_\nu\mathbf{n})\epsilon^{\mu\nu}, \tag{10.11}$$

where $\epsilon = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$.

Remark 10.12. Sometimes we also add a magnetic term $-\int sg\mathbf{B}\cdot\mathbf{n}|_{u=1}\,\mathrm{d}t$ to the action. ◂

Exercise 10.13. This is a bit of gory algebra, but check that if you introduce a specific coordimate system, letting $\theta(u,t) = u\theta(t)$ and keeping $\phi(t)$, then you get the coordinate-dependent action (10.7).

Now let's compute the variation of S_{WZW}.

$$\Delta S_{\mathrm{WZW}} = \frac{s}{2}\int_D \epsilon^{\mu\nu}\Big(\underbrace{3\delta n\cdot\partial_\mu n\times\partial_\nu n}_{(*)} + 2\partial_\mu(\mathbf{n}\cdot\delta\mathbf{n}\times\partial_\nu\mathbf{n})\Big). \tag{10.14}$$

A priori, there are additional terms, but they disappear because of the antisymmetrizer. I also missed another step in the computation; sorry about that.

The term $(*)$ vanishes, because δn cannot be linearly independent from the derivatives of n, and therefore we get only the boundary term. We're trying to collapse the u-direction, so we set $\mu = u$, forcing $\nu = t$. The result is

$$(10.14) = -s \oint \mathrm{d}t\, \delta n \cdot \left(\mathbf{n} \times \frac{\partial \mathbf{n}}{\partial t} \right). \tag{10.15}$$

So the mysterious extra coordinate u drops out as soon as we compute something physically relevant. If we added the magnetic term as in Remark 10.12, we get an extra factor of $\delta n g \mathbf{B}$.

Now let's study this a little more carefully. What if you and I picked different reference configurations $\mathbf{n}_0$ and $\mathbf{n}_0'$? Call the corresponding actions S_{WZW} and S'_{WZW}. Then

$$S_{\mathrm{WZW}} - S'_{\mathrm{WZW}} = \int_H L(\mathbf{n}) + \int_H L(\mathbf{n}'), \tag{10.16}$$

where H is a hemisphere and L is the Lagrangian density. We know the Lagrangians agree at $u = 1$, so we can glue these hemispheres together and conclude

$$= \int_{S^2} L(\mathbf{n}) = \frac{s}{2} \int_{S^2} n \cdot (\partial_\mu n \times \partial_\nu n) \epsilon^{\mu\nu}, \tag{10.17}$$

where n comes from the two choices of n and n' on the two hemispheres.

Now (10.17) isn't obviously zero, so let's look a little closer. The cross product is the oriented area of the parallelogram spanned by $\partial_\mu n$ and $\partial_\nu n$. Now, μ and ν are coordinates on the sphere, so (10.17) computes the surface area of the sphere, times the degree $N \in \mathbb{Z}$ of the map $S^2 \to S^2$ from the sphere with coordinates (u, t) to the Bloch sphere n.[14]

The key point is that the degree is a rigid invariant: if you wiggle the input data to (10.16), you get the same answer: the degree is a topological invariant of maps $S^2 \to S^2$, so it cannot change under small perturbations. Therefore the variation of the difference vanishes, which is good. It might seem weird that the actions differ by some number, but this is actually OK — we care about the partition function (and correlation functions) in order to actually compute physically meaningful information. This means we exponentiate the action: $Z \coloneqq e^{iS_{\mathrm{WZW}}}$. If Z' is the corresponding value for S'_{WZW},[15]

$$\frac{Z}{Z'} = e^{iS_{\mathrm{WZW}} - S_{\mathrm{WZW}'}} = e^{i4\pi s N}, \tag{10.18}$$

and since s is either an integer or a half-integer, this is always equal to 1.

When we care about imaginary time (to study thermal behavior), we replace t with $i\tau$, and replace $e^{iS_{\mathrm{WZW}}}$ with $e^{-S_{\mathrm{therm}}}$.

$$\sim \cdot \sim$$

Now let's study a spin chain. We'll have to tack on one of these WZW actions for every node in the chain. We look at a 1d spin chain with Hamiltonian $H = J \sum_i \mathbf{s}_i \cdot s_{i+1}$. If we perturb this, we get the AKLT phase as before; that's still true for this simpler Hamiltonian, at least for $J > 0$, which we assume. This Hamiltonian prefers states in which neighboring spins point in opposite directions. But this is the local description, and there might be a few blips in this globally.

The goal of quantum field theory is to cook out slowly fluctuating changes. So let's let $\mathbf{m}_j \coloneqq (-1)^j \mathbf{n}_j$: on even sites, it's positive on $\uparrow$ and negative on $\downarrow$, but on odd states the opposite is true. This is a better object than $\mathbf{n}_j$ to feed to field theory, because $\mathbf{n}_j$ changes at most sites, but changing $\mathbf{m}_j$ reflects a change in something physical in the system. So we rewrite the Hamiltonian as

$$H = -J \frac{s^2}{2} \sum_i \left((\mathbf{m}_i - \mathbf{m}_{i+1})^2 - \mathbf{m}_i^2 - \mathbf{m}_{i+1}^2 \right) \approx -J s^2 a \int \mathrm{d}x \left(\frac{\partial \mathbf{m}}{\partial \mathbf{x}} \right)^2, \tag{10.19}$$

[14] In differential topology, one can prove that a smooth map $S^n \to S^n$ is almost d-to-one for some integer d: for a set of full measure in the codomain, the preimage of a point is d points, counted with sign (meaning that oppositely oriented points cancel out, for example). This d is called the *degree* of the map.

[15] Well, really, we should sum over paths, but the point is that once we exponentiate, the difference disappears.

as the sum approximates an integral and the difference approximates a derivative. So we have

$$\sum_j S_{\mathrm{WZW}}(\mathbf{n}_j) = \sum_j (-1)^j S_{\mathrm{WZW}}(\mathbf{m}_j) \tag{10.20}$$

$$= \frac{s}{2}\int_D \sum_{j \text{ even}} (m_{j+1}\cdot \partial_\mu m_{j+1} \times \partial_\nu m_{j+1} - m_j \cdot \partial_\mu m_j \times \partial_\nu m_j) \tag{10.21}$$

$$= a\frac{s}{2}\sum_{j \text{ even}} \int \mathrm{d}\tau \left(\frac{\partial \mathbf{m}}{\partial x}\right)\cdot\left(\mathbf{m}\times\frac{\partial \mathbf{m}}{\partial \tau}\right) \tag{10.22}$$

$$= \frac{s}{4}\int \mathrm{d}x\,\mathrm{d}\tau\, \epsilon_{\mu\nu}\mathbf{m}\cdot(\partial_\mu \mathbf{m}\times\partial_\nu \mathbf{m}), \tag{10.23}$$

where $\mu,\nu \in \{x,t\}$. Thus the partition function is

$$Z = \sum_{\mathbf{m}} e^{i\frac{s}{2}\Gamma_{\mathrm{WZW}}[\mathbf{m}]} e^{-\int \mathrm{d}\tau\,\mathrm{d}x\,\frac{1}{2g}(\partial \mathbf{m}/\partial x)^2}. \tag{10.24}$$

This will have a contribution from the degree again, and will behave differently for integer and half-integer spins. We'll then study how this behaves under symmetries and how to see that this is a topological phase from this perspective.

Lecture 11.

The path integral for multiple spins: 10/8/19

Last time, we began discussing the path integral formalism for a single spin, leading to the partition function

$$Z = \operatorname{tr} e^{-\beta H} = \int D\mathbf{n}\, e^{-S_{\mathrm{WZ}}[n] - \int_0^\beta \mathrm{d}\tau\, H(\mathbf{n})}, \tag{11.1}$$

where S_{WZ} is the Wess-Zumino-(Witten) term, which is fictitious in that it's an integral over a disc, not really physical, whose boundary is the circle in question. This involves making choices, and we showed last time that these choices do not affect any physically observable quantity. Precisely, the Wess-Zumino-Witten term is

$$S_{\mathrm{WZ}} = is\int_{D^2}\epsilon^{\mu\nu}\left(\frac{\mathbf{n}\cdot\partial_\mu\mathbf{n}\times\partial_\nu\mathbf{n}}{2}\right). \tag{11.2}$$

We computed the total derivative, which led to computing the difference between the actions for two different choices of extensions to the disc:

$$S_{\mathrm{WZ}} - S'_{\mathrm{WZ}} = is\int_{S^2}\epsilon^{\mu\nu}\left(\frac{\mathbf{n}\cdot\partial_\mu\mathbf{n}\times\partial_\nu\mathbf{n}}{2}\right). \tag{11.3}$$

This looks the same as (11.2), but on a sphere, this is a topological invariant, equal to $4\pi i s k$, where $k\in\mathbb{Z}$. Therefore, if the spin is an integer or half-integer, then when we exponentiate this term disappears.

The next step is to apply this to a spin chain. We'll use the spin chain with Hamiltonian

$$H = J\sum_i \mathbf{S}_i\cdot\mathbf{S}_{i+1} = \frac{1}{2g}\int \mathrm{d}x\left(\frac{\partial\mathbf{m}}{\partial x}\right)^2. \tag{11.4}$$

The Hamiltonian encourages neighboring sites to have alternating spins, which is difficult to feed into the continuum formalism; thus, let $\mathbf{m}_j \coloneqq (-1)^j\mathbf{n}_j$. Now the Hamiltonian encourages neighboring sites to have the same $\mathbf{m}_j$, which is easier to deal with. The Wess-Zumino term is

$$S_{\mathrm{WZ}} = -si\sum_{j\text{ even}}\int_{D^2}\left(\frac{m_{j+1}\cdot\partial_\mu m_{j+1}\times\partial_\nu m_{j+1}}{2} - \frac{m_j\cdot\partial_\mu m_j\times\partial_\nu m_j}{2}\right)\epsilon_{\mu\nu}. \tag{11.5}$$

Let's continuum-ify this. Write

$$m_{j+1} = m(x+a) \approx m(x) + \frac{\partial m}{\partial x}a + \dots \tag{11.6}$$

Then we can rewrite (11.5) as

$$S_{\mathrm{WZ}} = -si\frac{1}{2a}\int \mathrm{d}x \int \mathrm{d}\tau\, a\left(\frac{\partial m}{\partial x} m \times \partial_\tau m\right)\epsilon^{\mu\nu}. \tag{11.7}$$

Therefore the partition function is

$$Z = \int D\mathbf{m}\, \exp\left(\frac{is}{2}\int \mathrm{d}x\, \mathrm{d}\tau\, \mathbf{m}\cdot\left(\frac{\partial_\mu \mathbf{m}\times\partial_\nu \mathbf{m}}{2}\right)\epsilon^{\mu\nu} - \int \mathrm{d}\tau \int \mathrm{d}x\, \frac{1}{2g}\left(\frac{\partial m}{\partial x}\right)^2\right). \tag{11.8}$$

Again there's a discrete topological term — but now, it's of the form $2\pi isk$ for $k \in \mathbb{Z}$. For integer spins, this disappears when we exponentiate. But for half-integer spins, this exponentiates to ± 1 — and this sign factor is the source of significant differences in the physics of integer spin chains and half-integer spin chains. For integer chains, the terms with interesting winding numbers (a topological event) all have the same sign, and gradient terms are penalized for small g. But for half-integer spins, the sign factor kills topological events with odd winding number, and the even-degree terms are more complicated. (TODO: there were more physical consequences that I missed.)

We want to argue that this is an SPT, and in particular that there is interesting physics at the edge of the system. To that effect, let's try to generalize the action a little bit: let

$$S_\theta = \frac{i\theta}{2\pi}\int \mathrm{d}^2x\, \mathbf{m}\cdot(\partial_\mu \mathbf{m}\times\partial_\nu \mathbf{m})\epsilon_{\mu\nu}. \tag{11.9}$$

Time-reversal acts on this system by $\theta \mapsto -\theta$. Also, $\theta \mapsto \theta + 2\pi$ affects the value of S_θ by an integer times $2\pi i$, hence disappears in the partition function and has no physically detectable consequences. Therefore we can think of $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.[16] The system has time-reversal symmetry precisely when $\theta = -\theta$, hence $\theta = 0, \pi$.

We can also imagine letting θ vary over space, giving a function $\theta(x)$ and an action

$$S_{\mathrm{WZ}} = i\int \mathrm{d}\tau \int \mathrm{d}x\, \frac{\theta(x)}{2\pi}\left(\frac{m\cdot\partial_\mu m\times\partial_\nu m}{2}\right)\epsilon^{\mu\nu}. \tag{11.10}$$

For example, one could choose θ to be 0 for $x < -\varepsilon$ and π for $x > \varepsilon$, where $\varepsilon > 0$ is small. We will understand (11.10) by considering variations in the action.

$$\Delta S_{\mathrm{WZ}} = i\int \frac{\theta(x)}{2\pi}\partial_\mu(\delta n\cdot m\times\partial_\nu m)\epsilon^{\mu\nu} \tag{11.11}$$

$$= i\int \partial_\mu\left(\frac{\theta(x)}{2\pi}\delta m\cdot m\times\partial_\nu m\right)\epsilon^{\mu\nu} \tag{11.12}$$

$$= -i\int \frac{\partial_\mu\theta(x)}{2\pi}(\delta m\cdot m\times\partial_\nu m)\epsilon^{\mu\nu} \tag{11.13}$$

$$= -\frac{i}{2}\int \delta m\cdot(\mathbf{m}\times\partial_\tau\mathbf{m})|_{x=0}\,\mathrm{d}\tau. \tag{11.14}$$

Now suppose we keep time-reversal symmetry, so $\theta = 0$ or $\theta = \pi$. One can read off of this variation that in the $\theta = \pi$ phase, we'll have a dangling edge state that's identical to the one in the AKLT state, and indeed these two systems are in the same phase.

$$\sim\cdot\sim$$

The AKLT phase has been experimentally realized. There are several compounds in nature which are three-dimensional, but split into long chains separated by other molecules. For example, "NENP" (chemical formula $\mathrm{Ni}(\mathrm{C_2H_8N_2})_2\mathrm{NO_2ClO_4}$) has the structure of a bunch of chains of nickel molecules bridged by $\mathrm{NO_2}$s, together with perchlorate ($\mathrm{ClO_4^-}$) ions. Nearby chains do weakly interact, which is not quite what we studied, but you can account for it in the Hamiltonian.

We'd like to determine if this is in an SPT phase. This means showing it's gapped in the bulk and has edge modes at the boundary. There are various methods to do this, e.g. neutron scattering, which would work but requires a neutron beam, hence a nuclear reactor!

A thing you can probably do in this building is to investigate the susceptibility of the system to a magnetic field at low temperatures — if the system is gapped, the graph of spin susceptibility as a function

[16] For spin-1/2 systems, there is a 4π periodicity.

of temperature should decay exponentially to zero at low temperatures. And in fact, you can approximately determine the gap from this data, and this is how researchers first noticed that these are gapped phases.

It's trickier to design experiments that will measure the topological edge states. One approach, studied in the early 1990s, is to dope the chain with other materials, introducing impurities which add lots of additional boundaries into the chains. That is, replacing the nickel ions with copper ions (dangling spin-1/2) or zinc, cadmium, or mercury (no dangling spin) has the effect of breaking the spin chain. It's possible the two new edges could interact weakly through this extra ion, but it won't be enough to make a big difference. Now introduce a magnetic field and hit the system with microwave radiation whose wavelength matches (something here?) about the spins. Some of the microwaves will be absorbed by the system, and as you vary the magnetic field strength, there's a clear peak in the absorbance. It's a little unusual that introducing impurities makes it easier to detect the topological behavior of the system! So people went ahead and tried this, and it worked.

$\sim \cdot \sim$

We've now seen four perspectives on 1d SPTs: solvable lattice models, matrix product states, continuum field theory, and experiments. Next we'll discuss the quantum Hall effect and its use to describe fractionalization and 2d topological phases.

Lecture 12.

The integer quantum Hall effect: 10/10/19

Recall that we've been discussing spin chains from a continuum perspective. The action is

$$S = \frac{i\theta}{2\pi} \int_M \left(\frac{\mathbf{m} \cdot \partial_\mu \mathbf{m} \times \partial_\nu \mathbf{m}}{2} \right), \tag{12.1}$$

so that

$$\delta S = \frac{i\theta}{2\pi} \int_{\partial M} \delta \mathbf{m} \cdot (\mathbf{m} \times \partial_\mu \mathbf{m}). \tag{12.2}$$

Therefore the spin in the edge state is $\theta/2\pi$.

This system has a symmetry $\theta \mapsto \theta + 2\pi s$, which doesn't change the system. But time-reversal sends $\theta \mapsto -\theta$, so the only values of θ compatible with time-reversal symmetry are $\theta = 0$ and $\theta = \pi s$ (since $\theta \in \mathbb{R}/2\pi s$). With $\theta = 0$, there are no edge states, so the system is trivial, but for $\theta = \pi s$, there are spin-1/2 edge states and therefore the system is in a nontrivial SPT phase, the same as the AKLT phase, and this is still an SPT if we restrict to a smaller subgroup of the spin rotation symmetry, or just the time-reversal symmetry. If you consider a spin-2 version of this chain, it's only an SPT if you preserve a full spin-rotation symmetry; otherwise one can trivialize the spin-1 edge states.

A similar θ-periodicity term will appear later in this class, in a discussion about topological insulators.

Remark 12.3. Sometimes people study these systems using anomalies, though we're not going to do that in this class. ◂

This SPT phase isn't inevitable: it happened to occur in a relatively simple Hamiltonian in this setting, but that's not true in general. If you impose all ambient symmetries (time-reversal, spin-rotation, translation), there are no nontrivial phases.

We saw that in 1d fermionic phases, there was a system (the Majorana chain) which is nontrivial, but doesn't preserve an additional symmetry. In bosonic phases, that's not true – we must impose some additional symmetry, so the topology is more fragile.

We will now step up to dimension 2, where there are a lot more options, including nontrivial phases with and without additional symmetry, as well as something new involving fractionalized symmetries. Our first example will be the integer quantum Hall effect; if you haven't seen this before, David Tong's lecture notes [Ton16] are great.

In dimension 2, the physics looks like it depends on which gauge one chooses with respect to the magnetic field; this isn't actually the case, but it's a bit subtle. A standard Hamiltonian, say on a torus of size $L_x \times L_y$, looks like

$$H = \frac{(\mathbf{p} + e\mathbf{A})^2}{2m} + V(\mathbf{r}), \tag{12.4}$$

where V is a potential. We will choose the gauge $\mathbf{A} = -By\hat{x}$, where the magnetic field is

$$\mathbf{B} = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{z} = B\hat{z}. \tag{12.5}$$

This has broken translation symmetry in the y-direction, which is OK because (TODO: I missed that part). The wavefunction is

$$\psi(\mathbf{r}) = \frac{Pe^{ik_x x}}{\sqrt{L_x}}\phi_{k_x}(y) \sim \exp\left(-(y + \ell_B^2 k_x)^2/4\ell_B^2\right). \tag{12.6}$$

So our specific Hamiltonian is

$$H_{(k_x)} = \frac{(k_x - y/\ell_B^2)^2}{2m} + \frac{p_y^2}{2m}. \tag{12.7a}$$

We can simplify this by rescaling $y \mapsto y + \ell_B^2 k_x$, yielding

$$H = \frac{(y/\ell_B^2)^2}{2m} + \frac{p_y^2}{2m} = \frac{1}{2}\left(\frac{1}{m\ell_B^2}\right)\left(\left(\frac{y}{\ell_B}\right)^2 + (\ell_B p_Y)^2\right). \tag{12.7b}$$

This looks a lot like a harmonic oscillator – in particular, $\tilde{y} \coloneqq y/\ell_B$ and $\tilde{p}_y \coloneqq \ell_B p_Y$ are dimensionless. Let $\omega_C \coloneqq 1/m\ell_B^2$, which we call the *cyclotron frequency*, from back when people studied these systems in cyclotrons; it's also true that $\omega_C = eB/m$.

Written as in (12.7b), this Hamiltonian is easy to diagonalize:

$$H = \frac{1}{2}\omega_C(\tilde{y}^2 + \tilde{p}_y^2 = \omega_C(\hat{n} + 1/2), \tag{12.8}$$

where $\hat{n} = a^\dagger a$ for

$$a \coloneqq \frac{1}{\sqrt{2}}(\tilde{y} + i\tilde{p}_y) \tag{12.9a}$$

$$a^\dagger \coloneqq \frac{1}{\sqrt{2}}(\tilde{y} - i\tilde{p}_y). \tag{12.9b}$$

So the spectrum once again has a band structure as k_x varies; the bands are indexed by $\hat{n}$, separated by a value we call ω_L, and are called *Landau levels*. The minimum energy level is 0, and the maximum is L_y/ℓ_B^2. In general, we get $2\pi j/L_x$ (TODO: may have confused some numbers).

If we add some electrons, we'll fill from the lowest levels up, and this will be a gapped phase of matter, but this is still not terribly novel. But if you look at this system with a boundary, something weird happens. We'll introduce a potential V; to keep the quantum number k_x, we'll ensure that V only depends on the y-coordinate, and as an additional simplifying assumption, we'll ask for V to change slowly with respect to the magnetic length, or more precisely that

$$\left|\frac{\ell_B}{V}\frac{\partial V}{\partial y}\right| \ll 1. \tag{12.10}$$

This makes the system behave as if it has a wall where the potential gets large. The energy levels are

$$\varepsilon_n(k_x) \approx \omega_L(n + 1/2) + V(-\ell_B^2. \tag{12.11}$$

Semiclassically, the electrons in the bulk are spinning in circles, under the influence of the magnetic field. But on the two walls at small and large y, electrons are moving to the right in one wall and to the left in the other! There will be n right-movers and n left-movers, where n is the number of Landau levels occupied (i.e. less than the chemical potential). This leads to some funny transport properties,

If you wanted to measure this in an experiment, you'd place the sample on a square, and add two contacts on the left and right (so x-direction) of ordinary metal. Along the y-direction, we place two contacts with a voltage difference $\mathcal{V}_y$ smaller than the energy gap. No current will flow in the middle of the system, because the gap is too big to cause an excitation. But we can get a current on the boundary, coming from the moving electrons on the edge – and there will be a current flow, because more will flow one way than the other.

Suppose just a single Landau level is occupied. Then the current flow in the x-direction will be

$$I_x = \int \frac{\mathrm{d}k_x}{2\pi} f(\varepsilon_n(k_x) - \mu(y))(-e)\left(\frac{\partial \varepsilon_n}{\partial k}\right) = -\frac{e}{2\pi}\int_{y_-}^{y_+} \mathrm{d}y\, \frac{\partial V}{\partial y}. \tag{12.12}$$

where μ is the chemical potential. If more than one Landau level is occupied, you can just add up the contributions independently.

The physics here is unusual: we have different phases indexed by the number of Landau levels filled. They all look the same in the bulk, but are distinguished by the *Hall conductance* $\sigma^{xy} \coloneqq e^2 N_{LL}/2\pi$.

When the potential varies in both x and y, k_x is no longer a good quantum number, and we can only graph the energy against TODO. There are critical energies with extended states that aren't fully understood yet (including phenomena with no exact solution), and they are the walls in a wall-and-chamber description of the phases: away from these critical energies, the Hall conductance is locally constant. Therefore the graph of quantized conductance as a function of the magnetic field strength looks like a bunch of steps — and famously, this is still true if there are some impurities in the system! Impurities will keep some electrons trapped near them, but not affect the overall behavior of the bulk.

We can also imagine putting the system on a cylinder, where the periodic direction is the x-direction. We've now left the realm of experiment: to do this in a real system would require a serious breakthrough in nanofabrication. In this case take the gauge $\mathbf{A} = -Byx + (\Phi_x/L_x)\hat{x}$. The magnetic field goes through the middle of the cylinder. Now $k_x \mapsto k_x + \Phi_x/L_x$ and $y \mapsto y - \ell_B^2 \Phi_y/L_x$.

The total charge that moves through any cut of the cylinder is

$$\Delta Q = \int \mathrm{d}x\,\mathrm{d}t\,\frac{\sigma^{xy}}{L_x}\frac{\partial \Phi_x}{\partial t}. = e\sigma^{xy}\Delta\Phi. \tag{12.13}$$

Lecture 13.

2d interacting systems: 10/15/19

TODO: because of computer issues, I missed roughly the first 35 minutes of information. Sorry about that.

Last time, we talked about the integer quantum Hall effect, which takes place in a 2d noninteracting system. We placed it on a rectangle, which represented a possible experimental setup, then discussed more of a *gedankenexperiment* where it's placed on a cylinder or torus. In each of these cases, we can observe that the Hall conductance is quantized. There's a related thermal quantity which is also quantized (this is called the *thermal Hall effect*, I think).

The thesis of today's lecture, I think, is that similar arguments apply to 2d interacting systems (but because of the above issues, I'm not completely sure).

Anyways, we have the system on a torus, with x- and y-directions. There are quantities Φ_x and Φ_y (which are…? TODO) You can use them to define something called the *Berry curvature* $F(\Phi)$, and it's a nontrivial fact that this equals the Hall conductance. One can apply another argument showing that on any closed 2-manifold, the flux is quantized.

$$\sim \cdot \sim$$

We've applied continuum arguments to show these systems are in topological phases, but one can also produce lattice models which realize the same phases. These are known as *Chern insulators*. Some of the details are gory, but it's possible for us to give an overview.

Consider a 2d lattice and a translation-invariant system on it. This gives us enough information to define a Brilloun zone 1BZ and momenta $\mathbf{k} \in \mathrm{1BZ}$. Bloch's theorem tells us that the wavefunctions satisfy

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{13.1}$$

The transformations $\boldsymbol{\phi} \mapsto \boldsymbol{\phi} + \Delta\boldsymbol{\Phi}$ and $\mathbf{A} \mapsto \mathbf{A} + \Delta\boldsymbol{\Phi}/L$ are equivalent to $\mathbf{k} \mapsto \mathbf{k} + \Delta\boldsymbol{\Phi}/L$. Now the Hall conductance is

$$\sigma_H^{xy} = \epsilon^{iy}\frac{\partial}{\partial \Phi_i}\langle \psi(\phi) \mid i\partial_{\phi_j} \mid \psi(\phi)\rangle. \tag{13.2}$$

and the Berry phase is

$$\text{Berry phase} = \sum_n{}' \sum_{\mathbf{k}} \Delta k_x \int_{[0,1]^2} u_{n,\mathbf{k}}(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}} i\frac{\partial}{\partial k_x}\left(u_{n,\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}\right)\mathrm{d}\mathbf{r} \tag{13.3a}$$

$$= \sum_{n,\mathbf{k}}{}'\left(\frac{2\pi}{L}\right)\int_{[0,1]^2}\left(u_{n,\mathbf{k}}^* i\frac{\partial u_{n,\mathbf{k}}}{\partial k_x} - x|u_{n,\mathbf{k}}(\mathbf{r})|^2\right)\mathrm{d}\mathbf{r}. \tag{13.3b}$$

We also call this quantity A_x (TODO: why?). If $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{F}$, then

$$F(\Phi) = \frac{1}{\Delta\phi^x \Delta\phi^y} \sum_{n,\mathbf{k}} \left(u^*_{n,\mathbf{k}} i \frac{\partial u_{n,\mathbf{k}}}{\partial k_x} - x |u_{n,\mathbf{k}}(\mathbf{r})|^2 \right). \tag{13.4}$$

We can simplify this by expressing the terms inside the integral in a way using $\Delta\phi^x$ and $\Delta\phi^y$, leading to the *TKNN formula*

$$F(\Phi) = \sum_n{}' \int_{1\text{BZ}} \frac{\mathrm{d}^2\mathbf{k}}{(2\pi)^2} F_{n,k} = -\sigma^{xy}. \tag{13.5}$$

This formula is named after its discoverers Thouless, Kohmoto, Nightingale, and Nijs [TKNdN82].

Anyways, we see that there is a unique ground state and a gap to the excited states. This leads to a loophole in the quantum Hall effect: if $\sigma^{xy} = (1/2\pi)(1/3)$, there is an issue with the above argument (TODO). The resolution is that there are multiple ground states. This has something to do with adiabatic flux insertion along a rectangle: first U_x, then U_y, then U_x^{-1}, then U_y^{-1}. If this equals the Hall conductance, then these flux insertion operators acting on the ground state don't commute, and instead pick up a phase. The only way for this to make sense is for U_x and U_y to be noncommuting matrices. The specific details of the system mean the space of ground states must be at least 3-dimensional! We'll see more about this in a few lectures.

One can couple this system to a gauge field; this is always possible given a conserved charge. One modifies the Hamiltonian $H = \sum_{i,j} \psi_i^\dagger t_{ij} \psi_j$ to something such as $\psi_i^\dagger t_{ij} e^{iA_{ij}} \psi_j$. This will be invariant under gauge transformations of ψ, e.g. $\psi_i \mapsto e^{i\chi_i}\psi_i$, where χ_i is a phase that may be site-dependent; and it's also invariant under

$$A_{ij} \longmapsto A_{ij} + (\chi_i - \chi_j). \tag{13.6}$$

Lecture 14.

Dirac cones and anomalies: 10/17/19

We've been discussing the 2d integer quantum Hall effect, with Hall conductance $\sigma^{xy} = N/2\pi$. One could consider other values of the Hall conductance. The adiabatic theorem applies, showing that the system is gapped for any $\sigma^{xy} \in \mathbb{R}$, and that there's a unique ground state for $N/2\pi$. However, for $\sigma^{xy} = p/(2\pi q)$, where $q > 1$, there can be degenerate ground states.

We also saw that it made sense to couple this system to a gauge field. This is not true in general — the obstruction is something called an *anomaly*.

Example 14.1. Our first example of an anomaly will be for a 2d Dirac fermion, with Hamiltonian

$$H = \int \mathrm{d}^2 r\, \psi^\dagger (v\boldsymbol{\sigma} \cdot \mathbf{p}) \psi, \tag{14.2}$$

where $\psi = \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}$. The spin is coupled to the momentum, in that if we graph k_x against $\varepsilon(k)$, the possible combinations form two cones, opening upward and downward.

This system has a time-reversal symmetry $\mathcal{T}\psi \coloneqq i\sigma^y K\psi$. This symmetry precludes nontrivial Berry curvature.

There's zero Hall conductance, which is fine so far… but let's add a small mass term to the Hamiltonian:

$$H = \int \mathrm{d}^2 r\, \psi^\dagger (v\boldsymbol{\sigma} \cdot \mathbf{p} + m\sigma^2) \psi, \tag{14.3}$$

This breaks time-reversal symmetry, since time-reversal symmetry would send $m \mapsto -m$. The energy satisfies

$$\varepsilon = \pm\sqrt{(vp)^2 + m^2}, \tag{14.4}$$

giving two parabolas. At $k = 0$, the spin points up on the positive branch, and down on the negative branch. If $k \neq 0$, the spin begins pointing a little sideways: if the momentum is large, the mass is irrelevant and the spin points sideways.

Now what happens to the Hall conductance? We have

$$\sigma^{xy} \int F_k \frac{\mathrm{d}^2 k}{(2\pi)^2} = \left(-\frac{1}{2} \operatorname{sign}(m) \right) \frac{1}{2\pi}, \tag{14.5}$$

which, in particular, is not $N/2\pi$ for $N \in \mathbb{Z}$! But that was a requirement to be a continuum limit of a lattice model, so this can't be one. This is the anomaly. And indeed, in general systems one only ever sees Dirac cones in pairs — indeed, one continuum approach to these systems is a pair of Dirac cones with equal mass, giving a Hall conductance of 1. ◂

Let's tallk about that continuum approach via a pair of Dirac fermions. Now

$$\psi = \begin{pmatrix} \psi_{\uparrow,1} \\ \psi_{\downarrow,1} \\ \psi_{\uparrow,2} \\ \psi_{\downarrow,2} \end{pmatrix}. \tag{14.6}$$

The Hamiltonian is

$$H = \int \mathrm{d}^2 r\, \psi^\dagger (v\mathbf{p} \cdot \boldsymbol{\sigma}\tau^z + m_1\sigma^z + m_2\sigma^z\tau^z)\psi. \tag{14.7}$$

Here σ is a spin Pauli matrix and τ is some other quantum number not related to spin.

The Hall conductance depends on the two masses. Change a basis to $e_1 \coloneqq m_1 + m_2$ and $e_2 \coloneqq m_1 - m_2$.

- In the first quadrant $\sigma_H = -1/2\pi$: the two half-integer pieces appear with opposite signs and cancel out.
- In the second and fourth quadrants, the Hall conductance vanishes.
- In the third quadrant, $\sigma_H = 1/2\pi$.

Haldane figured out how to explicitly realize this system on a hexagonal lattice; the details are a little tricky but the ideas are clear in retrospect.

Let's allow the masses to vary in the y-direction (in the (e_1, e_2) coordinates, e_1 decreases from positive to negative, but e_2 stays constant and positive). The $\varepsilon = 0$ solution is

$$\int \mathrm{d}^2 r\, \psi^\dagger_{\sigma,+}(v(\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}) + m(y)\sigma^z)\psi_{\sigma,+} \tag{14.8a}$$

$$(-iv\sigma^y\partial_y + m(y)\sigma(z))\chi = 0 \tag{14.8b}$$

$$\chi(y) = \exp\left(\int_0^y f(\hat{y})\,\mathrm{d}\hat{y}\right)\chi_0 - iv\sigma^y f(y), \tag{14.8c}$$

where $f(y) \coloneqq cm(y)/v$. So we have a solution, and in particular an interface in space where $\sigma^{xy} = 0$ ($e_1 = 0$). We also have

$$m(y)(-i\sigma^y + \sigma^z)\chi_0 = 0 \tag{14.9a}$$

$$(-i\sigma^z\sigma^y + 1)\chi_0 = 0(1 - \sigma^x)\chi_0 = 0, \tag{14.9b}$$

so we see that

$$\chi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{14.9c}$$

TODO: another calculation, after this, which I missed.

Let's return to the case of a single Dirac fermion, this time in 1d. If we let the momentum vary in time, one of the takeaways is that $\frac{\partial k_x}{\partial t} = -\varepsilon_x$. We also get a *chiral mode*, which can only move in one direction. This takes place, e.g., in the presence of a background electric field. This field causes momenta to increase over time — we start with a bunch of filled states moving into filled states, but above the chemical potential, filled states move into nonfilled states. Over time, the total charge Q of the system increases, and

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\frac{L}{2\pi}E_x. \tag{14.10}$$

That is, this perfectly inocuous system, which appears to obey charge conservation, actually doesn't. This is another signal that a single Dirac fermion can't be formulated on a lattice. In general, the violation of charge conservation is an integer:

$$\frac{\mathrm{d}\varrho}{\mathrm{d}t} = -\frac{1}{2\pi}E_x(N_R - N_L), \tag{14.11}$$

where N_R is the number of right-moving modes and N_L is the number of left-moving modes. Thus, like the Hall conductance, this anomaly coefficient is an integer.

In Example 14.1, this anomaly looked like it was about time-reversal symmetry, and followed a $\mathbb{Z}/2$ classification. Here it looked related to charge conservation, and has a $\mathbb{Z}$ classification. The anomaly is even worse than that, in fact — when coupled to a temperature gradient, this system doesn't even conserve *energy*.

The fact that anomalies are valued in some abelian group looking like $\mathbb{Z}$ or $\mathbb{Z}/2$ is reminiscent of our classification of topological phases — in both cases, we have a discrete classification, and an addition coming from something like stacking. Indeed, the previous example showed a relationship between anomalies of a 1d system and the 2d quantum Hall effect.

So the next logical question is: does the 2d anomaly, related to time-reversal symmetry, connect to a 3d topological phase? This would be a 3d phase, protected by time-reversal symmetry and a U_1 charge symmetry. Time-reversal complex-conjugates the charge, so the symmetry group is $\mathrm{U}_1 \ltimes (\mathbb{Z}/2)^T$.

Let's see how the 3d Hamiltonian

$$H = \int \mathrm{d}^2 r\, \psi^\dagger (\mathbf{p}\cdot\boldsymbol{\sigma}\tau^z + m_1\tau^x + m_2\tau^y)\psi. \tag{14.12}$$

behaves under time-reversal symmetry. (Again ψ is as in (14.6).) We need the Hamiltonian to be invariant, so real terms, such as m_1, m_2, and τ^x, are fine, but τ^y is complex, so the $m_2\tau^y$ term goes away. The $\mathbf{p}\cdot\boldsymbol{\sigma}$ term is also fine.

So we see a parameter that allows us to distinguish the two time-reversal-symmetric phases. TODO: I didn't figure out which it was; maybe it's m_1? The next step was to let m_1 vary in z, behaving roughly like $\arctan(z)$, so that we should expect an interface at $z = 0$, where the mass changes sign. The solution is

$$\chi = e^{i\mathbf{k}\cdot\mathbf{r}} \exp\left(-\int_0^z \frac{m(\tilde{z})}{v}\,\mathrm{d}\tilde{z}\right)\chi_0, \tag{14.13}$$

and $(1+\sigma^z\tau^y)\chi_0 = 0$. There is a two-dimensional space of solutions to this latter equation (idea: σ^z and τ^y must have opposite signs, but we could do $+-$ or $-+$). We can therefore write down Pauli matrices $\tilde{\sigma}$ for this two-dimensional vector space and insert this term into the Hamiltonian.

The upshot is that there are low-energy solutions, and they're bound to the interface. Moreover, these interface modes look a lot like 2d Dirac fermions with time-reversal synmmetry.

That is, we are able to realize the anomalous theory of a 2d Dirac fermion by setting it as the boundary of a 3d theory. Fermion doubling still applies: the 3d slab has two surfaces, and we have a Dirac fermion on each surface, but now they're spacelike separated. These systems have been studied experimentally, and there are probably over 100 different compounds experimentally realizing this.

This correspondence between bulk theories and anomalous theories in one dimension lower holds in great generality: every anomalous theory that we know of has a realization as a boundary theory. This was a surprise when people first studied examples: anomalies are a concept from high-energy physics, and the fact that this is the same thing as something from solid-state physics is kind of cool.

Next week, we'll study 2d topological phases with ground state degeneracy, together with fractionalized excitations and exotic statistics.

Lecture 15.

The Kitaev honeycomb model: 10/22/19

Today, we'll study a 2d system on a hexagonal lattice, written down and solved by Kitaev [Kit06], and called *Kitaev's honeycomb model*. For notation, choose an orientation of each plaquette (face) and then label the vertices, in cyclic order, x, y, z, x, y, z, so that the three edges adjacent to a given vertex are labeled x, y, and z. Let X denote the set of edges labeled x, Y denote the set labeled y, and Z denote the set labeled z.

Introduce a spin at each vertex. The Hamiltonian is

$$H := -\sum_{ij\in X} J^x \sigma_i^x \sigma_j^x - \sum_{ij\in Y} J^y \sigma_i^y \sigma_j^y - \sum_{ij\in Z} J^z \sigma_i^z \sigma_j^z. \tag{15.1}$$

We also have local operators associated with each plaquette P:

$$B_P := \sigma_1^y \sigma_2^z \sigma_3^x \sigma_4^y \sigma_5^z \sigma_6^x, \tag{15.2}$$

where the subscripts denote the edges of the hexagon in cyclic order. These B_P square to the identity.

Kitaev solved this model by rewriting the spins as if they were spins of fermions. There are different ways to do this, e.g.

$$\mathbf{S} = \frac{1}{2} v C_\alpha^\dagger \sigma_{\alpha\beta} C_\beta, \tag{15.3}$$

but Kitaev did something different: introducing four Majorana fermions f^x, f^y, f^z, γ (Majorana means these operators all square to 1), and writing $\boldsymbol{\sigma}_i = i\mathbf{f}\gamma$. This is consistent, because $(\sigma_i^\alpha)^2 = 1 = (if^\alpha\gamma)^2$. You can think of γ as living at the vertex and f^x, f^y, f^z as living slightly further away along the x, y, and z edges respectively.

Some other useful algebra:

$$\sigma_i^x \sigma_i^y = i\sigma_i^z \tag{15.4a}$$

$$\sigma_i^x \sigma_i^y \sigma_i^z = i. \tag{15.4b}$$

We can also calculate that

$$\sigma_i^x \sigma_i^y \sigma_i^z = i^3 f^x \gamma f^y \gamma f^z \gamma = i f^x f^y f^z, \tag{15.5}$$

so we artificially impose the constraint

$$G_i \coloneqq f_i^x f_i^y f_i^z \gamma_i = 1. \tag{15.6}$$

There are advantages and disadvantages of this approach: $\mathbf{f}$ and γ are weakly fluctuating degrees of freedom, which is nice, but dealing with (15.6) is a little annoying.

Remark 15.7. One can produce an analogy with electromagnetism, and introduce background electric or magnetic fields. In this case, the field lines will always be closed loops. ◁

Now we rewrite (15.2) in terms of the Majorana operators. In the end we'll see that we should rewrite the operators again!

$$B_P = \prod_{j=1}^{6} i f_j^y \gamma_j \tag{15.8a}$$

$$= (if_1^x f_1^z)(if_2^x f_2^y)(if_3^y f_3^z)\cdots \tag{15.8b}$$

so if we define

$$s_{ij} \coloneqq i f_i^{\alpha_{ij}} f_j^{\alpha_{ij}} \in \{\pm 1\}, \tag{15.9}$$

then

$$B_P = \prod_{ij \in \partial P} s_{ij}. \tag{15.10}$$

Now let's rewrite the Hamiltonian in terms of $\mathbf{f}$, γ, and s_{ij}.

$$H = -\sum_{ij \in X} J^x i f_i^x \gamma_i i f_j^x \gamma_j + (y \text{ and } z \text{ terms}) \tag{15.11a}$$

$$= \sum_{ij \in X} (J^x s_{ij})(i\gamma_i\gamma_j) + (y \text{ and } z \text{ terms}). \tag{15.11b}$$

This looks pretty similar to a tight-binding model of electromagnetism, a particle moving in a gauge potential, which has the Hamiltonian

$$\sum_{\text{edges } ij} \psi_i^\dagger t_{ij} e^{A_{ij}} \psi_j + \text{h.c.}, \tag{15.12}$$

where $A_{ij} \coloneqq \int_{x_i}^{x_j} \mathbf{A} \cdot \mathrm{d}\boldsymbol{\ell}$, so $e^{iA_{ij}} \in \mathrm{U}_1$.

Here, though, we've constrained $s_{ij} \in \{\pm 1\} \cong \mathbb{Z}/2$ — but everything else stays the same. In fact, if we try to write down the electromagnetic story but replacing U_1 with $\mathbb{Z}/2$ everywhere, and $e^{iA_{ij}}$ with s_{ij}, we'll get to the rewritten honeycomb model. That is, this is a sort of $\mathbb{Z}/2$ gauge theory. The J operators are like hopping amplitudes for the different particles.

In electromagnetism, magnetic flux around a loop ℓ is measured by

$$\exp\left(i \oint \mathbf{A} \cdot \mathrm{d}\boldsymbol{\ell}\right) = \prod_{\partial \ell} e^{iA}, \tag{15.13}$$

and the right-hand side looks a lot like (15.10) — that is, B_P measures the magnetic flux on the boundary of a plaquette P.

The constraint operator G_i from (15.6) now has a convenient interpretation, as a generator of the $\mathbb{Z}/2$ gauge transformations. Indeed,

$$G_i \gamma_i G_i^\dagger = -\gamma_i \tag{15.14a}$$

$$G_i \mathbf{f}_i G_i^\dagger = -\mathbf{f} \tag{15.14b}$$

$$G_i s_{ij} G_i^\dagger = -s_{ij}. \tag{15.14c}$$

The fact that we can rewrite this model as a gauge theory isn't too special — you can do that for any spin system. What is special is that we get the gauge theory in its deconfined state, making it easier to solve.

We'll eventually see that for every choice of the coupling constants J^x, J^y, J^z, the ground states are those without any magnetic flux. Intuitively, the fluxes frustrate hopping, because they cause destructive interference, and this is penalized in the Hamiltonian: being able to move lowers the kinetic energy of the particle.

The phase diagram is three-dimensional, given by J^x, J^y, J^z. To simplify the analysis, let's assume $J^x + J^y + J^z = 1$ and each J^μ is nonnegative. This can be done without loss of generality, because we don't care about the energy scale, and because the system is symmetric under (TODO: might not be exactly this) flipping the signs of the Js. See Figure 2 for a picture of the phase diagram.

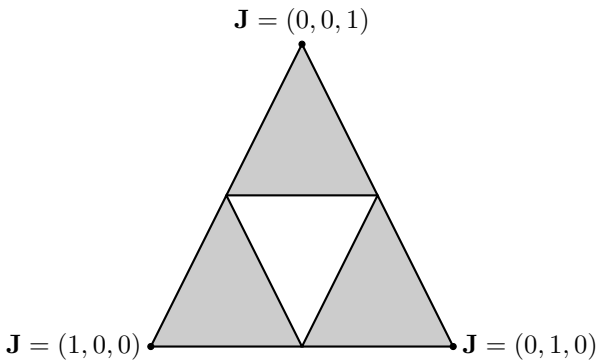


FIGURE 2. The phase diagram of the Kitaev honeycomb model, when the coupling constants are nonnegative and sum to 1. The gray zones are all in the same gapped phase; the colorless zone in the middle is gapless.

To actually see this, let's rewrite the Hamiltonian as

$$H = \frac{i}{4} \sum_k \gamma^\dagger(k) \begin{pmatrix} 0 & \alpha(k) \\ -\alpha^\dagger(k) & 0 \end{pmatrix} \gamma(k). \tag{15.15}$$

The fact that $\gamma^\dagger(k) = \gamma^{\mathrm{T}}(-k)$ is a kind of particle-hole symmetry, similar to the one we saw near the beginning of class. The energy of a state is $E = |\alpha(k)|$.

There is a curious Aharonov-Bohm effect at low energies: there are fermions ψ and $\mathbb{Z}/2$-magnetic fluxes m, and these have mutual statistics: if you move one around another, you pick up a -1. This is unusual. Another unusual thing is that, well, we began with bosons and somehow fermions popped out! This is an instance of fractionalization. However, the fermion parity is preserved: there is no local operator that can create an odd number of fermions.

The analysis can get a bit technical, and we'll eventually find a particle with nonabelian statistics! But let's begin near a corner of the phase diagram: $J^x = 1 - 2\varepsilon$, and $J^y = J^z = \varepsilon$, where ε can be zero (!) — the point is that J^x is basically the only important coupling constant.

We can use the standard trick to rewrite Majoranas into complex fermions, with $\psi_{ij} = (1/2)(\gamma_i + i\gamma_j)$. These are the gapped fermion excitations mentioned above. But the flux sector is massively degenerate. This is a bit sad, so let's perturb the flux system: a weak perturbation will fix this degeneracy and not affect the perturbations.

Now let's suppose $0 < \varepsilon \ll 1$. To fix the degeneracy, we will have to go to relatively high-order perturbation theory, which might be less familiar and will require some new technology. A preview of what's going to happen: the new states we'll mix in will come from old states, but acted on by various powers of the perturbation.

(TODO: something happened) and the upshot is that creating a pair of fermions takes energy, but then we can separate them for free. This is unlike, say, QCD, where separating two quarks costs energy. But the takeaway, I think, is that there are virtual particles that can be created and annihilated, and these are responsible for the need to include higher-order terms in the perturbation theory.

Lecture 16.

Perturbation theory for the Kitaev honeycomb model: 10/24/19

TODO: missing a lot of context because I was a bit too tired to effectively follow.

Recall that we've been discussing the Kitaev honeycomb model, and the way to rewrite it as a $\mathbb{Z}/2$ gauge theory, but with a fermion (an instance of fractionalization). We will use the same notation from last time.

We considered the strong J^x limit (so $J^x \gg J^y, J^z$). If $J^y = J^z = 0$, which forces $\sigma_x^1 = \sigma_x^2$, then there are multiple ground states, and they're mixed, but when J^y and J^z are small and nonzero, we will be able to pick out a distinguished ground state.

Consider the operators $X^{12} = \sigma_1^x = \sigma_2^x$, $Y_{12} = \sigma_1^y\sigma_2^z = -\sigma_1^z\sigma_2^z$ and $Z_{12} = \sigma_1^z\sigma_s^y = \sigma_1^y\sigma_2^z$. These satisfy the commutation relations of Pauli matrices, which can be useful.

To study the system in this regime, we'll need to do some higher-order perturbation theory. We will decompose the Hamiltonian into something easier to understand (the J^x terms) and something with a small perturbation (the J^y and J^z terms). We begin with the Green's function

$$G_{\text{tot}}(\varepsilon) = \frac{1}{\varepsilon \cdot \text{id} - H}. \tag{16.1}$$

We will write $H = H_0 + V$, where H_0 contains the J^x terms and V contains the J^y and J^z terms. In this case, H_0 is diagonal, with the first $2^{\#P}$ entries equal to zero. Here $\#P$ is the number of plaquettes. After that we have terms of the form $2J^z$, $4J^z$, etc. However, V has plenty of off-diagonal terms, and our goal in perturbation theory is to fix this. Write

$$G(\varepsilon) = \Pi_X G_{\text{tot}}(\varepsilon)\Pi_X \tag{16.2}$$

$$= \Pi_X\left(\frac{1}{\varepsilon - H_0 - V}\right)\Pi_X \tag{16.3}$$

$$= \Pi_X\left(\frac{1}{(\varepsilon - H_0)(1 - (\varepsilon - H_0)^{-1}V)}\right)\Pi_X. \tag{16.4}$$

So if we let $G_0(\varepsilon) \coloneqq 1/(\varepsilon - H_0)$, this is

$$G(\varepsilon) = \Pi_X(G_0^{-1}(1 - G_0V))^{-1}\Pi_X \tag{16.5}$$

$$= \Pi_X(1 - G_0V)^{-1}G_0\Pi_X \tag{16.6}$$

$$= \Pi_X\left(\sum_{n=0}^{\infty}(G_0V)^nG_0\right)\Pi_X \tag{16.7}$$

$$= \Pi_X\left(1 + \underbrace{\sum_{n=1}^{n_{\max}}(G_0V)^n}_{=:T}\right)G_0\Pi_X \tag{16.8}$$

$$= \Pi_X((1 + \Pi_XT\Pi_X)^{-1})^{-1}\Pi_X \tag{16.9}$$

$$= \Pi_X(\varepsilon - H_0 + \Pi_XG_0^{-1}T\Pi_X)^{-1} \tag{16.10}$$

This allows us to obtain an effective Hamiltonian (TODO: several more equations on the board that I didn't understand), as a function of $n_{\max}$. This tells us information about the ground states.

- If $n_{\max} = 1$, the constraint is $\Pi_X V \Pi_X = 0$.
- For $n_{\max} = 2$, the constraint is $\Pi_X V (1/4J^x) V \Pi_X = c$ for some constant c. This is a second-order correction to the ground-state energy but doesn't completely solve everything.
- The third-order ($N_{\max} = 3$) constraint also doesn't suffice: it involves three V terms, but in a way that you can use them to produce excited states.
- At 4^{th}-order, we're good. Because

$$V(G_0 V)^3 = \frac{(J^y)^2 (J^z)^2}{(4J^x)^3} \sigma_2^y \sigma_3^y \sigma_3^z \sigma_4^z \sigma 5^z \sigma_6^z \sigma_6^y \sigma_1^y = \frac{(J^y J^z)^2}{64 (J^x)^3} \sigma_3^x \sigma_6^x \sigma_2^y \sigma_5^y, \tag{16.11}$$

  we finally have an interacting terms between four spins, providing new information. The effective Hamiltonian is just

$$H_{\text{eff}} = -\frac{1}{16} \frac{(J^y J^z)^2}{(J^x)^2} B_P. \tag{16.12}$$

  So this tells us we'd prefer not to have flux around any plauette.

Higher-order terms behave similary: we can have larger loops, and obtain contributions from σ^y and σ^z terms in strings of length $n_{\max}$ – but the contribution decreases exponentially in $n_{\max}$. The general effective Hamiltonian will include energy penalties for all loops, but with a penalty that decreases exponentially in the circumference of the loop.

Now we can rewrite the effective Hamiltonian (16.12):

$$H_{\text{eff}} = -\kappa \sum X_L X_R Z_T Z_B = \sum \tilde{X}_1 \tilde{X}_2 \tilde{X}_2 \tilde{X}_4 + \sum \tilde{z}_1 \tilde{Z}_2 \tilde{Z}_3 \tilde{Z}_4. \tag{16.13}$$

TODO: these sums are over something that I didn't understand.

This looks a lot like the Hamiltonian for the toric code, which is a Hamiltonian on a square lattice. It's really nice, in that it's a sum of commuting projectors, so we just have to understand the product of the X terms and the product of the Z terms for each plaquette.

The excitations of this system behave a little surprisingly: excitations are deconfined: they can be created in pairs and can be arbitrarily separated without energy cost. However, you can't just create one at a time: they appear as edges of strings. The ground state is an eigenstate of closed strings of various lengths. You can think of the string operators as moving excitations between plaquettes.

If the string has a bunch of Z operators, call it an e particle, and if it has X strings, call it an m particle. Both have energy 2κ, and they have relative statistics of -1. They do behave a little like electric and magnetic flux, respectively — in the original honeycomb model, they're the fluxes through two different kinds of hexagonal faces. The translation symmetry of the honeycomb model turns one into the other, which is a little weird.

Lecture 17.

Particle statistics in the Kitaev honeycomb model: 10/29/19

In the study of the Kitaev honeycomb model, we saw that it looked a lot like a model on a square lattice, called the toric code. Color the plaquettes in the square lattice like a chessboard, so alternating black and white. Put a spin on each edge, and let X_v, Y_v, and Z_v denote the Pauli operators associated to any vertex of the lattice, i.e. acting on each of the spins on the edges touching the vertex. Then the Hamiltonian is

$$H = -\kappa \left(\sum_{\square} \prod_{\text{vertices } i \text{ of } \square} X_i - \sum_{\blacksquare} \prod_{\text{vertices } i \text{ of } \blacksquare} Z_i \right). \tag{17.1}$$

This is a nicer system to solve, because these terms pairwise commute and are projectors. On the infinite plane, the system is gapped and there is a unique ground state. And the excitations (TODO: I missed a bit of the argument because I was late) are given by string operators, acting by Z_v on each vertex v in the string. This creates two excitations localized at the ends of the string, and by acting with another string, we can move one or the other around at no cost — but we can't create only one at a time. The same is true if we

use strings of X operators instead of Z operators. You can think of the ground state as a fluctuating soup of closed strings, and that open strings correspond to excited states.[17]

There are two kinds of excitations: those created at the ends of Z strings, which we'll call e, and those created at the ends of X strings, which we'll call m. They are particle-like and point-like, lumps of energy, but they are also ends of strings, with the caveat that the strings themselves are not observable; there's no measurement that can detect the difference between two strings with the same starting and ending points (and the same operator X or Z). You can think of e as kind of an electric charge, and m as sort of a magnetic flux.

This story remains broadly true in the honeycomb model, though there's also a third kind of excitation, a fermion ψ. Before that, though, let's see what e and m look like in the honeycomb model. You can't create only a single flux in a single plaquette, but like in the toric code, you can create a pair and move one far away. But when you look more carefully into it, you can only move excitations over an even number of columns at a time. Therefore there are two kinds of excitations: those on even columns, and those on odd ones. The e excitations correspond to even columns, and m to the odd ones. The bond excitations, which we pushed away to infinite energy in the toric code, are the fermions.

If you act just by a z on an e excitation, you can produce an m excitation and a fermion. The reverse is also true: you can act on a ψ to produce an e and an m. These are local operations. You can think of this as ψ decaying into an e and an m. Since e and m each have energy 2κ, but $J^x \gg \kappa$, this is a little strange, but if you ignore this and focus on just topological properties, you'll see that ψ has the same exchange statistics as e and m together.

Remark 17.2. If this is confusing, you can gain a lot of insight by just writing down some string operators and seeing what happens. ◁

Since ψ behaves like a fermion, and a local operation can turn ψ into e and m, then e and m, together, should act like a fermion. This is an instance of something weird: separately, e and m are each bosons, but their bound state acts like a fermion.

To understand these statements, we want to determine the sign in $|\bullet_1\bullet_2\rangle = \pm|\bullet_2\bullet_1\rangle$. Of course, since this is a phase, we can't observe it directly; instead, we can let the particles evolve in time, where in one setup they don't exchange positions, and in another they do. Then one measures the relative phase, e.g. by using an interferometer and measuring whether there's constructive or destructive interference. The upshot is that this is a time-dependent process.

Let's work in the toric code for simplicity. Given an open string whose vertices are labeled by Z operators, the product of these Z operators acting on the ground state is $|e_1e_2\rangle$, and this leads to the conclusion that e is a boson. Running the same argument with X operators shows m is also a fermion.

What will be weird is when we try this with one e and one m. Even though they are different kinds of particles, they will have nontrivial mutual statistics, which is not something we're used to happening. That's because it cannot happen in three dimensions: the braid that corresponds to moving one particle around the other might be nontrivial, but two copies of it (moving one particle around another twice) can be untangled. Therefore the statistics must square to 1: bosons give 1 and fermions give -1.

Here, though, we're in dimension 2, so that argument doesn't apply. Consider the following process, starting with the ground state.

(1) Create a pair of e excitations from the vacuum and move them apart.
(2) Create a pair of m excitations along a string crossing the e string.
(3) Shrink the e string to zero, annihilating the e excitations.
(4) Shrink the m string to zero, annihilating the m excitations.

This reproduces the ground state, though it might have picked up a phase. If you draw the worldlines of the particles, you get two interlinked squares (called a *Hopf link*), one square corresponding to the e particles and one corresponding to the m particles. We have strings of Zs on one square and strings of Xs on the other. That reduces this to an algebraic question: take a sequence of X operators and a sequence of Z operators which share an odd number of sites, and compare with the completely unlinked worldlines. This can be done by unlinking one crossing, and this operator multiplies by -1. Therefore the mutual statistics of e and m is -1.

[17] This suggests that if you write this system on a manifold with nontrivial π_1, the space of ground states is different. This is true, and we'll learn more about this later.

The Majorana description of the ψ excitations means that it also has fermionic statistics, so the nontrivial mutual e-m statistics above and the fact that ψ can decay into an e and an m should be related. From the $\mathbb{Z}/2$-gauge theory perspective, this is an Aharonov-Bohm effect: e is the "electric particle", m is the "magnetic flux", and the fermions feel these, explaining these mutual statistics.

The spin-statistics theorem is an important result in physics relating the statistics of a particle and its rotational symmetry (its spin). In relativistic theories, so in high-energy physics, fermions are spin-1/2 particles, so rotating by 2π picks up a -1 sign in the wavefunction; correspondingly, bosons are integer-spin particles, so there's no extra sign under a full rotation.

The theorem is most often stated in relativistic settings, but that's strictly speaking not necessary. Consider the diagram in Figure 3, as a pair of framed links, i.e. we keep track of the orientation as we walk around the link. In the left-hand figure, the particles are created and annihilated without any fuss, but in the right-hand figure, the framing switches at the crossing, and therefore the particles are exchanged. But the right-hand figure also corresponds to performing a full rotation of the particle.

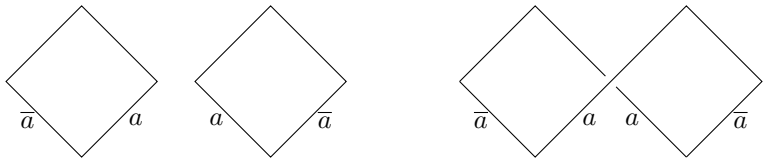


FIGURE 3. The relative phase between these two worldlines of particles is the statistics θ_a of the particle a.

The upshot is that if θ_a denotes the statistics of particle a, then

$$\theta_{e\times m} = \theta_e\theta_m\theta_{e,m} = (1)(1)(-1) = -1. \tag{17.3}$$

Here $\theta_{e,m}$ is the mutual statistics of excitation e and m.

In a general 2d system, we obtain the following facts about exchange statistics.

- We could consider the mirror image to the rotation above, applied to a particle a, and this mirror image has statistics θ_a^*.
- The mutual statistics of particles a and b satisfies $\theta_{a,b} = \theta_{b,a}^*$.

These statistics are invariant under smooth deformations of the path in both position and time used to compute them, hence are topological! The mutual statistics are also invariant under local perturbations. This suggests that we should consider two excitations to be "topologically equivalent" if they cannot be distinguished from far away. For example, in the honeycomb model, ψ and $e \times m$ are equivalence in this sense. The equivalence classes of particles have the fun name *topological superselection sectors*. Self and mutual statistics are preserved under this operation. In the toric code, there are four classes: e, m, ψ, and 1 (no excitations at all). Topologically, $[e \times e] \sim [1]$: a pair of es next to each other looks the same from far away as no excitations at all. The same is true for a pair of $[m]$s.

Remark 17.4. The list of topological superselection sectors, as well as their self and mutual exchange statistics, turns out to characterize a certain class of phases in 2d, called *abelian topological phases*, up to equivalence. Here, "abelian" means that particle exchange statistics doesn't depend on the order of the two particles. Not all phases are abelian. ◂

Lecture 18.

Anyons and the toric code: 10/31/19

We've been studying the toric code, with its four particles 1, e, m, and $\psi = e \times m$ (formally, putting an e and an m particle very close together and considering them to be one; this is called *fusion*). These particles have interesting statistics: in general, the statistics θ_a of a particle a is the phase ratio between switching the two particles and not switching them. (TODO: I think this is just for abelian anyons?)

In an experiment, however, you'll pick up other phases than θ_a, including a term corresponding to the geometry of the path and including Aharonov-Bohm effects. If you want to detect these phases in an experiment, you'll have to arrange for these extra phases to cancel.

The fusion rules in the toric code are: we already know $\psi = e \times m$, but $e \times e = m \times m = \psi \times \psi = 0$, since a pair of e or m particles can be created by a local operator from the vacuum, and for ψ we create both a pair of es and a pair of ms from the vacuum.

The mutual statistics of two particles a and b, $\theta_{a,b}$ is the phase ratio between moving b in a counterclockwise path around a, in a full circle, versus not moving it. If $a = b$, it's not necessarily true that $\theta_{a,a} = 1$, but it is true that $\theta_{a,a} = \theta_a^2$ (so for the toric code, $\theta_{a,a} = 1$ for all a, but not more generally). It is always true that $\theta_{a,b} = \theta_{b,a}$, which you can argue with a picture.

In the toric code, as long as $a, b \neq 1$ and $a \neq b$, $\theta_{a,b} = -1$. It's as if each particle sees all other particle types as flux sectors.

In general, you can prove that $\theta_{a\times b} = \theta_a\theta_b\theta_{a,b}$: one shows this by considering what happens if one moves an a and a b particle around another a and a b particle. You can use this to show that $m \times \psi$ has the same statistics as an e particle: it has $+1$ with 1 and e and -1 with m and ψ — in fact, it is an e particle, though this isn't a proof. In this case, the fusion product is abelian, so we can argue that

$$m \times \psi = m \times (e \times m) = m \times m \times e = e. \tag{18.1}$$

In the toric code, every particle is its own antiparticle. In general, if you can create a particle a with a local operator, it must also come with its antiparticle $\overline{a}$, since the local degrees of freedom are just spins. In particular, a is a boson, and

$$1 = \theta_{a\overline{a}} = \theta_a\theta_{\overline{a}}\theta_{a,\overline{a}}. \tag{18.2}$$

You can convince yourself that $\theta_a = \theta_{\overline{a}}$, so $\theta_{a,\overline{a}} = \theta_a^{-2}$.

$$\sim \cdot \sim$$

Now let's return to discussing the ground state of the toric code, where there's more structure to discuss. Let's consider a black square on the checkerboard, which the Hamiltonian assigns a product of Z operators on the edges. We could therefore have all four spins down, or all four spins up, orn two up and two down. This limits the choices for neighboring plaquettes if we want to stay in the ground state sector, since an even number of spins must be down on every black plaquette.

The second term in the Hamiltonian imposes additional constraints, with a product of X operators around a white plaquette. For example, we can't just put all four spins up; instead, we'll have to take a superposition of that and all four spins down.

So the Z terms tell us that the sites with down spins have to form a closed loop, and the X terms tell us that we must consider an equal superposition of *all* such loops. That's a ground state; does the state we get depend on how we began constructing it? The answer actually depends on the boundary conditions, or on which manifold we place the system on.

For example, on a sphere, any loop can be decomposed into small loops, so acting by plaquette operators can reach any kind of loop, and therefore the ground state is a superposition of all possible loops, and there is a unique ground state. But on a torus, there are loops which cannot be accessed from the empty configuration, or said differently, there are loops that cannot be created by applying plaquette operators. An example is a cycle that goes around the torus. In the end, we'll get a four-dimensional ground state, corresponding to the four mod 2 homology classes of loops in the torus: there's a ground state which is an equal superposition of all loops that can be created by plaquette operators starting with the empty state, and another where we begin with a fundamental cycle of the torus, and take an equal superposition of all configurations of loops that can be created from that cycle. These loops are all distinct from the ones in the other ground state, and in particular, these two ground states are orthogonal. The remaining two ground states are the same, but starting with one of the other two elements of $H_1(T_2; \mathbb{Z}/2)$. Any local operator cannot distinguish between any of these ground states, because locally on some patch of the torus, the configurations look the same.

Interestingly, you can deform away from the exactly solvable model, and the ground state degeneracy will persist.

On a genus-g surface Σ_g, the space of ground states has dimension 4^g, ultimately because one can product 2^g inequivalent cycles in the surface (as this is the order of $H_1(\Sigma_g; \mathbb{Z}2)$). Here, "inequivalent" means that one cannot be transformed into the other by applying local plaquette operators.

Even though local operators cannot distinguish the ground states, some nonlocal operators can. There will ultimately be a close link between the fact that we have particles e and m that don't commute and the ground state degeneracy on the torus.

Let ℓ^x be the fundamental cycle in the x-direction in the torus. Then, given a particle a, we obtain a nonlocal operator ℓ_a^x which creates an a and an $\overline{a}$, then moves them around ℓ^x and annihilates them. This permutes the ground states: ℓ_e^x sends the ground state $|0\rangle$ corresponding to the empty configuration to the state $|\ell^x\rangle$ corresponding to ℓ^x, and similarly $|\ell^y\rangle \mapsto |\ell^x + \ell^y\rangle$.

These operators commute with the Hamiltonian, and satisfy the equation

$$\ell_e^x \ell_m^y = \theta_{e,m} \ell_m^y \ell_e^x. \tag{18.3}$$

Similarly, $\{\ell_m^x, \ell_e^y\} = -1$. Two loops around the same cycle will commute.

If we perturb away from the exactly solvable limit, we'd expect a gapped phase with a finite, nonzero correlation length ξ; the toric code has $\xi = 0$, meaning there's no correlation between the spins on plaquettes that aren't immediately next to each other. Anyways, if the system is smaller than the correlation length, you wouldn't expect the same properties of the ground state degeneracy.

If you create an e excitation on a plaquette, then time-evolve with the Hamiltonian, it stays put. This is a little strange — in the real world, if you take a particle and let it evolve, it will propagate and diffuse out. But there is a perturbation-theoretic argument hopping e and m particles around, the matrix elements connecting the different ground states on the torus are exponentially small in the length. Thus in the limit $L \to \infty$, the ground state degeneracy is preserved, but at finite length, the states are related. But the gap is still preserved, and so we can still talk about the system as having ground state degeneracy, even at finite length.

We've seen that the honeycomb model and the toric code, which are equivalent in the regime when the honeycomb model is gapped, have a unique ground state on the plane, but can have ground state degeneracy on closed surfaces. The different ground states cannot be distinguished by local operators — and as a corollary, there's no way to adiabatically connect the nontrivial ground states (such as $|\ell^x\rangle$) to the trivial state: the trivial Hamiltonian has a unique ground state, and the toric code on a torus doesn't. A more careful argument says that if you do find such a map, you'd have to map the four different states onto different trivial product states — and since adiabatic time evolution acts by a finite-depth unitary circuit, it can't propagate information arbitrarily far, so the local evolution can't know which of the four ground states it's in, and therefore what to assign the ground state to. Thus the toric code is in a nontrivial phase, which we already knew, since there are excitations with nontrivial mutual statistics. One upshot is that there is entanglement in the toric code that cannot be untangled with a finite computation.

Next time, we'll characterize this nonlocal aspect of entanglement with the measures of entanglement entropy that we considered in previous classes.

Lecture 19.

Entanglement and the toric code: 11/5/19

The toric code has a duality in that it doesn't matter whether we switch which colors of the checkerboard map to X and Z operators: you can just translate one square in a cardinal direction.[18] Therefore the Hamiltonians might look different between different classes, but they describe the same system.

Recall that on the plane, the ground state is an equal superposition of all configurations of closed loops. (Well, there are infinitely many such configurations on the infinite plane, which is a problem, so we choose something like the sphere in order to obtain a finite superposition.) On a torus, there is ground state degeneracy, because not all configurations of loops can be created from the empty configuration by local operators; there are four equivalence classes of loop configurations, hence a fourfold ground state degeneracy. This does not depend on the size of the lattice on the torus.

Excitations in the toric code are created by open string operators labeled by Xs or Zs. Applying this string operator to the ground state gives you a superposition of all string and loop configurations beginning and ending at the ends of the string, where we consider all configurations which can be created from the string by plaquette operators.

TODO: after that, an interpretation of how creating pairs of e or m particles is like an insertion of a flux line. I didn't really follow.

[18] This is only true if the coupling constants for the two pieces of the Hamiltonian are equal, as they are when we obtain this model from the honeycomb model.

$\sim \cdot \sim$

Now, let's look at the entanglement structure of the ground state (let's work on the sphere, so we can say "the" ground state). Let's look at a region A large with respect to the edge length but still very far away from the boundary. Entanglement is measured by the extent that we can factor states into a product of states on A and on A^c; the more we can do this, the less entangled the ground state is.

Given an arbitrary state $|\psi\rangle$, one can always perform a Schmidt decomposition

$$|\psi\rangle = \sum_n \lambda_n |\psi_{A,n}\rangle \otimes |\psi_{A^c,n}\rangle. \tag{19.1}$$

Then the entanglement entropy is

$$S_A = -\sum_n \rho_n \log \rho_n, \tag{19.2}$$

where $\rho_n \coloneqq |\lambda_n|^2$. We can also write

$$\rho_A \coloneqq \mathrm{tr}(|\psi\rangle\langle\psi|) = \sum_n \rho_n |\psi_{A,n}\rangle\langle\psi_{A,n}|. \tag{19.3}$$

Let's apply this to the ground state. This is a superposition of loops entirely in A, loops entirely not in A, and loops which cross ∂A, and the last option can happen in different ways. So we decompose the ground state $|\psi_{\mathrm{GS}}\rangle$ as:

- We sum over all the different ways the loops can intersect ∂A. Concretely, we specify the set of points in which the loop configuration intersects ∂A, and sum over all such possibilities. Let S denote the intersection set.
- Inside that sum, we sum over the interior loops which are either closed or terminate at S, then also over the exterior loops which are closed or terminate at S.

This is the Schmidt decomposition, so the key thing we have to do is determine all of the possible boundary conditions S. The key constraint is that $\#S$ is even: all loops in $|\psi_{\mathrm{GS}}\rangle$ are closed, so cannot intersect ∂A an odd number of times, or they couldn't close up. The number of loop configurations for a specific boundary condition is always the same, no matter the boundary condition.

Suppose there are ℓ links exiting the boundary; then ℓ scales with the perimeter of A. Each of the ℓ links could have a link exiting at A. That gives 2^ℓ options, but half of these have S odd, so we throw them out, and the answer is $2^{\ell-1}$. Therefore $\lambda = 1/(2^{\ell-1})$, and therefore

$$S_A = -\sum_{n=1}^{2^\ell - 1} \frac{1}{2^{\ell-1}} \log\left(\frac{1}{2^{\ell-1}}\right) \tag{19.4}$$

$$= \log 2^{\ell-1} \sim |\partial A| \log 2 - \log 2. \tag{19.5}$$

The second term is telling us something topological.

One thing that might seem weird is that under a small petrurbation of the boundary, the coefficient $|\partial A|$ is a finer invariant than the phase: it can change while we're in the same phase. One fixes this by finding a way to leave it out, and focusing on the $-\log 2$, which does not depend on local perturbations.

The way to fix this was discovered at about the same time by Kitaev and Preskill [KP06] and Levin and Wen [LW06]. The idea is to consider multiple different regions and their entropies in such a way that the non-topological terms cancel out. There are different ways to do this; Kitaev and Preskill consider four regions A, B, C, and D as in Figure 4, and define

$$S_{\mathrm{top}} \coloneqq S_A + S_B + S_C - S_{A\cup B} - S_{B\cup C} - S_{A\cup C} + S_{A\cup B\cup C}. \tag{19.6}$$

The point of this definition is that local operators don't change it. For example, let's time-evolve by a Hamiltonian which is a finite sum of local terms:

$$H(t) \coloneqq \sum_\alpha h_\alpha(t). \tag{19.7}$$

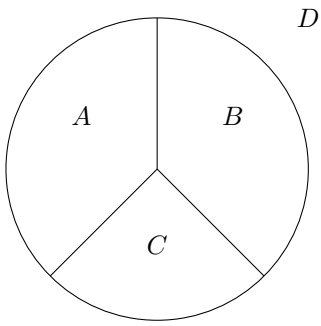


FIGURE 4. Subdividing a large region in the plane to compute topological entanglement entropy, following Kitaev and Preskill [KP06].

Then unitary evolution by time t is

$$U(t) = \exp\left(-i\int_0^t H(t)\right) \approx \prod_{i=1}^{N} \exp(-iH(t_i)\Delta t_i) = \prod_{i=1}^{N}\prod_{\alpha} \exp(-ih_\alpha(t_i)\Delta t). \tag{19.8}$$

There are quadratic corrections hidden in that "≈," but for Δt small enough, this is not a problem. The point is that this is some composition of actions by local unitary operators, so it suffices to consider these operators. Moreover, time-evolution can introduce local correlations, but by the Lieb-Robinson theorem, the correlation length is finite.

The upshot is that we can reduce to considering local operators happening in a small region R around, say, $\overline{A} \cap \overline{B}$ and far from C. This affects the terms S_A, S_B, $S_{A\cup C}$, and $S_{B\cup C}$: it only affects regions where R is at the boundary.

Now, the changes in S_A and $S_{A\cup C}$ are almost exactly the same: under our approximation they're exactly the same. Since S_A and $S_{A\cup C}$ appear in (19.6) with opposite signs, these two changes cancel. The same argument applies to S_B and $S_{B\cup C}$, and therefore S_{top} is unchanged.

The case where R is near the triple junction of A, B, and C is trickier: now S_A, S_B, S_C, $S_{A\cup B}$, $S_{B\cup C}$, and $S_{A\cup C}$ are affected. We would like to know that $\Delta S_A \approx \Delta S_{B\cup C}$, where "≈" is, as before, up to exponentially small corrections. What we do know is that $S_A = S_{A^c} = S_{B\cup C\cup D}$, and therefore

$$\Delta S_A = \Delta S_{B\cup C\cup D} \approx \Delta S_{B\cup C}, \tag{19.9}$$

as desired. In the same way, $\Delta S_B \approx \Delta S_{A\cap C}$ and $\Delta S_C \approx \Delta S_{A\cap B}$. All three of these pairs have opposite signs in (19.6), and therefore S_{top} does not change.

One can argue similarly for the case where R is at the boundary of A and D (or B and D, or C and D), or at the triple point of A, B, and D (etc.), and therefore S_{top} is an invariant of the phase. For example, for the toric code, we'll get $-\log 2$.

One way this is useful is as a way of determining whether some physical system is in a topological phase: it's easier when the system is exactly solvable, but this isn't always true, so you can use the topological entanglement entropy as an easier measure. However, this is not a complete invariant of topological phases: there are examples of distinct phases with the same topological entanglement entropy.

You can also use other entropic quantities, such as the *Rényi entropies*

$$S^{(n)} \coloneqq \frac{1}{1-n}\log(\operatorname{tr}(\rho - A^n)). \tag{19.10}$$

This is not easy to compute in general, but there are some simplifications when $n \in \mathbb{Z}$.

Next time, we'll discuss how to use the toric code to build a fault-tolerant quantum memory (at least in theory), using the ground states as qubits. The trick is error correction, including measuring whether errors are happening without completely collapsing the wavefunction.

Lecture 20.

The toric code as a quantum error-correcting code: 11/7/19

Today, we'll discuss how the toric code is a topological quantum error-correcting code, meaning roughly that it's a topological phase that can (at least theoretically) provide error correction for some systems of quantum computing. Before we dig into the toric-code-specific details, we'll provide an overview of what quantum error correction is and some simpler examples. For the toric code, the idea is that, if we place it on a torus, so there's ground state degeneracy, errors might change what ground state we're in: they can act on the loops in a nonlocal enough way that we could end up with a homologically nontrivial loop, and this is an indication that the system has made an error.

And now on to the overview. Quantum computing is tricky — errors can pop in due to decoherence and noise. Decoherence arises from the fact that it's extremely difficult to separate a quantum system from its environment, leading to situations where a pure state $|\psi_{\text{system}}\rangle \otimes |\psi_{\text{environment}}\rangle$ can under time evolution transform into a mixed state

$$\sum_\alpha \lambda_\alpha |\psi^\alpha_{\text{system}}\rangle \otimes |\psi^\alpha_{\text{environment}}\rangle. \tag{20.1}$$

Partial-tracing out the environment state doesn't solve the problem: we might still be in a mixed state, and not the one we want.

The basic idea of quantum computing is that if we have a system of N qubits (i.e. the state space is a 2^N-dimensional complex Hilbert space) prepared in the state

$$\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes N}, \tag{20.2}$$

a classical computer has to act on all N bits of information simultaneously, but we can act on quantum systems by unitary operators, which will act on all of the qubits simultaneously. This generally results in a superposition, and one then makes a measurement. The end result therefore isn't deterministic, but for some interesting questions, these are algorithms where quantum operations give the right answer with high probability, and much faster (asymptotically, at least) than corresponding classical algorithms, even nondeterministic ones.

The second issue, noise, comes from the difference between theory and reality. First off, whatever particular implementation of the unitary transformation we used might just be a close-enough approximation, and errors can pop up in that way. Second, the actual details of the experiment can cause errors, e.g. we might not have fine enough control of the laser to precisely fix the operator.

At present, decoherence is not a problem: the best systems stay coherent for about 10^6 times as long as the amount of time it takes to implement a gate (basic local unitary). Noise is more of a problem: if you define the *fidelity* of an operation by a unitary U_{gate} to be

$$F \coloneqq \int |\langle \psi_{\text{target}} \mid U_{\text{gate}} \mid \psi_{\text{initial}} \rangle|^2 \, \mathrm{d}\psi_{\text{initial}}. \tag{20.3}$$

Let $\varepsilon \coloneqq 1 - F$; this is a measure of error. For 2-qubit gates, which are basic building blocks for quantum computers, here are the current bounds on ε:

- for quantum computers based on superconducting qubits, $\varepsilon \sim 5 \times 10^{-2}$;
- for quantum computers based on ions, $\varepsilon \sim 1 \times 10^{-3}$.

From the perspective of experimental physics, this is really good — but compare to the fidelity of classical computers, which is orders of magnuitude better. And if we perform a computation with N gates with $N\varepsilon \sim 1$, then there's a very good chance of an error. Current bounds on ε are not good enough to implement interesting computations with negligible chances of error.

We could treat this as an engineering question to hammer away on, but there will always be error, so we could try a different tack: is it possible to make accurate computations in the presence of error? This leads to the idea of error correction.

Even for classical computers, you can imagine putting your computer in, e.g., space, where it can get hit by cosmic rays and make errors. It would be good to know whether an error occurred; one way to do this is to produce redundant data, and then check that it agrees later. This provides an example of an error-correcting code, whose efficacy is measured by three numbers:

- the *number of physical bits*, which is the total number of bits in the system (including the redundancy), denoted n;
- the *number of logical bits*, which is the number of bits that can be encoded, denoted k; and
- the *distance*, i.e. the total number of changes that must be made for the error-correcting bits to change.

Then the maximum number of errors we can detect is $\lfloor (d-1)/2 \rfloor$.

Example 20.4. A checksum is a simple example: given a string of bits, store an extra bit which is the sum of the previous bits. This encodes some redundancy, but if there are two changes, we won't notice, so this isn't the best approach. ◁

Quantum error correction is generally harder. Sometimes people cite the no-cloning theorem, which prevents us from duplicating an arbitrary unknown state, but it doesn't actually stand in the way: we can still build redundancy because we can prepare several copies of a known state (e.g. the initial state of the computation), then run the experiment on all of them. However, the problem arises when we try to compare these bits: a direct comparison, measuring both states, would destroy them. So we have to be craftier.

Example 20.5 (Quantum bit-flip code)**.** Consider a system with two logical qubits $|0_L\rangle$ and $|1_L\rangle$, where the logical qubit is represented by three physical qubits: $|0_L\rangle = |000\rangle$ and $|1_L\rangle = |111\rangle$. Now suppose an error occurs, flipping the first physical qubit, so that we're in state

$$|\psi\rangle = \alpha|100\rangle + \beta|011\rangle. \tag{20.6}$$

We can detect this with some Pauli Z-operators: let Z_i denote the Pauli Z-operator on the first qubit. Then we can measure $Z_1 Z_2 = -1$ without collapsing the outcome of the computation, informing us that an error occurred! We even can tell which bit it occurred on, because $Z_2 Z_3$ cannot see an error. ◁

Some terminology: the operators $Z_1 Z_2$, etc., that look for errors are called *stabilizers*, and the measurement outcome of a stabilizer is called a *syndrome*. Stabilizers are often built from Pauli operators. Generally, the system might be in a complicated state, but using the syndrome, you can project into simpler states.

Example 20.5 as written only detects bit flips; if we want to detect phase flips, we should change basis, writing

$$|0_L\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes 3} \qquad \text{and} \qquad |1_L\rangle = \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)^{\otimes 3}, \tag{20.7}$$

and using pairs of neighboring X operators to detect. If we want to detect both bit flips and phase flips, we should concatenate these two codes, producing a 9-qubit code.

∽ · ∼

Now let's discuss the toric code as a quantum error-correcting code. If you've been absent for the past two weeks, the toric code is a lattice Hamiltonian system, where the lattice is a square lattice with a checkerboard and there are spins on the vertices. The Hamiltonian is -1 times a sum over all black squares of products of Z operators on the edges, and over all white squares of the products of X operators on the edges.

The stabilizers in the toric code will be these products of Z and of X operators on the plaquettes; the Z operators will detect bit flips, and the X operators will detect phase flips. The logical states of the code are the ground states of the toric code; in particular, we consider the toric code on a torus, so there are four different ground states, hence $k = 2$ logical qubits. We've already seen what the syndromes are, albeit in different language: suppose an error introduces a bit flip at a vertex v. Then the black plaquettes adjacent to the vertex are no longer in the $+1$ eigenstate; instead, they're in the -1 eigenstate. We interpreted this as an excitation, producing two particles localized to these two plaquettes; today, we think of these as error syndromes.

And, as we've seen, if there are multiple errors, there are a few different possibilities: two bit flips could produce two particles separated by more than one plaquette. In general, the error syndromes are boundaries of error strings; the bit flips are detected by the Z stabilizers, and the phase flips by the X stabilizers.

Another potential that can happen is a closed string error, e.g. flipping the bit on all vertices in a plaquette. This is not a problem, though, because it's in the same ground state. Similarly, if there is an open string error, we can only see the ends of the string, where the particles are. If we apply any other string in the same

topological equivalence class, we correct the error, as composing the strings leads to a topologically trivial loop, which preserves the ground states.

It's possible that we guess wrong, and apply a string operator which, say, wraps around the torus. Maximal confusion would occur for a string half the length of the torus. In practice, large string errors are less likely, which is good. If the torus has the same length L in both directions, this implies the code distance is L, and the maximal number of one-qubit errors we can correct is about $L/2$: once strings get to about half the length of the torus, we get confused.

In summary, this is an $(L^2, 2, L)$-code: L^2 physical qubits encode 2 logical qubits, with code distance L.

Digging in, how do you actually use the toric code to correct errors?

(1) First, prepare a ground state $|\psi_{\text{GS}}\rangle$, which is the result of some quantum computation.
(2) Then, measure the stabilizers everywhere, and infer the mostly likely error strings that caused this, or at least the topological equivalence classes of these error strings.
(3) Then, apply correction strings in the same topological equivalence classes of the error strings.

Then, repeat: compute, look for errors, fix them. But each step has a few additional things to think about. For example, when we measure the stabilizers, we somehow need to measure four qubits at once. One failure could cause simultaneous errors on all four qubits. In particular, unless you do something clever with your circuit design, errors occur on 4 qubits at once, which means that actually $L/4$ errors are needed to corrupt your computation, rather than $L/2$.

Also, step (2) is difficult: finding the maximally likely set of (equivalence classes of) strings describing an error configuration is in general exponentially hard. But if you're willing to accept a "close enough" solution, there are several different good algorithms, relating the problem to, e.g., graph matching. One such algorithm is called the *blossom algorithm*. Crucially, these algorithms are both efficient and parallelizable — error detection must be fast, because the quantum system is still evolving while we're checking. People therefore are working on finding algorithms which have greater upfront complexity, but run quickly during the computation, such as neural net-based approaches. This is an active area of research.

Another issue is that measurements of states could be wrong — there's always a small chance of a false positive. Some of this amounts to implementation details, e.g. if your qubit is given by whether an atom is in a ground state or an excited state, which you detect via an emitted photon, what happens if you, for example, miss the photon? There are other ways this system can fail, and no system is infalliable. The standard way to fix this is to make redundant measurements, but the clock is still running, and there could be additional errors in between measurements. This is OK, in that things still work, but this degrades the code performance significantly.

Suppose the physical qubits can fault with independent and identical probabilities ϵ. Then, the odds of the system failing is $\epsilon^{L/2}$ times some polynomial in L. If ϵ is small and $L \to \infty$, then this is good: the probability of failure is small. But as ϵ grows, there's a phase transition to a regime where making a bigger system doesn't help. So there are two possibilties.

- If $0 < \epsilon \ll 1$, errors are rare, and with high probability, the errors are on short strings relative to L, and are easy to correct.
- If $0 \ll \epsilon < 1$, e.g. $\epsilon = 1/2$, there are lots of errors, and guessing the strings is much harder, making it more likely that we fail to fix the errors.

There is a sharp transition between these two regimes, and the logical parameter will act like a parameter in a continuous phase transition, which is in the same universality class as a 2d random-bond Ising model, though when we account for repeated measurements, we get a 3d model (two space dimensions, and one time dimension); then, the model is in the same universality class as a $\mathbb{Z}/2$ gauge theory. In a finite-size system, we get an approximation to this curve, and as $L \to \infty$, the approximation approaches the actual curve. The specific details of the curve, and the value of ϵ at which the transition occurs, depends on implementation details: the algorithm used to correct errors, the time taken between repeated measurements, etc. But the power laws governing how the error rate behaves near the phase transition are independent of all of this, and this is important if you want an accurate error-correcting code.

For simple models, if we do not account for measurement errors, $\epsilon \approx 0.1$. If we do account for measurement errors, $\epsilon \approx 0.001$. And the toric code has several other implementation advantages: there's a relatively easy way to simulate it on a torus using certain boundary conditions on flat space, and it scales nicely. The key disadvantage is that it only has two qubits; there are competing codes with greater numbers of logical bits.

Thinking of the toric code as an error-correcting code solves another mystery about the toric code: its Hamiltonian looks kind of bizarre, the kind of thing that wouldn't appear in a physical system. But for quantum error correction, that's okay; we can just project into the ground state using the stabilizers. This is not true for all phases, e.g. it doesn't apply for the integer quantum Hall effect phases.

In the last few minutes, let's discuss how to implement gates on the qubits. For example, the logical X operator for a qubit can be implemented by a string of X operators along a homologically nontrivial cycle; the logical Z operator for the same qubit is implemented by the dual (i.e. perpendicular) string for Z operators. This is hard, in that we have to rotate L qubits at once; we could make a very long circuit to entangle the physical qubits in question, but lots of gates means lots of opportunities for errors to emerge.

The set of operations we can apply without issue, which are sometimes called fault-tolerant operations, is relatively simple in the toric code, and not enough to enable arbitrary computation. The solution is to use a different code which has its own advantages and disadvantages; then, one combines it with the toric code in a clever way. That's a nice thing, but when one crunches the numbers, the problem is that one needs thousands of physical qubits to encode one logical qubit, and that is not really workable. So implementing this in practice is quite far-off, and trying to find better systems that are more tractable is an active subject of research.

Lecture 21.

Boundary conditions in the toric code: 11/12/19

Note: I was 15 minutes late to class, so some of the beginning material is missing. Sorry about that.

Today we're studying boundary conditions for the toric code. The idea here is that if you want to formulate the toric code on a finite-area checkerboard lattice, you'll need to choose boundary conditions on each component of the boundary. To understand this carefully, we'll have to figure out what to do with the lattice at the boundary — we want all of the vertices to be tetravalent, but at the boundary, *a priori* they're only trivalent. The solution is to add an extra edge between neighboring vertices. There are two ways in which we can do this:

- We can add this extra edge above all the black squares, and none over the white squares. Then the extra edges are labeled with Z operators.
- We can add this extra edge above the white squares but not the black squares; then the extra edges are labeled with X operators.

These create two kinds of boundary conditions. The first kind is the *e boundary condition*, because it modifies the allowed Z-strings in the ground state. Specifically, the ground state is a superposition not just of closed loops in the interior, but also open strings which terminate at the boundary. All of these loops are topologically trivial, so we obtain a unique ground state.

Correspondingly, the second kind of boundary condition is called the *m boundary condition*, and it modifies the ground state in the analogous way; the ground state is again unique.

If you think back to earlier in the class, something weird is going on: we proved, under some assumptions, that any nontrivial topological phase has nontrivial edge modes, but the toric code doesn't. The resolution is that one of our assumptions was that there was a unique ground state under any topology. The toric code, of course, has ground state degeneracy on a torus, and more generally on any closed, connected surface other than the sphere. So the criterion on edge modes is really about invertible phases.

Another less fun consequence of the unique ground state on the plane is that it's much harder to use the toric code for quantum error correction — most actual realizations of qubits are extremely hard to put on something like a torus, and the things that make quantum error correction work on the torus don't apply to the plane. That's a little sad, but we still have anyonic excitations, which is cool.

Now let's consider a slightly different system: put the toric code on a large square, but use e boundary conditions on the north and south boundaries, but m boundary conditions on the east and west boundaries. Now, there are some more interesting string operators: a string of Z operators from the north to the south boundary, and a string of X operators from the east to the west boundary. These two operators both commute with the Hamiltonian and anticommute with each other, and, crucially, they can't be contracted, because you'd have to cross the other kind of boundary, and adding small loops can't work around that. This implies that, in this system, there is ground state degeneracy: we can't have two nontrivial and anticommuting

operators on a one-dimensional vector space. Indeed, there is a twofold ground state degeneracy, and defects on the four corners of the square. This is depicted in Figure 5.

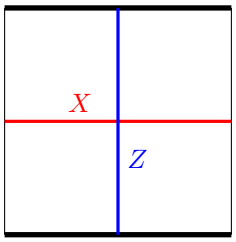


FIGURE 5. A picture of the toric code with e boundary conditions on the thick edges and m boundary conditions on the thin edges. The red line indicates an X string operator, and the blue line as a Z string operator.

But we don't have to stop here — you could consider a hexagon, coloring alternate edges with e and m boundary conditions. In this case, there are six defects, at the corners of the hexagon, and four string operators: two Z strings between two of the three pairs of e boundaries and two X strings between two of the three pairs of m boundaries. (The remaining pairs of edges can be obtained from superpositions of the other two pairs.) Thus there is a fourfold ground state degeneracy.

You can generalize to an n-gon as long as n is even; in this case there are n defects at the vertices and the ground state degeneracy is $2^{n/2-1}$. This means that you could try to build a quantum error-correcting code using one of these boundary conditions. In this case, the excitations behave a little like Majoranas, which is reminiscent of the honeycomb model that the toric code emerged from.

Remark 21.1. It's possible to describe the toric code's phase as an abelian Chern-Simons theory, with gauge group $\mathrm{U}_1 \times \mathrm{U}_1$ and Lagrangian

$$\mathcal{L} = ia_I \cdot \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}_{IJ} \frac{\mathrm{d}a_J}{4\pi}. \tag{21.2}$$

The e and m particles correspond to the two factors of U_1, and the pairing between them in the K-matrix leads to their nontrivial mutual statistics. ◁

Kitaev's honeycomb model is more than just the gapped phase that gave us the toric code; there's another gapless phase. We can introduce a magnetic field and obtain another gapped phase, which will turn out to contain nonabelian anyons. In this limit, the Hamiltonian looks like

$$H = -\sum_{\alpha \in x,y,z} J^\alpha \sum_{\langle ij \rangle \in \alpha} \sigma_i^\alpha \sigma_j^\alpha = -\sum_\alpha J^\alpha i \gamma_i s_{ij} \gamma_j. \tag{21.3}$$

This has a translation symmetry, which allows us to take a Fourier transform and obtain a Brilloun zone and a band structure. The way this works is that the unit site for this translation invariance is a pair of vertices of the hexagon; see Figure 6. Explicitly, the translation symmetry is generated by the vectors

$$\mathbf{a}_1 = \left(\frac{\sqrt{3}}{2}, \quad \frac{1}{2} \right) \tag{21.4a}$$

$$\mathbf{a}_2 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2} \right). \tag{21.4b}$$

We end up getting that

$$\widetilde{\gamma}_k = \sum_{\mathbf{R}_i} \frac{e^{i\mathbf{k}\cdot\mathbf{R}_i}}{\sqrt{\mathrm{vol}}} \begin{pmatrix} \gamma_{A,i} \\ \gamma_{B,i} \end{pmatrix}, \tag{21.5}$$

so we can rewrite the Hamiltonian (21.3) as an off-diagonal operator:

$$H = \sum_k i \widetilde{\gamma}_{-k}^{\mathrm{T}} \begin{pmatrix} 0 & f(k) \\ -f(k) & 0 \end{pmatrix} \widetilde{\gamma}_k, \tag{21.6}$$

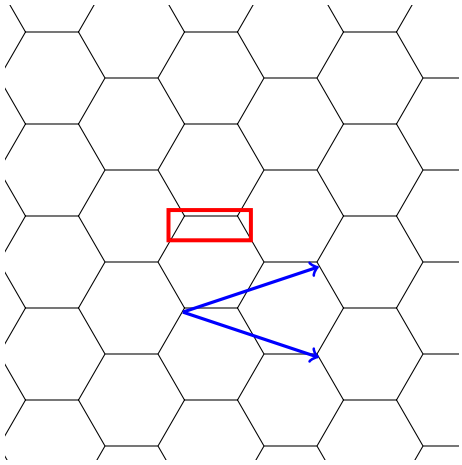


FIGURE 6. Red: the basic unit site for the translation symmetry of the Hamiltonian Equation (21.3) associated to the honeycomb model. Blue: the generating translation symmetry of (21.4).

where

$$f(k) = \frac{1}{2}\left(J_y e^{i\mathbf{k}\cdot\mathbf{a}_1} + J_z e^{i\mathbf{k}\cdot\mathbf{a}_2} + J_x\right). \tag{21.7}$$

Here the $\gamma_{A,i}$ and $\gamma_{B,i}$ are Hermitian and $\widetilde{\gamma}^{\mathrm{t}}_{-k} = \gamma^{\dagger}_{k}$. The eigenvalues are

$$E_k = \sqrt{|f(k)|^2}. \tag{21.8}$$

As we discussed, whenever one of J_x, H_y, or J_z is much bigger than the rest, we're in the same phase as the toric code. But when none dominates the others, we're in a gapless phase. This can be understood from a perspective similar to one appearing in the study of graphene (but TODO: I missed this perspective). There are values $\pm K$ of k with zero energy, which will (TODO I think) correspond to massless Majoranas.

In general, whenever something is forbidden, there's an underlying symmetry. We've used the translation symmetry, but there are other symmetries. For example, since the basic site consists of two underlying sites, we have a time-reversal symmetry exchanging them, and the Majorana points are protected by a combination of time-reversal and translation symmetries.

Next time, we will discuss what can happen when we perturb (21.3) while preserving the time-reversal symmetry. We'll see that this just shifts around the Dirac cone, and cannot give the Majorana modes a mass. Thus, to obtain a gapless topological phase, we'll have to break the time-reversal symmetry, which we will accomplish with a magnetic field.

References

[AKLT88] Ian Affleck, Tom Kennedy, Elliott H. Lieb, and Hal Tasaki. Valence bond ground states in isotropic quantum antiferromagnets. *Comm. Math. Phys.*, 115(3):477–528, 1988. 15

[Hal83a] F. D. M. Haldane. Nonlinear field theory of large-spin Heisenberg antiferromagnets: Semiclassically quantized solitons of the one-dimensional easy-axis Néel state. *Phys. Rev. Lett.*, 50:1153–1156, Apr 1983. 15

[Hal83b] F.D.M. Haldane. Continuum dynamics of the 1-D Heisenberg antiferromagnet: Identification with the O(3) nonlinear sigma model. *Physics Letters A*, 93(9):464–468, 1983. 15

[Hal85] F. D. M. Haldane. "θ physics" and quantum spin chains (abstract). *Journal of Applied Physics*, 57(8):3359–3359, 1985. 15

[Kit01] A. Yu. Kitaev. Unpaired Majorana fermions in quantum wires. *Physics-Uspekhi*, 44(10S):131, 2001. https://arxiv.org/abs/cond-mat/0010440. 5

[Kit06] Alexei Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321(1):2–111, 2006. January Special Issue. https://arxiv.org/abs/cond-mat/0506438. 32

[KP06] Alexei Kitaev and John Preskill. Topological entanglement entropy. *Phys. Rev. Lett.*, 96:110404, Mar 2006. https://arxiv.org/abs/hep-th/0510092. 41, 42

[LW06] Michael Levin and Xiao-Gang Wen. Detecting topological order in a ground state wave function. *Phys. Rev. Lett.*, 96:110405, Mar 2006. https://arxiv.org/abs/cond-mat/0510613. 41

[TKNdN82] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs. Quantized Hall conductance in a two-dimensional periodic potential. *Phys. Rev. Lett.*, 49:405–408, Aug 1982. 30

[Ton16] David Tong. Lectures on the quantum hall effect. 2016. http://www.damtp.cam.ac.uk/user/tong/qhe.html. 27