Physics 230: Quantum phases of matter
Spring 2024

Lecturer: McGreevy
These lecture notes live here. Please email corrections and questions to mcgreevy at physics dot ucsd dot edu.

Last updated: July 17, 2024, 14:16:35

Contents

0.1 Introductory remarks and goals

The study of phases of matter is a topology problem. Let me explain. We wish to divide the set

$$\{\text{macroscopic piles of stuff, with some interactions}\}$$

into equivalence classes. The equivalence relation is roughly: two interacting piles of stuff are regarded as being in same phase if their observable properties are adiabatically connected under varying the interactions and adding in more non-interacting, gapped stuff. So phases of matter are essentially elements of $\pi_0(\text{piles of stuff})$. [1]

A *topological invariant* is a quantity that does not change under such continuous variations, for example a quantity that is guaranteed to be an integer. Such invariants are wonderful because they provide labels on our equivalence classes. The simplest example of a topological invariant labelling a phase of matter is the (integer!) number of groundstates of an Ising magnet: it is 2 in the ordered phase, and 1 in the disordered phase; thus these two phases must be distinct. So you see that the use of topology in condensed matter physics is not just for 'topological phases'.

Topological phases are those that are distinguished from others, say from the trivial state, by properties other than ordinary symmetry breaking. (A good representative of the trivial state is an *atomic insulator*, where each particle is 600 miles from its nearest neighbor and never even says 'hello'. More generally, the trivial phase is one that has a product state representative that breaks no symmetries.) By now there is a large variety of known ways in which phases can be topological, some of which are pretty fancy mathematically. Some of them have even been found in Earth rocks. My main goal in this course will be to try to explain some of these phenomena, and the topologically-invariant labels we can attach to them, as concretely as possible.

To do this, it will occasionally be necessary to use some mathematics. A few years ago, I taught a course whose purpose was to develop some tools of algebraic topology, at least some such tools that are realized in toy models of physical systems. I will not assume this background. I am going to do my best to make our discussion here self-contained, while not making this a math course. At times I may have to ask you to do some extra background reading or to take some statements on faith. As a hopefully-useful resource, I've posted a summary of the mathematical highlights; I don't expect you to absorb every detail of this, but rather to use it as a resource as needed.

Likely, not everyone taking this class took the previous two quarters of the condensed matter series. If it seems like I am assuming some knowledge you don't have, please do not hesitate to ask.

[1] π_q of this space for $q > 0$ is also interesting but much less well-explored so far.

For a list of topics we might cover, see the table of contents of this document, or this administrative handout. There is a lot one could say on this subject and we will have to make some selections. Your input is encouraged.

0.2 Conventions

For some of us, eyesight is a valuable and dwindling commodity. In order not to waste it, I will often denote the Pauli spin operators by

$$\mathbf{X} \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \mathbf{Y} \equiv \begin{pmatrix} 0 & -\mathbf{i} \\ \mathbf{i} & 0 \end{pmatrix} \quad \mathbf{Z} \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(rather than $\sigma^{x,y,z}$, which hides the important information in the superscript) in the Z basis. I'll write $|0\rangle$, $|1\rangle$ for the Z eigenstates, $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$ and $|\pm\rangle$ for the states with $X|\pm\rangle = \pm|\pm\rangle$.

I use ijk for spatial indices, $\mu\nu\rho$ for spacetime indices. d is the number of space dimensions and $D = d + 1$ is the number of spacetime dimensions (it's bigger).

$\equiv$ means 'equals by definition'. $A \stackrel{!}{=} B$ means we are demanding that $A = B$. $A \stackrel{?}{=} B$ means A probably doesn't equal B.

The convention that repeated indices are summed is always in effect unless otherwise indicated.

A useful generalization of the shorthand $\hbar \equiv \frac{h}{2\pi}$ is

$$đk \equiv \frac{\mathrm{d}k}{2\pi}.$$

I will also write $\not\delta(q) \equiv (2\pi)^d \delta^d(q)$.

I try to be consistent about writing Fourier transforms as

$$\int \frac{\mathrm{d}^d k}{(2\pi)^d} e^{ikx} \tilde{f}(k) \equiv \int đ^d k \; e^{ikx} \tilde{f}(k) \equiv f(x).$$

WLOG $\equiv$ without loss of generality.
IFF $\equiv$ if and only if.
RHS $\equiv$ right-hand side. LHS $\equiv$ left-hand side. BHS $\equiv$ both-hand side.
IBP $\equiv$ integration by parts.
$+\mathcal{O}(x^n) \equiv$ plus terms which go like x^n (and higher powers) when x is small.
iid $\equiv$ independent and identically distributed.

We work in units where $\hbar$ and k_B are equal to one unless otherwise noted.

Please tell me if you find typos or errors or violations of the rules above.

0.3 Anticipated Sources

This list will grow with the notes.

N. Mermin, *The Topological Theory of Defects in Ordered Media*. The classic account of the use of homotopy theory to understand excitations of ordered media that, because of topology, cannot disappear.

G. Volovik, *Exotic properties of superfluid ^3He*. A nice discussion of the Landau-Ginzburg theory of this remarkable system.

R. Moessner, J. Moore, *Topological Phases of Matter*, Cambridge, 2021. I just got a copy of this book and it seems to have some nice things in it, with some overlap with our goals.

S. Girvin, K. Yang, *Modern condensed matter physics*, Cambridge, 2019.

M. Nakahara, *Geometry, Topology and Physics*. I was not a big fan of this book when I was a student because I thought it was superficial. Looking at it again now, I see its virtues more clearly. It has useful things in it and it is mostly written for physicists.

Nash and Sen, *Geometry and Topology for Physicists*. This book has the virtue of brevity.

D. Tong, *Lectures on the Quantum Hall Effect*.

D. Arovas, Lecture Notes on Quantum Hall Effect (A Work in Progress)

X.-G. Wen, *Quantum Field Theory of Many-Body Systems*, Oxford, 2004.

X.-G. Wen, *Topological orders and Edge excitations in FQH states*.

A. Zee, *Quantum Hall Fluids*.

G. Moore, *Quantum Symmetries and Compatible Hamiltonians*.

G. Moore, *Introduction to Chern-Simons Theories*.

J. Harvey, *Lectures on Anomalies*.

E. Witten, *Three Lectures On Topological Phases Of Matter*.

E. Witten, *Fermion Path Integrals And Topological Phases*.

A. Turner and A. Vishwanath, *Beyond Band Insulators: Topology of Semi-metals and Interacting Phases*.

T. Senthil, *Symmetry Protected Topological phases of Quantum Matter*.

C. Z. Xiong, *Classification and Construction of Topological Phases of Quantum Matter*.

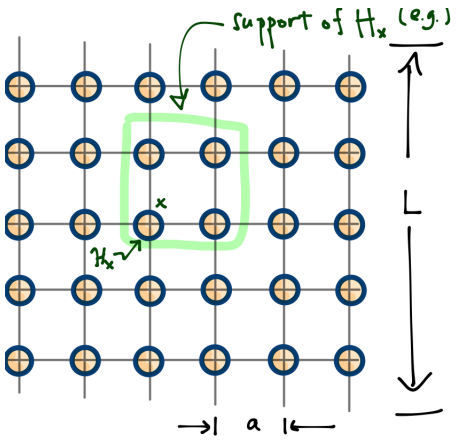
1 Introductory remarks about quantum matter

This is an introductory section on basic notions of quantum phases of matter, and how to think about them in terms of entanglement. For further reading, I suggest:

- Zeng, Chen, Zhou, Wen, *Quantum Information Meets Quantum Matter: From Quantum Entanglement to Topological Phase in Many-Body Systems*
- These TASI lectures.
- For more on applications of quantum information theory to many body physics from a similar viewpoint, see the lecture notes and problems posted here.

1.1 States of matter, classified by level of desperation

In this class we are going to talk about *extensive quantum systems*. A quantum system can be specified by its Hilbert space and its Hamiltonian. By the adjective *extensive* I mean that the Hilbert space is defined by associating finite-dimensional Hilbert spaces $\mathcal{H}_x$ to chunks of *space*, labelled by[2] some coordinates x. Then couple them by a local Hamiltonian, $H = \sum_x H_x$, where H_x acts only on the patch at x and not-too-distant patches (and as the identity operator on the other tensor factors in $\mathcal{H}$).



The phenomena whose study we will find most fulfilling only happen in the *thermodynamic limit*, where the number of patches grows without bound. I will use L to

[2] We can allow the local Hilbert space to be infinite-dimensional (as for rotors or bosons) if we add terms to the Hamiltonian that leave a finite-dimensional set of low-energy states. In practice, when simulating such systems for example, we can always truncate the Hilbert space to some finite value.

denote the linear size of the system. For a cubic chunk of d-dimensional hypercubic lattice, there are $\left(\frac{L}{a}\right)^d$ patches, where a is the size of the patches. So the thermodynamic limit is $L \to \infty$, or more precisely $L \gg a$. In the mysterious first sentence of this paragraph, I am referring to *emergent* phenomena: qualitatively new effects which can never be accomplished by small systems, such as spontaneous symmetry breaking (magnetism, superconductivity, the rigidity of solids), phase transitions, topological order, and all the other things we have not thought of yet because we are not very smart.[3] [4]

Perhaps the most basic question we can ask about such a system is: how many degrees of freedom are there at the lowest energies (lower than any interesting scale in the problem, in particular in the Hamiltonian)? By degrees of freedom (dofs) I mean excitations that can be created by a local operator, as in an experiment where we scatter particles (neutrons, photons...) off the material. There are essentially three possibilities:

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

As we proceed down this classification, our level of understanding rapidly decreases.

A more informative tour through that list goes like this. To get started let me make the assumption that the system has (at least discrete) translation invariance, so we can label the excitations by momentum.

1. **None:** Such a system has an energy gap ('is gapped'): the energy difference $\Delta = E_1 - E_0$ between the first excited state and the groundstate is nonzero, even in the thermodynamic limit. Note that Δ is almost always nonzero in finite volume. (Recall, for example, the spectrum of the electromagnetic field in a box

[3] In case you doubt that characterization, ask yourself this: How many of the items on this list were discovered theoretically before they were found to occur in Earth rocks by our friends who engage in experiments? The answer is **none**. Not one of them! Let us be humble. On the other hand: this is a source of hope for more interesting physics, in that the set of Earth rocks which have been studied carefully so far is likely to represent a very small sample of the possible emergent quantum systems.

[4] Can you think of other elements I should add to this list? One possibility (thanks to Ibou Bah for reminding me) can be called *gravitational order* – the emergence of dynamical space (or spacetime) (and hence gravity) from such ordinary-seeming quantum systems. The best-understood example of this is AdS/CFT, and was discovered using string theory. I was tempted to claim this as a victory for theorists, but then I remembered that we discovered gravity experimentally quite a while ago.

of linear size L: $E_n \sim \frac{n}{L}$.) The crucial thing here (in contrast to the case of photons) is that this energy stays finite even as $L \to \infty$.

The excitations of such a system are generally massive particles[5].

Actually, it is useful to allow a finite number of states below the gap (which are split by something of order $e^{-\alpha L}$ and therefore rapidly become degenerate in the thermodynamic limit) in our definition of a gapped system.

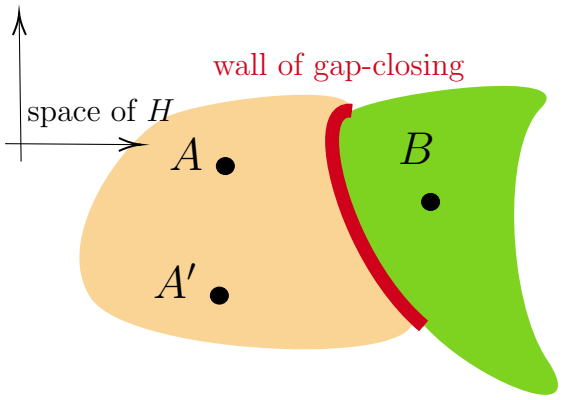
2. **Some:** An example of what I mean by 'some' is that the system can have excitations which are massless particles, like the photon.

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

   In this category I also put the gapless fluctuations at a critical point. It's not necessarily true that $\omega \sim k^{\text{integer}}$ and those excitations are not necessarily *particles*. But they are still at $k = 0$[6].

3. **A lot:** What I mean by this is things like Fermi surfaces, where there are excitations on a whole codimension-one locus in momentum space. This includes not just free fermions or adiabatic continuations of free fermions (Landau Fermi liquid theory). Such systems exist, for example in the half-filled Landau level and in the strange metal regime of cuprate superconductors.

Let's reconsider the case of gapped systems and define the notion of a quantum phase. Different gapped states are in different phases if we can't deform the Hamiltonian to get from one to the other without closing the gap. So a gapped phase is an equivalence class of Hamiltonians. In the figure at right, $[A] = [A'] \neq [B]$.



[5] Verstraete et al proves a version of this statement. I think it is worth looking for loopholes here.
[6] or some other isolated points in momentum space.

[78]You might be bothered by the following: it is hard to imagine checking that there is no way around the wall of gaplessness. It is therefore important to find sharp characterizations of such states, like integer labels, which cannot change smoothly. This is the very definition of topology. An important goal in condensed matter physics is to figure out labels that can be put on states which can distinguish them in this way as distinct phases of matter.

Here are two classes of examples (even in the absence of symmetry) of topological labels. These are the two main ways in which a state of matter can be topological: protected edge states and topological order. Fractional quantum Hall systems exhibit both of these phenomena, while integer quantum Hall states only exhibit the former.

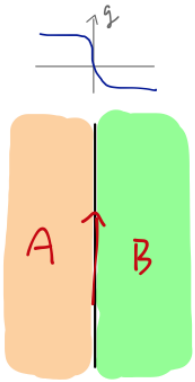
1) Edge modes. There exist nontrivial topological phases without topological order. Another (distinct!) possibility is that even if the system in infinite space has an energy gap, if we cut the space open, new stuff can happen; for example there may be gapless edge modes. There is a class of topological states which are characterized by their edge modes, generally called *invertible states* (special cases include Symmetry-Protected Topological states (SPTs) and Topological Insulators). The edge modes of such a state must carry some property which cannot be recreated locally on the surface; such a property is called an *anomaly*[9].

[7]Actually, there is an important extra equivalence relation that we must include: We don't care if on top of some nontrivial phase of matter someone sprinkles a dust of decoupled qubits which are totally inert and do nothing at all. This modification represents the same phase of matter. (A good example to keep in mind is the K-shell electrons that sit passively next to the nucleus while the conduction electrons do the hard work of forming some strongly-correlated phase.) Then, further, we are allowed to adiabatically deform the hamiltonian, including these decoupled bits, so that they can interact with the original degrees of freedom. So: in addition to allowing adiabatic variation of couplings, we also allow the addition of decoupled qubits.

[8]Note that the closing of the gap does not by itself mean a quantum critical point: at a first order transition, just the lowest two levels cross each other.

[9]This definition of SPT as a state characterized by its anomalous edge modes may be imperfect. There are examples of distinct states protected by lattice symmetries which seem not to have interesting edge modes. See appendix A of this paper or this paper. Thanks to Mike Hermele for bringing this phenomenon to my attention. A related exception would seem to be "higher-order topological insulators," which have no edge states, but have excitations localized to sharp corners of the boundary.

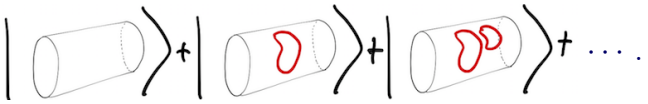
A reason to think that an interface between the vacuum and a gapped state of matter which is distinct from the trivial one might carry gapless modes is that the couplings in the hamiltonian are forced to pass through the wall where the gap closes. In fact there are important exceptions to this conclusion. For example, the toric code admits gapped boundaries (interface with vacuum).
We'll come back to the possibility of distinguishing phases of matter by their edge states in §3.



2) Topological Order. Even the lowest-energy (even below the gap) physics of gapped systems can be deeply fascinating. For example, it may be that the *number* of groundstates depends on the topology of the space on which we put the system. Since this is an integer, it cannot vary continuously and can only jump when the gap closes. This is a symptom of the phenomenon is called *topological order*[10].

As an example of a state with topological order, consider the toric code, aka $\mathbb{Z}_2$ gauge theory. A representative groundstate wavefunction (the fixed point one) locally has the form

$$|\mathrm{gs}, 0\rangle = \sum_{\text{closed loops},C} |C\rangle =$$



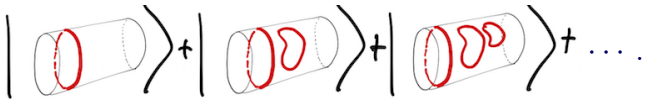
Here it is useful to think of the degrees of freedom as qubits living on the links of a lattice; a basis of the Hilbert space of each link is labelled 'yes string' or 'no string', and accordingly we do or do not draw a little red line segment over the link. The (local) Hamiltonian picks out only configurations where these links form closed loops, and picks out a uniform superposition of such loops. (We can talk about a specific form of such a Hamiltonian later.)

I say 'locally' because on a space with non-contractable loops (like a cylinder), we get orthogonal groundstates by including (or not) the loops that wind around the non-contractible cycles. In the picture above, I tried to indicate that space is a cylinder,

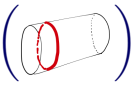
[10] In order for this degeneracy to be stable, it must be that no local operator maps one of these groundstates to another. Suppose our hamiltonian H_0 happens to annihilate two states $|\psi_{1,2}\rangle$ (*i.e.* WLOG, set the groundstate energy to zero). If we perturb H with any ΔH such that $\langle\psi_1|\,\Delta H\,|\psi_2\rangle$, the degeneracy will be split – the levels repel each other.

This property makes the groundstate subspace of a system with topological order into a quantum error-correcting code, with a *code distance* (the number of errors that can be corrected) of order the system size L.

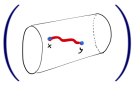
and I included only contractable loops in the sum. An orthogonal groundstate is

$|\mathrm{gs}, 1\rangle =$  $+ \ldots$.

So the number of groundstates depends on the topology of space. States in these different sectors are only related by an operator that creates a whole large loop winding around the nontrivial cycle, hence not by any local operator:

$$|\mathrm{gs}, 1\rangle = \mathcal{O}\left(\text{  }\right)|\mathrm{gs}, 0\rangle$$

The anyonic excitations arise by allowing the loops to end,

$$|\text{anyons at } x \text{ and } y\rangle = \mathcal{O}\left(\text{  }\right)|\mathrm{gs}\rangle$$

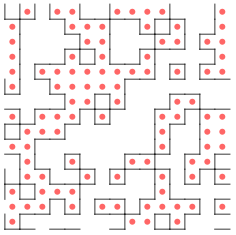
(where we can act on any of the groundstates). Such a state can arise as the groundstate of a spin system – it is a (gapped) *spin liquid* state.[11]

The low-energy physics of gapped phases is often described by a (unitary) topological field theory; this is a theory of groundstates, and it can provide a way to distinguish states of matter. When this is case, the phase is called a *liquid*. Exceptions include fracton topological phases, where the lattice is not forgotten by the topological groundstates.

So topological order means a robust spacetime-topology-dependent groundstate degeneracy. Associated with this phenomenon is also a fractionalization of the quantum numbers of the microscopic constituents. That is, the emergent quasiparticle excitations carry quantum numbers (statistics, spin, charge) which are rational fractions of those of the constituents. Particles of fractional spin are called *anyons*. In three space dimensions, there are also string excitations with robust and interesting fractional properties. In the example of the toric code above, although the microscopic degrees of freedom are all bosonic (just spins), there is an excitation that is a fermion,

[11] To see that topological order is a subtle thing that one might miss if asking the wrong questions, consider the following model of spins at the sites of the square lattice: $\mathbf{H} = \sum_i X_i$, where X_i is the Pauli σ^x operator.

Its groundstate, written in the σ^z eigenbasis, is (the product state) $\otimes_i |\rightarrow\rangle_i \propto \prod_i \sum_{s_i=\uparrow,\downarrow} |\{s_i\}\rangle$. We can also visualize this as a sum over closed loops: draw a loop around each region of $s_i = \downarrow$ (the red dots in the figure at right). They are closed loops because they are the boundaries of a region. Isn't this a toric code groundstate? No! One way to see the difference is that since these loops are defined as the boundaries of regions, they are always contractable.



and there are excitations that are mutual semions, meaning that they get a minus sign when braided around each other.

These two symptoms of topological order are not independent. A better way to define anyons (or more generally topological excitations) is: localized excitations that cannot be created by any local operator. But then these extended operators that create the anyons can do other things, too. That is, the fractional statistics of the quasiparticles implies a groundstate degeneracy on *e.g.* the torus: Pair-create a quasiparticle-antiquasiparticle pair, move them around a spatial cycle, and then re-annihilate them. This process is accomplished by an operator $\mathcal{F}_x$ maps one groundstate to another. But $\mathcal{F}_x$ does not commute with some $\mathcal{F}_y$ (an analogous operator around the other direction), by the anyonic statistics. The space of groundstates must represent the algebra of these operators. Conversely, a *robust* groundstate degeneracy requires that the different groundstates are related by the action of non-local operators. In cases where the groundstate degeneracy is independent of geometry[12], these non-local operators describe the transport of fractionalized excitations around the cycles of the space.

[End of Lecture 1]

1.2 Toric code

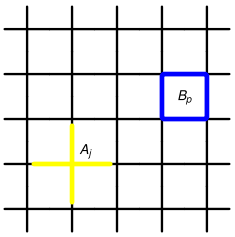
Here's the toric code. It emerges $\mathbb{Z}_2$ gauge theory from a local Hilbert space. There is a sense in which it exists in certain forms of artificial condensed matter (cold atoms in optical lattices, trapped ions).

[12] In contrast, in type-2 fracton models, the operators taking one topological groundstate to another are supported on fractals, and so are hard to interpret as transporting anyons. In such models the groundstate degeneracy depends (in a complicated way) on the system size, and not just on the topology.

Consider a 2d cell complex. This means a graph (a set of vertices who know with whom they share an edge) with further information about plaquettes, who know which edges bound them):

$$\text{plaquettes} \xrightarrow{\partial} \text{edges} \xrightarrow{\partial} \text{vertices} \tag{1.1}$$

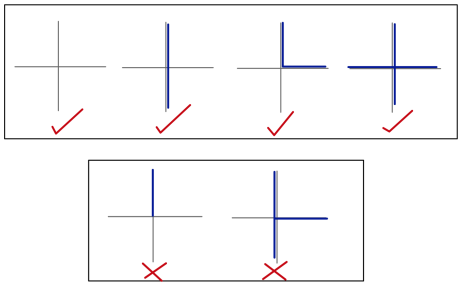
where ∂ is the boundary map, and I really mean formal linear combinations of these objects (we'll see a natural quantum mechanical realization of these linear combinations). For example, consider the square lattice at right. Now place a qubit on each *edge*. Now let's make the terms in the Hamiltonian. Associate to each plaquette a plaquette operator or 'flux operator', $B_p = \prod_{\ell \in \partial p} Z_\ell$, and to each vertex a star operator or 'gauss law operator', $A_v = \prod_{\ell \in \partial^{-1} v} X_\ell$. (The former names just describe the support of the operators on the graph. The latter names are natural if we consider Z to be related to a gauge field by $Z \sim e^{\mathbf{i}A}$, and X is its electric flux. For more on the translation to gauge theory see §5.2 here.) These definitions are not special to the square lattice and work for any cell complex, in any dimension.



[Fig by D.Ben-Zion, after Kitaev]

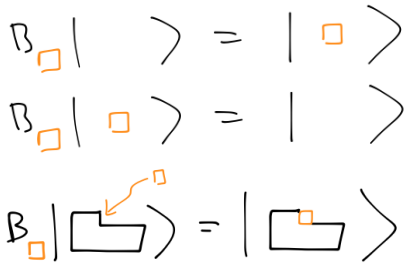
The hamiltonian is $\mathbf{H}_{\text{TC}} = -\Gamma_m \sum_p B_p - \Gamma_e \sum_v A_v$. These terms all commute with each other (since each vertex and plaquette share zero or two links), and they each square to one, so the Hamiltonian is easy to diagonalize. Let's find the groundstate(s).

Which states satisfy the 'gauss law condition' $A_v = 1$? In the X basis there is an extremely useful visualization: we say a link l of $\hat{\Gamma}$ is covered with a segment of string (an electric flux line) if $\mathbf{e}_l = 1$ (so $X_l = -1$) and is not covered if $\mathbf{e}_l = 0$ (so $X_l = +1$):  $\equiv X = -1$. In the figure at right, we enumerate the possibilities for a 4-valent vertex. $A_v = -1$ if a flux line ends at v.



So the subspace of $\mathcal{H}$ satisfying the gauss law condition is spanned by closed-string states (lines of electric flux which have no charge to end on), of the form $\sum_{\{C\}} \Psi(C) \, |C\rangle$.

Now we look at the action of B_p on this subspace of states:



$B_p = \prod_{\ell \in \partial p} Z_\ell$ creates and destroys strings around the boundary of the plaquette p:

$$B_p \, |C\rangle = |C + \partial p\rangle \ .$$

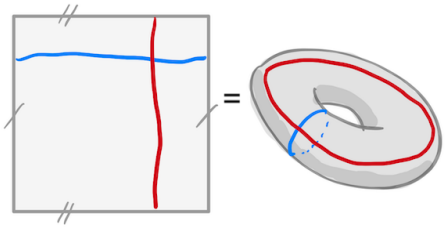
The argument of the ket is to be understood mod two. The condition that $B_p \, |\mathrm{gs}\rangle = |\mathrm{gs}\rangle$ is a homological equivalence. In words, the eigenvalue equation $\mathbf{B}_\square = 1$ says $\Psi(C) = \Psi(C')$ if C' and C can be continuously deformed into each other by attaching or removing plaquettes.

If the space is simply connected (like a sphere) – if all curves are the boundary of some region contained in the lattice – then this means the groundstate

$$|\mathrm{gs}\rangle = \sum_C |C\rangle \tag{1.2}$$

is a uniform superposition of all loops.

Topological order. If the space has non-contractible loops, however, then the eigenvalue equation does not determine the relative coefficients of loops of different topology! The two-dimensional torus obtained by considering periodic boundary conditions in x and y is an example of such a space:



(1.3)

On a space with $2g$ independent non-contractible loops, there are 2^{2g} independent groundstates. (In fact, the above is the very definition of the simplicial homology of the space, with $\mathbb{Z}_2$ coefficients; more generally the number of independent groundstates is 2^{b_1} where $b_1 \equiv \dim H^1(M, \mathbb{Z}_2)$. For more on the connection with homology and algebraic topology in general, see these notes.)

No local operator mixes these groundstates. This makes the topological degeneracy stable to local perturbations of the Hamiltonian. The degenerate groundstates are instead connected by the action of (Wegner-Wilson) loop operators:

$$W_C = \prod_{\ell \in C} Z_\ell \quad V_{\check{C}} = \prod_{\ell \perp \check{C}} X_\ell \ .$$

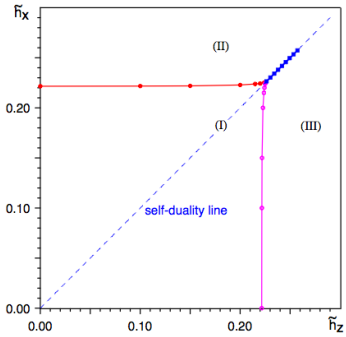
The second object is supported on a loop $\check{C}$ that lives in the *dual lattice*. In 2d, this is the lattice whose vertices are faces of the original lattice and vice versa. For the

square lattice this is another square lattice shifted by half a lattice spacing in each direction. (Notice that the loop operator for a single plaquette $W_{\partial\square} = B_p$ is the plaquette operator.) V, W commute with $\mathbf{H}_{\rm TC}$ and don't commute with each other (specifically W_C *anti*commutes with $V_{\check{C}}$ if C and $\check{C}$ intersect an odd number of times). This algebra must be represented on the groundstates, and it doesn't have any one-dimensional representations. In terms of our picture of strings, W_C *creates* a loop on C, and $V_{\check{C}}$ *detects* a loop intersecting $\check{C}$.

The deconfined phase. So far everything I've said works on any graph (actually: cell complex, since we need to know where the plaquettes are). And so far I've described the solvable limit, where $H = H_{\rm TC}$.

But the fact that the code distance goes like L (no local operator mixes the topological groundstates) is also the reason that the topological degeneracy is *robust*: adding local operators to the Hamiltonian will never split the degeneracy in perturbation theory. Therefore, this physics is characteristic of a phase of matter, and not just the special solvable Hamiltonian $H_{\rm TC}$. The toric code is a (special, RG fixed point, with zero correlation-length) representative of a phase of matter.

Perturbations such as $\Delta H = \sum_l (h_X X_l + h_Z Z_l)$ produce a nonzero correlation length. Let's focus on $D = 2+1$ for what follows. These couplings h_X and h_Z are respectively a string tension and a fugacity for the electric flux string endpoints: charges. Make these too big and the model is higgsed or confined, respectively. These fancy-sounding phenomena are actually adiabatically connected [Fradkin-Shenker]: Both are connected to the trivial state where e.g. $H = \sum_l X_l$ whose groundstate is a product $\otimes_l |{\rightarrow}_l\rangle$.



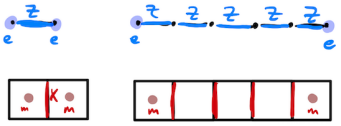
[from Tupitsyn-Kitaev-Prokof'ev-Stamp] The lower left corner is the toric code. The region on the lower left is the phase with topological order. Everything else is adiabatically connected to a product state.

Anyons. There are two kinds of elementary excited states of the toric code: violations of $A_s = 1$ and violations of $B_p = 1$. [13]

[13] Cultural note: The limit where the coefficient of the star term A_s goes to infinity is called 'pure $\mathbb{Z}_2$ gauge theory', where the condition $A_s = 1$, the Gauss' law constraint, is imposed exactly. The e particle defects cost infinite energy and hence are strictly forbidden in this theory.

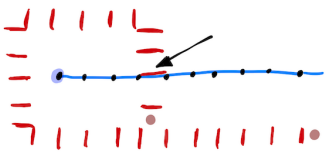
Here is how to make them. The defects are created by the endpoints of open Wilson lines. Again there are two kinds:

$$W(C) = \prod_{\ell \in C} Z_\ell, \quad V(\check{C}) = \prod_{\ell \perp \check{C}} X_\ell. \tag{1.4}$$



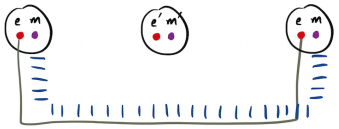
Here C is an open curve in the lattice, and $\check{C}$ is an open curve in the dual lattice. Endpoints of $W(C)$ violate A_s and endpoints of $V(\check{C})$ violate B_p.

These two kinds of particles have nontrivial mutual statistics, as you can see by moving one of them around the other and keep track of the strings trailing away from them. The process results in a net factor of (-1) on the state.

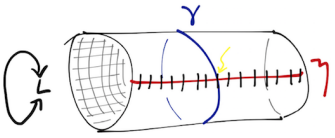


This has the further consequence that their bound state is a fermion, despite the fact that the model is entirely made from local, bosonic degrees of freedom. Hence, fractionalization.

To see this, observe that exchanging two particles can be accomplished by first rotating one around the other by a π rotation, and then translating both of them by their separation. As you can see in the figure, the first step requires the string creating the e particle to cross that creating the m particle on an odd number of links. (The second step is innocuous.)



Consider the cylinder. There is one nontrivial class of loops; call a representative γ. Let η be a line running along the cylinder. The two groundstates are generated by the action of the Wilson loop operator



$$V(\eta) \equiv \prod_{\ell \text{ crossed by } \eta} X_\ell$$

in the sense that

$$|\mathrm{gs}_2\rangle = V(\eta)\,|\mathrm{gs}_1\rangle \ .$$

This is also a groundstate (at $h_X, h_Z = 0$) since there is no plaquette with $\mathbf{B}_p = -1$ (more simply: $[\mathbf{H}_{h_X=h_Z=0}, V_x(\eta)] = 0$). They are *distinguished* by $W(\gamma) \equiv \prod_{l\in\gamma} X_l$ in the sense that the two groundstates are eigenstates of this operator with distinct eigenvalues:

$$W(\gamma)\,|\mathrm{gs}_\alpha\rangle = (-1)^\alpha\,|\mathrm{gs}_\alpha\rangle\,, \quad \alpha = 1, 2.$$

This follows since $W(\eta)V(\gamma) = -V(\gamma)W(\eta)$ – the two curves share a single link (the

one pointed to by the yellow arrow in the figure).

At finite h_X, h_Z (and in finite volume), there is tunneling between the topologically degenerate groundstates, since in that case

$$[\mathbf{H}, \prod_{l \in \gamma} X_l] \neq 0.$$

This means that for some n

$$\langle \mathrm{gs}_2 | \, \mathbf{H}^n \, | \mathrm{gs}_1 \rangle \neq 0.$$

The process that mixes the groundstates requires the creation of magnetic flux on some plaquette (*i.e.* a plaquette P with $B_P = -1$, which costs energy $2\Gamma_m$), which then must hop (using the h_X term in $\mathbf{H}$) all the way along the path η, of length L, to cancel the action of $V(\eta)$. The amplitude for this process goes like

$$\Gamma \sim \frac{\langle \mathrm{gs}_2 | \, (hX_1)\,(hX_2) \cdots (hX_L) \, | \mathrm{gs}_1 \rangle}{2\Gamma_m \cdot 2\Gamma_m \cdot \ldots 2\Gamma_m} \sim \left(\frac{h}{2\Gamma_m} \right)^L = e^{-L|\ln 2\Gamma_m/h|}$$

which is *extremely tiny* in the thermodynamic limit. The way to think about this is that the Hamiltonian is itself a local operator, and cannot distinguish the groundstates from each other. It takes a non-perturbative process, exponentially suppressed in system size, to create the splitting.

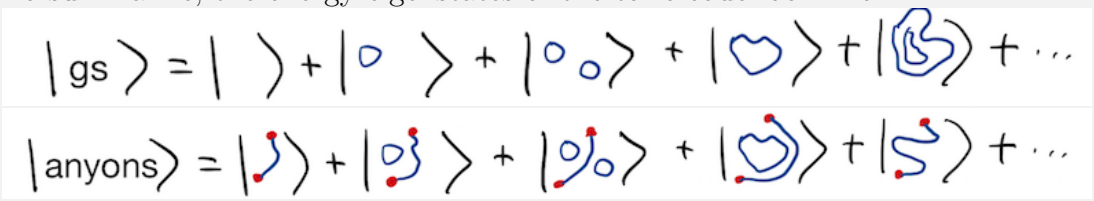
- I've focussed on the case of two spatial dimensions, but the toric code is well-defined on an arbitrary cell complex, in particular on a lattice in any number of dimensions. It has various generalizations:
- For example instead of the putting the dofs on the links, we can put them on the p-cells. Instead of using qubits, we can use $\mathbb{Z}_N$ clock and shift variables.
- The version on the p-cells with $\mathbb{Z}_N$ variables computes $H_p(C, \mathbb{Z}_N)$, the pth homology of the cell complex, as its groundstate subspace.
- With the dofs on the links, the model can be generalized to any finite group G (in fact this step was already taken in Kitaev's original paper). This is usually called the quantum double model.
- I haven't emphasized the connection to gauge theory above. The toric code is (a limit of) G lattice gauge theory with the gauss law condition imposed energetically, meaning that the low energy states satisfy the gauss law condition. At low energies it is governed by a TQFT called G gauge theory or sometimes it is called Dijkgraaf-Witten theory.

- By attaching various phases to the plaquette operators, we can make twisted gauge theory, as introduced by Dijkgraaf and Witten. A better framework for making explicit solvable lattice models for such states is the string net models developed by Levin and Wen here (see here for more).

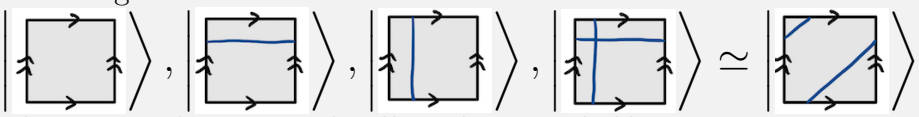
[End of Lecture 2]

The toric code.

To summarize, the energy eigenstates of the toric code look like:



The groundstates on the torus can be labelled as:



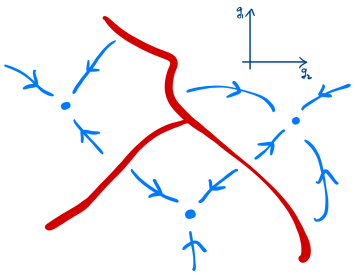
These groundstates are *locally indistinguishable*:

$$\left\langle \; \right| \mathcal{O}_x \left| \; \right\rangle = 0 \;\forall \text{ local operators } \mathcal{O}_x.$$

1.3 Back to the big overview

A word about the virtues of model Hamiltonians. In the previous subsection, we spent some time talking about a particular special Hamiltonian. It is pretty artificial-looking, in the sense that the terms in the Hamiltonian involve four or more spins at a time, and in that all the terms commute with each other, a very non-generic situation. Why do we think we can learn universal lessons from such a special system? The answer comes from the renormalization group.

Each gapped phase contains a single special point which is an attractive fixed point of the renormalization group, with zero correlation length. The fact that there can only be one such point in each gapped phase can be proved by contradiction: if we suppose there were more than one, we would not be able to draw the flow lines without discovering a repulsive fixed point separating them, but such a point must lie on the wall of gap-closing. So every gapped phase contains a special representative with zero correlation length. This special representative is easier to understand because all the irrelevant microscopic details have already been coarse-grained away. So we might as well use it to learn about the phase.



A third essential symptom of topological order is long-range entanglement in the groundstate wavefunction. I claim that the groundstate of a system with topological order cannot be made from a product state by a finite-depth local unitary circuit. I will explain this point in §1.4. Actually the converse of this is not true: there are some exceptional states with this property but which nevertheless do not have topological order, namely integer quantum Hall states and a few other examples we'll discuss below. A sharper diagnostic is the topological entanglement entropy, defined in terms of the entanglement entropy of a subsystem, which vanishes if and only if there are no anyons. This means that a state with topological order is (in a certain sense) far from a product state.

Let's enshrine these symptoms of topological order in a list:

1. Fractionalization of quantum numbers.

2. Groundstate degeneracy that depends on the topology of space.

3. Long-ranged entanglement.

I emphasize that the quantum numbers of the anyons (their statistics and (if there are global symmetries) charges) characterize the phase of matter. Especially in $D = 2 + 1$, the theory of anyons (their statistics and fusion rules) is a highly-developed mathematical edifice called topological field theory (TQFT), and more specifically in the case of 2+1 dimensions, unitary modular tensor category (UMTC) theory. Perhaps now is a good time to mention the most elementary distinction, between abelian and non-abelian topological order. By fusion of anyons, I mean the following. An anyon is a particle whose presence can be detected from a distance, by circling some other excitation around it and measuring the change of the resulting state. Given two anyon types a and b, I can consider circling other excitations around both of them. If I have

a complete basis of all the anyon types in the topological order under study, the result must look like one of them, but which one we get need not be uniquely determined:

$$a \times b = c_1 + c_2 + \cdots . \tag{1.5}$$

If the fusion rules look like the special case

$$a \times b = c$$

(a group law) for all the anyons, we say the topological order is abelian. Braiding such particles merely acts by a phase on the resulting unique state. In contrast, fusion rules like (1.5) require that the lowest-energy state in the presence of a and b is degenerate; in this case, braiding the two particles involves not only a phase, but a whole unitary matrix acting on this degenerate subspace.

If you have studied conformal field theory (CFT), you will notice a formal similarity between (1.5) and the operator product expansion. This is not a coincidence – a 2d CFT also defines a UMTC. In fact, the structure was defined first in that context, by Moore and Seiberg.

It was believed for a long time that all gapped phases are described at the lowest energies by a TQFT. However, this is not true. An interesting special case of topologically ordered states is *fracton phases*. A fracton phase has excitations (called fractons) that cannot be *moved* by any local operator (perhaps only in some directions of space). This is a strictly stronger condition than topological order, since an excitation can effectively be moved by annihilating it and creating it again elsewhere. This means that a fracton phase has a number of anyon types that grows with the system size – fractons in different places are really distinct anyons, since they are not related by any local operator. A consequence of this defining property is a groundstate degeneracy whose logarithm grows linearly with system size, and a subleading linear term in the scaling of the entanglement entropy of a region with the size of the region. Thus, the system knows about the geometry of space, and not just the topology, and hence cannot be described by an ordinary TQFT.

Gaplessness is something special that needs to be explained. An energy gap (and no topological order or special edge modes) should probably be the generic expectation for what happens if you pile together a bunch of degrees of freedom and couple them in some haphazard (even translation invariant) way. At the very least this follows on general grounds of pessimism: if you generically got something interesting by doing this, physics would be a lot easier (or more likely: we wouldn't find it interesting anymore). More seriously, gaplessness is an extreme case of a finite degeneracy: if allowed local operators could mix the the low-lying states, the levels would repel and create a gap under generic perturbations of the Hamiltonian.

Here is a list of some possible reasons for gaplessness (if you find another, you should tell me):

1. tuning to a critical point – notice that this requires some agent to do the tuning, and will only occur on some subspace of the space of couplings of nonzero codimension.

2. spontaneously-broken continuous symmetry (Goldstone bosons).

3. continuous unbroken gauge invariance (*e.g.* photons). Actually, this is a special case of item 2 for continuous one-form symmetries. If you are eager to learn more about this point of view, take a look here.

4. Fermi surface (basically only in this case do we get gapless degrees of freedom at some locus of dimension greater than one in momentum space)

5. edge of an invertible topological phase. Here the gaplessness or degeneracy is protected by an anomaly.

6. a symmetry that forbids mass terms in some other way. This is called 'technical naturalness'. An example is unbroken chiral symmetry, which forbids fermion masses. This is an explanation of the lightness of quarks and leptons in the Standard Model compared to the Planck scale. Supersymmetry (where chiral symmetry prevents fermion masses, and supersymmetry relates boson masses to fermion masses) goes in this category.

7. CFT with no relevant operators. I am not sure if there are examples of this which are not examples of item 3. Sometimes this is called self-organized criticality. See here for a recent search.

Each entry in this list is something to be understood[14]. If you encounter a gapless model and it does not fit into this list then I will bet you $5 that it is fine tuned, meaning that its creator simply didn't add enough terms to the Hamiltonian.

We can no longer define the boundary of a gapless phase by a wall of gap-closing. Instead, a useful definition that works for both gapped and gapless cases is that perturbation theory (in the difference of Hamiltonians) works within the phase. A phase

[14]Note that the masslessness of the graviton is a mystery not obviously solved by an element of this list.

is thus an equivalence class of hamiltonians such that within the phase, physics (thermodynamics, local operator expectations) varies smoothly[15].

Refinement by symmetry. Another important axis along which we may organize states of matter is by symmetry. Specifically, we can label states according to the symmetry group G that acts on their Hilbert space, and restrict our discussion to the space of Hamiltonians to those which are preserved by G. Here I am speaking about what are called *global symmetries*, that is, symmetries (not redundancies of our labelling, like gauge transformations).

Now, in the presence of a symmetry, for each H, we can ask whether the symmetry preserved by the groundstate? If not, this is called spontaneous symmetry breaking, and provides a label that distinguishes that state from the trivial state. This is the old story of classifying phases with symmetries.

But there is something else that can happen. There can be phase boundaries (walls of gaplessness) that cannot be circumvented within the space G-symmetric Hamiltonians, but which we can go around if we allow G-breaking terms in H. The example of SSB is already of this form: if we explicitly break the symmetry, there is no longer a sharp distinction between the paramagnetic phase and the broken phase, and we can go around the phase transition. (Consider the phase diagram of the Ising model including a longitudinal field.) If a G-symmetric state is connected to the trivial phase in the space of all H but not in the space of G-symmetric Hamiltonians, and does *not* spontaneously break G, it is called an SPT (symmetry-protected topological) state (protected by G). Such a state is very nearly trivial and we may wonder how it can be distinguished from the trivial state. The answer is that the edge theory realizes the symmetry in an interesting way, more precisely in an *anomalous* way. More on this later.

We can also have both topological order and anomalous edge modes. Various such G-symmetric states with topological order are called SET (symmetry-enriched topological) states.

Entanglement. An important perspective for organizing our understanding of quantum phases of matter – on which we focus for the rest of this chapter – is the amount and structure of entanglement in the groundstate.

A pure state is completely unentangled if it is a product state, $\otimes_x |s_x\rangle$. In a phase

[15] An annoying fact is that sometimes within a phase there are observables which vary non-analytically across a point where the thermodynamics and all local observables are perfectly smooth. A classic example is the roughening transition of Wilson loops in lattice gauge theory. So when I say 'physics varies smoothly' I really mean local observables. For non-local Hamiltonians, we do not even know how to define a notion of phase.

with a product-state representative, mean field theory applies. (And when it applies it is very useful: it predicts phase transitions and the associated critical theory of the order parameter, and even the excitation spectrum (*e.g.* spin waves).) Such states can be distinguished only by symmetries acting independently on each site. This problem is approximately solved (from the point of view of the experiment-free discussion of condensed matter physics we are having – of course the questions about energetic competition, which we are completely ignoring, can be very interesting from other points of view) by the representation theory of G. Elsewhere in such a phase, the effects of entanglement are a perturbative correction to non-universal quantities.

We've just defined what is 'unentangled'. We will want to be more quantitative about entanglement between A and its complement $\bar{A}$; here $\mathcal{H} = A \otimes \bar{A}$ is any bipartition of the Hilbert space, such as a region of space. Since we are talking about pure states (as opposed to *e.g.* thermal density matrices), we can do this using the entanglement entropy:

$$\boldsymbol{\rho}_A \equiv \mathrm{tr}_{\bar{A}} |\psi\rangle\langle\psi| \qquad S_A \equiv -\mathrm{tr} \boldsymbol{\rho}_A \log \boldsymbol{\rho}_A \quad .$$

This vanishes for product states. It equals $\log 2$ for $|\psi\rangle_{AB} = (|00\rangle + |11\rangle)/\sqrt{2}$, a single Bell pair. The dependence of S_A on the size and shape of A provides a lot of information about a state, some of which is universal, meaning a property of the phase, independent of the representative state.

So: highly-entangled and mean-field are antonyms. The description in terms of weakly-interacting waves above an ordered groundstate breaks down when the entanglement matters. The frontier of our understanding is states of matter where quantum mechanics is essential, not just a correction that can be included perturbatively. This is now a big industry (some interesting reviews are 1210.1281, 1302.0899) and I will try to give some flavor of it. The states of interest here are distinguished instead by their patterns of quantum entanglement. Furthermore, since such new states of matter are distinguished by different new kinds of orders, the phase transitions which separate them go beyond those described by fluctuations of local symmetry-breaking order parameters. This leads to new renormalization-group fixed points and new conformal field theories (CFTs).

1.4 Adiabatic continuation and local unitary circuits

[Zeng, Chen, Zhou, Wen, chapter 7] A useful alternative characterization of a gapped phase motivates this entanglement-based point of view.

First, a quantum phase is actually a property of the groundstate. This statement is a version of the Principle of Entanglement Bootstrap: all the universal data about

a phase of matter can be extracted from a single wavefunction, even just the density matrix of a ball. For example, the groundstate seems to know whether it is gapped or gapless, because in the former case the equal-time correlation functions will fall off exponentially. In the case of topological order, all of the data about the characteristic anyon excitations are encoded in the groundstate wavefunctions on a torus (see here) or indeed even a single wavefunction (see here or, most elegantly, here). We'll see some evidence below.

Claim: Two groundstates are representatives of the same phase[16] iff there is a quasi-local unitary circuit U of finite depth (the depth of a circuit is the (maximum) number of elementary gates acting on each site, 'finite' means independent of L, and I will explain quasi-local below) which maps one state to the other[17]. In symbols, $[\mathbf{H}_0] = [\mathbf{H}_1] \Leftrightarrow |\psi(1)\rangle = U\,|\psi(0)\rangle$.

[End of Lecture 3]

$\boxed{\Rightarrow}$ Suppose there is a path $\mathbf{H}(s)$ in the space of Hamiltonians starting at $\mathbf{H}_0$ (whose groundstate is $|\psi(0)\rangle$) and ending at H_1 (whose groundstate is $|\psi(1)\rangle$), with a gap for every s in between. In finite volume, the adiabatic theorem says we can construct a unitary which *probably* maps $|\psi(0)\rangle$ to $|\psi(1)\rangle$, namely slow-enough time-evolution along the path $\mathbf{H}(s)$,

$$\mathcal{T} e^{\mathbf{i}\int_0^1 dt \mathbf{H}(t)}\,|\psi(0)\rangle \propto |\psi(0)\rangle + \cdots\,. \tag{1.6}$$

Slow enough means compared to the timescale set by the gap, $1/\Delta$. Since the gap is independent of L, the required duration is too. The failure rate (the amplitude for the $\cdots$ in (1.6), however, is extensive. This problem can be fixed by a procedure called *quasi-adiabatic filtering* introduced by Hastings (a review is here) – one can construct a modified family of Hamiltonians $\tilde{\mathbf{H}}(s)$ which are almost as local[18] but precisely map groundstates to groundstates (the idea is to filter out the contributions from the excited states to which non-adiabatic transitions can happen)

$$|\psi(1)\rangle = \mathcal{T} e^{\mathbf{i}\int_0^1 dt \tilde{\mathbf{H}}(t)}\,|\psi(0)\rangle\,.$$

(So really the title of this section should have been 'quasi-adiabatic continuation...'.)

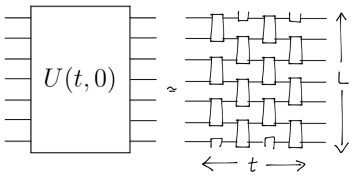
[16] In this discussion, we assume that the hamiltonians have a unique groundstate. So if we are talking about a phase with TO, we study it on a simply-connected space. The notion of phase is a local property.

[17] Actually, if you look back at our definition of gapped phase, we should also allow ourselves to tensor in ancillas in a product state before acting with the unitary.

[18] I've oversimplified the discussion here. Actually, there is a trade-off between locality of the filtered $\tilde{\mathbf{H}}$ and the precision with which the groundstates are mapped to each other. In fact, in order to precisely map the groundstates to each other, the operators $\tilde{\mathbf{H}}_x$ must have some tails, that is they have a profile which behaves like $e^{-r^{1-\delta}}$ where r is the distance from the point x – not quite exponential decay. This is the meaning of the modifier 'quasi-local'. Approximations to the exact map which are just as good for practical purposes can be made with $\tilde{H}_x$ which are strictly local.

Now this continuous time evolution can be *Trotterized*. That is, we can approximate it by a circuit, by breaking the time evolution into tiny steps. The crucial ingredient is the Baker-Campbell-Hausdorff formula, in the form

$$e^{\mathbf{i}\sum_x H_x \Delta t} = \prod_x e^{\mathbf{i} H_x \Delta t} + \mathcal{O}(\Delta t^2). \qquad (1.7)$$



The range of the terms in the Hamiltonian then determines the range of the individual unitary gates. The crucial point is that finite time evolution (independent of L) means a finite number of layers of elementary gates – this is a finite-depth circuit.

Thus we can regard circuits and unitaries from continuous-time evolution by local Hamiltonians as equivalent. A key point which is visible from the circuit picture is that there is a useful notion of lightcone, even in non-relativistic systems. Consider the domain of influence of a given input qubit in the circuit. It cannot effect output qubits that are arbitrarily far away, because the information about it only propagates by the local gates. The rigorous version of this statement is called the Lieb-Robinson bound.

Lieb-Robinson bound. Even non-relativistic theories have lightcones. Given a local Hamiltonian $\mathbf{H} = \sum_Z H_Z$ where the terms H_Z are supported on a subset Z and $\| H_Z \|$ shrinks rapidly with the diameter of Z (exponentially is good), then we can bound the correlations of local operators (A_X is supported on a set X and $A_X(t) = e^{-\mathbf{i}\mathbf{H}t} A_X e^{\mathbf{i}\mathbf{H}t}$ is its time evolution by $\mathbf{H}$):

$$\| [A_X(t), B_Y] \| \leq c_3 e^{-c_1 d_{XY}} \left(e^{2c_2 t} - 1 \right)$$

where $d_{XY} = \min_{i \in X, y \in Y} |i - j|$ is the distance between the sets X, Y and $c_3 = 2\| A_X \| \| B_Y \| |X|$, c_1, c_2 are constants. The quantity $2c_2/c_1$ is the *Lieb-Robinson velocity*.
You can find a relatively accessible proof (and many important applications) here.

$\boxed{\Leftarrow}$ Given a circuit $U = \mathcal{T} e^{\mathbf{i}\int_0^1 dt \tilde{\mathbf{H}}(t)}$ that accomplishes $|\psi(1)\rangle = U|\psi(0)\rangle$, we can define $U(s) \equiv \mathcal{T} e^{\mathbf{i}\int_0^s dt \tilde{\mathbf{H}}(t)}$ (just truncate the circuit at time s) and a family of states $|\psi(s)\rangle = U(s)|\psi(0)\rangle$. These states are the gapped groundstates of

$$\tilde{\mathbf{H}}(s) = \sum_x U(s) \tilde{\mathbf{H}}_x U(s)^\dagger$$

(gapped because the spectrum is independent of s) where $\mathbf{H}(0) = \sum_x \mathbf{H}_x$ is local, meaning that each $\mathbf{H}_x$ has finite range ξ (independent of L); the range $\tilde{\xi}$ of the terms in the filtered Hamiltonian $\tilde{\mathbf{H}} = \sum_x \tilde{\mathbf{H}}_x$ is still effectively finite (in the sense that there is still a Lieb-Robinson bound on its growth of correlations). But then the range of $U(s)\tilde{\mathbf{H}}_x U(s)^\dagger$ is bounded by $\tilde{\xi} + s v_{\max}$, where $v_{\max}$ is the maximum speed of propagation of correlations via $\tilde{\mathbf{H}}(t \leq s)$, which is again (according to the Lieb-Robinson bound) independent of L. ■

Notice that there is a lot of freedom in defining the unitary U that relates the two groundstates – we're actually only specifying its action on a single vector. What it does to the excited states (for example, the fact that it preserves the spectrum) is largely meaningless.

Here is an important consequence of this result. Recall that by a trivial phase we'll mean one with a representative groundstate which is a product state. This result implies that any groundstate in a nontrivial phase *cannot* be made from a product state by a finite-depth circuit. An example is a toric code ground state $\sum_{\text{loops},C} |C\rangle$.

1.5 Entanglement, short and long

Mean field theory is product states, which means there is no entanglement between regions of space at all. The next level of complication and interest to consider for possible groundstates of quantum many body systems is the case of states obtained by acting with a short-ranged quantum circuit of small depth on a product state. Let us consider such states, which are called short-range-entangled. What does their entanglement entropy of subregions look like and how do we distinguish which bits might be properties of a phase?

Let us focus on $d = 2$ space dimensions for definiteness. If the entanglement is short-ranged, we can construct a local 'entanglement entropy density' which is supported along the boundary of the region A [Grover-Turner-Vishwanath], and follow the logic of Landau theory to determine its form:

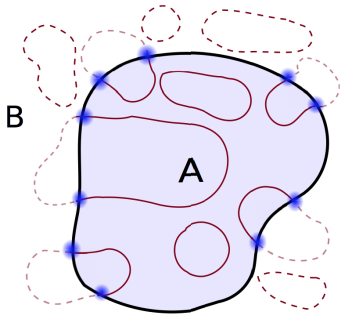
$$S_A = \oint_{\partial A} s d\ell = \oint \left(\Lambda + bK + cK^2 + ...\right) d\ell = \Lambda \ell(\partial A) + \tilde{b} + \frac{\tilde{c}}{\ell(\partial A)} + ...$$

In the first step, we use the fact that the entanglement is localized at the boundary between the region and its complement. In the second step we parametrize the local entropy density functional in a derivative expansion; K is the extrinsic curvature of the boundary. Since the total system is in a pure state, $S(A) = S(\bar{A})$, which implies $b = 0$: since interchanging A and $\bar{A}$ reverses the orientation of the boundary, the extrinsic curvature cannot contribute. This means that the subsystem-size-independent term

cannot come from terms local on the boundary; it is universal in the sense that it cannot be changed by changing the UV regulator (*e.g.* by rearranging lattice details). Where can such a term come from? For the example of the groundstate of $\mathbb{Z}_2$ gauge theory (the toric code), a closed string that enters the region A must leave again. This is one missing bit of freedom for the reduced density matrix of A, which means a contribution to the EE that is independent of the size of A:

$$S_A = |\partial A|\Lambda - \log 2 \equiv |\partial A|\Lambda - \gamma \tag{1.8}$$

where the area-law coefficient Λ is some short-distance-dependent junk and γ is a universal characterization of the nature of the topological order.



[fig: Tarun Grover]

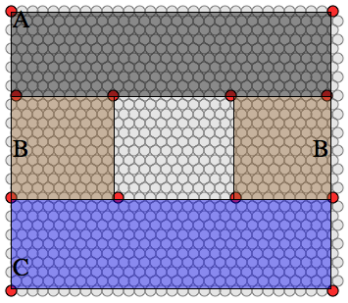
This is true for each component of the boundary of A individually, so the generalization of (1.8) to regions with $b_0(\partial A)$ boundary components is $S(A) = |\partial A|\Lambda - \gamma b_0(\partial A)$.

The universal constant term γ is called the topological entanglement entropy (TEE)[19]. For more general topological orders, γ can related to the spectrum of anyons; for Abelian states γ is $\frac{1}{2}\log$ (#torus groundstates). A beautiful argument for this is the Kitaev-Preskill wormhole construction (see their Fig. 2).

It is instructive to try to combine entropies of different regions to isolate the TEE from the area-law junk.

[19] It was introduced for $d = 2$ by Hamma-Ionicioiu-Zanardi, Kitaev-Preskill, Levin-Wen; the higher-dimensional generalizations are explained in the Grover et al paper linked above.

If the entanglement is indeed all short-ranged, then for collections of regions where the boundaries cancel out, $\partial(AB) + \partial(BC) = \partial(B) + \partial(ABC)$, (such as in the figure at right) nothing will be left. Let $S(x)$ be the EE of the subregion x in the state in question.



The regions should be large compared to the lattice spacing and the correlation length.

$$I(A:C|B) := S(AB) + S(BC) - S(B) - S(ABC)$$

is the conditional mutual information – correlations between variables A and C if we knew B. In general this combination of entropies satisfies a deep inequality called Strong Subadditivity (SSA), $I(A:C|B) \geq 0$. In general gapped phases in 2d, for the arrangement of regions at right, $I(A:C|B) = 2\gamma$, where γ is the subleading term to the area law defined in (1.8). The area-law contributions cancel out pairwise (notice that the corners cancel too).

When $\gamma = 0$, SSA is saturated. $I(A:C|B) = 0$ means $\boldsymbol{\rho}_{ABC}$ is a 'quantum Markov chain,' a state which can be reconstructed from its marginals $\boldsymbol{\rho}_A, \boldsymbol{\rho}_B, \boldsymbol{\rho}_C$ (by a formula due to Petz). So the quantity γ is an obstruction to this automatic reconstruction of the global state from local data.

The above argument shows that the TEE is not a short-distance artifact, but is it a property of a phase for any choice of A, B, C? And is it only nonzero for states with topological order? Almost. The papers linked above argue – assuming that the system is a liquid – that the TEE is independent of small changes in the regions (using $S_A = S_{\bar{A}}$ for pure states) and therefore insensitive to changes in the Hamiltonian that keep the correlation length short. There is, however, an important exception if the phase is not a liquid, whereby small changes of the regions lead the TEE to jump, and to give nonzero answers in states without TO. A consolation is that the correct value minimizes the answers you can get for γ.

In $d = 3$, ∂A is characterized by its number of components b_0 and its number of noncontractable loops b_1; these are related by $\chi = 2b_0 - b_1 = V - E + F = \frac{1}{2\pi}\int_{\partial A} R$ (the Gauss-Bonnet theorem) to the integral of a local density. The EE of A is linear in b_0 and b_1 (see Appendix E of the Grover-Turner-Vishwanath paper) but only one combination of them is a signature of long-range entanglement. Again this 3d TEE can be extracted by combining regions whose boundaries and corners cancel.

The TEE is only one number characterizing the nature of the topological order, and by no means uniquely characterizes it. For example, the double semion state is a distinct topological order from the toric code in $d = 2$, whose representative wavefunction is $\sum_{\text{closed loops},C}(-1)^{b_0(C)}|C\rangle$ (where $b_0(C)$ is the number of components

of the loops). As you can see from the form of the wavefunction it also has four groundstates on the torus and hence the same TEE. However, by now humans have learned to extract a great deal of the data specifying a given topological order from the entanglement properties of a single wavefunction, the most advanced incarnation of which is the *entanglement bootstrap*.

2 Some quantum Hall physics

By popular request, and because it still provides the best experimental examples of all the most interesting quantum topology phenomena, let's spend some time talking about effective descriptions of quantum Hall physics, both fractional and integer. I warn you that we will start from a distinctly macroscopic perspective, and will not say as much as others do about the important microscopic questions of energetic competition. For a perspective on those questions, I recommend chapters 12-14 of the textbook by Girvin and Yang (or these lectures by Leggett). I learned about this aspect of the subject from these notes by Girvin. For a great account of the whole subject I recommend David Tong's lectures.

2.1 Electromagnetic response of gapped states in $D = 2 + 1$

Let's think about a gapped state of matter made of some stuff in $D = 2 + 1$, out of which we can construct a conserved $\mathsf{U}(1)$ current j_μ (if you like, think of it as the current that keeps track of electron number). This means we can couple this current to an external, background, non-dynamical gauge field $\mathcal{A}_\mu$, by adding to the action functional like so:

$$S_{\text{microscopic}}[\text{the stuff}, \mathcal{A}] = S_{\text{microscopic}}[\text{the stuff}] + \int j^\mu \mathcal{A}_\mu + \cdots$$

where $\cdots$ is whatever other terms are needed to make this action fully gauge invariant. Here we'll treat $\mathcal{A}$ as a background field that we control[20]. The theory with this deformation is gauge invariant because the current is conserved $\partial_\mu j^\mu = 0$ (integrate by parts).

Integrate out the stuff to see the electromagnetic response:

$$e^{\mathbf{i}S_{\text{eff}}[\mathcal{A}]} \equiv \int [D\text{stuff}] e^{\mathbf{i}S[\text{stuff},\mathcal{A}]}.$$

[20] Notice that what we've done here is *not* gauging the $\mathsf{U}(1)$ symmetry. We are not changing the Hilbert space of the system. The background gauge field here just describes a particular collection of coupling constants.

The fact that the term linear in $\mathcal{A}$ gives the current density:

$$\langle j^\mu(x)\rangle = \frac{\delta}{\delta \mathcal{A}_\mu(x)} S_{\text{eff}} \tag{2.1}$$

where the RHS is evaluated on the configuration of background fields of interest, which could be $\mathcal{A} = 0$. Terms quadratic in $\mathcal{A}$ encode linear response:

$$\langle j^\mu(x) j^\nu(y)\rangle = \frac{\delta^2}{\delta \mathcal{A}_\mu(x)\delta \mathcal{A}_\nu(y)} S_{\text{eff}}.$$

Recall that $\langle jj\rangle$ is the main ingredient in Kubo's formula for the conductivity.

Because the stuff is gapped, S_{eff} is local. By the Landau-Ginzburg-Wilson logic, we can then determine $S_{\text{eff}}[\mathcal{A}]$ in a derivative expansion, as follows. To figure out the power counting, note that $\mathcal{A}$ is a gauge field, which is something that we can add to a derivative to make it a covariant derivative; therefore $\mathcal{A}$ has dimension 1, it counts the same as a derivative.

$$S_{\text{eff}}[\mathcal{A}] = \int \left(\underbrace{0 \cdot \mathcal{A}^2}_{\text{no symmetry breaking}} + \frac{\nu}{4\pi} \mathcal{A} \wedge \mathcal{F} + \frac{1}{g^2} \mathcal{F}_{\cdot\cdot} \mathcal{F}^{\cdot\cdot} \right) + \cdots . \tag{2.2}$$

($\mathcal{F} = \mathrm{d}\mathcal{A}$.) The $\mathcal{A}^2$ term is forbidden by gauge invariance[21]. With time-reversal symmetry (and only one gauge field), $\nu = 0$. If $\nu \neq 0$, Maxwell is less important than the term with ν, the Chern-Simons (CS) term. (Actually, without Lorentz invariance we can have non-vacuum dielectric constant and magnetic permittivity ϵ, μ, but this won't affect our story.) g is an energy scale that we can take to be large compared to our energies of interest. [End of Lecture 4]

The Kubo formula then says that the Hall conductivity is:

$$\sigma^{xy} = \lim_{\omega\to 0} \frac{1}{\mathbf{i}\omega} \underbrace{\langle j^x j^y\rangle}_{= \frac{\delta}{\delta \mathcal{A}_x(k)} \frac{\delta}{\delta \mathcal{A}_y(k)} S_{\text{eff}}[\mathcal{A}]} |_{k=0} = \frac{\nu}{2\pi} \overset{\text{in experimenter's units}}{=} \nu \frac{e^2}{h}.$$

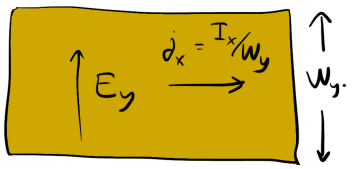
The analogous Kubo formula for the longitudinal conductivity σ^{xx} says that it is zero. This is one sense in which the system is an insulator. Note that one could argue with this characterization, since $\rho_{ij} = (\sigma^{-1})_{ij}$ has $\rho_{xx} = 0$, like a perfect conductor. While it's true that there is no dissipation (since the current is perpendicular to the voltage drop), no charge moves in the direction of the electric field, so I think it's safe to call it an insulator. Also, there's an energy gap (by assumption).

[21] To make it gauge invariant, we would have to add more (gapless) degrees of freedom, in particular the Goldstone mode ϕ for the broken $\mathsf{U}(1)$ symmetry, which would appear in the gauge-invariant combination $(\partial_\mu \phi + \mathcal{A}_\mu)^2$.

Next we'll show that, if there is no fractionalization, ν is quantized to be an integer. As a result, different values of ν label distinct phases of matter, since an integer can't change continuously. (Note that there could be further distinctions – states with the same ν could be distinct.)

Notice that 2d is special here because the conductivity is essentially dimensionless, and moreover the resistance is independent of the width of the sample:

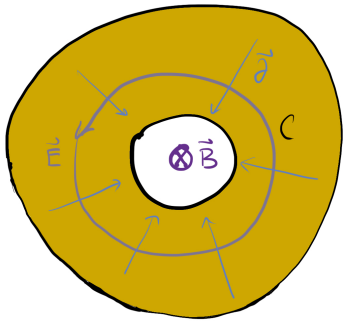
$$R_{xy} = V_y/I_x = E_y W_y/I_x = E_y/j_x = \frac{1}{\sigma_{xy}}$$



(Here W_y is the width of the sample in the y direction, perpendicular to the current, along the direction of voltage drop.) So it makes sense to say that σ^{xy} is quantized (in units of the quantum of conductivity $\frac{e^2}{h}$).

Flux-threading. I must emphasize that the following argument involves an important special case of studying the response of the system to background fields, called *flux-threading*. Rather than thinking about static background EM fields, we consider (still externally-fixed) EM fields that slowly vary in time – in a loop.

Consider the system on an annulus (sometimes called, in this context, the 'Corbino geometry'). Adiabatically thread 2π worth of magnetic flux through (a solenoid in) the hole in the annulus. This means we slowly vary the magnetic field in the hole, so that the change in flux is the flux quantum $\Phi_0 \equiv \frac{hc}{e} = 2\pi$.



In the following equation only I restore un-natural units:

$$\Phi_0 = \Delta\Phi = \int \mathrm{d}t \partial_t \left(\int_{\text{hole}} \mathrm{d}\vec{a} \cdot \vec{B} \right) \overset{\text{Faraday}}{=} -c \int \mathrm{d}t \oint_C \vec{E} \cdot \mathrm{d}\vec{\ell} \overset{j_r = \sigma_{xy} E_\varphi}{=} -\frac{c}{\sigma_{xy}} \underbrace{\int \mathrm{d}t I_r}_{=\Delta Q}$$

where C is a curve going around the hole, and $I_r = \oint_C j_r$ is the total radial current passing through the curve C. Note that Faraday's law is the spatial components of the equation $dF = 0$ stating the absence of magnetic monopoles, so does not depend on the form of the effective action for the EM field.

We conclude that an amount of charge

$$\Delta Q = \frac{\Phi_0}{c} \sigma_{xy} = \nu e \tag{2.3}$$

(e is the charge of the electron) is transferred from one edge of the cylinder to the other. Because of the energy gap, we can do the flux-threading adiabatically. Moreover, the initial and final Hamiltonians are related by a gauge gransformation:

$$H(\Phi = 0) \cong H(\Phi = 2\pi).$$

(We detect a magnetic field by moving a charged particle around a loop and acquiring a phase $e^{\mathbf{i}q \oint_C A}$; since the charge is quantized to be an integer, 2π flux is the same as no flux.) They have the same spectrum. Moreover, the work done on the system is $\int I d\Phi \propto \int dt \left(\frac{d\Phi}{dt}\right)^2$ which goes to zero in the thermodynamic limit, if our process is adiabatic. Therefore the initial and final states must be degenerate in the thermodynamic limit. But the adiabatic deformation can take one state to another. The states differ in that an amount of charge ν has been moved from one boundary to the other. Since, in the absence of fractionalization, charge is carried only by electrons, localized objects with integer charge, we conclude $\mathbb{Z} \ni \nu = \Delta Q/e = \sigma_{xy}$.

If we make the further assumption that the states can be labelled with the same labels as free-electron states, *i.e.* electron occupation numbers, we can say more. We've identified two different states related by the flux threading. The single-particle states whose occupation numbers have changed must lie near the Fermi level. Since we've assumed the bulk is gapped, we conclude that there must be gapless edge states.

The following argument implies further that a gapped system with $\sigma^{xy} = \nu\frac{e^2}{h}$ hosts a particle excitation with charge νe and exchange statistics $\pi\nu$.

Now consider the system on the plane. Adiabatically thread 2π worth of localized magnetic flux through some localized region R of the sample (as in the $\otimes$ at right). To do this, we have to stick a really thin solenoid through the 2d surface on which the system lives. This means as above that (I now return to units with $\hbar = c = e = 1$)

$$2\pi = \Delta\Phi = \int \mathrm{d}t \partial_t \left(\int_{R|\partial R = C} \mathrm{d}\vec{a} \cdot \vec{B} \right) \overset{\text{Faraday}}{=} - \int \mathrm{d}t \oint_C \vec{E} \cdot \mathrm{d}\vec{\ell} \quad \overset{j_r = \sigma_{xy} E_\varphi}{=} \quad -\frac{1}{\sigma_{xy}} \underbrace{\int \mathrm{d}t j_r}_{=\Delta Q}.$$

We conclude that the inserted flux sucks in an amount of charge

$$\Delta Q = \nu e.$$

Because of the energy gap, we can do this adiabatically. And because the flux is a multiple of 2π we end up with another state of the same system – the inserted flux is not an extrinsic defect[22]. This object is a localized excitation of the system – it can

[22]This is slightly less obvious than in the case where the flux was in a hole in the system. Put the

move around, it's a particle[23]. And if $\nu \notin \mathbb{Z}$, it has fractional charge νe.

But now consider two of them. From the Bohm-Aharonov effect, each has statistics angle $\pi\sigma_{xy}$ (for $\sigma_{xy} = \frac{e^2}{h}$, this is a fermion). Therefore, no topological order (hence no fractional statistics) implies (again) that $\sigma_{xy} \in \mathbb{Z}\frac{e^2}{h}$.

This argument has a stronger consequence for a system made only of bosons: In a gapped state with no fractionalization, all particles, including this one, must have the same statistics as the microscopic constituents. For a non-fractionalized state made from fermions, we can conclude that $\nu \in \mathbb{Z}$, since odd ν will produce a fermionic particle. But for a system made only of bosons, without topological order, σ_{xy} must be an *even* multiple of $\frac{e^2}{h}$. For a careful recent discussion of these arguments, see here.

Roles of topology. Quantum Hall insulators provide examples that are topological in two distinct ways. The Hall conductivity (apply small electric field in x direction, measure current in y direction, take ratio)

$$\sigma^{xy} = \frac{p}{q}\frac{e^2}{h} \tag{2.4}$$

is a rational number – $p, q \in \mathbb{Z}$ – despite (in fact, with the help of)[24] disorder.

1. The integer quantum Hall effect (IQHE) is governed by (2.4) with $q = 1$. We know such a phase with $p \neq 0$ is not adiabatically connected to the trivial phase

solenoid at the origin and choose the gauge $A = \Phi_0 \frac{d\varphi}{2\pi}$, where φ is the azimuthal coordinate in the plane (this is the field involved with the flux-threading, in addition to any magnetic field responsible for supporting the quantum Hall state). The gauge transformation that removes A is $g = e^{\mathbf{i}\varphi}$, which is singular at the origin of polar coordinates. This singularity has no effect, since it just changes the overall phase of the wavefunction, $\Psi \to \Psi \prod_j e^{\mathbf{i}\varphi_j}$.

[23] Here's something I'm confused about at the moment. Unlike in the case of the 'Corbino geometry' above, where the flux was inserted in a hole in the sample, in this case we do expect that the final state has a different energy than the initial state – the quasiparticle we've created has some rest mass larger than zero, and (by charge conservation) the gap should be of order twice this value. Where does the argument about the work done break down?

[24] If the system is translation invariant, one can show that the Hall conductivity must be strictly linear in the continuously-variable filling fraction $\nu = \rho\Phi_0/B$ (ρ is the electron density, $\Phi_0 = hc/e$ is the flux quantum), so there can be no quantized plateaux. One possibility for breaking the symmetry leading to this conclusion is disorder; this is the sense in which it helps. (It is then less obvious that it doesn't just make the Hall conductivity zero by Anderson localizing all the states.) Another possibility is that the symmetry could be broken spontaneously, as discussed here. Another possibility is the presence of a lattice – the effective description of QHE in terms of Chern-Simons gauge theory works just as well for Chern insulators, *i.e.* a tight-binding model of fermions hopping a lattice where the filled bands have a nonzero Chern number. We will talk about Chern insulators more later on. It seems to me that another possibility could be boundaries of the sample.

because this integer cannot vary continuously – it's definitely a distinct phase of matter. This happens for free electrons filling Landau levels or Chern bands. The quantization $p \in \mathbb{Z}$ arises because of topology of single-particle orbits; p is the Chern number. This does *not* exhibit topological order. This is an example of a 'topological insulator'. It is a band insulator – the electrons completely fill some bands, and therefore there is an energy gap, measured by the energy difference to the next band (or to the next Landau level, *i.e.* the cyclotron frequency eB/m).

What's special about such a phase? We'll see below that it has gapless edge modes.

Chern bands. To understand the possibility of IQHE without a magnetic field, first consider the possibility that our electrons in a 2d electron gas (2DEG) in a big magnetic field may be constrained to live at the sites of a lattice. (In fact this could be true of all the degrees of freedom of the Standard Model for all we know, if the lattice spacing is small enough.) That is, it could be described by a tight-binding model, like

$$H_{B=0} = -\sum_{ij} t c_i^\dagger c_j + h.c. \tag{2.5}$$

where i, j label sites of a lattice (say the square lattice). This hamiltonian is solved by going to Fourier space

$$c_k \propto \sum_i e^{\mathbf{i}\vec{k}\cdot\vec{r}_i} c_i \tag{2.6}$$

so that $H = \int_{\mathrm{BZ}} c_k^\dagger c_k \epsilon(k)$. If we expand $\epsilon(k)$ about the minimum, it will generically be quadratic and we get back a continuum description of non-relativistic electrons.

How do we include the magnetic field in such a description? The answer is that we replace the hopping parameter $t \to t e^{\mathbf{i}a_{ij}}$ so that

$$H_B = -\sum_{ij} t e^{\mathbf{i}a_{ij}} c_i^\dagger c_j + h.c. \tag{2.7}$$

with

$$a_{ij} = \frac{2\pi}{\Phi_0} \int_{\vec{r}_i}^{\vec{r}_j} \vec{A}(r) \cdot d\vec{r}. \tag{2.8}$$

(This is called Peierls' substitution.)

Consider the case with $\vec{B} = \vec{\nabla} \times \vec{A} = B\hat{z}$ uniform. The description depends a bit on whether $\Phi_\square \equiv \oint_{\partial\square} \vec{A} \cdot d\ell$, the flux through a single plaquette, is a rational

multiple of Φ_0. For simplicity assume that it is, say $\Phi_\square = \Phi_0/k, k \in \mathbb{Z}$. Then we can choose the phases so that there is a finite unit cell (with k sites) as follows: We'll choose an analog of $\vec{A} = -Bx\hat{y}$ gauge. This is accomplished by choosing

$$t_{(x,y),(x,y+1)} = \omega^x, \quad \omega \equiv e^{\frac{2\pi \mathbf{i}}{k}}. \tag{2.9}$$

With this choice of phases, the holonomy around each square $\prod_{\ell \in \partial \square} t_\ell$ is equal to ω. (Note that $t_{ij} = t^\star_{ji}$.)
Again we can diagonalize the resulting Hamiltonian in momentum space, but now the unit cell has k sites, so there will be k bands:

$$c_{k\alpha} \propto \sum_i e^{\mathbf{i}\vec{k}\cdot\vec{r}_i} c_{i+\alpha\hat{x}} \tag{2.10}$$

so that $H = \int_{\mathrm{BZ}} c^\dagger_{k\alpha} c_{k\alpha} \epsilon_\alpha(k)$. As k grows, these bands get flatter and flatter and it becomes a better approximation to the continuum Landau levels.
But what's special about these bands that if I fill some of them I get a nonzero quantized Hall response? The answer is given by the following formula, called the TKNN formula. For any free fermion system, with some number of fully-filled bands, the Hall conductivity is

$$\sigma_{xy} = -\frac{e^2}{h} \sum_{\text{occupied bands},\alpha} C_\alpha \tag{2.11}$$

where C_α is the *Chern number* of band α, defined as follows:

$$C_\alpha \equiv \frac{1}{2\pi} \int_{\mathrm{BZ}} \mathcal{F}_\alpha, \quad \mathcal{F}_\alpha \equiv \left(\vec{\nabla}_k \times \mathcal{A}_\alpha\right)_z, \tag{2.12}$$

where $\mathcal{F}$ is the Berry curvature, and

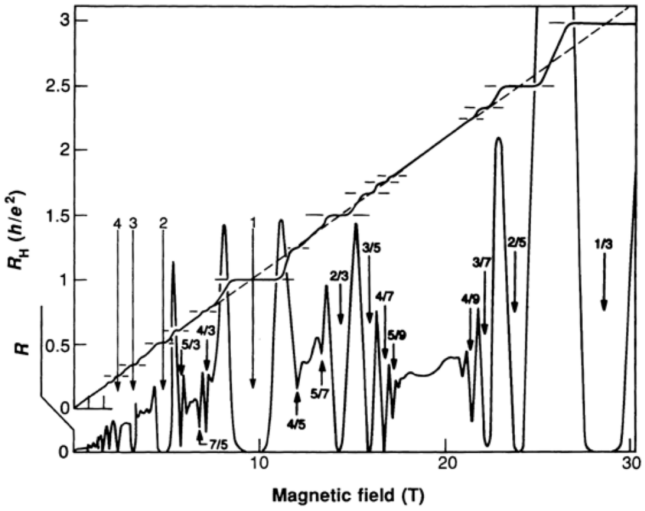
$$\mathcal{A}^\alpha_\mu \equiv \mathbf{i} \left\langle u^\alpha(k) \right| \partial_{k^\mu} \left| u^\alpha(k) \right\rangle \tag{2.13}$$

is the Berry connection. Here $|u^\alpha(k)\rangle$ is the wavefunction at momentum k of the αth band. The numbers C_α are always integers if the bands don't touch each other. We'll talk later about why this formula is true, and how to think about it so that it is robust to including interactions.

2. The fractional quantum Hall effect (FQHE) is described by $q \geq 2$, and requires interactions. There is necessarily topological order. $q \in \mathbb{Z}$ because of the topology of *many-body* wave function. The electron *fractionalizes*: as we'll see excitations

have charge $1/q$, fractional statistics.

Perhaps I should pause to emphasize that so far we have *assumed* that a bunch of stuff in $D = 2 + 1$ with a conserved charge but without time-reversal symmetry (such as a 2d electron gas in a magnetic field) can form a state with an energy gap. What we've shown, using the Landau-Ginzburg-Wilson logic, is that if this does happen, the system exhibits a quantized Hall conductivity. It is a remarkable fact, not at all obvious from anything we've said here, that this actually happens. As evidence, I include the classic plot:



In this plot, the electron density is fixed, and the horizontal axis varies B. The key point is that around the value of B where the actual filling fraction (number of electrons per flux quantum) hits certain microscopically-preferred rational values, the measured σ_{xy} shows a plateau (along which $\sigma_{xx} = 0$), in striking contrast with the prediction from translation symmetry[25]. When deforming away from the middle of the plateau,

[25] Theorem: translation symmetry implies $\sigma^{xy} = \frac{\rho ec}{B}$, where ρ is the electron density. Consider the Hamiltonian H for a 2d collection of electrons in a uniform magnetic field $\vec{B} = B\hat{z}$, with $\vec{E} = E\hat{y}$ uniform as well. Choose the gauge $\vec{A}(r,t) = (-cEt + Bx)\hat{y}$. Gauge invariance means that the momentum of the ith electron appears in H only in the combination $\vec{\Pi} = \vec{p}_i + \frac{e}{c}\vec{A}(r_i, t)$. Remove the Et term by changing variables to $\vec{r}'_i = \vec{r}_i - c\frac{E}{B}t\hat{x}, \vec{p}'_i = p_i, t' = t$. The current in the new frame is $j'_x = -\rho e \left\langle \dot{X}' \right\rangle$ where X' is the center-of-mass position in the $\hat{x}$ direction. This is time-independent. The center-of-mass momentum in the y direction $P'_y = P_y \equiv \sum_j (p_j)_y$ is conserved. Therefore if $\left\langle \dot{X}' \right\rangle \neq 0$, $\Pi'_y = \sum_i ((p'_i)_y + \frac{e}{c}BX'$ would blow up, and therefore so would the kinetic energy. We conclude that $0 = \left\langle \dot{X}' \right\rangle \propto j'_x = 0$. But $\vec{j}'(r') = \vec{j}(r) + ec\frac{E}{B}\rho\hat{x}$ and therefore

$$j_x = -\rho ec\frac{E}{B} = \sigma^{xy}E^y$$

from which we conclude $\sigma^{xy} = -\frac{\rho ec}{B}$ without any further assumptions. A closely-related argument

the extra electrons must somehow be prevented from participating in the transport – they are somehow localized, either by disorder or by forming a (Wigner) crystal.

2.2 Abelian Chern-Simons theory

[Wen's book or this review; Zee, Quantum Hall Fluids; Zee's QFT book §VI.2] I want to explain an example of how properties 1 and 2 can be realized in a simple physical system, using the EFT (effective field theory) that describes the canonical examples of topologically-ordered states: (abelian) *fractional* quantum Hall states in $D = 2 + 1$.

The low-energy effective field theory is Chern-Simons-Witten gauge theory, whose basic action is:

$$S_0[a_I] = \sum_{IJ}^{n} \frac{K_{IJ}}{4\pi} \int a_I \wedge \mathrm{d}a_J \tag{2.14}$$

a^I are a collection of abelian gauge fields.

Where did these gauge fields come from? We'll discuss some perhaps-more-informative possibilities below, but one very simple way to motivate their introduction is as follows. By assumption, our system has a conserved $\mathsf{U}(1)$ current, J^μ, satisfying $\partial_\mu J^\mu = 0$. In $D = 2 + 1$, we can *solve* this equation by introducing a field a and writing

$$J^\mu \propto \epsilon^{\mu\nu\rho}\partial_\nu a_\rho. \tag{2.15}$$

The continuity equation is automatic if J can be written this way (for nonsingular a) by symmetry of the mixed partials. (The equation could also be solved by a sum of such terms, as we write below. This ambiguity reflects some of the enormous multiplicity of different quantum Hall states.) Then we must guess what dynamics should govern a. Here we just add all terms allowed by the symmetries, as usual. When it's not forbidden by time-reversal symmetry or parity, the Chern-Simons term is the most important term at low energies.

Notice that we wrote this action in a coordinate-invariant way without needing to mention a metric. This is a topological field theory. In the absence of charges, the equations of motion say simply that $0 = \frac{\delta S_0}{\delta a} \propto f = da$. Unlike Maxwell theory, there are no local, gauge invariant degrees of freedom. And, by Legendre transformation, the Hamiltonian is just zero. It is a theory of groundstates.

[End of Lecture 5]

Consider the simplest case of (2.14) with a single such field a, $S_0[a] = \int \frac{k}{4\pi} a \wedge da$. As we'll see, this describes *e.g.* the Laughlin state of electrons at $\nu = 1/k$ for k an

assuming boost invariance appears in many places, in particular in the Girvin lectures.

odd integer. (More general K describe the so-called hierarchy states, and give some understanding of the pattern of plateaux that appear.)

When I say there are no local dofs, I am thinking of the limit where we totally ignore the Maxwell term. The Maxwell term is irrelevant: its effects go away at low energies. Let's add it back in and look at the spectrum of fluctuations with the action:

$$L = \frac{k}{4\pi}\epsilon^{\mu\nu\rho}a_\mu\partial_\nu a_\rho + \frac{1}{4M}f_{\mu\nu}f^{\mu\nu}$$

where M is some microscopic energy scale above which the Maxwell term matters. The equation of motion is

$$0 = \frac{\delta S}{\delta a_\lambda} = \frac{k}{2\pi}\epsilon^{\lambda\rho\nu}f_{\rho\nu} + \frac{\partial_\mu f^{\mu\lambda}}{M}. \tag{2.16}$$

In terms of $f^\lambda \equiv \epsilon^{\lambda\rho\sigma}f_{\rho\sigma}$ this is

$$\epsilon^{\mu\nu\rho}\partial_\nu f_\rho + \frac{Mk}{2\pi}f^\mu = 0. \tag{2.17}$$

Taking curl of the BHS ($\epsilon_{\mu\alpha\beta}\partial^\alpha$(BHS)) gives

$$\left(\partial_\mu\partial^\mu - \left(M\frac{k}{2\pi}\right)^2\right)f_\rho = 0. \tag{2.18}$$

This is the dispersion relation for an excitation of mass $\frac{Mk}{2\pi}$. As $M \to \infty$, the excitation goes off to infinite energy.

If we demand that (2.14) is invariant (or rather $e^{\mathbf{i}S_0}$ is invariant) under $\mathsf{U}(1)^n$ gauge transformations, including large gauge transformations, then k must be an integer[26].

[26]Note that when k is odd, our flux-threading argument shows the existence of gauge-invariant fermionic excitations, and so can only arise from a theory with microscopic fermions.

To see the mysterious factor of two, here is my advice:

1. Recall that $\int a_1 \wedge da_2 = \int a_2 \wedge da_1$ by IBP. This is why the K-matrix is symmetric.
2. Consider the system on a spacetime of the form $S^1 \times S_g$, where S_g is a compact Riemann surface. An arbitrary connection on such a spacetime is of the form $a = a_1 + a_2$ where $a_{1,2}$ are polarized along the first and second factor of the spacetime respectively. Then

$$\int a \wedge da = \int (a_1 \wedge da_2 + a_2 \wedge da_1)\,. \tag{2.19}$$

   Note that the second term need not be zero because a_1 can still depend on the coordinates of S_g. Using the first item, this gives

$$\frac{1}{4\pi}\int a \wedge da = \frac{1}{2\pi}\int a_1 \wedge da_2. \tag{2.20}$$

3. Now consider the variation $a \to a + \mathbf{i}g^{-1}dg, g = e^{\mathbf{i}\varphi}$, where φ is the coordinate along the circle.

Thanks to Aidan Sheckler for help with this.

From the point of view of (2.15), the demand that the gauge group is really $\mathsf{U}(1)$, and the concomitant quantization of flux of da, comes from demanding that the charge of the current J_μ is quantized (in units of the charge of the electron). It's pretty interesting that this seemingly-metaphysical microscopic information that all charges come in integer multiples of the electron charge has such strong consequences for the low-energy description of macroscopic quantum phases.

Flux quantization. A crucial ingredient in the above argument for the quantization of the level is the fact of *flux quantization*. This is a consequence of the fact that the gauge gauge group is compact, in this case $\mathsf{U}(1)$ (as opposed to $\mathbb{R}$). Flux quantization is also the reason that the TKNN invariant (aka Chern number of a band) (2.12) is an integer, so I really should have explained it earlier. Let me stop to say a few words about it. The basic statement is that the integral of the flux F over a closed submanifold is a topological property.

I'm just going to explain the basic example of the phenomenon. To begin, let's think about a configuration of the electromagnetic field in 3-space with a magnetic monopole at the origin, *i.e.* satisfying $\vec{\nabla} \cdot B = 4\pi g \delta^3(x)$. The LHS of this equation is the time component of the 1-form $\star dF = \star d^2 A$, so A must not be globally well-defined or else this would vanish. A way out is to cover space with patches. Actually, all the action happens on the unit sphere surrounding the monopole (that is, $\mathbb{R}^3 \setminus \{0\} \simeq S^2$), so let's just think about that. We cover this two-sphere with two patches U_N and U_S consisting of everything but the south and north poles respectively. The overlap deformation retracts to the equator. On U_N and U_S respectively we take the gauge potential to be[a]

$$A^N = g(1-\cos\theta)d\varphi, \quad A^S = g(-1-\cos\theta)d\varphi = A^N - 2gd\varphi = A^N + \mathbf{i}g_{NS}^{-1}dg_{NS}$$

where

$$g_{NS}(\theta,\varphi) \equiv e^{\mathbf{i}2g\varphi}$$

is a function on $U_N \cap U_S$. Notice that $F = dA^N = dA^S = g\sin\theta d\theta \wedge d\varphi$ are both proportional to the volume form on the 2-sphere, consistent with the demand that magnetic flux is coming out from the origin in a spherically-symmetric way. g_{NS} has two names here: Mathematically, it is the transition function for a complex vector bundle of rank one between our two patches on the 2-sphere. Physically it is a function parameterizing a gauge transformation

$$A \to A^g = g^{-1}\left(A - \mathbf{i}d\right)g. \tag{2.21}$$

between two choices of gauge for the vector potential.

What is the structure group G of the vector bundle in question? g_{NS} is single valued under $\varphi \cong \varphi + 2\pi$ and therefore $G = \mathsf{U}(1)$ iff

$$4\pi g \in 2\pi\mathbb{Z}. \tag{2.22}$$

This is the Dirac quantization condition. Notice that we didn't actually say anything about quantum mechanics. We just demanded that there was some $\mathsf{U}(1)$-valued transition function connecting the gauge potentials on the two patches. Quantum mechanics comes in for example if we put a charged particle in this background field, because the phase of the wavefunction of a charged particle transforms under a gauge transformation by multiplication by the transition function; if this isn't single-valued, the wavefunction is not well-defined. In that context, this is the physical reason we want the structure group to be $\mathsf{U}(1)$ and not $\mathbb{R}$.

This conclusion leads to *flux quantization*: A vector bundle with a compact structure group has quantized fluxes, $\oint_S \frac{F}{2\pi} \in \mathbb{Z}$, where S is any compact 2d submanifold of B. In this example, we have

$$\begin{aligned} \oint_{S^2} \frac{F}{2\pi} &= \frac{1}{2\pi}\left(\int_{H_N} dA^N + \int_{H_S} dA^S\right) \overset{\text{Stokes}}{=} \frac{1}{2\pi}\oint_{\text{equator}} \left(A^N - A^S\right) \\ &= -\frac{1}{2\pi}\oint_{\text{equator}} \mathbf{i} g_{NS}^{-1} dg_{NS} = \frac{1}{2\pi}\int_0^{2\pi} 2g\, d\varphi = 2g \overset{(2.22)}{\in} \mathbb{Z} \end{aligned} \tag{2.23}$$

where H_N and H_S are the north and south hemispheres, which lie respectively in U_N and U_S and which have $\partial H_{N/S} = \pm$equator.

Although I explained it in the context of a $\mathsf{U}(1)$ bundle on S^2 this phenomenon is general for compact gauge groups. In the case of the TKNN invariant, the Bloch wavefunctions define a $\mathsf{U}(1)$ bundle on the Brillouin zone, which is a 2-torus. In this case the structure group is $\mathsf{U}(1)$ because the gauge transformations are associated with rephasing the (normalized) wavefunction $\psi(k) \to g(k)\psi(k)$, so $g(k) = e^{\mathbf{i}\phi(k)}$ is a phase.

Let me be more explicit about this bundle over the BZ defined by some Bloch bands in 2d. The eigenfunctions are defined by $h(k)u_\alpha(k) = \epsilon_\alpha u_\alpha(k)$. Because space is a lattice, k takes values in a 2-torus. Let's think about the special case when there are two bands, $\alpha = \pm 1$, so the hamiltonian takes the form

$$h(k) = a_0(k)\mathbb{1} + \sum_{i=x,y,z} \sigma^i a^i(k). \tag{2.24}$$

The $a_0(k)$ bit doesn't affect the eigenfunctions so forget it. Assuming the gap is nonzero for all k requires $X \equiv \sum_i (a^i(k))^2 > 0$. Without changing the topology

we can divide h by X so that (the new) $a^i(k)$ defines a unit vector, *i.e.* a point on S^2. So $a^i(k)$ is a map from T^2 to S^2. The a_i determine the wavefunctions by

$$u_-(k) = \begin{pmatrix} \cos\frac{\theta(k)}{2} \\ \sin\frac{\theta(k)}{2} e^{\mathbf{i}\varphi(k)} \end{pmatrix} \tag{2.25}$$

where $\theta(k), \varphi(k)$ are the polar coordinates on S^2 labelling $a^i(k)$. From this information you can check that the Chern numbers of the two bands are just plus or minus the number of times the T^2 winds around the S^2.

[a] I'm using polar coordinates on the unit sphere where $x = \sin\theta\cos\varphi, y = \sin\theta\sin\varphi, z = \cos\theta$.

More generally, K must be a symmetric matrix (don't forget the sign from integration by parts) of integers.

Two more ingredients are required for this abelian CS theory to describe the low-energy EFT of a quantum Hall state:

(1) We must say how the stuff is coupled to the EM field. Notice that these gauge fields imply conserved currents $j^I_\mu = \frac{1}{2\pi}\epsilon_{\mu\nu\rho}\partial_\nu a^I_\rho$. This is automatically conserved by antisymmetry of $\epsilon_{\mu\nu\rho}$, as long as a is single-valued. In its realization as the EFT for a quantum Hall state, a linear combination of these currents is coupled to the external EM field $\mathcal{A}_\mu$:

$$S_{EM}[a_I, \mathcal{A}] = \int \mathcal{A}^\mu t_I j^I_\mu \ ,$$

i.e. the actual EM current is $J_\mu = \sum_I t_I j^I_\mu$. The normalization is determined so that flux quantization implies quantization of charge.

(2) Finally, we must include information about the (gapped) quasiparticle excitations of the system. Creating a quasiparticle excitation costs some energy of order the energy gap, and their dynamics is not included in this ultra-low-energy description. As I described above, however, the quantum numbers of these excitations is a crucial part of the data specifying the topological order. This is encoded by adding (conserved) currents minimally coupled to the CS gauge fields:

$$S_{qp} = \int a_I j^I_{qp}.$$

Alternatively, we can think of this as inserting Wilson lines $e^{\mathbf{i}\oint_W a^I q_I}$ along the trajectories W of a (probe) anyon of charge q^I.

Now let's show item 1, fractional statistics, in the simplest case with a 1×1 K-matrix. In this case, the quasiparticles are anyons of charge e/k. The idea of how this

is accomplished is called flux attachment. The CS equation of motion is $0 = \frac{\delta S}{\delta a_\mu} \sim -\epsilon^{\mu\nu\rho} f_{\nu\rho} \frac{k}{2\pi} + j^\mu_{qp}$, where j_{qp} is a quasiparticle current, coupling minimally to the CS gauge field. The time component of this equation $\mu = t$ says $b = \frac{2\pi}{k}\rho$ – a charge gets $2\pi/k$ worth of magnetic flux attached to it. Then if we bring another quasiparticle in a loop C around it, the phase of its wavefunction changes by (the ordinary Bohm-Aharonov effect)

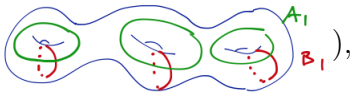
$$\Delta\varphi_{12} = q_1 \oint_C a = q_1 \int_{R,\partial R = C} b = q_1 \frac{2\pi}{k} q_2.$$

Hence, the quasiparticles have fractional braiding statistics[27] .

Now topological order property 2: # of groundstates = $|\det(K)|^{\text{genus}}$. Consider the simplest case, where $K = k$, and put the system on a torus $T^2 = S^1 \times S^1$. The gauge-invariant operators acting on the Hilbert space of the CS theory on a torus are of the form $\mathcal{F}_x \equiv e^{\mathbf{i} \oint_{C_x} a}, \mathcal{F}_y \equiv e^{\mathbf{i} \oint_{C_y} a}$ and *integer* powers of these operators. These are the operators that transport the anyons around the cycles of the torus. The restriction to integers comes from the demand that they are invariant under large gauge transformations, which take $\oint_C a \to \oint_C a + 2\pi\mathbb{Z}$. According to the CS action, a_x is the canonical momentum of a_y. Canonical quantization then implies that

$$[a_x(r), a_y(r')] = \frac{2\pi\mathbf{i}}{k} \delta^2(r - r')$$

and hence (by the BCH formula) that these flux-insertion operators satisfy a Heisenberg algebra: $\mathcal{F}_x \mathcal{F}_y = \mathcal{F}_y \mathcal{F}_x e^{2\pi i/k}$. The smallest irrep of this algebra is k dimensional, where $\mathcal{F}_x$ and $\mathcal{F}_y$ look like clock and shift matrices.

If space is a Riemann surface with g handles (like this: 

), then there are g pairs of such operators, so g independent Heisenberg algebras, all of which commute with the Hamiltonian, and hence k^g groundstates.

It is also possible to show that CS theory also exhibits the third property of long-range entanglement. See here.

This description shows a quasiparticle with charge e/k: If we stick in a quasiparticle at the origin, the equations of motion become

$$0 = \frac{\delta S}{\delta a_0(x)} = \frac{k}{2\pi} f_{xy} - \delta^2(x). \tag{2.26}$$

[27] The fractional statistics of the charge-$\frac{1}{3}$ quasiparticles of the $\nu = 1/3$ Laughlin state were finally observed experimentally just recently. Their charge had been measured using shot-noise measurements long ago.

From the relation $J^\mu = \frac{e}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu a_\rho$, the actual electric charge is then

$$\rho = e\frac{1}{2\pi}f_{xy} = \frac{e}{k}\delta^2(x). \tag{2.27}$$

Finally, we can do the (gaussian!) path integral over a to produce an effective action for $\mathcal{A}$ of the form (2.2). (Complete the square.) We find a rational Hall conductivity

$$\sigma^{xy} = t_I \left(K^{-1}\right)^{IJ} t_J \frac{e^2}{h}. \tag{2.28}$$

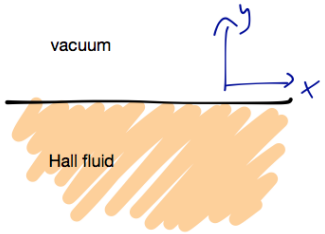
In the simplest case of $K = k, t = 1$, this is $\sigma^{xy} = \frac{1}{k}\frac{e^2}{h}$. The fact that the Hall conductivity is not an integer is not a problem – $e^{\mathbf{i}S_{\text{eff}}[A]}$ does not need to be invariant under large gauge transformations, since there are k groundstates on the torus, which are permuted by flux-threading.

[End of Lecture 6]

So far, we've shown that abelian CS theory reproduces the bulk phenomenology of some fractional quantum Hall states. Now here is a bonus: we can see what it does when the sample has a boundary in space (which actual samples in the laboratory tend to have).

Edge physics. Consider $\mathsf{U}(1)$ CS theory living on the lower-half plane.

$$S = \frac{k}{4\pi}\int_{\mathbb{R}\times\text{LHP}} a \wedge \mathrm{d}a$$



Let's work in $a_0 = 0$ gauge. We must still impose the equations of motion for a_0, which say $0 = f_{ij} = \epsilon_{ij}\partial_i a_j$. This is solved by $a = \mathbf{i}g^{-1}\mathrm{d}g = \mathrm{d}\phi$ ($g = e^{-\mathbf{i}\phi}, \phi \simeq \phi + 2\pi$), where d is the exterior derivative in just the spatial directions. This looks like a gauge transformation.

Only gauge transformations that approach $\mathbb{1}$ at the boundary preserve S_{CS}. This implies that the would-be-gauge-parameter ϕ is dynamical on the boundary. (Or equivalently, we must add a degree of freedom identical to ϕ to cancel the gauge variation of the action.)

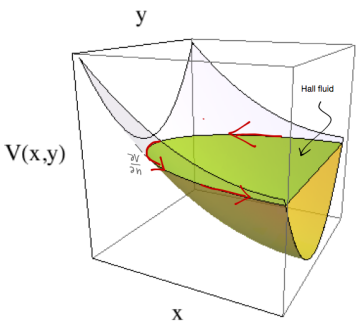
A good choice of boundary condition is: $0 = a - v(\star_2 a)$ i.e. $a_t = v a_x$. The velocity v is some non-universal UV data; it arises from a gauge invariant local boundary term, $\Delta S = \int_{\partial LHP}\frac{kv}{4\pi}a_x^2$. Plugging back into the CS action and adding the boundary term,

we find[28]

$$S_{CS}[a = \mathrm{d}\phi] = \frac{k}{4\pi}\int dtdx\left(\partial_t\phi\partial_x\phi + v\left(\partial_x\phi\right)^2\right). \tag{2.32}$$

Conclusion: ϕ is a chiral boson. $kv > 0$ is required for stability. The sign of k determines the chirality.

For the case of IQHE ($k = 1$), the microscopic picture in terms of free fermions is at right. For free fermions in a magnetic field, the velocity of the edge states is determined by the slope of the potential which is holding the electrons together. (This can be understood by considering the motion of a classical charged particle in a large enough magnetic field that the inertial term can be ignored: $q\vec{v}\times\vec{B} = -\vec{\nabla}V$, solve for v.) It is clearly not universal information.



The Hamiltonian H depends on the boundary conditions; the Hilbert space $\mathcal{H}$ does not.

I have to emphasize that a chiral theory like this cannot be realized from a local lattice model in $D = 1+1$ dimensions. There are more powerful arguments for this statement, but a viscerally appealing argument is simply to draw the bandstructure arising from any lattice Hamiltonian of free fermions. Each band is periodic in momentum space. This means that an even number of bands cross the Fermi level, and moreover that each band that crosses with positive slope must cross again with negative slope to return to its starting point. This is the essence of the *Nielsen-Ninomiya fermion doubling theorem*. An analogous argument applies in any number of dimensions. In fact, interactions provide a real loophole in the case of $D = 3+1$. But in $D = 1+1$, a nonzero chiral central charge (which in the simple examples we've discussed is just the number of right-movers minus the number of left-movers) is associated with a *gravitational anomaly*. A lattice model has zero gravitational anomaly, and this is a scale-independent quantity that must agree between the microscopic description and the EFT. The real obstruction to making a local lattice model is the anomaly.[29]

[28] In more detail, let $\tilde{d}$ denote the exterior derivative in just the spatial directions.

$$S_0[a = \tilde{d}\phi] = \frac{k}{4\pi}\int_{\mathbb{R}\times LHP} a\wedge\left(dt\partial_t + \tilde{d}\right)a = \frac{k}{4\pi}\int_{\mathbb{R}\times LHP}\tilde{d}\phi\wedge dt\partial_t\tilde{d}\phi \tag{2.29}$$

$$= \frac{k}{4\pi}\int_{\mathbb{R}\times LHP}\tilde{d}\left(\phi\wedge dt\partial_t\tilde{d}\phi\right) \overset{\text{Stokes}}{=} \frac{k}{4\pi}\int_{\mathbb{R}\times\partial LHP}\phi dt\partial_t\tilde{d}\phi \tag{2.30}$$

$$= \frac{k}{4\pi}\int_{\mathbb{R}\times\partial LHP} dxdt\phi\partial_t\partial_x\phi \overset{\text{IBP}}{=} -\int_{\mathbb{R}\times\partial LHP} dxdt\partial_x\phi\partial_t\phi. \tag{2.31}$$

[29] If, however, we break time translation symmetry, we can evade this outcome even in $D = 1+1$: for example, in a floquet system, where $H(t+T) = H(t)$, the set of energy eigenvalues is also periodic,

In the case with general K matrix,

$$S = \frac{K^{IJ}}{4\pi} \int_{\mathbb{R}\times\mathrm{LHP}} a_I \wedge \mathrm{d}a_J$$

$$S_{CS}[a^I = \mathrm{d}\phi^I] = \frac{1}{4\pi} \int dt dx \left(K^{IJ} \partial_t \phi^I \partial_x \phi^J + v_{IJ} \partial_x \phi^I \partial_x \phi^J \right).$$

(v is a positive matrix, non-universal.) This is a collection of chiral bosons. The number of left-/right-movers is the number of positive/negative eigenvalues of K.

Abelian Chern-Simons theory of the toric code. Consider now the following theory of two gauge fields with a *mutual* Chern-Simons term:

$$S[a, b] = \frac{k}{4\pi} \int d^3x \, (a\partial b + b\partial a) \ .$$

So the K-matrix is $\begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix}$. The argument above suggests that a boundary of this model should have one left-mover and one right-mover, altogether an ordinary boson in $1+1$d. In this case, we can add local, single-valued, gauge-invariant terms to the boundary (such as $\cos\phi$) to kill the edge mode. Notice that unlike the generic abelian CS theory, this system has a time-reversal symmetry acting by $a \leftrightarrow b$.

So the TO described by this K matrix allows a gapped boundary. In fact it is an effective field theory of a familiar system. To see this, consider the anyon types: they can labelled by their electric charges under the two gauge fields (a, b). Because of the CS term, the electric charge of a gets k units of magnetic flux of b attached to it, and vice versa. The well-defined operators (ferrying these anyons around) are

$$W_C = e^{\mathbf{i} \oint_C a}, \quad V_{\check{C}} = e^{\mathbf{i} \oint_{\check{C}} b}.$$

Because of the Aharonov-Bohm phase, if we place the curves in a fixed-time slice, they satisfy

$$W_C V_{\check{C}} = \omega^{\#\check{C}\cap C} V_{\check{C}} W_C.$$

These are the operators that ferry the e and m particles of the $\mathbb{Z}_k$ toric code.

As with the CS theories above, the anyons themselves are not dynamical degrees of freedom here, they are static external objects. However, we can include their dynamics, for example, by adding massive scalar fields that are charged under the CS gauge group. By varying the parameters of the potential for such scalar fields, we can move around the phase diagram. When their mass-squared passes through zero, they can condense

so we can have a band that starts below the Fermi level and ends above it, separated by $2\pi/T$ from its starting energy.

and take us to a new phase. In the toric code example, condensing a field charged under a higgses a and confines b. This takes us to the trivial phase, as condensing e should do.

Given a K-matrix theory with equal numbers of left-movers and right-movers, when can we gap out the boundary? The question is whether we can add local operators that give them a mass. For a chiral mode, $e^{\mathbf{i}\phi_R} + h.c. = \cos\phi_R$ is not a local operator because of the commutation relations of ϕ_L determined from (2.32). But $\cos(\phi_R+\phi_L)/2$ is local. A keyword for the answer is 'Lagrangian subalgebra'. Roughly, this is a subset of the anyons with the property that they are all (self and mutual) bosons, and that they braid nontrivially with everyone else. The idea is that we can *condense* these anyons (since they are bosons), and the result, because of the nontrivial braiding, is to confine everyone else, producing a trivial phase. If we do this condensation in the upper half plane, the real axis is a gapped boundary between the original TO and the trivial phase. In the case of the toric code, $a + b = \partial\phi_R, a - b = \partial\phi_L$, and both $\cos\left(\frac{1}{2}(\phi_R + \phi_L)(x)\right) = e^{\mathbf{i}\int^x a} + h.c.$ and $\cos\left(\frac{1}{2}(\phi_R - \phi_L)(x)\right) = e^{\mathbf{i}\int^x b} + h.c.$ are local. These two choices correspond to boundaries on the toric code where e and m are condensed, respectively.

Thermal Hall conductivity. Above I mentioned the notion of 'central charge' at some point. This is a concept from 1+1d CFT, which also appears in another universal observable. This observable is important for example when we are studying a system that does not couple directly to electromagnetism (such as spins) so doesn't have a Hall conductivity. So even if there is no $\mathsf{U}(1)$ symmetry, if the system is time-translation invariant, there will be energy conservation, and we can measure an energy current. One way to make an energy current is by applying a temperature gradient. The thermal Hall conductivity is defined as κ_{xy} in the linear response equation

$$I_y = \kappa_{xy}\vec{\nabla}_x T \ . \tag{2.33}$$

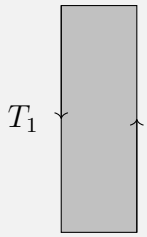
In a system with a free boson chiral edge mode like the one we found above, we can compute this. For each edge, any excitation moves along to the right with velocity v, so the energy current is $I = v\epsilon$, where ϵ is the energy density. In thermal equilibrium at temperature T, for a real chiral scalar field, this is

$$\epsilon(T) = \int_0^\infty đqvqb_T(vq) \tag{2.34}$$

where $b_T(\varepsilon) \equiv \frac{1}{e^{\varepsilon/T}-1}$ is the Bose distribution, we used $\omega = vq$ is the dispersion relation, and because the field is real $\phi_{-q} = \phi_q^\star$, the independent modes are

labelled only by positive momenta. Therefore each edge gives

$$I(T) = v\epsilon(T) = \int_0^\infty đE E \frac{1}{e^{E/T}-1} = T^2 \frac{1}{2\pi}\int_0^\infty \frac{xdx}{e^x-1} = \frac{\pi}{12}T^2 \tag{2.35}$$



Now place the left and right edges at slightly different temperatures as in the figure at left, so the total current is

$$I_y = \frac{\pi}{12}\left(T_1^2 - T_2^2\right) \simeq \frac{\pi}{6} T \Delta_x T. \tag{2.36}$$

$$\kappa_{xy}/T = c_- \frac{\pi}{12}\frac{k_B^2}{\hbar}. \tag{2.37}$$

In the last step I restored human-centric units (the $\hbar$ comes from the density of states), and placed c_- (which is 1 for this example) where it belongs in general.

To give some evidence that this is where c_- belongs, let's consider an example of a different edge theory. Suppose instead the edge is inhabited by a chiral majorana fermion field η, with action

$$S[\eta] = \int dxdt\ \eta\left(\partial_t - v\partial_x\right)\eta. \tag{2.38}$$

The critical Ising model in $D = 1{+}1$ is described by two copies of this theory, one right-moving and one left-moving. This chiral version arises as the edge theory of several interesting gapped 2+1d phases, including the $p + \mathbf{i}p$ superconductor (which has a free-fermion description and therefore is completely understood), and the non-abelian phase of the Kitaev honeycomb model. I will talk about both of those more later.

Again all the stuff just moves to the right with velocity v, and again the dispersion relation is $\omega = vq$, so each edge now gives

$$I(T) = v\int_0^\infty đqvqf_T(vq) \tag{2.39}$$

where now $f_T(\varepsilon) \equiv \frac{1}{e^{\varepsilon/T}+1}$ is the Fermi distribution. Following the same steps,

$$I(T) = T^2\frac{1}{2\pi}\int_0^\infty \frac{xdx}{e^x+1} = \frac{\pi}{24}T^2. \tag{2.40}$$

Therefore, we again find (2.37), but with $c_- = 1/2$, which is indeed the chiral central charge of a single right-moving majorana mode.

More, generally the central charge is a measure of the number of degrees of freedom. In fact, two chiral majoranas (with $c_- = \frac{1}{2} + \frac{1}{2} = 1$) can be related to a single chiral boson by bosonization.

Non-abelian CS theory. So far we've talked about CS theory with gauge group $\mathsf{U}(1)^n$. CS theory with more general gauge groups G, such as a non-abelian Lie group, can also arise as an EFT for states of matter. The non-abelian CS action looks like[30]

$$S_{\rm CS}[a] = \frac{k}{4\pi}\int_M \mathrm{tr}\left(a\wedge da + \frac{2}{3}a\wedge a\wedge a\right)$$

where now a is a Lie-algebra-valued one-form, *i.e.* $a = \sum_{A=1}^{\dim G} a^A T^A$ where T^A are generators of the Lie algebra, say in the fundamental representation.

Again invariance under large gauge transformations, $g : M \to G$, requires that k is quantized. The variation of the CS Lagrangian

$$\mathcal{L}_{CS} = \frac{k}{4\pi}\mathrm{tr}\left(a\wedge da + \frac{2}{3}a\wedge a\wedge a\right)$$

under $a \to gag^{-1} - \partial g g^{-1}$ is

$$\mathcal{L}_{CS} \to \mathcal{L}_{CS} + \frac{k}{4\pi}d\left(\mathrm{tr} dg g^{-1}\wedge a\right) + \frac{k}{12\pi}\mathrm{tr}\left(g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg\right).$$

The first term is a total derivative and integrates to zero on a closed manifold. Over any closed surface, the second term integrates to the winding number of the map $g : M \to \mathsf{G}$, and therefore the integral of the second term is an integer. We conclude that $e^{\mathbf{i}S_{CS}}$ is gauge invariant if $k \in \mathbb{Z}$.

A similar story holds for the edge modes on $M = \mathbb{R}\times\Sigma$ with $\partial\Sigma \neq 0$. Again we work in $a_0 = 0$ gauge, and the constraint $0 = \frac{\delta S}{\delta a_0} \propto f = da + a\wedge a$ is solved by $a = g^{-1}\tilde{d}g$, where $\tilde{d}$ is the spatial exterior derivative. Only g that approach $\mathbb{1}$ at the boundary of Σ are gauge redundancies, and so the boundary value of g is a physical degree of freedom. Plugging into the action, and adding a local boundary term because

[30] Full disclosure: in treating a as a Lie-algebra-valued one-form I am assuming that it is a connection on a trivial G-bundle on M. More generally, M must be covered by patches between which a is related by a gauge transformation. One way to robustly define the CS action is to realize $M = \partial N$ as the boundary of some 4-manifold N and use the fact that $\frac{1}{4\pi^2}\mathrm{tr} f\wedge f = d\omega_{\rm CS}$. Therefore the integral $\int_N \frac{1}{8\pi^2}\mathrm{tr} f\wedge f = \int_M \omega_{\rm CS} = S_{CS}[a]$ is perfectly well-defined. One shortcoming of this method is that not every M is the boundary of some N. For example, if M has a boundary, then it cannot be the boundary of something. From this point of view, the quantization of the level comes from demanding that the result for S_{CS} is independent of *which* 4-manifold we choose: the exp of the difference between the result for M and M' is

$$e^{\mathbf{i}k\left(\int_M p_1 - \int M' p_1\right)} = e^{\mathbf{i}k\int_W p_1} \tag{2.41}$$

where $W = M\cap M'$ is now a 4-manifold without boundary, and $p_1 \equiv \frac{1}{4\pi^2}\mathrm{tr} f\wedge f$. A generalization of flux quantization says that this integral is an integer multiple of 2π.

you can't stop me,

$$S_{\mathrm{CS}}[a = g^{-1}\tilde{d}g] + \int_{\partial\Sigma\times\mathbb{R}} kv\mathrm{tr}a_x^2 = k\mathrm{tr}\left(\int_{\partial\Sigma\times\mathbb{R}} (g^{-1}\partial_t g g^{-1}\partial_x g + v g^{-1}\partial_x g g^{-1}\partial_x g) + \int_{\Sigma\times\mathbb{R}} \frac{1}{12\pi} g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg\right).$$

The first two terms are just like in the abelian case. The third term is still written as a 3d integral, but it only depends on the boundary value of g. It is called a WZW term. The resulting $1+1$d field theory is a conformal field theory (CFT) called a chiral G_k WZW model. The central charge for $G = \mathsf{SU}(N)$ at level k is

$$c = \frac{k\dim G}{k+N}.$$

[End of Lecture 7]

For non-abelian G, G_k CS theory (at least for $k > 1$) realizes non-abelian topological order. For example, $\mathsf{SU}(2)_2$ is a description of the (non-abelian) Moore-Read state that seems to occur at filling $\nu = 5/2$ in GaAs (see *e.g.* p. 45 of this useful review). In case you missed it, the class I mentioned in the introduction concludes with a discussion of some bulk observables in non-abelian CS theory.

2.3 Representative wavefunctions

You'll notice that I haven't said very much about microscopic energetic questions. Quantum Hall states have been realized by now in many very different materials (in semiconductor heterojunctions, in other sandwiches made from semiconductors and insulators and metals, in graphene), and these energetic questions are pretty different in each case. Part of the reason to avoid that discussion is that it is different in each of these platforms. Quantum Hall states can even be realized in lattice models without any external magnetic field, namely Chern insulators. This just means that the hopping matrix elements are such that the bands have nonzero Chern number. A filled band with Chern number one has the same effect on the EM response as a filled Landau level.

But there is a valuable perspective more microscopic than CS theory, but still more universal than lattice details, namely representative wavefunctions. By this I mean a groundstate wavefunction *somewhere* in the same phase. There is mounting evidence that the groundstate wavefunction of a gapped phase contains all the universal data of the phase. This is particularly valuable in the case of quantum Hall states where there is a topological obstruction to exactly solvable models with exactly zero correlation length (like the toric code).

How to write down a wavefunction in the right phase? One way is to make an educated guess, which is what Laughlin did. For times when we are feeling less inspired, here is a reliable method.

Parton construction. [I recommend Sung-Sik Lee's TASI 2010 lectures] Here is a strategy for writing down wavefunctions that represent a phase with topological order. It also produces a candidate effective field theory, and has many other virtues. It is widely regarded with suspicion.

A practical point of view on what I'm going to describe here is a way to guess variational wavefunctions for fractionalized groundstates. A more ambitious interpretation is to think of the parton construction as a low-energy duality between a model of interacting electrons (or spins or bosons or ...) and a gauge theory of (candidate) 'partons' or 'slave particles'. Like any low-energy duality, it is a guess for useful low-energy degrees of freedom. The goal is to describe states in roughly the same Hilbert space[31] as the original model, in terms of other (hopefully better!) variables. The appearance of gauge fields (perhaps only discrete ones) is an inevitable side effect when there is frac-

[31] I don't mean exactly the same Hilbert space. The construction takes advantage of our ability to add in ancillary, decoupled, inert bits in changing our representative of a phase. Sometimes condensed matter physicists use the phrase "the same Hilbert space" to mean up to this equivalence, and it is in this sense that we mean it here.

tionalization of quantum numbers (spin-charge separation, fractional charge ...) in $D > 1+1$.

I will describe the construction in two steps. For definiteness, let's focus on the following example. Suppose c to be the annihilation operator for a (spinless) electron. Suppose we are interested in the (difficult) model with

$$H = \sum_{\langle ij \rangle} \left(t_{ij} c_i^\dagger c_j + h.c. \right) + \sum_{\langle ij \rangle} V n_i n_j \tag{2.42}$$

Comments:

1. We're going to talk about spinless electrons. This can be reasonable in a big magnetic field, which implies a big Zeeman splitting, so that the wrong-pointing spin states are high energy states we can ignore. (However, see the section of Girvin's review about QH ferromagnetism – the Zeeman splitting in GaAs is not that big.)

2. We can suppose that the hopping terms t_{ij} include some lattice version of the magnetic field, so $t_{ij} = te^{\mathbf{i}\mathcal{A}_{ij}}$. If you like, you could think of my lattice model here as just a discretization of electrons in the continuum in a magnetic field.

3. This kind of 'Hubbard-V interaction' is the shortest range interaction we can have for spinless fermions (since the density $n_i = c_i^\dagger c_i$ is zero or one and so satisfies $n_i^2 = n_i$).

4. To fully specify the system (2.42), we need to specify the filling – how many electrons are there per site. If the electrons fully fill some bands and $V = 0$, the system is an insulator; since there's a gap we expect this fact to persist even for nonzero V. If the filled bands have nonzero Chern number, this is a Chern insulator, and there is a quantized Hall response. It is just a lattice version of the IQHE.

5. If we partially fill some Chern bands, without V the system would be a metal. Interactions have a chance to change that. Indeed such a model can produce fractional quantum Hall groundstates. On the lattice, such a thing is called a fractional Chern insulator; as of just recently, these exist in various twisted bilayer systems.

Parton construction: step 1 of 2 (Kinematics)

Relabel states of the many-body $\mathcal{H}$ with new, auxiliary variables.

For example, a parton ansatz appropriate to the $\nu = \frac{1}{3}$ Laughlin FQH state is

$$e.g. \quad c = f_1 f_2 f_3 = \frac{1}{3!} \epsilon_{\alpha\beta\gamma} f_\alpha f_\beta f_\gamma$$

fs are complex fermion annihilation operators (they must be fermionic in order that three of them make up a grassmann operator).

Not all states made by fs are in $\mathcal{H}$. There is a redundancy: if we change

$$f_1 \to e^{i\varphi(x)} f_1, f_2 \to e^{-i\varphi(x)} f_2, f_3 \to f_3, \text{ or } f_1 \to f_1, f_2 \to e^{i\varphi(x)} f_2, f_3 \to e^{-i\varphi(x)} f_3 \ , \tag{2.43}$$

then the physical variable c is unchanged. In fact, there is an $SU(3)$ redundancy $f_\alpha \to U_\alpha^\beta f_\beta, \quad c \to \det U c$ (of which (2.43) is a Cartan (maximal abelian) subgroup). We are making the ansatz that c is a baryon.

In any state in $\mathcal{H}$, the number of actual electrons is equal to the number of partons of each color, since $c^\dagger$ creates one of each. The Lagrange multipliers imposing

$$f_1^\dagger f_1 = f_2^\dagger f_2 = c^\dagger c = \text{number of } e^-; f_2^\dagger f_2 = f_3^\dagger f_3 \tag{2.44}$$

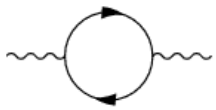
are the time components a_0 of a gauge field.

To write an action for the fs that is covariant under this redundancy, introduce the spatial components of the gauge field, a_i. Perhaps you don't like this idea since it seems like we added degrees of freedom. Alternatively, we can think of it as arising from e^- bilinears, in decoupling the $c_x^\dagger c_x c_{x+i}^\dagger c_{x+i}$ interaction by the Hubbard-Stratonovich trick. What I mean by this is:

$$\begin{aligned} e^{\mathbf{i}V \int dt n_i(t) n_j(t)} &= e^{\mathbf{i}V \int dt c_i^\dagger(t) c_i(t) c_j^\dagger(t) c_j(t)} \\ &\overset{(2.44)}{=} e^{\mathbf{i} \int dt \frac{V}{9} \sum_\alpha f_{i\alpha}^\dagger(t) f_{i\alpha}(t) \sum_\beta f_{j\beta}^\dagger(t) f_{j\beta}(t)} \\ &= \int [\prod_{\alpha,\beta} D\eta_{ij}^{\alpha\beta}] e^{\mathbf{i} \int dt \ \sum_{\alpha\beta} \left(\frac{9|\eta_{ij}^{\alpha\beta}|^2}{V} + f_{i\alpha}^\dagger(t) f_{j\beta}(t) \eta_{ij}^{\alpha\beta} + h.c. \right)} \end{aligned} \tag{2.45}$$

where η is a new complex (auxiliary) bosonic field on each link. Now let $\eta_{ij} = |\eta_{ij}| e^{\mathbf{i}a_{ij}}$ (for each $\alpha\beta$) and ignore the (massive) fluctuations of the magnitude $|\eta_{ij}| = t_{ij}$. *Voilà* the gauge field, and the parton kinetic term.

How does the practical viewpoint of constructing possible wavefunctions arise? Guess weakly interacting partons: $H_{\text{partons}} = -\sum_{ij} t_{ij} f_i^\dagger e^{ia_{ij}} f_j + h.c.$ Then fill bands of f and project onto the gauge invariant subspace.

But what about the fluctuations of a (*i.e.* we still have to do the a integral)? Microscopically, a has no kinetic term; in that sense the partons are surely strongly coupled and confined at *short* distances (of course they are – the system is made of electrons if you look closely enough). a only gets a kinetic term from the parton fluctuations, by processes like this:  . The hope is that with enough

other partons around, they can be shared and juggled amongst the electrons, so that which parton is in which electron fluctuates.

Parton construction: step 2 of 2 (Dynamics)

Such a rewrite is always possible, and there are many possibilities. The default result of such a rewriting is that the gauge theory also confines the partons at low energies. By a confining state, I mean one in which the energy cost to separate colorful partons is much larger than other scales in the problem, namely, the gap, or the inverse lattice spacing, or energies associated with chemistry (gasp). A picture of a confined state might be something like this:

In this picture, each of the sites has one of each color of parton localized to it – this is just an electron. This means there is no fractionalization and no topological order and usually leads us back to the microscopic description in terms of the microscopic degrees of freedom. (It doesn't mean the parton description is useless however; see §3).

Pure 2+1d gauge theory (without a CS term) likes to do this. Recall that the Maxwell or Yang-Mills kinetic term is an irrelevant operator according to naive dimensional analysis, if we treat the gauge field as a connection (*i.e.* something we can add to a spatial derivative). This is true even of (compact) $U(1)$ gauge theory: In terms of the dual photon σ, defined by $\partial_\mu \sigma \equiv \frac{1}{2\pi}\epsilon_{\mu\nu\rho}\partial_\nu a_\rho$, the gas of monopole instantons produces an effective potential of the form

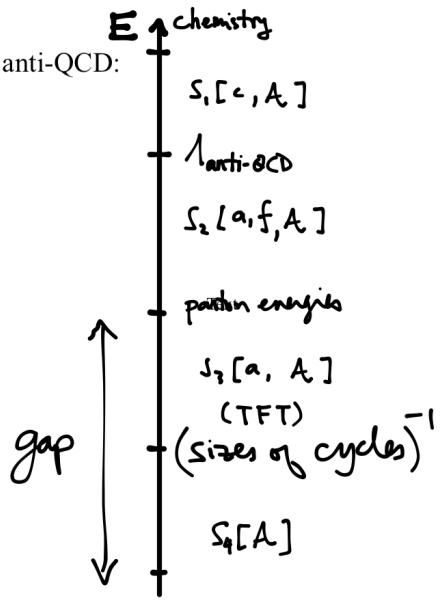
$$V_{\text{eff}} = \Lambda^3 e^{i\sigma} + h.c. = \Lambda^3 \cos\sigma.$$

Expanding around the minimum of this potential, we find a mass for σ, and hence for the photon a_μ. The statement that abelian gauge theory with compact gauge group in $D = 2+1$ likes to confine is due to Polyakov[32].

[32] See section 7.2 of these notes for more details about this from the point of view of a regularization on the lattice.

Let me emphasize again that it's *deconfined* states of parton gauge theories that are most interesting here. So we are looking for gauge theories that behave oppositely to QCD, really like *anti*-QCD, where the partons are deconfined *below* the confinement scale $\Lambda_{\text{anti-QCD}}$, as in the figure at right. Interesting states we can make this way correspond to interesting phases of gauge theory, a worthy subject.

Our discussion in this section has followed this diagram starting from the highest energies (chemistry!) and guessing the lower-energy degrees of freedom that result from the interactions of the constituents. (This dialectic between high-energy physics and condensed matter physics, of GUT and anti-GUT, is described vividly by Volovik.)



Like gaplessness, deconfinement requires an explanation. Known exceptions which allow for this:

- enough dimensions that the Maxwell term becomes marginal or relevant and we can have a Coulomb phase.
- partial Higgsing to $\mathbb{Z}_n$. Condensing electric charge makes monopoles heavy.
- lots of charged degrees of freedom at low energy. One way to describe their effects is that they produce zeromodes on the monopole configuration, and the monopoles only contribute to higher-dimension operators involving insertions of the light fields. (Interesting constraints on how many modes is enough, from strong-subadditivity of the entanglement entropy, were derived here.) Partons that are gapless at points in k-space inhabit phases called algebraic (something) liquids; the 'something' is whatever visible quantum numbers they carry, *e.g.* spin; if they happened in the model (2.42), it would be charge. If the partons form a Fermi surface, that is certainly enough (Sung-Sik Lee reviews his proof of this in the notes linked above). This is a kind of spin liquid which may have been observed in various materials in the past decade or so.
- in $D = 2+1$: the Chern-Simons term $a \wedge da$ is marginal, and can gap out gauge dynamics, as we saw in §2.2, producing a stable, deconfined, topological phase. Mutual CS terms can accomplish the same goal.

If I've forgotten some please tell me.

Parton construction of Laughlin state. Let's pursue the Laughlin example a bit further, but let's retreat to the continuum. So consider a pile of electrons in 2+1 dimensions, on a space of area A with periodic boundary conditions, in large uniform B. And let's put only enough electrons to fill a third of the lowest Landau level. That is, the number of electrons per lowest Landau level state is

$$\frac{1}{3} = \nu_e \equiv \frac{N_e}{N_\Phi(e)} = \frac{N_e}{eBA/(hc)}.$$

The fact that this band is partially filled means that if the electrons are free, the system is gapless. But this degeneracy is fragile. Any interaction between the electrons will split the degeneracy somehow.

If, on the other hand, the electron fractionalizes as $c = f_1 f_2 f_3$, then f_α carries charge $1/3$[33]. Consider then each f_α in the same external field B, and suppose the partons are free (as a first approximation). Their filling fraction is:

$$\nu_f = \frac{N_f}{N_\Phi(e/3)} = \frac{N_e}{N_\Phi(e/3)} = 3\nu_e = 1 \quad .$$

The wonderful thing about this guess is that the partons can now form a gapped state: that is, we can pretend they are free and fill their bands, so that they make a band insulator. However, because they are filling a Landau level, this band insulator is an integer quantum Hall (IQH) state. Then, integrating out the gapped partons produces a (nonsingular but nontrivial) contribution to the effective action for the gauge field: the IQH nature of the bands means that there is a Hall response for any gauge fields to which they are coupled, just as we've discussed above. This is encapsulated precisely by the CS term! [34]

[33] Actually, it is completely arbitrary how we divide up the electron charge amongst the partons; different choices differ by relabelings of the gauge group which cannot affect the physics.

[34] We showed that QHE means a CS term earlier. The massive Dirac fermion in 2+1 dimensions also has a Hall response. The mass term $m\bar{\psi}\psi$ breaks parity in $D = 2+1$. This slightly-more-microscopic calculation can be done in just the same manner as the path integral calculation of the chiral anomaly, and the $\epsilon_{\mu\nu\rho}$ arises for the same reason:

$$\begin{aligned}
\log \int D\psi D\psi^\dagger \; e^{-\int d^3x \bar{\psi}\left(\mathbf{i}\not{D}-m\right)\psi} &= \log\det\left(\mathbf{i}\not{D}-m\right) = \mathrm{Tr}\;\log\left(\mathbf{i}\not{D}-m\right) \\
&\equiv \mathrm{Tr}\;\log\left(1-\mathbf{i}\not{D}/m\right)e^{-\Box/M^2} + \mathrm{cst} \qquad \Box \equiv (\mathbf{i}\not{D})^2 = -(\partial+a)^2 - \frac{1}{2}\Sigma_{\mu\nu}f^{\mu\nu} \\
&= -\mathrm{Tr}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{\mathbf{i}\not{D}}{m}\right)^n e^{-\Box/M^2} \; .
\end{aligned}$$

where $\Sigma_{\mu\nu} \equiv \frac{1}{2}[\gamma_\mu, \gamma_\nu]$ is the rotation generator. Now expand the regulator exponential as well and

Here is the simplest route to the low-energy theory[35]. We saw above that a useful description of the IQHE is in terms of a dynamical $\mathsf{U}(1)$ gauge field in terms of which the current describing the dofs forming the IQH state takes the form $j = \frac{1}{2\pi}\epsilon\partial a$, and the Lagrangian has a term $\frac{ada}{4\pi}$. So let's introduce such a gauge field for each of the three species of parton:

$$j^{(\alpha)\mu} = \frac{1}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu b^{(\alpha)}_\rho.$$

We also have gauge fields associated with the parton gauge redundancy (2.43)[36], that I'll call a^1 and a^2. The full effective action is then

$$4\pi L = \sum_\alpha b^\alpha db^\alpha + 2A\sum_\alpha q_\alpha db^\alpha + 2a^1(db^1 - db^2) + 2a^2(db^2 - db^3) \tag{2.46}$$

where q_α are the electric charges of the partons, which satisfy $\sum_\alpha q_\alpha = 1$. $a^{1,2}$ are just Lagrange multipliers setting $b^1 = b^2$ and $b^2 = b^3$. Setting $a = b^1 = b^2 = b^3$, then, the action becomes

$$4\pi L = 3ada + 2Ada\sum_\alpha q_\alpha = 3ada + 2Ada$$

which is the effective action we advertised above.

Hence we arrive at a CS theory, like (2.14), for some particular choice of K, determined by the QH response of the partons, *i.e.* by their charges-squared times the Chern numbers of their bands.

The Hall conductivity is just sums of the contributions of the partons:

$$\sigma^{xy} = \frac{(e/3)^2}{h}\times 3 = \frac{1}{3}e^2/h.$$

The parton groundstate is $|\Phi_{mf}\rangle = \mathbb{P}\,|\text{free parton state}\rangle$, where $\mathbb{P}$ is the projection onto the gauge invariant subspace, and the free parton state is obtained just by filling the lowest Landau levels of the partons. The electron wavefunction is

$$\Psi(r) = \langle 0|\prod_i c(r_i)\,|\Phi_{mf}\rangle = \left(\underbrace{\prod_{i<j}^N z_{ij}e^{-\sum_i^N |z_i|^2/(4\ell_B^2(e/3))}}_{\nu=1\text{ slater det of charge 1/3 fermions}}\right)^3 = \prod_{i<j} z_{ij}^3 e^{-\sum_i^N |z_i|^2/(4\ell_B^2)}$$

in $D = 3$, the term that survives the trace over Dirac matrices is

$$\text{Tr}\,\frac{\mathbf{i}a_\rho\gamma^\rho}{m}\frac{1}{M^2}\left((\partial + a)^2 + \frac{1}{2}\Sigma_{\mu\nu}f^{\mu\nu}\right) = \text{sign}(m)\underbrace{\text{tr}\gamma_\rho\Sigma_{\mu\nu}}_{=\epsilon_{\rho\mu\nu}}\int d^3x\frac{1}{2}\mathbf{i}a^\rho f^{\mu\nu}\ .$$

Please see my QFT notes from s22 if that was too fast.

[35] One place in the literature where it appears is section IV of this paper.

[36] You can ask: what happened to the rest of the $\mathsf{SU}(3)$? One possible answer is that it is spontaneously broken down to this $\mathsf{U}(1)^2$ subgroup in the state we are describing.

where $z_{ij} \equiv z_i - z_j$. This is the Laughlin wavefunction. (Note that $\ell_B^2(e) = \frac{\hbar}{eB}$, so $\ell_B^2(e/3) = 3\ell_B^2$.)

Another route to the low-energy theory is as follows. Just couple the partons to an $\mathsf{SU}(3)$ gauge field a and

$$\text{integrate out the gapped partons:} \qquad \int [Df]\ e^{i\int L(f,a)} = e^{ik\mathrm{CS}(a)+\cdots}$$

The resulting low-energy effective field theory of a is $\mathsf{SU}(3)_1$ CS theory (with gapped fermionic quasiparticles). It's a non-trivial fact that $\mathsf{SU}(3)_1$ CS theory with gapped fermionic quasiparticles is dual to the $U(1)_3$ CS theory that we found earlier – they have the same groundstate degeneracy and anyon types and bulk response theory. I hope to explain more about this duality after we speak about invertible phases, since the derivation essentially involves subtracting an invertible phase from the BHS.

The Laughlin quasiparticle is the parton f with a Wilson line to make it gauge invariant.

$D = 2 + 1$ is kind of cheating from the point of view of emergent gauge fields. This is because the Chern-Simons term is a self-coupling of gauge fields that gives the photon a mass without the addition of degrees of freedom. We have seen above (in the toric code example) that this does not necessarily require breaking parity symmetry.

For this reason partons work extremely well to describe QH physics, but they are also useful for other kinds of quantum matter with strong correlations. For more about parton gauge theory I heartily recommend Sung-Sik Lee's TASI 2010 lectures. In his lectures 2 and 3, he applies this method to bosons and to spins and provides a great deal of insight.

For a long time I thought that gauge fields were only interesting for condensed matter physics when deconfinement could be somehow achieved, *i.e.*, when there is topological order. We'll see examples in §3 where even confined emergent gauge fields can do something interesting!

Attempted parable. The parton construction is a method for 'solving' non-holonomic constraints, like inequalities. In what sense were we solving such a constraint above? Suppose that the nearest-neighbor repulsion V is the biggest scale in the problem. Then we want the number of electrons on each pair of neighboring sites to be ≤ 1.

Here is a much simpler example: I can solve the condition $y > 0$ by writing $y = x^2$. So we can do a 0-dimensional path integral (integral) over $y > 0$ in terms of an

unconstrained variable x by writing

$$\int_0^\infty dy\ e^{-S(y)} = \frac{1}{2}\int_{-\infty}^\infty dx\ e^{\log|2x| - S(x^2)}.$$

In this model, the operation $x \mapsto -x$ is a gauge redundancy. In this case, it is a finite dimensional gauge group and we account for it by the factor of $\frac{1}{2}$ out front.

The extra $\log|2x|$ term in the action from the Jacobian is like a contribution from the gauge fluctuations. If I were clever enough I would illustrate deconfinement here, but I guess that isn't going to happen in zero dimensions.

[End of Lecture 8]

The parton construction makes possible

- new mean field ansatze,
- candidate many-body groundstate wavefunctions,
- good guesses for low-energy effective theory,
- accounting of topological ground-state degeneracy and edge states,
- an understanding of transitions to nearby states. (I'll give an example below.)

It has the following difficulties:

- making contact with microscopic description,
- its use sometimes requires deciding the IR fate of strongly coupled gauge theories.

Plasma analogy. [For more on this subject and the next please see Dan Arovas' QHE notes.] So the Laughlin wavefunction at filling $\nu = \frac{1}{k}$ is

$$\Psi_k(z) = \prod_{i<j} z_{ij}^k e^{-\sum_i^N |z_i|^2/(4\ell_B^2)}\ . \tag{2.47}$$

For odd k this is a fermionic wavefunction, and for even k this is a bosonic wavefunction. It has a number of very interesting properties. One is that correlation functions of operators diagonal in position space can be computed by a certain auxiliary classical system of N particles in 2d. The N-particle probability density, which determines all such correlations is:

$$|\Psi_k(z)|^2 = e^{-\beta h(\vec{r}_1 \cdots \vec{r}_N)}, \qquad \text{with } \beta = 1/k \tag{2.48}$$

and

$$h(\vec{r}_1 \cdots \vec{r}_N) = -2k^2 \sum_{i<j} \log|r_i - r_j| + \frac{k}{2\ell_B^2}\sum_i r_i^2. \tag{2.49}$$

This is the Boltzmann distribution for a two-dimensional one-component plasma: N point charges of charge $\sqrt{2}k$ interacting via the 2d Coulomb potential ϕ that satisfies

$$\vec{\nabla}^2\phi(r) = 2\pi\delta^2(r) \implies \phi(r) = -\log r \tag{2.50}$$

with a uniform background charge

$$\rho = -\frac{1}{\sqrt{2}\pi\ell_B^2} \qquad \Longrightarrow \qquad \phi_{\rm BG} = \frac{kr^2}{2\ell_B^2}. \tag{2.51}$$

To minimize energy the charges form a unifrom density (at least at large enough β), to cancel out the background charge $nk\sqrt{2} + \rho = 0$, so that $n = \frac{1}{2\pi k\ell_B^2}$. With N particles, they fit in a disk of radius R with $\pi R^2 n = N$, so $R = \sqrt{2kN}\ell_B$.

An important such observable is the *pair density distribution*:

$$ng(r) \equiv \frac{1}{N}\left\langle \sum_{i\neq j}^{N} \delta^2(\vec{r} + \vec{r}_j - \vec{r}_i) \right\rangle. \tag{2.52}$$

It is proportional to the probability of finding two particles separated by the vector $\vec{r}$. $g(r)$ determines the static structure factor (measurable in scattering experiments) by

$$\hat{s}(k) = 1 + n\int d^2r\,(g(r) - 1)\,e^{-\mathbf{i}\vec{k}\cdot\vec{r}}. \tag{2.53}$$

One funny thing is that at some large value of $\beta = 1/k$, this plasma crystallizes! That is, a delta-function peak forms in $\hat{s}(k)$. This is a result of Monte Carlo simulations (and the critical value of $k \sim 70$). This is a completely different phenomenon from the energetic competition between the Laughlin state and a possible Wigner crystal state of the electrons – it says that the Laughlin wavefunction itself exhibits translation-symmetry breaking. Note that this shows that topological order and ordinary symmetry breaking are not necessarily inimical.

Fermi statistics demands that $g(r)$ vanish as $r \to 0$. But the Laughlin wavefunction does even better:

$$g_{Laughlin}(r) \overset{r\to 0}{=} c_k\left(\frac{r}{\ell_B}\right)^{2k} = c_{k+1}\left(\frac{r}{\ell_B}\right)^{2(k+1)} + \cdots. \tag{2.54}$$

Parent Hamiltonians. The behavior of the pair correlation function (2.54) in the Laughlin states can be used to identify a family of exact parent Hamiltonians for it, that is, nice local Hamiltonians with (very) short-ranged interactions whose groundstate

is the Laughlin state. They take the uncontroversial form $H = H_{\rm kin} + H_{\rm int}$, where $H_{\rm kin} = \frac{1}{2m}\int c^\dagger(\vec{\nabla} + A)^2 c$ is the usual kinetic term and

$$H_{\rm int} = \int d^2r_1 d^2r_2 n(r_1)n(r_2)v(|\vec{r}_1 - \vec{r}_2|). \tag{2.55}$$

In this expression I emphasize that the interaction potential is central, *i.e.*$v(r) = v(|r|)$. This means that its Fourier transform is of the form

$$\hat{v}(k) = \int d^2r v(r)e^{-\mathbf{i}\vec{k}\cdot\vec{r}} = 2\pi\int_0^\infty dr r v(r)J_0(k) = \sum_{j=0}^\infty A_j(-k^2\ell_B^2)^j. \tag{2.56}$$

Therefore back in real space

$$v(r) = \sum_{j=0}^\infty A_j(\ell_B^2\vec{\nabla}^2)^j\delta^2(r). \tag{2.57}$$

For a state made only from lowest Landau-level orbitals, the kinetic energy doesn't contribute. The interaction energy of any state is

$$E_{\rm int}/N = \langle\Psi|\, H_{\rm int}\,|\Psi\rangle = \frac{n}{2}\int d^2r v(r)g(r) = \frac{n}{2}\sum_{j=0}^\infty A_j(\ell_B^2\vec{\nabla}^2)^j g(r)|_{r=0}. \tag{2.58}$$

For fermions, A_j doesn't contribute. For the Laughlin state at filling $\nu = 1/k$, the A_j with $j < k$ don't contribute. Thus, for any Hamiltonian with $A_1, \cdots A_{k-1} \neq 0$, but $A_k = A_{k+1} = \cdots = 0$, Ψ_k has $E_{\rm int} = 0$ exactly. But we can choose the nonzero A_j so that $H_{\rm int} > 0$, and thus Ψ_k is a groundstate.

Furthermore, we can argue that it's the unique groundstate (on the plane or sphere), the only homogeneous wavefunction at $\nu = 1/k$ with $E_{\rm int} = 0$. The idea is that any such state must vanish at least as fast as z_{ij}^k as any two particles approach, and therefore must have a factor of $V(z)^k$, where $V(z) \equiv \prod_{i<j} z_{ij}$ is the vandermonde determinant. But including any higher-degree polynomial $\Psi = e^{-\frac{|z|^2}{4\ell_B^2}}V(z)^k\tilde{P}(z)$ will dilute the particles below $\nu = 1/k$. (The filling is $\nu = \frac{N(N-1)}{J}$, where J is the total angular momentum, which for a holomorphic wavefunction is the degree in λ under $z_i \to \lambda z_i$. A single power of $V(z)$ has degree $\frac{1}{2}N(N-1)/2$.)

One can interpolate at fixed ν between the $H_{\rm exact}$ described above and the actual Coulomb interaction Hamiltonian governing particles in a 2DEG made from an insulating heterojunction. The claim is that there is no phase transition.

Lowest-Landau-Level Projection. To actually do this it is best to project into the LLL. Let me remind you what this means. Recall that in first-quantized notation,

in a uniform magnetic field $\vec{B} = -B\hat{z}$,

$$H_{\rm kin} = \frac{1}{2m}\Pi^2 = \hbar\omega_c(a^\dagger a + \frac{1}{2}). \tag{2.59}$$

where

$$\begin{aligned} \vec{\Pi} \equiv \vec{p} + \frac{e}{c}\vec{A}, \quad \Pi \equiv \Pi_x + \mathbf{i}\Pi_y &= -\frac{\mathbf{i}\sqrt{2}\hbar}{\ell_B}a \\ \vec{\kappa} \equiv \vec{p} + \frac{e}{c}\vec{A} - \frac{e}{c}\vec{B}\times\vec{r} \quad \kappa \equiv \kappa_x + \mathbf{i}\kappa_y &= \mathbf{i}\frac{\sqrt{2}\hbar}{\ell_B}b^\dagger. \end{aligned} \tag{2.60}$$

Here $[b, b^\dagger] = 1 = [a, a^\dagger], [a, b] = 0 = [a, b^\dagger]$. The kinetic energy depends only on the Landau level quantum number, and is independent of the so-called 'guiding center' operators b. For future reference, the complex coordinates are related to these operators by

$$z = \frac{\mathbf{i}\ell_B^2}{\hbar}(\Pi - \kappa) = 2\ell_B(a + b^\dagger), z^\dagger = 2\ell_B(a^\dagger + b)\ . \tag{2.61}$$

The derivatives are

$$\partial = \frac{1}{\sqrt{8}\ell_B}(b - a^\dagger)\bar{\partial} = \frac{1}{\sqrt{8}\ell_B}(a - b^\dagger). \tag{2.62}$$

The LLL projection $\mathcal{O} \to \mathbb{P}\mathcal{O}\mathbb{P} \equiv \mathcal{O}_L, \mathbb{P} \equiv |0\rangle\langle 0|, a\,|0\rangle = 0$. The resulting operator $\mathcal{L}_L$ is still an operator acting on the guiding center dofs. Although $[\hat{x}, \hat{y}] = 0$ in the full Hilbert space,

$$[\hat{x}_L, \hat{y}_L] = \mathbf{i}\ell_B^2\ . \tag{2.63}$$

A better way to think about this projection is just to expand the full electron annihilation operator

$$c_x = \sum_{nk} \Phi_{nk}(x)c_{nk}, \quad \mathbb{P}c_x\mathbb{P} = \sum_k \Phi_{0k}(x)c_{0k}. \tag{2.64}$$

Recall that the LLL wavefunctions are holomorphic functions times a gaussian factor. In the LLL, the operator $\bar{z}_i$ can be replaced by $2\ell_B^2\partial_{z_i}$.

2.4 Composite fermions and hierarchy states

The following line of thought, which allows us to understand other abelian FQH that actually occur (at fractions besides $\nu = \frac{1}{m}$) can be regarded as an important special case of the parton construction.

Consider what happens as we move away from the center of the plateau where the filling fraction is exactly $\frac{1}{m}$, say by varying the external magnetic field as in the famous

plot. The lowest-energy way to add charge is to create some charge-e/m quasiparticles, so when we change the filling by a finite amount, we produce some nonzero density of these objects, still in a large magnetic field. What do they do? They interact with each other somehow. If they are localized by disorder (*i.e.* form an Anderson insulator) or form a Wigner crystal, the Hall conductivity stays at the plateau value. Indeed at the lowest densities, charged particles with Coulomb interactions do form a Wigner crystal. But at higher densities, what's to stop them from forming their own FQH state?

There are several nice ways to describe this. One is called composite fermions, where the idea is to think of the factors in the Laughlin wavefunction z_{ij}^m as $z_{ij}z_{ij}^{m-1}$, one IQH wavefunction, and one boson $\nu = \frac{1}{m-1}$ (bosonic) Laughlin wavefunction – of the same variables. Fermi statistics of the electron require $\Psi \propto z_{ij}$, but the extra $m-1$ powers are something else. Regard the particle whose wavefunction this is as a fermion (the one in the IQH state) with $m-1$ units of some kind of flux attached (we'll see precisely what flux in a moment). Such a particle experiences a reduced magnetic field:

$$B^\star = B - (m-1)\rho\Phi_0. \tag{2.65}$$

Since the number of such 'composite fermions' is the same as the number of electrons, we have (if $B^\star > 0$)

$$\rho = \frac{\nu B}{\Phi_0} = \frac{\nu^\star B^\star}{\Phi_0} \tag{2.66}$$

and hence the actual filling ν is related to the filling of composite fermions $\nu^\star$ by

$$\nu = \frac{\nu^\star}{(m-1)\nu^\star + 1}. \tag{2.67}$$

If we let the composite fermions fill $\nu^\star \in \mathbb{Z}$ Landau levels (or fill bands with total Chern number $\nu^\star$), we get a gapped state with Hall conductivity given by (2.67). For $m = 3$, and $\nu^\star = 1, 2, 3\cdots$ this is

$$\nu = \frac{\nu^\star}{2\nu^\star + 1} = \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}\cdots. \tag{2.68}$$

You can see these plateaux in the famous plot. This picture suggests a very successful trial wavefunction for these fillings[37], namely:

$$\tilde{\Psi}_\nu(z) = \mathcal{P}_{LLL}\prod_{i<j} z_{ij}^2 \tilde{\Psi}_{\nu^\star}(z, \bar{z}). \tag{2.69}$$

Here $\mathcal{P}_{LLL}$ is the projector to the lowest Landau level: If we are filling multiple Landau levels, $\Psi_{\nu^\star>1}$ is no longer holomorphic; to make wavefunctions for electrons at $\nu < 1$,

[37] Define $\Psi \equiv \tilde{\Psi} e^{-\sum_i \frac{|z_i|^2}{4\ell_B^2}}$ so we don't have to write the annoying gaussian factor.

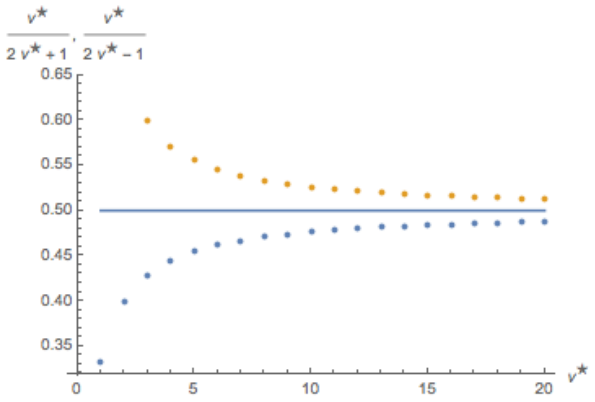
which we expect to be made just from LLL orbitals, we project this thing into the LLL by the replacement $\bar{z}_i \mapsto 2\ell_B^2 \partial_{z_i}$.

Note that there are also such 'hierarchy states' with fillings larger than $\frac{1}{m}$, which means $B^\star < 0$. In that case (2.66) is replaced by

$$\rho = \frac{\nu B}{\Phi_0} = -\frac{\nu^\star B^\star}{\Phi_0} \tag{2.70}$$

so

$$\nu = \frac{\nu^\star}{2\nu^\star - 1} = 1, \frac{2}{3}, \frac{3}{5}, \frac{4}{7}, \frac{5}{9} \cdots . \tag{2.71}$$



If we plot the fillings we achieve by this construction (2.68) and (2.71) as a function of $\nu^\star$, we get the following plot. This will be useful later. (Notice that if instead we took $m-1$ to be some other even number as our starting point, the asymptote would be $\nu = \frac{1}{m-1}$ instead.)

Now here is an explanation of the above numerology in terms of partons. We make the following parton ansatz

$$c = fb. \tag{2.72}$$

c is the electron destruction operator, and f and b are fermionic and bosonic partons respectively. This fractionalization leads to a $\mathsf{U}(1)$ gauge field which let's call a_1, under which f and b have charges 1 and -1. What are the charges of f and b under the external A? I claim that it does not matter, as long as they add up to one. Now (here we make a choice), let's put b into a $\nu = 1/2$ Laughlin state. One way to do that is to write $b = d_1 d_2$ in terms of two more fermionic partons, introducing a second gauge field a_2 under which $d_{1,2}$ have opposite charge (note that it again doesn't matter how we distribute the a_1 charge between d_1 and d_2), and let each of d_1 and d_2 fill a Chern band. The full table of charges is

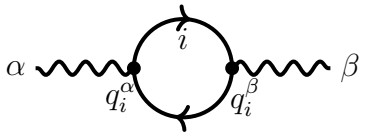
	A	a_1	a_2
d_1	q_1	0	1
d_2	q_2	1	-1
f	q_f	-1	0

Now let's integrate out d_1 and d_2. In general, integrating out a field d_i with charges q_i^α under gauge field a^α filling chern bands with total chern number c_i leads to an effective

Lagrangian

$$L_i = \frac{c_i}{4\pi} \sum_{\alpha\beta} (q_i^\alpha a_\alpha)\, d \left(q_i^\beta a_\beta \right). \tag{2.73}$$

A useful mnemonic is the following diagram:



For convenience of writing I'll call $a_0 \equiv A$, so the couplings to the external gauge field are included in (2.73). We ignore any other interactions between the partons, so the contributions of multiple fields just add. So integrating out $d_{1,2}$ leaves us with

$$L_{\text{eff}} = \frac{c_1}{4\pi} (a_2 + q_1 A)\, d\, (a_2 + q_1 A) + \frac{c_2}{4\pi} (a_1 - a_2 + q_2 A)\, d\, (a_1 - a_2 + q_2 A) + L(f, a) \tag{2.74}$$

We can think of these CS terms as attaching flux to the remaining parton f, which is the composite fermion.

Now we can decide what to do with f; depending on what we do, we'll find different states, with different Hall response. If we let f fill Chern bands with total Chern number $\nu^\star$, we can integrate out f, too, in the same way, and get $L_{\text{eff}} = \sum_i L_i$ with L_i as in (2.73). If we set $\nu^\star = 1$ we reproduce exactly our earlier construction of the Laughlin state. To find the general Hall response, we can just solve the equations of motion for $a_{1,2}$; these are linear equations that determine $a_{1,2}$ in terms of A. Plugging back into (2.74) then gives an action of the form $L = \frac{\nu}{4\pi} A dA$ with, setting $c_1 = c_2 = c$[38],

$$\nu = \frac{c\nu^\star}{c + 2\nu^\star}. \tag{2.75}$$

If I set $c = 1$, this is exactly (2.68). And the parton trial wavefunction is just (2.69).

Hierarchy and K-matrices. There is another way to construct the hierarchy, by combining the logic with which we began §2.2 with the picture with which we began our discussion of the hierarchy states. So we've argued that the EFT in terms of CS gauge fields is an inevitable consequence of $\mathsf{U}(1)$ symmetry, a gap, and broken time-reversal symmetry. That's the same situation we're in when we add a density of quasiparticles to move away from the center of the plateau (assuming we make a gapped state). So the same logic suggests that we write also the quasiparticle current

$$j^\mu_{qp} = \frac{1}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu \tilde{a}_\rho \tag{2.76}$$

[38] In this calculation, it is an extremely useful check to leave the electric charges of the partons arbitrary, and make sure that the answers only depend on the sum of their charges, which is fixed by the charge of the electron.

in terms of a (new) CS gauge field! You can't stop me. The minimal-coupling term $a_\mu j^\mu_{qp}$ becomes a mixed CS term. The full action (epsilon tensors are implicit) is

$$4\pi L = kada + 2Ada + 2ad\tilde{a} + \tilde{k}\tilde{a}d\tilde{a} + \cdots ,$$

where I've added a CS term to describe the dynamics of the quasiparticles, by the same logic as before. This is of the form (2.14) with a 2×2 K-matrix, $K = \begin{pmatrix} k & 1 \\ 1 & \tilde{k} \end{pmatrix}$ with charge vector $t = (1, 0)$. If we integrate out $\tilde{a}$ and a in this action, we'll find Hall conductivity

$$\nu = \frac{1}{k - \frac{1}{\tilde{k}}}.$$

If $\tilde{k} = 2$, we reproduce our previous hierarachy states. These FQH states of composite fermions are observed here. This CS description determines the groundstate degeneracy (to be $|\det K|^g$ on a genus-g surface) and the charges and statistics of the quasiparticle excitations above this state: minimally couple particles to a and $\tilde{a}$. The quasiparticle charges that one computes this way are observed experimentally, *e.g.* at $\nu = \frac{2}{5}$.[39]

Exercise: explain in this language how to get the states with filling larger than a half that we described earlier.

Now, why should we stop here? Moving away from the middle of the plateau of one of these states, there is a density of those quasiparticles whose charges I just mentioned. Why can't they, too, form a QH state? Then we would write their current in terms of yet a third CS gauge field, and we would arrive at a description with a 3×3 K-matrix, and a Hall conductivity that was a continued fraction with a third level. You can also see these states in clean samples.

[End of Lecture 9]

This construction gives K-matrices with integers on the diagonal and ones on the next-to-diagonal. The K-matrix $K = \begin{pmatrix} m_1 & n \\ n & m_2 \end{pmatrix}$ and charge vector $t = (1, 1)$ also describes states that exist (named after Halperin), and whose filling fraction, quasiparticle data and GSD you can now compute.

We can relate theories with different K-matrices by a relabelling of fields, $a^I \to W^I_J a^J$. But in order to preserve the flux quantization, the matrix W must have $|\det W| = 1$, that is $W \in \mathsf{GL}(n, \mathbb{Z})$. Thie resulting equivalence relation acts by

$$K^{IJ} \to W^I_K K^{KL} W^J_L, \quad t^I \to W^I_J t^J. \tag{2.77}$$

[39] A nice resource for the literature on experimental studies of FQHE up to 2004 is the slides of Willett here.

Quantum Hall Metal. Now comes some real magic: look again at the limit as $\nu^\star \to \infty$ of the plot above. What happens when $\nu = \frac{1}{2}$ (or more generally $\nu = \frac{1}{m-1}$) in the composite fermion construction? Look at (2.65): when $\nu = \frac{1}{m-1}$, the composite fermion sees no magnetic field at all! This is despite the fact that half-filling happens (for realistic electron densities in a 2DEG) at enormous magnetic fields like 25 Tesla. This suggests that in our parton ansatz (2.72) the composite fermion f can just fill a Fermi sea, and would move in straight lines in response to an external electric field. This would be a gapless, metallic state. And indeed that is what is seen near $\nu = \frac{1}{2}$! A trial wavefunction for this amazing state is then just

$$\tilde{\Psi}(z) = \mathcal{P}_{LLL} \prod_{i<j} z_{ij}^2 \det_{ij} e^{\mathbf{i} k_i \cdot r_j} \tag{2.78}$$

where the $\{k_i\}$ label the N lowest energy modes. I think this wavefunction was first written down by Read and Rezayi, but Halperin, Lee and Read explained a lot of the physics, so I call it the HLR state. A similar story occurs (in experiments as well) at other fillings of the form $\nu = \frac{1}{m-1}$ for m an odd number.

But there's more: let's start at the HLR state at $\nu = \frac{1}{2}$ and vary the magnetic field (at fixed electron density) away from the special value where the composite fermions see no field $B = B_{\nu=\frac{1}{2}} + \delta B$. What happens when we subject a Fermi surface to a magnetic field? We get quantum oscillations: various quantities, including the conductivity, are periodic with period $1/\delta B$. But this periodic structure is exactly where the hierarchy plateaux appear! (That is: if $\frac{1}{\delta B} = \pm \frac{\nu^\star}{\rho \Phi_0}$ for some integer $\nu^\star$, the actual filling fraction is $\nu = \frac{\nu^\star}{2\nu^\star \pm 1}$.) This means we can regard all of the hierarchy states, including the original Laughlin states, as extreme manifestations of quantum oscillations.

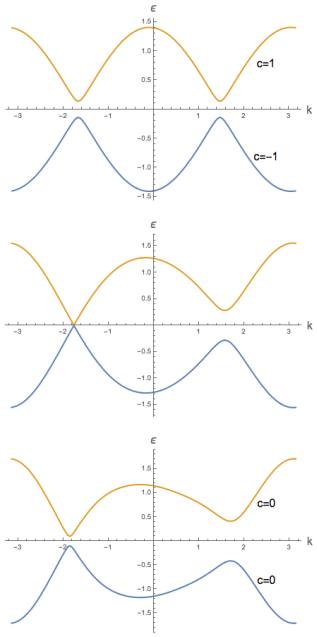
Notice that the composite fermions at the Fermi surface still interact with a Chern-Simons gauge field. This is a non-Fermi liquid: a metallic system that is not described by ordinary Fermi liquid theory.

Incompressible states at even denominators. Finally, one thing that fermions at a Fermi surface like to do is pair up and superconduct. What happens if the composite fermions in the HLR state form a superconductor? Well, by Fermi statistics, it has to be a p-wave superconductor because the composite fermions are spinless (we're assuming the electrons are completely spin-polarized). If they form a $p + \mathbf{i}p$ (nodeless) superconductor, the projection of the BCS wavefunction is the Moore-Read state, which is a non-abelian topological order. (Perhaps more on this wavefunction later. This is not how it is was first discovered.) Whether or not the composite fermions pair up depends on the consequences of the CS interactions, and on microscopic details. (For an analysis of this drama, see here.) In the half-filled lowest Landau level, we see a metallic state, but in the half-filled *third* Landau level (*i.e.* at $\nu = \frac{5}{2}$), there is indeed

an incompressible state, which could be the Moore-Read state. (At $\nu = \frac{3}{2}, \frac{7}{2}$, one finds instead states that break translation symmetry – stripes. For why this happens, see Misha Fogler's article here.)

Recently there has been some further development in our understanding of what happens near $\nu = \frac{1}{2}$. The state (2.78) makes a distinction between a half-full and a half-empty Landau level – it is not symmetric under the interchange of particles and holes. But experiments show that the physics is particle-hole invariant. Son suggested a way to write down a particle-hole symmetric state by starting with Dirac composite fermions, rather than the non-relativistic ones that I've been talking about. This line of inquiry has also had some important consequences for our understanding of the gapped and likely non-abelian state at $\nu = \frac{5}{2}$. Recent measurements (of thermal Hall conductivity (which counts the number of edge modes, including neutral ones) and of interfaces with other states) favor a particle-hole symmetric version of the Moore-Read state (called the PH-Pfaffian).

Transitions to neighboring phases. What happens if the Chern numbers of the bands occupied by the partons $d_{1,2}$ are not both 1 (let's call it $(c_1, c_2) = (1, 1)$)?



It is not hard to describe a transition where Chern numbers change: just vary the bandstructure until the bands touch at a Dirac point. Moving past that point, the two touching bands exchange Chern numbers. (I like to think of it like a crossover event during meiosis. The pictures are very similar.) If one is filled and one is empty, this will change the resulting contribution to the Hall response. For example, imagine that in the figures at right the blue band is filled. And we see that a theory of the transition is a Dirac fermion coupled to Chern-Simons gauge fields.

Consider just the theory of a boson b, with the ansatz $b = d_1 d_2$, where d_1 and d_2 fill Chern bands with Chern numbers (c_1, c_2). The effective Lagrangian for the gauge field a_1 gluing them together is

$$L = \frac{c_1}{4\pi} (a_1 + q_1 A)\, d\, (a_1 + q_1 A) + \frac{c_2}{4\pi} (-a_1 + q_2 A)\, d\, (-a_1 + q_2 A) \ . \tag{2.79}$$

Integrating out a_1 gives Hall response

$$\nu_b = \frac{c_1 c_2}{c_1 + c_2}. \tag{2.80}$$

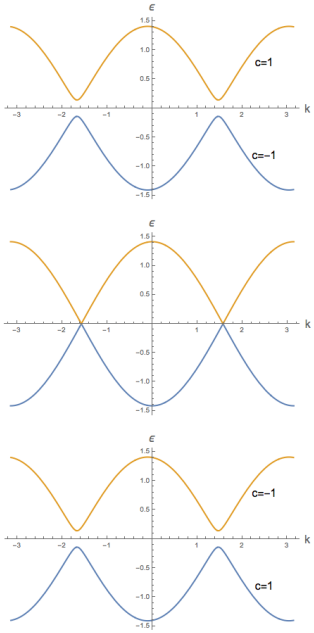
Several cases are interesting. If $(c_1, c_2) = (1, 0)$, the calculation above gives $\nu = 0$, no Hall response. But the state of the boson $b = d_1 d_2$ is still an insulator, there is still an energy gap. This is a (weird!) description of a featureless Mott insulator.

Another very interesting case is when $(c_1, c_2) = (1, -1)$, in which case (2.80) blows up. This is a signal that the theory is actually gapless. In fact it is a dual description of the superfluid phase of the boson. To see this, go back to (2.79). When $c_1 = -c_2 = 1$, the CS term for a_1 cancels and it reduces to

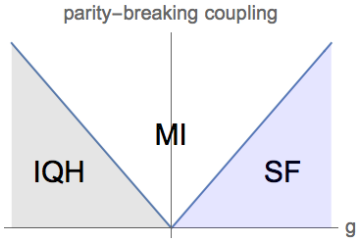
$$L = \frac{1}{2\pi} a dA + f^2 \tag{2.81}$$

which says that there is no CS term for a, and I've added back the Maxwell term, since it's now the leading term governing the dynamics of a. Earlier I said that pure compact $\mathsf{U}(1)$ gauge theory in $D = 2 + 1$ confines because of monopole instantons. But the first term in (2.81) says that magnetic flux of a carries charge under the $\mathsf{U}(1)$ global symmetry. This means the operator $e^{i\sigma}$ is not $\mathsf{U}(1)$ symmetric (σ shifts) and can't be added to the action. This describes a gapless theory, where the photon is the goldstone boson, $\partial a = \epsilon \partial \sigma$, for spontaneously breaking the $\mathsf{U}(1)$ symmetry.[40]

In the presence of some lattice symmetry (like parity, $k \to -k$) that forces two band-touching points at the same point in the phase diagram, it can force the Chern number of the filled band to change from $c = -1$ to $c = 1$. If this happens to d_1 (fixing $c_2 = 1$) we have a direct transition from the Laughlin $\nu = \frac{1}{2}$ state to a superfluid. Notice that the critical theory involves two species of Dirac fermions.



We can thereby produce a boson phase diagram that includes these three phases.



(The state with $(c_1, c_2) = (-2, 1)$ is also interesting – it is a boson IQH state with $\nu = 2$.)

Now we can use this to make a theory of electronic transitions out of the HLR phase, for example to a Mott insulator. In our description of the HLR phase, the Hall response comes entirely from that of b. When b goes from the $\nu = \frac{1}{2}$ state to the Mott

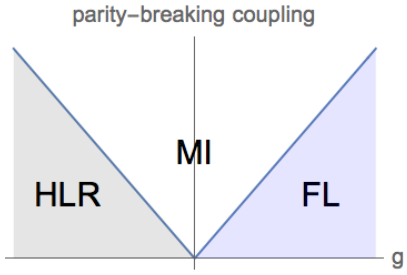
[40] Another way to think about (2.81) is to integrate out a and see what effective action we get for A. The equation of motion says roughly $\partial^2 a \sim \partial A$, and plugging back in we get $L_{\text{eff}}[A] \sim A\partial\left(\frac{1}{\partial^2}\right)\partial A \sim A^2$, a Meissner mass for A, as we should have in a superfluid.

insulator, the full electron state goes from HLR to a (fermionic) Mott insulator. (And indeed, the general formula, if f also fills Chern bands with Chern number $\nu^\star$, is

$$\nu = \frac{c_1 c_2 \nu^\star}{(c_1 + c_2)\nu^\star + c_1 c_2}. \tag{2.82}$$

Again if $(c_1, c_2) = (1, 0)$, we find no Hall response.)

With the fractionalization $c = bf$, if b instead forms a superfluid state, it completely higgses the $\mathsf{U}(1)$ gauge field gluing together b and f. The result is that we can forget about both b and the gauge field, and we get an ordinary Fermi liquid. So in this way we can describe a phase diagram containing the HLR state, Mott insulator and ordinary metal. You could imagine moving around in such a phase diagram by applying a periodic potential to a 2DEG, at fixed electron density and magnetic field. Recently such a transition seems to actually have been realized in twisted bilayer dichalcogenides. See here for an update of the theory.

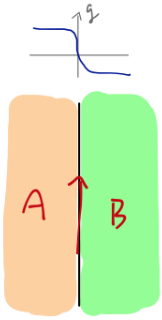


3 Symmetry-protected topological phases

We begin with a low-tech example-based response-theory point of view on symmetry-protected topological (SPT) states. Sources for this discussion include Senthil's review. Useful and brief is the second part of this review. See also this Journal Club for Condensed Matter Physics commentary by Matthew Fisher.

We are interested in possible ways to distinguish phases of quantum matter (without having to check every possible adiabatic path between them in the infinite-dimensional space of Hamiltonians). Here is a simple yet still-interesting question (which has been very fruitful in the past decade or so): how do we distinguish phases of matter that preserve their symmetry group G but *don't* have topological order?

One possible answer: put them on a space with boundary, *i.e.* an interface with vacuum or with each other. Quantized (hence topological) properties of the surface states can be characteristic of distinct phases.

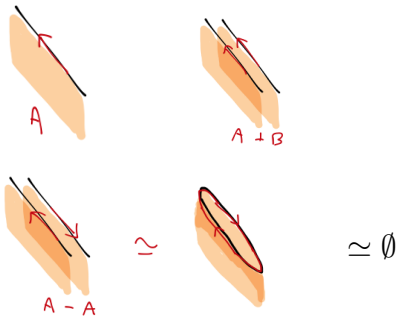


The rough (and not entirely correct) idea is: just like varying the Hamiltonian in time to another phase requires closing the gap $\mathbf{H} = \mathbf{H}_1 + g(t)\mathbf{H}_2$, so does varying the Hamiltonian in space $\mathbf{H} = \mathbf{H}_1 + g(x)\mathbf{H}_2$. (We've already seen an exception to this in the interface between the toric code and vacuum.)

Notice that the possibility of characterizing a state in this way requires that the $D-1$-dimensional edge theories involved must not be realizable intrinsically in $D-1$ dimensions, consistent with the symmetries involved. If they could be so realized, preserving G, we could paint them on the surface of any bulk state, and they therefore could not be characteristic of the bulk state. Therefore there is a correspondence between SPTs in D dimensions and obstructions to symmetry-preserving regulators of quantum field theory (QFT) in $D-1$ dimensions.

Definition: a gapped groundstate of some local Hamiltonian $\mathbf{H}$ preserving G which is distinct from any trivial product state in the space of G-symmetric Hamiltonians, but which is trivial in the space of all Hamiltonians, is called a SPT (symmetry-protected topological) state with respect to G. (Instead of demanding that the state is trivial if we break G, why don't we just say 'without TO'? There are some subtleties about this to which we'll return.) Inevitably I will also use the term 'SPT' for the phase of matter of which such a groundstate is a representative.

The deformation classes of SPT states form a group. The group law is *stacking*. Stacking A and B means we make the system whose Hilbert space is the tensor product of those of A and B, and whose Hamiltonian is the sum. (Then we are allowed to deform the Hamiltonian, preserving the gap, as usual.) The inverse of a state A, which we can call $-A$, is (usually) the mirror image. A cartoon of why this is the case is given at right.

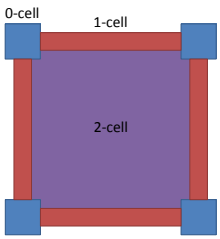


To be clearer, the definition of the inverse state of A is a state with the property that there exists a finite-depth unitary circuit U such that

$$U\,|\psi_A\rangle \otimes |\psi_{-A}\rangle = \text{product state}. \tag{3.1}$$

Please don't take the cartoon too seriously; its main point is to show that the edge modes can be cancelled. In fact, if we consider a state with gapless edge modes on a space with boundary, the operation (3.1) cannot be done with a finite-depth circuit (by the Lieb-Robinson bound) – it's only on a closed spatial manifold that we demand (3.1). Note that with topological order, even if we can gap out the edge states, there is still stuff going on (*e.g.* anyons) in the bulk, and no such ψ_{-A} or U exist: states with topological order do not form a group under stacking. [End of Lecture 10]

Naturally, a state that has an inverse is called *invertible*. It would have been wise of me to *define* SPTs to be invertible states, and I will do so below. In any case, one can ask whether the absence of topological order, for example the vanishing of the topological entanglement entropy (TEE), is sufficient to imply the existence of an inverse state. A sketch of why this might be the case is the following: a nonzero TEE is an obstruction to the reconstruction of the global groundstate from the density matrices on subsystems. If all TEEs vanish, in a system with no edge states (such as A plus its mirror image with gapped boundary conditions), we can therefore assemble the full groundstate wavefunction by the following procedure, starting from the state on an array of balls (each large compared to the correlation length), which we'll call the 0-skeleton.



Then we can connect these balls by tubes (also topologically balls) acting like 1-cells in a cell complex; then we can glue in 2-cells, and so on. The resulting process defines a finite-depth unitary circuit that produces the state from a product state, and the idea is that TEE and edge modes are the only obstructions to doing this. (More details can be found in §3.1 here. A related construction is described here, in section 1.2.).

How to characterize these states more precisely? Many possibilities have been explored so far. To get started:

1. If $\mathsf{G} \supset \mathsf{U}(1)$, we can study the response to an external electromagnetic (EM) field. Just like the integer quantum Hall effect, this will involve some quantized coefficients.

Without a $\mathsf{U}(1)$ (or other continuous symmetry), we need something else. (And two states with the same quantized EM response may be distinct for some other reason.)

2. What happens if you gauge G? In general this produces a new state, with TO (or gapless). That state can be used as a label on the SPT. (This works both ways: labelling TO or gapless states is the hard part.) I'm not going to say more about this here.

3. Weird, seemingly-forbidden stuff at the surface. For example, the surface could have half-integer Hall response, but a gap and no topological order. Or the surface can have patterns of fractionalization that violate associativity constraints. The general name for these phenomena, which mean that the effective action of the edge theory is not gauge invariant, is *anomalies*. And the phenomenon where the presence of the bulk theory allows for would-be-impossible things on the boundary, is called *anomaly inflow*.

This concept of anomaly has many manifestations. A useful definition of an anomaly is any obstruction to gauging a symmetry, that is, to treating the background fields as dynamical variables. The name comes from high energy physics. The perspective there is that there can be symmetries of the classical action that are not symmetries of the path integral (*i.e.* of the path integral measure). What this really means is that such a field theory cannot be realized on a lattice in such a way that the symmetry can be gauged. We'll have an occasion to review some examples of this in a bit.

The more precise definition of anomaly is the following. A system has a G anomaly if it has a G symmetry, but when we couple it to background fields $\mathcal{A}$ for G, the resulting partition function is not gauge invariant

$$Z[\mathcal{A} + \mathbf{i}g^{-1}dg] = e^{\mathbf{i}\varphi[g]}Z[\mathcal{A}] \tag{3.2}$$

in a way that can't be fixed by adding local functionals of $\mathcal{A}$ to the action[41]. Here is one reason that people love this concept. Notice that directly from this definition you can

[41] About these last qualifying words: consider the example of a superfluid, which has a $\mathsf{U}(1)$ symmetry with current $j_\mu = \partial_\mu \varphi$, where $\varphi \to \varphi + \theta$ under the symmetry. The naive minimal coupling $\int \mathcal{A}_\mu j_\mu$ is *not* gauge invariant. But this is not anomaly, because we can cancel the variation by adding a term proportional to $\int \mathcal{A}^2$. (This is sometimes called a 'contact term' for some reason.) Altogether, the gauge-invariant Lagrangian looks like $L = (\partial_\mu \varphi + \mathcal{A}_\mu)^2$.

see that anomaly is an RG-invariant notion, since by definition an RG transformation preserves the partition function. Thus, any candidate IR description of some given UV theory must have the same anomalies.

Since anomaly is defined to mean some failure of gauge invariance of the effective action, you can already see that it has a close connection to the LSMOH ideas. In particular, a theory with nontrivial anomalies cannot have a trivially-gapped symmetric groundstate.

The definition above gives a group of SPTs for each choice of symmetry group G and dimension d. This group is abelian. What can it be? Essentially using the Landau what-else-can-it-be method for suitably-general choices of background fields, this group has been determined. It is the Anderson dual of the G-equivariant bordism group in $d+1$ dimensions. Later I will try to make some of these words meaningful. For now, I want to convey the nature of this classification: what's been classified is the set of possible effective actions of the background fields. The classification we have does not produce representative wavefunctions.

Perhaps I should emphasize here that there is an important divide between SPTs that are made from only microscopic bosons, where the Hilbert space is a tensor product, and SPTs with microscopic, gauge-invariant fermions, where distant local operators are allowed to anticommute. With even free fermions, there are many examples of SPTs (called topological insulators) realized by topological bandstructure. In contrast, SPT states of bosons (especially in $D = 3+1$) are particularly interesting partly because they *require* interactions. (Non-interacting bosons just form a boring superfluid.) Note that 'states of bosons' also means states of spins by a simple mapping of states.

Why study SPTs? For me, the point of studying SPT states is several-fold. The general question being asked here is basically: how can symmetries be realized on many-body systems? Clearly this is an interesting question of broad applicability.

One motivation is that by understanding all the ways in which symmetries can be realized in quantum systems, we can learn how to gauge them, and thereby make new states with topological order. Are *all* topologically-ordered states obtained this way? Maybe not. (*e.g.* how to think about fibonacci anyons in this language? Actually the parton construction produces a gauge theory that can be regarded as gauging an SPT.)

A second reason is that each SPT state in D dimensions represents an obstruction to regularizing a QFT in $D-1$ dimensions with certain properties. Such a QFT does not come from a local, symmetric lattice model. A famous example of such an obstruction is the Nielsen-Ninomiya fermion-doubling theorem. Some new anomalies have been found this way. This perspective has also added some new vigor to the quest to define the Standard Model (a chiral gauge theory) on the lattice.

One outcome of this direction is that the surface of an SPT facilitates the emergence of supersymmetry from a lattice model. Supersymmetry is a great idea still looking for its place in nature. The basic idea is that the supersymmetric fixed point has two relevant deformations, but one of them is forbidden by a funny (SPT-edge) realization of time-reversal symmetry.

Finally, a zeroth-order reason to study SPTs is that they exist. If we are going to understand all phases of matter, we need to understand them.

3.1 EM response of SPT states protected by $\mathbf{G} \supset \mathbf{U}(1)$

3.1.1 $D = 2+1$, $\mathbf{G} = \mathbf{U}(1)$

This is what we did in §2.1. Our conclusion was $\sigma^{xy} = \lim_{\omega\to 0} \frac{\langle j^x j^y\rangle|_{k=0}}{\mathbf{i}\omega} = \nu\frac{e^2}{h} = \frac{\nu}{2\pi}$, and ν is quantized if there is no fractionalization (and ν is even in a model of bosons without fractionalization).

On the other hand, even without TO, the integer ν provides a label on a phase of matter, since it cannot vary continuously.

3.1.2 K-matrix construction of SPT states in D=2+1

Let's talk about effective-field-theory realizations of such states.

(This discussion is from here.) Recall our description of abelian FQH states, with effective action

$$S[a_I] = \sum_{IJ} \frac{K_{IJ}}{4\pi} \int a_I \wedge \mathrm{d}a_J.$$

The actual particle (electron or boson) current, representing the global $\mathsf{U}(1)$ symmetry, is $J_\mu = \frac{1}{2\pi}\epsilon_{\mu\nu\rho}\partial_\nu a^I_\rho t_I$. That is: coupling to the external gauge field is $\Delta L = \mathcal{A}t^I\epsilon\partial a^I/(2\pi)$. We showed that this produces a Hall response

$$\sigma^{xy} = \frac{1}{2\pi} t^{-1}K^{-1}t.$$

The number of groundstates on a genus-g surface is $|\det K|^g$. If we choose K with $|\det K| = 1$, this suggests that there is no topological order. We can check that this is the case by examining the spectrum of quasiparticles. Quasiparticles are labelled by an integer vector l^I specifying their coupling to a^I.
self (exchange) statistics: $\theta = \pi l^T K^{-1} l$.
mutual statistics: $\theta_{12} = 2\pi l_1^T K^{-1} l_2$.

external U(1) charge of quasiparticle l: $Q = t^T K^{-1} l$.

To make an SPT state, we must ensure that all these quantum numbers are multiples (not fractions!) of those of the microscopic constituents.

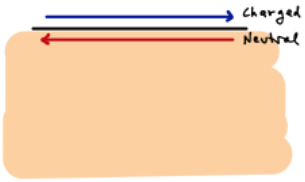
To describe a boson IQH state, consider $K = \sigma^x$. Think of the two states as like two 'layers' or species of bosons, so we can take statistics vectors $l_1 = (1,0), l_2 = (0,1)$. These are self bosons and mutual bosons. If we take the charge vector to be $t = (1,1)$ (both species carry the charge) then this state has $\nu = 2$.

For $\nu = 2$ boson IQH: $K = \boldsymbol{\sigma}^x, t = (1,1)$. Let's put the system on the lower-half plane and look at the edge theory. Writing it in terms of the eigenvectors $\phi^\pm$ of K, we have

$$S_{CS}[a^I = \mathrm{d}\phi^I] = \frac{1}{4\pi}\int dtdx \left(\partial_t\phi^+\partial_x\phi^+ - \partial_t\phi^-\partial_x\phi^- + v\sum_\pm \left(\partial_x\phi^\pm\right)^2\right)$$

This shows that $\phi^\pm \equiv \frac{1}{\sqrt{2}}\left(\phi^1 \pm \phi^2\right)$ are left- and right-moving respectively.

Conclusion: it's just a non-chiral free boson (at the SU(2) radius). This is relatively ordinary in the sense that it arises as the low-energy effective theory of the (gapless) spin-half Heisenberg chain.



How is the thing at the edge of the $\nu = 2$ boson IQHE special? The specialness arises in the way the U(1) symmetry is realized – in the coupling to the external gauge field: since $t = (1,1)$,

$$L \ni \mathcal{A}^\mu\partial_\mu\left(\phi^1 + \phi^2\right) \propto \mathcal{A}^\mu\partial_\mu\phi^+ \ .$$

Specifically, although $c_L = c_R$, only the left mover ϕ_+ carries the U(1) charge. This means that preserving U(1), we can't backscatter, that is we can't add to the action (local) terms like $\Delta S = g_\pm \cos\left(\phi^+ \pm \phi^- + \alpha\right)$ (α is a constant) which would lift the edge states. (Such terms made from just ϕ^+ would not be local.) This means the U(1) *protects* the edge states.

The paper linked above suggests some possible microscopic realizations of this state, which seem to be borne out by numerics.

3.1.3 $D = 3+1$ and $\mathbf{G} = \mathbf{U}(1) \ltimes \mathbb{Z}_2^T$

The effective field theory for any 3+1d insulator, below the energy gap, has the following form

$$S_{\text{eff}}[\vec{E}, \vec{B}] = \int \mathrm{d}^3x\mathrm{d}t \left(\epsilon \vec{E}^2 - \frac{1}{\mu}\vec{B}^2 + \frac{\theta e^2}{4\pi^2}\vec{E}\cdot\vec{B} + \mathcal{O}(\vec{E},\vec{B})^4 \right) \tag{3.3}$$

where ϵ, μ are the dielectric constant and permittivity, and α is the fine structure constant. (In saying that the next corrections go like the fourth power of E, B I am assuming that we can approximate the material as isotropic and using rotation invariance. Without this assumption, I should have written $\epsilon_{ij}E_iE_j$ and so on. Moreover, if the system has a nontrivial ferromagnetic or ferroelectric response, we could have linear terms: $\vec{m}\cdot\vec{B} + \vec{p}\cdot\vec{E}$ (and cubic terms).) Flux quantization implies that

$$\frac{e^2}{4\pi^2}\int_{M_4} d^4x \vec{E}\cdot\vec{B} = \frac{e^2}{8\pi^2}\int_{M_4} F\wedge F \in \mathbb{Z}$$

is an integer for any closed 4-manifold M_4 ($\partial M_4 = \emptyset$) that admits a spin structure, *i.e.* on which we can put fermions. This means that the partition function of a fermionic system on such a manifold is periodic

$$Z(\theta + 2\pi) = Z(\theta)$$

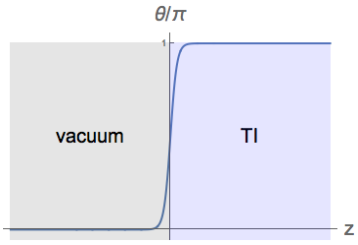
and hence the spectrum on a closed 3-manifold ($M_4 = S^1 \times X_3$ is spin) is periodic in θ. (As we will discuss, shifting θ by 2π is not so innocuous on a space with boundary or for the wavefunction.)

Time reversal acts by

$$\mathcal{T} : (\vec{E}, \vec{B}) \to (\vec{E}, -\vec{B})$$

which means $\theta \to -\theta$, which preserves the spectrum only for $\theta \in \pi\mathbb{Z}$. So fermionic time-reversal-invariant insulators are labelled by a quantized 'magneto-electric response' $\theta/\pi = 1$.

Now consider what happens on a space with boundary, like for any actual chunk of material. The interface with vacuum is a domain wall in θ, between a region where $\theta = \pi$ (TI) and a region where $\theta = 0$ (vacuum).



[End of Lecture 11]

Letting θ depend on space, the electromagnetic current derived from (3.3) is

$$j^\mu_{EM} = -\frac{e^2}{8\pi^2}\epsilon^{\mu\alpha\gamma\delta}\partial_\alpha\theta\partial_\gamma A_\delta + \cdots \tag{3.4}$$

where the $\cdots$ indicate contributions to the current coming from degrees of freedom at the surface that are not included in (3.3). If we may ignore the $\cdots$ (for example because the edge is gapped), and approximate $\partial_z \theta = \Delta\theta\delta(z)$, then we find a surface Hall conductivity (restoring units $2\pi = h$)

$$\sigma_{xy} = \frac{e^2}{h}\frac{\Delta\theta}{2\pi} = \frac{e^2}{h}\left(\frac{1}{2} + n\right) \tag{3.5}$$

where $\Delta\theta$, the change in θ between the two sides of the interface, is a half-integer multiple of 2π.

In terms of the edge theory, the periodicity in $\theta \simeq \theta + 2\pi$ for the fermion TI can be understood from the ability to deposit an (intrinsically 2+1 dimensional) integer quantum Hall system on the surface. This changes the integer n in the surface Hall response (3.5).

But the half-integer Hall response is not something that can ever happen in a microscopically well-defined 2+1d system. We've already seen that it definitely can't happen in a *gapped* 2+1d system without TO, but this is a much stronger statement, and here's why it's true. Such a system can be coupled to a background $\mathsf{U}(1)$ gauge field $\mathcal{A}$, and the resulting partition function must be gauge invariant. On the other hand, the Hall response in (3.5) comes from a CS term

$$S_{\text{eff}}[\mathcal{A}] \overset{?}{\ni} \frac{\nu}{4\pi}\int \mathcal{A}d\mathcal{A} \tag{3.6}$$

which (as we've seen) is definitely not gauge invariant when $\nu = n + \frac{1}{2}$.

Parallel with polarization. There is a nice parallel between this connection between the bulk ambiguity in θ and the possibility of depositing IQHE layers on the surface and the theory of electric polarization. The polarization P of a (say 1d) crystal is a periodic variable, because shifting every electron by a lattice spacing takes the crystal back to itself in bulk. In the presence of a surface, this shift removes charge from one surface and deposits it on the other surface. Indeed, a good understanding of polarization for free fermions is in terms of the Berry connection[42]. Very roughly, this is because polarization is the density of electric dipoles $e\langle\vec{r}\rangle \sim e\int_k \left\langle \mathbf{i}\vec{\nabla}_k \right\rangle = e\int_k \vec{\mathcal{A}}(k)$.

And indeed, there is a $D = 1+1$ analog of the story above: Consider an insulator in $D = 1+1$ with $\mathsf{U}(1)$ symmetry and a $\mathbb{Z}_2$ charge-conjugation symmetry, under which $j_\mu \to -j_\mu$. The effective action for the background gauge field can have a term of the form

$$S_\theta[A] = \int_{X_2} \frac{\theta}{2\pi} F. \tag{3.7}$$

You can see that $P = \frac{\theta}{2\pi}$, since the coupling of an electric dipole is $\vec{d}\cdot\vec{E}$, and P should be the dipole density.

[42] More precisely, the polarization can be obtained as an integral over time of the current that produces it: $\frac{d\vec{P}}{dt} = \vec{j}$. Now imagine varying some parameter λ in H to produce adiabatically such a polarization. Consider now free fermions in a lattice, so we can speak about filling bands. First order perturbation theory says the change in the state of the nth band is

$$|\delta\psi_{nk}\rangle = -\mathbf{i}\hbar\dot{\lambda}\sum_{m\neq n}\frac{\langle\psi_{mk}|\partial_\lambda\psi_{nk}\rangle}{E_{nk}-E_{mk}}|\psi_{mk}\rangle\,.$$

(In the first step we replaced δH with H because the extra terms give back $|\psi_{nk}\rangle$ which is orthogonal to $|\psi_{mk}\rangle$.) The contribution to the resulting current is

$$\vec{j}_n = \frac{d\vec{P}_n}{dt} = \left\langle\frac{e\vec{p}}{m}\right\rangle_n = \frac{\mathbf{i}\hbar e}{m}\dot{\lambda}\sum_{m\neq n}\oint_{\mathrm{BZ}}\mathrm{đ}^d k\frac{\langle\psi_{nk}|\,\vec{p}\,|\psi_{mk}\rangle\langle\psi_{mk}|\partial_\lambda\psi_{mk}\rangle}{E_{nk}-E_{mk}} + h.c.$$

Now use Bloch's theorem to write $\psi_{nk}(r) = e^{\mathbf{i}k\cdot r}u_{nk}(r)$, with $u_{nk}(r)$ periodic, which satisfy $H_k|u_{nk}\rangle = E_{nk}|u_{nk}\rangle$ with $H_k = \frac{(p+\hbar k)^2}{2m} + V$. Therefore $\vec{p} = \frac{m}{\hbar}\vec{\nabla}_k H_k - \vec{k}$. Differentiating the eigenvalue equation with respect to k and eliminating $\partial_k H$ leads to some juicy cancellations and gives

$$\frac{d\vec{P}}{d\lambda} = \mathbf{i}e\sum_{n,\mathrm{occupied}}\oint_{\mathrm{BZ}}\mathrm{đ}^d k\left\langle\vec{\nabla}_k u_{nk}|\partial_\lambda u_{nk}\right\rangle + h.c.$$

The polarization is therefore

$$\vec{P} = \int_0^1 d\lambda\frac{d\vec{P}}{d\lambda} = e\mathrm{Im}\left(-\mathbf{i}\sum_{n,\mathrm{occupied}}\oint_{\mathrm{BZ}}\mathrm{đ}^d k\vec{\mathcal{A}}_{nn}(\vec{k})\right)$$

where $\vec{\mathcal{A}}_{mn}(k) \equiv \mathbf{i}\langle u_{mk}|\vec{\nabla}_k|u_{nk}\rangle$ is the Berry connection.

Because of flux quantization, $Z(\theta) = Z(\theta + 2\pi)$ on a closed manifold. Microscopically, $\theta = \oint đk \sum_{n,\text{occupied}} \mathcal{A}_{nn}$ is just the Berry connection. It is periodic because under a rephasing of the wavefunction $u_{nk}(r) \to g(k)u_{nk}(r)$, $\mathcal{A} \to \mathcal{A} + g^{-1}\partial_k g$ changes by a gauge transformation. The choice $g(k) = e^{\mathbf{i}mak}$ for $m \in \mathbb{Z}$ (where a is the lattice spacing) shifts θ by $2\pi m$.

Under the charge conjugation symmetry, $F \to -F$. The allowed values of θ, then, are $\theta = 0, \pi$. In the presence of a domain wall connecting the two phases, the EM current has a contribution $j_\mu = \epsilon^{\mu\nu}\frac{\partial_\nu \theta}{2\pi}$, which says that there is a half-unit of charge localized at the wall. This happens in polyacetylene.

Note, by IBP, the similarity between the polarization term (3.7) and $S_\nu[\theta, A] = \int \frac{\nu}{2\pi} d\theta \wedge A$ in our theory of elasticity! In fact, the most current, many-body understanding of electric polarization in general dimension d is in terms of a term just like our d-dimensional term

$$S_\nu[\theta^I, A] = \frac{\nu}{(2\pi)^d d!} \int A \wedge d\theta^{I_1} \wedge \cdots \wedge d\theta^{I_d} \epsilon_{I_1 \cdots I_d}. \tag{3.8}$$

The role of the polarization vector is played by $\vec{P} = \frac{\vec{\theta}}{2\pi}$[43]. More generally, the point in life of the term (3.8) is to encode the response to flux-threading. It is sometimes said that (3.8) is like an unquantized anomaly, since P itself need not be an integer. But, as we saw earlier, the coefficient ν is an integer if the action needs to be invariant under large gauge transformations (*i.e.* if there's no groundstate degeneracy). And this integer is the number of particles per unit cell.

While I'm talking again about S_ν I have to mention something else that it encodes, called the *Thouless quantized charge pump*. It is closely related to the presence of the fractional surface charge. Consider the case of $D = 1+1$. Suppose we can adiabatically vary the coupling parameter θ n times through its period, $\theta \to \theta + 2\pi n$, during some time interval T. What amount of charge Q passes a given point in the bulk? The current is $j^\mu = \frac{\nu}{2\pi}\epsilon^{\mu\nu}\partial_\nu\theta$, so

$$\Delta Q = \int_0^T dt j^x(t) = \int_0^T dt \partial_t \theta \frac{\nu}{2\pi} = \frac{\nu}{2\pi} \oint d\theta = \nu n. \tag{3.9}$$

So, under the by-now-familiar assumptions that quantize ν, the charge transported in an adiabatic loop is quantized. (Note that when the parameter $\frac{\theta}{2\pi}$ is interpreted as

[43] Note that in the reference above, the term is described in terms of $\mathbb{Z}$ *gauge fields*. These are somewhat-mysterious objects whose properties are identical to those of $d\theta^I$ (they encode the size and shape and defects of the lattice, in that *e.g.* $\oint_{C_i} d\theta^i = L_i$, and $\frac{d^2\theta^i}{2\pi}$ is the density of dislocations with Burgers' vector $\hat{i}$). The intended advantage of the formulation in terms of the more-mysterious $\mathbb{Z}$ gauge fields is that they carry *less* information than a configuration of θ^I, which also carries geometric data. In practice however I think it's always necessary to pick an analog of the θ^I anyway.

the polarization, this conclusion becomes a bit tautological.) We've already seen an example of this in Laughlin's argument around (2.3) – consider the limit where the annulus is a thin tube, so it looks like a 1d wire. (This picture is the subject of the first two sections of this famous paper.) Generalizations to higher dimensions, closely related to (3.8) are discussed here.

Polarization from monopole quantum numbers. Plugging in $\theta^I \to \theta^I + 2\pi P^I/\nu$ to (3.8) gives

$$S_{\text{pol}} = \sum_{i=1}^{d} (-1)^{i+1} \int P^i x^1 \wedge \cdots x^{i-1} \wedge dA \wedge x^{i+1} \wedge \cdots \wedge x^d \tag{3.10}$$

where $x^i \equiv \frac{d\theta^i}{2\pi}$ are the $\mathbb{Z}^d$ gauge fields associated to lattice translations. To recommend the interpretation of the P^i in this term as the components of the polarization vector, consider a perfect cubic lattice of size $L_x \times L_y \times ...$ where $\theta^i(x) = 2\pi x^i$. Then in a uniform background electric field in the i direction, the term evaluates to $S_{\text{pol}} = \frac{V}{L_i} P_i \int dt dx_i E_i$, so that regarded as a 1d system the total polarization is VP_i/L_i. If we let P_i depend on space and time, then we find a current density

$$j^\mu = \frac{\delta S_{\text{pol}}}{\delta A_\mu} = (-\vec{\nabla} \cdot \vec{P}, \partial_t \vec{P})^\mu \ . \tag{3.11}$$

Let's ask how this action responds to the insertion of 2π flux.

1. In $D = 1+1$, insertion of 2π flux is an event in spacetime described by $F = 2\pi\delta^2(x_\perp)$, and this produces a phase $S_{\text{pol}} = 2\pi P$. Thus we can infer the polarization from the variation of the action upon flux insertion.

2. In $D = 2+1$, 2π flux is a particle, with a worldline on which $F = 2\pi\delta^2(x_\perp)$ has support. The term evaluates to

$$S_{\text{pol}} = \int_{\text{worldline}} dt 2\pi \left(P^x \partial_t \theta^y - P^y \partial_t \theta^x\right). \tag{3.12}$$

What does this do? Consider the canonical momentum for θ^i:

$$\Pi^i \equiv \frac{\delta S}{\delta \theta^i(x)} = \cdots + 2\pi \epsilon^{ij} P^j \tag{3.13}$$

where the $\cdots$ are contributions from other terms. The generator of lattice translations in the i direction is

$$\mathcal{T}^i \equiv e^{\mathbf{i}\Pi^i} = \mathcal{T}_0^i e^{2\pi \mathbf{i} \epsilon^{ij} P^j}.$$

This says that 2π flux carries lattice momentum

$$\vec{k} = 2\pi(-P^y, P^x).$$

We can therefore infer the polarization by measuring the momentum (phase acquired by unit translation) of the state in the presence of 2π flux.[44]

3. In $D = 3+1$, the insertion of 2π flux is still codimension two, so it happens along the worldsheet of a string. The endpoint of such a string is a magnetic monopole. The polarization term evaluates to

$$S_{\text{pol}} = \int_{\text{worldsheet}} dtdx 2\pi\epsilon^{ijk} P^i \partial_t \theta^j \partial_x \theta^k. \tag{3.14}$$

This changes the canonical momentum of θ^j by

$$\Pi^j = \cdots + 2\pi\epsilon^{ijk} P^i \partial_x \theta^k.$$

The consequence is that acting on the magnetic monopole (the ends of the 2π-flux string), the translation operators satisfy

$$\mathcal{T}^x \mathcal{T}^y = \mathcal{T}^y \mathcal{T}^x e^{2\pi \mathbf{i} P^z}.$$

This is the magnetic translation algebra for a particle with charge q in a magnetic field B, with $2\pi P^z = qB$. The simple way to understand this is that the magnetic monopole experiences the displacement field created by the polarization as a magnetic field, and this is the ordinary projective realization of translations for a particle in a (dual) magnetic field.

This phenomenon is closely related to the magneto-electric response in (3.3), as follows. As I mentioned, one way to think about the θ angle is

$$\theta \propto \frac{\partial P}{\partial B} = \frac{\partial M}{\partial E}.$$

Upon turning on a magnetic field, we should have

$$\Delta\vec{P} = \frac{\theta}{4\pi^2}\vec{B}. \tag{3.15}$$

[44] Some more telegraphic comments connecting this point to other ideas: The fact that flux can carry lattice momentum was realized long ago by Haldane and by Read and Sachdev, and plays an important role in the Neel-to-VBS transition on the square lattice (the skyrmion of the AFM is realized in the NCCP^1 description as 2π flux, and the fact that it carries lattice momentum is why its condensation produces a state that breaks translation symmetry, the VBS phase). More generally, this idea was used by Song-He-Vishwanath-Wang (the same authors in an earlier paper) to provide an understanding of the lattice-symmetry-breaking phases of a system containing an algebraic spin liquid. Those phases are described by the condensation of monopoles; the monopole quantum numbers therefore determine the lattice symmetries that are broken.

But as we'll see below, in the presence of nonzero θ, a magnetic monopole carries electric charge $q = \frac{\theta}{2\pi}$. This means that when we turn on B^z, the monopole behaves like an electric charge in a magnetic field, and so the magnetic translation algebra should obtain, with an extra contribution

$$2\pi\Delta P^z = qB^z = \frac{\theta}{2\pi}B^z,$$

which is (3.15).

In $D = 2+1$ we had a very general argument that fractional Hall response required fractionalization. There is an analogous theorem about the magneto-electric response in $\mathcal{T}$-invariant systems: in the absence of fractionalization, $\frac{\theta}{\pi}$ must be an integer for a gapped $\mathcal{T}$-invariant system.

Here's the argument: Put the system on T^3, and apply a tiny magnetic field $\vec{B} = B\hat{z}$, uniform in x, y, so that $\int dxdyeB_z = 2\pi$, the minimal flux. I say that this is tiny to emphasize that this time-reversal-breaking cannot close the gap or otherwise destroy our effective description. Now adiabatically thread 2π flux through the z-circle, so $\oint_{C_z} A = \frac{2\pi t}{e}, t \in [0,1]$, and $\vec{E} = \hat{z}\frac{2\pi}{eL_z}$. From our effective action (3.3), the ground-state-to-groundstate amplitude for this process is

$$Z = Ce^{\mathbf{i}S_\theta[\vec{E},\vec{B}]} = Ce^{\mathbf{i}\theta}$$

where C is some real constant (since the other terms in the action are time-reversal invariant).

But our system is assumed to be $\mathcal{T}$-invariant (before we applied $\vec{B}$). This means that the response to the time-reversed background field configuration must be the same. The time-reversed configuration still has $\oint A = \frac{2\pi}{e}$, but the magnetic flux is reversed, $\int B = -\frac{2\pi}{e}$. We conclude that

$$Z^{\mathcal{T}} = Ce^{-\mathbf{i}\theta} \stackrel{!}{=} Z = Ce^{\mathbf{i}\theta}$$

and therefore θ must be 0 or π. If there were topological order, the flux-threading process could take one degenerate groundstate to another, and we find a quantization condition on θ related to the number of torus groundstates, as in the FQHE. ■

Bosons. Similarly, we argued above that a non-fractionalized system of bosons in 2+1d must have a Hall response that is an even integer. And we just said that shifting θ by 2π corresponds to gluing to the surface a layer of IQHE with unit Hall

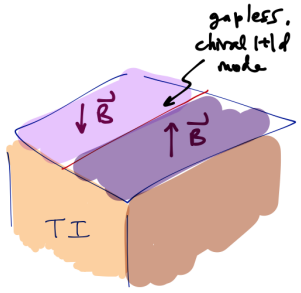
response. Therefore a (3+1)d boson TI has a θ parameter with period 4π, and the $\mathcal{T}$-invariant value $\theta = 2\pi$ is nontrivial. The fact that a surface layer of $\nu = 1$ IQHE would require fractionalization in a bosonic system mirrors (in a pretty remarkable way, I think) the mathematical fact that such a system can be put on a manifold that isn't spin, where $\int_X \frac{F\wedge F}{8\pi^2}$ can be a half-integer, so Z is only periodic in θ with period 4π, $Z(\theta) = Z(\theta + 4\pi)$. For a bosonic system, even $\theta = \pi$ is fractional.

[End of Lecture 12]

Gapped TI boundary. There is one subtlety in the above argument for the half-quantized surface Hall conductivity (3.5): it implicitly assumes that the EFT (3.3) is still valid even in the presence of the boundary, *i.e.* that there are no gapless boundary dofs.

To be able to gap out the edge states it is sufficient to break $\mathcal{T}$ symmetry at the surface, for example by applying a magnetic field. Here is a dramatic manifestation of the nontriviality of the bulk theory in this case.

There are two different ways of breaking $\mathcal{T}$, corresponding to the two directions in which the magnetic field can point. The 1+1d domain wall between these on the surface separates two regions whose Hall conductivity differs by $\frac{e^2}{h}$. This is just like the interface between IQHE and vacuum. Gauge invariance then requires that this wall must support a *chiral* edge mode.



With interactions or disorder other edge states are possible within the same bulk phase, including a gapped edge preserving $\mathcal{T}$. Such an edge must have topological order, and in fact *anomalous* topological order. These were found in a big flurry in June 2013. In fact, the edge is a time-reversal-invariant version of the Moore-Read Pfaffian state. I'll say more about such states later.

No matter what, the edge must somehow be *interesting*. I hope this sounds familiar to you from the discussion of LSMOH theorems. Something that may be confusing: when you thought about LSMOH theorems, you weren't thinking of a system living on the edge of a larger system. Why is it OK for those systems to live in a well-defined symmetric way without a bulk? The answer is that the symmetries involved there were spacetime symmetries (in particular, translations), which do not act in an on-site way: a translation, by definition, takes one site to another site. Such a symmetry can be anomalous, even in a well-defined lattice model. In particular, this means we can't use the usual minimal coupling prescription to couple to background gauge fields.

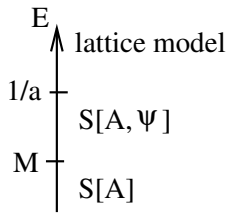
In the next bit, we'll briefly discuss a physical realization of the non-trivial case $\theta = \pi$ using non-interacting fermions. Later we'll get to the nontrivial case $\theta = 2\pi$ for

bosons.

3.1.4 Free fermion topological insulators in $D = 3 + 1$

[At this point, perhaps it would have been wise for me to follow these beautiful lectures by Witten. I particularly recommend them if you are not comfortable with the relativistic notation I am using.]

Free fermion TIs exist and are a realization of this physics with $\theta = \pi$. The simplest short-distance completion of this model is a single massive Dirac fermion:

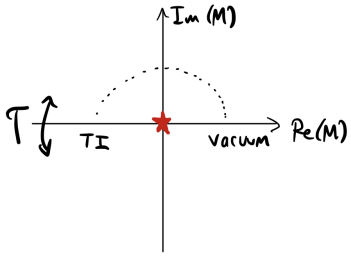


$$S[A,\psi] = \int \mathrm{d}^3x\mathrm{d}t\ \bar{\Psi}\left(\mathbf{i}\gamma^\mu D_\mu + m + \mathbf{i}\tilde{m}\gamma^5\right)\Psi. \tag{3.15}$$

Here the Lorentz symmetry is just to make things pretty. $\gamma^5 = \mathbf{i}\gamma^0\gamma^1\gamma^2\gamma^3$ is the chirality operator. The $\mathbf{i}$ in front of $\tilde{m}$ is required for reality of the action. It is convenient to denote $M \equiv m + \mathbf{i}\tilde{m}$. We can take time-reversal symmetry to act by[45]

$$\mathcal{T} : \mathbf{i} \mapsto -\mathbf{i}, \quad \psi(t,\vec{x}) \mapsto \gamma^1\gamma^3\psi(-t,\vec{x}), \qquad \Longrightarrow \qquad M \mapsto M^\star \tag{3.16}$$

so time reversal symmetry demands real M, but this allows for $M > 0$ or $M < 0$. As I will explain in the next subsection, these represent two inequivalent SPT phases. If we break time-reversal symmetry, of course we can adiabatically connect them just by rotating the phase of the mass.



A further short-distance completion of this massive Dirac fermion (as shown in the figure above) can come from discretizing the above action on a lattice. Consider the following Hamiltonian for 4-component fermions at each site x of a cubic lattice:

$$H = \sum_{x,\hat{\imath}=\hat{x},\hat{y},\hat{z}} c_x^\dagger \frac{\gamma^0 - \mathbf{i}\gamma^i}{2} e^{\mathbf{i}\mathcal{A}_{x,x+\hat{\imath}}} c_{x+\hat{\imath}} + m\sum_x c_x^\dagger\left(\sin\alpha\gamma^5 + (m + \cos\alpha)\mathbb{1}_x\right) \tag{3.17}$$

$$\overset{\alpha\in\{0,\pi\}}{=} \oint \mathrm{d}^3k c_k^\dagger d_\mu(k)\gamma^\mu(k)c_k \tag{3.18}$$

where

$$d_\mu(k) = (m + \sum_i \cos k_i + \cos\alpha, \sin k_x, \sin k_y, \sin k_z). \tag{3.19}$$

[45] Note that it looks like ψ transforms linearly, rather than antilinearly. But really it's just that $\mathbf{i} \to -\mathbf{i}$, $\mathrm{Re}\,\psi \to \mathrm{Re}\,\psi$ and $\mathrm{Im}\,\psi \to -\mathrm{Im}\,\psi$.

Here γ^μ are the same collection of gamma matrices as above, α is a parameter, and I set the lattice spacing to unity. Only for $\alpha \in \pi\mathbb{Z}$ is this time-reversal invariant. Shifting α acts like the chiral transformation $\Psi \to e^{\mathbf{i}\alpha\gamma^5}\Psi$. You can see that this looks like a discretization of the Dirac Lagrangian above[46].

More generally, filling an odd integer number of bands with a nontrivial Chern-Simons invariant of the Berry curvature produces $\theta = \pi$. That is,

$$\theta = -\frac{1}{4\pi}\int_{\mathrm{BZ}} d^3k\ \epsilon_{ijk}\mathrm{tr}\left(\mathcal{A}_i\partial_j\mathcal{A}_k - \mathbf{i}\frac{2}{3}\mathcal{A}_i\mathcal{A}_j\mathcal{A}_k\right) \tag{3.20}$$

where $\mathcal{A}_i(k)$ is the Berry connection. (The unfamiliar $\mathbf{i}$ in the second term is there because the three dimensions are Euclidean.) This paper gives a derivation of this formula by interpreting θ as the response of the electric polarization of the medium to a magnetic field: $\theta = \frac{2\pi h}{e^2}\frac{\partial P}{\partial B}$. (Recall that the polarization itself can be expressed in terms of the Berry connection. The above calculation requires going to second order in perturbation theory.)

Here is a field theory calculation that explains (3.20). Think about the variation of the action under an axial transformation, a shift of α above, under which $M \to e^{\mathbf{i}\alpha}M$ This comes from a triangle diagram with two insertions of $\mathcal{A}$ and one insertion of the axial current, and the fermion running in the loop. It gives

$$\delta_\alpha S = \alpha\frac{G_3}{4\pi}\int d^3x dt \epsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu \partial_\rho A_\sigma \tag{3.21}$$

where

$$\begin{aligned} G_3 &= -\frac{\pi}{6}\int \text{đ}^3k\,\text{đ}\omega\epsilon^{\mu\nu\rho\sigma}\mathrm{tr}G\partial_{q^\mu}G^{-1}G\partial_{q^\nu}G^{-1}G\partial_{q^\rho}G^{-1}G\partial_{q^\sigma}G^{-1}G\partial_\alpha G^{-1} \\ &= \frac{1}{4\pi}\int \text{đ}^3k\epsilon^{ijk}\mathrm{tr}\mathcal{F}_{\alpha i}\mathcal{F}_{jk}\ . \end{aligned} \tag{3.22}$$

G is the fermion propagator. To get to the last step, first observe that this object is some kind of winding number, and depends only on the set of filled bands and not on how they vary in energy; the model can be replaced with a model with flat bands. The full explanation is in appendix B of this paper. The last expression the variation of the Chern-Simons invariant of the Berry connection.

This is a beautiful formula, but it has the following two disadvantages: First, writing

[46] As I type this, it seems to me that from this naive discretization we actually get many (eight) copies of the massive Dirac fermion, from the band minima at each $k_i = 0, \pi$. I'm not actually going to use this discretized action for anything, and the important thing is just that some tight-binding model produces a single Dirac fermion.

this formula relies on translation symmetry, so that the BZ is well-defined. But the TI is a phase of matter protected by $\mathsf{U}(1) \ltimes \mathcal{T}$ – it is robust to the breaking of translation symmetry. Second, is a single-particle formula – it assumes non-interacting fermions. In fact, however, there is a way to interpret (3.20) that makes it robust to disorder and at the same time becomes a many-body formula: put the system on T^3 and reinterpret $k_{i=x,y,z}$ as a flux thread through the i direction, as in the work by Niu, Thouless, Wu.

Let me say a few more words about the important step taken by Niu, Thouless and Wu[47]. Put the system on T^d with $x^i \equiv x^i + L^i$ and impose boundary conditions

$$e^{\mathbf{i}\Pi^i_\alpha L_i}\Psi(x) \stackrel{!}{=} e^{\mathbf{i}q_\alpha \Theta^i}\Psi(x) \;\forall i = 1..d, \alpha = 1..N \tag{3.23}$$

where $\Psi(x)$ here is the full many-particle wavefunction, and $\Pi^i_\alpha \equiv -\mathbf{i}\frac{\partial}{\partial x^i_\alpha} + A_i(x^\alpha)$ is the canonical momentum of the αth particle. There's no sum over i or α in the exponent, so the monster acting on the wavefunction on the LHS is the (magnetic) translation operator that translates the αth particle all the way around the ith direction. q_α is the charge of the αth particle. On the one hand, this boundary condition can be removed by the unitary transformation

$$\Phi(x) \equiv e^{-\mathbf{i}\sum_i \frac{\Theta^i}{L^i}\sum_{\alpha=1}^N x^i_\alpha}\Psi(x) \; . \tag{3.24}$$

Acting on Φ, $-\mathbf{i}\partial_{x^i}$ gets replaced by $-\mathbf{i}\partial_{x^i} + \Theta^i$. This means that Θ^i appears wherever k^i would appear. As long as $H = \sum_i \frac{\Pi_i^2}{2m} + V(x)$, the perturbative calculation proceeds as before and gives the integrand of (3.20). On the other hand, (Niu et al argue that) in an insulating phase in the thermodynamic limit the response to background fields is independent of the boundary conditions, Θ (in fact such insensitivity to boundary conditions is Thouless' definition of an insulator). Therefore, we can average the result over Θ, and we obtain (3.20).

The Qi et al paper linked above describes various other physical consequences of nontrivial magneto-electric response, about which I'll say more below. One example is an extra contribution to the rotation of the polarization of light reflected off the surface; such effects can be understood entirely using the effective action (3.3) (with position-dependent ϵ, μ, θ).

The chiral anomaly in $D = 1 + 1$ from flux threading. [Polyakov, *Gauge Fields and Strings*, page 102; Kaplan 0912.2560 §2.1] The formula for the violation of the axial current is easy to understand in the special case of $D = 1 + 1$ in terms of

[47] Those authors had in mind the quantum Hall response in $D = 2 + 1$ (and its formula in terms of the Chern number of the Berry connection), but the same logic applies here, and to the polarization in $D = 1 + 1$.

flux-threading. Consider non-relativistic free (*i.e.* no 4-fermion interactions) fermions in 1+1 dimensions, *e.g.* with 1-particle dispersion $\omega_k = \frac{1}{2m}\vec{k}^2$. (Alternatively, we could have considered a tight-binding model of electrons hopping around on a chain, and here we are drawing just the bottom of the band.) The groundstate of N such fermions is described by filling the N lowest-energy single particle levels, up the Fermi momentum: $|k| \leq k_F$ are filled. Put them in a box of length L, so that $k_n = \frac{2\pi n}{L}$. (In Figure 1, the red circles are possible 1-particle states, and the green ones are the occupied ones.) The fields near these Fermi points in k-space satisfy the Dirac equation:

$$(\omega - v_F \delta k)\,\psi_L = 0, \quad (\omega + v_F \delta k)\,\psi_R = 0$$

where $\delta k \equiv k - k_F$. (The modes $\psi_{L/R}$ arise from the original fermion field as $\psi(x) \simeq \psi_L(x)e^{\mathbf{i}k_L x} + \psi_R(x)e^{\mathbf{i}k_R x}$.)

It would therefore seem to imply a conserved axial current – whose conserved charge is the number of left moving fermions minus the number of right moving fermions. But the fields ψ_L and ψ_R are not independent; with high-enough energy excitations, you reach the bottom of the band (near $k = 0$ here) and you can't tell the difference. This means that the numbers are *not* separately conserved.

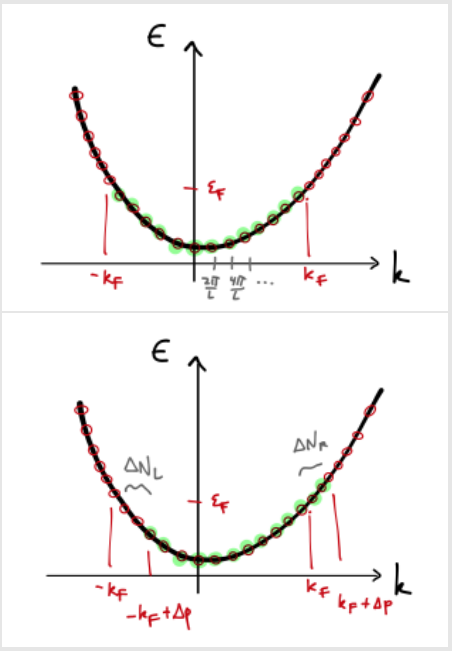


Figure 1: Green dots represent occupied 1-particle states. Top: In the groundstate. Bottom: After applying $E_x(t)$.

We can do better in this 1+1d example and show that the amount by which the axial current is violated is given by the anomaly formula I claimed above. Consider subjecting our poor 1+1d free fermions to an electric field $E_x(t)$ which is constant in space and slowly varies in time. That is, we adiabatically thread flux $\oint_0^L A_x dx$ through the circle. Suppose we gradually turn it on and then turn it off; here gradually means slowly enough that the process is adiabatic. Then each particle experiences a force $\partial_t p = eE_x$ and its net change in momentum is

$$\Delta p = e \int \mathrm{d}t E_x(t).$$

This means that the electric field puts the fermions in a state where the Fermi surface $k = k_F$ has shifted to the right by Δp, as in the figure. Notice that the total number of fermions is of course the same – charge is conserved.

Now consider the point of view of the low-energy theory at the Fermi points. This theory has the action

$$S[\psi] = \int \mathrm{d}x\mathrm{d}t\bar{\psi}\left(\mathbf{i}\gamma^\mu\partial_\mu\right)\psi \ ,$$

where γ^μ are 2×2 and the upper/lower component of ψ creates fermions near the left/right Fermi point. In the process above, we have added N_R right-moving particles and taken away N_L left-moving particles, that is *added* N_L left-moving holes (aka anti-particles). The axial charge of the state has changed by

$$\Delta Q_A = \Delta(N_L - N_R) = 2\frac{\Delta p}{2\pi/L} = \frac{L}{\pi}\Delta p = \frac{L}{\pi}e\int \mathrm{d}tE_x(t) = \frac{e}{\pi}\int \mathrm{d}t\mathrm{d}xE_x = \frac{e}{2\pi}\int \epsilon_{\mu\nu}F^{\mu\nu}$$

On the other hand, the LHS is $\Delta Q_A = \int \partial^\mu J^A_\mu$. We can infer a local version of this equation by letting E vary slowly in space as well, and we conclude that

$$\partial_\mu J^\mu_A = \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu}.$$

This means that under a chiral rotation by angle α, the action shifts by

$$\delta_\alpha S = \int dxdt\alpha\frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu}, \tag{3.25}$$

the (1+1)d analog of (3.28).

If we had only the right-moving fermion (impossible in a $D = 1+1$ lattice model), the anomaly would be half as big.

[End of Lecture 13]

3.2 Anomaly inflow and fermion zeromodes on topologically-protected defects

A word on nomenclature: By 'topologically-protected defects' in the title of this subsection I mean field configurations that cannot be continuously deformed to a constant configuration, such as domain walls or vortices. Such things used to be called 'topological defects.' They are topological in the sense that their presence is protected by the topology of the field space or the vacuum manifold. Another completely different class of things which are now called 'topological defects' in field theory are operators whose support may be continuous varied without changing their effects on observables; such operators describe symmetries. This property is not shared by the objects we study

here – their location really matters.

[Harvey] An elaboration of the example above is the context where the notion of anomaly inflow was first discovered. Consider coupling the bulk fermion to a complex scalar field, in addition to the gauge field. Essentially we are just making the mass M into another background field Φ.

So consider the following action in $D = 3 + 1$:

$$S[A, \Phi, \psi] = \int d^4x \left(\bar{\psi} \mathbf{i} \gamma^\mu D_\mu \psi - \bar{\psi} \left(\Phi_1 + \mathbf{i} \gamma^5 \Phi_2 \right) \psi - \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + \partial_\mu \Phi^\star \partial^\mu \Phi - U(|\Phi|^2) \right). \tag{3.26}$$

I've written the action as if A and Φ were dynamical variables, including their kinetic terms, but actually we will regard them as fixed background fields throughout this discussion, so I put them in a ghostly color. For some application of these ideas, they are indeed dynamical. We're going to regard Φ as a proxy for the bandstructure of a TI. We're going to be interested in configurations of Φ which are nonzero most everywhere, so we'll assume $U(x)$ has a minimum at $|\Phi| = M > 0$.

The action (3.26) is invariant under the transformation

$$\Phi \to e^{\mathbf{i}\alpha} \Phi, \quad \psi \to e^{-\mathbf{i}\gamma^5 \alpha/2} \psi, \tag{3.27}$$

which I'll call $\mathsf{U}(1)_A$, with associated Noether current j^μ_A. This *chiral symmetry* is anomalous in the sense that, although S is invariant, the measure of the fermion path integral is not. As a result, the variation of the effective action under a chiral transformation (3.27) is not zero, but (in flat space)

$$\delta_\alpha S = \int \alpha \frac{F \wedge F}{8\pi^2} = \int \alpha \partial_\mu j^\mu_A. \tag{3.28}$$

This is the chiral anomaly. For a derivation, see here, pages 14-15.

If Φ were dynamical, the $\mathsf{U}(1)_A$, in addition to being violated by the chiral anomaly, would be spontaneously broken, and there would be a goldstone mode θ, where $\Phi = Me^{\mathbf{i}\theta}$, called an *axion*. Its coupling to the fermions $\bar{\psi} M e^{\mathbf{i}\theta\gamma^5} \psi$ can be removed by a redefinition of ψ by a chiral rotation (3.27). The anomaly then implies that the axion reappears in a term of the form (3.28) with $\alpha = \theta$. We can then write a simpler effective action for just the massless fields (called 'axion electrodynamics'):

$$S_{\text{eff}}[A, \theta] = \int d^4x \left(\frac{\theta}{16\pi^2} F_{\mu\nu} \tilde{F}^{\mu\nu} - \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + \frac{M^2}{2} \partial_\mu \theta \partial^\mu \theta \right). \tag{3.29}$$

Here $\tilde{F}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$[48].

[48] Let me state my beliefs about the factors of two:

$$\int d^4x \frac{1}{2} \frac{\epsilon^{mnpq} F_{mn} F_{pq}}{8\pi} = \int d^4x \frac{F_{mn} \tilde{F}^{mn}}{8\pi} = \int \frac{F \wedge F}{4\pi} = \int d^4x \frac{1}{2\pi} \vec{E} \cdot \vec{B}$$

Now let's return to the description of a $\mathcal{T}$-invariant TI, where $\Phi = \pm M$ is real, and only the fermion is dynamical. We conclude that $M > 0$ and $M < 0$ produce effective actions whose value of θ differ by π, as promised. That is, the phase with $M < 0$ has $\theta = \pi$ and hence represents a non-trivial topological insulator phase, distinct from the vacuum. (It is not really so meaningful which is which. Such a phenomenon is called a *relative* invariant. Within a chunk of the phase, we can't tell which is which, but when putting them next to each other, we can tell the difference. In condensed matter physics we can always compare to the vacuum so there is an absolute answer.)

Now that we've derived the effective action (3.3) from a more microscopic theory, let's think harder about its consequences. The equations of motion for the EM field are $\partial^\mu F_{\mu\nu} = J_\nu$ with (repeating (3.4))

$$J_\nu = \frac{e^2}{4\pi^2}\partial^\mu\left(\theta\tilde{F}_{\mu\nu}\right) \ . \tag{3.30}$$

Consider a situation where the axion field θ varies in spacetime. To make it more visceral, let's write the equations in terms of E and B:

$$\rho \sim \theta\vec{\nabla}\cdot\vec{B} + \vec{B}\cdot\vec{\nabla}\theta \qquad \vec{j} \sim \dot{\theta}\vec{B} + \vec{\nabla}\theta\times\vec{E}. \tag{3.31}$$

The first term in the first equation is there even if θ is constant, and says that a magnetic monopole carries electric charge θ/π; this is called the Witten effect. The second term in the $\vec{j}$ equation is responsible for the surface Hall effect. The second term in ρ is also interesting: it says that breaking $\mathcal{T}$ at the surface with a magnetic field results in a deposition of charge. More on the interplay between this and the Witten effect below.

A very interesting way to make θ vary (while preserving $\mathcal{T}$) is to have an interface between a TI and vacuum. An interface between the TI and vacuum is a domain wall in Φ between $+M$ and $-M$. Let's consider what the theta-term does in the presence of such a wall:

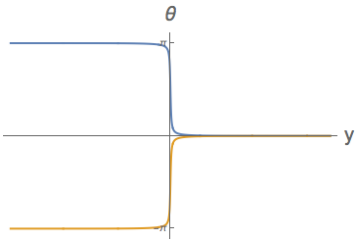
$$S_\theta = \frac{1}{8\pi^2}\int \theta F\wedge F \overset{\text{IBP}}{=} \int \frac{d\theta}{\pi}\wedge\frac{A\wedge F}{8\pi}. \tag{3.32}$$

Now think about the limit where the domain wall is very narrow, so θ is $\pm M$ almost everywhere. Then we can write $d\theta = \Delta\theta\delta(y)dy$, and we have

$$S_\theta = \frac{\Delta\theta}{\pi}\frac{1}{2}\int_{y=0}\frac{A\wedge F}{4\pi}. \tag{3.33}$$

are all integers times 2π on a spin manifold

Now what is $\Delta\theta$? Actually, the axion is ill-defined if Φ goes through zero. To regularize this, let's add a tiny imaginary part to Φ. We are explicitly breaking $\mathcal{T}$ by a small amount at the surface. The profile of θ looks like the figures at right for $\mathrm{Im}\,\Phi$ positive and negative, respectively. We conclude that $\Delta\theta = \pm\pi$. There is a 2π ambiguity in $\Delta\theta$.



Using this result in the expression (3.33) for the theta term, we see that

$$S_\theta = \pm\frac{1}{2}\int_{y=0}\frac{A\wedge F}{4\pi}. \tag{3.34}$$

the domain wall has a half-integer quantum Hall response! As we've discussed before, this is not gauge invariant. Such an effective action, for a $(2+1)$-dimensional system, requires either gapless modes or topological order. This is an LSMOH theorem for the edge of the TI – it can't be boring.

There are two immediate questions we can answer:
1) The full $(3+1)$-dimensional effective action must be gauge invariant! How does that work?
2) What is responsible for the half-integer Hall conductivity on the edge in the case of free fermions?

Let's think about the equations of motion for the fermions in the presence of the domain wall $\Phi(y)$:

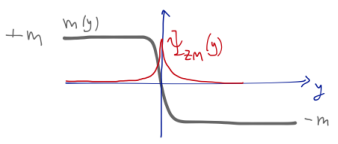
$$0 = \frac{\delta S}{\delta\bar{\psi}} = (\mathbf{i}\gamma^a\partial_a + \mathbf{i}\gamma^y\partial_y + \Phi(y))\,\psi. \tag{3.35}$$

Here $a = 0, 1, 2$ label the coordinates along the wall. I've turned off the background field A for a moment, it wouldn't change anything. Consider the ansatz $\psi = \eta_\pm(x^a)e^{\alpha(y)}$, where y is the coordinate normal to the wall. If we choose $\gamma^y\eta_\pm = \pm\mathbf{i}\eta_\pm$ (the γ^i are antihermitean matrices that square to -1), the Dirac equation separates:

$$\mp\,\partial_y\alpha/\alpha + \Phi(y) = 0, \quad \gamma^a\partial_a\eta_\pm = 0. \tag{3.36}$$

The solution of the first equation is $\alpha_\pm(y) = \pm\int_0^y dy'\Phi(y')$. Therefore if $\Phi \to +M$ as $y \to +\infty$, then e^{α_-} is exponentially localized on the wall.

We conclude that $\psi = \eta_-(x^a)e^{-\int_0^y dy'\Phi(y')}$ is a mode of the field exponentially localized on the wall, which satisfies the massless Dirac equation along the wall. Such a domain wall hosts a single 2+1d massless Dirac fermion. The spinor η_- satisfies $\gamma^y\eta_- = -\mathbf{i}\eta_-$, so has two independent components, as befits a spinor in $D = 2+1$.



There is no way to give a mass to a single Dirac cone in $D = 2+1$, while preserving $\mathsf{U}(1) \ltimes \mathcal{T}$ symmetry[49]. This is true even if we break Lorentz symmetry – the term $\mu\eta^\dagger\eta$ is a chemical potential the addition of which only makes the gaplessness more severe by creating a Fermi surface. Even breaking the $\mathcal{T}$ symmetry on the boundary (either explicitly or spontaneously) produces something interesting, as it must. That is, if we add a γ^5 mass (an imaginary part of M) localized at the wall, we give a mass to the Dirac fermion on the edge. But as shown in footnote (34), the Hall response of a single Dirac fermion of mass $m = \mathrm{Im}\, M$ is $\frac{m}{|m|}\frac{AdA}{8\pi}$ – exactly the kind of half-integer Hall response required to cancel the anomalous variation of the bulk action. Moreover, a single Dirac fermion (even a massive one) is not something that can arise from a local lattice model – this is the content of the Nielsen-Ninomiya fermion doubling theorem in $D = 2 + 1$. In fact, the very effective action we just wrote is a proof of this claim (at least in the presence of the $\mathsf{U}(1)$ symmetry): if there were a local lattice model, the effective action would be gauge invariant[50].

The TI is protected by $\mathsf{U}(1) \ltimes \mathcal{T}$, so there is another way to gap out the single Dirac cone, namely by adding the superconducting term

$$\Delta\eta_a\epsilon_{ab}\eta_b + h.c. \tag{3.37}$$

($a, b = \uparrow, \downarrow$ are the remaining spin indices) that breaks the $\mathsf{U}(1)$ symmetry. This is what happens if we stick a superconductor on top of the TI surface. This state, naturally, has some very interesting properties, too. One way to see them is to consider what happens in the presence of a superconducting vortex. The answer is: Majorana zeromodes, as we'll see below.

You can see that the sum of the contribution from S_θ and from the Hall response of the edge theory is gauge invariant. The ambiguity in S_θ is up to the stacking of an integer quantum Hall state on the domain wall. This is a well-defined non-anomalous $D = 2 + 1$ system that we can just paint on there.

This phenomenon of cancellation of anomalies between bulk and boundary is called anomaly inflow. Let's discuss some other examples of this phenomenon.

Anomaly inflow in $D = 2 + 1$. Let me explain the edge states of the IQHE in the same language. Consider a Dirac fermion in $D = 2 + 1$ whose mass we regard as a

[49] How does the $\mathcal{T}$ symmetry act on the zeromode? The matrix $\gamma^1\gamma^3$ appearing in the time-reversal transformation of ψ, (3.16), commutes with γ^2, so acts as a 2×2 matrix $\mathbf{i}\sigma^2$ within the eigenspace of γ^2.

[50] Another proof is to look at the Berry connection in the presence of single Dirac point. It has a singularity with nontrivial Berry flux coming out of it. But in a lattice model, the Brillouin zone is compact (a d-torus) and this flux has nowhere to go.

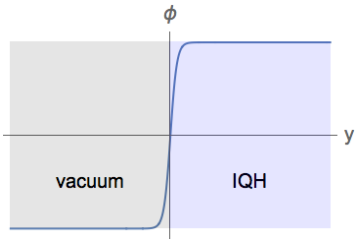
background (real) scalar field:

$$S = \int d^{2+1}x \left(\bar{\psi}\mathbf{i}\not{D}_A\psi - \phi\bar{\psi}\psi\right) + \int \frac{AdA}{8\pi} \tag{3.38}$$

In $D = 2+1$ the Dirac representation is irreducible and there is no chirality operator γ^5, so the mass (and hence ϕ) is real. (One subtlety here is visible in the last term: as we've seen a single Dirac fermion in $D = 2+1$ is not something that we can make with a lattice model in $D = 2+1$. So you could say that there is a second very heavy Dirac fermion, which is responsible for the half-integer Hall response term at the end of (3.38).)

Now consider a background configuration of $\phi(y)$ describing a static domain wall. Again think about the limit where the domain wall is very thin, so that everywhere with $y < 0$ has constant mass $-m$, and everywhere with $y > 0$ has constant mass $+m$. The Hall response is then described by

$$S_{\text{eff}}[A] \simeq \int (1 + \text{sign}(y)) \frac{1}{8\pi} AdA = \int_{y>0} \frac{AdA}{4\pi}. \tag{3.39}$$



We conclude that the left side is vacuum, and the right side is an IQH state.

What is the gauge variation of the action (3.39), under $A \to A + d\lambda$?

$$\delta_\lambda S_{\text{eff}} = \int_{y>0} \frac{1}{4\pi} d\lambda \wedge dA \overset{\text{stokes}}{=} -\int_{y=0} \lambda \frac{\epsilon_{ab}F^{ab}}{4\pi}. \tag{3.40}$$

This failure of gauge invariance is analogous to that of (3.34), but simpler because it happens even for small gauge transformations.

Again consider the Dirac equation in the presence of the domain wall:

$$0 = \frac{\delta S}{\delta\bar{\psi}} = \left(\mathbf{i}\gamma^a\partial_a + \mathbf{i}\gamma^y\partial_y - \phi(y)\right)\psi$$

where now $a = 0, 1$. Making a similar ansatz as above $\psi = \eta_\pm(x^a)e^{\alpha(y)}$, with $\gamma^{01}\eta_\pm = \pm\eta_\pm$, with $\gamma^{01} = \gamma^0\gamma^1$ (note that in 3d, the product of all the gamma matrices is a c-number, so constraining γ^{01} also constraints γ^y), we again separate the Dirac equation. But now we find a *chiral* fermion mode

$$\mathbf{i}\gamma^a\partial_a\eta = 0, \quad \gamma^{01}\eta = \eta. \tag{3.41}$$

Note that γ^{01} is the (1+1)d chirality operator, the analog of γ^5 along the string world-sheet. If we couple this system to the background gauge field A, a charged chiral

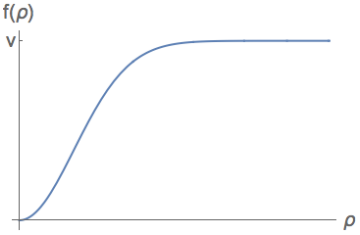
fermion mode produces a (1+1)d version of the chiral anomaly:

$$\delta_\lambda S_{2d}[A] = \int_{y=0} \lambda \partial^\mu j_\mu = \int_{y=0} \lambda \frac{\epsilon_{ab} F^{ab}}{4\pi} \tag{3.42}$$

(this is half the value you'll see quoted for the chiral anomaly of a (1+1)d Dirac fermion; that's because we have *only* the right-mover). This variation cancels the bulk contribution (3.40). In this case, there is no option of gapping out the edge mode.

Anomaly inflow in codimension two. Let's return to the $D = 3 + 1$ axion electrodynamics (3.29) related to the 3+1d TI. Now recall that $\pi_1(V = S^1) = \mathbb{Z}$ for this system, so there are topologically-protected vortex string defects. Since Φ, when real, is a proxy for the TI bandstructure, maybe you won't be too surprised that something like the physics below occurs on dislocation lines in a TI[51].

In cylindrical coordinates, where ρ is the distance from the string, such a vortex string has a profile like $\Phi = f(\rho)e^{i\varphi}$, where $f(\rho \to 0) \to 0, f(\rho \to \infty) \to M$. What we have to say is insensitive to the detailed form of f, which looks roughly like the figure at right.



Now apply an electric field along the string, $\vec{E} = E\hat{z}$. According to (3.31), this produces a current flowing radially in to the string,

$$\vec{J} = -\frac{e}{4\pi^2} E \frac{\hat{\rho}}{\rho}. \tag{3.43}$$

The divergence of this current satisfies $\vec{\nabla} \cdot \vec{J} = 0$ for $\rho \neq 0$, but

$$\int_R \vec{\nabla} \cdot \vec{J} \overset{\text{stokes}}{=} \oint_{\partial R} d\vec{n} \cdot \vec{J} = \int_0^{2\pi} d\varphi \rho J_\rho = -\frac{eE}{2\pi}.$$

But charge conservation then requires that the charge has to go somewhere – the string acquires charge, and the charge per unit length accumulating on the string (per unit time) is:

$$\frac{dQ}{dt} = -\int_R \vec{\nabla} \cdot \vec{J} = \frac{eE}{2\pi}. \tag{3.44}$$

We conclude

$$\partial^\mu J_\mu = \delta^2(x_\perp) \frac{1}{4\pi} \epsilon^{ab} F_{ab} \tag{3.45}$$

[51] Actually, what happens on a dislocation line is a helical mode: a pair of complex fermion modes with opposite chirality. In fact (see here, chapter 1, for a discussion), a dislocation line can be regarded as a dipolar boundstate of disclination lines, this strongly suggests that a *disclination line* in a TI should carry a chiral mode like we are seeing here. I'm not sure if this has been studied.

where $x_\perp$ are the two directions transverse to the vortex.

(Cliff-hanger!)

I claim that the resolution is that the string has a normalizable charged chiral fermion zeromode. That is, the Dirac field has a mode which instead of being a plane wave like in empty space is exponentially localized on the vortex, and satisfies the 2d chiral Dirac equation. But now think about the world from the point of view of the worldsheet of the string. It is a QFT in $D = 1 + 1$ with a chiral fermion, like the edge of an integer QH state, and this chiral fermion carries charge under the gauge field A. But then the chiral anomaly in $D = 1 + 1$ implies that the 2d chiral current is violated by a definite amount in a background EM field:

$$\partial^a j_a = \frac{1}{4\pi}\epsilon^{ab}F_{ab} \tag{3.46}$$

where $a = 0, 1$ labels coordinates along the string, and j is the current restricted to the string worldsheet at $x_\perp = 0$. This violation of charge conservation exactly accounts for the violation of charge conservation by (3.45). (Notice how tightly constrained everything was! Even I could get the factors of two right eventually.)

So this is an example where the anomaly of an 'edge theory' is cancelled by a bulk theory with two more dimensions.

Let me outline the discovery of the zeromode in this case. The Dirac equation in the vortex background, assuming azimuthal symmetry, is

$$\left(\mathbf{i}\gamma^a\partial_a + \mathbf{i}\gamma^2\left(\underbrace{\cos\varphi + \mathbf{i}\gamma^{23}\sin\varphi}_{=e^{\mathbf{i}\gamma^{23}\varphi}}\right)\partial_\rho\right)\psi_\mp = f(\rho)e^{\pm\mathbf{i}\varphi}\psi_\pm. \tag{3.47}$$

Here we've decomposed $\gamma^5 = \gamma^{01}\gamma^{23}$, with $\gamma^{01} \equiv -\gamma^0\gamma^1, \gamma^{23} = \mathbf{i}\gamma^2\gamma^3$ are respectively the chirality operators along and transverse to the string, and we've written the equation in a chiral basis where $\gamma^5\psi_\pm = \pm\psi_\pm$. The weird second term just comes from the derivative terms normal to the string. Note that we've ignored the possibility of a gauge field profile as a result of the vortex. Such a profile can lower the energy of the vortex. But the zeromode will exist either way – we can absorb the gauge field bits into a field redefinition of ψ, so let's just ignore it. In fact, the existence of this zeromode is guaranteed by an index theorem (this version is due to Callias), relating the number of zeromodes of the Dirac operator to integrals of curvatures.

This equation (3.47) is solved by

$$\psi_+ = -\mathbf{i}\gamma^2\psi_-, \quad \psi_- = \eta(x^a)e^{-\int_0^\rho f(\rho')d\rho'} \tag{3.48}$$

if

$$\mathbf{i}\gamma^a\partial_a\eta = 0, \quad \gamma^{01}\eta = -\eta. \tag{3.49}$$

This last condition (3.49) is exactly the 2d Dirac equation for a *chiral* fermion. Restoring the coupling to A, because ψ is charged under A, so is η coupled to $A(x_\perp = 0)$.

Using the form of the 2d chiral anomaly (3.42), we conclude that

$$\delta_\lambda \left(S_{\text{bulk}} + S_{\text{string}}\right) = \int_\Sigma \lambda \partial^\mu J_\mu + \int_\Sigma \lambda \frac{\epsilon_{\alpha\beta} F^{\alpha\beta}}{4\pi} = 0, \tag{3.50}$$

where Σ is the 2d worldsheet of the string.

WZ descent. Let's formalize a little bit the cancellation of the anomaly in terms of what is called the *Wess-Zumino descent procedure* that relates anomalies in various dimensions. The cancellation follows from the relations

$$I_{D+2} = dI^{(0)}_{D+1}, \quad \delta I_{D+1} = dI^{(1)}_D.$$

Here I_k is a k-form. δ indicates gauge variation, and d is exterior derivative. In the above, we have $D = 2$,

$$I_4 = \frac{F \wedge F}{8\pi^2}, \qquad I^{(0)}_3 = \frac{A \wedge F}{8\pi^2}, \qquad I^{(1)}_2 = \frac{\lambda}{2\pi} \frac{F}{4\pi}.$$

The axion electrodynamics action (3.29) contains the interesting term

$$S_\theta = \int_{M_4} \theta I_4 \overset{\text{IBP}}{=} - \int_{M_4} d\theta \wedge I^{(0)}_3$$

if $\partial M_4 = 0$. Then its gauge variation is

$$\delta \left(- \int_{M_4} d\theta \wedge I^{(0)}_3 \right) = - \int_{M_4} d\theta \wedge \delta I^{(0)}_3 = - \int_{M_4} d\theta \wedge dI^{(1)}_2 = \int_{M_4} d^2\theta \wedge I^{(1)}_2.$$

Now the axion winds around the string, so

$$\oint_{\text{around string}} d\theta = \Delta\theta = 2\pi = \int_{\text{disc around string}} d^2\theta$$

so we can approximate $d^2\theta = 2\pi\delta_2(\Sigma)$, where Σ is the string worldsheet. Therefore

$$\delta_\lambda S_\theta = 2\pi \int_\Sigma I^{(1)}_2$$

which is the formula for the 2d anomaly.

The reason to go through this trouble of formalizing is twofold: first, the same story relates local anomalies in any even dimensions D and $D + 2$. Second, it's now easy for

me to say what happens if we redo this in a fixed curved spacetime background: we just add some terms to the Is that depend on the curvature:

$$I_4 = \frac{F \wedge F}{8\pi^2} - \frac{\mathrm{tr} R \wedge R}{48 \cdot (2\pi)^2}.$$

$I_3^{(0)}$ and I_2 then get extra terms involving the gravitational CS term and R respectively. In fact in every even dimension, these objects are all pieces of the same thing:

$$I = \widehat{A}(TM) \wedge \mathrm{tr} e^{\frac{F}{2\pi}}.$$

This is a formal sum of differential forms of different (even) degree; the chiral anomaly in D dimensions, I_D, is the D-form piece of it. $\widehat{A}(TM)$ is called the A-roof genus and is a polynomial in the curvature form. The quantity I is the integrand of the RHS of the Atiyah-Singer index theorem counting zeromodes of the Dirac operator. As we'll see below, this is not an accident.

Zeromodes on vortices. Essentially the same calculation shows that a vortex of a $p + \mathbf{i}p$ superconductor carries a majorana zeromode. This is the basic mechanism by which the Moore-Read state (or any other realization) exhibits Ising topological order. But let's consider a vortex in the Fu-Kane superconductor on the surface of a TI, preserving $\mathbb{Z}_2^{\mathcal{T}}$ but breaking the $\mathsf{U}(1)$ symmetry. The action for the edge Dirac mode is

$$S[\eta, \Delta] = \int d^{2+1}x \left(\bar{\eta} \mathbf{i} \gamma^\mu D_\mu \eta + \Delta \eta_a \mathbf{i} \epsilon_{ab} \eta_b + \Delta^\star \eta_a^\star \mathbf{i} \epsilon_{ab} \eta_b^\star \right). \tag{3.51}$$

a, b are (2+1)d spinor indices. You can check that this action is real. Here we regard the superconducting pairing function Δ in the same way we regarded the background field Φ earlier. We allow it to depend on space, for example in a vortex configuration, $\Delta = e^{\mathbf{i}\varphi} f(\rho)$. Just to be clear: this is a continuum description of a vortex in a superconductor stuck to the surface of a TI.

So the Dirac equation is

$$0 = \frac{\delta S}{\delta \bar{\eta}} = \mathbf{i} \gamma^\mu D_\mu \eta - \Delta^\star \sigma^2 \eta^\star \tag{3.52}$$

Let's choose the gamma matrices to be $\gamma^1 = \sigma^2 = \mathbf{i}\epsilon, \gamma^2 = \sigma^3, \gamma^0 = \mathbf{i}\sigma^1$ (mostly plus convention); this choice can't matter. If we look for solutions independent of the azimuthal coordinate φ, we have

$$\left(\mathbf{i}\gamma^0 \partial_t + \mathbf{i}\gamma^1 \underbrace{\left(\cos\varphi + \mathbf{i}\gamma^{12} \sin\varphi \right)}_{= e^{\mathbf{i}\gamma^{12}\varphi}} \partial_\rho \right) \eta = \Delta^\star \gamma^1 \eta^\star \tag{3.53}$$

where $\gamma^{12} = \mathbf{i}\gamma^1\gamma^2$ is the chirality operator in the two space directions. Now let's look for a solution independent of t (a zeromode). Choosing $\gamma^{12}\eta_\pm = \pm\eta_\pm$, we have

$$e^{\pm\mathbf{i}\varphi}\mathbf{i}\partial_\rho\eta_\pm = f(\rho)e^{-\mathbf{i}\varphi}\eta_\pm^\star \ . \tag{3.54}$$

To match the φ-dependence, there can only be a solution for η_-.

A crucial point: (3.54) is not a linear equation in η over the complex numbers. It *determines* the phase of η. Make the ansatz $\eta_- = e^{\mathbf{i}\alpha}\eta_0$ for some *real* chiral spinor η_0. Then we have

$$e^{\mathbf{i}\left(\frac{\pi}{2}+2\alpha\right)}\partial_\rho\eta_0 = f(\rho)\eta_0.$$

To have a normalizable real solution

$$\eta_0(\rho) = e^{-\int_0^\rho d\rho' f(\rho')},$$

this equation must take the form $\partial_\rho\eta_0 = -f(\rho)\eta_0$, which means

$$\frac{\pi}{2} + 2\alpha \stackrel{!}{=} \pi$$

meaning that $\alpha = \frac{\pi}{4}$.

The particular answer for the phase is not important and would change if we changed representation of gamma matrices, but the point is that the phase is fixed: we have found a *majorana zeromode*. This is a fermion zeromode satisfying a reality condition. [End of Lecture 14]

The mode expansion of the quantum field $\eta(x^\mu)$ is then something like

$$\eta(x^\mu) = e^{\mathbf{i}\alpha}\eta_0(\rho)\gamma + \sum_k \left(a_k u_k^+(x) + b_k^\dagger u_k^-(x)\right) \tag{3.55}$$

where the second term is the usual thing with operators that create and annihilate particles and antiparticles in states that look like plane waves far from the vortex, $u_k^\pm(x) \sim e^{\pm\mathbf{i}k_\mu x^\mu}$ for large $|x|$. Canonical commutation relations then tell us that the mode operator γ anticommutes with the other modes and squares to 1.

What's the big deal? Suppose we put N well-separated vortices. Then each vortex $A = 1..N$ will have its own real zeromode (that they are exponentially localized guarantees that they are exponentially close to being zeromodes[52]), η_0^A, and associated majorana zeromode operator γ^A. They anticommute with all the other mode operators (and hence they commute with the Hamiltonian (they are zeromodes)) and amongst themselves satisfy the (Clifford) algebra

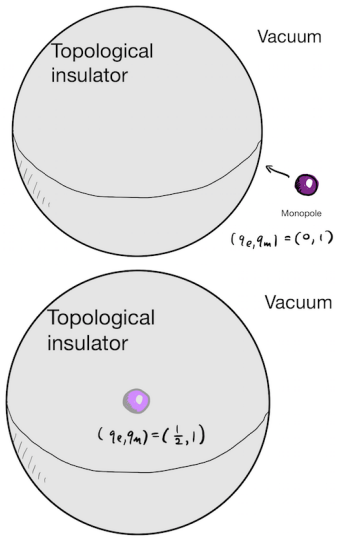
$$\{\gamma^A, \gamma^B\} = \delta^{AB}. \tag{3.56}$$

[52] Assuming an extra discrete symmetry, the Callias index theorem guarantees that they are exact zeromodes. See Appendix B of this paper.

What are the representations of this algebra? For N zeromodes there are $2^{\lfloor \frac{N}{2} \rfloor}$ states (where $\lfloor x \rfloor$ denotes the floor of x, the largest integer smaller than x). This grows like $\sqrt{2}^N$ at large N. It is as if there are $\sqrt{2}$ states per vortex. The Hilbert space therefore cannot be stored locally on the vortices. The operations of adiabatically moving the (identical) vortices around each other and exchanging them act on the Hilbert space by interesting unitary operators – it gives a representation of the (infinite) braid group on N elements that is not just a representation of the (finite) permutation group, S_N.

In the Fu-Kane superconductor (and in a $p + \mathbf{i}p$ superconductor), the objects carrying the majorana zeromodes are extrinsic defects rather than dynamical particle excitations of the system. This system does not have non-abelian topological order. The situation is different in the Moore-Read state, and in another surface termination of the TI that we'll construct momentarily.

More about the Witten effect. Consider a ball of TI. When we bring a magnetic monopole from the vacuum outside into the TI, it acquires electric charge $e/2$. (The charge is ambiguous modulo e, because an electron or hole can stick to it.) Charge conservation then requires that a compensating charge $-e/2$ is stuck to the boundary of the TI somehow. Each of the many possible boundary conditions accomplishes this in its own way, and it's worth examining these mechanisms.



The simplest is the Fu-Kane superconductor, $\Delta L = \eta\sigma^2\eta + h.c.$: a superconductor breaks the $\mathsf{U}(1)$ symmetry and can absorb arbitrary amounts of charge. End of story.

Next simplest is a time-reversal-breaking surface, $\Delta L = m\bar{\eta}\eta$. We saw already how moving a monopole through the surface is consistent with global charge conservation because of the two terms in the contribution to the electric charge density ρ from the theta term, (3.31) – the surface carries charge $e/2$ in the presence of 2π flux. Actually we knew this from the Laughlin flux-threading argument: threading localized 2π flux through a system with Hall conductivity $\sigma^{xy} = \nu\frac{e^2}{h}$ produces a quasparticle with charge $e\nu$, which here is $\pm\frac{1}{2}$ depending on the way in which we break time reversal. This charge then gets transported around the surface by the Hall effect due to the moving monopole.

What happens when there is a gapless edge? Think about the case where the boundary of the TI is a nice round S^2. In that case, the Dirac field η on the surface

is living on an S^2 with a single magnetic monopole inside. This is a classic problem. The answer is that η has zeromodes. These are eigenvectors of the Dirac operator on S^2 with one unit of flux $\oint_{S^2} F = 2\pi$, with eigenvalue zero. Their number is the subject of the Atiyah-Singer index theorem, which says that the number of right-handed zeromodes minus the number of left-handed zeromodes of the Dirac operator on an even-dimensional manifold X is

$$n_+ - n_- = \int_X \widehat{A}(TX) \wedge \mathrm{tr} e^{\frac{F}{2\pi}} \overset{X=S^2}{=} \int_{S^2} \frac{F}{2\pi} = 1. \tag{3.57}$$

This quantity is topological because nonzero eigenvalues of $\mathcal{D} \equiv \mathbf{i}\gamma^\mu D_\mu$ come in pairs related by the action of the chirality operator. That is, nonzero-energy eigenstates of $H \equiv \{\mathcal{D}, \mathcal{D}^\dagger\}$ are all doubly-degenerate, with one state of each chirality. This means small deformations of $\mathcal{D}$ can only change the number of zeromodes by the addition or subtraction of one of each chirality. The statement (3.57) is some pretty heavy machinery, but for physics purposes it essentially follows from the QFT derivation of the chiral anomaly.

What is the consequence of the single zeromode η_0 of η guaranteed by the index? Unlike in the Fu-Kane discussion, this Dirac equation is linear, so this is a *complex* zeromode, which carries $\mathsf{U}(1)$ charge 1. It leads to a complex fermion zeromode operator $c^\dagger$ of charge 1 ($\eta(x) = \eta_0 c + \text{nonzeromodes}$), satisfying $\{c, c^\dagger\} = 1$ (and anticommuting with all the other fermionic nonzeromode operators, and commuting with the Hamiltonian). This means that there are two degenerate states $|-\rangle$ annihilated by c and $|+\rangle = c^\dagger |-\rangle$ annihilated by $c^\dagger$. These two states have electric charges $q_\pm$ that differ by 1 (or e): $q_+ - q_- = 1$, and they are related by charge conjugation symmetry, so $q_- = -q_+$. We conclude that these two states have charge $\pm\frac{e}{2}$. After the monopole enters the TI, the charge is spread out all over the surface according to the wavefunction η_0 of this zeromode.

Finally, there is another possibility for a surface state of the TI: there can be a gapped, symmetric surface. Think about what happens when a monopole enters a TI with such a surface. In such a gapped state, charge cannot propagate freely. The charge $e/2$ that gets stuck on the surface when the monopole enters must be carried by a gapped quasiparticle with charge $-e/2$. We conclude that the gapped symmetric surface must be fractionalized, *i.e.* must have topological order, since it hosts a quasiparticle of fractional charge. Next we turn to a simple way to describe such a state. [53]

[53] There is another, unphysical, possibility for the surface that sometimes occurs in solvable lattice models, called *symmetry extension*. That is, the protecting symmetry group G could be a subgroup of a larger symmetry $H \supset G$ in the presence of an edge, with the property that the anomaly for G is not an anomaly for H. Certainly it's possible for the edge theory to have some enlarged emergent

Symmetric gapped surface of $D = 3+1$ TI. Here is a way to motivate the $D = 2+1$ topological order that carries the same anomaly as a single Dirac cone, or as the Fu-Kane superconductor, or as the state with $\frac{1}{2}$-integer Hall response. Such states were first constructed in these four papers, and later understood better using $D = 2+1$ dualities in these papers. There is also a close relation with Son's particle-hole-symmetric description of the half-filled Landau level. The construction I'll describe comes from here.

Begin with the single Dirac cone

$$S[\eta] = \int d^{2+1}x \bar{\eta} \mathbf{i} \gamma^\mu D_\mu \eta. \tag{3.58}$$

Let's regard this as the Higgs phase of some new $\mathsf{U}(1)$ gauge theory. Essentially, we are making a parton construction, with ansatz

$$\eta = \chi w^n \ . \tag{3.59}$$

The charge assignments are:

	χ	ϕ	w
A	1	2	0
a	n	$2n$	-1

(3.60)

Here χ is a Dirac fermion with a kinetic term just like (3.58), and ϕ and w are complex scalar fields, with appropriate kinetic terms. n is a number that we'll fix below.

First, to see why we should care about this theory, observe that if we condense w (but not ϕ), we Higgs the new $\mathsf{U}(1)$, and we can forget about it, and we are back to the theory of η, (3.58), since if w is condensed then (3.59) says $\eta = \chi \langle w^n \rangle = \chi$. This already proves that the two theories have the same anomalies, whatever other phases they realize.

The role of ϕ is more interesting. We should think of ϕ as arising by decoupling a four-fermion interaction of χ in the s-wave BCS channel. This just means that with the given charge assignments, we can (and should) add a term that looks a lot like the Fu-Kane term, an s-wave pairing term for χ:

$$L_\phi = \bar{\phi} \chi_a \mathbf{i} \epsilon^{ab} \chi_b + h.c. \tag{3.61}$$

symmetry. But this is unphysical because it requires H to be an *exact* symmetry of the theory with a boundary, since $G \subset H$. In contrast, emergent symmetries are always broken by irrelevant operators. I recommend this talk by Witten for a clear discussion of an example where this happens.

Such a solvable model can be used to construct a gapped symmetric boundary state, by gauging H/G. In that case, the would-be extra symmetry is just a redundancy of description – certainly H/G gauge theory is allowed to emerge.

where the dynamical field $\bar{\phi}$ plays the role of the superconducting pair field Δ. But now think about what happens in a phase where we condense ϕ (but not w): since ϕ is charged under both a and A, condensing ϕ higgses a combination of them down to a diagonal subgroup, which we can regard as the new global $\mathsf{U}(1)$ symmetry to which A is coupled. (The mass term is $(2A + 2na)^2$; varying a will just set $a = -2A/2n$, and there is no mass term for A.) The global $\mathsf{U}(1)$ that survives is the subgroup under which ϕ is neutral, with generator $\tilde{q} = q_A - q_a/n$. The effective table of charges in this phase is then:

	χ	ϕ	w
A	0	0	$-1/n$

(3.62)

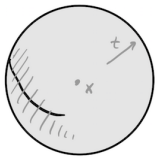
Moreover, since ϕ carries charge $2n > 1$ under $\mathsf{U}(1)_a$, a discrete ($\mathbb{Z}_{2n}$) subgroup of $\mathsf{U}(1)_a$ survives. This is a gapped state with topological order that preserves the protecting symmetry $G = \mathsf{U}(1) \ltimes \mathbb{Z}_2^{\mathcal{T}}$, so it must be interesting somehow.

Ignoring the fermions, this would just be $\mathbb{Z}_{2n}$ gauge theory, an abelian topological order. But now let's think about the vortices of ϕ. These are actual dynamical quasiparticles in the system. A vortex of ϕ has fractional flux $\int \frac{f}{2\pi} = \frac{1}{2n}$ – it is an m-particle or 'vison' of the $\mathbb{Z}_{2n}$ topological order. Ignoring the fermions, the quasiparticles would be labelled by e and m quantum numbers (k, v) and would satisfy the abelian braiding statistics $W_{kv}W_{k'v'} = W_{k'v'}W_{kv}e^{\frac{2\pi\mathbf{i}(kv'-k'v)}{2n}}$.

But the Dirac equation for χ is just the same as in our discussion of the Fu-Kane superconductor. We conclude that the vortex carries a majorana zeromode. There is a crucial physical difference, however: in the Fu-Kane case, the vortex was an external defect of a fixed background field Δ. In contrast, ϕ is a dynamical variable on the surface of the TI – its vortices are quasiparticles. (And they have finite energy, since the $\mathsf{U}(1)$ acting on the phase of ϕ is gauged.)

You could ask about charge $v > 1$ vortices, which seem to have v majorana zeromodes γ_i. But in the absence of extra discrete symmetries, nothing forbids quadratic terms $\mathbf{i}\gamma_i\gamma_j$ in the Hamiltonian which lift these zeromodes. So odd v vortices have a single majorana zeromode, and even v vortices have none.

So far we have not determined n. Let us think about the monopole operators of this $\mathsf{U}(1)$ gauge theory. A monopole operator is a kind of disorder operator, defined by a boundary condition in the path integral over a. The monopole operator $e^{\mathbf{i}\sigma(x)}$ is defined by cutting out a small ball around the spacetime point x, and demanding that on the boundary S_x^2 of this ball, we have $\oint_{S_x^2} f = 2\pi$, as if there were a magnetic monopole inside:



$$\left\langle e^{\mathbf{i}\sigma(x)}\cdots\right\rangle \equiv Z^{-1}\int_{\oint_{S^2_x} f=2\pi}[Da]e^{\mathbf{i}S}\cdots . \tag{3.63}$$

The quantum numbers of such an operator can be determined by regarding the radial direction away from the point x as time, so that the S^2 surrounding x is an equal-time slice. The quantum numbers of the operator are just those of the resulting groundstate on S^2 with flux[54].

The Dirac field χ has charge n, and so the same index theorem we used above predicts that it has n complex chiral zeromodes on S^2 with one unit of flux, whose associated operators we'll call $\chi_{i=1..n}$. Being a mode of a spin-half field, each χ_i has half-integer spin. They satisfy

$$\{\chi_i,\chi_j^\dagger\}=\delta_{ij},\quad \{\chi_i,\chi_j\}=0=\{\chi_i^\dagger,\chi_j^\dagger\}.$$

This algebra is represented by starting with a lowest-weight state $|\Downarrow\rangle$ annihilated by all the χ_is, and acting with $\chi_i^\dagger$s until we get to a highest-weight state $|\Uparrow\rangle$, annihilated by all the $\chi_i^\dagger$s. Let's determine the $\mathsf{U}(1)_A\times\mathsf{U}(1)_a$ charges and spins of these states: each $\chi_i^\dagger$ adds charge $\Delta q=(-1,-n)$, and a half unit of spin. The spectrum of charges must be symmetric about zero, so we must have

$$q_\Uparrow=-q_\Downarrow=q_\Downarrow+n\Delta q$$

from which we conclude that the charge of $|\Downarrow\rangle$ is $q_\Downarrow=\left(\frac{n}{2},-\frac{n^2}{2}\right)$. The states in the middle

$$\chi_{i_1}^\dagger\cdots\chi_{i_{n/2}}^\dagger|\Downarrow\rangle \tag{3.64}$$

have charge $\left(\frac{n}{2},-\frac{n^2}{2}\right)+\frac{n}{2}(-1,-n)=(0,0)$.

Now we can constrain n on physical grounds. First, if n is odd, the spectrum cannot be symmetric under charge conjugation. For example, if $n=1$, the states $|\Downarrow\rangle$ and $|\Uparrow\rangle=\chi_1^\dagger|\Downarrow\rangle$ have spins differing by a half-integer, and therefore cannot be related by a discrete symmetry. So we assume $n=2s$ is even. This means that there are gauge-invariant local monopole operators associated with the states (3.64). These states have integer or half-integer spin according to whether s is even or odd[55], *i.e.* the associated gauge-invariant monopole operator $e^{\mathbf{i}\sigma(x)}$ has spin $s/2$ mod $\mathbb{Z}$, and is neutral under $\mathsf{U}(1)_A$. But in a condensed matter system with all the degrees of freedom made

[54] More precisely, this statement uses the state-operator correspondence of conformal field theory. Conformal symmetry is valid at high energies in the system we are studying, and so we can use it to determine the microscopic quantum numbers of the monopole operators here.

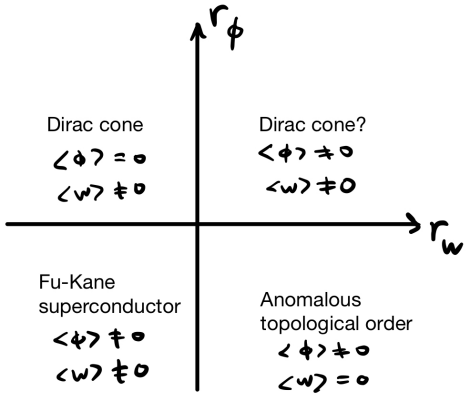
[55] Here I appeal to the fact that $|\Downarrow\rangle$ is a unique state annihilated by all the χ_i, which therefore must have spin 0.

from electrons, all half-integer-spin particles (fermions) carry odd charge under $\mathsf{U}(1)_A$. There is no way even the magic of emergence can change this property, it's a property of the Hilbert space. This 'spin-charge relation' is only satisfied by this gauge theory if s is even. We conclude that the minimal possible value of n is $n = 2s = 4$.

Being more explicit about the potential terms, we can write

$$L_r = u_w|w|^4 + r_w|w|^2 + u_\phi|\phi|^4 + r_\phi|\phi|^2 \qquad (3.65)$$

and we can think about varying the signs of r_w, r_ϕ[56] to condense or not each of w, ϕ. The phase diagram looks schematically like the figure at right. The walls are surface phase transitions, leaving the bulk SPT intact.



The upper right corner, where neither ϕ nor w is condensed is less easy to understand. Presumably the $\mathsf{U}(1)$ gauge theory confines by the Polyakov mechanism, because it has only a single Dirac fermion charged under it (I think the current consensus is that the critical number of flavors above which $\mathsf{U}(1)$ gauge theory in $D = 2+1$ is deconfined is $N_f = 2$). Does the fact that χ has charge 4 make a difference? Maybe. So we should look for gauge-invariant objects. The gauge invariant fermionic boundstate χw^n is then the right degree of freedom to consider. (Are there others? There are also monopole operators, the bosonic ones of which should be added to the action[57].) But this is just η, and it can't get a mass unless we break $\mathcal{T}$ or $\mathsf{U}(1)$. So maybe this is just the same phase as the upper left.

Exercises:

1. Apply the same technique to identify topologically-ordered symmetric gapped edge states for the case where the bulk is protected by $G = \mathsf{U}(1) \times \mathbb{Z}_2 \ltimes \mathbb{Z}_2$ where the extra $\mathbb{Z}_2$ is a chiral particle-hole symmetry as in this paper. The consequence of this extra symmetry is that an arbitrary integer number k of Dirac cones on

[56] In the figure here I wrote expressions like "$\langle\phi\rangle \neq 0$". This is a confusing shorthand. It is confusing because the expectation value of a non-gauge-invariant operator is always zero. Even when the minimum of the potential is far from the origin, it is always an orbit of the gauge group, and the value in a physical (gauge-invariant) state always averages to zero. (This is called Elitzur's theorem.) More precise statements are: $\langle|\phi|^2\rangle$ gets a big nonzero classical value, or we can choose a unitary gauge where we set ϕ to be real and positive.

[57] Note that these operators should also be added to the action in the phases with $\langle\phi\rangle \neq 0$. But there they don't do much besides explicitly break the flux conservation symmetry which is not a microscopic symmetry.

the surface is protected – the classification is $\mathbb{Z}$ rather than $\mathbb{Z}_2$.

2. Apply the same technique to understand the edge states of a $D = 3{+}1$ topological superconductor, like ^{3} He-B.

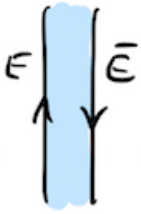
[End of Lecture 15]

3.3 Coupled-layer construction

The simplifying assumptions defining SPT states allow them to be built from lower-dimensional parts, actually in several different ways. (So it is a good idea to understand low-dimensional examples well.)

The idea of this subsection comes under many names. It is sometimes called the coupled-layer construction, or network construction, of SPTs.

The idea is to start with just the edge theory E and its inverse $\bar{E}$ in $D-1$ dimensions. This is what we would have if we had a very thin slab of the SPT. Call this a *layer*. Since $\bar{E}$ by definition has the opposite anomaly to E, a layer has no anomaly, and can be realized symmetrically by a lattice model in $D-1$ dimensions.



Now take a bunch of layers and make a one-dimensional chain of them. This makes a trivial state in D dimensions. But in this chain, each E is next to the $\bar{E}$ from the next layer, and there is a natural coupling between them. If we make this coupling larger than the one within the layer, we can drive a transition to a new D dimensional state that manifestly has E at one end and $\bar{E}$ at the other, which is the SPT.

Here are many examples.

- **AKLT/Haldane chain.** An anomalous system in $D = 0 + 1$ with $\mathsf{G} = \mathsf{SO}(3)$ symmetry is a single qubit. To see this, we can study the partition function for the system coupled to a background gauge field in the time direction for the $\mathsf{U}(1)$ generated by S^z: $\theta \equiv \oint_{S^1} A$. The Hamiltonian is just zero, so

$$Z(\theta) = \mathrm{tr} e^{\mathrm{i}\theta S^z} = \mathrm{tr} e^{\mathrm{i}\theta \frac{\sigma^z}{2}} \ . \tag{3.66}$$

  Another way to think about this function is as the expectation value of the operator $D(g(\theta))$ representing the symmetry transformation of a rotation by θ about the z direction: $Z(\theta) = \mathrm{tr} D(g(\theta))$. Either way, because of the crucial factor of $1/2$ in the exponent, you can see that

$$Z(\theta + 2\pi) = -Z(\theta), \tag{3.67}$$

despite the fact $g(\theta + 2\pi) = g(\theta)$ in $\mathsf{SO}(3)$. This is a fancy way of saying that spin-half is a projective representation of $\mathsf{SO}(3)$ that puts it in context of other anomalies. [For more discussion of this example and this perspective, see this paper by Seiberg and Cheng or Nati's related lectures at TASI 2023.]

So, an example of an edge for a $D = 1+1$ SPT for $G = \mathsf{SO}(3)$ is a single spin-$\frac{1}{2}$. A pair of spin-$\frac{1}{2}$s is a linear representation of $\mathsf{SO}(3)$, not a projective representation, since the two signs in the 2π rotation cancel out: $D_1(g(\theta+2\pi)) \otimes D_2(g(\theta+2\pi)) = D_1(g(\theta)) \otimes D_2(g(\theta))$. Call this a layer and label the two spin operators $\vec{S}$ and $\vec{\tilde{S}}$. It can be destroyed by the symmetric hamiltonian

$$H_0 = \sum_i \vec{S}_i \cdot \vec{\tilde{S}}_i \qquad (3.68)$$

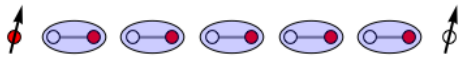
within the layer. What I mean by 'destroyed' is that the groundstate of H_0 is

$$|\text{trivial}\rangle = \otimes_i \left| \uparrow_i \tilde{\downarrow}_i - \downarrow_i \tilde{\uparrow}_i \right\rangle / \sqrt{2}. \qquad (3.69)$$

We can depict the groundstate of H_0 as at right, where each small circle represents a spin one-half, and the blobs represent singlets.

Now there are naturally two distinct ways of gapping out the bulk. One is with H_0, and the other is with the pairing

$$H_g = g \sum_{i=1}^{N-1} \vec{S}_i \cdot \vec{S}_{i+1} \qquad (3.70)$$



between the right spin of one layer and the left spin of the next. (This system is in the same phase as the spin-1 Heisenberg model, if we regard each pair of sites here as a single site; the difference is just the singlet in $\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3}$, which plays no role.)

When $g \gg 1$ we are in a gapped phase distinguished from the trivial one by the fact that the left spin-$\frac{1}{2}$ of first layer and the right spin-$\frac{1}{2}$ of the last are unpaired, as depicted in (3.70) for the extreme limit, where the wavefunction takes the form

$$|\text{AKLT}\rangle = \otimes_i \left| \tilde{\uparrow}_i \downarrow_{i+1} - \tilde{\downarrow}_i \uparrow_{i+1} \right\rangle / \sqrt{2}. \qquad (3.71)$$

The edges of the chain thus lie in *projective* representations of the symmetry group $\mathsf{SO}(3)$. This is the general property of the edge states of SPTs in $D = 1+1$. One is led to believe, then, that they are classified by the group cohomology group $H^2(G, \mathsf{U}(1))$ that classifies projective representations of G.

Notice that there must be a phase transition at some point in between as we increase g. It happens when $g = 1$, so that we get a homogeneous spin-half

Heisenberg chain. This model is described at low energies by the $\mathsf{SU}(2)_1$ WZW model CFT with $c_L = c_R = 1$. The model is also exactly solvable by Bethe ansatz.

A very similar story obtains with the same degrees of freedom if instead of $\mathsf{SO}(3)$, we choose the protecting symmetry to be $G = \mathbb{Z}_2^{\mathcal{T}}$, time-reversal symmetry. Recall Kramers' theorem: on a spin-half, $\mathcal{T}^2 = -1$. Such a pair of states is called a Kramers' doublet, whose degeneracy can't be split while preserving $\mathcal{T}$[58]. On a pair of spin-halfs, in contrast, $\mathcal{T}^2 = 1$. So if we make a 1d array of an even number Kramers' doublets, we again have two choices for how to pair them (for example by the Heisenberg interaction), and the interesting one leaves behind a single Kramers' doublet at each end. Notice that this is a $\mathbb{Z}_2$ classification, since a pair of such doublets can be destroyed by the symmetric hamiltonian (3.70) (or just $X\tilde{X} + Z\tilde{Z}$).

[58]Here is the proof that $|\psi\rangle$ and $\mathcal{T}|\psi\rangle$ are orthogonal when $\mathcal{T}^2 = -1$:

$$\langle\psi|\mathcal{T}|\psi\rangle \overset{\text{antiunitary}}{=} (\langle\mathcal{T}\psi|\psi\rangle)^\star \overset{\text{antiunitary}}{=} \mathcal{T}\langle\mathcal{T}\psi|\psi\rangle = \langle\mathcal{T}^2\psi|\mathcal{T}\psi\rangle \overset{\mathcal{T}^2=-1}{=} -\langle\psi|\mathcal{T}\psi\rangle = 0.$$

I learned this argument from Yi Li.

Projective representations. In quantum mechanics a state is often only defined up to a multiple of a phase $e^{\mathrm{i}\phi}$. So you might think that we can relax a bit the definition of a representation to allow

$$D(g)D(h) = \omega(g,h)D(gh) \tag{3.72}$$

where $\omega(g,h) \in \mathsf{U}(1)$ is a phase $|\omega(g,h)| = 1$. We must still demand associativity:

$$D(g)D(h)D(k) = (D(g)D(h))D(k) = \omega(g,h)D(gh)D(k) = \omega(g,h)\omega(gh,k)D(ghk)$$
$$= D(g)(D(h)D(k)) = D(g)\omega(h,k)D(hk) = \omega(g,hk)\omega(h,k)D(ghk)$$

from which we conclude

$$1 = \frac{\omega(g,h)\omega(gh,k)}{\omega(h,k)\omega(g,hk)}. \tag{3.73}$$

A function $\omega : G \times G \to \mathsf{U}(1)$ satisfying this associativity condition is called a *cocycle*. A set of $D(g)$s with nontrivial cocycle is called a *projective representation*.

When are two such things equivalent? By rephasing the generators $D(g) \mapsto \gamma(g)D(g)$, with $\gamma : G \to \mathsf{U}(1)$, (3.72) becomes $\gamma(g)D(g)\gamma(h)D(h) = \omega(g,h)\gamma(gh)D(gh)$, so

$$\omega(g,h) \mapsto \omega(g,h)\frac{\gamma(gh)}{\gamma(g)\gamma(h)}.$$

So if $\omega(g,h) = \frac{\gamma(g)\gamma(h)}{\gamma(gh)}$ for some function $\gamma : G \to \mathsf{U}(1)$, then this is actually equivalent to an ordinary (linear) representation of G.

To crystallize what we've just learned, define $\Omega^p \equiv \Omega^p(G, \mathsf{U}(1)) \equiv$ maps from $\underbrace{G \times G \times \cdots G}_{p \text{ times}} \to \mathsf{U}(1)$. An element of Ω^p is called a p-cochain. We can construct a (co)chain complex:

$$\Omega^1 \xrightarrow{\delta_1} \Omega^2 \xrightarrow{\delta_1} \Omega^3 \tag{3.74}$$

$$\gamma \mapsto \delta_1\gamma(g,h) = \frac{\gamma(h)\gamma(g)}{\gamma(gh)} \tag{3.75}$$

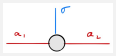
$$\omega \mapsto \delta_2\omega(g,h,k) = \frac{\omega(g,h)\omega(gh,k)}{\omega(h,k)\omega(g,hk)}.$$

You can check that $\mathrm{Im}\,\delta_1 \subset \ker\,\delta_2$. So $\omega \in \ker\,\delta_2$ is a cocycle, defining a projective representation. The equivalence relation is the map δ_1. Therefore, inequivalent projective reps correspond to elements of the quotient

$$\frac{\ker\,\delta_2 \subset \Omega^2}{\mathrm{Im}\,\delta_1 \subset \Omega^2} \equiv H^2(G, \mathsf{U}(1)).$$

This object is called the 2d group cohomology of G. It is a group under multiplication.

Classification of 1+1d SPTs. More generally, we can invoke a fact about gapped states in 1d: they can be written as a *matrix product state*:


$$= \sum_{a_{1,2\ldots}=1}^{\chi} \cdots M^{\sigma_1}_{a_1 a_2} M^{\sigma_2}_{a_2 a_3} \cdots |\cdots \sigma_1, \sigma_2 \cdots\rangle$$

$$\equiv M^{\sigma}_{a_1 a_2} \tag{3.76}$$

χ, the range of the auxiliary index, is called the *bond dimension*. This encodes the groundstate (a vector with d^L components,) in terms of χ^2 numbers (times L if the Ms for different sites are different). In such a state, each site is manifestly entangled with the rest of the system only through its neighbors. The statement that *any* gapped groundstate of a 1d local Hamiltonian can be written this way is a result of Hastings.

We want this state to be invariant under the action of $U = \prod_x u$ (so that it is a paramagnet).

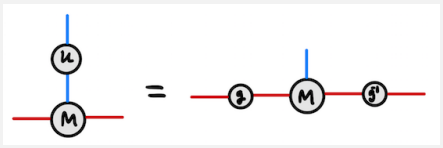
$$U \sum_{a_{1,2\ldots}=1}^{\chi} \cdots M^{\sigma_1}_{a_1 a_2} M^{\sigma_2}_{a_2 a_3} \cdots |\cdots \sigma_1, \sigma_2 \cdots\rangle = \sum_{a_{1,2\ldots}=1}^{\chi} \cdots M^{\sigma_1}_{a_1 a_2} u_{\sigma_1 \sigma_1'} M^{\sigma_2}_{a_2 a_3} u_{\sigma_2 \sigma_2'} \cdots |\cdots \sigma_{1'}, \sigma_{2'} \cdots\rangle$$

Clearly it would be invariant if at each site

$$M^{\sigma_1}_{a_1 a_2} u_{\sigma_1 \sigma_1'} \stackrel{?}{=} M^{\sigma_1'}_{a_1 a_2}.$$

But this is more than we need. Suppose instead that at each site we can factorize the effects of u as

$$M^{\sigma_1}_{a_1 a_2} u_{\sigma_1 \sigma_1'} = g_{a_1 a_1'} M^{\sigma_1'}_{a_1' a_2'} \left(g^{-1}\right)_{a_2' a_2} .$$



The gs may form a projective representation of G, since the phases will cancel in the previous expression – it's like the action of u is *fractionalized*. Then for a closed chain, the effects of the transformation would cancel between each pair of Ms, by  But if there were a boundary, there will be dangling gs. This leads to a classification of 1d (bosonic) SPTs in terms of $H^2(G, \mathsf{U}(1))$. Note that the group operation on the cohomology maps to the stacking operation.

For more on this I recommend these notes. You might guess that SPTs in d dimensions are classified by $H^{d+1}(G, \mathsf{U}(1))$, and this is almost but not quite true.

- **Kitaev chain.** Consider a single complex fermion mode c in $0+1$ dimensions, with $\{c, c^\dagger\} = 1$. Regard this as two majorana modes $c = \frac{\gamma + \mathbf{i}\tilde{\gamma}}{\sqrt{2}}$, with $\{\gamma, \tilde{\gamma}\} = 0, \gamma^2 = 1, \tilde{\gamma}^2 = 1$, and think of γ and $\tilde{\gamma}$ as generating the edge states E and $\bar{E}$ respectively.

  To specify what I'm talking about, I should say what is the protecting symmetry I have in mind. Let's consider two cases: $G = \mathbb{Z}_2^{\mathcal{T}}$ and no symmetry. $\mathbb{Z}_2^{\mathcal{T}}$ acts by $c \to c, \mathbf{i} \to -\mathbf{i}$, so $\gamma \to \gamma, \tilde{\gamma} \to -\tilde{\gamma}$. In either case, we can add

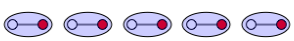
$$H_0 = \sum_i c_i^\dagger c_i = \mathbf{i}\gamma_i\tilde{\gamma}_i \qquad (3.77)$$

  in each layer, which has a unique groundstate ($c_i |0\rangle = 0$). We can depict the groundstate of H_0 as at right, where now each small circle represents a majorana mode, and the blobs represent complex fermion modes whose number operator appears in H_0.

  The chain of layers is then just a 1d chain of spinless fermions. With either choice of symmetry, we can add both fermion hopping terms, as well as p-wave pairing terms $c_i c_{i+1}$. Let's consider the coupling

$$H_g = g \sum_{i=1}^{N-1} \mathbf{i}\tilde{\gamma}_i\gamma_{i+1}. \qquad (3.78)$$

  For $g < 1$, the gap stays open and we remain in the trivial phase. When $g = 1$, the gap closes, and we have a translation-invariant majorana chain with half the unit cell. A massless $D = 1+1$ (non-chiral) majorana fermion propagates along the chain. (For a derivation of this statement see p. 54 here.) This is locally the same as the critical Ising model, a CFT with $c = \frac{1}{2}$. For $g > 1$, locally it looks the same as the trivial phase, in that each majorana in the bulk is paired with a neighbor. But now it is the neighbor from the next layer. This leaves out one majorana from the first layer and one from the last as unkillable edge modes. The extreme limit $g \gg 1$ is depicted in (3.78). This is the Kitaev chain (reviewed here). Since the two unpaired majorana modes are separated by a distance of order system size, they produce a protected qubit

$$\text{span}\{|0\rangle, c^\dagger |0\rangle, c \equiv \gamma_1 + \mathbf{i}\tilde{\gamma}_N\} \qquad (3.79)$$

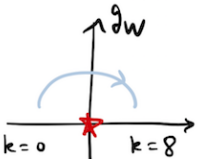
  with degeneracy exponentially small in the system size.

  In $H_0 + H_g$, we've chosen some special couplings, but the gapped phase we've discovered occupies an open set in the space of all couplings of a p-wave superconductor of spinless electrons in 1d.

The difference between the two symmetry choices arises when we consider stacking multiple copies of the Kitaev chain. If we have no symmetry, then when we stack a pair of these chains, nothing prevents us from pairing up the two end majoranas at each end. There is a $\mathbb{Z}_2$ classification in this case.

But with the $\mathbb{Z}_2^{\mathcal{T}}$ symmetry acting as above, the coupling $\mathbf{i}\gamma_1^{(1)}\gamma_1^{(2)}$ (where the index in parentheses labels which chain the mode comes from) or $\mathbf{i}\tilde{\gamma}_N^{(1)}\tilde{\gamma}_N^{(2)}$ is forbidden. This suggests that an arbitrary number of chains produces an arbitrary number of majorana modes which can't be destroyed by adding mass terms, leading to a $\mathbb{Z}$ classification.

This is true for free fermions. However, Kitaev and Fidkowski showed that allowing interactions reduces this to a $\mathbb{Z}_8$ classification. They explicitly identify a 4-fermion term W that can lift the degeneracy of 8 majorana modes in a $\mathbb{Z}_2^{\mathcal{T}}$-invariant way, *i.e.* it has a unique groundstate.
Adding this term in the bulk provides a route to go around the phase transition of the free fermion theory that would separate k chains and $k + 8n$ chains. Their construction involves some lovely group theory, featuring the triality symmetry of $\mathsf{SO}(8)$.



- **Integer quantum Hall states.** Now an example where layers of $D = 1 + 1$ systems produce a gapped state in $D = 2 + 1$. To make the fermion IQHE, a layer is just a single non-chiral boson mode ϕ, where the right-mover is E and the left-mover is $\bar{E}$. These can be gapped out by the backscattering interaction $\sum_i t_0 \cos\phi^i = \sum_i t_0 \cos\left(\phi_L^i + \phi_R^i\right)$. But if we add

$$\sum_i t_e \cos\left(\phi_R^i + \phi_L^{i+1}\right)$$

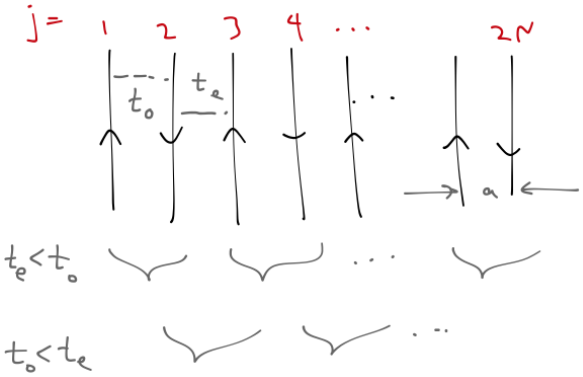
we can gap out all the modes in the bulk in a different way, leaving behind a leftmover at one end and a rightmover at the other.

Alternatively, here is a model of a transition in a system of (spin-polarized) electrons across which the Hall conductance changes by $\frac{e^2}{h}$, just in terms of free fermions.

$$S = \int \mathrm{d}x\mathrm{d}\tau \sum_{j=1}^{2N} \mathcal{L}_j$$

$$\mathcal{L}_j = \underbrace{c_j^\dagger \left(\partial_\tau - \mathbf{i} s_j \partial_x\right) c_j}_{\text{IQHE edge}} - \underbrace{t_j \left(c_{j+1}^\dagger c_j + c_j^\dagger c_{j+1}\right)}_{\text{interlayer hopping}}$$

$$t_j = \begin{cases} t_e, & j \text{ even} \\ t_o, & j \text{ odd} \end{cases} \quad , \qquad s_j = -(-1)^j \ .$$

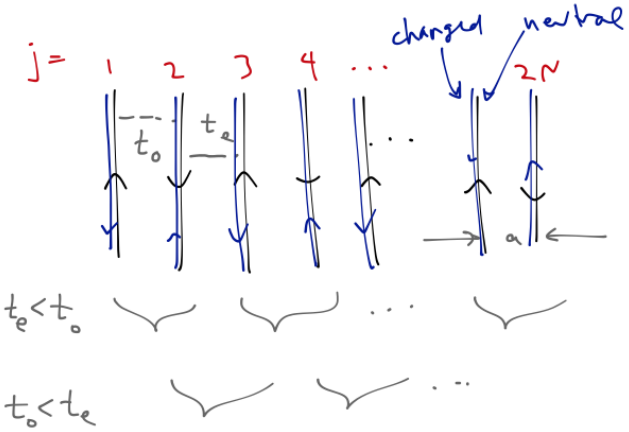


For each j we have one chiral mode, that is, a $D = 1+1$ complex chiral fermion. The t_j term is some back-scattering by which these layers hybridize and eat each other.

For $t_e < t_0$, all layers are paired and there's nothing left; $\sigma_{xy} = 0$.

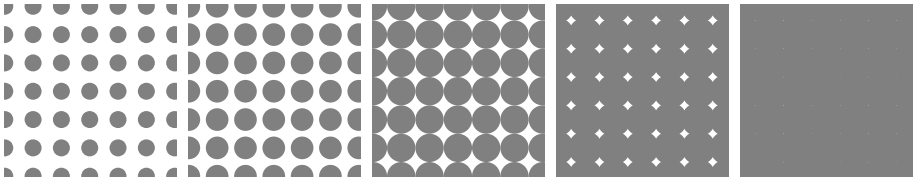
For $t_e > t_0$, there are leftover chiral modes at the edges, $\sigma_{xy} = \frac{e^2}{h}$. The transition occurs at $t_e = t_o$, where we restore some extra translation symmetry. The critical point is a discretization of a massless 2+1d Dirac fermion.

A similar example for the case of the boson IQHE is shown in the picture at right.



In general, the way the edge excitations emerge in each of these examples is just like in the classic picture of edge charge from polarization of an insulator; the role of the polarization angle is played by $\arctan\left(\frac{t_e}{t_o}\right)$, where $t_{e/o}$ are the couplings between the layers on the even and odd links respectively.

Building things from flat layers is not so important. We could just as well study a collection of decoupled droplets of IQH liquid (which has $\nu = 0$), and then gradually increase the couplings between the droplets:



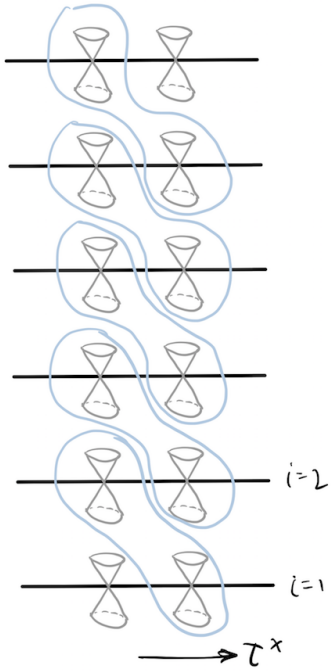
- $D = 3+1$ **TI, with** $G = \mathbf{U}(1) \ltimes \mathbb{Z}_2^{\mathcal{T}}$**.** [From here]

Here each layer $E\bar{E}$ is *two* Dirac cones in $D = 2 + 1$. For example, we could take the layer hamiltonian to be

$$H_0 = \sum_i \oint \mathrm{đ}^2 k c_i^\dagger(k) h_0(k) c_i(k), \quad h_0(k) = \tau_x \left(k_x \sigma_y - k_y \sigma_x\right)$$

where τ acts on a 'valley degeneracy' labelling the two cones, which both lie at $(k_x, k_y) = (0, 0)$ in this approximate description. Note that my bandstructure is not periodic in k, but what I've written is just an approximation near the Dirac points; the important thing is that it *could* be made periodic, consistently with the doubling theorem.
We could add a $\mathcal{T}$-invariant mass that pairs them up, and get a trivial insulator, but let's not. Instead couple the layers by

$$H_g = \sum_{i=1}^{N-1} \left(\oint \mathrm{đ}^2 k c_i^\dagger(k) h_+ c_{i+1}(k) + h.c. \right), \quad h_+ \equiv \tau_z + \mathbf{i}\tau_y \ .$$



The $(3+1)$d bandstructure described by $H_0 + H_g$ is a discretization of a massive $D = 3 + 1$d Dirac fermion. The resulting single-particle edge hamiltonians are $h_{\text{edge}} = \pm\vec{k} \times \vec{\sigma}$, and can no longer be paired up with each other because they are separated by a distance of order system size. The transition between the two phases hosts a massless 3+1d Dirac fermion.

- $G = \mathbb{Z}_2^{\mathcal{T}}$ **boson SPT in** $D = 3 + 1$**.** [From here] The edge state we wish to construct is the all-fermion toric code. This is a $\mathbb{Z}_2$ topological order where all the nontrivial anyon types, e, m, ϵ, are fermions. One way to get TO with this spectrum of anyons is $\mathsf{U}(1)^4$ CS theory with K matrix equal to the Cartan matrix of $\mathsf{SO}(8)$. But that theory is chiral (it has four chiral boson edge modes, $c_- = 4$) and clearly breaks $\mathcal{T}$ symmetry.

  One way to argue that the $\mathcal{T}$-invariant all-fermion toric code is anomalous, *i.e.* cannot be realized in a local bosonic system in 2+1d preserving $\mathcal{T}$, is the following. In a local bosonic system with TO, there is a relation called the Gauss-Milgrom formula:

  $$\frac{\sum_a d_a e^{2\pi \mathbf{i} s_a}}{\sqrt{\sum_a d_a^2}} = e^{2\pi \mathbf{i} c_-/8} \tag{3.80}$$

  between the spectrum of anyons (their quantum dimensions d_a and their

topological spins s_a) and the chiral central charge $c_- = c_L - c_R$ mod eight. [See Kitaev's honeycomb paper appendix E for some discussion of this formula.]

In a model of abelian anyons, all $d_a = 1$ and the total quantum dimension, $\mathcal{D} = \sqrt{\sum_a d_a^2}$, is simply the square root of the number of anyon types (including the identity). The fact that the central charge is only determined mod 8 is not an accident. The E_8 state of bosons has no anyonic excitations but has chiral central charge $c_- = 8$, hence we may add layers of the E_8 to any anyon model without changing the anyon content but shifting the chiral central charge by 8.

For the familiar $\mathbb{Z}_2$ gauge theory in which charges and vortices are bosons, we have $a \in \{1, e, m, em\}$, $d_a = 1$, $s_1 = s_e = s_m = 0$, and $s_{em} = 1/2$. Hence (3.80) gives

$$e^{2\pi i c_-/8} = \frac{3 + (-1)}{2} = 1 \tag{3.81}$$

hence $c_- = 0 \bmod 8$. In other words, the minimal $\mathbb{Z}_2$ gauge theory has no chiral edge states. However, if we consider the all-fermion gauge theory, then we find

$$e^{2\pi i c_-/8} = \frac{1 + 3(-1)}{2} = -1 \tag{3.82}$$

hence $c_- = 4 \bmod 8$. Thus the all-fermion gauge theory must have chiral edge states and hence must indeed break $\mathcal{T}$. The reason why this state can be realized in a $\mathcal{T}$-invariant manner at the surface of a $\mathcal{T}$-invariant $3+1$ bulk state is that in this case it is impossible to create an edge of the boundary at which the chiral edge states can be exposed!

[End of Lecture 16]

But two copies of this TO is related to two copies of the ordinary toric code, by a relabelling. Each one has six fermionic anyons and 10 bosonic anyons (including the trivial anyon), with the same braiding data. That is certainly realizable in a $\mathcal{T}$-invariant way.

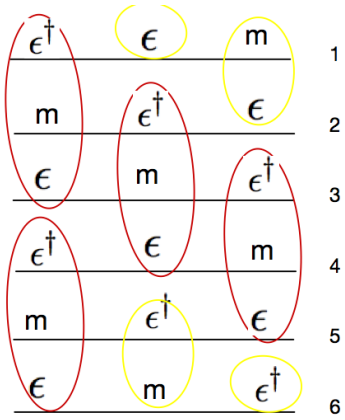
So a layer in this construction will be two copies of the ordinary toric code. We'll stack some number $N/2$ of such layers. For convenience let me label the individual toric codes $i = 1..N$. Now add a bulk hamiltonian that couples the layers and has the consequence that the objects

$$B_i \equiv \epsilon_i m_{i+1} \epsilon_{i+2}$$

condense. B_i is a boson (since it's made of two fermions and a boson with trivial mutual statistics), so it makes sense to condense it. Moreover, different B_i have

trivial mutual statistics, so we can condense them at the same time. (We could just add $\Delta H = -\sum_{i=1}^{N-2} B_i$.)

In the diagram at right, for $N = 6$, the objects circled in red are Bs. You can ignore the daggers on $\epsilon = \epsilon^\dagger$. Condensing B destroys all the topological order in the bulk, since each bulk anyon has nontrivial braiding with something in the condensate. The nontrivial anyons that survive (circled in yellow) are ϵ_1 and $m_1\epsilon_2$ at one end (and their boundstate), and $\epsilon_{N-1}m_N$ and ϵ_N (and their boundstate) at the other. But these particles are all fermions!



This is an example of an SPT that lies outside those realized by group cohomology. Note that if there are gauge-invariant fermionic particles available the whole story above goes away. For example, in that case we can cancel even $c_- = 1/2$ by making a $p + \mathbf{i}p$ state of the fermions without changing the anyon content, so the Gauss-Milgrom formula is out the window.

One consequence of this construction is the following conclusion: we can see that the cost of preserving time-reversal symmetry in this system is the conservation of fermion parity within each edge. Consider a fermion at the top layer, say ϵ_1. Since the boson $b_1 = \epsilon_1 m_2 \epsilon_3$ is condensed, we can create this fermion by acting with $\epsilon_1 b_1$ instead. The same is true of $b_3 = \epsilon_3 m_4 \epsilon_5$. Thus, up to multiplying by condensed bosons, we have

$$f_{\text{top}} \equiv \epsilon_1 b_1 b_3 \cdots f_{N/2} = m_2 m_4 \cdots m_N f_{\text{bottom}} \tag{3.83}$$

where $f_{\text{bottom}} = \epsilon_{N-1} m_N$. That is, by multiplying by a string of bulk magnetic charges, we can convert a fermion at the top layer into a fermion at the bottom layer. More concretely, there is a nonzero amplitude for a fermion to tunnel from the top layer to the bottom layer:

$$\begin{aligned} \langle \text{gs} |\, f_{\text{bottom}} m_2 m_4 \cdots m_N f_{\text{top}}^\dagger \,| \text{gs} \rangle &= \langle \text{gs} |\, f_{\text{top}} f_{\text{top}}^\dagger b_1 b_3 \cdots b_{N/2} \,| \text{gs} \rangle \\ &= v^{(N-2)/2} \langle \text{gs} |\, f_{\text{top}} f_{\text{top}}^\dagger \,| \text{gs} \rangle \neq 0. \end{aligned} \tag{3.84}$$

So from the point of view of someone living in the top layer, the fermion number can simply disappear.

A very similar construction produces a different SPT with the same symmetry, where all the quasiparticles are Kramers' doublets.

3.4 Wavefunctions for SPTs

Decorated defects. If we are interested in making paramagnetic states, one way to do it is to begin with the ordered state, and then destroy the order by condensing (topologically-protected) defects, like domain walls and vortices of the order parameter. This point of view in terms of defects in the would-be-bose-condensate turns out to be very fruitful. The completely trivial Mott insulator of the bosons with $\mathsf{U}(1)$ symmetry is described in this language as a condensate of featureless vortices (in 2+1d; in 3+1d they are vortex loops). (With $\mathbb{Z}_2$ symmetry, we would talk about domain walls instead.) More interesting states are obtained instead if the defects in the order carry some kind of decoration. A scheme for understanding SPT states then follows by characterizing possible decorations of the defects. I will not pursue it explicitly here.

BF theory for 3+1d boson SPTs. Consider the following $D = 3 + 1$ analog of CS theory

$$S[B, a] = \sum_I \frac{1}{2\pi} B^I \wedge da^I + \vartheta \sum_{IJ} \frac{K_{IJ}}{4\pi^2} da^I \wedge da^J$$

Note that the theta angle ϑ here is not the same as the θ in the magnetoelectric resonse, but time-reversal symmetry requires $\vartheta = 0$ or π.

This is called BF theory because the lagrangian is B times $f \equiv da$. It is topological, like CS theory, in that we didn't need to introduce the metric to integrate the action covariantly. In $D = 3 + 1$ we need the form degrees to add up $2 + 1 + 1 = 4$. We can add analogs of Maxwell terms (for both B and a), but just like in $D = 2 + 1$ they are irrelevant, *i.e.* they merely introduce new UV physics, they don't change the IR.

(The coefficient of the $B \wedge f$ term is chosen so that there is no bulk topological order. If we multiply it by k we get a description of $\mathbb{Z}_k$ gauge theory. Note that the more general seeming thing with a more general matrix coupling f and B can be removed by an integer-valued field redefinition that changes nothing.)

Briefly, who are these variables? Focus on the case $K = \boldsymbol{\sigma}^x$. In $D = 2 + 1$: the flux of the CS gauge field was some charge density. Here, each B^I implements a 3+1d version of charge-vortex duality, where for each boson current

$$j_\mu^{I=1,2} = \frac{1}{2\pi} \epsilon_{\mu \cdots \partial \cdot} B^I_{\cdot\cdot}$$

which has $\partial \cdot j^I_{\cdot} = 0$ as long as B is single-valued. The point of a is to say that B is flat, so that there are no local bulk degrees of freedom. You can also see from the equations of motion that the magnetic field lines of a^I_μ are the vortex lines of the microscopic bosons b_I.

One virtue of this effective action is that it reproduces the EM response we expect

of a topological insulator. If we couple to an external $\mathsf{U}(1)$ gauge field $\mathcal{A}$ by

$$\Delta\mathcal{L} = \mathcal{A}_\mu \, (j_1^\mu + j_2^\mu)$$

then

$$\log \int [DaDB] e^{\mathbf{i}S[a,B,\mathcal{A}]} = \int \frac{2\vartheta}{8\pi^2} \mathrm{d}\mathcal{A} \wedge \mathrm{d}\mathcal{A} + \cdots$$

that is, the magneto-electric response is $\theta_{EM} = 2\vartheta$. So $\vartheta = \pi$ will be a nontrivial boson TI.

Ways of slicing the path integral. Now let's think about the path integral for a QFT with a theta term. Examples include the BF theory above, and many non-linear sigma models which arise by coherent-state quantization of spin systems. In general what I mean by a theta term is a term in the action which is a total derivative, and where the object multiplied by theta evaluates to an integer on closed manifolds. The following point of view has been vigorously emphasized by Cenke Xu.

When spacetime is closed $Z(\theta + 2\pi) = Z(\theta)$. On a closed spacetime manifold M_D

$$Z_\theta(M_D) \equiv \int [D\text{stuff}] e^{-S} = \sum_{n\in\mathbb{Z}} e^{\mathbf{i}\theta n} Z_n$$

and $Z_\theta(M_D) = Z_{\theta+2\pi}(M_D)$. In particular, we can take $M_D = S^1 \times N_{D-1}$ to compute the partition function on any spatial manifold N_{D-1}. This means the bulk spectrum is periodic in θ with period 2π.

With boundaries, it not so in general. A boundary in space produces edge states. We've already said a lot about these.

A boundary in time in the path integral means we are computing wavefunctions. For quantum mechanics of a single variable $q(t)$, this is manifested in the Feynman-Kac formula for the position-space groundstate wavefunction:

$$\psi(q) = \int_{q(t_0)=q} \prod_{t\in(-\infty,t_0)} \mathrm{d}q(t) \; e^{-S_{\text{euclidean}}[q]} \; .$$

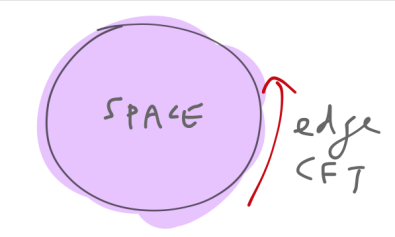
For a field theory, 'position-space wavefunction' means a wavefunctional $\Psi[\phi(x)]$, in

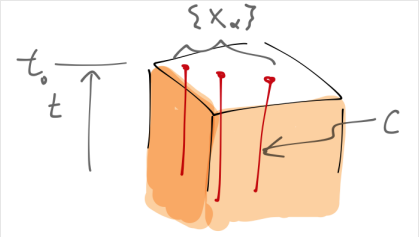
$$|\Psi\rangle = \int [D\phi(x)] \Psi[\phi(x)] \, |\phi(x)\rangle$$

where x labels *spatial* positions, and $|\phi(x)\rangle$ are coherent states for the field operator $\hat{\phi}(x)$. Which wavefunction? If the path integral is over a large euclidean time T before

reaching the boundary, this is a *groundstate* wavefunction, since the euclidean time propagator $e^{-T\mathbf{H}}$ is a (un-normalized) projector onto lowest-energy states.

Semi-philosophical digression: An important guiding concept in the study of interesting gapped states is that it is the *same stuff* living at a spatial boundary (edge modes) as at a temporal boundary (the wavefunction). This perspective first arose (I think) in the context of quantum Hall states where one can write groundstate and several-quasiparticle wavefunctions as correlation functions of certain operators in a 1+1d conformal field theory (CFT), which is the same CFT that arises at a spatial edge. Why should this be true? It's because the bulk can be described by a path integral for a Chern-Simons gauge theory, which has a certain CFT (chiral bosons for abelian CS theory, more generally a WZW model) living at its boundaries, wherever they are. For a spatial boundary, it produces a copy of that CFT at the boundary (roughly the group-valued CFT field g is related to the CS gauge field by $a = g^{-1}\mathrm{d}g$).





For a temporal boundary, the path integral expression for the wavefunctional (with some Wilson line insertions at the positions of the electrons) takes the form

$$\Psi[g(x)] = \int_{a(t_0,x)=g^{-1}\mathrm{d}g} e^{\mathbf{i}S[a]} W[C] = \left\langle \prod_\alpha V_\alpha(x_\alpha) \right\rangle_{WZW} . \tag{3.85}$$

A too-brief explanation of this rich formula: the Wilson line insertion is $W[C] = \mathrm{tr}_R \mathcal{P} e^{\mathbf{i}\oint_C a}$ where R is a representation of the gauge group G and $\mathcal{P}$ is path ordering. C is a collection of curves ending at the points x_α. In a TQFT describing a FQH state made of electrons, there is a quasiparticle with the quantum numbers of the electron; the representation R is the one corresponding to the electrons. In the rightmost expression, (3.85), x_α are the locations where the curve C intersects the fixed-$t = t_0$ surface, and V_α are some operators in the CFT the appropriate representations R of G.

For the case of the Laughlin state, $\mathsf{G} = \mathsf{U}(1)_q$, and the boundary value of the gauge field is $a(t_0, x) = g^{-1}\mathrm{d}g$, $g = e^{\mathbf{i}\phi(\vec{x})}$, where d is the spatial exterior derivative. The WZW model is just a free boson theory of the field ϕ. The Wilson lines are $e^{\mathbf{i}q\int_{C_\alpha} A} = e^{\mathbf{i}q\phi(x_\alpha)}$.

The correct version of the rightmost expression is

$$\left\langle \prod_{\alpha=1}^{N} e^{\mathbf{i}q\phi(x_\alpha)} e^{-\mathbf{i}\int d^2x \rho_0 \phi(x)} \right\rangle = \mathcal{N} \prod_{\alpha<\beta} (x_\alpha - x_\beta)^q e^{-\frac{1}{4}\sum_\alpha |x_\alpha|^2} \tag{3.86}$$

where the last factor on the LHS is a background charge $\rho_0 = \sqrt{q}N$, necessary to get a nonzero answer (but I haven't understood how it arises from the CS theory). This is the Laughlin wavefunction! In (3.86), we used the free boson CFT propagator

$$\langle \phi(z)\phi(w)\rangle = -\frac{1}{q}\log(z-w) \tag{3.87}$$

(the $1/q$ comes from the CS term $\frac{q}{4\pi}\int_{M_3} ada|_{a=d\phi} = \int_{\partial M_3} \frac{q}{2\pi}\partial_z\phi\partial_{\bar{z}}\phi$) and perhaps some sleight of hand. See the paper by Moore and Read for more on this step. Some comments and confessions:

- I haven't seen the step from the CS path integral (3.85) to (3.86) anywhere in the literature. It seems important. This paper has a related discussion for the $p + \mathbf{i}p$ state.
- In this step, I don't see yet where the background charge term comes from; obviously it has to be there or else the integral over the zeromode of ϕ produces a delta function saying that $Nq = 0$.
- Moore and Read say that the cross-terms between background charge and the vertex operators produces the $\sum_\alpha |x_\alpha|^2/4$ term, but I didn't understand this yet. Some of the details of this calculation are explained more in this review by Hansson et al, which I learned about a few months after this class ended.
- In (3.86) the self-interaction of the background charge goes into the multiplicative constant $\mathcal{N}$, which goes away in the normalized state.

Side remark: the canonical application of this story is to the *Haldane chain* – a chain where each site carries a representation of SO(3). At low energies, such chains are described by an NLSM with a theta term. $\theta = 0$ is trivial and gapped. $\theta = 2\pi$ is gapped and trivial in the bulk but the edge states are spin $\frac{1}{2}$s – a projective representation of SO(3). See the homework.

Let's apply this picture to BF theory for 3+1d boson SPT states. The analogous bulk EFT is, instead of CS gauge theory, some weird BF theory or strongly-coupled

sigma model. At a spatial edge, we have some vortex excitations in $D = 2 + 1$. Correspondingly, the bulk wavefunctions will turn out to have a nice representation in a basis of states labelled by vortex loop configurations in $D = 3 + 1$.

In contrast to the case of a closed manifold, if we compute the path integral on a space with a boundary say at $\tau = 0$, then θ does matter, not just mod 2π. Think of the space as the lower-half of euclidean space, so we are computing the groundstate wavefunctional.

Choose $a_\tau = 0$ gauge. Since a and B are conjugate variables, the analog of position space here is $|\vec{a}(x)\rangle$. For the same reason, we can only specify BCs on one or the other:

$$\int_{\vec{a}(x,\tau=0)=\vec{a}(x)} [D\vec{a}(x,\tau)DB(x,\tau)]e^{-S[\vec{a}(x,\tau),B]} = \left\langle \vec{a}(x)\middle|\mathrm{gs}\right\rangle = \Psi[\vec{a}(x)] \ . \tag{3.88}$$

Notice that in expressions for functionals like $S[a(x,\tau)]$ I am writing the arguments of the function a to emphasize whether it is a function at fixed euclidean time or not. The fact that the theta term is a total derivative

$$f^I \wedge f^J = \mathrm{d}\left(a^I \wedge f^J\right) \equiv 4\pi^2 \mathrm{d}w(a)$$

means that the euclidean action here is

$$S[\vec{a}(x,\tau)] = \int_{M_D} \frac{\mathbf{i}}{2\pi} B \wedge f + \mathbf{i}\vartheta \int_{\partial M_D} w(\vec{a}(x)).$$

The θ term only depends on the boundary values, and comes out of the integral in (3.88).

The integral over B^I is

$$\int [DB] e^{\mathbf{i}\frac{1}{2\pi}\int B^I \wedge f^I} = \delta[f^I].$$

The delta functional on the RHS here sets to zero the flux of the gauge field for points in the interior of the cylinder. After doing the integral over B, there is nothing left in the integral and (3.88) gives:

$$\Psi[\vec{a}_I(x)] = \exp \underbrace{\mathbf{i}\frac{\vartheta}{4\pi^2}\int_{\text{space}} a^I_{\cdot} \partial_{\cdot} a^J_{\cdot} \epsilon^{\cdots} K^{IJ}}_{K = \sigma^x \mathbf{i}\vartheta(\text{linking \# of } 2\pi \text{ magnetic flux lines})} \tag{3.89}$$

What does this mean? Label configurations of a by the flux loops (i.e. the field lines of the vector field). When $\vartheta = \pi$, this wavefunction is $(-1)^{\text{linking number of the 1-loops and the 2-loops}}$.

If we break the U(1) $\times$ U(1) symmetry, the flux lines of 1 and 2 will collimate (by the Meissner effect) into vortex strings.

Claim: in the presence of an edge, these flux lines can end. The ends of these flux lines are fermions. (Warning: doing this right requires a framing of the flux lines – *i.e.* they shouldn't collide.) Conclusion: on the surface of this SPT state of bosons there are *fermionic* vortices.

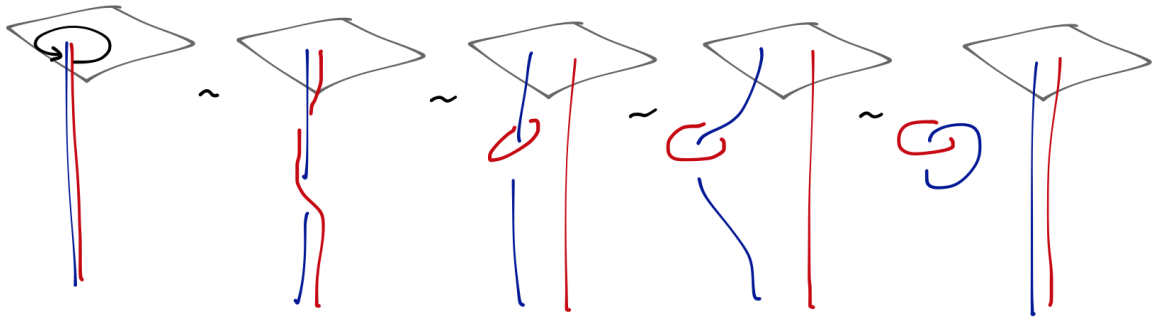


Figure 2: The end of the ribbon is a fermion: In the first step, we rotate the red string around the blue one. The squiggles mean that the states associated with these configurations have the same amplitude in the groundstate, according to (3.89).

Note that the BF theory describes a very strongly confined abelian gauge theory in the following sense: the flux gets set to zero by the B term. (With a string source for B the flux gets localized to the source.)

Comment on Kodama state of gravity. This wavefunction (3.89) actually solves the Schrödinger equation for quantum Maxwell theory at *finite* coupling. There is even a non-Abelian version of it for which this is also true. There is even an analog for gravity called the 'Kodama state'! What's the catch? It's not normalizable as a wavefunction for photon fields; attempting to quantize the model about this ground-state gives negative energy for one of the two circular polarization states. But as a wavefunction for the confining phase of the gauge theory it's fine.

Note that this paper does an analogous thing for very strongly coupled sigma models with theta terms; they just set the kinetic term to zero (!) and find wavefunctions closely analogous to (3.89). They would have the same problem as Witten points out if they thought of their wavefunctions as wavefunctions for gapless magnons. But for the disordered phase of the sigma model (gapped and analogous to confinement) it is just fine. Some of these are the same wavefunctions we'll arrive at in the following discussion.

Group cohomology SPTs. Here is a nice construction of a large class of bosonic SPTs.

To begin, suppose that spacetime is chopped up into a simplicial complex. This means that it is divided up into volumes, each of which is a D-simplex, and each of the faces is a $D-1$-simplex and so on. We will construct a path integral. The degrees of freedom are elements of G living on the 0-simplices (vertices). The path integral $Z = |G|^{-n_v} \sum_{\{g\}} \mathcal{A}(\{g\})$ is a sum over these group variables of an amplitude $\mathcal{A}(g) \equiv e^{-S(\{g\})}$ with the following properties:

1. It is G-symmetric, *i.e.* $\mathcal{A}(\{gg_i\}) = \mathcal{A}(\{g_i\})$, $g \in \mathsf{G}$.
2. The amplitude for any configuration $e^{-S(\{g\})} = 1$ on a closed spacetime manifold.
3. The amplitude $e^{-S(\{g\})}$ is subdivision invariant. That is, it represents a fixed-point of the renormalization group.

[End of Lecture 17]

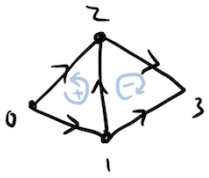
The essential ingredient is a group cocycle, which is first of all a map:

$$\nu_D : G^{D+1} \to \mathsf{U}(1)$$

– given $D+1$ elements of G, it determines a phase. Notice that a D-simplex is specified by $D+1$ vertices. The amplitude is a product of a bunch of these phases, one for each D-simplex:

$$Z = \frac{1}{|G|^{n_v}} \sum_{\{g\}} \prod_{\text{simplices, } [i_0 \cdots i_D]} \nu_D^{s_{i_0 \cdots i_D}}(g_{i_0} \cdots g_{i_D}). \tag{3.90}$$

Here $s_{i_0 \cdots i_D} = \pm 1$ depending on the orientation of the simplex. To define the amplitude, actually we need to specify a *branching structure* of the simplicial complex: each edge gets an orientation, and no loops are allowed. This means that the vertices in a simplex can be ordered by the number of incoming arrows, and they are all different. A simplex has $s_{i_1 \cdots i_D} = +1$ if this order matches the orientation in spacetime. For example: in the figure, the simplex [012] has $s_{012} = +1$ since 012 are traversed anticlockwise, but [123] has orientation $s_{123} = -1$ since 123 are traversed anticlockwise.



The properties claimed above follow from the two further conditions that are part of the definition of a group cocycle. First, ν is G-invariant:

$$\nu_D^{s(g)}(g_0, g_1 \cdots g_D) = \nu_D(gg_0, gg_1 \cdots gg_D) \tag{3.91}$$

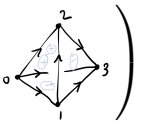
where $s(g) = \pm 1$ according to whether g is unitary or antiunitary (an antiunitary transformation reverses the orientations of the spacetime simplices). And secondly, ν is a cocycle, $(\delta\nu) = 1$, or more explicitly

$$\prod_{i=0}^{D+1} \nu_D^{(-1)^i}(g_0 \cdots g_{i-1}, g_{i+1} \cdots g_{D+1}) = 1. \tag{3.92}$$

Notice that the replacement $\nu_D \to \nu_D\,(\delta\mu)$ for some $(D-1)$-cochain μ would work just as well. But the factors of μ cancel out in the product (3.90) (on a closed manifold).

For example, the cocycle condition (3.92) in $D = 2$ is exactly the statement that the summand of the partition function of the model on a tetrahedron is equal to one:

$$1 \overset{(3.92)}{=} \frac{\nu(123)\nu(013)}{\nu(023)\nu(012)} \leftrightarrow 1 = Z\left(\right. \tag{3.93}$$



$$\left. \right).$$

(Here I am using the numbers $0 \cdots 3$ to stand for the group elements $g_0 \cdots g_3$.) We can make more general closed 2-manifolds by gluing tetrahedra along their faces (the connect sum operation). A similar statement to (3.93) holds in higher dimensions.

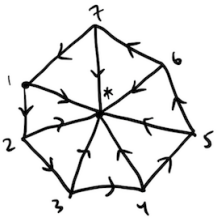
The data defining Z is therefore an element of the group cohomology group $[\nu] \in H^D(G, \mathsf{U}(1)) \equiv$ closed D-cocycles modulo exact D-cocycles.

There are several ways to see that nontrivial ν produces a nontrivial phase of matter, protected by G. The first is to couple this system to background G-gauge fields on the links of the complex. This just means G-valued variables on the links. If we make these variables dynamical, this is exactly the partition function of the Dijkgraaf-Witten discrete G-gauge theory twisted by the group cocycle ν. The cocycle condition (3.93) (more generally (3.92)) is the condition that the bulk DW partition function is gauge invariant on a closed manifold. For example, in $D = 2 + 1$ this procedure produces two different $\mathbb{Z}_2$ gauge theories; one is the toric code (associated with K-matrix $K = 2\sigma^x$), whose fixed-point groundstate wavefunction is a uniform superposition of closed loops. The other, associated with the nontrivial cocycle in $H^3(\mathbb{Z}_2, \mathsf{U}(1))$, is the double semion model (associated with K-matrix $K = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}$). Its fixed-point groundstate wavefunction is a sum over closed loops weighted by $(-1)^{\text{number of loops}}$. So if you believe that these theories are different for different ν, then the SPTs must be different.

A second, perhaps more direct way is to look at the groundstate wavefunction that's defined by this path integral. That is, consider the path integral on the lower-half plane with boundary conditions on the real line which produces the groundstate wavefunction. Equivalently, we can replace the half-plane with the ball. But since the

model is subdivision invariant, we can coarse-grain until there is only a single interior vertex in the ball. (This can be called a *minimal diagram*.)
Denoting by h the boundary group elements, and labelling the interior vertex $\star$, the resulting wavefunction is

$$\Psi_\nu(\{h\}) = \frac{1}{|G|}\sum_{g^\star}\prod_{\text{simplices},i_1\cdots i_D\star}\nu_D^{s_{i_1\cdots i_D\star}}(h_{i_1},\cdots h_{i_D},g^\star)\ . \qquad (3.94)$$



Here is the minimal diagram for $D = 1+1$.

But by property (3.92), the summand is actually independent of $g^\star$. So we can instead write

$$\Psi_\nu(\{h\}) = \prod_{\text{simplices},i_1\cdots i_D\star}\nu_D^{s_{i_1\cdots i_D\star}}(h_{i_1},\cdots h_{i_D},g^\star) \qquad (3.95)$$

with no sum.

Now, what is the Hilbert space where this wavefunction lives? To each vertex (0-simplex), associate a Hilbert space which is a regular representation of G, that is, the Hilbert space of site i is $\mathcal{H}_i = \text{span}\{|g\rangle_i, g\in G\}$, and these states are orthonormal. Define $|1\rangle_i \equiv \sum_{g\in G}|g\rangle_i$ (I don't normalize it just to avoid clutter later). Then the Hamiltonian

$$H_0 \equiv -\sum_i |1\rangle\langle 1|_i$$

has a unique groundstate $|\mathbb{1}\rangle \equiv \otimes_i |1\rangle_i$, a trivial product state. Its wavefunction in the group-element basis is $\Psi(\{g\}) \equiv \langle\{g\}|\mathbb{1}\rangle = 1$, that is, $|1\rangle_i \equiv \sum_{g\in G}|g\rangle_i$.

From this trivial state, we can make the state (3.95) using the following local, finite-depth *but not G-symmetric* unitary:

$$U_{g^\star} \equiv \sum_{\{h\}}\prod_{\text{simplices},i_1\cdots i_D\star}\nu_D^{s_{i_1\cdots i_D\star}}(h_{i_1},\cdots h_{i_D},g^\star)|\{h\}\rangle\langle\{h\}|.$$

So $H \equiv U_{g^\star}H_0U_{g^\star}^\dagger$ has groundstate $U_{g^\star}|\mathbb{1}\rangle = |\Psi_\nu\rangle$ with wavefunction (3.95). (If we sum over $g^\star$ to make U manifestly symmetric we get something that's not unitary.) This H is local (because U is a local circuit and H_0 is local) and *is* G-symmetric (this takes some work to show). Certainly one thing you can see is that U is not a product of symmetric gates – the individual factors of $\nu(h\cdots h,g^\star)$ are not at all G invariant. It's only the product that could be symmetric. So this is not an approximation to time evolution through G-symmetric local hamiltonains, and $|\Psi_\nu\rangle$ represents a nontrivial G-SPT phase.

Haldane chain example. Thinking about the Haldane chain as a $\mathbb{Z}_2^\mathcal{T}$ SPT, we can write it in this language as follows. Name the group elements $\mathbb{Z}_2^\mathcal{T} = \{e,t\}$, with

$t^2 = e$. The not-one elements of the nontrivial cocycle are:

$$\nu_2(e,t,e) = \nu_2(t,e,t) = -1$$

– the associated projective representation is $U(t)^2 = -1$. This wavefunction doesn't look quite like the AKLT wavefunction (3.71). But it is related to it by local rearrangements, as follows.

First copy (in the g basis) each site into two sites $g_i = h_i^r = h_{i+1}^l$, so that the wavefunction is $\Psi_\nu(\{h_i^r = h_{i+1}^l = g_i\}) = \prod_i \nu_2(g_i, g_{i+1}, g^\star)$. So far this was adding in ancillary bits and acting with a local unitary. Now regard h_i^r and h_i^l as making up a single site, as the labelling suggests. And do a local basis transformation by the unitary $\otimes_i u_i$ with

$$u_i \equiv \nu_2(h_i^l, h_i^r, g^\star)|h_i^l, h_i^r\rangle\langle h_i^l, h_i^r| = \nu_2(g_{i-1}, g_i, g^\star)|h_i^l, h_i^r\rangle\langle h_i^l, h_i^r|.$$

This gets rid of the ν_2 in the wavefunction and makes it exactly the state (3.71).

The same would have been true if ν_2 were equal to one. The nontrivial bit in the new basis is how the symmetry acts. In the original basis, the nontrivial element acted by $\mathbf{i} \to -\mathbf{i}$ combined with $e \leftrightarrow t$, *i.e.*

$$W_g\left(a\,|\{h\}\rangle\right) = a^\star\,|\{gh\}\rangle\,.$$

In the new basis you can check (using both defining properties of ν) that it acts by $W = \otimes W_i$ with

$$W_i(g)\left|h_i^l h_i^r\right\rangle = \nu_2(h_i^r, \bar{g}g^\star, g^\star)\nu_2^\star(h_i^l, \bar{g}g^\star, g^\star)\left|h_i^l h_i^r\right\rangle.$$

It factorizes in its action on the two parts of the site! This means that at the ends of the chain we'll be left with just one factor – a projective representation with exactly the phase ν_2.

Cluster state example. In $D = 1+1$ with $G = \mathbb{Z}_2 \times \mathbb{Z}_2$, we have $H^2(G, U(1)) = \mathbb{Z}_2$. Let's represent the group elements by $g = (s_e, s_o)$, $s_{e/o} = 0, 1$. A representative of the nontrivial class is specified by

$$\nu(1, g_1, g_2) = \begin{cases} -1 & ,\text{if } g_1 = g_2 = (1,1) \\ +1 & ,else \end{cases}. \tag{3.96}$$

The Hilbert space is then a tensor product of two qubits, and the whole Hilbert space is then a chain of (an even number of) qubits, where the symmetry is generated by $X_{e/o} \equiv \prod_{i,\text{even/odd}} X_i$. Up to multiplying by a coboundary, the result is

$$H = UH_0U^\dagger = -\sum_i Z_{i-1}X_iZ_{i+1}, \quad U = \prod_i \mathsf{CZ}_{i,i+1}, \tag{3.97}$$

where CZ_{ij} is the control-Z operation, and $H_0 = -\sum_i X_i$. The groundstate is $U\prod_i |+\rangle_i$.

An example in $D = 2+1$ with $G = \mathbb{Z}_2$ (unitary) is developed in detail in this paper. Gauging the $\mathbb{Z}_2$ symmetry results in the double-semion theory. Its edge theory is the same one we found for the boson IQHE: only the left-mover is charged, so the $\mathbb{Z}_2$ symmetry acts by T-duality on the edge mode, $(\phi_L, \phi_R) \mapsto (\phi_L, -\phi_R)$.

3.5 Classification of SPTs, part one

[I found this document very helpful.] Stacking is something we can do with anything. Given two physical systems A and B, we can make $A+B$ just by putting A and B on top of each other. If A and B are specified by G-preserving, gapped groundstates of G-symmetric Hamiltonians with an energy gap, then $A+B$ is, too.

To be a little more precise, let's say $\mathcal{M}_G^d$ is the space of deformation classes of triples $(\mathcal{H}, \rho, H)$, where $\mathcal{H}$ is a Hilbert space made as a tensor product (or, for fermions, a graded tensor product) on a d-dimensional lattice, ρ is a representation of G on the Hilbert space $\mathcal{H}$, and H is a local hamiltonian on $\mathcal{H}$. By deformation classes, we can mean adding ancillas and adiabatic variation without closing the gap.

So stacking means $\mathcal{H}_{A+B} = \mathcal{H}_A \oplus \mathcal{H}_B$, $\rho_{A+B} = \rho_A \oplus \rho_B$, and $H_{A+B} = H_A + H_B$, and then we are allowed to deform while preserving the gap. Stacking with the trivial phase does nothing. So stacking gives $\mathcal{M}_G^d$ the structure of a 'monoid' – a group except for inverses.

A better definition of SPTs with G symmetry in d dimensions is

$$\mathrm{SPT}_G^d \equiv \text{invertible elements of } \mathcal{M}_G^d \ . \tag{3.98}$$

Thus SPT_G^d is a group, essentially by definition. Moreover, it is not too hard to see that $\mathcal{M}_G^d$ is associative and commutative. Basically it's just because $\mathcal{H}_1 \oplus \mathcal{H}_2 = \mathcal{H}_2 \oplus \mathcal{H}_1$. Therefore SPT_G^d is an abelian group. Discrete abelian groups are pretty simple. How can we resist asking: for a given G, d (and specification of the microscopic constituents, *i.e.* is it made of bosons or fermions) which one is it?

But first, some important comments about this definition. This definition is different from (a definition of SPTs that was made in the earlier literature)

$\mathrm{SRE}_G^d \equiv$ G-symmetric non-SSB systems
that *can* be deformed to the trivial state while preserving the gap,
but not through G-symmetric Hamiltonians.

SRE stands for 'short-range-entangled', and my notation here is not standard. SRE_G^d is a subgroup of SPT_G^d (and in fact a direct summand).

Why is definition (3.98) is better? The condition of invertibility is weaker than deformability to the trivial phase in the absence of symmetry, but it still implies the absence of anyons (since obviously you can't get rid of anyons by adding more anyons). It should therefore also imply a unique groundstate on any closed manifold. These properties are easier to check than deformability to the trivial phase.

Nontrivial invertible states with no symmetry. Unlike SRE^d_G, our definition of SPT^d_G is nontrivial even when $G = \{e\}$ is the group with one element. Such systems are protected instead by a gravitational anomaly. For fermions, examples of such states are the following:

- In $D = 0+1$, an odd number of majorana modes. You can see that there is a $\mathbb{Z}_2$ classification of such states, since, given two majorana modes, the perturbation $H = \mathbf{i}\gamma_1\gamma_2$ lifts both of them.
- In $D = 1+1$, the Kitaev chain. Oh there is a lot to say about this.
- In $D = 2+1$, a $p + \mathbf{i}p$ superconductor, two copies of which are an IQH system after the $\mathsf{U}(1)$ symmetry is broken. In this case, the central charge of the edge modes indicates the gravitational anomaly.

For bosons, the first example is the E_8 state in $D = 2+1$. Our description of this state is abelian CS theory with K-matrix equal to the Cartan matrix of the Lie algebra E_8. What's special about E_8? The E_8 lattice has the remarkable property that it is *even* and *self-dual*. Lattices with both of these properties only occur in dimensions that are multiples of 8. (These are the same reasons that E_8 plays an important role in string theory.) These properties imply in particular that $|\det K| = 1$, so there is a unique groundstate on any Riemann surface, and that all the quasiparticles are bosons. So there is no topological order.

Another example is the state in $D = 4+1$ described by the analog of CS theory with $k = 1$. It is a theory of two two-form potentials with action $S[B, C] = \frac{k}{2\pi}\int B \wedge dC|_{k=1}$. For k an integer larger than one, this is a description of $\mathbb{Z}_k$ gauge theory, when $k = 1$ there is no topological order, but it is a nontrivial invertible phase. One description of the edge theory is as electrodynamics where both the electric charge and the magnetic monopole are fermions. I'll say more about how to see the anomaly below.

Functorial properties. Think of SPT^d_G as a machine that eats a group G and spits out a discrete abelian group. In trying to figure out what this machine does, the following observation is valuable: In fact $\mathrm{SPT}^d_\bullet$ is a (contravariant) functor, from groups to groups. That means that it maps the structure of the victim to the structure of the output: given a group homomorphism $\varphi : H \to G$, there is automatically a

group homomorphism $\varphi^\star : \mathrm{SPT}_G^d \to \mathrm{SPT}_H^d$. (The fact that the arrows get reversed is the reason for the adjective 'contravariant'.)

Despite the intimidating words, this is simple. First consider the case where the homomorphism is just inclusion of a subgroup: $H \overset{i}{\hookrightarrow} G$. Then $[A] \in \mathrm{SPT}_G^d$ is *also* $[A] \in \mathrm{SPT}_H^d$, since anything G-symmetric is also $H \subset G$-symmetric. (Note that a nontrivial G-SPT could be trivial as an H-SPT.) More generally, given a group homomorphism $\varphi : H \to G$ and a G-SPT, $(\mathcal{H}, \rho, H)$, we can make an H-SPT by

$$\varphi^\star(\mathcal{H}, \rho, H) = [(\mathcal{H}, \rho \circ \varphi, H)]. \tag{3.99}$$

Notice that

$$\rho \circ \varphi : H \overset{\varphi}{\to} G \overset{\rho}{\to} \mathcal{U}(\mathcal{H}) \tag{3.100}$$

(where $\mathcal{U}(\mathcal{H})$ means unitaries on $\mathcal{H}$) is a rep of H.

3.6 Global anomaly inflow

[We'll follow Witten for a bit. I found this paper by García Extebarria and Montero very useful.]

How to discover that anomalies are encoded in an extra-dimensional theory. Let me return once more to the gapless edge theory of a $(3+1)$D TI with $G = \mathsf{U}(1) \ltimes \mathbb{Z}_2^{\mathcal{T}}$. A mass term for the Dirac fermion in 2+1d breaks time-reversal symmetry. With two such modes, $\int \mathbf{i}\bar{\eta}_1\eta_2$ would be $\mathcal{T}$-invariant, so there is a $\mathbb{Z}_2$ classification of anomalous theories. One way to think about this is that if we can add a symmetric mass term, we can use Pauli-Villars regularization to define the path integral in a manifestly symmetric way so there can be no anomaly.

Let's think about trying to define the path integral for η on some oriented spacetime X:

$$Z_X[A, g] = \int D\eta e^{\int d^{2+1}x \bar{\eta}\mathcal{D}\eta} = \det(\mathcal{D}), \quad \mathcal{D} \equiv \mathbf{i}\gamma^\mu D_\mu.$$

This dirac operator $\mathcal{D} = \mathcal{D}^\dagger$ is hermitian and so has real eigenvalues, so $Z_X = \prod_i \lambda_i \in \mathbb{R}$. This is a consequence of $\mathcal{T}$-invariance: in general, for an oriented manifold X, $Z^\star(X) = Z(\bar{X})$, where $\bar{X}$ has the reversed orientation. With two Dirac fermions, the partition function would be $Z_X^2 = \prod_i \lambda_i^2 > 0$ would be positive, but with one copy the sign of Z_X is not fixed:

$$\mathrm{sign} Z_X = (-1)^{\#\text{of negative eigenvalues of } \mathcal{D} \text{ mod two}}.$$

In the continuum, the object in the exponent is infinity. Note that if we are willing to break $\mathcal{T}$ and add a mass, then there is no problem defining the sign of Z_X.

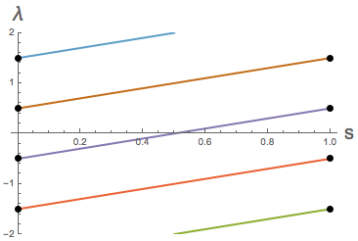
You might not think to care about the sign of Z_X, and think that we can just declare it to be positive. But here's why it's non-trivial: just declaring it to be positive is not a gauge-invariant choice. Z_X here is really a functional of the background fields (g, A), where A is the background $\mathsf{U}(1)$ gauge field and g is the metric on X. Suppose we declare $Z(g_0, A_0) > 0$ for some reference background. Then we define the sign of $Z(g, A)$ for general (g, A) by finding a path from (g_0, A_0) to (g, A) in the space of background fields. Z changes sign whenever an eigenvalue λ crosses 0 along this path.

But: let ϕ be some gauge transformation (or coordinate change, *i.e.* diffeomorphism), under which $(g_0, A_0) \to (g_0^\phi, A_0^\phi)$. Then

$$A_s = (1-s)A_0 + sA_0^\phi, \quad g_s = (1-s)g_0 + sg_0^\phi$$

is an allowed background for $s \in [0, 1]$ (note that $g_s > 0$ if $g_{0,1} > 0$). Now we can ask: as we vary s, how many eigenvalues cross 0?

Gauge invariance of our definition requires an even number. The fact that ϕ is a gauge transformation means that the spectrum of $\mathcal{D}$ is the same at $s = 0, 1$. But the fact that the spectrum goes off to infinity means that a given eigenvalue need not return to itself – there can be *spectral flow*.



There is an index formula for the net number Δ of eigenvalues crossing $\lambda = 0$, so

$$Z(g_0, A_0) = (-1)^\Delta Z(g_0^\phi, A_0^\phi).$$

The formula is $\Delta = \mathfrak{I}$ where $\mathfrak{I}$ is the index of a $(3+1)$D Dirac operator $\hat{\mathcal{D}}$ on the following 4-manifold, Y_ϕ. Y_ϕ is called the *mapping torus*:

$$Y = I \times X/(0, x) \sim \phi(1, x)$$

where ϕ here is the action of ϕ on the metric. The same identification on A_s determines a $\mathsf{U}(1)$ connection (and bundle) on Y_ϕ. (If ϕ acts only as a gauge transformation and not a diffeomorphism, then the space Y_ϕ is just $S^1 \times X$.)

The conclusion here is that $Z_X(g, A)$ is not a *function* on the space of background fields, but rather a section of some bundle with structure group $\mathbb{Z}_2$. The nontriviality of this bundle is what makes the theory anomalous. (In this example, it says either we have to give up $\mathcal{T}$ or we have to give up gauge invariance, an easy choice.)

So we have an association between our anomaly (manifesting itself as a difficulty in defining the sign of the path integral on X) in $D-1$ dimensions, and a field theory in D dimensions. One way to think about the association is that if we coupled the two of them together, there would be no anomaly – this is anomaly inflow. So to

understand Z_X, we can find a manifold Y with $\partial Y = X$ and extend the gauge bundle on X to Y and put the 'anomaly theory' on Y. In this case the 'anomaly theory' is a $(3+1)D$ massive Dirac fermion with $M < 0$. Think about the bulk path integral on a background configuration with Dirac index $\mathfrak{I}$: this means, generically that there are $\mathfrak{I}$ left-handed zeromodes ψ_+^i (and their conjugates $\bar{\psi}_-^i$). The path integral is then of the form

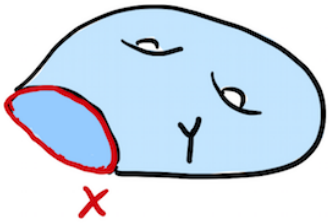
$$Z_{\text{bulk}} = \int [D\Psi] e^{\int \bar{\Psi}(i\hat{\not{D}} - M)\Psi + \cdots} \propto \int d\psi_+^1 \cdots d\psi_+^{\mathfrak{I}} d\bar{\psi}_-^1 \cdots d\bar{\psi}_-^{\mathfrak{I}} e^{M\bar{\psi}_-\psi_+} = M^{\mathfrak{I}}.$$

In the TI phase, $M < 0$, so the sign of this path integral is indeed $(-1)^{\mathfrak{I}}$.

[End of Lecture 18]

Now let's generalize. We want to identify possible anomalies of a system in $D-1$ spacetime dimensions: failures of the effective action (as a functional of general background fields, on general manifolds) to be gauge invariant. Such a system is then a candidate for a characteristic surface state of an SPT in D dimensions. The construction here will incorporate the D dimensional theory quite organically.

To begin, suppose we wish to understand the partition function Z_X of our putative anomalous theory on X, a $D-1$ manifold. Z_X is a section of a nontrivial bundle on the space of background fields. The first step is to find a D-manifold Y with $\partial Y = X$[59]. Any structure that we care about, such as a background gauge field or a spin structure on X we must also extend to Y. The key idea is that there is a quantity associated with Y that's a section of the same bundle – it's the partition function of the associated 'anomaly theory', *i.e.* the bulk SPT.



In the case where the anomalous edge theory in question is just free fermions, the 'anomaly theory' is a massive fermion in one extra dimension. The dynamics of the massive fermion are not important at all here, and sometimes the 'theory' is described as just an association of a phase to each D-manifold, called the η-invariant:

$$e^{2\pi i \eta_Y} \equiv \text{partition function of SPT on Y}$$

(with boundary conditions compatible with X).

To get started at appreciating this idea, note that it incorporates all the perturbative (or 'local') anomalies of fermions by taking $Y = \partial Z, Z = X \times D$ (D is a disk). Then

[59] In fact, not all X admit such a Y. What to do in that case I hope to explain below. If I fail, see section 6 of the García-Etxebarria and Montero paper.

the APS index theorem for manifolds with boundary says that the index of the Dirac operator on Z is

$$\mathbb{Z} \ni \mathrm{Ind}\,(\mathbf{i}\mathcal{D}_Z) = \eta_Y + \int_Z \widehat{A}(R) \wedge \mathrm{tr} e^{\frac{F}{2\pi}}$$

which says

$$e^{2\pi\mathbf{i}\eta_Y} = \exp\left(2\pi\mathbf{i}\int_Z \left(\widehat{A}(R)\wedge \mathrm{tr} e^{\frac{F}{2\pi}}\right)_{D+2}\right) = e^{2\pi\mathbf{i}\int_Z I_{D+1}} \tag{3.101}$$

whose variation, by the WZ descent procedure, reproduces the chiral anomaly in $D-1$ dimensions.

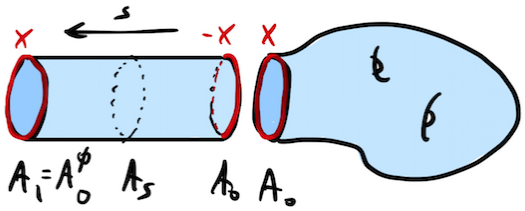
Any anomaly in a gauge transformation not continuously connected to the identity is called a *global anomaly*. Such a gauge transformation is a map from spacetime X to G. Homotopy classes of such maps are denoted $[X, G]$. For example, the one-point compactification of $\mathbb{R}^d$ is $X = S^d$, and $[S^d, G] \sim \pi_d(G)$ (under some assumptions about G). For example, in a system in $d = 3+1$, with $G = \mathsf{SU}(2)$, we have $\pi_4(\mathsf{SU}(2)) = \mathbb{Z}_2$ and there is an interesting possibility of an $\mathsf{SU}(2)$ anomaly.

The η invariant has some nice properties that we want for any SPT: If we can decompose $M = Y_1 \cup Y_2$ as the result of gluing two open manifolds along cancelling boundaries ($\partial Y_1 = -\partial Y_2$), then

$$e^{2\pi\mathbf{i}\eta_{Y_1\cup Y_2}} = e^{2\pi\mathbf{i}\eta_{Y_1}} e^{2\pi\mathbf{i}\eta_{Y_2}}. \tag{3.102}$$

So for example the change of phase from $Z[A_0]$ to $Z[A_0^\phi]$ is

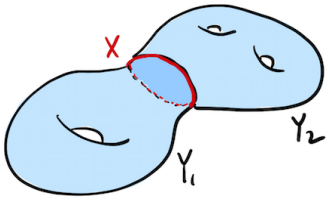
$$Z[A_0^\phi] = e^{2\pi\mathbf{i}\eta_{X\times I}} Z[A_0].$$



This means that this idea incorporates the mapping-torus global anomaly we discussed above.

More general anomalies. Now if the partition function of the anomaly theory on a mapping torus is not equal to one, it clearly represents a failure of gauge invariance of the theory on X. But what about the partition function of the anomaly theory on other manifolds?

Take any closed D-manifold Y and chop it into two parts: $Y = Y_1 \cup Y_2$ with $\partial Y_1 = X = -\partial Y_2$. Then Y_1 and Y_2 each represent a possible extension of X that could be use to define the partition function of the edge theory on X. But if $e^{\mathbf{i}2\pi\eta(Y_1 \cup Y_2)} \neq 1$, it means that the answer for the partition function on X depends on the extension to Y! So this is also an anomaly. (Sometimes this case is called a *Dai-Freed* anomaly.)



For the case of free fermions, if the local anomalies vanish (this means the anomaly polynomial $I = 0$), then η_Y is topological, and moreover is a bordism invariant. To see this, suppose Y_1 is bordant to Y_2. This means that there exists Z with $\partial Z = Y_1 \amalg \bar{Y}_2$ (that symbol $\amalg$ means disjoint union). Now use (3.102) to write

$$e^{2\pi\mathbf{i}\eta_{Y_1 \amalg \bar{Y}_2}} = \frac{e^{2\pi\mathbf{i}\eta_Y}}{e^{2\pi\mathbf{i}\eta_{Y_2}}} \overset{(3.101)}{=} e^{2\pi\mathbf{i}\int_Z I_{D+2}} = 1 \tag{3.103}$$

by the assumption of no local anomalies. In the first step of (3.103) we used the fact that $e^{2\pi\mathbf{i}\eta_{\bar{Y}}} = e^{-2\pi\mathbf{i}\eta_Y}$.

3.7 Spin structures and fermions

In unitary, continuum, relativistic theories, fermions are always created by spinor fields. This is the spin-statistics theorem. I have two points to address on this subject. First, in the preceding discussions, we've often found it useful to consider putting our systems in spaces or spacetimes that are not just flat space. The ability to put spinor fields on a particular manifold M is not automatic – the manifold must have a *spin structure*. Essentially, this is a consistent set of choices of sign for each loop in M, in recognition of the fact that a 2π rotation should produce a minus sign on a fermion. Some manifolds admit many different spin structures, and some do not admit any spin structure at all.

Second, in condensed matter physics, we often give up the adjectives 'continuum' and 'relativistic'. Does a theorem relating spin and statistics still hold? The answer is yes, if it is satisfied microscopically.

[The subject of spin structures is often treated in a very abstract way. Here is a paper that can help make it concrete. Another useful reference is the lectures by Preskill on vortices. Also these lectures.]

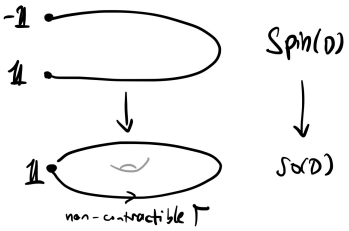
To put a spin structure on a manifold M, we need a notion of parallel transport of not just vectors (local sections of the tangent bundle, which has structure group $\mathsf{SO}(D)$ for an oriented manifold), but spinors. The structure group for the bundle of which

spinors are local sections is $G = \mathfrak{Spin}(D)$, the double cover of $\mathsf{SO}(D)$. It's a double cover because a 2π rotation of a spinor acts by $-\mathbb{1}$.

The simplest example to keep in mind is $M = T^D$, where we must choose a boundary condition for fermions around each 1-cycle of the torus, periodic or antiperiodic. These two choices are associated with the names Ramond and Neveu-Schwarz respectively. So there are 2^D possible choices of spin structure on T^D. Think about S^1 for a moment. It is the periodic (Ramond) boundary condition that's the nontrivial one, since it can't be extended to a disc whose boundary is the circle.

More generally, there can be an obstruction to the existence of such a double cover of the tangent bundle, if M contains non-contractible spheres. Given such a sphere, consider a family of closed loops $\gamma(t, s)$ (with the same base point p), starting and ending at the trivial loop ($\gamma(t, 0) = \gamma(t, 1) = p$ for all t). Parallel transporting vectors along each loop $\gamma(\cdot, s)$ produces an element of $\mathsf{SO}(D)$ for each s. Therefore $\gamma(t, s)$ produces a closed loop Γ in $\mathsf{SO}(D)$.

Now let's think about the loop Γ, which begins and ends at the identity, and therefore defines an element of $\pi_1(\mathsf{SO}(D))$. Γ may not be contractible, since $\pi_1(\mathsf{SO}(D)) = \mathbb{Z}_2$ (the double cover $\mathfrak{Spin}(D)$ is the universal cover). If Γ is not contractible, its lift to $\mathfrak{Spin}(D)$ will be an open path, starting at $\mathbb{1}$ and ending at $-\mathbb{1}$.



But that would mean that a spinor would have to change sign under an infinitesimal loop. There could then be no smooth spinor field on M.

Notice that $[\Gamma] \in \pi_1(\mathsf{SO}(D)) = \mathbb{Z}_2$ is a property of M. It classifies (up to homotopy) the $\mathsf{SO}(D)$ bundle on S^2 obtained by the restriction of the tangent bundle to S^2. If $[\Gamma] \neq 1$, this bundle describes a $\mathbb{Z}_2$ monopole. Note the similarity with the argument for Dirac quantization by considering the holonomy of the electron wavefunction around the equator.

An example of a manifold that admits no spin structure is $\mathbb{CP}^2 = \mathsf{SU}(3)/\mathsf{U}(2) = G/H$. Using this description as a coset, you can identify a non-contractible loop in $\mathsf{SO}(4)$.

This similarity with the monopole suggests a cure for the problem in the case where the spinor field is coupled to a gauge field. We can cancel the offending sign by a choice of wrong quantization for the flux of that gauge field. A spin$_c$ structure is obtained by

putting a monopole of charge g inside each sphere for which there is a problem, giving the spinors charge e with the wrong Dirac quantization condition $2eg = n + \frac{1}{2}, n \in \mathbb{Z}$. (Notice that by the spin-charge relation, systems made from electrons can be put on an arbitrary spacetime by introducing such a spin$_c$ structure.)

Stiefel-Whitney classes.. Here's a more systematic and general way of thinking about the possible obstruction to a spin structure we just encountered. Consider a good cover of our manifold M. This means a collection of open sets $\{U_\alpha\}_\alpha$, each of which is topologically a ball, whose intersections are also balls. We'll denote double-overlaps by $U_{\alpha\beta} \equiv U_\alpha \cap U_\beta$, and triple overlaps as $U_{\alpha\beta\gamma}$ and so on. The transition functions of the tangent bundle TM of a D-dimensional manifold M are maps on the double-overlaps

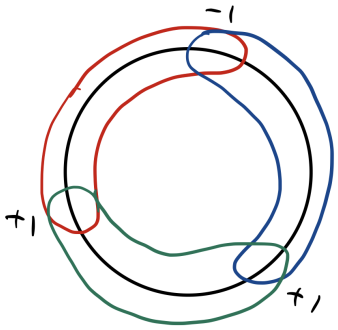
$$g_{\alpha\beta} : U_{\alpha\beta} \to \mathsf{O}(D). \tag{3.104}$$

If the manifold were oriented, these maps would live in $\mathsf{SO}(D)$, the component connected to the identity. In general then, to each $U_{\alpha\beta}$ we can associate a sign

$$(w_1)_{\alpha\beta} : U_{\alpha\beta} \to \mathbb{Z}_2 = \mathsf{O}(D)/\mathsf{SO}(D) \tag{3.105}$$

that says in which component of $\mathsf{O}(D)$ the map lives on that patch. (The map is just $(w_1)_{\alpha\beta} = \det g_{\alpha\beta}$.) I claim that $(\delta w_1)_{\alpha\beta\gamma} \equiv (w_1)_{\alpha\beta}(w_1)_{\beta\gamma}(w_1)_{\gamma\alpha} = 1$ on triple overlaps, *i.e.* w_1 is a Cech cocycle, and therefore defines a cohomology class $[w_1] \in H^1(M, \mathbb{Z}_2)$. If this class is the trivial class, then $(w_1)_{\alpha\beta} = (\delta s)_{\alpha\beta} = s_\alpha s_\beta$ is exact. Then $s_\alpha : U_\alpha \to \mathbb{Z}_2$ specifies an orientation of M. $w_1(TM) = [w_1]$, called the first Stiefel-Whitney class of M, therefore represents an obstruction to orienting the manifold.

For example, at right is a good cover of the circle. Three patches are required so that each pair overlaps only in one disk. I've indicated a set of transition functions $w_{\alpha\beta}$ representing a nontrivial element of $H^1(S^1, \mathbb{Z}_2)$. You can recognize them as defining the transition functions of the (not orientable) möbius bundle.



Now suppose that $[w_1] = 0$, and moreover that we've chosen an orientation $\{s_\alpha\}_\alpha$ of M, so that the $g_{\alpha\beta}$ live in $\mathsf{SO}(D)$. A spin structure on M requires a lift of these maps to the spin group

$$\hat{g}_{\alpha\beta} : U_{\alpha\beta} \to \mathfrak{Spin}(D).$$

On triple intersections, recall that the transition functions of a vector bundle satisfy the cocycle condition $(\delta g)_{\alpha\beta\gamma} \equiv g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = 1$ on $U_{\alpha\beta\gamma}$ (*i.e.* $g_{\alpha\beta}$ specifies an $\mathsf{SO}(D)$-valued

1-cocycle). This means that

$$w_{\alpha\beta\gamma} \equiv \hat{g}_{\alpha\beta}\hat{g}_{\beta\gamma}\hat{g}_{\gamma\alpha} : U_{\alpha\beta\gamma} \to \mathbb{Z}_2$$

is just a sign. I claim that w is a Cech cocycle, $(\delta w)_{\alpha\beta\gamma\delta} = 1$ (I don't know how to check this in a non-ugly way), and therefore $w_2(TM) \equiv [w] \in H^2(M, \mathbb{Z}_2)$ represents an element of Cech cohomology. This is the second Stiefel-Whitney class. If it is trivial, *i.e.* if $(w)_{\alpha\beta\gamma} = (\delta f)_{\alpha\beta\gamma} = f_{\alpha\beta}f_{\beta\gamma}f_{\gamma\alpha}$ is exact, for some $f_{\alpha\beta} : U_{\alpha\beta} \to \mathbb{Z}_2$, then we can just take as the spin structure

$$\hat{g}_{\alpha\beta} = g_{\alpha\beta}f_{\alpha\beta}. \tag{3.106}$$

But if w_2 is not exact, it is an obstruction to the existence of such an f.

3.8 Characteristic classes and classifying spaces

Stiefel-Whitney classes of TM are examples of a more general notion of characteristic classes of a bundle. There are many ways to think about characteristic classes. Here is one: The topology of a G-bundle on M can be specified by a map $\gamma : M \to BG$. This is the point in life of the *classifying space* BG. Such a space exists for each G. But then each cohomology class of BG produces a cohomology class on M via the pullback:

$$\gamma^\star : H^n(BG, A) \to H^n(M, A) \tag{3.107}$$

(for any abelian group A). The images of this map are characteristic classes.

Classifying space and universal bundle. One way to think about this is that there is a *universal G-bundle* with total space EG

$$\begin{array}{ccc} G & \longrightarrow & EG \\ & & \downarrow \xi \\ & & BG \end{array} \tag{3.108}$$

with the property that any bundle on X is the pullback of ξ via the map $\gamma : X \to BG$. (The fiber over the point $x \in X$ is $\xi^{-1}(\gamma(x))$.) BG is determined (up to homotopy) by demanding that EG is contractible and admits a free G action. Essentially, it's always some kind of infinite-dimensional sphere. For example, for $G = \mathsf{U}(1)$, it's the limit of the Hopf fibration as the dimension goes to ∞.

$$\begin{array}{ccc} \mathsf{U}(1) & \longrightarrow & S^{2\infty+1} \\ & & \downarrow \\ & & \mathbb{CP}^\infty = BU(1) \end{array} \qquad \begin{array}{ccc} \mathbb{Z}_2 & \longrightarrow & S^\infty \\ & & \downarrow \\ & & \mathbb{RP}^\infty = B\mathbb{Z}_2 \end{array} \qquad \begin{array}{ccc} SU(2) & \longrightarrow & S^{4\infty+1} \\ & & \downarrow \\ & & \mathbb{HP}^\infty = BSU(2) \end{array}$$

Note that for bundles on X of dimension $< k$ we actually only really care that $\pi_{q<k}(EG)$ vanishes, and so in practice we can just think about high-dimensional spheres and they'll do just as well.

By the way, here's a nice use of the classifying space to define the Dijgraaf-Witten path integral for a discrete G gauge theory with a cocycle $\nu \in H^D(G, \mathsf{U}(1))$. The group cohomology we defined above is in fact related to the *ordinary* cohomology of BG [60]. For discrete G the relation is just

$$H^D(BG, \mathsf{U}(1)) = H^D(G, \mathsf{U}(1)) \tag{3.109}$$

[60] In general, the relation is $H^D(G, \mathsf{U}(1)) = H^{D+1}(BG, \mathbb{Z})$; for discrete G, this is the same as $H^D(BG, \mathsf{U}(1))$ by the long exact sequence associated to $\mathbb{Z} \to \mathbb{R} \to \mathsf{U}(1)$, using the fact that for discrete G, $H^\bullet(BG, \mathbb{R}) = 0$.

(these very similar-looking objects have very different definitions!). So if spacetime M is divided up into a simplicial complex (with a branching structure), we specify a gauge field configuration by a map $\gamma : T \to BG$ for each simplex T, and the cocycle $\nu \in H^D(BG, \mathsf{U}(1))$ associates a phase $\nu(\gamma(T))$. The amplitude (summand of the partition function of the DW theory) is then just

$$W[\gamma] = \prod_T \nu^{s(T)}(\gamma(T)) \tag{3.110}$$

where $s(T)$ is the same sign choice as above[61]. Note that $W[\gamma]$ is (a discrete representative of) the characteristic class defined by the pullback of the class $\nu \in H^D(BG, \mathsf{U}(1))$ via the map γ. In fact, the relation between (3.110) and the group cohomology SPT construction above is the best way to understand the equation (3.109).

A second perspective, more obviously related to what we said in the discussion of spin structures, is from obstruction theory: try to construct a nowhere-zero section of the bundle V in question. Approximate M by a cell complex. First find a solution on the 0-cells. Then try to extend this solution from the p-cells to the $p+1$-cells: suppose given a vector field on $S^p = \partial B^{p+1}$, that is, a map $S^p \to \mathbb{R}^m$ (where m is the rank of the bundle, and I'm supposing that it's a bundle over $\mathbb{R}$ for example), can it be extended to the interior of the ball? The bundle itself can trivially be extended to B^{p+1}, since $\pi_\bullet(B) = 0$. But demanding that the vector field is nowhere zero means that this map is homotopic to a map $S^p \to S^{m-1}$. For $p+1 < m$, this is fine, but for $p+1 = m$, there is a winding number, the degree of the map, which if it's nontrivial obstructs the extension to the interior. So we have an integer for each m-cell, an element of $H^m(M, \mathbb{Z})$. This thing mod two is $w_m(V)$, the Euler class of the bundle V. Or: This failure to construct a nowhere-zero section means that any actual section has zeros. This locus of zeros gives the Poincaré dual of $w_m(V)$.

If we instead try to construct q linearly-independent nowhere-zero sections, we can orthogonalize the sections and produce a map $S^p \to S^{m-q}$, which gives an element of $H^{m-q}(M, \mathbb{Z}_2)$, which is the $w_{m-q}(V)$. The Poincaré dual to this is the locus of points where the sections fail to be linearly independent.

[61] Actually there is one more step: we must show that the partition function actually only depends on a configuration of link variables g_σ specifying a flat connection. The key idea is that a configuration of the link variables specifies a map $\gamma : M \to BG$ because of the homotopy equivalence of ΩBG and G. (This is consistent with $\pi_1(BG) = \pi_1(EG/G) = G = \pi_0(G)$ and with $BG = K(G,1)$ and $G = K(G,0)$.) Define γ to take all 0-cells to the base point of BG. A link-variable configuration $\{g_\sigma\}_{\sigma \in M_1}$ associates an element of G to each edge, and hence a homotopy class of closed paths in BG (starting and ending at the base point). Then the flatness of the connection allows us to extend this map to all the higher-dimensional cells.

I haven't shown that this definition is the same as the definition in terms of Cech cohomology above.

Actually there are some very simple examples of characteristic classes. Let's think about the case of a complex vector bundle V over X. Then choose a connection A on V. Let's call $G \subset \mathsf{U}(N)$ the structure group of V. Now make a G-invariant polynomial in $F = dA + A \wedge A$, where G acts by $F \to g^{-1}Fg$. For example,

$$c(F) \equiv \det\left(1 + \frac{\mathbf{i}F}{2\pi}\right) = 1 + c_1(F) + c_2(F) + \cdots$$

is the sum of symmetric polynomials in the eigenvalues of F called the *total Chern class of* F. $c_j(F)$ is the degree-j part of $c(F)$ and is a $2j$-form. Some properties of these objects that I will not prove:

1. $dc(F) = 0$, so these forms all cohomology classes on the base manifold X.

2. Their integrals $\oint_S c(F)$ over closed cycles $S \subset X$ depend only on the transition functions of the bundle V and not on the choice of connection. This is because given two connections $A_{0,1}$ one can take a convex combination of them ($A_t = A_0 + t(A_0 - A_1), F_t = dA_t + A_t^2$) and show that the difference of their Chern classes is exact:

$$c_m(F_1) - c_m(F_0) = md \int_0^1 dt \mathrm{tr}(A_0 - A_1) \wedge F_t^{m-1}.$$

   For an elementary proof of this statement, see Harvey's anomaly lectures.

3. The total Chern class behaves nicely under the direct sum operation on bundles:

$$c(V \oplus V') = c(V) \wedge c(V'). \tag{3.111}$$

   So for example, $c_2(V \oplus V') = c_2(V) + c_2(V') + c_1(V) \wedge c_1(V')$. Notice that this relation (3.111) implies that the Chern classes are *stable* under the addition of trivial bundles $c\,(V \oplus \text{trivial}\,) = c(V)$.

For real bundles, F is a real antisymmetric matrix, and it is better to define instead

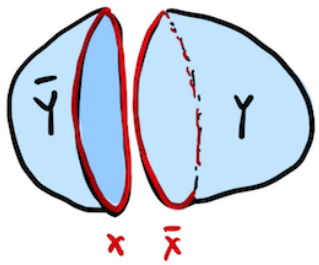
$$p(F) = \det\left(1 - \frac{F}{2\pi}\right) = 1 + p_1 + p_2 + \cdots$$

where p_j is a $4j$ form. p is for Pontryagin.

3.9 Classification of SPTs, part two

[Kapustin, Freed, Xiong. See this last paper for more references.] Maybe we should expect this property (3.103) for the effective action of any SPT (of finite order) – it should be a bordism invariant. Note that this is a stronger condition than just topological invariance.

So there should be an association between SPTs and $\mathsf{U}(1)$-valued bordism invariants, satisfying the cutting and gluing rules like (3.102), and unitarity $\langle \Phi_{X=\partial Y} | \Phi_{X=\partial Y} \rangle = Z_{Y \cup \bar{Y}} \overset{!}{\geq} 0$ (where $|\Phi_{X=\partial Y}\rangle$ is obtained from the path integral on Y as a functional of the boundary conditions on $X = \partial Y$, and the norm is the path integral on the space obtained by gluing along the opposite boundaries).



Bordism classes form an abelian group under disjoint union, $[Y_1] + [Y_2] = [Y_1 \amalg Y_2]$. This is true whatever additional structure we wish to include, such as a G-bundle or a spin structure, or a map to some space W. For example,

$$\Omega_D^{\text{spin}}(W) \equiv \text{equivalence classes of } D\text{-dimensional spin manifolds with a map to } W$$

with the equivalence relation $Y_1 \equiv Y_2$ if $\exists Z$ with $\partial Z = Y_1 \amalg \bar{Y}_2$ consistent with all the structure.

Gauge theory as classifying space sigma model. Now why would we want to care about maps to some space W? One reason is if our theory were a non-linear sigma model with target space W. This perspective leads to an improved topological understanding of WZW terms. Here is a second reason: we saw above that for any space Y and group G, there is a space BG, the classifying space, with the property that

$$G\text{-bundles on } Y/\sim \quad \leftrightarrow \quad [Y, BG].$$

So one way to include the data of a G-bundle in our bordism groups is to specify the bundle by a map to the classifying space $W = BG$. (Note that we are actually asking for bundles with connection, but the space of connections is contractible so does not contribute to this homotopy classification.) So the eta invariant (or more generally an SPT partition function) is a group homomorphism

$$e^{2\pi \mathbf{i} \eta} : \Omega_{D+1}^{\text{Spin}}(BG) \to \mathsf{U}(1). \tag{3.112}$$

If the bordism group Ω vanishes, then there can be no anomaly, and the bordism group classifies possible anomalies, and hence possible SPTs.

Bordism invariants and SPT effective actions. Consider a bosonic SPT α with respect to some symmetry G. The partition function of α on a closed oriented manifold Y with background fields for G is $e^{2\pi \mathbf{i} S_Y^\alpha[A]}$. Some properties of this functional follow from our assumptions:

- Since the phase α is gapped, S_Y^α is a local functional of the background gauge field A and the metric, and it's gauge invariant if Y is closed. Terms whose coefficient we can vary continuously while preserving these properties (like a Maxwell term) don't label the phase.
- Since α has a unique groundstate on any closed spatial slice, the amplitude $e^{2\pi \mathbf{i} S_Y^\alpha[A]}$ must be a phase[62].
- $S_{\bar{Y}}^\alpha[A] = -S_Y^\alpha[A]$ by the CPT theorem.
- $S_{Y_1 \amalg Y_2} = S_{Y_1}^\alpha + S_{Y_2}^\alpha$.
- $S_Y^{\bar{\alpha}}[A] = S_{\bar{Y}}^\alpha[A]$ where $\bar{\alpha}$ is the inverse SPT phase.

What terms can appear in $S_Y^\alpha[A] = \int_Y \sum_i \mathcal{L}_i^\alpha$? If the variation of a term in S with respect to $g_{\mu\nu}$ vanishes, then that term is said to be topological. There can certainly be non-topological terms, like a Maxwell term. A less innocuous non-topological term is a gravitational Chern-Simons term. Such terms occur in $D = 4n-1$; in $D = 3$, such a term produces a thermal Hall response κ_{xy}. In the SPT group law, any two SPTs with the same κ_{xy} differ by an SPT with $\kappa_{xy} = 0$. So let's first classify SPTs with $\kappa_{xy} = 0$.

So let's focus on the topological terms in $S_Y^\alpha[A]$. How can the effective action depend on the topology of Y and the G-bundle on it?

- **No symmetry.** First, suppose the protecting symmetry is nothing, $G = \{e\}$. The fact that there is no time-reversal symmetry means that Y must be oriented. S^α then depends on the topology of Y through local integrals. This must be a sum of integrals of products of characteristic classes of TY. These are Pontryagin classes (in degrees that are multiples of 4) and Stiefel-Whitney classes (in all degrees, but just mod two). Such sums of integrals of their products are called Pontryagin and Stiefel-Whitney numbers of Y. These numbers depend only on

[62] To see this, slice an arbitrary D-dimensional Y along some closed $D-1$ manifold Σ into $Y = Y_1 \cup_\Sigma Y_2$, so we can write $Z_Y = \langle Y_1 | Y_2 \rangle$, in terms of some normalized states $|Y_{1,2}\rangle$. Since $\dim \mathcal{H}_\Sigma = 1$, we have, using the magic of 1d linear algebra,

$$|Z_Y|^2 = \langle Y_1|Y_2\rangle \langle Y_2|Y_1\rangle = \langle Y_1|Y_1\rangle \langle Y_2|Y_2\rangle = 1.$$

the bordism class of Y. In fact, Thom proved that they also determine the bordism class of Y.

So in this case the effective action is a group homomorphism

$$e^{2\pi \mathbf{i} S_Y^\alpha} : \Omega_{D,\mathsf{SO}}(\mathrm{pt}) \to \mathsf{U}(1).$$

The SO subscript indicates that the structure group of Y is $\mathsf{SO}(D)$. The fact that S_Y is additive under disjoint union implies that it is a group homomorphism. The group of homomorphisms from Ω_D to $\mathsf{U}(1)$ can be called the *cobordism* group

$$\Omega^D_{\mathsf{SO}}(\mathrm{pt}, \mathsf{U}(1)) \equiv \mathrm{Hom}(\Omega_{D,\mathsf{SO}}(\mathrm{pt}) \to \mathsf{U}(1)). \tag{3.113}$$

But the abelian group $\Omega_{D,\mathsf{SO}}(\mathrm{pt})$ can include a free part. The associated terms are like theta terms with continuously-variable coefficient that don't label the phase. So we expect bosonic SPT phases with no symmetry to correspond to

$$\mathrm{bSPT}^D_{G=\{e\}} = \Omega^D_{\mathsf{SO}}(\mathrm{pt}, \mathsf{U}(1))/\Omega^D_{\mathsf{SO}}(\mathrm{pt}, \mathbb{R}).$$

These are the SPTs missing from the SRE subgroup.

[End of Lecture 19]

- **Time-reversal symmetric phases.** If $G = \mathbb{Z}_2^{\mathcal{T}}$, then Y can be unoriented. This also means that $\alpha = \bar{\alpha}$, and $2S_Y^\alpha[A] = 0 \bmod 1$, that is, $S_Y^\alpha[A] = 0, \frac{1}{2} \bmod 1$, so all phases are order two. The bordism invariants are just the Stiefel-Whitney classes, also all mod two, so there is no free part. So we expect

$$\mathrm{bSPT}^D_{G=\mathbb{Z}_2^{\mathcal{T}}} = \Omega^D_{\mathsf{O}}(\mathrm{pt}, \mathsf{U}(1)).$$

- **General unitary G.** The background gauge field specifies a map $\gamma : Y \to BG$, and so we can demand that our cobordisms include an extension of this map:

$$\mathrm{bSPT}^D_G = \Omega^D_{\mathsf{SO}}(BG, \mathsf{U}(1))/\Omega^D_{\mathsf{SO}}(BG, \mathbb{R}).$$

- **General symmetry.** If the symmetry includes time-reversing elements, we can specify them by a homomorphism $\rho : G \to \mathbb{Z}_2$, with $\ker \rho$ is the unitary part of the group. ρ induces a map $BG \to B\mathbb{Z}_2$, and hence a pullback Z of the universal $\mathbb{Z}_2$-bundle to BG. A background gauge field configuration γ for G must have the property that the pullback $\gamma^\star(Z)$ is the orientation bundle of Y.

Let's think about the case $G = \mathbb{Z}_2^{\mathcal{T}}$ in more detail. The formal sum of unoriented bordism groups over dimension is a nice graded algebra of polynomials $\Omega_{\bullet \mathsf{O}}(\mathrm{pt}) =$

$\mathbb{Z}_2[\{x_j\}]$ with generators x_j of degree $j = 2, 4, 5, 6, 8...$ (that is, not $2^i - 1$). This is a result of Thom that I will not prove. This means that the groups are:

D	1	2	3	4	5
$\Omega_{D,\mathsf{O}}(\mathrm{pt})$	0	$\mathbb{Z}_2$	0	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2$

We can associate them with particular combinations of Stiefel-Whitney classes. We've seen the associated phases already in this class.

- The result in $D = 1 + 1$ comes from the invariant $\int_Y w_1^2 = \int_Y w_2 =$ the euler character of Y mod two. This means that $Z_Y = e^{2\pi \mathbf{i}\left(\frac{1}{2}\int_Y w_2(Y)\right)} = e^{\mathbf{i}\pi\chi(Y)} = \pm 1$ according to whether or not Y is orientable. This is the Haldane chain.

- In $D = 2 + 1$ the SW classes satisfy the relations $w_1 w_2 = w_3 = w_1^3 = 0$ for all closed 3-manifolds. This is consistent with the statement that every such 3-manifold is null-bordant – the boundary of some 4-manifold.

- In $D = 3 + 1$, the relations $w_3 w_1 = w_2 w_1^2 = 0$ and $w_4 + w_2^2 + w_1^4 = 0$ leave two independent invariants that we can take to be $\frac{1}{2}w_1^4$ and w_2^2. This means three nontrivial bosonic SPTs for time-reversal symmetry (two nontrivial generating phases and their stacking). These are just the phases we saw from the coupled-layer construction. The one corresponding to w_2^2 is the one we discussed more explicitly, whose surface is the all-fermion toric code (this phase is called $efmf$ by Wang and Senthil). This makes sense because if there are neutral fermions, we can only put the system on a spin manifold, which has $w_2 = 0$: hence the anomaly disappears in the presence of neutral fermions. The other one, corresponding to $\frac{1}{2}w_1^4$, is the one visible in group cohomology, and has a surface which is the toric code where e and m are both Kramers' doublets (called $e\mathcal{T}m\mathcal{T}$ by Wang and Senthil).

- Finally in $D = 4{+}1$, the nontrivial invariant is $\frac{1}{2}\int w_2 w_3$, whereas $H^4(B\mathbb{Z}_2, \mathsf{U}(1)) = 0$. This phase survives the breaking of time-reversal symmetry. It has a surface given by all-fermion electrodynamics, and an effective description in terms of 2-form CS gauge theory with action $S[B, C] = \int_Y \frac{B \wedge dC}{2\pi}$.

Generalized cohomology and loop spectrum. These bordism groups are an example of a *generalized cohomology theory*. There is a set of seven axioms due to Eilenberg and Steenrod defining what is a cohomology theory. It is a contravariant functor from topological spaces to abelian groups, just like SPT_D^G. The usual de Rham and Cech theories satisfy all of these axioms. A generalized cohomology theory fails to satisfy the last axiom, which demands that $H^\bullet(\text{point})$ is trivial. This is where the non-SRE states come in.

A *spectrum* (or loop spectrum or Ω-spectrum) is a machine for producing a generalized cohomology theory. It is a family of topological spaces K_d, labelled by $d \in \mathbb{Z}$ with property that $K_d \simeq \Omega K_{d+1}$, where ΩX means the loop space of X, and the relation is homotopy equivalence. Given such a spectrum $K_\bullet$, the generalized cohomology groups associated with $K_\bullet$ are defined to be

$$H^d(X, K) \equiv [X, K_d]$$

homotopy classes of maps into K_d[63].

To make contact with something more familiar, an example of an Ω-spectrum is $K^A_\bullet$, with $K^A_d = K(A, d)$, the Eilenberg-Maclane spaces for an abelian group A, defined by $\pi_q(K(A,d)) = \begin{cases} A, & d = q \\ 0, & \text{else} \end{cases}$. These are to homotopy what spheres are to (reduced) cohomology. (Note that BG is a $K(G, 1)$ by the long exact sequence on homotopy from (3.108).) And in fact $H^d(X, K^A) = H^d(X, A)$, the ordinary cohomology groups with coefficients in A.

Perhaps I should mention that $H^d(\bullet, K)$ is indeed a contravariant functor, in the sense that given $f : X \to Y$, then we can define a pullback operation $f^\star : H^d(Y, K) \to H^d(X, K)$ by pre-composing maps: $[c] \to [c \circ f]$. And there is indeed a group structure on $H^d(X, K) \ni [c_{1,2}]$. $c_{1,2} : X \to \Omega K_{d+1}$, where the group law is $[c_1] + [c_2] = [c_1 + c_2]$, where $c_1 + c_2$ means concatenate loops for each point in X. This addition is abelian since $H^d(X, K) = [X, K_d] = [X, \Omega K_{d+1}] = [X, \Omega^2 K_{d+2}]$, so the order of concatenation doesn't matter (the same reason that $\pi_{q>1}(X)$ is abelian).

The rough idea for making contact with physics is that K_d is the space of G SPTs in d dimensions, and ΩK_d is the space of maps from the unit cell to K_d. The whole thing is a big generalization of the Thouless pump.

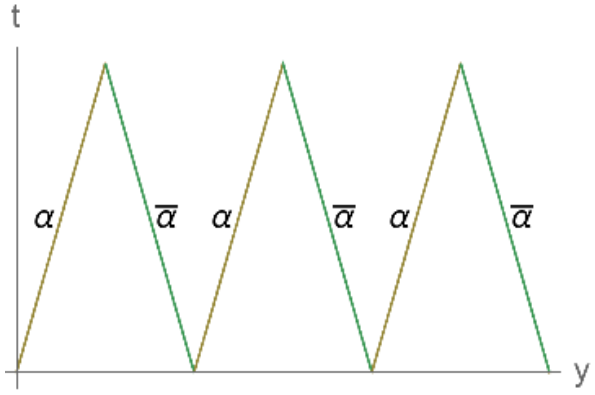
Looping and de-looping. It remains to construct a loop spectrum from SPTs. This requires a relation between SPTs in d dimensions and loops of SPTs in $d + 1$ dimensions. Let $F_d \equiv \{$ d-dimensional invertible states$\}$. Let's define a map called 'pumping' that takes

$$f : \begin{array}{ccc} F_d & \to & \Omega F_{d+1} \\ \alpha & \mapsto & f(\alpha)_t \end{array} \tag{3.114}$$

with $f(\alpha)_0 = f(\alpha)_1$. It is best to just look at the picture, which shows a loop in the

[63] When speaking about homotopy classes of maps there is always the annoying question of whether the maps fix a base point. Here if we demand that our maps fix a base point, we get $\tilde{H}^d(X, K)$, the reduced cohomology. Recall that the idea of reduced cohomology is to mod out by the cohomology of a point.

space of invertible states in $d+1$ dimensions:



Here y is the extra dimension.

And we can define a map going the other way

$$g : \begin{array}{ccc} \Omega F_{d+1} & \to & F_d \\ \beta(t) & \mapsto & g(\beta) \end{array} \tag{3.115}$$

by cutting the system open at some particular point in x. $g(b)$ is the d-dimensional SPT that is pumped past the cut to the right.

Claims: f and g are homotopy inverses. $g \circ f = 1$ by construction, and $f \circ g \simeq 1$.

Concatenation of loops $\beta_1(t), \beta_2(t)$ in ΩF_{d+1} maps to stacking of d-dimensional SPTs, $g(\beta_{1,2})$.

So the current understanding is that

$$\mathrm{SPT}^d_G = H^d(BG, K) \tag{3.116}$$

for some Ω-spectrum K. As we discussed above, the group cohomology $H^\bullet(G, \mathsf{U}(1))$ is related to the *ordinary* cohomology of BG. So just as the homotopy groups are different from the cohomology groups of a space, this improved classification will both add and remove SPTs relative to the states that were explicitly constructed in 3.4.

If we take K to be the *Thom spectrum* (which I'm not going to say what it is), we get the cobordism classification motivated above. A nice thing about this perspective is that as our understanding of the space of invertible states F_d grows, we can improve our classification.