# Lattice Gauge Theory

March 4, 2025

## Contents

# 1 CZX

Each site: four spins, suppose in state $|abcd\rangle_s$, where $a, b, c, d$ are either spin $|0\rangle$ or spin $|1\rangle$, and we introduce the notation $\bar{a}$ to denote the negative of $a$, and the subscript $s$ to remind ourselves that the spins are in a single site.

Now define a unitary transformation for any on-site state $|abcd\rangle_s$: define $U_X = \sigma_1^x\sigma_2^x\sigma_3^x\sigma_4^x$ which flips all the four spins of the site, i.e.

$$U_X|abcd\rangle_s = |\bar{a}\bar{b}\bar{c}\bar{d}\rangle_s.$$

Then we efine $U_{CZ} = z_{12}z_{23}z_{34}z_{41}$ where each $z = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 01| - |11\rangle\langle 11|$. We have

$$U_{CZ}|abcd\rangle_s = z_{12}z_{23}z_{34}(-1)^{ad}|abcd\rangle_s = ... = (-1)^{ab+cd+bc+ad}|abcd\rangle_s = (-1)^{(a+c)(b+d)}|abcd\rangle_s,$$

note since the effect of each $z_{ij}$ is to attach a $Z_2$ phase, all the four $z_{ij}$ commute and $U_{CZ}$ does not necessarily have to be written in the order of $z_{12}z_{23}z_{34}z_{41}$. Therefore we see that

$$U_{CZX}|abcd\rangle_s = U_X U_{CZ}|abcd\rangle_s = (-1)^{(a+c)(b+d)}|\bar{a}\bar{b}\bar{c}\bar{d}\rangle_s.$$

We see that $U_X^2 = U_{CZ}^2 = U_{CZ}^2 = 1$, $U_X$ and $U_Z$ commute. We also see that in order for a state $|\Psi\rangle = \sum_{a,b,c,d} c_{abcd}|abcd\rangle_s$ to have this symmetry, $|abcd\rangle_S$ and $|\bar{a}\bar{b}\bar{c}\bar{d}\rangle_s$ must appear at the same time, with their coefficient opposite if precisely $a, c$ opposite spin and at the same time $b, d$ opposite spin, and their coefficient the same otherwise.

Each plaquette: the $vh$ corner contains the $\bar{v}\bar{h}$ of the spin in that site, where $v$ = up or down, $h$ = left or right, and again we define $\bar{v}$ or $\bar{h}$ to be the negativ of $v$ or $h$. The Hamiltonian for this plaquette $p$ does not only act on this plaquette, but also on the four neighboring plaquettes:

$$H_p = -X_4 P_2^u P_2^d P_2^l P_2^r, \tag{1}$$

where $X_4 = |0000\rangle_{pp}\langle 1111| + |1111\rangle_{pp}\langle 0000|$ acts on the four spins of the plaquette $p$. (We assume the spins are in the order of $ur, ul, dl, dr$.) The projections $P_2 = |00\rangle\langle 00| + |11\rangle\langle 11|$ act on the two spins of the neighboring half plaquette. Therefore, each Hamiltonian term $H_p$ acts essentially on four sites (i.e. 16 spins) surrounding the plaquette $p$. Note again that, as the case for $zzzz$ about, here the four projections $PPPP$ can also be reordered since they commute with each other. We know write the most general form of the state in on such a four site:

$$|\psi\rangle = |a_1 b_1 \underline{c_1} d_1\rangle_{s_1} \otimes |a_2 b_2 c_2 \underline{d_2}\rangle_{s_2} \otimes |\underline{a_3} b_3 c_3 d_3\rangle_{s_3} \otimes |a_4 \underline{b_4} c_4 d_4\rangle_{s_4},$$

where by our convention, $1, 2, 3, 4$ denote the four sites around the plaquette $p$ in the order of $ur, ul, dl, dr$, and $a, b, c, d$ denote the four spins in a site in the order of $ur, ul, dl, dr$. Then, the Hamiltonian action is

$$\begin{aligned}
H_p|\psi\rangle &= -X_4 P_2^u P_2^d P_2^l P_2^r |a_1 b_1 \underline{c_1} d_1\rangle \otimes |a_2 b_2 c_2 \underline{d_2}\rangle \otimes |\underline{a_3} b_3 c_3 d_3\rangle \otimes |a_4 \underline{b_4} c_4 d_4\rangle \\
&= -X_4 P_2^u P_2^d P_2^l \delta_{d_1 a_4} |a_1 b_1 \underline{c_1} d_1\rangle \otimes |a_2 b_2 c_2 \underline{d_2}\rangle \otimes |\underline{a_3} b_3 c_3 d_3\rangle \otimes |a_4 \underline{b_4} c_4 d_4\rangle \\
&= \cdots \\
&= -X_4 \delta_{d_3 c_4}\delta_{c_2 b_3}\delta_{b_1 a_2}\delta_{d_1 a_4} |a_1 b_1 \underline{c_1} d_1\rangle \otimes |a_2 b_2 c_2 \underline{d_2}\rangle \otimes |\underline{a_3} b_3 c_3 d_3\rangle \otimes |a_4 \underline{b_4} c_4 d_4\rangle \\
&= -\delta_{c_1=d_2=a_3=b_4}\delta_{d_3 c_4}\delta_{c_2 b_3}\delta_{b_1 a_2}\delta_{d_1 a_4} |a_1 b_1 \bar{c}_1 d_1\rangle \otimes |a_2 b_2 c_2 \bar{d}_2\rangle \otimes |\bar{a}_3 b_3 c_3 d_3\rangle \otimes |a_4 \bar{b}_4 c_4 d_4\rangle,
\end{aligned} \tag{2}$$

we see that in order to make the four-site state $|\psi\rangle$ to be the ground state of the Hamiltonian term $H_p$, we must have $c_1 = d_2 = a_3 = b_4 = 0$ and 1 showing up on the same (positive) weight and the two spins in each of the four neighboring half plaquettes have the same spins. Now define the total Hamiltonian to be

$$H = \sum_p H_p \tag{3}$$

to be the sum of all plaquette Hamiltonian terms. first thing to note is that different Hamiltonian terms commute with each other: this is easy to see since the only overlapping part of the two such terms is $X_4$ on one plaquette $p$ and the $P_2$ acting on one bond of $p$ coming from a neighboring plaquette $p'$; and they commute since any of them acting first on a state selects the bond to have the same spin, and on such a Hilbert subspace they commute. Therefore the ground state of each $H_p$ must also be the ground state of $H$, i.e.

$$|GS\rangle = \otimes_{p\in P}(|0000\rangle_p + |1111\rangle_p),$$

where $P$ are the set of the plaquettes of the model. Note we are overlooking subtleties about boundaries. Also note that the projection $P_2$' are ineffective when getting the ground state: what this means is that $H = \sum_p H_p$ and $H = \sum_p \widetilde{H}_p$, where $\widetilde{H}_p = -X_4$, will give the same ground state.

Now we have to prove that this (many-body) ground state respects the onsite symmetries $U_X$ and $U_{CZ}$, respectively. Note here by symmetry we mean the symmetry action simultaneously on all the sites, or rigorously $\prod U_X$, or $\prod U_{CZ}$. But instead we will just use the notation $U_X$ and $U_{CZ}$. It is easy to see that the ground state respects $U_X$, since $U_X = \prod \sigma_1$ is simply the flip operator for all the spins. For any tensor basis $|abc....\rangle$ it simply gives $|\bar{a}\bar{b}\bar{c}...\rangle$, and therefore for any states it just flips all the spins, therefore

$$U_X\colon |GS\rangle = \otimes_{p\in P}(|0000\rangle_p + |1111\rangle_p) \to \otimes_{p\in P}(|1111\rangle_p + |0000\rangle_p) = |GS\rangle.$$

Then for $U_{CZ}$: for any $|abc...\rangle$, $U_{CZ} = \prod_s U^s_{CZ}$, and each $U^s_{CZ}$ gives a factor of $(-1)^{l_1+l_2+l_3+l_4}$, where as shown before $l_i$ are the product of the value of the state on the bond $i$ of the site $s$. Look at bond $i$ which consists of spin $ab$ in the site $s$: $a$ belongs to plaquette $p$ and $b$ belongs to plaquette $p'$; and we find $a'$ in $p$ and $b'$ in $p'$, where the bond $aa'$ and $bb'$ are parallel. Then, it is easy to see that the ground state satisfies the condition that $b = b'$ and $a = a'$, which is just a natural property of the ground state, therefore when $U_{CZ}$ acting on $ab$ we get a phase $(-1)^{ab}$, and when $U_{CZ}$ acting on $a'b'$ we get another phase $(-1)^{a'b'} = (-1)^{ab}$, which gives unity.

Looking at a star $s$. Therefore four plaquettes involved, and the relevant part of the wavefunction is

$$\otimes_{p\in s}(|0000\rangle_p + |1111\rangle_p),$$

where $p \in s$ means the four nearest plaquettes with $s$ a corner. Or written more explicitly,

$$(|000\underline{0}\rangle_{ur} + |111\underline{1}\rangle_{ur}) \otimes (|000\underline{0}\rangle_{ul} + |111\underline{1}\rangle_{ul}) \otimes (|\underline{0}000\rangle_{dl} + |\underline{1}111\rangle_{dl}) \otimes (|\underline{0}000\rangle_{dr} + |\underline{1}111\rangle_{dr}),$$

where $ur, ul, dl, dr$ are the locations of the respective plaquette relative to the site $s$, and the underlined spin belongs to site $s$.

At this point we still have not seen the importance of the projections $PPPP$: we have just seen that they are ineffective in selection the ground state and that the ground state has the symmetry $U_X$ and $U_{CZ}$. In proving that $U_{CZ}$ is a symmetry we used the ground state property that $a = a'$ and $b = b'$; in fact, any state satisfying this property will have $U_{CZ}$ symmetry, and this property is exactly gauranteed by the $PPPP$ part. Therefore, the importance of the $PPPP$ part is to let the Hamiltonian $H$ to have the symmetrey $U_X$ and $U_{CZ}$, not only the ground state; note that $H = \sum_p \widetilde{H}_p$ will not have this symmetry.

Now we prove $H = \sum_p H_p$ is $U_X$ symmetric and $U_{CZ}$ symmetric. Therefore we prove the $[H, \prod U_X] = 0$ and $[H, \prod U_{CZ}] = 0$, i.e we prove the $[H_p, \prod U_X] = [H_p, \prod U_{CZ}] = 0$ for any plaquette $p$. As analyzed before, $H_p$ acts on the 16 spins on the four sites at the corner of the plaquette $p$. The most general state is the superposition of

$$|\psi\rangle = |a_1b_1\underline{c_1}d_1\rangle_{s_1} \otimes |a_2b_2c_2\underline{d_2}\rangle_{s_2} \otimes |\underline{a_3}b_3c_3d_3\rangle_{s_3} \otimes |a_4\underline{b_4}c_4d_4\rangle_{s_4},$$

and we just need to prove $H_p \prod U_{CZ}|\psi\rangle = \prod U_{CZ} H_p|\psi\rangle$ and $H_p \prod U_X|\psi\rangle = \prod U_X H_p|\psi\rangle$. Note we have, using (2),

$$\prod U_{CZ} H_p|\psi\rangle = -(-1)^{\phi}\delta_{c_1=d_2=a_3=b_4}\delta_{d_3c_4}\delta_{c_2b_3}\delta_{b_1a_2}\delta_{d_1a_4}|a_1b_1\bar{c}_1d_1\rangle \otimes |a_2b_2c_2\bar{d}_2\rangle \otimes |\bar{a}_3b_3c_3d_3\rangle \otimes |a_4\bar{b}_4c_4d_4\rangle,$$

where $\phi = (a_1 + \bar{c}_1)(b_1 + d_1) + (a_2 + c_2)(b_2 + \bar{d}_2) + (\bar{a}_3 + c_3)(b_3 + d_3) + (a_4 + c_4)(\bar{b}_4 + d_4)$, and

$$H_p \prod U_{CZ}|\psi\rangle = -(-1)^{\widetilde{\phi}}\delta_{c_1=d_2=a_3=b_4}\delta_{d_3c_4}\delta_{c_2b_3}\delta_{b_1a_2}\delta_{d_1a_4}|a_1b_1\bar{c}_1d_1\rangle \otimes |a_2b_2c_2\bar{d}_2\rangle \otimes |\bar{a}_3b_3c_3d_3\rangle \otimes |a_4\bar{b}_4c_4d_4\rangle,$$

where $\phi = (a_1 + c_1)(b_1 + d_1) + (a_2 + c_2)(b_2 + d_2) + (a_3 + c_3)(b_3 + d_3) + (a_4 + c_4)(b_4 + d_4)$, we have $\phi - \widetilde{\phi} = c_1(b_1 + d_1 + a_2 + c_2 + b_3 + d_3 + a_4 + c_4)$ in the $Z_2$ sense, using the fact $c_1 = d_2 = a_3 = b_4$, and using $d_3 = c_4$, $b_1 = a_2$, $c_2 = b_3$, $d_1 = a_4$ this difference vanishes in the $Z_2$ sense. We see that $H_p$ commutes with $U_{CZ}$. Then for $U_X$:

$$\prod U_X H_p|\psi\rangle = -\delta_{c_1=d_2=a_3=b_4}\delta_{d_3c_4}\delta_{c_2b_3}\delta_{b_1a_2}\delta_{d_1a_4}|\bar{a}_1\bar{b}_1c_1\bar{d}_1\rangle \otimes |\bar{a}_2\bar{b}_2\bar{c}_2d_2\rangle \otimes |a_3\bar{b}_3\bar{c}_3\bar{d}_3\rangle \otimes |\bar{a}_4b_4\bar{c}_4\bar{d}_4\rangle,$$

$$H_p \prod U_X|\psi\rangle = -\delta_{\bar{c}_1=\bar{d}_2=\bar{a}_3=\bar{b}_4}\delta_{\bar{d}_3\bar{c}_4}\delta_{\bar{c}_2\bar{b}_3}\delta_{\bar{b}_1\bar{a}_2}\delta_{\bar{d}_1\bar{a}_4}|\bar{a}_1\bar{b}_1c_1\bar{d}_1\rangle \otimes |\bar{a}_2\bar{b}_2\bar{c}_2d_2\rangle \otimes |a_3\bar{b}_3\bar{c}_3\bar{d}_3\rangle \otimes |\bar{a}_4b_4\bar{c}_4\bar{d}_4\rangle,$$

which is obviously the same. Therefore $H_p$ commutes with $U_X$.

To summarize, we define symmetry $X$ and $CZ$ has siultaneously acting $U_X$ or $U_{CZ}$ on all sites. Then we can prove that each plaquette $H_p$ has this symmetry. Since the total Hamiltonian consists of mutually commuting terms, $H$ also has this symmetry.

The density matrix of the ground state can be written as the tensor product of the density matrix of each plaquette $\rho_4 = (|0000\rangle + |1111\rangle)(\langle 0000| + \langle 1111|)$. It is easy to trace out half of the degrees of freedom from $\rho_4$ to obtain the reduced density matrix on a bond, $\rho_2 = \mathrm{Tr}_2\rho_4 = |00\rangle\langle 00| + |11\rangle\langle 11|$, and the reduced density matrix on one site, $\rho_1 = \mathrm{Tr}_1\rho_2 = |0\rangle\langle 0| + |1\rangle\langle 1|$. Note this says that the two spins on a half plaquette spans a two-dimensional subspace, and these two spins can be in either $|00\rangle$ or $|11\rangle$ rather than just one state. One can then define an effective spin that live on the middle of the bond, $|\widetilde{0}\rangle = |00\rangle$ and $|\widetilde{1}\rangle = |11\rangle$. On these effective spins, the symmetry operation, $U_{CZX}$, acting on a string of half-plaquettes, can now be writte as $\widetilde{U}_{CZX} = \prod_{i=1}^N \widetilde{X}_i \prod_{i=1}^N \widetilde{CZ}_{i,i+1}$ where $N + 1 \equiv 1$ (periodic boundary condition), where $\widetilde{X}$ flips the effective spin and $\widetilde{CZ}$ acts on neighboring effective spins (on the two spins on the original site),

## 2 Statistical 1D and 2D Ising model

Befor talking about 2D, we look at the case of statistical 1D Ising model: The model is just

$$H = -J\sum_i \sigma_i^z\sigma_{i+1}^z - H\sum_i \sigma_i^z,$$

with partition function

$$Z = \sum_{\{\sigma_i^z\}=\pm 1} e^{-\beta H} = \sum_{\{\sigma_i^z\}=\pm 1} e^{K\sum_i \sigma_i^z\sigma_{i+1}^z + h\sigma_i^z},$$

where we have defined $K = \beta J$ and $h = \beta H$. The basic conclusion of this model is that it is magnetized only at $T = 0$ and is disordered at any finite temperature. This is easily understood using the energy-entropy argument, see the begining of the Section for Wegner's Ising gauge theory; one can also use transfer matrix to solve this case. This is nicely taught in Altland and Simons, Section 8.1 and we recap here. We write

$$Z = \sum_{\{\sigma_i^z\}=\pm 1} e^{K\sum_i \sigma_i^z\sigma_{i+1}^z + \frac{h}{2}(\sigma_i^z+\sigma_{i+1}^z)} = \left(\prod_i \sum_{\sigma_i^z=\pm 1}\right) e^{\sum_i L(\sigma_i^z,\sigma_{i+1}^z)},$$

since the value of $\sigma_i^z$ will determine both $L(\sigma_{i-1}^z, \sigma_i^z)$ and $L(\sigma_i^z, \sigma_{i+1}^z)$, we see that this is essentially a matrix multiplication: for $L(\sigma_i^z, \sigma_{i+1}^z)$, the two values of $\sigma_i^z$ labels the row of $L$ and the two values of $\sigma_{i+1}^z$ labels the column of $L$, which gives a $2\times 2$ matrix $L$; and $e^{L(\sigma_i^z,\sigma_{i+1}^z)}$ is just element-wise exponentiate $L$, which gives

$$T = \begin{pmatrix} e^{K+h} & e^{-K} \\ e^{-K} & e^{K-h} \end{pmatrix},$$

therefore

$$Z = \mathrm{Tr}\prod_i T_i = \mathrm{Tr}T^N,$$

where we have assumed periodic boundary condition for the chain. Now, since we have the form of $T$, we can calculate everything: $Z$, $F = -\beta^{-1}\ln Z$, or magnetization $M = \partial_h Z$. The spontaneous magnetization can be obtained by calculating $M$ then sending $h$ to zero, which always give

$$M_{\text{spontaneous}} = \lim_{h\to 0} M = \lim_{h\to 0}\partial_h Z = \lim_{h\to 0}\frac{h}{\sqrt{\sinh^2 h + e^{-4K}}} = 0,$$

confirming our previous claim that there is no spontaneous magnetization in te statistical 1D Ising model. This is in constrast to the statistical 2D Ising model, which does have a finite temperature phase transition as presented below. Note we can also derive the 0+1D quantum mechanical model in the so called $\tau$-continuum limit, where $T = e^{-\delta\tau\hat{H}} = 1 - \delta\tau\hat{H}$. The quantum mechanical model simple reads (up to constant)

$$\hat{H} = -2K\sigma^x + h\sigma^z,$$

this can be easily tested, or see Fradkin and Susskind's paper `https://journals.aps.org/prd/pdf/10.1103/PhysRevD.17.2637`.

Before going to the 2D case, let us summarize the lessons here:

- For classical statistical systems with discrete symmetries, the discrete symmetry cannot be spontaneously broken in (spatial dimension) $d \leq 1$ (Stanley's result);
- For classical statistical systems with continuous symmetries, the continuous symmetry cannot be spontaneously broken in (spatial dimension) $d \leq 2$ (Mermin-Wagner theorem).

The (classical) statistical 2D Ising model: again the derivation below can be found in `https://journals.aps.org/prd/pdf/10.1103/PhysRevD.17.2637`. partition function is

$$Z = \sum_{\{\sigma_n^z\}=\pm 1} e^{-S},$$

where

$$S = -\sum_n \beta_\tau \sigma_{n+\tau}^z\sigma_n^z + \beta\sigma_{n+x}^z\sigma_n^z,$$

note here the model is allowed to be anisotropic. Here we have defined the two direction as $\tau$ and $x$ for later use (transfer matrix). We have also put the temperature $\beta$ into $S$, where we defined the exchange constant $J = 1$ for the $x$ direction and $J_\tau$ for the $\tau$ direction as in $\beta_\tau = \beta J_\tau$. $n$ runs over all sites of the 2D lattice. The model can be written as, up to constant,

$$S = \frac{1}{2}\beta_\tau \sum_n (\sigma^z_{n+\tau} - \sigma^z_n)^2 - \beta \sum_n \sigma^z_{n+x}\sigma^z_n, \tag{4}$$

note that to use transfer matrix, we must have a matrix, i.e. we should have two set of variables labeling the row and column of the transfer matrix. We should always keep in mind that the essense is

$$\langle x'|\hat{T}|x\rangle = T(x', x),$$

where $x$ and $x'$ denote the two sets of variables. Here there are two procedures: first, $T(x', x)$ is the target form one (needs some cleverness) tries to arrive when given a classical partition function $Z$:

$$Z = \int \prod_i [dx_i T(x_{i+1,i})],$$

and if such $T(x', x)$ is found, one (again needs some cleverness) tries to find a quantum operator $\hat{T}$ such that the matrix element $\langle x'|\hat{T}|x\rangle$ is precisely $T_{x',x}$. If both steps are achieved, one is then allowed to write

$$Z = \mathrm{Tr}\hat{T}^N$$

assuming periodic boundary condition in the classical system. Now we try to achieve these two steps for the action (4). (notice this is first done, as referenced by Kogut, by Fradkin and Susskind in `https://journals.aps.org/prd/pdf/10.1103/PhysRevD.17.2637`; there the level of detailedness is more or less the same as Kogut. Notice this paper talks both 2D Ising and 4D gauge systems) First, we can use $s^z$ for odd rows and $\sigma^z$ for even rows, i.e. we write

$$S = \sum_{n_r} \sum_{m\in n_r} \frac{1}{2}\beta_\tau (s^z_m - \sigma^z_m)^2 - \frac{1}{2}\beta\sigma^z_{m+1}\sigma^z_m + s^z_{m+1}s^z_m,$$

where $n_r$ labels the rows, and for each row $n_r$, $m$ sums over the sites in this row. Note when $n_r$ is odd (even) $s^z_m$ is above (below) $\sigma^z_m$. Now we can define

$$S = \sum_{n_r} L_{n_r}(s, \sigma),$$

where

$$L_{n_r}(s,\sigma) = \sum_{m\in n_r} \frac{1}{2}\beta_\tau (s^z_m - \sigma^z_m)^2 - \frac{1}{2}\beta\sigma^z_{m+1}\sigma^z_m + s^z_{m+1}s^z_m,$$

and the partition function now becomes

$$Z = \left( \prod_{n_r/2} \sum_{\{s^z_m\}\in n_r/2} \sum_{\{\sigma^z_m\}\in n_r/2} \right) e^{-\frac{1}{\hbar}\sum_{n_r} L_{n_r}(s,\sigma)},$$

or this can be written as

$$Z = \prod_{n_r} \left( \sum_{\{s^z_m\}\in n_r/2} \sum_{\{\sigma^z_m\}\in n_r/2} e^{-\frac{1}{\hbar}L_{n_r}(s,\sigma)} \right), \tag{5}$$

where $n_r/2$ means we count rows by increment of 2, and for each two rows we look at all possibilities of the variables $\{s^z_m\}$ and $\{\sigma^z_m\}$ in these two rows. I admit the above formula is not written in very good notation because the set of values, say, $\{\sigma^z_m\}$ will determine both $L_{n_r}(s,\sigma)$ and the next row; but I assume this is understood. The $n_r/2$ should be understood as counting two rows every time. Suppose there are $M$ sites in a row. Note that the two sums gives all the possible configurations for $s$ and $m$ in a two-row, and for each such configuration, $e^{-\frac{1}{\hbar}L_{n_r}(s,\sigma)}$ will spit out a number. Note that if for the $(i, i+1)$ rows we have $(s^z, \sigma^z)$, then for the $(i+1, i)$ rows we have $(\sigma^z, s^z)$, therefore the part in the large parenthesis in (5) is precisely matrix product, i.e. we write

$$Z = \prod_{n_r} T_{n_r} = \mathrm{Tr}T^N, \tag{6}$$

where $N$ is the number of rows and we have assumed periodic boundary condition for the $\tau$ direction,

$$T_{n_r} = \sum_{\{s^z_m\}\in n_r/2} \sum_{\{\sigma^z_m\}\in n_r/2} e^{-\frac{1}{\hbar}L_{n_r}(s,\sigma)}.$$

Now, to work out the form of the $2^M \times 2^M$ matrix $T_{n_r}$, we only need to work out the $2^M \times 2^M$ matrix of $L_{n_r}$, and element-wise exponentiate $L_{n_r}$: $x \to e^{-\frac{1}{\hbar}x}$. The matrix $L_{n_r}$ actually has a simple structure: assume its row is labeled by $\sigma_m^z$ and column by $s_m^z$, then suppose we are looking at the entry $(\{\sigma_m^z\}, \{s_m^z\})$, where the $\{\sigma_m^z\}$, and $\{s_m^z\}$ pick a particular configuration, we will see that the first term of $L_{n_r}$ only depends on the difference of (say) the number of spin-ups between $\sigma$ and $s$. Suppose this difference is $n_{s\sigma}$, then the first term contributes $\frac{1}{2}\beta_\tau(2n_{s\sigma})^2 = 2\beta_\tau n_{s\sigma}^2$. Note this alreay completes the solution of $\hat{T}$. Let us write in a more explicit way:

$$\begin{aligned} \langle\{\sigma_m^z\}\hat{T}|\{s_m^z\}\rangle\Big|_{n_{s\sigma}=0} &= e^{\beta\sum_m \sigma_{m+1}^z\sigma_m^z}, \\ \langle\{\sigma_m^z\}\hat{T}|\{s_m^z\}\rangle\Big|_{n_{s\sigma}=n} &= e^{-2n\beta_\tau}e^{\frac{1}{2}\beta\sum_m[\sigma_{m+1}^z\sigma_m^z + s_{m+1}^z s_m^z]}, \quad n \geq 1, \end{aligned} \tag{7}$$

And this is already enough to give the desired transfer matrix form of the classical statistical partition function, (6).

Now, suppose we want to go one step further: we want to write the classical partition function $Z = \mathrm{Tr}\hat{T}^N$ as some quantum Hamiltonian $Z = \mathrm{Tr}e^{-\tau\hat{H}}$. Tis way, we must have

$$\hat{T} = e^{-\delta\tau\hat{H}} = 1_{2^M\times 2^M} - \delta\tau\hat{H},$$

where we defined $\delta\tau = \tau/N$. Since $\delta\tau$ is small we have expanded the exponential to be an identity matrix and a small part. Note again the above equation should be understood as a matrix equation. Now we simply need to compare $\hat{T} = 1 - \delta\tau\hat{H}$ with the form of matrix elements of $\hat{T}$ in Eq. (7): we see that to identify $\hat{T}$ with $1 - \delta\tau\hat{H}$ we must have: from the diagonal element, we must have $\beta \sim -\delta\tau$; then, from the $n = 1$ off-diagonal element we read that $e^{-2\beta_r} = \delta\tau$. [Note actually I have some doubt how to justify this argument; Kogut does not explain superclearly. I will just follow his reasoning here.] Note due to $n$ appear on $e^{-2n\beta_r}$, for $n \geq 2$ actually the term is so small so that we can actually ignore them. This way, $\hat{H}$ is only nonzero for $n = 0$ (i.e. diagonal) and $n = 1$. It is not hard then to show that $\hat{H}$ is of the form

$$\hat{H} = -\sum_m \hat{\sigma}_m^x + \lambda\hat{\sigma}_{m+1}^z\hat{\sigma}_m^z,$$

which is simply the $1 + 1$ quantum (transverse field) Ising model. Note, we have introduced the parameter $\lambda$. It is introduced in this way: note we just mentioned (with doubt) that $\delta\tau \sim \beta \sim e^{-2\beta_r}$, this actually says that, in order to get a quantum Hamlitonian (i.e. a smooth $\tau$-continuum theory, in the language of Kogut), we must have $\beta$ and $e^{-2\beta_r}$ be proportional to each other, i.e. $\beta = \lambda e^{-2\beta_r}$ for some $\lambda$. Kogut emphasize that this has the interpretation that this means that, in order to get the same quantum Hamiltonian (meaning $\lambda$ is kept fixed) from the classical model, the temporal coupling must grows large when we let the spatial coupling becomes weak, and it seems to be a general feature. But note that by varying $\lambda$ we are indeed able to cover all the $(\beta_\tau, \beta)$ choices for the 2D statistical Ising model. This is an example that a $d$-dimensional statistical model is equivalent to a $(d-1)+1$ quantum mechanical model: note in statistical model one never talks about temperal dimension so $d$ is pure spatial dimention. We see in our Ising model example that we have to separate one dimension in the statistical model and regard that direction as the temperal direction of the quantum mechanical model.

The Kramers-Wannier duality: we use the 1D quantum model: define $\mu_n^x = \sigma_{n+1}^z\sigma_n^z$, $\mu_n^z = \prod_{m\leq n}\sigma_m^x$.Notice that $\sigma$'s on different sites commute, due to the tensor structure, and only $\sigma^x$ and $\sigma^z$ on the same site anticommute. It easy to check that we have $\mu_n^x\mu_n^z = -\mu_n^z \otimes \mu_n^x$, and $\mu$ on different sites commute (easy to show that $\mu_n^x, \mu_m^x$ commute, and $\mu_n^z, \mu_m^z$ commute. Note we can think of $n$ in $\mu_n^x$ as $(n, n+1)$; $m$ in $\mu_m^z$ as $[1, ...., m]$; the only nontrivial case therefore is when $n+1 = m$: but this case the two $\sigma^z$'s anticommute with the two $\sigma^x$'s and we have $(-1)^2 = 1$). This way the Hamiltonian is just

$$H = -\sum_n \mu_n^z\mu_{n+1}^z - \lambda\mu_n^x,$$

i.e. we have

$$H(\sigma, \lambda) = \lambda H(\mu, \lambda^{-1}), \tag{8}$$

which states that the statistical 2D Ising model is self-dual, i.e. the high temperature and the low temperature phase of the (very same) statistical 2D Ising model is dual to each other. We also have $E(\lambda) = \lambda E(\lambda^{-1})$. Therefore, if the critical point is unique, we must have $\lambda = 1$. Note $\lambda = 1$ is deduced from the assumption that the critical point is unique (therefore $\lambda = \lambda^{-1}$), not from the fact that at the critical point $E(\lambda) = E(\lambda^{-1})$: this is because at critical point $E = 0$ therefore even if $E = \lambda E$ $\lambda$ is not determined.

Note here by duality we mean that for the 1D quantum (transverse field) Ising model $H$, $H(\lambda)$ is dual to $H(\lambda^{-1})$. Now the question is: when translating the language back to the statistical 2D Ising model what does the duality mean? To answer this, we first look at the self-duality of the isotropic statistical 2D Ising model:

$$Z(K) = \sum_{\{\sigma_i\}=\pm 1} e^{K\sum_{\langle ij\rangle}\sigma_i\sigma_j},$$

where $i$ runs over all the 2D lattice sites, $K = \beta J$. Note $K > 0$, so that Hamiltonian is actually $H = -J\sum \sigma\sigma$ with $J > 0$, and the ground state is ferromagnetic. At high temperature, $K$ is small, we use $e^{K\sigma} = \cosh K + \sigma \sinh K = \cosh K(1 + \sigma \tanh K)$ to do the expansion:

$$Z(K) = \cosh^{2MN} K \sum_{\{\sigma_i\}=\pm 1} \prod_{\langle ij\rangle} (1 + \sigma_i\sigma_j \tanh K),$$

where $M$ is the number of columns and $N$ the number of rows. Using the fact $\sum_{\sigma=\pm1} \sigma = 0$ and $\sum_{\sigma=\pm1} 1 = 2$ we see that only the terms in the sum with $\sigma_i$'s appearing all even number of times will contribute, and these are precisely the ones where $\langle ij\rangle$ extends to a closed loop. Depending on the number of the bonds in a loop, call it $n$ (notice $n$ must be even), we will get a term $f_n(MN)\tanh^n K$, where $f_n(MN)$ is a degree-$n$ polynomial of $MN$. Note we also have a factor of $2^{MN}$ recording the number of configurations for the spins. Now we write down the first few terms of $f$: we have $f_0 = 1$, $f_2 = 0$, $f_4 = MN$, $f_6 = 2MN$, $f_8 = \frac{1}{2}MN(MN - 5)$, etc., i.e. we have

$$Z(K) = 2^{MN}\cosh^{2MN} K \left(1 + MN\tanh^4 K + 2MN\tanh^6 K + \frac{1}{2}MN(MN-5)\tanh^8 K + \cdots\right).$$

then, consider the low-temperature expansion: since at zero temperature all spins are alighed, the low temperature expansion can be done with respect to the number of flipped spins. Note that if one spin is flipped, four bonds are broken which gives energy difference of $8K$, and there are $MN$ ways of choosing that spin; if two spins are flipped, then six spins are flipped if they are neighbors; and eight bonds are broken otherwise, Creating energy difference of $12K$ and $16K$, and whose way of choosing configurations are $\sim N$ and $\sim N^2$, respectively. Therefore the low energy expansion gives

$$Z(K) = e^{2MNK}\left(1 + MNe^{-8K} + 2MNe^{-12K} + \frac{1}{2}MN(MN-5)e^{16K} + \cdots\right),$$

the above two equations cam be mapped term-by-term. This is actually reasonable: for any flipped spin, or more rigorously the connected region of flipped spins, one can surround it by a square. Therefore the way of choising flipped spins are exactly mapped to the way of chosing squares. We see that, for any given $K$ and the model $Z(K)$ in the high temprature phase, we can fabriate another model $K^*$ and $Z(K^*)$ which stays in the low temperature phase, where

$$\tanh K = e^{-2K^*},$$

define $\exp(K) = x$ and $\exp(K^*) = y$, this equation is $\frac{x-1/x}{x+1/x} = 1/y^2$, or $(x^2-1)y^2 = (x^2+1)$, or $x^2y^2 - x^2 - y^2 - 1 = 0$, or $x^2y^2 - 1 = x^2 + y^2$, square both sides we get $x^4y^4 + 1 - 2x^2y^2 = x^4 + y^4 + 2x^2y^2$, or $x^2y^2 - \frac{x^2}{y^2} - \frac{y^2}{x^2} + 1 = 4$, or

$$\sinh 2K \sinh 2K^* = 1, \tag{9}$$

and we have $\frac{Z(K^*)}{(e^{2K^*})^{MN}} = \frac{Z(K)}{2^{MN}(\cosh^2 K)^{MN}}$ which can be simplified to

$$\frac{Z(K^*)}{\sinh^{MN/2}(2K^*)} = \frac{Z(K)}{\sinh^{MN/2}(2K)}.$$

now the duality is manifest: a temperature phase at $K$ corresponds to a low temperature phase at $K^*$ provided $\sinh 2K \sinh 2K^* = 1$. Note the word duality simply means a mapping between two points in the phase space (here the phase space is one dimensional $K = J/T$) such that they have the same partiion function form (or, better, the same form of Hamiltonian). Now, again, if we assume the model has a unique critical point (as is true for the 2D Ising model) it must be at a temperature such that $K = K^*$, therefore the critcal temperature is given by $\sinh^2(2K_c) = 1$, or $x^2 - 1/x^2 = 2$, or $x^2 = 1 \pm \sqrt{2}$ (of course we must pick plus sign since $K_c > 0$, therefore

$$K_c = J/k_BT_c = \frac{1}{2}\ln(\sqrt{2}+1). \tag{10}$$

As mentioned by Kogut, this result was found by Kramers and Wannier long before the Ising model was solved by Onsager.

We have just found the critical point of the isotropic statistical 2D Ising model. We have derived before that the anisotropic statistical 2D Ising model can be mapped to 1D quantum transverse field Ising model. So how about the critical points (or we should say line) of the anistropic statistical 2D Ising model? Kogut mentioned that we can use the same method (low and high temperature expansion) to obtain, after a little effort,

$$\sinh 2\beta_\tau \sinh 2\beta = 1. \tag{11}$$

One should not confuse this equation with Eq. (9) although they have the same form: Eq. (9) is the relation for $K$ and $K^*$ between the dual models; while Eq. (11) is the condition for the model to be self-dual, i.e. at critical point. Eq. (10) is really a special point of Eq. (11) in the case of $\beta_r = \beta$.

Now we organize some things more straightly. We review what we have done: we have mapped the anisotropic statistical 2D Ising model into a quantum transverse field 1D Ising model, using the transfer matrix technique. There we introduced a parameter $\lambda$ such that $\beta = \lambda e^{-2\beta_\tau}$. In the quantum transverse field 1D Ising model, we should that the model at $\lambda$ is dual to the model at $\lambda^{-1}$ and hence the self-dual critical point is at $\lambda = 1$, this means that for the anisotropic statistical 2D Ising model the self-dual critical point (line) happens at $\beta = e^{-2\beta_\tau}$, setting $\lambda = 1$. On the other hand, using low- nd high-temperature expansion we were able to find the duality between the isotropic statistical 2D Ising model at $K$ and that at $K^*$, and we solved the critical point which is $K_c = K_c^* = \frac{1}{2}\ln(\sqrt{2}+1)$; and we mentioned that the same method can be shown to give critical point for the anisotropic statistical 2D Ising model: $\sinh 2\beta_\tau \sinh 2\beta = 1$. What is missing here is the relationship between the result $\beta = e^{-2\beta_\tau}$ and $\sinh 2\beta_\tau \sinh 2\beta = 1$.

It is easy to see that the former equation is simply a limiting case of the latter equation, when $\beta_\tau \to \infty$; in this ase, we must have $\beta$ very small, and therefore $\sinh 2\beta \sim \beta$, and we have $\sinh 2\beta_\tau \sim e^{2\beta_\tau}$. This is saying that the condition for the critical point obtained from the quantum case, $\beta = e^{-2\beta_\tau}$, is valid only in the $\beta_\tau \to \infty$ limit. This is exactly true: we have mentioned when using the transfer matrix technique that we were aiming for a smooth $\tau$-continuum theory. A smooth $\tau$-continuum is a synonym for a quantum mechanical model, since in a quantum mechanical model temporal direction is always continuous (except for quantum gate, etc.). To emphasize once more: the transfer matrix method writes a (classical) statistical partition function $Z$ as $Z = \mathrm{Tr}T^N$, where this $N$ does not have to be large; but to regard $Z = \mathrm{Tr}T^N$ as a quantum mechanical problem, we must send $N \to$, therefore $\delta\tau = \tau/N$ is small, due to the fact that path integral in quantum mechanics must be continuous in the $\tau$ direction. More concretely, the starting point of writing transfer matrix towards a smooth $\tau$-continuum theory (i.e. quantum path integral) is the moment we define $\beta \sim \delta\tau \sim e^{-2\beta_\tau}$, and this is assumping $\beta \to 0$ and $\beta_\tau \to \infty$. As mentioned by Kogut, this is a general feature when constructing the $\tau$-continuum theory, mapping a classical partition function to a quantum path integral: the couplings must be adjusted so that the temporal coupling grows large while the spatial coupling becomes weak.

As we now have learned, the $\tau$-continuum formulation of the anisotropic statistical 2D Ising model is just the path integral for the quantum transverse 1D Ising chain, and the quantum transverse 1D Ising chain is exactly soluble, using Jordan-Wigner transformation. The Hamiltonian eventually becomes a fermionic BdG model, and the dispersion can be solved to be $2\sqrt{1+2\lambda\cos k+\lambda^2}$. Note the lowest energy excitation happens at $k = \pm$: when $\lambda \neq 1$ (i.e. not critical), the excitation is gapped; when $\lambda = 1$, the spectrum is gapless and disersion is linear around $k = \pi$, i.e. Lorentz symmetry is restored. The zero field is self-charge conjugate, hence a Majorana field.

# 3 Wegner Ising lattice gauge theory

2D Ising model: block of perimeter $L$ corresponds to partition function $Z_L = \mu^L e^{-2\beta L}$ where $\mu^L$ counts the number of blocks with perimeter $L$, where $\mu$ can be estimated using random walks, and $2L$ is the excitation energy. We see that as $L$ grows, region of spin flip is not energetically favored, therefore magnetization is not destroyed at low temperature; at high temperature though, the entropy term $\mu^L$ grows much faster than $e^{-2\beta L}$ and spin flip regions are favored; hence phase transition and magnetization vanishes. By the same spirit, if we look at 1D Ising model, the energy for spin flip interval with size $L$ is constant $e^{-2\beta}$ but entropy scales with $\mu^L$, therefore magnetization is destroyed at any finite temperature.

Wegner's motivation: have a model which does not magnetize but have nontrivial phase diagram. This is the Ising lattice gauge theory. Note as Leon writes in his PHY217 lecture notes, Kitaev's toric code is "largely a reinterpretation of the Ising lattice gauge theory introduced much earlier by Wegner". Let us use Toric code's language to start with since it's simple.

First the crucial starting point is that spins now live on the bonds of a $d$-dimensional cubic lattice, i.e. now we put a spin state on the middle of each bond. This spin state of course has most general form of $a|\uparrow\rangle + b|\downarrow\rangle$. Define $P_p = \prod_{i\in p}\sigma_i^z$ to be product of $\sigma^z$ of a plaquette $p$, and $S_s = \prod_{i\in s}\sigma_i^x$ to be product of $\sigma^x$ of a star $s$. Wegner's Ising lattice gauge model then has the form

$$S = -J\sum_p P_p.$$

Note that for any $s, p$, we have $[S_s, P_p] = 0$, this is because $S_s P_p S_s^{-1}$ will flip either 0 or 2 bonds (i.e. flip the spin state) in a plaquette $p$, and the product is not changed. To be more precise: for any state $|\psi\rangle$ we have to prove $S_s P_p|\psi\rangle = P_p S_s|\psi\rangle$: since $P_p$ only acts on $p$ therefore for $i \notin p$ the action is simply just $S_s$; now we just need to look at the four spins states on the plaquette; note such a state is just a linear superposition of $|abcd\rangle$ where $a, b, c, d \in \{\uparrow, \downarrow\}$, therefore we just have to prove that $S_s P_p|abcd\rangle = P_p S_s|abcd\rangle$. since $S_s$ flips 0 or 2 spins of $\{a, b, c, d\}$ and $P_p|abcd\rangle = abcd|abcd\rangle$, we see that $[P_p, S_s]$ clearly holds.

We stress here that Wegner's Ising lattice gauge theory is classical gauge theory, in the sense that only the plaquette term $P_p$ is involved therefore there is no nontrivial commutation relation.

Elitzur's theorem: now we compute such a quantity

$$\langle \sigma_i^z \rangle \equiv \frac{\sum_{\text{spin config}} \sigma_i^z e^{\beta \prod_p P_p + h \sum_j \sigma_j^z}}{\sum_{\text{spin config}} e^{\beta \prod_p P_p + h \sum_j \sigma_j^z}},$$

the important thing to notice is that the $\langle \sigma_i^z \rangle$ on the left is just telling to calculate the expectation value of $\sigma^z$ on bond $i$, and has different meaning from the $\sigma_i^z$ on the right; the $\sigma_i^z$ on the right is actually a dummy variable: it changes as different spin configurations are picked from the sum.

define $\widetilde{\sigma}_j^z = -\sigma_j^z$ for $j \in s$ where $s$ is a star containing $i$., and $\widetilde{\sigma}_j^z = \sigma_j^z$ otherwise. We have

$$\langle \sigma_i^z \rangle = \frac{\sum_{\text{spin config}} -\widetilde{\sigma}_i^z e^{\beta \prod_p \widetilde{P}_p + h \sum_{j \notin s} \widetilde{\sigma}_j^z - h \sum_{j \in s} \widetilde{\sigma}_j^z}}{\sum_{\text{spin config}} e^{\beta \prod_p P_p + h \sum_j \sigma_j^z}},$$

note the crucial point is that, 1) the left side contains not a variable; if we write the left side as $-\langle \widetilde{\sigma}_i^z \rangle$, we will not be able to get the desired result; 2) we do not change the denominator since it is just a sum; otherwise we will not get desired result either. Following the above, change the notation back $\widetilde{\sigma}_j^z \to \sigma_j^z$, we have

$$\langle \sigma_i^z \rangle = -\frac{\sum_{\text{spin config}} \sigma_i^z e^{\beta \prod_p P_p + h \sum_{j \notin s} \sigma_j^z - h \sum_{j \in s} \sigma_j^z}}{\sum_{\text{spin config}} e^{\beta \prod_p P_p + h \sum_j \sigma_j^z}}$$

since $-\sigma_j^z = \sigma_j^z - 2\sigma_j^z$, we have

$$\langle \sigma_i^z \rangle = \langle -\sigma_i^z e^{-2h \sum_{j \in s} \sigma_j^z} \rangle,$$

we see that as $h \to 0$, $\langle \sigma_i^z \rangle = \langle -\sigma_i^z \rangle$ and therefore $\langle \sigma_i^z \rangle = 0$.

The mathematical reason for this to vanish is all hidden above. The physical interpretation is that even in the presence of $h$, the configuration with $\sigma_i^z$ positive and the other with $\sigma_i^z$ negative only differ by a finite energy barrier (the two configurations are related by the operator $S_s$), thus this is essentially a quantum mechanics problem with only a few local degrees of freedom. We have seen that symmetry is never broken in a finite quantum system. For similar argument, see Nagaosa below Eq. (3.4.9). To summarize: the local expectation value vanishes in a gauge system because only local gauge transformation is needed to change this local expectataion value, unlike in a non-gauge system with global symmetry, in which the change of all degrees of freedom is needed to change the local expectation value.

The significance of Elitzur's theorem is that gauge systems cannot have local order parameters. In other words, if a gauge system exhibits different phases, the phase transition cannot be discribed by broken symmetries.

Below we will study 2D Ising gauge theory and 3D Ising gauge theory. We will see that the 2D Ising gauge theory is equivalent to a (classical) statistical 1D Ising model, and that the 3D Ising gauge theory is equivalent to a quantum (transverse field) 3D Ising model. The 2D claim is easy: as the case of statistical 2D Ising model, we again define two directions as the temporal $\tau$ and spatial $x$ directions. We can then use the gauge freeom to fix $\sigma_{n,\tau}^z = 1$, where now we are labeling the Ising variable on a bond $(n, n+\tau)$ as $\sigma_{n,\tau}^z$. This way, the action becomes

$$S = -J \sum \sigma_{n,x}^z \sigma_{n+\tau,x}^z,$$

we see that the remaining two coupled Ising variables are neighbors in the temperal direction of the bonds; different spatial positions are therefore decoupled, meaning that Wegner's (classical) 2D Ising gauge theory is equivalent to (decoupled and classical) statistical 1D Ising model. We have proved before that a statistical 1D Ising model is always in disordered phase for any finite temperature $T > 0$, so is the 2D (classical) Ising gauge theory, i.e. it only has one phase.

Now we study the Wegner's (classical) 3D Ising gauge theory. Again we use the temporal gauge, i.e. we set $\sigma_{n,\tau}^z = 1$; but note we are not exhausing gauge freedom: if we apply the same in-layer gauge transformation for all the spatial layers, i.e. if we are performing a $\tau$-independent gauge transformations, clearly $\sigma_{n,\tau}^z = 1$ is still maintained. Now, we classify all the plaquette terms in $S$ into two types: the vertical/temporal one (i.e. the ones containing a $\tau$ direction) and the horizontal/spatial one (i.e. the one that lies in the spatial layer), and give them different coupling constant:

$$S = -\beta_\tau \sum_{p_\tau} P_{p_\tau} - \beta \sum_{p_x} P_{p_x},$$

where $P_{p_\tau}$, due to the temporal gauge, can be written as $P_{p_\tau} = \sigma_{n,x}^z \sigma_{n+\tau,x}^z = -\frac{1}{2}(\sigma_{n,x}^z - \sigma_{n+\tau,x}^z)^2 + 1$, therefore, up to some constant,

$$S = \frac{1}{2}\beta_\tau \sum_{p_\tau} (\sigma_{n,x}^z - \sigma_{n+\tau,x}^z)^2 - \beta \sum_{p_x} P_{p_x},$$

by completely analogous method of using transfer matrix method to rewrite the (classical) anisotropic statistical 2D Ising model to the quantum Hamiltonian (i.e. the $\tau$-continuum theory), here again we have

$$\hat{H} = -\sum_{n,i} \sigma^x_{n,i} - \lambda \sum_{p_x} P_{p_x}, \tag{12}$$

where in taking the $\tau$-continuum limit we have assumed

$$\beta \sim \delta\tau \sim e^{-2\beta_\tau} \to 0.$$

Now we have a 2+1D quantum Ising model $\hat{H}$. Now we see the remaining gauge freedom is precisely exercised by the star operator $S_s$ acting on the site of the 2D lattice: we have probed that $[S_s, P_p] = 0$, or $S_s^{-1} P_p S_s = P_p$, or

$$S_s^{-1} \hat{H} S_s = \hat{H}.$$

This means that if $|\psi\rangle$ is an eigenstate of $\hat{H}$, then $S_s|\psi\rangle$ is also an eigenstate. We further conclude that $S_s|\psi\rangle$ is just $|\psi\rangle$ (which contains serious assumption/argument which must be justified!), as Kogut does. This means that $\langle\psi|\sigma^z_{n,i}|\psi\rangle = \langle\psi|S_n^{-1}\sigma^z_{n,i}S_n|\psi\rangle = \langle\psi|(-\sigma^z_{n,i})|\psi\rangle = 0$. This can be regarded as the quantum version of the Elitzur;s theorem.

Now we proceed further to show that the Wegner's 3D Ising gauge theory is equivalent to the statistical 3D Ising gauge theory. We first show the former is equivalent to the quantum transverse field 2D Ising model. By analogy of the equivalence between quantum transverse field 1D Ising chain and the statistical 2D Ising model, one should believe that a quantum transverse field 2D Ising model is equivalent to a statistical 3D Ising model (again, in the highly anisotropic, $\tau$-continuum limit).

We start by using the remaining gauge fixing freedom to eliminate $\sigma^z_{n,y}$. For detailed justification that this is possible and that this exhausts all the remaining gauge freedom, see Kogut's argument therein; however I an not very familiar with his language: he was doing the elimination on the operator label; but the only thing I can understand is doing gauge fixing in the original classical Ising gauge model. Therefore I will hand-wave my argument, which I believe is essentially equivalent to Kogut's language: we have mentioned before that in writing down the quantum model (12) not all gauge freedoms are fixed. Therefore we can first fix all the gauge freedom and then write down the quantum model: we can gauge fix $\sigma^z_{n,\tau} = \sigma^z_{n,y} = 1$ and leave $\sigma^z_{n,x}$ undetermined, and this has exhaused all the gauge freedom. Therefore, when using transfer matrix to write the quantum model, what we will get is simply

$$\hat{H} = -\sum_{n,x} \sigma^x_{n,x} - \lambda \sum_{P_x} \sigma^z_{n,x}\sigma^z_{n+y,x}. \tag{13}$$

Now we proceed to define a dual lattice: sites $n^*$ of the dual lattice is associated with (i.e. at the center of) the plaquette $P$. We then define

$$\mu^1_{n^*} = P_p = \sigma^z_{n,x}\sigma^z_{n+y,x}, \qquad \mu^z_{n^*} = \prod_{n' \geq 0} \sigma^x_{n-n'y,x},$$

where we have chosen $n$ to be, e.g. the lower-left corner of the plaquette $p$. It is easy to see that $\mu^x_{n^*}\mu^z_{n^*} = -\mu^z_{n^*}\mu^x_{n^*}$, and that $\mu^x_{n^*}\mu^z_{m^*} = \mu^z_{m^*}\mu^x_{n^*}$ (both in the tensor product sense) when $m$ not $n$. On the other hand, we have $\mu^z_{n^*}\mu^z_{n^*-y} = \sigma^x_{n,x}$, we see that the Hamiltonian (13) can be written as

$$H = -\sum_{n^*,i} \mu^z_{n^*}\mu^z_{n^*+i} - \lambda \sum_{n^*} \mu^x_{n^*}.$$

This again in turn is the quantum Hamiltonian formulation of the statistical 3D Ising model with $1/\lambda$ (remember that from $d$-dimensional statistical to $(d-1)+1$-dimensional quantum, $\lambda$ always appear in the $\sigma^z\sigma^z$ term and not the transverse field term), therefore we have

$$H_{\text{Wegner's classical 3D Ising gauge}}(\lambda) = \lambda H_{\text{Statistical classical 3D Ising}}(\lambda^{-1}), \tag{14}$$

therefore we have a mapping between the high- (low-) temperature properties of the gauge system and the low- (high-) temperature properties of the statistical 3D Ising model; and the latter is of course well studied.

This map gives more intuitive characterization of the gauge theory: in the low temperature phase of the statistical 3D Ising model, $\langle 0|\mu^z_{m^*}|0\rangle$ is an order parameter and is nozero. This term is mapped to $\langle 0|\prod_{n\geq 0}\sigma^x_{m,n-y,x}|0\rangle$ of the 3D Ising gauge theory, which is nonzero at high temperature not zero at low temperature. Therefore we call the latter a nonlocal disorder parameter. It can be thought of as a kink operator, and the intuitive picture for the two phases of the gauge theory is as follows: at low temperature phase, the gauge theory is free of kinks, while for $T$ above a critical $T_c$, there is a kink condensate. The kink language talks about the string operator, Alternatively, we can use a gauge-invariant language in which the fundamental objects are the plaquette term $P_p = \pm 1$ (note it is a classical object so we can talk about its

value): if $P_p = -1$ we call it frustrated. We see tat if the operator $\prod_{n\geq 0}\sigma^x_{m,n-y,x}$ is appled to the $T=0$ state of the gauge theory, which is free of frustrated plaquettes, then a frustration is made at $m^*$. Therefore, the phase transition from the low temperature to the high temperature phase can be viewed as a frustration condensation. Note that kink and frustration are just two languages of the same object; this object is fundamentally nonlocal (which is better seen in the kink langauge), and therefore can develop vacuum expectation value, without violating Elitzur's theorem.

So far we have established the equivalence between the gauge 2D Ising and the statistical 1D Ising, and between the gauge 3D Ising and the statistical 3D Ising. Note these two equivalences are totally different in flavor, since in the first equivalence temperature map is $T$ to $T$, and in the secnd equivalence the temperature map is $T$ to $T^{-1}$. One must bear this mind.

A part left unstudied is the scaling of (gauge invariant) correlation functions in the gauge theories. We will talk about it when I have time, and for now we refer to Kogut, IV D. We only state the most important conclusion: for gauge theory with two phases,

- At high temperature, we have the area law:
$$\left\langle \prod_{i\in C}\sigma^z_l \right\rangle \sim e^{-A};$$
- At low temperature, we have the perimeter law:
$$\left\langle \prod_{i\in C}\sigma^z_l \right\rangle \sim e^{-P}.$$

It is the behavior of this correlation function that determines a gauge theory has two phases. The two scaling behaviors can be obtained using high- and low-temperature expansions, taking into account the complication due to gauge freedoms. Note the argument for the low temperature phase does not work for 2D because the expansion has vanishing radius of convergence, consistent with the fact that the 2D gauge theory only has a high temperature phase. The area law is hand-in-hand with the kink condensation: the rough argument is that, when calculating $\langle\prod_C \sigma^z\rangle$, there can be $N_c$ kinks inside the contour $C$, and $N_c \sim A$ the area enclosed by $C$. This picture applies to Ising gauge theories in any dimension. The unique property of the 2D is that kinks cost finite energy to be produced, and therefore it is very easy to produce kinks and obtain area law, as soon as temperature is inifinitesimally above zero.

The 4D Ising gauge theory is self-dual, with self-dual point $\lambda = 1$, with two phases; the phase transition at $\lambda = 1$ is first-order. For details, see the paper by Fradkin and Susskind which we alrady mentioned before: `https://journals.aps.org/prd/pdf/10.1103/PhysRevD.17.2637`.

# 4 $Z_2$ and U(1) spin/gauge models: summary and generalization

As can be sensed, Kogut's review actually discussed several types of duality. Now we give a summary and comment on the different types of dualities. This is following the nice Review of Robert Savit, `https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.52.453`.

First of all, let us review the statistical Ising model. Quote Savit: these models can be regarded either as classical statistical mechanical systems in $d$ space dimensions, or as Euclidean cutoff quantum field theories in $d$ spacetime dimensions. In the former case, the argument of the partition function should be thought of as the exponential of the Hamiltonian of the classical statistical system, while in the latter case the partition function's argument is the exponential of the Lagrangian of the quantum field theory.

Let first adopt the first point of view. Then we can try to map the classical Hamiltonian into another classical Hamiltonian, by defining new variables on the dual lattices: this is the Kramers-Wannier duality. Savit has made a nice summary of the duality of the Ising and Ising gauge theories. Remember so far we have spins defined on either sites (Ising theories) or links (Ising gauge theories), which are spins defined on 0D and 1 objects. We can then generalize the idea to define spins on $r$-dimensional objects, where $r = 0$ and 1 for the case just mentioned, i.e. we write Ising spin as $\sigma_{(x_1,x_2,\ldots,x_r)} = \pm 1$. Note in Ising spin (defined on 0D sites), each term is defined on 1D link, and in Ising gauge spin (define on 1D link), each term is defined on 2D plaquette. Therefore we define terms for the general $r$-dimensional spin on a $s = r+1$ dimensional object which we call plaquette, or a better name $s$-simplex $\Delta_s$, i.e. each term is $\Sigma_{\Delta_s} \equiv \prod_{i=1}^{2s}\sigma_{\partial_i\Delta_s}$, where note $\partial_i\Delta_s$ are $i$-th $r = s-1$-dimensional faces of the $s$-simplex, and for hypercubic lattices there are $2s$ of them. The classical or statistical action then is $Z = e^{-\beta H}$, with

$$H = -J\sum_{\Delta_s\in\mathbb{Z}^d}\Sigma_{\Delta_s}.$$

From these construction, it is evident that whenever $s \geq 2$, we have $r = s - 1 \geq 1$, and therefore the object on which Ising spin $\sigma_{(x_1,...,x_r)}$ sits, $\partial_i \Delta_s$, also has boundaries, which we call $\partial_j \partial_i \Delta_s$, where $j = 1, 2, ..., r-1$. The Hamiltonian $H$ then has local symmetry due to the transformation of these $r-1$ objects, due to the fact that each term $\sum_{\Delta_s}$ is gauge invariant. Now, we recap that

- Ising theory: $r = 0$, $s = r + 1 = 1$;
- Ising gauge theory: $r = 1$, $s = r + 1 = 2$.

By analogy of our previous experience in writing the statistical Ising model using the transfer matrix method, we see that the dimension of $\Delta_s$, which is $s$, must be smaller than the dimension of the hypercubic lattice, which is $d$, by 1, i.e. $s \leq d - 1$, in order to separate out one lattice direction to make it a temporal direction. This direction then can be made continuous if we assume the coupling along this direction, which we call $J_\tau$, is different from those along other "spatial directions", which we call $J_x$. Define $\beta_\tau = J_\tau / k_B T$ and $\beta = J_x / k_B T$ (which amounts to setting $J_x = 1$), our previous lesson tells us that the $\tau$-continuum limit, or a quantum formulation, is given in the highly anisotropic case

$$\beta \sim \delta\tau \sim e^{-2\beta_\tau} \to 0.$$

On the other hand, when $s = d$, one cannot go to the $\tau$-continuum limit, i.e. there is no dynamical degrees of freedom. The example of the $s = d$ cases we have studied are the Ising model in 1D and the Ising gauge model in 2D (Note we just commented on the latter in the last Section without studying it in detail; but our study of the 2D Abelian lattice gauge theory carefully derived in the section does apply to the Ising gauge theory in a complete analogous way, which leads to the same conclusion that the Wilson loop always have area law, and that there is only one confining phase), the partition function is just a simple product of numerical ($\beta$-dependent) factors.

We just mentioned that we have a gauge theory whenever $s \geq 2$. In this case, the natural gauge invariant correlation function has its fundamental block as $\partial \Delta_s$, which is an $s-1$ object; more generally any $s-1$-dimensional closed surface (or we should say hyperloop) is gauge invariant.

The classical duality transformation maps a $(s, d, \beta)$ Ising model to a $(\widetilde{s}, d, \widetilde{\beta})$ model, with

$$\widetilde{s} = d - s, \qquad \sinh 2\widetilde{\beta} \sinh 2\beta = 1.$$

Note we have studied the $d = 3$ case, at which the $(s, d) = (1, 3)$ Ising model is dual to the $(\widetilde{s}, d) = (2, 3)$ Ising lattice theory; this is exactly what we exhibited in the last section with the conclusion written there in Eq. (14). Note the relation further implies that, if the dimension $d$ is even, then the model $(s, d)$ with $s = d/2$ is self-dual. We have studied such self-dual models in the lowest dimensions, $(s, d) = (1, 2)$ and $(s, d) = (2, 4)$, where the former is the 2D Ising model studied with the conclusion Eq. (8), and the latter is the 4D Ising gauge theory which we commented at the end of the last section.

However, one must note that the following dualities that we have mentioned do not fall into the above (i.e. Kramers-Wannier) duality: we have mentioned briefly (since the duality is easy to estalish) that the 2D Ising lattice gauge theory is dual to 1D Ising chain [see Kogut Eq. (5.35)]; we have also mentioned in passing that the 2D Abelian lattice gauge theory is dual to 1D planar spin chain [see Kogut Eq. (6.46)]. Note these two dualities are not like Kramers-Wannier duality: the K-W duality, is a statement about two theories in the same (spacetime/space) dimension $d$, and the ordered/disordered phase of one theory maps to the disorder(confined)/ordered(deconfined) phase of the other theory. However, the two dualities we mentioned here does not map inversely (i.e. does not map temperature of one theory to the inverse temperature of the other theory), and what's more the duality is about two theories in different dimensions (so far we have $d = 2$ for one theory and $d = 1$ for the other). Such a duality was first noticed by Migdal and then Kadanoff. The statement is basically that

> "Most 4D gauge theories are similar to their 2D spin system counterparts in the sense that the renormalization-group equations have the same structure in both systems, and they exhibit the same kind of instantons, etc."

quoted from Fradkin & Susskind, PRD 17 2637. Kogut mentioned that more generally there is correspondence between a $d$-dimensional gauge system and $d/2$-dimensional spin systens for gauge groups $Z_N$ ($Z_2$), $U(1)$ and $SU(N)$; the corresponding spin system has $Z_N$ ($Z_2$), $U(1)$ and $SU(N) \times SU(N)$ spin rotation symmetry. The claim of Fradkin & Susskind is the case with $d = 4$. Note Kogut in his review has studied briefly this statement in $d = 2$: the duality between a 2D Ising lattice gauge theory ($d = 2$, $s = 2$) and a 1D Ising chain ($d = 1$, $s = 1$), and the duality between a 2D Abelian lattice gauge theory and a 1D planar spin chain. As mentioned above, the duality maps confinement/deconfinement to disorder/order, which is opposite to the mapping direction in Kramers-Wannier duality.

Then, let us see what we adopt the latter interpretation of regarding the model as a Euclidean quantum field theory in $d$ space-time dimension. Quote Savit: If we adopt the latter interpretation, we can rework the lattice theory so that it is expressed in terms of the Hamiltonian of the quantum field theory. (This is not the same as the Hamiltonian of the

classical statistical system described above.) This is exactly the $\tau$-continuum formlation of a quantum Hamiltonian, which is a formalism developed by Kogut and Susskind; this is referred to by Savit as the lattice Hamiltonian formalism. Go on quoting Savit: In the Hamiltonian formulation of latticized quantum field theory, one space-time direction is identified as time, and the lattice spacing in that direction is sent to zero. Thus the $d$-dimensional lattice becomes a $(d-1)$-dimensional lattice plus one continuous time axis. In order to have a chance of retaining the same (large distance) physics when the time direction lattice spacing is sent to zero, it is necessary to allow lattice coupling constants in the time direction to vary with the lattice spacing so that the effective interaction over some fixed distance is unchanged.$\cdots$ Since time is now continuous, it is relatively straightforward to define ommenta which are canonically conjugate to the coordinates of the theory. The Hamiltonian can then be written, in the usual way, in terms of these coordinates and momenta. Thus, unlike the Lagrangian or the classical statistical mechanics Hamiltonian, this fielf-theoretic Hamiltonian contains noncommuting operators. Example would be the (quantum) transverse field Ising model obtained in Kogut's review.

Furthermore, the quantum Hamiltonian so obtained can also have several dualties. To start with, the 1D transverse Ising model is self-dual, reflecting the fact that the classical statistical 2D Ising is self-dual. Thinking reversely, since we know that the classical statistical 4D Ising gauge theory is self-dual, it must also be possible to show that the quantum lattice Hamiltonian corresponding to the 4D Ising gauge theory is also self-dual. This is briefly mentioned at the end of the last section.

We go on following Savit to give a summary of the dualities for the U(1) theories: again dimension is $d$ and simplex number $s$, with $s=1$ a spin model and $s\geq 2$ a lattice gauge theory. The dual theory is a $Z_\infty$ invariant theory, also in $d$ dimensions, with simplex number $\widetilde{s}=d-s$, and analogous to the Kramers-Wannier duality, the high- (low-) temperature region of the original theory is mapped into the low- (high-) temperature region of the dual theory. Following the derivation of KT theory, we see that the dual theory can be easily converted into a third form, in which the physical degrees of freedom appearing are spin waves and the topological excitations of the original U(1) invariant spins. The topological excitations exist on closed manifolds if dimension $d-s-1$.

# 5 Abelian lattice gauge theory

Wegner's (classical/statistical) Ising lattice gauge theory is generalized to (classical/statistical) continuous (Abelian) lattice gauge theories by Wilson and Polyakov. The starting point is a planar (XY) model: let spins have only $XY$ components and write $\boldsymbol{s}_n=(\cos\theta_n,\sin\theta_n)$, then the action is

$$H=-J\sum_{n,\mu}\boldsymbol{s}_n\cdot\boldsymbol{s}_{n+\mu}=-J\sum_{n,\mu}\cos(\theta_n-\theta_{n+\mu})=-J\sum_{n,\mu}\cos\Delta_\mu\theta_n, \tag{15}$$

where we introduced the finite difference $\Delta_\mu\theta_n=\theta_{n+\mu}-\theta_n$. Now we specify two cases: the original XY model is defined on the sites of a (square) lattice. In this case, the model has a global U(1) symmetry and is just called the XY model. Although the symmetry is just a global one and there is no gauge freedom, the model itself is interesting in the sense that it has two phases which are both disordered by the creed of Mermin-Wigner theorem; and the phase transition is induced by vortex condensation rather than broken symmetry (no symmetry is broken as required by M-W theorem). This is the famous KT transition and will be studied in the next section. In this section, we study the other case in which the XY variables $\theta$ are located on the links of a lattice, and this immediately introduced local continuous symmetry. We then denote these variables by $\theta_\mu(n)$, with the consistency condition $\theta_\mu(n)=-\theta_{-\mu}(n+\mu)$, and define curl$\theta(n)=\theta_{\mu\nu}(n)=\Delta_\mu\theta_\nu(n)-\Delta_\nu\theta_\mu(n)=\theta_\nu(n+\mu)-\theta_\nu(n)-\theta_\mu(n+\mu)+\theta_\mu(n)=\theta_\mu(n)+\theta_\nu(n+\mu)+\theta_{-\mu}(n+\mu+\nu)+\theta_{-\nu}(n+\nu)$, it is easy to see that curl$\theta(n)$ is invariant under transformation $\theta_\mu(n)\to\theta_\mu(n)+\chi_{n+\mu}-\chi_n=\theta_\mu(n)+\Delta_\mu\chi_n$, and we define the action of the Abelian lattice gauge theory to be

$$S=J\sum_{n,\mu\nu}(1-\cos\theta_{\mu\nu}(n)).$$

Note that everything here parallel's the classical analysis of Maxwell's theory of electromagnetism. Under proper condition which we will analyze below, the cos term can be expanded and we have $S=J\sum\frac{1}{2}\theta_{\mu\nu}^2\to J\int\frac{d^dx}{a^d}\frac{1}{2}\theta_{\mu\nu}^2$, and the correspondence to the action $S=\frac{1}{4}\int d^4xF_{\mu\nu}F_{\mu\nu}$ can be identified. Here we will introduce $J=1/2g^2$, where $g$ is the lattice coupling constant.

Now we look at phases of the Abelian lattice gauge theory by looking at the scaling law of gauge invariant correlation functions. Note again Elitzur's theorem applies so that the ground-state expectation value of say $\cos\theta_\mu(r)$ vanishes, and we have to use the gauge invariant quantity for correlation function. We use the Wilson loop variable $e^{i\oint_C\theta_\mu(r)}$, and therefore the gauge invariant correlation function is

$$\left\langle e^{i\oint_C\theta_\mu(r)}\right\rangle=\frac{1}{Z}\prod_{r,\mu}\int_0^{2\pi}d\theta_\mu(r)e^{i\oint_C\theta_\mu(r)}e^{-S}.$$

First we suppose $J \ll 1$, or $g^2 \gg 1$, which amounts going to the high temperature regime, and then we can use the high-temperature expansion. Again we use the fact that $\int d\theta e^{i\theta} = 0$, and $\int d\theta = 2\pi$. Notice that the action term

$$e^{\frac{1}{2g^2}\sum \cos\theta_{\mu\nu}} = \prod_{r,\mu\nu} e^{\frac{1}{2g^2}\cos\theta_{\mu\nu}(r)} = \prod_{r,\mu\nu} e^{\frac{1}{4g^2}(e^{i\theta_{\mu\nu}(r)}+e^{-i\theta_{\mu\nu}(r)})} = \prod_{r,\mu\nu}\sum_n \frac{1}{n!}\frac{1}{(4g^2)^n}\left(e^{i\theta_{\mu\nu}(r)}+e^{-i\theta_{\mu\nu}(r)}\right)^n,$$

now remeber that $\theta_{\mu\nu}(r)$ is the sum of four $\theta$ variables on the four links around a plaquette. Since $1/4g^2 \ll 1$, we may want to just look at the $n = 0$ and $n = 1$ terms: we are forced to look at $n = 1$ terms because of the term $e^{i\oint_C \theta_\mu(r)}$ in the correlation function: we must pair it with $e^{-i\oint_C \theta\mu(r)}$ which comes from $n = 1$ terms of the expansion of the action. Notice the number of $n = 1$ terms equals the number of plaquettes inside the contour: the cartoon picture is simply that we are drawing little squares inside the contour, until the little squares can tile up the entire contour; each little square is contributed by a $n = 1$ term in the action which is accompanied with a $1/4g^2$; and apparently the number of squares is roughly the area $A$ inside the contour, which gives $(1/4g^2)^A$. To summarize, we have

$$\left\langle e^{i\oint_C \theta_\mu(r)}\right\rangle \simeq \left(\frac{1}{4g^2}\right)^A = e^{-A\ln 4g^2}.$$

What is important but left unproved is that the high-temperature expansion has a finite radius of convergence. Therefore if we also include the higher $n$ contributions, we should have

$$\left\langle e^{i\oint_C \theta_\mu(r)}\right\rangle \simeq \left(\frac{1}{4g^2}\right)^A = e^{-Af(g^{-2})}$$

where $f(g^{-2})$ is a finite function for $g^2 \ll 1$ with leading term $\ln 4g^2$. The detailed form of $f$ depends on the details of the model, but as we see, at zeroth order the argument and result is the same for both Abelian lattice gauge theory and Ising lattice theory.

Then, suppose $J \gg 1$, or $g^2 \ll 1$, i.e. the lattice coupling constant is small. We guess without proof that a naive continuum limit could be taken and the action could be replaced by the Gaussian approximation, which in 4D is simply the Maxwell action for electromagnetism, $\frac{1}{4}\int d^4x F_{\mu\nu}^2$. Therefore we write, at small $g^2$,

$$\left\langle e^{i\oint_C \theta_\mu(r)}\right\rangle \to \left\langle e^{ig\oint_C A_\mu dx}\right\rangle = \frac{1}{Z}\prod_{r,\mu}\int dA_\mu(r) e^{-\frac{1}{4}\int d^4x F_{\mu\nu}^2 + ig\oint A_\mu dx},$$

where the integral of $A_\mu$ goes from $-\infty$ to $\infty$. We then follow Kogut to point out two technical problems of this equation: first is that the right hand side can be evaluated only after a particular gauge is chosen, and here we will choose Feyman's gauge; the second is that, when evaluating the right hand side, the gauge field propagator will appear; but note that we cannnot directly use the continuum propagator since we are after all doing lattice gauge theory and hence a lattice cutoff exists; what we should use is a lattice propagator. Considering that the lattice propagator and the continuum propagator only differ much at lattice cutoff, we will simply use the continuum propagator while separating the (short distance) part which must be substituted by a lattice propagator: i.e. we substitute $\langle A_\mu(x)A_\nu(0)\rangle = \delta_{\mu\nu}\Delta(x)$ by

$$\Delta(x-y) = \Delta(0)\delta_{x,y} + \widetilde{\Delta}(x-y), \qquad \widetilde{\Delta}(x) = \begin{cases} \frac{1}{2\pi^2}\frac{1}{|x|^2}, & |x| > a, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

A digression: let us review some basics of path integrals: note by virtue of $ax^2 + bx = a(x + b/2a)^2 - b^2/4a$, we have $-\frac{1}{2}ax^2 + bx = -\frac{a}{2}(x-b)^2 + b^2/2a$. We have adjusted the coefficient of the quadratic term to be $\frac{1}{2}a$ due to the following reason: note the Gaussian integral is then

$$\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}ax^2+bx} \propto e^{\frac{b^2}{2a}},$$

The lesson from here is that: the propagator is $\Delta = a^{-1}$, and if we write the Gaussian term as half times $\frac{1}{2}\psi\Delta^{-1}\psi$, and if $\psi$ couples to $J$, then after integrating $\psi$ we get $\frac{1}{2}J\Delta J$. Therefore we see that, as long as write the Gaussian term with the $1/2$ in the coefficient, we can simply write the result after the integral, simply replacing $\psi$ in the gaussian term by $J$, and replacing $\Delta^{-1}$ by $\Delta$, and the $1/2$ in the coefficient is retained. Therefore, we have

$$\int D\psi e^{\int -\frac{1}{2}\psi\Delta^{-1}\psi + J\psi} \propto e^{\int \frac{1}{2}J\Delta J}.$$

The only thing one should be careful is the final sign, because in practice we do not always have the above sign choice; in the above, we have $-\frac{1}{2}ax^2 + bx$, but we may also have $\frac{i}{2}ax^2 + ibx$, $-\frac{1}{2}ax^2 + ibx$, and so on. Then one simply have to adjust the result from $\frac{b^2}{2a}$ to $\frac{i}{2}\frac{b^2}{a}$, $-\frac{b^2}{2a}$, and so on.

Next we review some basics about photon propagators. Note the rigorous derivation takes care of the complication due to gauge choices, see e.g. Srednicki for detail; but here we just give a very naive derivation. We have $e^{-S}$, where $S = \frac{1}{4}\int d^4x F_{\mu\nu}^2 = \frac{1}{4}\int d^4x(\partial_\nu A_\mu - \partial_\mu A_\nu)^2$, define $A_\mu(x) = \int d^4k e^{ik_\nu x_\nu} A_\mu(k)$ we have $\frac{1}{4}\int d^4x F_{\mu\nu}^2 = \frac{1}{4}\int d^4x \int d^4k \int d^4k' e^{ikx}(k_\nu A_\mu(k) - k_\mu A_\nu(k))e^{ik'x}(k'_\nu A_\mu(k') - k'_\mu A_\nu(k')) = \frac{1}{4}\int d^4k(k_\nu A_\mu(k) - k_\mu A_\nu(k))(-k_\nu A_\mu(-k) + k_\mu A_\nu(-k)) = \frac{1}{4}\int d^4k A_\mu(k)A_\mu^*(k)(-2k^2) + 2\int d^4k A_\mu(k)A_\nu^*(k)k_\mu k_\nu = \frac{1}{2}\int d^4k A_\mu(k)\Delta_{\mu\nu}^{-1}(k)A_\nu^*(k)$, where $\Delta_{\mu\nu}^{-1}(k) = \delta_{\mu\nu}k^2 - k_\mu k_\nu$, therefore $\Delta_{\mu\nu}(k) = \frac{1}{k^2}P_{\mu\nu}$ where $P_{\mu\nu} = \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}$, or more rigorously we should write $\Delta_{\mu\nu}(k) = \frac{g_{\mu\nu} - k_\mu k_\nu/k^2}{k^2 - i\epsilon}$, and therefore $\Delta_{\mu\nu}(x-y) = \int \frac{d^4k}{(2\pi)^4} e^{ik(x-y)} \frac{P_{\mu\nu}}{k^2 - i\epsilon}$. We now want to evaluate $\Delta_{\mu\nu}(k)$, but this is constrained by the gauge fixing we use. The detail can be found in Srednicki; here we just say that due to $\int \frac{d^3k}{(2\pi)^3}\frac{e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})}}{\boldsymbol{k}^2} = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}$, we will get the usual Coulomb interaction (if we use Coulomb gauge).

Now go back to our problem. One immediately notices that our scaling of $\Delta(x-y)$ is different from the Coulomb law: from Eq. (16) we see that we have $\Delta(x-y) \sim \frac{1}{|x-y|^2}$, but the Coulomb law gives $\Delta(x-y) \sim \frac{1}{|x-y|}$, as mentioned above. The technical reason is that for our case, we are doing the integral $\int_0^\infty k^3 dk \frac{e^{ikr}}{k^2} \sim 1/r^2$, and the Coulomb case, due to the Coulomb gauge, we are doing the integral $\int_0^\infty k^2 dk \frac{e^{ikr}}{k^2} \sim 1/r$. The physical reason should be answered when I come back to look at this issue once again.

Now go back to our low temperature correlation function $\left\langle e^{ig\oint_C A_\mu dx}\right\rangle$: from the gaussian integral we just reviewed, we see that here only for the sites on the contour $C$ do we have a linear term, and therefore only these terms will affter the result after the Gaussian integral on $A$. Therefore we can directly write the result

$$\left\langle e^{ig\oint_C A_\mu dx}\right\rangle = e^{-\frac{1}{2}g^2\oint_C\oint_C \Delta(x-y)_{\mu\nu}dx_\mu dy_\nu} = e^{-\frac{1}{2}g^2\oint_C\oint_C \Delta(x-y)dx_\mu dy_\mu},$$

where we have used the fact that there is a $\delta$ function in the propagator (which again is not proved). The final minus sign is due to the "$i$" appear in the coefficient on the exponential.

We expect to get perimeter law from this equation. If $\Delta(x-y)$ is short range (e.g. if we have a Yukawa potential with some screening distance) then the two contour integrals will only interact when the two contours are close to each other, which can give perimeter law. However now $\Delta$ is power law and not short range, therefore detailed calculation is needed. Now we actually evaluate the double integrals on the exponent: we start by choosing for ourselves a contour along which the calculation is easy: we choose a rectangular contour, and call the two directions $\tau$ and $r$. Notice that since the integral element is $dx_\mu dy_\mu$ with the same $\mu$, we see that the lines along which $x$ and $y$ changes must not be perpendicular; and therefore we are only left with these cases: call the four sids of the rectangular contour $I, II, III$ and $IV$, then we must have $(x,y) \in \{I,I\}, \{II,II\}, \{III,III\}, \{IV,IV\}, \{I,III\}, \{II,IV\}$, i.e. $(x,y)$ are on either the same side or the opposite sites. In the first case, say we just look at $(x,y) \in \{I,I\}$ which is a $\tau$ direction, the integral gives

$$I_{I,I} = \int_0^T\int_0^T \delta_{|x-y|>a}\widetilde{\Delta}(x-y)dx_\mu dy_\mu = 2\int_0^T dy\int_0^{y-a} dx\widetilde{\Delta}(x-y) = \frac{2}{2\pi^2}\int_0^T dy\int_0^{y-a} d\frac{1}{(y-x)^2} = \frac{1}{\pi^2}(T/a - \ln(T/a)),$$

and the second case, say we just look at $(x,y) \in \{I,III\}$ which again are both along $\tau$ direction, we have

$$I_{I,III} = \int_0^T\int_T^0 \widetilde{\Delta}(x-y)dxdy = -\frac{1}{2\pi^2}\int_0^T dx\int_0^T dy\frac{1}{R^2+(x-y)^2} = -\frac{1}{\pi^2}\left[\frac{T}{R}\arctan\frac{T}{R} - \frac{1}{2}\ln(1+\frac{T^2}{R^2})\right],$$

Note the first integral is very easy to evaluate; the second integral can be evaluated by Mathematica (with assumptions that $T>0, R>0$). We specialize to the case $T \gg R$, and we finally have $\oint_C\oint_C \Delta(x-y)dx_\mu dy_\mu \simeq (\Delta(0) + \frac{1}{\pi a^2})P - (1/\pi)T/R - 4/\pi^2\ln(R/a)$, where $P = 2(T+R)$ is the perimeter of $C$, therefore we have

$$\left\langle e^{ig\oint_C A_\mu dx}\right\rangle \simeq e^{-\frac{1}{2}g^2cP + \frac{g^2}{2\pi}\frac{T}{R} + \frac{2g^2}{\pi^2}\ln\frac{R}{a}},$$

we see that the long-range character of the massless propagator $\Delta(x)$ has generated additional long-range effects in addition to the perimeter law. Our core claim, that the correlation behavor at weak and strong coupling is different, can now be clearly seen. We admit that the weak coupling limit is rather naive, but this result strongly suggests that the Abelian lattice gauge theory in 4D is a two-phase system.

After this, Kogut went on to show actually the gauge invariant correlation function (i.e. Wilson loop) actually measures the force law between static charged quanta which can be placed into the system. This can be seen by noting that $\oint_C A_\mu dx_\mu = \int d^4x \delta^\mu(x \in C)A_\mu$, where the $\delta$ function means that it picks nonzero value whenever $x \in C$. This way, we can regard the $\delta$ function as some kind of external current $J_\mu(x) = \delta^\mu(x \in C)$, Therefore

$$\left\langle e^{ie\oint_C A_\mu dx_\mu}\right\rangle = \left\langle e^{ie\int A_\mu(x)J_\mu(x)d^4x}\right\rangle = Z(J)/Z(0),$$

where $Z(J)$ is simply the partition function in the presence of the coupling term between the external current $J_\mu$ and the potential $A_\mu(x)$.

Kogut then went on to use the correspondence between free energy density and ground-state energy density to argue that we can actually read off the force law $V(r)$ between the static charged quanta, which is given by

$$V(R) = -\lim_{T\to\infty} \frac{1}{T} \ln\langle e^{-ie\oint A_\mu dx_\mu}\rangle,$$

and conclude that at strong coupling we have the area law $V(R) \sim R$, which leads to quark confinement, and at weak coupling we have $V(R) \sim$ const., which means the quarks are deconfined, and what's more from the concrete calculation above we have $V(R) \sim$ const. $- e^2/R$, i.e. even Coulomb's law can be recovered.

The above completes the analysis of the 4D (classical) Abelian lattice gauge theory. The 3D (classical) Abelian lattice gauge theory is not studied in the review. Now we look at the 2D (classical) Abelian lattice gauge theory, the gauge invariant correlation function reads

$$\langle e^{i\,\overline{\sum}_C\,\theta_\mu}\rangle = \frac{1}{Z}\int \prod_{n,\mu} d\theta_\mu(n) e^{\beta\sum_{n,\mu\nu}\cos\theta_{\mu\nu} + i\,\overline{\sum}_C\,\theta_\mu}.$$

Note we can write $\overline{\sum}_C\,\theta_\mu(n) = \sum_{\{P\}\in C}\theta_{\mu\nu}(n)$, which is simply the Storke's theorem, therefore the action is only a function of $\theta_{\mu\nu}$. Now the integral variable is still $\theta_\mu(n)$. We try to see if we can also express $\theta_\mu(n)$ in terms of $\theta_{\mu\nu}(n)$, under proer choce of gauge.

This can be achieved: we choose the temporal gauge so that $\theta_\tau(n) = 0$, then actually we have $\theta_x(n) = \sum_{j=0}^{\infty}\theta_{x(n+j\tau)} - \theta_{x(n+(j+1)\tau)} = \sum_{j=0}^{\infty}\theta_{\mu\nu}(n+j\tau)$, therefore we have successfully written $\theta_x(n)$ also in terms of $\theta_{\mu\nu}(n)$, which can serve as new integral variable. To summarize, we have

$$\langle e^{i\,\overline{\sum}_C\,\theta_\mu}\rangle = \frac{\int\prod_{\{P\}\in C} d\theta_{\mu\nu}(n) e^{\beta\sum_{n,\mu\nu}\cos\theta_{\mu\nu} + i\sum_{\{P\}\in C}\theta_{\mu\nu}}}{\int\prod_{\{P\}\in C} d\theta_{\mu\nu}(n) e^{\beta\sum_{n,\mu\nu}\cos\theta_{\mu\nu}}},$$

the most important thing is that it says that the correlation function for each elementary plaquette is actually decoupled, i.e. we have

$$\langle e^{i\,\overline{\sum}_C\,\theta_\mu}\rangle = \langle e^{i\sum_{\{P\}\in C}\theta_{\mu\nu}}\rangle = \prod_{\{P\}\in C}\langle e^{i\theta_P}\rangle = \left(\frac{I_1(\beta)}{I_0(\beta)}\right)^A,$$

where $A$ is the number of elementary plaquettes inside the contour $C$, or the area of $C$, and we have written $\theta_P = \theta_{\mu\nu}$, and we have

$$\frac{I_1(\beta)}{I_0(\beta)} = \frac{\int_0^{2\pi} d\theta_P e^{\beta\cos\theta_P + i\theta_P}}{\int_0^{2\pi} d\theta_P e^{\beta\cos\theta_P}} = \frac{2\pi\mathrm{BesselI}(0,\beta)}{2\pi\mathrm{BesselI}(1,\beta)},$$

the last equality is obtained via Mathematica. Note the denominators (numerators) across the equalities are equal. When $g^2 \gg 1$ we have $\beta = 1/2g^2 \ll 1$ therefore $I_1(\beta)/I_0(\beta) \simeq \beta/1 = 1/4g^2$, this means that the original Wilson loop

$$\lim_{g^2\to\infty}\langle e^{i\,\overline{\sum}_C\,\theta_\mu}\rangle \simeq e^{-\ln(4g^2)A}.$$

when $g^2 \ll 1$, we have $I_1(\beta)/I_0(\beta) \simeq 1 - 1/2\beta = 1 - g^2$, therefore

$$\lim_{g^2\to 0}\langle e^{i\,\overline{\sum}_C\,\theta_\mu}\rangle \simeq e^{\ln(1-g^2)A} \simeq e^{-g^2 A}.$$

In summary the 2D model confines for all $g$. The strength of the interquark potential is a smooth function of $g$ which extends between $g^2R$ at weak coupling and $\ln(4g^2)R$ at strong coupling. We also see that the 2D Abelian lattice gauge model is exactly solvable, due to the fact that, after gauge fixing, the plaquette variables are related to the remaining link variables (which cannot be fixed even after all gauge fixing) by linear transformation, and therefore one can choose the plaquette variables as independent variables. Note in 3D, we actually have $\sum_{P\in\mathrm{Cube}}\theta_P = 0$, i.e. the sum of the six plaquette variables on the six faces of the cube vanishes, showing that all the plaquette variables cannot be independent. Therefore the 2D method leading to exact solution cannot be generalized to higher dimensions.

Note that in the above derivation, in choosing the temporal gauge we have the correspondence that the 2D Abelian lattice gauge theory corresponds to 1D planar spin model.

Finally, we use transfer matrix method to study the connection between the Abelian lattice gauge theory and the $\tau$-continuum quantum Hamiltonian. First, separate the action for the Abelian lattice gauge theory into two terms, one involving only plaquettes along $\tau$ direction, and the other involving plaquettes just in the spatial direction:

$$S = \beta_\tau\sum_{n,k}(1-\cos\theta_{0k}(n)) - \beta\sum_{n,jk}\cos\theta_{jk}(n),$$

where we have labeled the $(d-1)$ spatial directions by $j,k$, and temporal direction $\tau$ with index 0. Note we have also assigned anisotropic couplings as before, which will result in $\beta \sim \delta\tau \sim e^{-2\beta_\tau} \to 0$ in the quantum (i.e. $\tau$-continuum) limit. We then choose the gauge $\theta_0(n) = 0$, and the $\tau$-independent gauge transformations are still manifest local symmetries of the system (i.e., there are still gauge freedoms unfixed, which are those which transform each spatial layer in the same way, i.e. $\tau$-independent). This way, $\theta_{0k}(n) = \theta_k(n+\tau) - \theta_k(n)$. In the $\tau$-continuum limit, $\theta_{0k}(n)$ becomes small and slowly vary, therefore we have $1 - \cos\theta_{0k} \simeq \frac{1}{2}\theta_{0k}^2 \simeq \frac{1}{2}a_\tau^2(\partial\theta_k/\partial\tau)^2$, where $a_\tau$ denotes the lattice spacing in the $\tau$ direction. We then have $\sum_{n,k} \to \frac{1}{a_\tau}\int d\tau \sum_{m,j}$ where $j$ is just along spatial direction, and $m$ is some spatial site. The action becomes

$$S \to \int d\tau \left( \frac{1}{2}\beta_\tau a_\tau \sum_{m,j} \dot{\theta}_j^2(\tau, m) - \frac{1}{a_\tau}\beta \sum_{m,jk} \cos\theta_{jk}(\tau, m) \right).$$

Therefore to obtain a sensible limit, $\beta_\tau$ must scale as $a_\tau^{-1}$ and $\beta$ as $a_\tau$. The constant of proportionality will be identified with the coupling constant $g^2$ which is held fixed, $\beta_\tau = g^2/a_\tau \to \infty$, and $\beta = a_\tau/g^2 \to 0$. Now we write the action in terms of a Hamiltonian: to do so we must set up a Hilbert space for each fixed $\tau$ slice. We promote $\theta_k(m)$ to a quantum field, and define its conjugate momentum $L_j(m)$ with

$$[L_j(m), \theta_{j'}(m')] = i\delta_{jj'}\delta_{mm'}.$$

And the quantum Hamiltonian, after the transfer matrix derivation, gives

$$aH = \sum_{m,j} \frac{1}{2}g^2 L_j^2(m) - \frac{1}{g^2}\sum_{m,jk} \cos\theta_{jk}(m).$$

Define $G(\chi) = e^{i\sum_m \sum_{\pm j} L_j(m)\chi(m)} = \prod_m e^{i\sum_{\pm j} L_j(m)\chi(m)} = \prod_m G_m(\chi)$, we have

$$G(\chi)\theta_j(m)G^{-1}(\chi) = \theta_j(m) + \chi(m+j) - \chi(m) = \theta_j + \Delta_j\chi,$$

we see that we have

$$G(\chi)HG^{-1}(\chi) = H.$$

Then we define $\theta_j(m) = agA_j(m)$ and $E_j(m) = (g/a^2)L_j(m)$, we have

$$[E_i(r), A_j(r')] = i\delta_{ij}\delta(r - r').$$

Note since $L_j(m)$ is an angular momentum operator, i.e. it is conjugate to the angular variable $\theta_j(m)$, its spectrum is discrete: $L_j(m) = 0, \pm1, \pm2, ...$, which implies the electric flux on a link, $a^2E_j(m)$, is quantized in units of the charge $g$. The Hamiltonian becomes

$$H = a^2 \sum_{m,j} \frac{1}{2}E_j^2(m) - \frac{1}{g^2 a}\sum_{m,jk} \cos\theta_{jk}(m).$$

For the second term we can define $\theta_{jk} = a^2 g B_i$ where $(i,j,k) = (x,y,z)$,$(y,z,x)$ or $(z,x,y)$, therefore we have, after expanding cosine,

$$H = a^3 \sum_{m,j} \frac{1}{2}E_j^2(m) + \frac{1}{2}B_j^2(m).$$

# 6 XY model and KT transition

We follow Nagaosa's book and Kogut's review paper. The XY model is written in Eq.(32). Here we specify the model to a 2D square lattice, and write it as

$$H = -J\sum_{\langle ij\rangle} \cos(\theta_i - \theta_j).$$

It is obtained from 2D planar (i.e. XY) Heiseiberg model by using $e^{\pm i\theta} = S^\pm$. We know look at the spin-spin correlation function at high and low temperature $T$ to show that it has two phases:

$$\langle e^{i\theta_i} e^{-i\theta_j}\rangle = \frac{1}{Z}\int \prod_m \theta_m e^{i(\theta_i - \theta_j)} e^{-\frac{J}{k_B T}\sum_{kl}\cos(\theta_k - \theta_l)},$$

- At high temperature, $T$ is large and we can expand according to the small parameter $J/k_BT$. The zeroth order term is just 1 which vanishes due to the integral $\int_0^{2\pi} d\theta_m e^{i\theta_m} = 0$ for $m = i$ or $j$. The first nonzero order is contributed by paths connecting $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$: each bond $\langle kl\rangle$ on this path corresponds to a phase $e^{i(\theta_k-\theta_l)}$, and all these phases along the bond will cancel (the first and the last will cancel $e^{i\theta_i}$ and $e^{-i\theta_j}$, respectively). Each bond will contribute a factor of $2\pi J/k_BT$ due to $\int_i^{2\pi} d\theta = 2\pi$. Note there are infinite such paths, but the dominant ones (least powers of $1/T$) are the ones are the shorted paths, and we get $\langle e^{i\theta_i} e^{-i\theta_j}\rangle \sim \left(\frac{J}{2k_BT}\right)^{|\boldsymbol{R}_i-\boldsymbol{R}_j|}$.
- At low temperature, thermal fluctuation should be small, suggesting $\theta_i$ varies slowly and smoothly in the system, therefore it should be valid to expand the cosine to $\Delta\theta_i$. We get $\langle e^{i\theta_i} e^{-i\theta_j}\rangle = |\boldsymbol{R}_i - \boldsymbol{R}_j|^{-k_BT/2\pi J}$.

Two things we leared: low temperature there is no magnetization (i.e. $\langle e^{i\theta_i}\rangle \to 0$), which is in accord with Mermin-Wagner theorem; second, the low energy phase is critical (algebraic). This phase support another form of excitation, not considered in the conventional Ginzburg-Landau theory: vortices. In the continuum limit $H = \frac{J}{2}\int(\nabla\theta)^2 d^2\boldsymbol{R}$, whose equation of motion $\nabla^2\theta = 0$ supports vortex configuration $\theta = \arctan y/x$. We have $\nabla\theta = \frac{(-y,x)}{R^2}$, $\nabla^2\theta = 0$, therefore $E_{\text{vortex}} = \frac{J}{2}\cdot\int_0^{2\pi} d\theta \int d_a^{R_c} RdR \cdot \frac{1}{R^2} = \pi J \ln R_c/a$, where $a$ is lattice spacing and $R_c$ system size. Free energy $F = U - TS$ where $S = \ln W$, where $W$ is number of configurations of putting one vortex in the sample which is roughly $W = R_c^2/a^2$, therefore $F = (\pi J - 2T)\ln R_c/a$. This is the KT picture: at low temperature, the only low energy excitations are spin waves (not Goldstone modes! since in 2D the continuous symmetry is not broken), vortex-antivortex pair can appear but they are in bound state; free vortex cannot appear. As temperature grows, size of the vortex-antivortex pair grows until the size diverges at $T_c$.

The above picture/argument is coarse, in the sense that only one vortex is discussed. Below we will use the duality transfomation (periodic Gaussian model) to give a full analysis. We will find: 1. vortex sector and spin wave sector decouple (i.e. with no interaction); the vortex sector is equivalent to a 2D Coulomb gas (Sine-Gordan model).

We use the following approxiation (due to Jacques Villain) and Poisson summation formula

$$e^{\lambda\cos\theta} \simeq \sum_{m=-\infty}^{\infty} e^{\lambda} e^{-\frac{\lambda}{2}(\theta-2\pi m)^2}$$

and

$$\sum_{m=-\infty}^{\infty} h(m) = \sum_{l=-\infty}^{\infty}\int_{-\infty}^{\infty} d\phi h(\phi) e^{2\pi i l\phi},$$

combine the two and do the integral for $\phi$ we get $[-\frac{\lambda}{2}(2\pi\phi-\theta)^2 + il(2\pi\phi-\theta) + il\theta = -\frac{\lambda}{2}(2\pi\phi-\theta-il/\lambda)^2 - l^2/2\lambda + il\theta$ ], note that the Gaussian integral gives $\sqrt{\pi}$ and that we are doing integral variable substitution $\phi \to \sqrt{\lambda 2\pi^2}\phi = \phi'$,

$$e^{\lambda\cos\theta} \simeq \frac{1}{\sqrt{2\pi\lambda}} \sum_{l=-\infty}^{\infty} e^{\lambda} e^{il\theta} e^{-l^2/2\lambda}. \tag{17}$$

Note another way of showing this approximate identity, according to Kogut, is that we have the rigorous Fourier decomposition

$$e^{\lambda\cos\theta} = \sum_{l=-\infty}^{\infty} e^{il\theta} I_l(\lambda)$$

where $I_l(\lambda)$ is Bessel function (of some kind), and we have the limiting form $\lim_{\lambda\to\infty} I_l(\lambda) = \frac{1}{\sqrt{2\pi\lambda}} e^{-l^2/2\lambda}$ of Gaussian form. This formula will be used all the time. Now we start from the original definition of the partition function

$$Z = \int d\theta_1 \cdots \int d\theta_N e^{\beta J\sum_{\langle ij\rangle}\cos(\theta_i-\theta_j)},$$

using the formula above we get

$$Z = \int d\theta_1 \cdots \int d\theta_N \sum_{\{l_{ij}\}} e^{\sum_{\langle ij\rangle} il_{ij}(\theta_i-\theta_j) - l_{ij}^2/2\beta J},$$

by writing $\sum_{\langle ij\rangle} l_{ij}(\theta_i-\theta_j) = \sum_i \theta_i \sum_\mu l_{i,\mu}$, the integral over $\theta$'s can be evaluated and we get a constraint in $l_{i,\mu}$ which is $\sum_\mu l_{i,\mu} = 0$, this is saying that $\text{div} l = 0$ i.e. $l$ is divergenceless. We now have

$$Z = \sum_{\{l_{ij}\}} e^{-l_{ij}^2/2\beta J}\delta_{\text{div} l_i = 0},$$

Analgous to the continuous case $\boldsymbol{\nabla}\cdot\boldsymbol{B}=0\Rightarrow\exists\boldsymbol{A}$ s.t. $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$ and that $\boldsymbol{\nabla}\cdot\boldsymbol{B}=\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{A})=0$ satisfies automatically, now we there should also exists $n_i$ living on the sites such that the condition $\mathrm{div}l=0$ is automatically satisfied. Suppose the $x$ bond has the condition $l_{i,x}=n(i+y)-n(i)$. Note the lesson we now learn is that the $l_{i,y}$ condition is entirely constained by the $l_{i,x}$ condition, and must be $l_{i,y}=-n(i+y)+n(i+y-x)$ [e.g. it cannot be $l_{i,y}=n(i)-n(i+x)$ or anything other than that]. According to these definition we have $l_{i,-y}=-l_{i-y,y}=-(-n(i)+n(i-x))$, and $l_{i,-x}=-l_{i-x,x}=-(n(i-x+y)-n(i-x))$, it is easy to check that $\mathrm{div}l=0$ is automatically satisfied.

Now that we have find a way to get rid of the $\delta$ function using the new variable $n$, we write $Z$ in terms of the new variables:

$$Z=\sum_{\{n_i\}}e^{-\frac{1}{2\beta J}\sum_{i,\mu}(n(i)-n(i-\mu))^2},$$

now we again use Poisson summation formula and write

$$Z=\int_{-\infty}^{\infty}\prod_i d\phi_i\sum_{\{m(i)\}=-\infty}^{\infty}e^{-\frac{1}{2\beta J}\sum_{i,\mu}(\Delta_\mu\phi(i))^2+2\pi i\sum_i m(i)\phi(i)},$$

Kogut then claims that, upon some analysis, we should identify $\phi(i)$ with the spin waves of the originl model and $m(i)$ with its vortices. We can now do Fourier transform to $\phi$: define $\phi_i=\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}$, we have $\Delta_x\phi(i)=\sum_{\boldsymbol{k}}(e^{ik_x}-1)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}$, and therefore $\sum_i(\Delta_x\phi(i))^2=\sum_i\sum_{\boldsymbol{k}}(e^{ik_x}-1)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}\sum_{\boldsymbol{k}'}(e^{ik'_x}-1)e^{i\boldsymbol{k}'\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}'}=\sum_{\boldsymbol{k},\boldsymbol{k}'}(e^{ik_x}-1)(e^{ik'_x}-1)\delta_{\boldsymbol{k},-\boldsymbol{k}'}\phi_{\boldsymbol{k}}\phi_{-\boldsymbol{k}}=\sum_{\boldsymbol{k}}(e^{ik_x}-1)(e^{-ik_x}-1)\phi_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}=\sum_{\boldsymbol{k}}(2-2\cos k_x)\phi_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}$, therefore $\sum_{i,\mu}(\Delta_\mu\phi(i))^2=\sum_{\boldsymbol{k}}(4-2\cos k_x-2\cos k_y)\phi^*_{\boldsymbol{k}}\phi_{\boldsymbol{k}}$. Therefore the exponential becomes $A\equiv-\frac{1}{2\beta J}\sum_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}(4-2\cos k_x-2\cos k_y)\phi_{\boldsymbol{k}}+2\pi i\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\sum_i m(i)\phi_{\boldsymbol{k}}$, using the $ax^2+bx=a(x+b/2a)^2-b^2/4a$, we have $A=\sum_{\boldsymbol{k}}-\frac{4-2\cos k_x-2\cos k_y}{2\beta J}|(\phi_{\boldsymbol{k}}+\frac{2\pi i\sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}m(i)}{-\frac{1}{\beta J}(4-2\cos k_x-2\cos k_y)})|^2-\frac{(2\pi i\sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}m(i))^2}{-\frac{2}{\beta J}(4-2\cos k_x-2\cos k_y)}$, and after integrating $\phi_{\boldsymbol{k}}$, we get

$$Z=Z_{\text{Spin Wave}}\cdot\sum_{\{m(i)\}=-\infty}^{\infty}e^{-2\pi^2\beta J\sum_i\sum_{i'}m(i)\sum_{\boldsymbol{k}}\frac{e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_{i'})}}{4-2\cos k_x-2\cos k_y}m(i')},$$

or we define the spin wave propagator $G(r)=\sum_{\boldsymbol{k}}\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{4-2\cos k_x-2\cos k_y}=\int_{\mathrm{BZ}}\frac{d^2\boldsymbol{k}}{(2\pi)^2}\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{4-2\cos k_x-2\cos k_y}$, we have

$$Z=Z_{\text{Spin Wave}}\cdot\sum_{\{m(i)\}=-\infty}^{\infty}e^{-2\pi^2\beta J\sum_i\sum_{i'}m(i)G(i-i')m(i')}.$$

where $Z_{\text{Spin Wave}}$ is the part after integrating $\phi_{\boldsymbol{k}}$. We see that the spin wave part and the vortex part decouples. Since spin wave does not give transition, the transtion must be induced by the vortex part.

At large enough $r$, we have $G(r)\simeq-\frac{1}{2\pi}\ln\frac{r}{a}-\frac{1}{4}$; at $r=0$, we have $G(0)\simeq\frac{1}{2\pi}\ln\frac{R}{a}$. The former shows that vortices interact through a logarithmic potential; the latter is consistent with the self energy of a single vortex. Then it is better to introduce $G(r)=G'(r)+G(0)$ where $G'(r=0)=0$. This way, the exponential of $Z$ gives $-2\pi^2\beta J$ times $[\sum_i m(i)]^2G(0)+\sum_i\sum_{i'\neq i}m(i)G'(i-i')m(i')$, wheere the first term indicates that the biggest contribution to $Z$ is the configurations with $\sum_i m(i)=0$, i.e. the vortices are neutral. We can therefore write

$$Z=Z_{\text{Spin Wave}}\cdot\sum_{\{m(i)\}=-\infty}^{\infty}{}'e^{-2\pi^2\beta J\sum_{i,i'\neq i}m(i)G'(i-i')m(i')},$$

where the sum $\sum'$ means neutral configurations of vortices only. We now plug in the asymptotic form of $G(r)=-\frac{1}{2\pi}\ln\frac{r}{a}-\frac{1}{4}$: we get

$$Z=Z_{\text{Spin Wave}}\cdot\sum_{\{m(i)\}=-\infty}^{\infty}{}'e^{\frac{\pi^2\beta J}{2}\sum_{i,i'\neq i}m(i)m(i')+\pi\beta J\sum_{i,i'\neq i}m(i)\ln\frac{|r-r'|}{a}m(i')},$$

we can use $0=\sum_{i,i'}m(i)m(i')=\sum_{i,i'\neq i}m(i)m(i')=\sum_i m^2(i)$ to get

$$Z=Z_{\text{Spin Wave}}\cdot\sum_{\{m(i)\}=-\infty}^{\infty}{}'e^{-\frac{\pi^2\beta J}{2}\sum_i m^2(i)+\pi\beta J\sum_{i,i'\neq i}m(i)\ln\frac{|r-r'|}{a}m(i')},$$

The first term gives the chemical potential of each vortex and the second term gives the logarithmic interactions between different vortices. This is exatly the partition function for Coulomb gas in 2D. The phases of it is known: when $\beta$ is

large (low temperature), the chemical potential suppresses the vortices therefore there is only spin waves; the logarithmic potential says that, at low temperature, even if vortices populate, they appear in vortex-antivortex pairs (bound state). As temperature is raised, vortices are not suppressed; they are still not free but the interactions are effectively screened.

In fact, we can follow Savit and generalize this to any lattice theory with a U(1) invariance. Savit has done this rigorously for the d=3 planar (XY) model in a 1978 paper. But here we just sketch the general form that he give in his review paper: on a $d$ dimensional lattice, remember we define the simplex number $s$ to be the dimension that each term occupies (and therefore Ising/XY model has $s=1$ which is the bond dimension and Ising gauge/abelian gauge has $s=2$ which is the plaquette dimension). On each simplex of dimension $s-1$, place a complex phase $Q_{\mu_1...\mu_{s-1};i}=e^{i\theta_{\mu_1\cdots\mu_{s-1};i}}$, these phases interaction according to the form

$$I_{\mu_1..\mu_s;i}=\cos\left(\frac{1}{(d-s)!}\epsilon_{\mu_1...\mu_s\alpha_1...\alpha_{d-s}}\epsilon_{\alpha_1..\alpha_{d-s}\beta\gamma_1...\gamma_{s-1}}\Delta_\beta\theta_{\gamma_1..\gamma_{s-1};i}\right)=\cos(\epsilon\epsilon\Delta\theta),$$

where the $\epsilon$ is just the fully antisymmetric tensor. Now we try to understund this term a little better: note $i$ is a $d$ dimensional vector which labels the site in the $d$ dimensional space/spacetime lattice; for each site $i$, in order to define the variable $\theta_{\mu_1...\mu_{s-1};i}$ which lives on a $s-1$ simplex (i.e. boundary of the $s$ simplex), we just have to define the variable direction, which is specified by $s-1$ directions, $\mu_1,...,\mu_{s-1}$. Now, remember that in the $XY$ model, for each site $i$, there are actually two terms, depending on the direction of the bond $x$ or $y$, where bond is simply the $s=1$ simplex. This means that, in our generalization, fixing $i$ we can also obtain differen terms; and these terms are distinguished by there simplex directions $\mu_1...\mu_s$, and each term is exactly what we write down here, $I_{\mu_1...\mu_s;i}$. Therefore we stress that, each term corresponds to one simplex (i.e. for any chosen $s$ simplex, a term should be written down). Let us understand the part in the parenthesis better: as an example, in the usual Abelian lattice gauge theory, $s=2$, and the variable lives on bonds, i.e. $\theta_{\mu;i}$, where $i$ is the starting point of the bond and $\mu$ is the direction of the bond. Then, $\Delta_\beta\theta_{\gamma_1...\gamma_{s-1};i}$ takes derivative along direction $\beta$, where $\beta$ is restricted to the only free direction in the $s$ simplex after fixing $\gamma_1...\gamma_{s-1}$, in other hand, we have

$$\Delta_\beta\theta_{\gamma_1...\gamma_{s-1};i}=\theta_{\gamma_1...\gamma_{s-1};i+\gamma_s}-\theta_{\gamma_1...\gamma_{s-1};i},\qquad \beta=\gamma_s.$$

Note the introduction of $\epsilon\epsilon$ sums over the codimensions, which gives $(d-s)!$ terms; these are redundant and that's why we need to divide by $(d-s)!$. But through this $\epsilon\epsilon$ we can indeed introduce the correct sign convention when taking derivative $\Delta_\beta$ along different coordinate $\beta$. One then follows the KT transformations to proceed.

# 7 Non-abelian lattice gauge theory

Again we follow Kogut. As summarized by Kogut: the lattice action has simple properties at strong coupling (area law, linear potential) which may be properties of the strong interactions. The continuum quantum theory has simple short-distance characteristics (as lattice constant $a\to 0$ we have $g\to 0$, i.e. the interacting theory is weakly coupled at short distances). The immediate question is whether these two theories are related: the idea case is that they lie on the two ends of one renormalization group trajectory and that the theory of non-Abelian gauge fields enjoys asymptotic freedom on fine lattices and confinement on coarse ones.

The action for SU(2) theory is

$$S=-\frac{1}{2g^2}\sum_{n,\mu\nu}\mathrm{tr}U_\mu(n)U_\nu(n+\mu)U_{-\mu}(n+\mu+\nu)U_{-\nu}(n+\nu)+h.c.$$

where

$$U_\mu(n)=\exp[iB_\mu(n)],\qquad B_\mu(n)=\frac{1}{2}ag\tau_iA^i_\mu(n).$$

We can taylor expand the action to recover Yang-Mills theory $S=\frac{1}{4}\int d^4x(F^i_{\mu\nu})^2$.

Choosing temporal gauge, the 2D SU(2) lattice gauge action can be mapped to an $SU(2)\times SU(2)$ Heisenberg spin chain. Here $SU(2)\times SU(2)$ is due to the fact that any $U_\mu(n)$ lives on the link therefore it admits a $SU(2)$ on one end and another $SU(2)$ on the other end. We know that $SU(2)$ is equivalent to a four-component vector $\vec{s}=(a,b,c,d)$ in the sense of $SU(2)=a+i(b\sigma^1+c\sigma^2+d\sigma^3)$, therefore we have $\mathrm{tr}U_x(n)U_x^{-1}(n+\tau)=2\vec{s}(n)\cdot\vec{s}(n+\tau)$, which is disordered at all temperature due to Mermin-Wagner.

Kogut then talked about renormalization group flow of the $O(3)$ model: $Z=\int\prod_n d\vec{s}(n)\exp(-\frac{1}{2g}\int\partial_\mu\vec{s}\cdot\partial_\mu\vec{s}d^2r)$. Then he gives the generalized version of the RG equation, which is $a\frac{dg}{da}=\frac{1}{2\pi}(N-2)g^2$ for $O(N)$ model. In fact, Altland & Simons has a more generalized version (Sec. 8.5): For $S[g]=\frac{1}{\lambda}d^dr\mathrm{tr}[\nabla g\nabla g^{-1}]$, we have $\frac{d\ln\lambda}{d\ln b}\simeq\frac{(N-2)}{4\pi}\lambda-\epsilon$, where $\epsilon=d-2$. We see that it is the same as above if we make the identification $2g=\lambda$ and $a=b$. We will always set $N>2$. We then see that, if $d<2$, the $\beta$ function (i.e. the RHS of the above RH equation for $\lambda$) is always positive. Note that $a=b$ is lattice spacing and therefore is the inverse of momentum cutoff. We see that, $\lambda$ (or $g$) is a decreasing function

of the momentum cutoff, i.e. a smaller coupling must be used with a finer space-time cutoff than with a coarse one. In other words, the long distance properties of the theory (i.e. $a \to 0$) are described by a strongly coupled Action (since $\lambda = 2g \to 0$ but its inverse is large therefore a strongly coupled action).

# 8 Ruben's paper

Relevant papers:

- Charles Stahl, Self-correction from higher-form symmetry protection on a boundary, `https://arxiv.org/pdf/2206.05294.pdf`
- Oliver Buerschaper, Siddhardh Morampudi, Frank Pollmann, Double semion phase in an exactly solvable quantum dimer model on the kagome lattice, `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.90.195148`
- Aleksander Kubica and Beni Yoshida, Ungauging quantum error-correcting codes, `https://arxiv.org/pdf/1805.01836.pdf`
- Yaodong Li, C.W. von Keyserlingk, Guangyu Zhu, Tomas Jochym-O'Connor, Phase diagram of the three-dimensional subsystem toric code, `https://arxiv.org/pdf/2305.06389.pdf`
- Ruben Verresen, Nathanan Tantivasadakarn, and Ashvin Vishwanath, Efficiently preparing Schr¨odinger's cat, fractons and non-Abelian topological order in quantum devices, `https://arxiv.org/pdf/2112.03061.pdf`
- Nathanan Tantivasadakarn, Ryan Thorngren, Ashvin Vishwanath, Ruben Verresen, Long-range entanglement from measuring symmetry-protected topological phases, `https://arxiv.org/pdf/2112.01519.pdf`
- Nathanan Tantivasadakarn, Ruben Verresen, Ashvin Vishwanath, The Shortest Route to Non-Abelian Topological Order on a Quantum Processor, `https://arxiv.org/pdf/2209.03964.pdf`

Any lattice as a graph is made out of 1-cells and 0-cells, which in the language of chain complex is denoted as $\mathcal{C}_1 \xrightarrow{\partial} \mathcal{C}_0$, where here $\mathcal{C}_i = \mathbb{F}_2[e_i^1, e_i^2, ...] = \otimes_{l=1,2,...}\mathbb{F}_2[e_i^l]$ is a vector space over $\mathbb{F}_2$, with basis $e_i^1, e_i^2...$ the $i$-dimensional objects. $\mathcal{C}_0$ and $\mathcal{C}_1$ are directly identified Hilbert spaces of qubits placed on the middle of the bonds or on the sites.

On such a graph there exists an important state, the cluster state. This state is a quantum state living in the Hilbert space $\mathcal{C}_1 \otimes \mathcal{C}_0$. It is defined in the stalizer language: according to the graph connectivity it is natural to consider the following two kinds of operator: the centered-star operator $Z_0 S_1^X$ and the boundary-bond operator $Z_1 L_0^X$, where 0 (1) denotes the fact that this operator lives on the 0-cell (1-cell), and $S_1^X = \prod X_1$ where all the $X_1$ lives on the bonds that connects to $Z_0$, and $L_0^X = X_0 X_0$ where the two $X_0$'s are the boundary of $Z_1$. Note that these two kinds of operators commute with each other, hence these two terms define a commuting projector Hamiltonian. Here we have abused the notation to use the operator to also label the corresponding underlying cell. The cluster state is defined as the state stabilized by all these operators: $Z_0 S_1^X |\psi\rangle = |\psi\rangle$. Now for 2D case, measuring $Z_0$ on all the 0-cells gives a toric code: the measurement on all the 0-cells defines a homomorphism group $\{M_{\mathbf{m}}^0\}$ labeled by measurement outcome $\mathbf{m}$, whose elements $M_{\mathbf{m}}^0$ are maps $M^0 \colon \mathcal{C}_0 \otimes \mathcal{C}_1 \to \mathcal{C}_1$. For each $M_{\mathbf{m}}^0$, the resuting state $|\psi_1\rangle = M_{\mathbf{m}}^0(|\psi\rangle)$ is a toric code state: the toric code state is defined as a state in $\mathcal{C}_1$, stabilized by stars $S_1$ and plaquettes $P_1$. The resulting state $|\psi_1\rangle$ is stabilized by all the $S_1$'s is evident; $|\psi_1\rangle$ is also stabilizied by all $P_1$'s due to the fac that $P_1|\psi_1\rangle = P_1 M_{\mathbf{m}}^0 |\psi_0\rangle = Z_0 S_0^X Z_0 S_0^X Z_0 S_0^X Z_0 S_0^X M_{\mathbf{m}}^0 |\psi_0\rangle = M_{\mathbf{m}}^0 Z_0 S_0^X Z_0 S_0^X Z_0 S_0^X Z_0 S_0^X |\psi_0\rangle = M_{\mathbf{m}}^0 |\psi_0\rangle = |\psi_1\rangle$, i.e. fundamentally due to the fact that measurement commputes with the plaquette operators inside $\mathcal{C}_1 \otimes \mathcal{C}_0$, and that $|\psi_0\rangle$ is stabilized by the boundary bond operators and hence by the plaquette operators. On the other hand, measuring $Z_1$ on all the bonds (i.e. measuring all the 1-cells) defines another homormorphism group $\{M_{\mathbf{m}}^1\}$ whose elements are maps $\mathcal{C}_0 \otimes \mathcal{C}_1 \to \mathcal{C}_0$. The resulting state $|\psi_0\rangle = M_{\mathbf{m}}^1 |\psi\rangle$ is no longer symmetric under local $Z_0$: as $Z_0|\psi_0\rangle = Z_0 M_{\mathbf{m}}^1 |\psi\rangle = M_{\mathbf{m}}^1 Z_0 |\psi\rangle = M_{\mathbf{m}}^1 S_1^X |\psi\rangle$ which does not have fixed value, but it has the global symmetry: $(\prod Z_0)|\psi_0\rangle = M_{\mathbf{m}}^1 (\prod Z_0)|\psi\rangle = M_{\mathbf{m}}^1 (\prod Z_0 S_0^X)|\psi\rangle = M_{\mathbf{m}}^1 |\psi\rangle = |\psi_0\rangle$, and also $|\psi_0\rangle = M_{\mathbf{m}}^1 |\psi\rangle = M_{\mathbf{m}}^1 Z_1 L_0^X |\psi\rangle = L_0^X (-1)^m |\psi_0\rangle$, meaning that $|\psi_0\rangle$ satisfies the property that the wavefunction is symmetric under flipping any two spins on the adjacent 0-cells and at the same time appending a phase $(-1)^m$ that is the measurement outcome on the 1-cell connecting these two 0-cells. This condition suffices to completely recover the state $|\psi_0\rangle$.

Using the cluster state technique the arXiv2011.03061 paper by Ruben et al looked also at preparing the toric code on the diamond lattice, see their Fig. C.2. Again, here one puts qubits on both the bond and sites of the diamond lattice: so strictly speaking $\mathcal{C}_1$ is associated with the pyrohclore lattice (see blue dots in their plot) and only $\mathcal{C}_0$ is the diamond lattice (see red dots in their plot), and the cluster state lives on the diamond-pyrochlore lattice whose Hilbert space is $\mathcal{C}_1 \otimes \mathcal{C}_0$. Using the logic explained above, measuring the diamond lattice qubits, i.e. the red cubits, i.e. those live in $\mathcal{C}_0$, gives the toric code on the "diamond" lattice, which is essentially a toric code whose qubits forms a pyrochlore lattice. Note that the main point of the paper arXiv2011.03061 is to estimate the realistic value of the Wilson-loop correlation function

using cold atoms, where realistic means that there is always longer range interaction than just the "nearest-neighbor" terms $Z_0 S_1^X$ and $Z_1 L_0^X$ that we talk about.

We can further attach 2-cells to the graph, so that we have $\mathcal{C}_2 \to \mathcal{C}_1 \to \mathcal{C}_0$. In the language of chain complex

# 9 1D physics

The following content in this section is copied from `Xiao_project_2`.

We now want to prove $\langle T_\tau e^{i\sum_j (A_j\phi(r_j)+B_j\theta(r_j))}\rangle = e^{-\frac{1}{2}\langle T_\tau[\sum_j (A_j\phi(r_j)+B_j\theta(r_j))]^2\rangle}$:

The term in the exponential can be written: $\sum_j A_j\phi(r_j)+B_j\theta(r_j) = \frac{1}{\beta\Omega}\sum_q [A(q)\phi(-q)+B(q)\theta(-q)]$, this is of course done using $\phi(r) = \frac{1}{\beta\Omega}\sum_q e^{iq\cdot r}\phi(q)$ and $\theta(r) = \frac{1}{\beta\Omega}\sum_q e^{iq\cdot r}\theta(q)$, therefore $\sum_j A_j\phi(r_j)+B_j\theta(r_j) = \frac{1}{\beta\Omega}\sum_q \phi(q)\sum_j e^{iq\cdot r_j}A_j + \frac{1}{\beta\Omega}\sum_q \theta(q)\sum_j e^{iq\cdot r_j}B_j = \frac{1}{\beta\Omega}\sum_q [A(-q)\phi(q)+B(-q)\theta(q)] = \frac{1}{\beta\Omega}\sum_q [A(q)\phi(-q)+B(q)\theta(-q)]$, where we defined $A(-q) = \sum_j e^{iq\cdot r_j}A_j$ and $B(-q) = \sum_j e^{iq\cdot r_j}B_j$. Note that it is defined that

$$q\cdot r_j = (k,\omega_n/u)\cdot(x_j,u\tau_j) = kx_j - \omega_n\tau,$$

therefore $A(q) = \sum_j e^{-iq\cdot r_j}A_j = \sum_j e^{-i(kx_j-\omega_n\tau)}A_j$. Note that $Z_0 = \int D\phi D\theta e^{-S_0/\hbar}$ with

$$-S_0 = \int_0^\beta d\tau \int dx \left[ i\frac{1}{\pi}\nabla\theta\partial_\tau\phi - \frac{1}{2\pi}\left(uK(\nabla\theta)^2 + \frac{u}{K}(\nabla\phi)^2\right)\right],$$

this gives, after Fourier transform, note that $\nabla\theta(r) = \nabla(\frac{1}{\beta\Omega}\sum_q e^{iq\cdot r}\phi(q)) = \nabla(\frac{1}{\beta\Omega}\sum_q e^{i(kx-\omega_n\tau)}\phi(q)) = \frac{1}{\beta\Omega}\sum_q ike^{iq\cdot r}\phi(q) = \frac{1}{\beta\Omega}\sum_q(-ik)e^{-iq\cdot r}\phi(-q)$ and $\partial_\tau\phi(r) = \frac{1}{\beta\Omega}\sum_q(-i\omega_n)e^{iq\cdot r}\phi(q)$, therefore

$$-S_0 = \frac{1}{(\beta\Omega)^2}\sum_{q,q'}\int d\tau dx \left[e^{i(-q+q')r}[i\frac{1}{\pi}(-ik)\cdot(-i\omega_n)\theta(-q)\phi(q') - \frac{1}{2\pi}\left(uKkk'\theta(-q)\theta(q') + \frac{u}{K}kk'\phi(-q)\phi(q')\right)\right],$$

completing the $d\tau dx$ integral gives $\int d\tau dx e^{i(-q+q')} = \delta_{q=q'}\beta\Omega$, therefore

$$-S_0 = \frac{1}{\beta\Omega}\sum_q \left[-\frac{ik\omega_n}{\pi}\theta(-q)\phi(q) - \frac{k^2}{2\pi}\left(uK\theta(-q)\theta(q) + \frac{u}{K}\phi(-q)\phi(q)\right)\right],$$

this is

$$-S_0 = -\frac{1}{\beta\Omega}\sum_q \begin{pmatrix}\theta_{-q} & \phi_{-q}\end{pmatrix}\begin{pmatrix}\frac{uK}{2\pi}k^2 & \frac{ik\omega_n}{2\pi}\\ \frac{ik\omega_n}{2\pi} & \frac{u}{2\pi K}k^2\end{pmatrix}\begin{pmatrix}\theta_q\\ \phi_q\end{pmatrix},$$

or we write $-S_0 = -\frac{1}{2\beta\Omega}\sum_q \begin{pmatrix}\theta_{-q} & \phi_{-q}\end{pmatrix} M^{-1}\begin{pmatrix}\theta_q\\ \phi_q\end{pmatrix}$ with $M^{-1} = \frac{1}{\pi}\begin{pmatrix}uKk^2 & ik\omega_n\\ ik\omega_n & \frac{u}{K}k^2\end{pmatrix}$. Note of course we have $\theta_{-q} = \theta_q^*$ and so for $\phi$. Now we have

$$\langle T_\tau e^{i\sum_j (A_j\phi(r_j)+B_j\theta(r_j))}\rangle = \frac{1}{Z_0}\int D\phi D\theta e^{-\frac{1}{2\beta\Omega}\sum_q\left[\begin{pmatrix}\theta_{-q} & \phi_{-q}\end{pmatrix}M^{-1}\begin{pmatrix}\theta_q\\ \phi_q\end{pmatrix} - i\left(\begin{pmatrix}B_{-q} & A_{-q}\end{pmatrix}\begin{pmatrix}\theta_q\\ \phi_q\end{pmatrix} + h.c.\right)\right]},$$

then we have

$$\langle T_\tau e^{i\sum_j (A_j\phi(r_j)+B_j\theta(r_j))}\rangle = e^{-\frac{1}{2\beta\Omega}\sum_q\begin{pmatrix}B_{-q} & A_{-q}\end{pmatrix}M\begin{pmatrix}B_q\\ A_q\end{pmatrix}},$$

where we used

$$\left(\prod_i \int \frac{du_i du_i^*}{2\pi i}\right) e^{-u^\dagger Au + h^\dagger u + u^\dagger h} = \frac{e^{h^\dagger A^{-1}h}}{\mathrm{Det}A}.$$

Note that a simple matrix inverse gives $M = \left[\frac{1}{\pi}\begin{pmatrix}uKk^2 & ik\omega_n\\ ik\omega_n & \frac{u}{K}k^2\end{pmatrix}\right]^{-1} = \frac{\pi}{k^2(u^2k^2+\omega_n^2)}\begin{pmatrix}\frac{u}{K}k^2 & -ik\omega_n\\ -ik\omega_n & uKk^2\end{pmatrix}$.

Now we proceed further to evaluate the expoential: the $AA$ term on the exponential on the right is $AA = -\frac{1}{2\beta\Omega}A_{-q}\frac{\pi uKk^2}{k^2(u^2k^2+\omega_n^2)}A_q$, using $A_q = \sum_j e^{-iq\cdot r_j}A_j$ we get $AA = -\frac{1}{2\beta\Omega}\sum_{i,j}\sum_q A_iA_j\frac{uK\pi}{u^2k^2+\omega_n^2}e^{iq\cdot(r_i-r_j)}$, the sin part in the exponential is odd with respect to $i,j$ so it cancels and we are left with $AA = -\frac{1}{2\beta\Omega}\sum_{i,j}\sum_q A_iA_j\cos(q\cdot(r_i-r_j))\frac{uK\pi}{u^2k^2+\omega_n^2}$. To deal with the divergent part, we define $F_1(r) = \frac{1}{\beta\Omega}\sum_q(1-\cos(q\cdot r))\frac{2\pi u}{u^2k^2+\omega_N^2}$, then $AA = -\frac{1}{4}\sum_{i,j}A_iA_jKF_1(r_i-r_j) + (\sum_i A_i)^2\frac{1}{2\beta\Omega}\sum_q\frac{uK\pi}{u^2k^2+\omega_n^2}$,

where we have separated the finite and infinite part. Since this $AA$ term appears on the exponential, we see that the infinte part will render the exponential zero; on other other hand, if $\sum_i A_i = 0$ (which is called neutral coefficients $A$), the would-be-infinite part actually vanishes, meaning the exponential is finite (instead of zero). Let us only look at these neutral cases: then we have $AA = -\frac{1}{2}K\sum_{i<j} A_iA_jF_1(r_i - r_j)$. Also define $F_2(r) = \frac{1}{\beta\Omega}\sum_q e^{iq\cdot r}\frac{-i2\pi\omega_n/k}{u^2k^2+\omega_n^2}$, then it can be shown that (also assuming $B$ is neutral: $\sum_i B_i = 0$)

$$\langle T_\tau e^{i\sum_j (A_j\phi(r_j)+B_j\theta(r_j))}\rangle = e^{-\frac{1}{2}\sum_{i<j}\left[(-A_iA_jK-B_iB_jK^{-1})F_1(r_i-r_j)+(A_iB_j+B_iA_j)F_2(r_i-r_j)\right]}. \tag{18}$$

Further using the fact that $\langle[\phi(r)-\phi(0)]^2\rangle = KF_1(r)$, $\langle[\theta(r)-\theta(0)]^2\rangle = K^{-1}F_1(r)$ and $\langle\phi(r)\theta(0)\rangle = \frac{1}{2}F_2(r)$, we have that the exponential is equal to $-\frac{1}{2}\sum_{i<j}\langle[(-A_iA_j(\phi(r_i)-\phi(r_j))^2-B_iB_j(\theta(r_i)-\theta(r_j))^2+2(A_iB_j\phi(r_i)\theta(r_j)+B_iA_j\phi(r_j)\theta(r_i)))]\rangle$, where we have used some translation symmetry; this is $-\frac{1}{2}\sum_{i<j}\langle -A_iA_j(\phi(r_i)^2+\phi(r_j)^2) - B_iB_j(\theta^2(r_i)+\theta(r_j)^2) + 2(A_iA_j\phi(r_i)\phi(r_j)+B_iB_j\theta(r_i)\theta(r_j)+A_iB_j\phi(r_i)\theta(r_j)+B_iA_j\phi(r_j)\theta(r_i))\rangle$, note that we have $\sum_{i<j}\langle A_iA_j(\phi(r_i)^2+\phi(r_j)^2)\rangle = \sum_{i<j}\langle A_iA_j\phi(r_i)^2\rangle+\sum_{i>j}\langle A_iA_j\phi(r_i)^2\rangle = \sum_{i\neq j}A_iA_j\langle\phi(r_i)^2\rangle = \sum_i\langle\phi(r_i)^2\rangle A_i\sum_{j\neq i}A_j = \sum_i\langle\phi(r_i)^2\rangle A_i(-A_i) = -\sum_i A_i^2\langle\phi(r_i)^2\rangle$. This way, finaly we see that the exponential is actually a square

$$\langle T_\tau e^{i\sum_j (A_j\phi(r_j)+B_j\theta(r_j))}\rangle = e^{-\frac{1}{2}\langle T_\tau[\sum_i A_i\phi(r_i)+B_i\theta(r_i)]^2\rangle}. \tag{19}$$

or more concisely,

$$\langle e^A\rangle = e^{\frac{1}{2}\langle A^2\rangle}.$$

$$S = \frac{1}{2\pi K}\int dxd\tau\frac{1}{u}(\partial_\tau\phi)^2 + u(\partial_x\phi)^2] + \frac{2g}{(2\pi\alpha)^2}\int dxd\tau\cos(\sqrt{8}\phi(x,\tau)),$$

define $r = (x, u\tau)$, $q = (k, \omega_n/u)$,

$$\frac{Z}{Z_0} = \frac{1}{Z_0}\int D\phi e^{-(S_0^>+S_0^<)}\left[1 - \frac{2g}{(2\pi\alpha)^2u}\int d^2r\cos(\sqrt{8}(\phi^>(r)+\phi^<(r))) + \frac{2g^2}{(2\pi\alpha)^4u^2}\int d^2r_1d^2r_2\cos(\sqrt{8}(\phi^>(r_1)+\phi^<(r_1)))\cos(\sqrt{8}(\phi\right.$$

integrate out the fast modes $\phi^>$: using $\frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>}\cos(\sqrt{8}(\phi^>+\phi^<)) = \frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>}\frac{1}{2}(e^{i\sqrt{8}(\phi^>+\phi^<)}+e^{-i\sqrt{8}(\phi^>+\phi^<)}) = \frac{1}{2}e^{i\sqrt{8}\phi^<}\frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>+i\sqrt{8}\phi^>} + \frac{1}{2}e^{-i\sqrt{8}\phi^<}\frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>-i\sqrt{8}\phi^>} = \frac{1}{2}e^{i\sqrt{8}\phi^<}\langle e^{i\sqrt{8}\phi^>(r)}\rangle + \frac{1}{2}e^{i\sqrt{8}\phi^<}\langle e^{-i\sqrt{8}\phi^>(r)}\rangle$, using Eq. (19) we get $\langle e^{-i\sqrt{8}\phi^>(r)}\rangle = \langle e^{i\sqrt{8}\phi^>(r)}\rangle = e^{-4\langle(\phi^>(r))^2\rangle}$, therefore $\frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>}\cos(\sqrt{8}(\phi^>+\phi^<)) = \cos(\sqrt{8}\phi^<)e^{-4\langle(\phi^>(r))^2\rangle}$.
But I do not understand: here the neutral condition is missing $\sum_i A_i \neq 0$, so how can we directly use the previous result?

Next, we use $\cos(\sqrt{8}(\phi^>(r_1)+\phi^<(r_1)))\cos(\sqrt{8}(\phi^>(r_2)+\phi^<(r_2))) = \frac{1}{2}\sum_{s=\pm}\cos\left(\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)+(\phi^>(r_1)+s\phi^>(r_2)))\right) = \frac{1}{4}\sum_{s=\pm}e^{i\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)+(\phi^>(r_1)+s\phi^>(r_2)))}+e^{-i\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)+(\phi^>(r_1)+s\phi^>(r_2)))}$, therefore $\frac{1}{Z_0^>}\int D\phi^> e^{-S_0^>}\cos(\sqrt{8}(\phi^>(r_1)+\phi^<(r_1)))\cos(\sqrt{8}(\phi^>(r_2)+\phi^<(r_2))) = \frac{1}{4}\sum_{s=\pm}\left(e^{i\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2))}e^{-4\langle(\phi^>(r_1)+s\phi^>(r_2))^2\rangle} + e^{-i\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2))}e^{-4\langle(\phi^>(r_1)+s\phi^>(r_2))^2\rangle}\right.$ $\frac{1}{2}\sum_{s=\pm}\cos(\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)))e^{-4\langle(\phi^>(r_1)+s\phi^>(r_2))^2\rangle}$, therefore we have

$$\frac{Z}{Z_0} = \frac{1}{Z_0^<}\int D\phi e^{-S_0^<}\left[1 - \underbrace{\frac{2g}{(2\pi\alpha)^2u}}_{\lambda}\underbrace{\int d^2r\cos(\sqrt{8}\phi^<)e^{-4\langle(\phi^>(r))^2\rangle}}_{A} + \underbrace{\frac{2g^2}{(2\pi\alpha)^4u^2}}_{\frac{1}{2}\lambda^2}\underbrace{\int d^2r_1d^2r_2\frac{1}{2}\sum_{s=\pm}\cos(\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)))e^{-4\langle(\phi^>(\right.}_{B}$$

Now we want to reexponentiate this to obtain the effective action. The terms inside bracktes are of the form $1-\lambda A+\frac{1}{2}\lambda^2B+o(\lambda^2)$ where we denoted $\lambda = \frac{2g}{(2\pi\alpha)^2u}$, then we have $1-\lambda A+\frac{1}{2}\lambda^2B+o(\lambda^2) = e^{-\lambda A+\frac{1}{2}\lambda^2B-\frac{1}{2}\lambda^2A^2+o(\lambda^2)}$, this is the cumulant expansion. Note it's important to add the term $-\frac{1}{2}\lambda^2A^2$ because $e^{-\lambda A-\frac{1}{2}\lambda^2A^2} = 1-\lambda A+\frac{1}{2}\lambda^2A^2-\frac{1}{2}\lambda^2A^2+o(\lambda^2) = 1-\lambda A+o(\lambda^2)$, which is the first lesson in cumulant expansion. This way we get

$$\frac{Z}{Z_0} = \frac{1}{Z_0^<}\int D\phi e^{-S_0^<-\delta S},$$

with

$$\begin{aligned}-\delta S = &-\frac{2g}{(2\pi\alpha)^2u}\int d^2r\cos(\sqrt{8}\phi^<(r))e^{-4\langle(\phi^>(r))^2\rangle} + \frac{g^2}{(2\pi\alpha)^4u^2}\int d^2r_1d^2r_2\sum_{s=\pm}\cos(\sqrt{8}(\phi^<(r_1)+s\phi^<(r_2)))e^{-4\langle(\phi^>(r_1)+s\phi^>(r_2))^2\rangle}\\ &-\frac{2g^2}{(2\pi\alpha)^4u^2}\int d^2r_1d^2r_2\cos(\sqrt{8}\phi^<(r_1))e^{-4\langle(\phi^>(r_1))^2\rangle}\cos(\sqrt{8}\phi^<(r_2))e^{-4\langle(\phi^>(r_2))^2\rangle},\end{aligned} \tag{20}$$

Originally $\phi$ has momentum cutoff $|\boldsymbol{q}| < \Lambda$; what we just did was to introduce a smaller cutoff $\Lambda' < \Lambda$ such that $\phi^>(\boldsymbol{q})$ lives in $\Lambda' < |\boldsymbol{q}| < \Lambda$. Now we have integrated out $\phi^>$ so that the remaining field $\phi^<$ has cutoff $|\boldsymbol{q}| < \Lambda'$. Now we want to do some rescaling such that the momentum cutoff of $\phi^<$ goes back to the original value, $\Lambda$. To do this, we define new momentum $\boldsymbol{q}_{\text{new}} = \frac{\Lambda}{\Lambda'}\boldsymbol{q}$, whose cutoff is $\Lambda$. This means that $r_{\text{new}} = (x_{\text{new}}, u\tau_{\text{new}}) = \frac{\Lambda'}{\Lambda}(x, u\tau) = \frac{\Lambda'}{\Lambda}r$, or $r = \frac{\Lambda}{\Lambda'}r_{\text{new}}$. Plug this into the linear term we have $-\delta S_1 = -\left(\frac{\Lambda}{\Lambda'}\right)^2 \int d^2 r_{\text{new}} \frac{2g}{(2\pi\alpha)^2 u} \cos(\sqrt{8}\phi^<(\frac{\Lambda}{\Lambda'}r_{\text{new}}))e^{-4\langle(\phi^>(\frac{\Lambda}{\Lambda'}r_{\text{new}}))^2\rangle}$, now define $\phi^<(\frac{\Lambda}{\Lambda'}r_{\text{new}}) = \phi_{\text{new}}(r_{\text{new}})$, then we hae $-\delta S_1 = -\left(\frac{\Lambda}{\Lambda'}\right)^2 \int d^2 r_{\text{new}} \frac{2g}{(2\pi\alpha)^2 u} \cos(\sqrt{8}\phi_{\text{new}}(r_{\text{new}})e^{-4\langle(\phi^>(\frac{\Lambda}{\Lambda'}r_{\text{new}}))^2\rangle}$, we see that effectively the coupling constant becomes

$$g_{\text{new}} = \left(\frac{\Lambda}{\Lambda'}\right)^2 e^{-4\langle(\phi^>(\frac{\Lambda}{\Lambda'}r_{\text{new}}))^2\rangle}.$$

`https://www.physik.uni-muenchen.de/lehre/vorlesungen/wise_17_18/TVI_TMP-TA4_-Condensed-Matter-Field-Theory exercises/ps11.pdf`

# 10 A Brief Introduction to CFT

The following content in this section comes from the `MBJC_talk_on_CFT`.

Start from massless Dirac fermion in 1+1D (which is a CFT): we know that in the quantum theory the vector and axial currents are both conserved ($0 = \partial_\mu J^\mu = \epsilon^{\mu\nu}\partial_\mu J_\nu$, or $\partial_- J_+ = \partial_+ J_- = 0$). In the bosonized theory, first try $\mathcal{L}_0 = \frac{1}{4\lambda^2}\text{Tr}\partial_\mu g \partial_\mu g^{-1}$: it is classically CFT but not quantumly since there are no two conserved currents. Turns out the correct one is $\mathcal{L} = \mathcal{L}_0 + k\Gamma[g]$, with $k$ integer and that $k = 4\pi/\lambda^2$. Note that at this $\lambda$ value, it has two conserved currents $J = g^{-1}\partial g$ (from EOM) and this $\lambda$ value is a gapless RG fix point. (Witten's Non-Abelian Bosonization in Two Dimensions).

Up to now we have not shown $\mathcal{L}$ is a CFT: if the action is invariant under a conformal coordinate transformation then the theory is a classical CFT; if furthermore the measure is invariant under a conformal coordinate transformation then the theory is also a quantum CFT. Here, since from $\mathcal{L}_0$ the field $g$ is dimensionless and that the term $\Gamma[g]$ is a top form and has no metric, the theory $\mathcal{L}_0$ is both a classical and quantum CFT, and so is the theory $\mathcal{L}$.

For a quantum CFT, (the expectation value of) all the observables are computed according to the conformal Ward identities $\partial_{\bar{z}}\left[\langle T(z,\bar{z})X\rangle - \sum_{i=1}^n \left(\frac{1}{z-w_i}\partial_{w_i}\langle X\rangle + \frac{h_i}{(z-w_i)^2}\langle X\rangle\right)\right] = 0$, and the equation that (bar)(unbar) all the barred (unbarred) variables), where $T = -2\pi T_{zz}$ and $\bar{T} = -2\pi T_{\bar{z}\bar{z}}$. By diagrammatics one can actually show that the energy-momentum tensor $\langle T(z)\rangle = \frac{1}{2k+C_\theta}\langle J^a(z)J^b(z)\rangle$(original reference?), which is different from the classical energy-momentum tensor $T(z) = \frac{1}{2k}$(exercise 15.1 of the CFT book).

Above we also have to assume $\gamma$ is simple Lie group $\bar{G}$-valued, which corresponds to $\bar{\mathfrak{g}}$. Using the conformal Ward identity (see CFT yellow book P621-623 or P181-121) gives $J^a(z)J^b(w) \sim \frac{k\delta_{ab}}{(z-w)}^2 + \sum_c if_{abc}\frac{J^c(w)}{(z-w)}$, $[T(z)J^a(w)] = (z-w)^{-2}J^a(w)$, and $[T(z)T(w)] = \frac{c}{2}(z-w)^{-4} + 2T(w)/(z-w)^2 + \partial T(w)/(z-w)$ and so forth. Then introduce $J^a(z) = \sum_{n\in\mathbb{Z}} z^{-n-1}J_n^a$, and$T(z) = \sum_{n\in\mathbb{Z}} z^{-n-2}L_n$, we get the set of algebra

$$\begin{aligned}
[J_n^a, J_m^b] &= \sum_c if_{abc}J_{n+m}^c + kn\delta_{ab}\delta_{n+m,0}, \\
[L_m, T_n^a] &= -nT_{m+n}^a, \\
[L_m, L_n] &= \frac{c}{12}m(m^2-1)\delta_{m+n,0} + (m-n)L_{m+n},
\end{aligned} \tag{21}$$

The bracket for $J_n^a$ defines an affine Lie algebra (call it $\mathfrak{g}$) at level $k = k^\vee$(since $(\theta,\theta) = 2$; the bracket for $L_n$ defines a Virasoro algebra $\mathfrak{v}$ (Note we can also obtain the copy for $\bar{J}$ which gives another copy of $\mathfrak{g}$, and similarly for $\bar{T}$ – every holomorphic object has an antiholomorphic copy). We thus see that the current and energy-momentum tensor together defined

$$\mathfrak{g} \oplus \mathfrak{v}, \tag{22}$$

which we call a chiral algebra (chiral in the sense that there is another copy for $\bar{J}$ and $\bar{T}$).

A natural question to ask: what is the relation between $\bar{\mathfrak{g}}$ and $\mathfrak{g}$? Is there a $\bar{\mathfrak{v}}$?

Answer: $\mathfrak{g}$ is obtained froom $\bar{\mathfrak{g}}$ via nontrivial central extension, and there is a $\bar{\mathfrak{v}}$, the so-called Witt algebra, whose nontrivial central extension is $\mathfrak{v}$. (This is how Witt algebra appears: note in 2D, CFT is equivalent to saying that the transformation must be of the form $z \to f(z)$, and $\bar{Z} \to \bar{f}(\bar{z})$, and define the generator $L_n^{(c)} = -z^{n+1}d/dz$ then the Lie algebra is $[L_m^{(c)}, L_n^{(c)}] = (m-n)L_{m+n}^{(c)}$ which is the Witte algebra). Note that $\mathfrak{g}$ is not directly the central extension of $\bar{\mathfrak{g}}$, but is the central extension of the loop algebra $\bar{\mathfrak{g}}_{\text{loop}}$. $\bar{\mathfrak{g}}_{\text{loop}}$ is the analytic mappings from the circle $S^1$ to $\bar{\mathfrak{g}}$. That $\mathfrak{g}$ is not

the central extension of $\bar{\mathfrak{g}}$ makes sense, first because a simple Lie group $\bar{\mathfrak{g}}$ has trivial central extension; second because the classical current $J$ is discribed by the loop algebra.

More generally, by promoting a Lie algebra from classical theory to quantum theory amounts to doing a (one-dimensional) central extention of the Lie algebra. If nontrivial extension exists, i.e. there is a new central element appearing on the RHS of Lie commutator which cannot be eliminated in anyway, we say the symmetry develops a quantum anomaly. When looking at the representation, this central element can only act with a constant, which is called the central charge or anomaly charge. We see above that $k$ and $c$ are the anomaly charge, called the level and the central charge, respectively, of the CFT.

As another rather trivial example: Lorentz algebra (as semisimple Lie algebra) only has trivial extension; it turns out Poincare algebra also only has trivial extension. But Galileo algebra has nontrivial extension in which another central element, $M$, can appear on the RHS of Lie brackets. The eigenvalue (i.e. the central charge) of $M$ is just the mass. In all the above cases where nontrivial central extension exist, the physical meaning of this nontrivial central extension is the existence of superselection rule: since different central charge label different representations, a quantum system, which can only carry one representation, can only have one central charge, and not the superposition of them.

Deriving the central extension: assume $[L_m, L_n] = (m-n)L_{m+n} + c(m,n)$, we must have $c(m,n) = -c(n,m)$; by the redefinition $L_m \to L_m + b(m)$ we have $c(m,n) \to c(m,n) + (m-n)a(m+n)$ therefore we set $c(m,0) = c(0,m) = c(1,-1) = 0$. Then the Jacobi Identity for $L_m, L_n, L_0$ gives $c(m,n) = a(m)\delta_{m+n,0}$ for some $a$, and then Jacobi Identity for $L_m, L_n, L_p$ gives $0 = (m-n)a(m+n) - (2m+n)a(n) + (m+2n)a(m)$, which gives $a(m) = \frac{1}{6}m(m^2-1)a(2)$. We then define $a(2) = c/2$ and we get Virasoro.

Note that from the Mathematical point of view, it is not natural to have a predefined Lie bracket between the generators of the affine Lie group $\mathfrak{g}$ and the Virasoro algebra $\mathfrak{v}$. However, as a physical theory, such a Lie bracket can be very naturally obtained in a WZW theory, since the energy-momentum tensor is written as the bilinear of the current.

Now formulate the above from a mathenatical point of view. the minimal data to start with is solely the affine Lie algebra $\mathfrak{g}$: one then defines the current $J^a(z)$ as the generating function for this algebra. The study of the representation theory of the algebra inevitably require the study of its universal enveloping algebra $\mathfrak{U}(\mathfrak{g})$: for example, the element $J^a(z)J^b(z)$ does not live in $\mathfrak{g}$ but lives precisely in $\mathfrak{U}(\mathfrak{g})$. The quantity $:J^a(z)J^b(z):$ cna be written in an antisymmetric part and a symmetric part, which must be of the form

$$:J^a(z)J^b(z):=\frac{1}{2}f^{ab}{}_c J^c(z) + \frac{\bar{\kappa}^{ab}}{\xi d}T(z),$$

and contracting with $\bar{\kappa}_{ab}$ we have $T(z) = \xi\bar{\kappa}_{ab} :J^a(z)J^b(z):$. We give this quantity a name, the energy-momentum tensor, which coincides that in a physical (i.e. WZW) theory. Then by defining $T(z) = \sum_{n\in\mathbb{Z}} z^{-n-2}L_n$ we automatically get a well-defined Lie algebra 22. The process

$$\mathfrak{g} \to \mathfrak{U}(\mathfrak{g}) \supset T \equiv \bar{\kappa}_{ab} :J^aJ^b: \to \mathfrak{g}\otimes\mathfrak{v}$$

is called the Sugawara construction. Along with this construction, the constant $c$ is seen to be defined by $c = 2kd\xi$. By choosing $\xi = 1/(2k+C_\theta)$ to agree with the Ward Identity. Therefore the parameter $c$ is fixed to be $c = \frac{2kd}{2k+C_\theta}$.

The class of CFT with such a chiral algebra $\mathfrak{g}\oplus\mathfrak{v}$ is called WZW theories. Such a chiral algebra signifies that the observable of the theory is current and energy-momentum tensor.

Definition of chiral algebra (Beilinson, Drinfeld, 2004) on a more general sense: the generators of symmetry are observables in a quantum theory (e.g. momentum, angular momentum). In CFT, the generators of the conformal symmetry are also a basis for the observables, in this sense the observables form a Lie algebra (the Lüscher-Mack theorem), called the symmetry algebra $\mathcal{W}$. In 2D, the symmetry algebra is direct sum of the holomorphic and antiholomorphic functions, $\mathcal{W}_{\text{tot}} = \mathcal{W}\oplus\overline{\mathcal{W}'}$, each direct summand called a chiral algebra. Since the stress tensor is always in the theory, we must have

$$\mathfrak{v}\in\mathcal{W}.$$

Formally speaking, the information of a CFT is contained in a chiral algebra and its representations. The classification of chiral algebras is obviously a too ambitious task. However, when the chiral algebra is of the form 22 (i.e. the chiral fields being just $T(z)$ and $J(z)$), i.e. the WZW theories, we have known a great deal.

The next biggest question is representation theory: The Lie algebra $\mathcal{W}$ acts on the Hilbert space $\mathcal{H}$, called a $\mathcal{W}$-module. With two conditions: 1. representation is unitary (i.e. the module is endowed with a positive definite product; 2. the energy is bounded from below.

$\mathcal{W} = \{W_{n\in\mathbb{Z}}^{i\in\mathbb{Z}_{\geq 0}}\}\cup\{C\}$

The virasoro: identified as $L_n \equiv W_n^0$. For at least $m\in\{0,\pm1\}$ and generator $W_n^i$ must satisfy $[W_m^0, W_n^i] \equiv [L_m, W_n^i] = ((\Delta_i - 1)m - n)W_{m+n}^i$; in particular, $[L_0, W_n^i] = -nW_n^i$, any vector in $\mathcal{H}$ will be anihilated by all $W_n^i$ with sufficient large $n$. We have the graded structure on $n$; according to $n>0$, $n=0$ and $n<0$ we have

$$\mathcal{W} = \mathcal{W}^-\oplus\mathcal{W}^0\oplus\mathcal{W}^+;$$

Therefore there most be an energy operator, which is $L_0 = -z\partial/\partial z$, it measures the scaling dimension of other operators.

## 10.1 Conformal group and algebra in classical theory and CFT

A CFT is a QFT invariant under the conformal group. The conformal group describes the symmetry of the underlying manifold. Although both invariant under the conformal group, the representations that they carry are different.

The conformal group acts on the metric as $g_{\mu\nu} \mapsto \sigma(x) g_{\mu\nu}(x)$. Under infinitesimal: $x^\mu \mapsto x^\mu + \epsilon^\mu(x)$, then we have, in the infinitesimal case, $\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = s(x) g_{\mu\nu}(x)$ where $s(x) = \sigma(x) - 1$ is some scalar function. taking trace (or contracting by $g^{\mu\nu}$) one gets $s(x) = \frac{2}{d}\partial_\rho \epsilon^\rho = \frac{2}{d} g^{\rho\lambda}\partial_\rho \epsilon_\lambda$. Therefore the first order equation for $\epsilon_\mu$ is

$$\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = \frac{2}{d} g_{\mu\nu} g^{\rho\lambda} \partial_\rho \epsilon_\lambda. \tag{23}$$

(Note that the equation below is not useful: On the other hand, taking spacetime is euclidean so $g_{\mu\nu} = \eta_{\mu\nu}$, and take another derivative we get $\partial_\mu s = \frac{2}{d}\partial_\rho\partial_\mu \epsilon^\rho = \frac{2}{d}\eta^{\rho\nu}\partial_\rho\partial_\mu\epsilon_\nu = \frac{2}{d}\eta^{\nu\rho}\partial_\nu\partial_\mu\epsilon_\rho \overset{\eta^{\nu\rho}\equiv\eta^{\rho\nu}}{=} \frac{1}{d}\eta^{\nu\rho}\partial_\mu(\partial_\nu\epsilon_\rho + \partial_\rho\epsilon_\nu) = \frac{1}{d}\eta^{\nu\rho}\partial_\mu(s\eta_{\nu\rho}) = \partial_\mu s$) Then from the original equation take derivative we get $\partial_\rho\partial_\mu\epsilon_\nu + \partial_\rho\partial_\nu\epsilon_\mu = \partial_\rho s \eta_{\mu\nu}$. Now we want $\partial_\rho\partial_\mu\epsilon_\nu =?$, the only way to get this is to write three equations by permuting $\rho$, $\nu$, and $\mu$ to get $2\partial_\mu\partial_\nu\epsilon_\rho = \eta_{\mu\rho}\partial_\nu s + \eta_{\nu\rho}\partial_\mu s - \eta_{\mu\nu}\partial_\rho s$, and then contracting with $\eta^{\mu\nu}$ we get $2\partial^2\epsilon_\rho = (2-d)\partial_\rho s$. Then we have $2\partial^2\partial_\nu\epsilon_\rho = (2-d)\partial_\rho\partial_\nu s$, we also have $2\partial^2\partial_\rho\epsilon_\nu = (2-d)\partial_\nu\partial_\rho s$ therefore we have $(2-d)\partial_\rho\partial_\nu s = \partial^2(\partial_\nu\epsilon_\rho + \partial_\rho\epsilon_\nu) = \partial^2 s\eta_{\nu\rho}$, contracting by $\eta^{\nu\rho}$ finally we get $(2-d)\partial^2 s = \partial^2 s$, or

$$(d-1)\partial^2 s = 0. \tag{24}$$

Note this equation looks convenient but it is actually loosing information from the original equation (23) because we have taken three derivatives on $\epsilon_\mu$. For example, it is seen from 23 that when $d = 2$ and euclidean geometry what we have is the Cauchy-Rieman equation $\partial_0\epsilon_0 - \partial_1\epsilon_1 = 0$, $\partial_0\epsilon_1 + \partial_1\epsilon_0 = 0$, but this cannot be obtained from Eq. (24).

therefore the finite conformal transformations are the direct product of the group of holomorphic coordinate transformations with that of antiholomorphic coordinate transformations. This is an infinite-dimensional Lie group.

Why Lie group? This is equivalent to asking why symmetry transformations form Lie group structure, i.e. why care the brackets. Then we just remember that the group satisfies composition rule $U(T(\theta))U(T(\xi)) = U(T(f(\theta,\xi)))$, which is the Baker–Campbell–Hausdorff formula, where $f$ is inferred from $\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + \cdots)$. Therefore when expanded by the generators we have $(1 + \theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \cdots)(1 + i\xi^a t_a + \frac{1}{2}\xi^b\xi^c t_{bc} + \cdots) = 1 + (\theta^a + \xi^a + f^a{}_{bc}\xi^b\theta^c + \cdots)t_a + \frac{1}{2}(\theta^b + \xi^b + \cdots)(\theta^c + \xi^c + \cdots)t_{bc}$, we see the lowest nontrivial order is $\theta^a\xi^b t_a t_b = f^a{}_{bc}\xi^b\theta^c t_a + \frac{1}{2}(\theta^b\xi^c + \xi^b\theta^c)t_{bc}$, which gives $t_b t_c = f^a{}_{bc}t_a + t_{bc}$, where we used the fact that $t_{bc} = t_{cb}$. Therefore we also have $t_c t_b = f^a{}_{cb}t_a + t_{cb}$, which gives $[t_b, t_c] = (f^a{}_{cb} - f^a{}_{bc})t_a$, which is the Lie bracket. We see the essense that leads to the analysis of Lie bracket is the Baker–Campbell–Hausdorff formula, which contains Lie bracket. This formula is fundamentally reflecting the fact that the Lie group multiplication is noncommutative. Here, of course, our Lie group multiplication is function composition, therefore is natually noncommutative, which justifies the appearance of a Lie bracket. This requires us to find the Lie algebra of holomorphic transformations, which are natually identified with the vector fields of the conformal transformation, which are $L_n^{(c)} = -z^{n+1}d/dz$. They satisfy Lie algebra $[L_m^{(c)}, L_n^{(c)}] = (m-n)L_{m+n}^{(c)}$, the Witt algebra.

The above is the classical theory, i.e. the conformal transformation was studied in the classical theory (i.e. manifolds) and not the quantum theory. In quantum theory, one studies the conformal transformation on the quantum states. As shown by Weinberg, this amounts to finding the central extension of the classical Lie algebra (in our case this is the Witte algebra). Of course there is always a trivial central extention, in the sense that the Lie algebra brackets are not modified. However, it turns out (??) that the central extension describing the quantum theories with conformal invariance must be nontrivial. An equivalent way of saying is that the quantum theory develops an anomaly. Now the Lüscher–Mach theorem (or see the analysis in Weinberg) gives the form of this extended algebra: it is exactly the Virasoro algebra:

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n,0}.$$

## 10.2 How Kac–Moody algebra emerge in a quantum theory

In this section, we show how a Kac–Moody algebra emerges in a physical theory.

Consider $N$-free massless Majorana fermions in 2D spacetime: the action is $S = \frac{1}{2}\int d^2x i\Psi\partial_x\Psi$ (usually there are two families, sitting at fermi point $k_+$ and $k_-$; now we are only considering one family of them), where $\Psi$ is an $N$-component field transforming under some representation $R$ of a simple Lie algebra $\bar{\mathfrak{g}}$. It is then natural, from a physical point of view, to define the current operator

$$J^a =: \frac{1}{2}\Psi^T R(T^a)\Psi :,$$

then from the canonical anticommutation relations of $\Psi$ we can deduce the commutation relations for $J^a$, which is

$$[J^a(x), J^b(y)] = \frac{1}{2\pi i}\bar{\kappa}^{ab}\delta'(x-y) + f^{ab}{}_c J^c(x)\delta(x-y),$$

transforming the coordinate to $z = \exp(2\pi i x/L)$ and with appropriate boundary conditions, the current $J^a(z)$ (which is defined after defining a series of other quantities)

$$\big[J^a(z)J^b(w)\big]_- = (z-w)^{-2}\bar{\kappa}^{ab}k - (z-w)^{-1}f^{ab}{}_c J^c(w),$$

with $k\bar{\kappa}^{ab} = \frac{1}{2}\mathrm{tr}(R(T^a)R(T^b))$, i.e. $k = \frac{1}{2}I$, $I$ is the Dynkin index of representation $R$. This shows that the current fields $J^a(z)$ generate an affine Lie algebra

In representation of affine lie algebra, the number $k^\vee := k \cdot \frac{2}{(\theta,\theta)}$ is called the level of the representation. We therefore see that $k^\vee = I/(\bar{\theta},\bar{\theta})$, which by the theory of highest weight modules of simple Lie algebra is integer (since the module here is orthogonal module), therefore the module is an irreducible integrable highest weight module: the fermions $\psi^i(z)$ correspond to the highest weight vector. What is the significance of the representation analysis? Well, the significance is that, we have just shown that a construction realizing irreducible integrable highest weight modules for arbitrary untwisted affine Lie algebras, at any level which is equal to the Dynkin index of a finite-dimensional module of the horizontal subalgebra exists!

Note: integrable highest weight modules have highest weight being dominant integral.

## 10.3 Bottom up: CFT from affine Lie algebra

Define: operator product: $\mathcal{R}(A(z)B(w))$; contraction: $[A(z)B(w)]_-$; normal ordered product: $:A(w)B(w):$;

The normal order of two currents can be split into an antisymmetric part and a symmetric part:

$$:J^a(z)J^b(z): := \frac{1}{2}f^{ab}{}_c J^c(z) + \frac{\bar{\kappa}^{ab}}{\xi d}T(z),$$

therefore $T(z) = \xi\bar{\kappa}^{ab}:J^a(z)J^b(Z):$. Evaluating this using the definition of contraction and the result for $\mathcal{R}(J^a(z)J^b(w))$, we have

$$[J^a(z)T(w)]_- = \xi(2K + C_\theta)(z-w)^{-2}J^a(W),$$

where $C_\theta$ is the quadratic defined by $C_\theta\delta^d_c = f^{ab}{}_c f_{ba}{}^d$. replace $K$ by its eigenvalue $k$ (remember that $K$ has same eigenvalue for all vectors in a highest weight module $R_\Lambda$), therefore if we choose $\xi = \frac{1}{2k+C_\theta}$ then we have $[J^a(z)T(w)]_- = (z-w)^{-2}J^a(w)$. Note that $2k + C_\theta = k^\vee(\theta,\theta) + g^\vee(\theta,\theta) = (k^\vee + g^\vee)(\theta,\theta)$ (see P117 and P60).

## 10.4 Free boson theory

Next, we study Luttinger liquids from the CFT point of view. Especially, we will derive the RG equation using operator product expansion. We will introduce basic concepts in CPF along the way. The basic references are of cours the big yellow book, Senechal's notes `https://arxiv.org/pdf/cond-mat/9908262.pdf`, and Fradkin's notes `http://eduardo.physics.illinois.edu/phys561/LL-11-17-09.pdf`.

In a free boson theory, also known as the gaussian model, the boson $\varphi(\boldsymbol{x})$ has scaling dimension $\Delta = \frac{d-2}{2}$ where $d$ is the Euclidean spacetime dimension [1], which is a consequence of the physical demand of scaling invariance of the massless case. The massless case is called the Gaussian fixed point. The mass term, then, is a relevant term at all dimensions to the Gaussian fixed point. We see that in 2D, a free boson has vanishing scaling dimension, $\Delta = 0$.

Given a field $\phi(z,\bar{z})$ with scaing dimension $\Delta$ and planar spin $s$. It is a primary field, if its variation under all local conformal transformation in 2D is given by

$$\phi'(w,\bar{w}) = \left(\frac{dw}{dz}\right)^{-h}\left(\frac{d\bar{w}}{d\bar{z}}\right)^{-\bar{h}}\phi(z,\bar{z}), \tag{25}$$

with $h = (\Delta + s)/2$, and $\bar{h} = (\Delta - s)/2$, called the holomorphic and antiholomorphic conformal dimensions.

The two-point and three-point correlation functions among primary fields $\phi_i$ are fully constrained to a fixed form by conformal invariance, and only depends on their conformal dimensions (see P117). The four-point function cannot be fully fixed, but has a universal scaling form (see P117).

The jargon "conformal fields" was used in a few places in the book (see e.g. P153) but was never explained. I suscept it is just another name for primary fields. From now on we assume this is true. Then any primary fields has a Laurent series

[1] Scaling dimension is defined as $\Delta$ in $\varphi(\boldsymbol{x}) \to \varphi'(\boldsymbol{x}/s) = s^\Delta\varphi(\boldsymbol{x})$ for $s > 1$.

representation (mode expansion), with the requirement that the modes in front of negative powers of $z$ or $\bar{z}$ annihilates the vacuum state. Here underlies that the theory has a vacuum state. Here also underlies that we are using the operator formalism (as opposed to path integral formalism), which requires definition of Hilber space, which in term underlies that a particular time direction is chosen (remember that we restrict ourselves to Euclidean spacetime and we have a lot freedom in choosing it).

The energy-momentum tensor $T_{\mu\nu}(z,\bar{z})$ with $\mu,\nu = z,\bar{z}$ can be reduced to just $T(z) = T_{zz}(z,\bar{z})$. By definition, an energy-momentum tensor has $\Delta = s = 2$ (exercise: check its transformation under scaling and rotation), so that $T(z)$ has conformal dimension $h = 2$ and $\bar{h} = 0$. the energy-momentum tensor is not a primary field because the above rule (25) does not hold for all conformal transformations as required. Note that one way of revealing the conformal dimension of a primary field $X$ is through the following Ward identity :

$$T(z)X = \frac{1}{z-w}\partial_w X + \frac{h}{(z-w)^2}X, \qquad X \text{ is primary,} \tag{26}$$

where above is in the sense of expecation values and only singular terms are shown; furthermore, the most general choice of $X$ is a product of primary fields (see P120).

The book presents that a generalized Ward identity can be written down, in which the field $X$ does not have to be primary, see Eq. (5.46) and P121-122. Also, the most general conformal symmetry allowed form of the OPE of energy-momentum tensor is written down (see Sec. 5.3.1-5.4, and Eq. (4.77) for motivation and reasoning):

$$T(z)T(w) \sim \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w},$$

whre $c$, defined as central charge, has a value $c = 1$ for free boson, $c = 1/2$ for free fermion, and $c = -26$ for the reparametrization ghosts. As the book emphasizes, the $c$ term is an anomalous term, in the sense that it does not exist in the classical theory; and that although the term is allowed by symmetry, the central charge $c$ cannot be determined solely from symmetry consideration and contains the UV (i.e. short-distance) data of the theory (note that this UV data is indeed in the continuous field theory).

We mentioned that the energy-momentum tensor $T(z)$ is not primary, so a natural question is how it deviates from a primary field under a local conformal transformation. This can be written down (see P136) using the generalized Ward identity mentioned above. Unsurprisingly, the deviation is measured by the central charge (product with the well-known Schwarzian derivative). The transformation rule of $T(z)$ un

Back to the massless free boson theory: one can show that $\langle\varphi\varphi\rangle \sim \ln|z-w|$ (direct calculation), $\langle\partial\varphi\partial\varphi\rangle \sim -\frac{1}{(z-w)^2}$ (take derivative in $\langle\varphi\varphi\rangle$), $T(z)\phi(w,\bar{w}) \sim \frac{h}{(z-w)^2}\phi(w,\bar{w}) + \frac{1}{z-w}\partial_w\phi(w,\bar{w})$ (using the Ward identity Eq. (26)), $T(z)\partial\varphi(w) \sim \frac{\partial\varphi}{(z-w)^2} + \frac{\partial_w^2\varphi(w)}{z-w}$ (using Ward identity above knowing that $h = 1$). Note that since this is a free theory, all these correlations can be obtained by just using Wick's theorem.

Following the usual canonical quantization, we can quantize the free boson theory (P159) on a cylinder (radial quantization). This choice of spacetime geometry was motivated on P151, and can be mapped to other geometries using conformal transformations. This directly gives the mode expansion of a boson field $\varphi(z,\bar{z})$, and shows that it is not a primary field (essentially, this is due to the existence of a boson zero mode, which contributes a term linear in $t$); but one immediately finds by taking derivative that $\partial\varphi$ and $\bar{\partial}\bar{\varphi}$ are primary. Using the fact that $\Delta = s = 0$ for the boson field $\varphi$, we see that $\partial\varphi$ has $h = 1$.

Define the so-called vertex operators $\mathcal{V}_\alpha(z,\bar{z}) =: e^{i\alpha\varphi(z,\bar{z})} :$, one can show that these fields are primary with conformal dimensions $h_\alpha = \bar{h}_\alpha = \frac{\alpha^2}{8\pi g}$ with $\alpha \in \mathbb{R}$ and $g \equiv \frac{1}{4\pi}$. One can further calculate (see P162) the OPE of $\partial\varphi\mathcal{V}_\alpha \sim -\frac{i\alpha}{4\pi g}\frac{\mathcal{V}_\alpha(w,\bar{w})}{z-w}$ , $T\mathcal{V}_\alpha \sim \frac{\alpha^2}{8\pi g}\frac{\mathcal{V}_\alpha(w,\bar{w})}{(z-w)^2} + \frac{\partial_w\mathcal{V}_\alpha(w,\bar{w})}{z-w}$ (which agrees with the Ward identity Eq. (26) and gives the conformal dimension), and $\mathcal{V}_\alpha\mathcal{V}_\beta \sim |z-w|^{-2\alpha^2}\delta_{\alpha=-\beta}$.

Using (See P162) : $e^{a\varphi_1} :: e^{b\varphi_2} :=: e^{a\varphi_1+b\varphi_2} : e^{ab\langle\varphi_1\varphi_2\rangle}$, we have:
$e^{i\alpha\varphi(z,\bar{z})}e^{-i\alpha\varphi(w,\bar{w})} = \frac{1}{(z-w)^{\alpha^2}(\bar{z}-\bar{w})^{\alpha^2}} : e^{i\alpha(\varphi(z,\bar{z})-\varphi(w,\bar{w}))} :$, so that $e^{i\alpha\varphi(z,\bar{z})}e^{-i\alpha\varphi(w,\bar{w})} + h.c. = \frac{1}{(z-w)^{\alpha^2}(\bar{z}-\bar{w})^{\alpha^2}} : 2\cos\alpha(\varphi(z,\bar{z})-\varphi(w,\bar{w})) := \frac{2}{(z-w)^{\alpha^2}(\bar{z}-\bar{w})^{\alpha^2}}\left(1-\frac{1}{2}(\varphi(z,\bar{z})-\varphi(w,\bar{w}))^2\right) = \frac{2}{(z-w)^{\alpha^2}(\bar{z}-\bar{w})^{\alpha^2}}\left(1-\frac{1}{2}((z-w)\partial\varphi(z,\bar{z}) + (\bar{z}-\bar{w})\bar{\partial}\varphi(w,\bar{w})\right.$
from this last form it is easy to derive the first equation of Eqs.(191) of Senechal's notes `https://arxiv.org/pdf/cond-mat/9908262.pdf`. The second and third equation of (191) is easy to prove. This should clear all the technical difficulties of the appendix A.

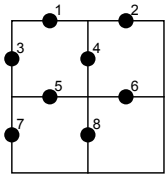



Figure 1: A minimal lattice of the toric code model. $S = 1/2$ spins live on the bonds (the black dots). Periodic boundary condition (PBC) are assumed (lattice is a torus).

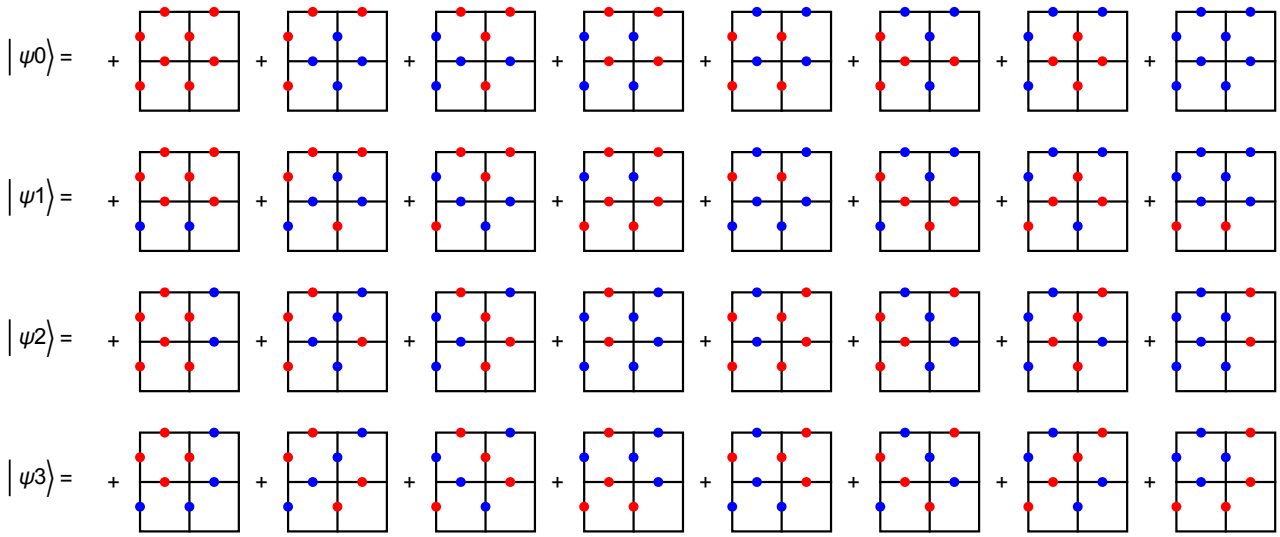



Figure 2: The four ground state of the minimal toric code model on the torus.

# 11 Quantum double models

## 11.1 Review of toric code

The Kitaev's toric code model defined on the minimal lattice (a torus under PBC) in Fig. 1 is written as

$$\begin{aligned} H_{\text{T.C.}} &= -\sum_{p\in\text{plaquettes}} P_p - \sum_{s\in\text{stars}} S_s \\ &= -(P_{1345} + P_{2346} + P_{1578} + P_{2678}) - (S_{1237} + S_{1248} + S_{3567} + S_{4568}) \\ &= -(\sigma_1^z\sigma_3^z\sigma_4^z\sigma_5^z + \sigma_2^z\sigma_3^z\sigma_4^z\sigma_6^z + \sigma_1^z\sigma_5^z\sigma_7^z\sigma_8^z + \sigma_2^z\sigma_6^z\sigma_7^z\sigma_8^z) \\ &\quad -(\sigma_1^x\sigma_2^x\sigma_3^x\sigma_7^x + \sigma_1^x\sigma_2^x\sigma_4^x\sigma_8^x + \sigma_3^x\sigma_5^x\sigma_6^x\sigma_7^x + \sigma_4^x\sigma_5^x\sigma_6^x\sigma_8^x), \end{aligned} \tag{27}$$

In some sense this is the simplest model for **topological order**. It only has 8 spins (=8 bonds), 4 plaquettes $P$ and 4 stars $S$.

**Basic feature of toric code on the torus:**

- Hamiltonian has a local (i.e. gauge) $\mathbb{Z}_2$ symmetry.
- Four degenerate ground states $\psi_{0,1,2,3}$ as shown in Fig. 2. All are "highly" entanglement states — i.e. each $\psi$ is the superposition of many copies of spin configurations. **This is one of the defining properties of topological order.** All states have the property that $P_p|\psi\rangle = S_s|\psi\rangle = |\psi\rangle$ for all plaquettes $p$ and starts $s$.
- Low energy excitations are **anyons**: $e$ ($S_s = -1$ for some $s$) and $m$ ($P_p = -1$ for some $p$) particles, although each must be created in pairs, can individually **freely propagate** (imagine the lattice infinitely large). **This is another defining property of topological order.** They satisfy nontrivial statistics (fusion and braiding), which can be understood in terms of **charge** (hence the name $e$ particle) and **flux** (hence the name $m$ particle).

Model:

$$H = -K\sum_p P_p - K'\sum_s S_s,$$

where $P_p = \prod_{i\in p}\sigma_i^z$ and $S_s = \prod_{i\in s}\sigma_i^x$. Easy to see that $[S_s, P_p] = 0$ Therefore ground state can be obtained this way: first choose a state with all $S_s = 1$, for example the state $|\psi_0\rangle = \prod_i |\sigma_i^x = 1\rangle$; Then act on $|\psi_0\rangle$ the projection operator $\prod_p Q_p$, where $Q_p = \frac{1+P_p}{2}$. The ground state

$$|0\rangle = \prod_p Q_p|\psi_0\rangle$$

can be viewed as massive superpositions of loop states, where by loop we mean a closed contour on the lattice, on the site of which we have $s_i = -1$.

A careful counting can show that the ground state must be four-fold degenerate, provided the lattice is a torus. What we care the most is excitations.

In the ground state $|0\rangle$, one can create excitations by either acting on it using $\sigma_i^z$ or $\sigma_i^x$ on a site $i$; the former changes two nearby $S_s$ and the latter changes two neareby $P_p$. We Using this idea, we can create the following state

$$|e_s, e_{s'}\rangle = \prod_{i\in l(s,s')}\sigma_i^z|0\rangle, \qquad |m_p, m_{p'}\rangle = \prod_{i\in l^*(p,p')}\sigma_i^x|0\rangle,$$

where $l(s, s')$ denote a line along lattice bonds connecting stars $s$ and $s'$, and $l^*(p, p')$ denote a line on the dual lattice intersecting lattice bonds, connecting plaquettes $p$ and $p'$. Note that $\sigma_i^z$ commutes with $\prod_p P_p$ therefore $|e_s, e_{s'}\rangle$ is still the ground state of the $K'$ term; for the $K$ term, only the terms $S_s$ and $S_{s'}$ flips sign, therefore the $K$ term acting on $|e_s, e_{s'}\rangle$ gives the ground state energy plus $4K$; i.e. we have

$$H|e_s, e_{s'}\rangle = (E_0 + 4K)|e_s, e_{s'}\rangle.$$

Similarly we have

$$H|m_p, m_{p'}\rangle = (E_0 + 4K')|m_p, m_{p'}\rangle.$$

We see that the excitations can extend infinitely far away yet remain finite energy. Therefore, although $e$ or $m$ must be created on pairs, we can imagine they are infinitely apart (with just finite energy) and each of them behaves as independent quasiparticles, which we call $e$ and $m$.

Now we explain why we call them $e$ and $m$: this is simply because $e$ corresponds to a start with $S_s = -1$ which is like a divergence, why $m$ corresponds to a plaquette $P_p = -1$ which is like a curl. Now what is important is the statistics: $e$ and $m$ are both bosons, in the sense that interchanging two $e$ or two $m$, the initial and final wave function are just the same. However, when moving an $e$ around an $m$ (vice versa) the wave function accumulates a $\pi$ flux. There is also the particle $\varepsilon$ which consists of a pair of nearby $e$ and $m$: it turns out $\varepsilon$ is a fermion, i.e. when interchanging two $\varepsilon$ the total wave function picks up a $\pi$ flux.

Language of confinement/condense: notice that in toric code, $e$, $m$, and $\varepsilon$ are finite energy excitations. Since $e$ and $m$ are boson and $\varepsilon$ is fermion, we can talk about condensation of $e$ and $m$: they are not condensed in toric code since as just said they are finite energy excitations. However, if we add large enough perturbation to the system, they may become gapless and may consequentely condense ($e$ or $m$ may condense first, depending on the perturbation).

Suppose $m$ condense first, then an individual $m$ particle is no longer well-defined since the number of $m$ particles in the ground state is uncertain. In terms of flux, the magnetic field lines now can end in arbutrary places in the sample, already in the ground state. We say that the $m$ particle is "Higgsed". Note the condensation language is easy to understand but saying that $m$ is Higgsed might be hard to understand, since $m$ becomes gapless, but this is just the language people use. For the $e$ particle, we can say that since the ground state is full of a superposition of different $m$ states, the mutual statistics of $e$ and $m$ particles means that the $e$ particle can no longer propagate coherently. Or, we can use another language: since $m$ condenses, therefore the magnetic field lines connecting $m$ particles become highly fluctuating. Since the magnetic field lines consists of $\sigma_i^z$ on the dual lattice becomes highly flucuating, the conjugate variable $\sigma_i^x$ becomes very sharply defined and do not fluctuate, meaning the electric field lines become concentrated and cannot spread out. Consequentely an $e$ particle emits a concentrated field line, which because it cannot spread, costs an energy proportional to its length, i.e. $e$ particle is confined. The $\varepsilon$ particle is also confined because it carries the electric charge.

We see that for toric code the Higgs phenomena and confinement are just two different languages of the same phase transition.

Now we will use toric code as a simple model to exhibit the algebraic theory of anyons. The very comprehensive and nice paper for this is `https://arxiv.org/pdf/1410.4540.pdf`. Anyon fusion: $a \times b = \sum_{c\in\mathcal{C}} N_{ab}^c c$, $F_d^{abc}$, $R^{ab}$. Here $\mathcal{C}$ is called a unitary braided tensor category (UBTC). The $F$-symbols and $R$-symbols completely specify a braded tensor category. They must satisfy the Pentagon and Hexagon equations. further requiring unitarity means $R$ is unitary. For any UBTC $\mathcal{C}$, one can consider $\mathrm{Aut}(\mathcal{C})$, which is called the topological symmetry group. Note this group is (so far) independent of any potential symmetry group $G$ of the microscopic Hamiltonian.

## 11.2 Quantum Double construction of anyon models

Q: what can we learn from toric code? Can we find a way to construction a family of anyons models?

Back to toric code. Perhaps not obviously: the anyons can be constructed entirely from the gauge group (in toric code is $\mathbb{Z}_2$).

**Central idea of quantum double models:** we can indeed construct a family of anyon systems, with only a single input: a group $G$ (serving as the gauge group). This construction is called quantum double model.

**Quantum double models: construct a system of anyons from a single input $G$.** (The explicit model Hamiltonian will be constructed in the next Friday to come.)

Recipe: a finite (possibly nonabelian) group $G$, and some intuitions from toric code and the Aharonov-Bohn experiment.

Comment: generally, anyons don't have to have the meaning of flux/charge. But in quantum double model they always do.

**Charges and fluxes are defined by how they can be detected:**

**Charges: (unitary) irrep of $G$.**

**Fluxes:** by definition, transform charges (think about an **Aharonov-Bohn (AB) experiment**). Therefore, **identified with the group element themselves.**

Caveat: fluxes are identified with **conjugacy classes** of $G$. A work understanding: $g \in G$ is a flux in the sense that $g = g_1 g_2 \cdots g_n$ (product of gauge transformations/gauge fields). Then this should be the same physically as $g_2 \cdots g_n g_1 = g_1^{-1} g g_1$.

**Tentative Summary: in quantum double model, anyon data = (charge, flux) = (irrep($G$), conjugacy class ($G$)).**

But this is still not quite correct: **not all $g \in G$ operation are well defined on the charges.** Charges can only be determined through the AB experiment, which requires the flux in the center to remain unchanged during the experiment. But if the charged anyon secretely has flux (which we don't know beforehand), this is not doable. **Therefore, the concept of charge exists only when flux are defined; one talks about charge with respect to some $g \in G$ (regarded as flux) only when the charged anyon itself has a flux $h \in Z(g)$,** here $Z(g) = \{h \in G | hg = gh\}$ is the centralizer group of $g$ in $G$.

**True summary: in quantum double model, anyon data = (charge, flux) = (irrep $Z(g)$, Conj. cls $(g)$).**

## 11.3 Table for anyons: examples of $\mathbb{Z}_2$ and $S_3$

The toric code model is exactly the quantum double model $D(\mathbb{Z}_2)$ for the gauge group $G = \mathbb{Z}_2$. Here $\mathbb{Z}_2$ is the group

$$\mathbb{Z}_2 = \{1, \bar{1}\}.$$

Table 1: Anyon table for the quantum double model $D(\mathbb{Z}_2)$.

| Superselection sector | Flux | Charge | Normalizer | Dimension |
|---|---|---|---|---|
| $\Gamma_1$ | $\{1\}$ | $A_+$ | $\mathbb{Z}_2$ | 1 |
| $\Gamma_2$ | $\{1\}$ | $A_-$ | $\mathbb{Z}_2$ | 1 |
| $\Gamma_3$ | $\{\bar{1}\}$ | $A_+$ | $\mathbb{Z}_2$ | 1 |
| $\Gamma_4$ | $\{\bar{1}\}$ | $A_-$ | $\mathbb{Z}_2$ | 1 |

We further consider the quantum double model $D(S_3)$ for the gauge group $G = S_3$. Here $S_3$ is the smallest nonabelian group

$$S_3 = D_3 = \{1, M_1, M_2, M_3, C_3, C_3^2\}.$$

Note: in the above two tables, Dimension = number of flux $\times$ number of charge; irreps are over complex field;

Meaning of **superselection sector:** global sectors that do not talk to each other.

Sum rule: as can be checked, $\sum \dim^2 = |G|^2$.

Topological spin? see below.

## 11.4 Statistics (fusion and braiding)

Next: consider statistics of anyons in a quantum double model, i.e. anyon fusion and braiding. Want to emphasize again why we are studying them: group $G$ is not enough! Any reasonable mathematical description of physical degree of freedom (particles) must contain these structures, simply because particles can combine (or, degenerate energy levels can split, which is exactly fusion), and particles can exchange (which in 2D is braiding).

Table 2: Anyon table for the quantum double model $D(S_3)$.

| Superselection sector | Flux | Charge | Normalizer | Dimension |
|---|---|---|---|---|
| $\Gamma_1$ | $\{1\}$ | $A_1$ | $N(1) = S_3$ | 1 |
| $\Gamma_2$ | $\{1\}$ | $A_2$ | $N(1) = S_3$ | 1 |
| $\Gamma_3$ | $\{1\}$ | $E$ | $N(1) = S_3$ | 2 |
| $\Gamma_4$ | $\{M_1, M_2, M_3\}$ | $A_+$ | $N(M_1) = \{1, M_1\}$ | 3 |
| $\Gamma_5$ | $\{M_1, M_2, M_3\}$ | $A_-$ | $N(M_1) = \{1, M_1\}$ | 3 |
| $\Gamma_6$ | $\{C_3, C_3^2\}$ | $A_0$ | $N(C_3) = \{1, C_3, C_3^2\}$ | 2 |
| $\Gamma_7$ | $\{C_3, C_3^2\}$ | $E_\omega$ | $N(C_3) = \{1, C_3, C_3^2\}$ | 2 |
| $\Gamma_8$ | $\{C_3, C_3^2\}$ | $E_{\bar{\omega}}$ | $N(C_3) = \{1, C_3, C_3^2\}$ | 2 |

**Fusion:**

for anyons $A, B, C, ...$,

$$A \otimes B = \bigoplus_C C,$$

or (physics notation)

$$A \times B = \sum_C N_{AB}^C C.$$

where the left hand side is tensor product, and right hand side is direct sum (i.e. direct product). Comment: difference between tensor product and direct product? tensor product is locally independent (local dof), whereas direct product is globally independent (global dof)!!

**Spin**: in the table above we chose one representative element in each conjugacy class $g$. Since $g$ commute with $Z(g)$, it is a constant matrix in the $d$-dimensional irrep $R_{Z(g)}$ (by Schur's lemma): $e^{2\pi s_g} 1_{d\times d}$, where 1 is identity matrix. We call the number $s_g$ the spin of this superselection sector. I did not find good physical intution for this.

**Braiding:**

braiding for fluxes $a, b \in G$: naively one would thought it is $(a, b) \to (b, a)$, but this is not compatible with fusion (when $G$ is nonabelian)! Turns out the correct rule is $(a, b) \to (aba^{-1}, a)$.

Braiding operator $R$: for $a, b \in G$, $R_{ab}$: $|a, b\rangle \to |aba^{-1}, a\rangle$.

Braiding is important in that **braiding twice, $R^2$, is the "AB" operator**. This is the process that detect charges! Using $R^2$, one can assign a representation of $G$ for all the fluxes, hence assigning charges to them. This is the way to find all the anyons.

## 11.5 Some data

For Toric code (i.e. $D(S_2)$), the fusion matrices $N_{bc}^a$ for $a = 1, 2, 3, 4$ in the basis above are

$$N^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, N^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, N^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, N^4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (28)$$

For $D(S_3)$, the fusion matrices $N^a_{bc}$ for $a = 1, 2, ..., 8$ in the basis above are

$$N^1, N^2, N^3, N^4, N^5, N^6, N^7, N^8 =$$

$$\begin{pmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1 \end{pmatrix}, \begin{pmatrix} 0&1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1 \end{pmatrix}, \begin{pmatrix} 0&0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 1&1&1&0&0&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&0&0&0&1&1\\ 0&0&0&0&0&0&1&1\\ 0&0&0&0&0&1&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 1&0&1&0&0&1&1&1\\ 0&1&1&0&0&1&1&1\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0 \end{pmatrix}$$

$$\begin{pmatrix} 0&0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&1&1&0&0&1&1&1\\ 1&0&1&0&0&1&1&1\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&1\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 1&1&0&0&0&1&0&0\\ 0&0&1&0&0&0&0&1\\ 0&0&1&0&0&0&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&1&0&0&0&0&1\\ 1&1&0&0&0&0&1&0\\ 0&0&1&0&0&1&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0\\ 0&0&0&1&1&0&0&0\\ 0&0&0&1&1&0&0&0\\ 0&0&1&0&0&0&1&0\\ 0&0&1&0&0&1&0&0\\ 1&1&0&0&0&0&0&1 \end{pmatrix} \tag{29}$$

For example, we have $\Gamma_4 \times \Gamma_4 = \Gamma_1 + \Gamma_3 + \Gamma_6 + \Gamma_7 + \Gamma_8$, as given in Preskill (9.47).

Useful reference: `https://web.physics.utah.edu/~lake/tqft/quantum_double.pdf`

`https://cqwbkpro.s3.eu-west-2.amazonaws.com/wp-content/uploads/2021/10/27144201/210625_CQ_Quantum-Double-A`
`pdf`

## 11.6 Kitaev honeycomb model

Kitaev Honeycomb model: see the original paper `https://arxiv.org/abs/cond-mat/0506438` for an introdcution. The corresponding Hamiltonian $H$, corresbonding Hilbert space has a direct product structure: $\mathcal{L} = \bigoplus_w \mathcal{L}_w$, where $w = \{w_1, ..., w_N\}$ is the flux structure of all hexagons (hexagon number is $N$).

When writing $\sigma^{x,y,z} \to \widetilde{\sigma}^{x,y,z} = ib^{x,y,z}c$, the Hilbert space is enlarged: $H \to \widetilde{H}$, $\mathcal{L} \subset \widetilde{\mathcal{L}}$. The projector can be written as $P = \prod_j P_j$, where $j = 1, 2, ..., 2N$ labels lattice sites and $P_j$ projects onto the space with $D_j \equiv b^x_j b^y_j b^z_j c_j = 1$ (note that $D_j = \pm 1$), so that $P_j = \frac{1}{2}(1 + D_j)$. Note that there is also a direct product structure for $\widetilde{\mathcal{L}}$: $\widetilde{\mathcal{L}} = \bigoplus_u \widetilde{\mathcal{L}}_u$, where each $\widetilde{\mathcal{L}}_u$ corresponds to one quadratic Hamiltonian $\widetilde{H}_u \sim ic^\dagger Ac$ with a definite sector for $b$'s. We have $\mathcal{L} = P\widetilde{\mathcal{L}}$.

A natural question is to ask what is the relation between $\mathcal{L}_w$ and $\widetilde{\mathcal{L}}_u$. To see this, note we have $w_{123456} \sim u_{12}u_{23}u_{34}u_{45}u_{56}u_{61}$ (with some minus sign). We see that a sector $\widetilde{\mathcal{L}}_u$ with definite $u$'s corresponds to a sector with definite $w$'s. On the contrary, a definite $w$ can correspond to different $u$'s. Now the question is what happens if we replace "corresponds" by te projector $P$, i.e., the question is, what is $P\widetilde{\mathcal{L}}_u \subset \mathcal{L}$? An easy thing to notice is that $w = \{w_1, ..., w_m\}$ commutes with $D_j$ for all site $j$, so that $w$ commutes with $P$. Therefore, if we starts with a definite $u$ which has a definite $w$ before projection, the projection $P$ will not alter this value, so we do have:

For a given $u$, which fixes a $\widetilde{\mathcal{L}}_u$ and determines a $w$ as described above, then $\emptyset \neq P\widetilde{\mathcal{L}}_u \subseteq \mathcal{L}_w$.

the next question is wehther $P\widetilde{\mathcal{L}}_u$ covers the full $\mathcal{L}_w$ space. Suppose $U = \{u_1, u_2, ...\}$ are all the $u$ configurations such that the hexagon flux corresponds to $w$. It is easy to see that any two configurations in $U$ can be mapped to each other by the product of $D_j$'s on some lattice sites. crucially, note that $D_j P = P$, so that really we have $P\widetilde{\mathcal{L}}_u = P\widetilde{\mathcal{L}}_{u'}$ for any $u, u' \in U$. Therefore, we have

For any $u$ and its corresponding $w$, $P\widetilde{\mathcal{L}}_u = \mathcal{L}_w$.

Now let's do some basic counting. Suppose the honeycomb lattice has $N$ hexagons. A hexagon is just a unit cell (or more precisely, equal in area to a diamond unit cell), so there are in total $2N$ sites, each site has a spin degree of freedom so the physical Hilbert space dimension is $|\mathcal{L}| = 2^{2N}$. there are $2^N$ different flux configurations for $w$, so the dimension of the Hilbert subspace for each flux configuration is $|\mathcal{L}_w| = 2^N$. Now the enlarged Hilbert space dimension is $|\widetilde{\mathcal{L}}| = 4^{2N} = 2^{4N}$ since it has four Majorama fermions on each site, which can be viewed as two complex fermions on each site. According to the above analysis, the enlarged Hilbert subspace corresponding to a fixed flux configuration should have dimension $2^{4N}/2^N = 2^{3N}$, i.e. $|\bigoplus_{u\in U} \widetilde{\mathcal{L}}_u| = 2^{3N}$, which can be checked by using $|\widetilde{\mathcal{L}}_u| = 2^{(2N/2)} = 2^N$ (each $\widetilde{\mathcal{L}}_u$ is

the Hilbert space for $c$'s, i.e. one Majorana fermion per site) and $|U| = 2^{2N}$ (the way of applying $D_j$'s) so that we do have $|\bigoplus_{u\in U}\widetilde{\mathcal{L}}_u| = |U|\cdot|\widetilde{\mathcal{L}}_u|$. So we actually have

$$|\widetilde{\mathcal{L}}_u| = |P\widetilde{\mathcal{L}}_u| = 2^N.$$

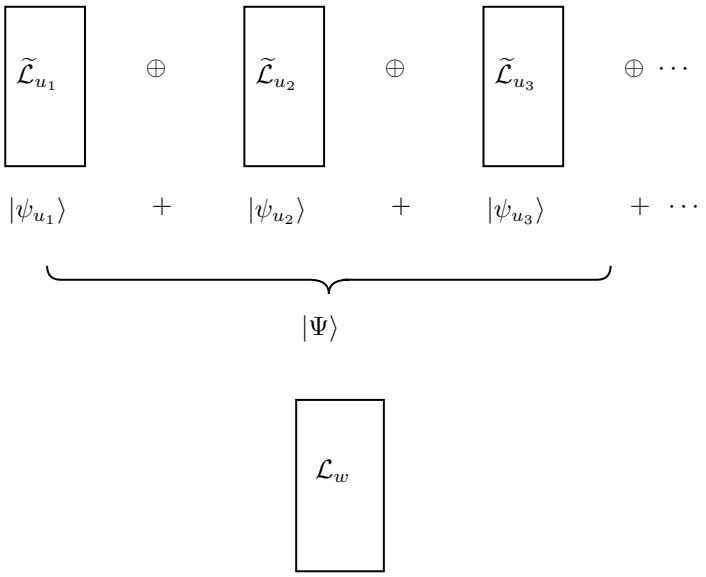
what this means is that, while $P$ is a true projector for $\widetilde{\mathcal{L}}$, it is not a true projector for subspace $\widetilde{\mathcal{L}}_u$, in particular, $|(P\widetilde{\mathcal{L}}_u)\cap\widetilde{\mathcal{L}}_u| = 1$, since $P$ consists of terms of product of $D_j$'s, and any non-identity product maps $\widetilde{\mathcal{L}}_u$ to some other subspace. So, $P$ actually has a nontrivial structure, hidden in the direct sum of the Hilbert subspace below

$$\bigoplus_{u'\in U}\widetilde{\mathcal{L}}_{u'} \supsetneq P\widetilde{\mathcal{L}}_u = \mathcal{L}_w \quad \forall u\in U. \tag{30}$$

The true meaning of this is on site. Define $\mathcal{D}$ to be the group generated by $D_j$', then $U = \{D(u)|D\in\mathcal{D}\}$, $|\mathcal{D}| = |U| = 2^{2N}$, $P = \prod_i P_i = \prod_i \frac{1+D_i}{2} = \frac{1}{|\mathcal{D}|}\sum_{D\in\mathcal{D}} D$. For an eigenstate $|\widetilde{\psi}_u\rangle$ of $\widetilde{H}_u$, we have $P|\widetilde{\psi}_u\rangle = \frac{1}{|\mathcal{D}|}\sum_{D\in\mathcal{D}}|\widetilde{\psi}_{D(u)}\rangle$, $|\widetilde{\psi}_{D(u)}\rangle\in\widetilde{\mathcal{L}}_{D(u)}$.

According to Kitaev, the operators $D_j$'s are the operators for gauge transformations. To make sense of this, let us summarize the structure of the problem we are considering:

Consider the situation where a physical state is the superposition of states, each of which lives in a different sector (in the sense of direct product) of the Hilbert space. The sectors are mapped to each other under some operators, and the states are mapped to each other under the same set of operators. In this situation, each sector is called a gauge-fixed sector, and the operators that map each other are called gauge transformations. The state in each gauge-fixed sector is not physical, but only the superposition gives the true physical state.




Next question: what is the ground state of the original Kitaev spin Hamiltonian? The Majorana model, $\widetilde{\mathcal{H}}_u$, is of course very easy to solve. Let's say we get its ground state, which is $|\widetilde{\psi}_u\rangle$. But what is the ground state of $H$? We know that $|\psi_w\rangle = P|\widetilde{\psi}_u\rangle\in\mathcal{L}_w$, but is $|\psi_w\rangle$ even an eigenstate of $H$? It seems that the Kitaev paper turned to analyzing $\widetilde{\mathcal{H}}_u$ thereafter, without going back to the very original (and important) question of what the ground state of the original spin Hamiltonian $H$ is.

According to `https://journals.aps.org/prb/abstract/10.1103/PhysRevB.84.165414`, the energies of projected (i.e. symmetrized) and unprojected (i.e. unsymmetrized) states are the same and the spin correlation functions (whose certain exact properties were discussed in Refs. 2 and 20) can be conveniently computed with unprojected eigenstates. Ref. 2 and 20 are:

`https://iopscience.iop.org/article/10.1088/1751-8113/41/7/075001/pdf`

`https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.98.247201`

We try to summarize from another perspective. We have

$$H = \mathrm{diag}(H_{w^{(1)}}, H_{w^{(2)}}, ..., H_{w^{(2^N)}}),$$

where each $w^{(i)}$ is a flux configuration for all $N$ hexagons (the $j$-th hexagon has flux $w_j^{(i)}$, with $j = 1, 2, ..., N$). Here $H$ is a $2^{2N}\times 2^{2N}$ matrix, and $H_{w^{(i)}}$ is a $2^N\times 2^N$ matrix. For each $w\in\{w^{(1)}, ..., w^{(2^N)}\}$, there are $2^{2N}$ $u$ configurations that

corresponds to $w$. Denote the set of these $u$ configurations as $U_w = \{u_w^{(1)}, u_w^{(2)}, ..., u_w^{(2^{2N})}\}$. Corresponding to Eq. (30), we have

$$H_w \cong P\,\mathrm{Diag}(\widetilde{H}_{u_w^{(1)}}, \widetilde{H}_{u_w^{(2)}}, ..., \widetilde{H}_{u_w^{(2^{2N})}})\,P,$$

where each $\widetilde{H}_u$ is a $2^N \times 2^N$ Hamiltonian. As mentioned before, $P = \frac{1}{|\mathcal{D}|}\sum_{D\in\mathcal{D}} D$, where $\mathcal{D} = \langle D_1, D_2, ..., D_{2N}\rangle$, and $|\mathcal{D}| = 2^{2N} = |U_w|$. So we have

$$P\widetilde{H}_{u_w}P = P\left(\frac{1}{2^{2N}}\sum_{i=1}^{2^{2N}}\widetilde{H}_{u_w^{(i)}}\right)P, \qquad \forall u_w \in U_w.$$

Note that all $2^{2N}$ Hamiltonians $\widetilde{H}_{u_w}$ have identical spectrum since they are related by unitary transformations (i.e. the gauge transformation $\mathcal{D}$).

Now we summarize a few results about gauge transformations. For each gauge transformation $D \in \mathcal{D}$ and each $u_w \in U_w = \{u_w^{(1)}, u_w^{(2)}, ..., u_w^{(2^{2N})}\}$, we have (note that $D^\dagger = D$)

$$D^\dagger \widetilde{H}_{u_w} D = \widetilde{H}_{u'_w}, \qquad \text{where } u'_w = D(u_w), \quad \text{so that } [D, \widetilde{H}_{u_w}] \neq 0.$$

This means that gauge transformation $D$ permutes the matter field Hilbert subspaces, and consequently, $[D, \widetilde{H}] \neq 0$. However, for the physical Hamiltonian $H$, we have (note that $D^\dagger = D$)

$$D^\dagger H D = H \qquad \Leftrightarrow \qquad [D, H] = 0,$$

this of course is not surprising at all since we have $H = P\widetilde{H}P$ and $[D, P] = 0$ for any $D \in \mathcal{D}$. We can also direct verify it in the extended Hilbert space $\widetilde{\mathcal{L}}$ using Majorana fermions: since each term in $H$ has the form of *cbbc* and has either zero or two Majorana fermion overlap with $D_j$, so $D$ and $H$ commute. From another perspective, the physical Hilbert space $\mathcal{L}$ can be viewed as defined by $[D, H] = 0$: define the subspace $\mathcal{L}' \equiv \oplus_{D\in\mathcal{D}} D\widetilde{\mathcal{L}}$, then we have $\mathcal{L}' = D\mathcal{L}'$ for any $D \in \mathcal{D}$ so that $\mathcal{L}' = \mathcal{L}$. This is exactly what we mean by "physical state must be invariant under gauge transformations".

Since $D$ commute with $H$, we might want to use the eigenvalue of $D$ to block diagonalize $H$. But actually this is trivial since any $|\xi\rangle \in \mathcal{L}$ satisfies $D_j|\xi\rangle = |\xi\rangle$ for all $j = 1, ..., 2N$ (see Eq. (11) of Kitaev's paper).

Now, we have (using Eq. (14) of Kitaev)

$$\begin{aligned} H|\Psi_w\rangle &= HP|\widetilde{\Psi}_u\rangle = P\widetilde{H}PP|\widetilde{\Psi}_u\rangle = P\widetilde{H}P|\widetilde{\Psi}_u\rangle = P\widetilde{H}\frac{1}{|\mathcal{D}|}\sum_{D\in\mathcal{D}} D|\widetilde{\Psi}_u\rangle = \frac{1}{|\mathcal{D}|}P\sum_{D\in\mathcal{D}}\widetilde{H}D|\widetilde{\Psi}_u\rangle \\ &= \frac{1}{|\mathcal{D}|}P\sum_{D\in\mathcal{D}}\widetilde{H}|\widetilde{\Psi}_{D(u)}\rangle = \frac{1}{|\mathcal{D}|}P\sum_{D\in\mathcal{D}}\widetilde{H}_{D(u)}|\widetilde{\Psi}_{D(u)}\rangle = \frac{1}{|\mathcal{D}|}P\sum_{D\in\mathcal{D}}E|\widetilde{\Psi}_{D(u)}\rangle = EP|\widetilde{\Psi}_u\rangle = E|\Psi_w\rangle, \end{aligned} \tag{31}$$

so actually the state Eq. (14) is an eigenstate of $H$. Furthermore, since the $H_w$ is $2^N \times 2^N$ and $\widetilde{H}_u$ is also $2^N \times 2^N$, the eigenstates of $H_w$ is actually in one-to-one correspondence with the eigenstates of $\widetilde{H}_u$ via projection. The low energy property (or, really, any property) of the physical states can therefore be inferred by just examining a single free Majorana model. This marks the solvability of the Kitaev honeycomb model.

Comment: quite often, $\widetilde{H}_u$ with gauge related $u$ are viewed as sharing the same Hilbert space, since they are just Hamiltonian for the same set of Majoranas $c$. But according to the above analysis, they are really different diagonal blocks of the totla Hamiltonian, so that the eigenstate of one and the eigenstate of the other are always orthogonal. Note that these two are really complementary view points: the first view point is the usual "gauge fixing" treatment, while the second view point has the advantage of presenting the structure more clearly and introducing gauge fluctuations. But at the end of the day, it is the gauge invariant state that is the physical one; and this is the state that "entangles" different Hamiltonian blocks.

## 11.7 Steven Simon's talk at les Houches

Ising anyons: $\psi \times \psi = I$, $\psi = \sigma$, $\sigma \times \sigma = I + \psi$, and $\sigma$ is its own antiparticle, so $\sigma = \bar{\sigma}$.

Fusion multiplicity:

$N_{ab}^c$ graphically has a,b down and c up; we can also have $a, b, \bar{c}$ in, $a, b$ fuse first to $c$ and then $c$ and $\bar{c}$ fuse to nothing (with nothing coming out).

$N_{ab}^c = N_{ba}^c = N_{ab\bar{c}} = N_{\bar{a}\bar{b}}^{\bar{c}}$.

Associativity constraint: (om)

$[N_a, N_c] = 0$, where we define $N_{ab}^c = [N_a]_b^c$. We have $N_{\bar{a}} = N_a^T$. Meaning we have a bunch of commuting matrices, which also commte with their transposes. (These are normal matrices, which means they are simultaneously diagonalizable by some unitary matrices:

$$UN_aU^\dagger = D_a,$$

These define modular anyon theories, where $U_{ab}$ can be roughly obtained as the linking of $a$ and $b$.

The "$F$" diagram for five $a$'s: to compute the number of outcomes, we calculate $\sum_{b,c,d} = N_{aa}^b N_{ab}^c N_{ac}^d N_{ad}^e = [N_a^4]_a^e$ as a matrix multiplication. More generally, the dim of $M$ $a$'s is $[N_a^{M-1}]_a^e$ which equals roughly the largest eigenvalue of $N_a$ (which we call $d_a$), to the $M$'th power.

Fibonacci: $\tau \times \tau = I + \tau$, $\{I, \tau\}$. We have $N_\tau = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, having largest eigenvalue $\phi := \frac{1+\sqrt{5}}{2}$.

Ising: $N_\psi = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $N_\sigma = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, with $d_\sigma = \sqrt{2}$ (the largest eigenvalue).

Hw: $d_a d_b = \sum_c N_{ab}^c d_c$.

"Quantum dim is the rep. of the fusion algebra."

Hw: No theory where $a \times a = a$ only, except $a = I$.

"Identity only appear on RHS once – an axiom."

Sphere $\to$ Sphere with two holes, trapping charges $a$ and $\bar{a}$, $\to$ torus $T^2$. Then conclude: Dim of $T^2$ = number of $a$'s.

The pants/fat/tube fusion diagram (of type 2-0), with input $a, b$ below and output $c$ up: Dim(the dim. of Hilbert space on this manifold) $= N_{abc}$.

Glueing a 0-3 and a 3-0 together, we get $\text{Dim}(g = 2) = \sum_{abc} N_{abc} N_{abc} N_{\bar{a}\bar{b}\bar{c}}$.

Hw: find the GS digeneracy of Fib anyons on $g = 4$.

Toric code to Kitaev ('97) quantum double: $V_\alpha = \prod_{star} V_z$, $P_\beta = \prod_{plaq} \sigma_x$. star operator $\tilde{V}_\alpha = (1 + V_\alpha)/2$ and plaquette operator $\tilde{P}_\beta = (1 + P_\beta)/2$, so that they are projectors.

Choose a discrete group $G$, and any graph with oriented edge (orientation specified by $g$ or $g^{-1}$).

$$H = -\sum_\alpha V_\alpha - \sum_\beta P_\beta$$

where (for e.g. a four valent star) $V_\alpha = \delta_{g_1 g_2 g_3 g_4 = id}$, and (for e.g. a square plaq.) $P_\beta$: first define $P(h)\cdot$Plaquette labeled by ordere Plaquette labeled by ordered$(hg_1, hg_2, hg_3, hg_4)$, where the meaning of $h$ is "inserting flux". Then, $P_\beta = \frac{1}{|G|} \sum_{h \in G} P_\beta(h)$. First, check that this is a projector.

To check that $V$ and $P$ comppute: define a star (with arrows all going out) and a plaquette made ouf the left and down edges. Note that $V_\alpha = \delta_{g_1 g_2 g_3 g_4 = id}$. $PV$ is to make $(g_4, g_3) \to (hg_4, g_3 h^{-1})$. (....)

Then: G.S. is the configuration with all $V$'s and $P$'s having value $+1$, which is the same as summing over all edge configs such that all $V$'s and $P$'s having value $+1$.

Imagine in the G.S., multiply one edge with $h$ ("creating defects".). Then for nonabelian group: inorder to saparate the two defects created by the $h$ multiplication (i.e. making one defect propagate), need to multiply the adjecent edge by $xhx^{-1}$ (rather than $h$), where $x \in G$ is the vertical bond right below the vertex connecting the two horizontal bonds that we are multiplying by $h$ and $xhx^{-1}$, respectively. These vertex excitations are called "fluxons"

Continue. For plaquette operator: for a series of vertical bonds that are adjacent horizontally, labeled by $g_1, g_2, g_3, ...$, $\rho^R_{n,n'}(g_4 g_3 g_2 g_1)$:

$\rho^{R_a} \otimes \rho^{R_b} = U^\dagger[\rho^{c_1} \oplus \rho^{c_2} \oplus \oplus ...]U$ (...) These are called "chargeons"

A procedure to reduce the plquettes and vertices to count the dim of G.S. on the torus: finally only two edges (with element $g_1, g_2$) and one plaquette left. (Note this reduction only works if we want to count ground states.) Then: $V = \delta_{g_1 g_2 g 1^{-1} g_2^{-1} = id}$, meaning that $g_1$ and $g_2$ must commute. For a general state $|a, b\rangle$, the plaquette operating on lt gives $P|a, b\rangle = \frac{1}{|G|} \sum_{h \in G} |hah^{-1}, hbh^{-1}\rangle$. According to Burnside's lemma (number of orbits is equal to $\frac{1}{|G|} \sum_{h \in G}$(number of objects in a set fixed by $h$), so we are looking for $h$ that commute with $a$ and $b$, where recall that $a$ and $b$ also commute. Therefore the number of ground state is equal to the number of commuting triples $(a, b, c)$, divided by $|G|$.

E.g.: for the group $S_3 = \{e, r, r^{-1}, x_1, x_2, x_3\}$, then $\{e, r, r^{-1}\}$ has 27 equivalences, $\{e, e, x_i\}$ has 9, and $\{e, x_i, x_i\}$ has 9, and $\{x_i, x_i, x_i\}$ has 9. So $48/|S_3| = 8$ G.S.'s.

We have that different basis are related by the $F$ symbols: $\sum_f [F_e^{abc}]_{df}$ ($F$ relates a left comb to a right comb). We usually choose s.t. $F_c^{Iab} = F_c^{aIb} = F_c^{abI} = 1$.

For Fibonacci: $F_I^{\tau\tau\tau} = 1$; where $F_\tau^{\tau\tau\tau} = \begin{pmatrix} \phi^{-1} & \phi^{-1/2} \\ \phi^{-1/2} & -\phi^{-1} \end{pmatrix}$. We can define states $|0\rangle$ = leftcomb of $\tau, \tau, \tau$ fusing to $\tau$ with an intermediate state of $I$, and $|1\rangle$ = leftcomb of $\tau, \tau, \tau$ fusing also to $\tau$ with an intermediate sate of $\tau$. We then define the right comb corresponding states as $|0'\rangle$ and $|1'\rangle$. Then we have that $\begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix} = F_\tau^{\tau\tau\tau} \begin{pmatrix} |0'\rangle \\ |1'\rangle \end{pmatrix}$.

$F$ satisfies the pentagon equation, associated with the two ways of changing a leftcomb to a rightcomb.

If the fission is described by gauge transformation that $Y_c^{ab}$(as a diagram) $= Y_{\tilde{c}}^{\tilde{a}\tilde{b}} U_c^{ab}$, then one can show that

$$(\widetilde{F}_e^{abc})_{df} = \frac{U_e^{af} U_f^{bc}}{U_d^{ab} U_e^{dc}} (F_e^{abc})_{df}.$$

Def. A set of particles with $N$'s and $F$'s is a uniatry fusion category.

Rigidity: (Oceano Ioanou): finite number of solutions to the pendagon equations.

Def. Braiding. A naive fission from $c$ to get $a, b$, denote the final state $|\psi\rangle$; a less naive fission with $a, b$ exchanging position (or rotating 180 deg) we call $R_c^{ab}|\psi\rangle$, or $[R_c^{ab}]^{-1}|\psi\rangle$, depending on the detailed convention.

Example in Fibonacci anyons: $R_\tau^{\tau\tau} = e^{3\pi i/5}$, and $R_I^{\tau\tau} = e^{-4\pi i/5}$.

Consistency condition between the $F$'s and $R$'s — the hexagon equation: $RFR = FRF$ (details see the Simon textbook)

Particles + $N$'s + $F$'s + $R$'s (where the $F$ and $R$'s are unitaries), called UBTC.

Consider the case where $f$ braids 360 around $a$. If this is the same as bo braiding $f$ around $a$, we call $f$ "transparent". If the only transparent object is $I$, then called modular tensor category. If all transparent, then called symmetric tensor category.

We can always undo any knots in 3D space, therefore the point particles in 3D is characterized by the symmetric tensor category.

Sign$[F_a^{aaa}]_{II}$: the Frobenius-Schur indicator.

Lein-Wen model: choose a UFC, so that each trivalent vertex is labeled by three particles $a, b, c$. we have $H = -\sum V - \sum P$, where $V$ basically this rule, and $P$ acting on the plaquette gives $\frac{1}{D^2}\sum_a d_a$times having an $a$ in the plaquette and we have to push this $a$ into the vertices using the $F$ symbols. We have $|GS\rangle = \sum_{\text{all labels satisfying } V} W(\text{labeling})|\text{labeling}\rangle$.

# 12 XY model and Berezinskii-Kosterlitz-Thouless (BKT) transition

## 12.1 References

- Section 3.3 of Nagaosa's book "Quantum Field Theory in Condensed Matter Physics".
- Section VII of the review paper by Kogut, "An introduction to lattice gauge theory and spin systems", `https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.51.659`,
- Section IV of the review paper by Savit, "Duality in field theory and statistical systems" `https://journals.aps.org/rmp/pdf/10.1103/RevModPhys.52.453`,
- The concept of charged BKT phase transition, whose elementary introduction can be found in `https://arxiv.org/pdf/1909.01820.pdf`

## 12.2 The model

XY model on a 2D square lattice:

$$H = -J\sum_{n,\mu} \boldsymbol{s}_n \cdot \boldsymbol{s}_{n+\mu} = -J\sum_{n,\mu} \cos(\theta_n - \theta_{n+\mu}) = -J\sum_{n,\mu} \cos \Delta_\mu \theta_n, \tag{32}$$

## 12.3 Phase analysis from correlation function

Spin-spin correlation function at high and low temperature $T$ shows it has two phases:

$$\langle e^{i\theta_i} e^{-i\theta_j}\rangle = \frac{1}{Z}\int \prod_m \theta_m e^{i(\theta_i-\theta_j)} e^{-\beta J\sum_{kl}\cos(\theta_k-\theta_l)},$$

- High $T$: expand according to the small parameter $\beta J$. The zeroth order term is just 1 which vanishes due to the integral $\int_0^{2\pi} d\theta_m e^{i\theta_m} = 0$ for $m = i$ or $j$. The first nonzero order is contributed by paths connecting $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$: each bond $\langle kl\rangle$ on this path corresponds to a phase $e^{i(\theta_k-\theta_l)}$, and all these phases along the bond will cancel (the first and the last will cancel $e^{i\theta_i}$ and $e^{-i\theta_j}$, respectively). Each bond will contribute a factor of $2\pi J/k_B T$ due to $\int_i^{2\pi} d\theta = 2\pi$. Note there are infinite such paths, but the dominant ones (least powers of $1/T$) are the ones are the shorted paths, and we get $\langle e^{i\theta_i} e^{-i\theta_j}\rangle \sim \left(\frac{J}{2k_B T}\right)^{|\boldsymbol{R}_i - \boldsymbol{R}_j|}$.

- Low $T$, thermal fluctuation should be small, suggesting $\theta_i$ varies slowly and smoothly in the system, therefore valid to expand the cosine to $(\Delta\theta_i)^2$. We get $\langle e^{i\theta_i}e^{-i\theta_j}\rangle = |\boldsymbol{R}_i - \boldsymbol{R}_j|^{-k_BT/2\pi J}$.

Two things we leared: low temperature there is no magnetization (i.e. $\langle e^{i\theta_i}\rangle \to 0$), which is in accord with Mermin-Wagner theorem; second, the low energy phase is critical (algebraic). This phase support another form of excitation, not considered in the conventional Ginzburg-Landau theory: vortices.

## 12.4 Single vortex analysis

In the continuum limit $H = \frac{J}{2}\int(\nabla\theta)^2 d^2\boldsymbol{R}$, whose equation of motion $\nabla^2\theta = 0$ supports vortex configuration $\theta = \arctan y/x$. We have $\nabla\theta = \frac{(-y,x)}{R^2}$, $\nabla^2\theta = 0$, therefore $E_{\text{vortex}} = \frac{J}{2}\cdot\int_0^{2\pi} d\theta \int d_a^{R_c} RdR \cdot \frac{1}{R^2} = \pi J \ln R_c/a$, where $a$ is lattice spacing and $R_c$ system size. Free energy $F = U - TS$ where $S = \ln W$, where $W$ is number of configurations of putting one vortex in the sample which is roughly $W = R_c^2/a^2$, therefore $F = (\pi J - 2T)\ln R_c/a$. This is the KT picture: at low temperature, the only low energy excitations are spin waves (not Goldstone modes! since in 2D the continuous symmetry is not broken), vortex-antivortex pair can appear but they are in bound state; free vortex cannot appear. As temperature grows, size of the vortex-antivortex pair grows until the size diverges at $T_c$.

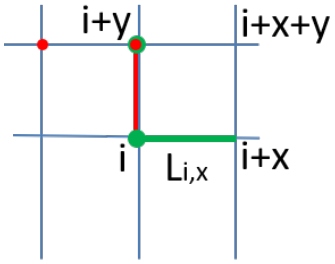
## 12.5 Duality transformation

The above picture/argument is coarse, in the sense that only one vortex is discussed.

Below is duality mapping. We will find: 1. vortex sector and spin wave sector decouple (i.e. with no interaction); the vortex sector is equivalent to a 2D Coulomb gas (Sine-Gordan model).

Now we start from the original definition of the partition function (letting $\beta J \to \beta$)

$$
\begin{aligned}
Z &= \int D[\theta]_1^N e^{\beta\sum_{i,\mu}\cos(\theta_i-\theta_{i+\mu})} \\
&= \int D[\theta]_1^N \sum_{\{l_{i,\mu}\}} e^{\sum_{i,\mu} i l_{i,\mu}(\theta_i-\theta_{i+\mu})} I_{l_{i,\mu}}(\beta) \\
&= \int D[\theta]_1^N \sum_{\{l_{i,\mu}\}} I_{l_{i,\mu}}(\beta) e^{i\sum_i\theta_i\sum_\mu l_{i,\mu}} \\
&= \sum_{\{l_{i,\mu}\}} e^{\ln I_{l_{i,\mu}}(\beta)}\delta_{\operatorname{div} l_i=0} \\
&= \sum_{\{l_{i,\mu}\}} e^{\ln I_{\epsilon_{\mu\nu}\Delta_\nu n(i)}(\beta)} \qquad \text{Physical analysis of two phases} \\
&= \sum_{\{n_i\}} e^{-\frac{1}{2\beta}\sum_{i,\mu}(n(i)-n(i-\mu))^2} \\
&= \int_{-\infty}^{\infty}\prod_i d\phi_i \sum_{\{m(i)\}=-\infty}^{\infty} e^{-\frac{1}{2\beta}\sum_{i,\mu}(\Delta_\mu\phi(i))^2 + 2\pi i\sum_i m(i)\phi(i)} \qquad \text{Interpretation of } m,\phi \\
&= Z_{\text{Spin Wave}}\cdot \sum_{\{m(i)\}=-\infty}^{\infty} e^{-2\pi^2\beta\sum_i\sum_{i'} m(i)\sum_{\boldsymbol{k}}\frac{e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_{i'})}}{4-2\cos k_x-2\cos k_y}m(i')} \\
&= Z_{\text{Spin Wave}}\cdot \sum_{\{m(i)\}=-\infty}^{\infty} e^{-2\pi^2\beta\sum_i\sum_{i'} m(i)G(i-i')m(i')} \\
&= Z_{\text{Spin Wave}}\cdot \sum_{\{m(i)\}=-\infty}^{\infty}{}' e^{-2\pi^2\beta J\sum_{i,i'\neq i} m(i)G'(i-i')m(i')} \qquad \sum_i m(i) = 0 \\
&= Z_{\text{Spin Wave}}\cdot \sum_{\{m(i)\}=-\infty}^{\infty}{}' e^{\frac{\pi^2\beta J}{2}\sum_{i,i'\neq i} m(i)m(i')+\pi\beta J\sum_{i,i'\neq i} m(i)\ln\frac{|r-r'|}{a}m(i')} \\
&= Z_{\text{Spin Wave}}\cdot \sum_{\{m(i)\}=-\infty}^{\infty}{}' e^{-\frac{\pi^2\beta J}{2}\sum_i m^2(i)+\pi\beta J\sum_{i,i'\neq i} m(i)\ln\frac{|r-r'|}{a}m(i')}
\end{aligned}
\tag{33}
$$




- Fourier decomposition

$$e^{\beta\cos\theta}=\sum_{l=-\infty}^{\infty}e^{il\theta}I_l(\beta) \tag{34}$$

  where $I_l(\lambda)=\frac{1}{\pi}\int_0^\pi e^{\beta\cos x}\cos lxdx$ is modified Bessel function of the first kind.

- $\sum_{n,\mu}l_{n,\mu}(\theta_n-\theta_{n+\mu})=\sum_n\theta_n\sum_\mu l_{n,\mu}$, integrating $\theta$'s gives $\sum_\mu l_{n,\mu}=0$, i.e. div$l_i=0$ (divergenceless at site $i$).

- continuous case $\boldsymbol{\nabla}\cdot\boldsymbol{B}=0\Rightarrow\exists\boldsymbol{A}$, i.e. $(B_x,B_y)=(\partial_yA_z,\partial_xA_z)$ s.t. $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$ and that $\boldsymbol{\nabla}\cdot\boldsymbol{B}=\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{A})=0$. By analogy: suppose $l_{i,x}=n(i+y)-n(i)$. then the $l_{i,y}$ condition is entirely constained by the $l_{i,x}$ condition, and must be $l_{i,y}=-n(i+y)+n(i+y-x)$ [e.g. it cannot be $l_{i,y}=n(i)-n(i+x)$ or anything other than that]. According to these definition we have $l_{i,-y}=-l_{i-y,y}=-(-n(i)+n(i-x))$, and $l_{i,-x}=-l_{i-x,x}=-(n(i-x+y)-n(i-x))$, it is easy to check that div$l=0$ is automatically satisfied.

- Physical analysis:

  - when $\beta\ll1$, expand using $\beta$: all $n(i)$ want to have the same value, see from the expansion of $I_p(\beta)=\delta(p)+\frac{\beta}{2}[\delta(p+1)+\delta(p-1)]+\frac{\beta^2}{4}[\delta(p+2)+\delta(p-2)+2\delta(p)]+\cdots$,

$$Z=\sum_{\{n\}}\prod_{i,\mu}(1+\beta^2/2)\delta[(\Delta_\mu n(i))^2]+\frac{\beta}{2}\delta[(\Delta_\mu\phi_j)^2-1]+\frac{\beta^2}{4}\delta[(\Delta_\mu n(i))^2-4]+O(\beta^3)+\cdots$$

    therefore $n$ is disorder parameter for the system.

  - when $\beta\gg1$, expand Eq. (34) in powers of $l^2$: that is to say, expand in powers of $(\Delta n)^2$ and only keep this leading term. This is gaussian model. Even this model has phase transition!

- Poisson summation formula[2]

$$\sum_{m=-\infty}^{\infty}h(m)=\sum_{l=-\infty}^{\infty}\int_{-\infty}^{\infty}d\phi h(\phi)e^{2\pi il\phi}.$$

- Define $\phi_i=\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}$, we have $\Delta_x\phi(i)=\sum_{\boldsymbol{k}}(e^{ik_x}-1)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}$, and therefore $\sum_i(\Delta_x\phi(i))^2=\sum_i\sum_{\boldsymbol{k}}(e^{ik_x}-1)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}}\sum_{\boldsymbol{k}'}(e^{ik'_x}-1)e^{i\boldsymbol{k}'\cdot\boldsymbol{r}_i}\phi_{\boldsymbol{k}'}=\sum_{\boldsymbol{k},\boldsymbol{k}'}(e^{ik_x}-1)(e^{ik'_x}-1)\delta_{\boldsymbol{k},-\boldsymbol{k}'}\phi_{\boldsymbol{k}}\phi_{-\boldsymbol{k}}=\sum_{\boldsymbol{k}}(e^{ik_x}-1)(e^{-ik_x}-1)\phi_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}=\sum_{\boldsymbol{k}}(2-2\cos k_x)\phi_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}$, therefore $\sum_{i,\mu}(\Delta_\mu\phi(i))^2=\sum_{\boldsymbol{k}}(4-2\cos k_x-2\cos k_y)\phi^*_{\boldsymbol{k}}\phi_{\boldsymbol{k}}$. Therefore the exponential becomes $A\equiv-\frac{1}{2\beta J}\sum_{\boldsymbol{k}}\phi^*_{\boldsymbol{k}}(4-2\cos k_x-2\cos k_y)\phi_{\boldsymbol{k}}+2\pi i\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\sum_i m(i)\phi_{\boldsymbol{k}}$, using the $ax^2+bx=a(x+b/2a)^2-b^2/4a$, we have $A=\sum_{\boldsymbol{k}}-\frac{4-2\cos k_x-2\cos k_y}{2\beta J}|(\phi_{\boldsymbol{k}}+\frac{2\pi i\sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}m(i)}{-\frac{1}{\beta J}(4-2\cos k_x-2\cos k_y)})|^2-\frac{(2\pi i\sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}m(i))^2}{-\frac{2}{\beta J}(4-2\cos k_x-2\cos k_y)}$ and $\phi_{\boldsymbol{k}}$ can be integrated out.

- spin wave propagator $G(r)=\sum_{\boldsymbol{k}}\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{4-2\cos k_x-2\cos k_y}=\int_{\text{BZ}}\frac{d^2\boldsymbol{k}}{(2\pi)^2}\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{4-2\cos k_x-2\cos k_y}$. $Z_{\text{Spin Wave}}$ is the part after integrating $\phi_{\boldsymbol{k}}$. We see that the spin wave part and the vortex part decouples. Since spin wave does not give transition, the transtion must be induced by the vortex part.

[2]Using Jacque Villain's approximation

$$e^{\lambda\cos\theta}\simeq\sum_{m=-\infty}^{\infty}e^{\lambda}e^{-\frac{\lambda}{2}(\theta-2\pi m)^2}$$

and Poisson summation formula, we can also obtain Eq. (**??**) by doing the integral for $\phi$ we get $[-\frac{\lambda}{2}(2\pi\phi-\theta)^2+il(2\pi\phi-\theta)+il\theta=-\frac{\lambda}{2}(2\pi\phi-\theta-il/\lambda)^2-l^2/2\lambda+il\theta\ ]$, note that the Gaussian integral gives $\sqrt{\pi}$ and that we are doing integral variable substitution $\phi\to\sqrt{\lambda2\pi^2}\phi=\phi'$.

- At large enough $r$, we have $G(r) \simeq -\frac{1}{2\pi}\ln\frac{r}{a} - \frac{1}{4}$; at $r = 0$, we have $G(0) \simeq \frac{1}{2\pi}\ln\frac{R}{a}$. The former shows that vortices interact through a logarithmic potential; the latter is consistent with the self energy of a single vortex. Then it is better to introduce $G(r) = G'(r) + G(0)$ where $G'(r = 0) = 0$. This way, the exponential of $Z$ gives $-2\pi^2\beta J$ times $[\sum_i m(i)]^2 G(0) + \sum_i \sum_{i'\neq i} m(i)G'(i-i')m(i')$, where the first term indicates that the biggest contribution to $Z$ is the configurations with $\sum_i m(i) = 0$, i.e. the vortices are neutral. The sum $\sum'$ means neutral configurations of vortices only. We now plug in the asymptotic form of $G(r) = -\frac{1}{2\pi}\ln\frac{r}{a} - \frac{1}{4}$; we also have $0 = \sum_{i,i'} m(i)m(i') = \sum_{i,i'\neq i} m(i)m(i') = \sum_i m^2(i)$.
- The first term gives the chemical potential of each vortex and the second term gives the logarithmic interactions between different vortices. This is exatly the partition function for Coulomb gas in 2D. The phases of it is known: when $\beta$ is large (low temperature), the chemical potential suppresses the vortices therefore there is only spin waves; the logarithmic potential says that, at low temperature, even if vortices populate, they appear in vortex-antivortex pairs (bound state). As temperature is raised, vortices are not suppressed; they are still not free but the interactions are effectively screened.

## 12.6 Interpretation of $m$ and $\phi$ as vortex and spin waves

Interpretation of $m$ and $\phi$ as vortex and spin waves: can only be seen in Villain approximation.

$$
\begin{aligned}
Z &= \int D[\theta]_1^N e^{\beta\sum_{i,\mu}\cos(\theta_i-\theta_{i+\mu})} \\
&= e^{2N\beta}\int D[\theta]_1^N \sum_{\{m_{i,\mu}\}} e^{-\frac{\beta}{2}\sum_{i,\mu} i l_{i,\mu}(\Delta_\mu\theta_i - 2\pi m_{i,\mu})^2} \qquad \text{Villain approximation} \\
&= (2\beta e^{2\beta})^N \sum_{l_{i,\mu}} \int D[\theta]_1^N \int D[\phi_{i,\mu}] e^{-\sum_{i,\mu}\frac{1}{2\beta}\phi_{i,\mu}^2 + i k_{i,\mu}(\Delta_\mu\theta_i - 2\pi l_{i,\mu})} \qquad \text{Poisson summation} \\
&\left[= (4\pi\beta e^{2\beta})^N \sum_{\{l_{i,\mu}\}}{}' \int D[\phi_{i,\mu}] e^{\sum -\frac{1}{2\beta}\phi_{i,\mu}^2 - i2\pi\phi_{i,\mu}l_{i,\mu}} \prod \delta_{\mathrm{div}\,\phi_{i,\mu}}\right] \\
&= \left[(4\pi\beta e^{2\beta})^N \sum_{\{l_{i,\mu}\}}{}' \int D[A]_N^i e^{\sum -\frac{1}{2\beta}(\Delta_\mu A_i)^2 - i2\pi\epsilon_{\mu\nu}\Delta_\mu A_i l_{i,\nu}}\right] \\
&= \left[(4\pi\beta e^{2\beta})^N \sum_{\{l_{i,\mu}\}}{}' \int D[A]_N^i e^{\sum -\frac{1}{2\beta}(\Delta_\mu A_i)^2 + i2\pi A_i\epsilon_{\mu\nu}\Delta_\mu l_{i,\nu}}\right] \\
&= (2\beta e^{2\beta})^N \int D[\theta]_1^N \sum_{\{l_{i,\mu}\}} e^{\sum -\frac{1}{2\beta}l_{i,\mu}^2 - i\Delta_\mu\theta_i l_{i,\mu}} \\
&= \sum_{\{l_{i,\mu}\}} e^{-\frac{1}{2\beta}l_{i,\mu}^2}\delta_{\mathrm{div}\,l_i=0} \\
&= \sum_{\{n_i\}} e^{-\frac{1}{2\beta}\sum_{i,\mu}(n(i)-n(i-\mu))^2}
\end{aligned}
\tag{35}
$$

we can identify

$$m(i) = \epsilon_{\mu\nu}\Delta_\mu l_{i,\nu},$$

this proves that the interpretation of $m_j$ is the vortex. $l_{i,\mu}$ represents change of $\theta_i$ along $i,\mu$; $\epsilon_{\mu\nu}\Delta_\mu l_\nu$ is the curl. More explicitly, if define

$$\Delta_\mu\Theta_i = \Delta_\mu\theta_i - 2\pi l_{i,\mu},$$

then

$$\oint_C \Delta_\mu\Theta_i = \sum m_i,$$

therefore $m_i$'s are just the vortices of the angle $\Theta_i$.

## 12.7 Renormalization group analysis of BKT transition

Starting from the line containing $\phi(i)$ and $m(i)$, we add chemical potential $y$. By keeping only $m(i) = 0, \pm 1$, we obtain a sine-Gordon model with $y = y_0 e^{-\pi^2\beta/2}$ being the chemical potential. Then the usual RG gives

$$d\mu = -\pi\beta\mu(d\Lambda/\Lambda), \qquad d\beta = -2\pi^4\alpha_2\mu^2\beta^3(d\Lambda/\Lambda^5),$$

From the flow defined by these equations, two phases are identified: one is $y = 0$ (note that physics stays the same on the flow line), whose physical meaning is no vortex, just spin wave; and a $y \to \infty$, i.e. flow to infinite temperature. More information can be extracted from these equations (see Kogut's paper).

## 12.8 General duality in U(1) theories

U(1) invariant lattice theories in $d$ dimensions with simplex number $s$; the dual theory is a $Z_\infty$ invariant theory in $d$ dimensions, with simplex number $\bar{s} = d - s$; phase diagram on $\beta$ axis is inverted (i.e. the low and high temperature phases of the original theory is mapped to the high and low temperature phases). The dual theory can be easily converted into a third form: it contains spin waves and topological excitations of the original U(1) invariant spins, the latter exist on a closed manifold of dimension $d - s - 1$. For KT model: $d = 2$, $s = 1$.

$$\begin{aligned}
Z &= \int D\theta e^{\beta \sum_s \cos(\epsilon\epsilon\Delta\theta)} \\
&= \int D\theta \prod_s \sum_{\{k\}=-\infty}^{\infty} I_k(\beta) e^{i\sum_s k\cdot\epsilon\epsilon\Delta\theta} \\
&= \sum_{\{k\}} \prod_s I_k(\beta) \int d\theta e^{-i\sum_t \theta\cdot\epsilon\epsilon\Delta k} \\
&= \sum_{\{k\}} e^{\sum_s \ln I_k(\beta)} \prod_t \delta(\epsilon\epsilon\Delta k) \\
&\approx e^{N_s\beta} \sum_{\epsilon\Delta\phi} e^{-\frac{1}{2\beta}\sum_s(\epsilon\epsilon\Delta\phi)^2} \qquad \text{low temperature approximation} \\
&= e^{N_s\beta} \int D(\epsilon\Delta\phi) \sum_{\{L\}=-\infty}^{\infty} e^{\sum -\frac{1}{2\beta}(\epsilon\epsilon\Delta\phi)^2 + i2\pi L\epsilon\Delta\phi} \\
&= e^{N_s\beta} \int D(\epsilon\Delta\phi) \sum_{\{L\}=-\infty}^{\infty} e^{\sum -\frac{1}{2\beta}(\epsilon\epsilon\Delta\phi)^2 + i2\pi\epsilon\Delta L\phi} \\
&= e^{N_s\beta} Z_0 \sum_{\{J\}}{}' e^{\beta\sum_{i,j} J_i V_{ij} J_j}
\end{aligned} \tag{36}$$

on a $d$ dimensional lattice, remember we define the simplex number $s$ to be the dimension that each term occupies (and therefore Ising/XY model has $s = 1$ which is the bond dimension and Ising gauge/abelian gauge has $s = 2$ which is the plaquette dimension). On each simplex of dimension $s - 1$, place a complex phase $Q_{\mu_1\ldots\mu_{s-1};i} = e^{i\theta_{\mu_1\cdots\mu_{s-1};i}}$, these phases interaction according to the form

$$I_{\mu_1..\mu_s;i} = \cos\left(\frac{1}{(d-s)!}\epsilon_{\mu_1\ldots\mu_s\alpha_1\ldots\alpha_{d-s}}\epsilon_{\alpha_1..\alpha_{d-s}\beta\gamma_1\ldots\gamma_{s-1}}\Delta_\beta\theta_{\gamma_1..\gamma_{s-1};i}\right) = \cos(\epsilon\epsilon\Delta\theta),$$

where the $\epsilon$ is just the fully antisymmetric tensor. Now we try to understund this term a little better: note $i$ is a $d$ dimensional vector which labels the site in the $d$ dimensional space/spacetime lattice; for each site $i$, in order to define the variable $\theta_{\mu_1\ldots\mu_{s-1};i}$ which lives on a $s-1$ simplex (i.e. boundary of the $s$ simplex), we just have to define the variable direction, which is specified by $s-1$ directions, $\mu_1, ..., \mu_{s-1}$. Now, remember that in the $XY$ model, for each site $i$, there are actually two terms, depending on the direction of the bond $x$ or $y$, where bond is simply the $s = 1$ simplex. This means that, in our generalization, fixing $i$ we can also obtain differen terms; and these terms are distinguished by there simplex directions $\mu_1...\mu_s$, and each term is exactly what we write down here, $I_{\mu_1\ldots\mu_s;i}$. Therefore we stress that, each term corresponds to one simplex (i.e. for any chosen $s$ simplex, a term should be written down). Let us understand the part in the parenthesis better: as an example, in the usual Abelian lattice gauge theory, $s = 2$, and the variable lives on bonds, i.e. $\theta_{\mu;i}$, where $i$ is the starting point of the bond and $\mu$ is the direction of the bond. Then, $\Delta_\beta\theta_{\gamma_1\ldots\gamma_{s-1};i}$ takes derivative along direction $\beta$, where $\beta$ is restricted to the only free direction in the $s$ simplex after fixing $\gamma_1...\gamma_{s-1}$, in other hand, we have

$$\Delta_\beta\theta_{\gamma_1\ldots\gamma_{s-1};i} = \theta_{\gamma_1\ldots\gamma_{s-1};i+\gamma_s} - \theta_{\gamma_1\ldots\gamma_{s-1};i}, \qquad \beta = \gamma_s.$$

Note the introduction of $\epsilon\epsilon$ sums over the codimensions, which gives $(d-s)!$ terms; these are redundant and that's why we need to divide by $(d-s)!$. But through this $\epsilon\epsilon$ we can indeed introduce the correct sign convention when taking derivative $\Delta_\beta$ along different coordinate $\beta$. One then follows the KT transformations to proceed.

### 12.9 Charged BKT

Charged superfliud or superconductor:

- Anderson-Higgs mechanism makes all excitations massive
- vortex excitations: winding counteracted by the penerating magnetic field, thus energy of vortex finite $\sim \ln \lambda_L/a$, $\lambda_L = \sqrt{\frac{m^*}{\mu_0 e^{*2}|\psi|^2}}$.
- From the vortex: it seems $T_{\mathrm{BKT}} \to 0$
- reality: dimensionality (vortex couple to 3D E&M): effective penetration depth $\lambda_{L,2D} = \lambda_L \frac{\lambda_L}{w}$ can still be very large if system thickness $w$ is small. Therefore still at finite temperature.

## 13 Quantum Spin Liquids and Gauge Theories

### 13.1 Spin systems that are exotic

Any system with odd number of electrons per unit cell: must a metal (gapless)!

Lieb-Schultz-Mattis Theorem: A local SU(2) invariant spin Hamiltonian on a lattce system with a total half-integer spin per unit cell and PBC, must have at least one low-lying state whose energy scales to zero with system size $L$.

It involves different mechanisms:

- Symmetry breaking systems:
  - Spin rotation symmetry
  - lattice symemtry breakng...
- Non-symmetry-breaking systems: bulk gapless excitation or ground state degeneracy: both exotic! (paramagnet is gapped and no ground state degeneracy)
  - Ground state degeneracy: (on a torus)+no Symmetry breaking, must be topologically ordered
  - Gapless bulk: not protected by Goldstone th, but by non-trivial entanglement of the ground satte!

Its generalizations: by Oshikawa, etc.
Instances corresponding to those mechanisms:

- Symmetry breaking systems:
  - Magnetic order: Ferromagnet
  - Valence bond solid
- Non-symmetry-breaking systems: bulk gapless excitation or ground state degeneracy: both exotic!
  - Toric code/$\mathbb{Z}_2$ lattice gauge theory, $\mathbb{Z}_2$ spin liquids
  - Gapless bulk: U(1) spin liquids

Hamiltonian and lattice corresponding those mechanisms:
Heisenberg model

$$H = \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} J_{\mathbf{r},\mathbf{r}'} \vec{S}_{\mathbf{r}} \cdot \vec{S}_{\mathbf{r}'}$$

- Symmetry breaking systems:
  - $J < 0$
  - $J > 0$, triangular lattice
- Non-symmetry-breaking systems?

## 13.2 Trick: partons

Write

$$S^i_{\mathbf{r}} = \frac{1}{2} f^\dagger_{\mathbf{r}\alpha} \sigma^i_{\alpha\beta} f_{\mathbf{r}\beta}$$

get

$$H = \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} -\frac{1}{2} J_{\mathbf{r},\mathbf{r}'} f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}'\alpha} f^\dagger_{\mathbf{r}'\beta} f_{\mathbf{r}\beta} + \text{Const.}$$

Problematic: enlarged the Hilbert space: each site has four states now. So need to impose (Baskaran et al., 1987)

$$f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}\beta} = 1, \quad f_{\mathbf{r}\alpha} f_{\mathbf{r}\beta} \epsilon_{\alpha\beta} = 0.$$

Goal: write into a mean field theory:

- Zeroth-order approximation:

$$\langle f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}\alpha}\rangle = 1, \qquad \langle f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}'\beta}\rangle = \chi_{\mathbf{r},\mathbf{r}'}$$

$$H_{\text{MF0}} = \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} -\frac{1}{2} J_{\mathbf{r},\mathbf{r}'} \left[ (f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}'\alpha} \chi_{\mathbf{r}',\mathbf{r}} + h.c.)\right] + \sum_{\mathbf{r}} a_0(\mathbf{r}) (f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}\alpha} - 1)$$

  and ignore fluctuations of $a_0(\mathbf{r})$ and $\chi_{\mathbf{r},\mathbf{r}'}$(i.e. time dependence)[3]

- First-order approximation: keep the fluctuations if $a_0$ and phase fluctuations of $\chi$:

$$H_{\text{MF1}} = \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} -J_{\mathbf{r},\mathbf{r}'} \left[ (f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}'\alpha} \bar{\chi}_{\mathbf{r}',\mathbf{r}} e^{-i a_{\mathbf{r}',\mathbf{r}}} + h.c.)\right] + \sum_{\mathbf{r}} a_0(\mathbf{r}) (f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}\alpha} - 1)$$

  , where $\bar{\chi}$ is the solution to zeroth-order MF problem. First-order MF has gauge symmetry:

$$a_{\mathbf{r},\mathbf{r}'} \to a_{\mathbf{r},\mathbf{r}'} + \theta_{\mathbf{r}} - \theta_{\mathbf{r}'}, \qquad f_{\mathbf{r}} \to f_{\mathbf{r}} e^{i\theta_{\mathbf{r}}}.$$

In summary:

- Zeroth-order MF: theory of free spinon excitations $f_{\mathbf{r}\alpha}$.
- First-order MF: theory of spinon excitations coupled to $U(1)$ gauge field! or: spinons interact via gauge field! –spinons exications in pair!

## 13.3 Partons: physical or not?

What we just did:

- 0-th MF: bosonic excitation ⇒ single spinon excitation (unphyisical?)
- 1st MF: gauge fields glue two spinon back in excitation, to give bosonic excitation.

It seems: 0-th is wrong; 1st did nothing more.
But in fact: 0-th can be correct, 1st has something more!

- 0-th MF: excellent effective description of the deconfined phase
- 1st MF: have two phases! confined and deconfined phases!

[3] A check: if we put fluctuations back we get

$$Z = \int \mathcal{D}f \mathcal{D}\left[a_0(\mathbf{r})\right] \mathcal{D}\chi_{\mathbf{r},\mathbf{r}'} e^{i\int dt \left(\mathcal{L} - \sum_{\mathbf{r}} a_0(\mathbf{r},t)(f^\dagger_{\mathbf{r}} f_{\mathbf{r}} - 1)\right)},$$

where

$$\mathcal{L} = \sum_{\mathbf{r}} f^\dagger_{\mathbf{r}} i\partial_t f_{\mathbf{r}} - \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} -\frac{1}{2} J_{\mathbf{r},\mathbf{r}'} \left[ (f^\dagger_{\mathbf{r}\alpha} f_{\mathbf{r}'\beta} \chi_{\mathbf{r}',\mathbf{r}} + h.c.) - |\chi_{\mathbf{r},\mathbf{r}'}|^2 \right].$$

Integrating $a_0$ and $\chi$ gives the original $H$. These fluctuations describe the collective excitations above the MF GS.

So, 0-th MF is physical, and its excitations are just as well contained in the original spin Hamiltonian!
Q: how do we see this:
A: Gutzwiller projection!

$$\Psi^{(\chi)}_{\text{spin}}(\{\alpha_{\mathbf{r}}\}) = \langle 0_f | \prod_{\mathbf{r}} f_{\mathbf{r}\alpha_{\mathbf{r}}} | \Psi^{\chi}_{\text{MF0}} \rangle,$$

where the physical spin wavefunction is just

$$|\Psi\rangle = \sum_{\{\alpha_{\mathbf{r}}\}} \Psi^{(\chi)}_{\text{spin}}(\{\alpha_{\mathbf{r}}\}) |\{\alpha_{\mathbf{r}}\}\rangle.$$

Also notice: two mean-field Ansatze, related by gauge tranformation, give the same physical spin state. So we call the whole procedure *projective construction*.
Back to the understanding of deconfined phase: how can we understand this?
Naively,

no symmetry breaking

$\Rightarrow$ highly degenerate ground state

$\Rightarrow$ highly entangled, non local ground state

$\Rightarrow$ support non-local excitation (non-local excitation can have only take finite amount of energy)

$\Rightarrow$ each spinon (unpaired spin) in the spinon pair excitation can be viewed as individual.

Summary:

- In $2D$, spinon excitations are simply unpaired spins. In ground state all spins are paired; in lowest excitation, there are two unapred spins, but they can be separated far away, due to the highly entangled nature of ground state.
- Partons are simply one way of finding highly entangled spin ground states.
- Quantum spin liquids can be defined as spin ground states that are highly entangled.

Definition of Spin liquids: highly entangled spin states, that support non-local excitations with fractionalized quantum numbers, non-abeial statistics, etc...

## 13.4 Types of spin liquids according to gauge theory

In the case of gapped spinons, we can integrate them out and get pure gauge theory.
Result: pure compact $U(1)$ gauge theory:

- 2D: confined (instanton effects)
- 3D: deconfined (coulomb phase) and confined

Result: there is no QSL in 2+1D with gapped spinon.
In the case of gapless spinons, in 2D, several cases:

- Algebraic SL: Dirac fermions: theory is $\text{QED}_3$

$$S = \int d\tau d^2x \left[ \sum_{i=1}^{N} \overline{\psi}_i \left( \partial_\tau - i a_0 - i(\vec{\nabla} - i\vec{a}) \cdot \vec{\sigma} \right) \psi_i + \frac{1}{4g} F^2 \right]$$

  For sufficiently large $N > N_c$: in a conformally invariant phase: algebraic spin liquid/Dirac SL. belieaved for small $N$ is unstable.

- Spinon surface SL: Fermi surface, more stable to monopoles
- Can use (Schwinger) boson to do parton construction, (boson at critical point).

## 13.5 Deconfined phase of compact $U(1)$ gauge theory: using $XY$ model

Now let us prove that there is no deconfined phase for 2D compact quantum $U(1)$ gauge theory, in condensed matter way.

There is a duality between the $U(1)$ gauge theory

$$Z_{U(1)} = \int \mathcal{D}a_\mu e^{i\int dtd^2x\mathcal{L}_{U(1)}},$$

$$\mathcal{L}_{U(1)} = \frac{1}{4g}F^2, \quad F_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu,$$

and the $XY$-model

$$Z_{XY} = \int \mathcal{D}\theta e^{i\int dtd^2x\mathcal{L}_{XY}},$$

$$\mathcal{L}_{XY} = \frac{\chi}{2}\left(\dot{\theta}^2 - (\vec{\nabla}\theta)^2\right).$$

These two theories have the same spectrum if

$$\chi = \left(\frac{qg}{2\pi}\right)^2,$$

and the same EOM if

$$\frac{q}{2\pi}b = \frac{1}{\chi}\dot{\theta}, \qquad \frac{q}{2\pi}\epsilon_{ij}e_j = \pm\frac{1}{\chi}\partial_i\theta,$$

where "+"("−") for real(imaginary) time, and $b = \partial_1 a_2 - \partial_2 a_1$ is the magnetic field.

A compact $U(1)$ gauge theory has quantized gauge charge. The $U(1)$ gauge theory can be coupled to charges

$$\mathcal{L} = \frac{1}{4g} - J_\mu a_\mu,$$

where $J_0 = q\delta(\vec{x})$ for a point charge at $\vec{x} = \vec{0}$, and from EOM we get the electric field

$$\vec{e} = \frac{g^2 q\vec{x}}{2\pi x^2},$$

thus using the duality we see that this $\vec{e}$ corresponds to a circular flow

$$\partial_i\theta = \frac{\epsilon_{ij}x^j}{x^2},$$

i,e, a quantized charge in the U(1) gauge theory is the quantized vortex in the $XY$-model.

- Unlike non-compact $U(1)$, in compact $U(1)$ gauge theory, there are quantized magnetic charge (monopoles). In $3+1D$, these charges are gapped excitations. In $2+1D$ however, they are not particles, but appear in the form of instantons (flux change in space-time), and flux not conserved.
- In imaginary time, a instanton of the $U(1)$ gauge theory at $x^\mu = 0$ is

  $$b(x^\mu) = \frac{1}{2q}\frac{x^0}{|x|^3}, \qquad e_i(x^\mu) = \frac{1}{2q}\frac{x^i}{|x|^3},$$

  which changes the folux by $2\pi/q$: $\left(\int_{x^0>0} - \int_{x^0<0}\right) d^2x b = 2\pi/q$,
- In presense of a finite cut-off scale (lattice gauge theory), the path integral should not only include the smooth fluctuations of the gauge field, but also instanton fluctuations. Then it is not pure gauge theory anymore.
- Use the duality, we can write

  $$\mathcal{L}_{U(1)} = \frac{\chi}{2}(\partial_\mu\theta)^2,$$

  the effect of instanton is to creats $2\pi/q$ amount of flux, and the the dual $XY$ model language, it creates/annihilates a single particle (duality: flux vs particle!). In $XY$-model, it is $e^{\pm i\theta}$ that creates/annihilates a single particle, so we should include them with same weight, so we get a term $2\cos\theta$, and now

  $$\mathcal{L} = \frac{\chi}{2}(\partial_\mu\theta)^2 - K\cos\theta,$$
- When $\chi$ is large, the fluctuations of $\theta$ around $\theta = 0$ is small, so we expand it at this point to get

  $$\mathcal{L} = \frac{\chi}{2}(\partial_\mu\theta)^2 + \frac{1}{2}K\theta^2,$$

  which is gappped. Thus instantons open up a gap.
- As a consequence, interaction grow linearly with distance – confined!

### 13.6 Mean field states – Symmetry considerations

- We mentioned before that QSLs typically do not break any symmetry of the lattice. Naturally, we want to know how is symmetry properties of the mean field Hamiltonian.
- It turns out, due to the gauge redundacy, the partons carry a "projective representation" of the lattice symmetry:

$$U_a(g)U_a(h) = \omega(g,h)U_a(gh),$$

  and consequently, the gauge equivalent Hamiltonian may transform into each other under symmetry.
- This defines an equivalence relation between all set of mean field Hamiltonians. We can use this relation to classify all possible mean field Hamiltonians that preserve lattice symmetry.
- This is called Projective Symmetry Group classification. It is a pure symmetry consideration.
- The nontrivial symmetry operations on partons is called symmetry fractionalization. It is associated to any statistics, etc.

To understand this more, let us give a simple example: four bonds $s_{12}, s_{23}, s_{34}, s_{41}$ on a square, taking value from $\{\pm 1\}$. Define gauge transformation as

$$s_{ij} \to W_i s_{ij} W_j^{-1},$$

where $W_i$ take values in $\{\pm 1\}$, and $i = 1, 2, 3, 4$ are the four sites.

Under this definition of gauge transformation, how many gauge inequivalent classes are there? (2)

Concept of IGG(invariant gauge group)

### 13.7 Reference

- *Quantum Field Theory of Many-body Systems*, (Chapt. 7&9), Xiao-Gang Wen
- *Quantum Spin Liquids*, arXiv 1601.03742, Lucile Savary, Leon Balents
- *Quantum Spin Liquid States*, arXiv 1607.03228, Yi Zhou, Kazushi Kanoda, Tai-Kai Ng

## 14 Two-band toy models for Chern insulators and Weyl semimetals

In this section, we shall use two band models to elucidate as much topological physics as possible. We will be careful whenever a statement is particular to the two-band situation.

### 14.1 Chern number

$H(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma}$, $\boldsymbol{d} = (d_1, d_2, d_3) = |\boldsymbol{d}|(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $E_\pm(\boldsymbol{k}) = \pm|\boldsymbol{d}| = \pm d$, $u_\pm = \frac{1}{\sqrt{2d(d\mp d_3)}}\begin{pmatrix} d_1 - id_2 \\ -d_3 \pm d \end{pmatrix} = \frac{1}{\sqrt{2d(d\mp d_3)}}\begin{pmatrix} d_1 - id_2 \\ \pm(d \mp d_3) \end{pmatrix} = \sqrt{\frac{d\pm d_3}{2d}}\begin{pmatrix} \frac{d_1 - id_2}{\sqrt{d_1^2+d_2^2}} \\ \pm\sqrt{\frac{d\mp d_3}{d\pm d_3}} \end{pmatrix} = \sqrt{\frac{1\pm\cos\theta}{2}}\begin{pmatrix} e^{-i\phi} \\ \pm\sqrt{\frac{1\mp\cos\theta}{1\pm\cos\theta}} \end{pmatrix} = \cos\frac{\theta}{2}\begin{pmatrix} e^{-i\phi} \\ \tan\frac{\theta}{2} \end{pmatrix}$ for $u_+$ and $\sin\frac{\theta}{2}\begin{pmatrix} e^{-i\phi} \\ -\cot\frac{\theta}{2} \end{pmatrix}$ for $u_-$, where we have used the fact that $\sin\frac{\theta}{2} = \sqrt{\frac{1-\cos\theta}{2}}$ and $\cos\frac{\theta}{2} = \sqrt{\frac{1+\cos\theta}{2}}$ for $\theta \in [0, \pi]$. So that

$$u_+ = \begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2} \end{pmatrix}, \qquad u_- = \begin{pmatrix} \sin\frac{\theta}{2}e^{-i\phi} \\ -\cos\frac{\theta}{2} \end{pmatrix}.$$

Another way to think about this is: we need to rotate the $\boldsymbol{d}$ vector to $\hat{z}$, so we need to rotate it along $\boldsymbol{n} = \frac{1}{d_1^2+d_2^2}(d_2, -d_1, 0) = (\sin\phi, -\cos\phi, 0)$ by an angle of $\theta$, where $\boldsymbol{n}$ is the vector that the projection of $\boldsymbol{d}$ onto the $xy$ plane, rotated clockwise by $\pi/2$ (so that $\boldsymbol{n} \perp (d_1, d_2, 0)$. By SU(2) rotation identitiy, define

$$U = e^{-i\frac{\theta}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}} = \cos\frac{\theta}{2} - i\boldsymbol{n}\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2}e^{-i\phi} \\ -\sin\frac{\theta}{2}e^{i\phi} & \cos\frac{\theta}{2} \end{pmatrix}, \qquad U^\dagger = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2}e^{i\phi} & \cos\frac{\theta}{2} \end{pmatrix},$$

where the minus sign is following the usual convention, then one can check that

$$U\boldsymbol{d}\cdot\boldsymbol{\sigma}U^\dagger = d\sigma^3.$$

So we have $U^\dagger = (e^{i\phi_+}u_+, e^{i\phi_-}u_-)$, where $\phi_\pm$ are the U(1) phase that one is free to choose. In our choice, we have $\phi_+ = \phi$ and $\phi_- = \pi$.

$\boldsymbol{A} = i\langle u_\pm|\boldsymbol{\nabla_k}|u_\pm\rangle = i(\cos\frac{\theta}{2}(-\frac{1}{2}\sin\frac{\theta}{2})\boldsymbol{\nabla_k}\theta + \cos^2\frac{\theta}{2}(-i)\boldsymbol{\nabla_k}\phi + \sin\frac{\theta}{2}\cdot\frac{1}{2}\cos\frac{\theta}{2}\boldsymbol{\nabla_k}\theta) = \cos^2\frac{\theta}{2}\boldsymbol{\nabla_k}\phi = \frac{1+\cos\theta}{2}\boldsymbol{\nabla_k}\phi$. note that this is subject to a gauge redundancy. $\oint \boldsymbol{A}\cdot d\boldsymbol{k} = \oint \frac{1+\cos\theta}{2}(\partial_{k_x}\phi, \partial_{k_y}\phi)\cdot(dk_x, dk_y) = \oint \frac{1+\cos\theta}{2}d\phi$. If $\cos\theta = d_3/d$ is constant on the path, we get $\oint \boldsymbol{A}\cdot d\boldsymbol{k} = \pi(1+\cos\theta) = \pi(1+d_3/d)$,

**Dirac Hamiltonian:** $\phi = \arctan k_y/k_x$, so $\boldsymbol{\nabla_k}\phi = \frac{1}{k_x^2+k_y^2}(-k_y, k_x)$.

## 14.2 Chern Insulator: QWZ model, analytical solution of edge states

Qi-Wu-Zhang Model: the bulk model in the basis $(c_{\boldsymbol{k},A}, c_{\boldsymbol{k},B})^T$ (where $A$ and $B$ are sublattice index) reads

$$H(\vec{k}) = (\mu - \cos k_x - \cos k_y)\sigma_3 + \sin k_x\sigma_1 + \sin k_y\sigma_2, \tag{37}$$

or

$$H(\vec{k}) = \begin{pmatrix} \mu - \cos k_x - \cos k_y & \sin k_x - i\sin k_y \\ \sin k_x + i\sin k_y & -\mu + \cos k_x + \cos k_y \end{pmatrix},$$

bulk energy

$$E = \pm\sqrt{\sin^2 k_x + \sin^2 k_y + (\mu - \cos k_x - \cos k_y)^2},$$

band gap closes at $(k_x, k_y) = (0, 0)$ when $\mu = 2$.

Put on a finite bar: $x$-direction is infinite still, $y$ direction is finite $N$ but $N$ very large. $k_y$ is no longer a good quantum nubmer and should be transformed back to real space: $\cos k_y \to \frac{1}{2}(\delta_{y,1} + \delta_{y,-1})$ and $i\sin k_y \to \frac{1}{2}(\delta_{y,1} - \delta_{y,-1})$, therefore the $(k_x, y)$-space Hamiltonian becomes

$$H(k_x) = \begin{pmatrix} A_1 & B & & & & \\ B^\dagger & A_2 & B & & & \\ & B^\dagger & A_3 & B & & \\ & & B^\dagger & A_4 & B & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{38}$$

where the empty elements are zero, and

$$A_i = \begin{pmatrix} \mu_i - \cos k_x & \sin k_x \\ \sin k_x & -\mu_i + \cos k_x \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Note that we have assigned spatial index to the mass: $\mu_i$ (and hence the $2\times 2$ matrix $A$), so it can be spatially dependent.

If there is low energy modes, it must be of the form

$$H(k_x)\begin{pmatrix} \mathbf{b} \\ \lambda_1\mathbf{b} \\ \lambda_1\lambda_2\mathbf{b} \\ \vdots \end{pmatrix} = E(k_x)\begin{pmatrix} \mathbf{b} \\ \lambda_1\mathbf{b} \\ \lambda_1\lambda_2\mathbf{b} \\ \vdots \end{pmatrix},$$

therefore we must have

$$(A_1 + B\lambda_1)\mathbf{b} = E\mathbf{b}, \quad (B^\dagger + A_i\lambda_{i-1} + B\lambda_{i-1}\lambda_i)\mathbf{b} = E\lambda_{i-1}\mathbf{b}, \quad i = 1, 2, ...,$$

therefore we must have

$$|A_1 - E + B\lambda_1| = |B^\dagger + \lambda_{i-1}(A_i - E + B\lambda_i)| = 0,$$

or

$$\begin{aligned} E^2 - \lambda_1\mu_1 + (\lambda_1 + 2\mu_1)\cos k_x - \mu_1^2 - 1 = 0, \\ \lambda_{i-1}\left[\lambda_{i-1}\left(E^2 - \mu_i^2 - 1\right) - \lambda_{i-1}\lambda_i\mu + (\lambda_{i-1}\lambda_i + 2\lambda_{i-1}\mu_i + 1)\cos k_x - \mu_i - \lambda_i\right] = 0, \end{aligned} \tag{39}$$

the solutions are

$$\lambda_i = 0 \text{ for all } i, \quad E = \pm\sqrt{1 + \mu_1^2 - 2\mu_1\cos k_x},$$

or

$$\lambda_i = \cos k_x - \mu_i, \quad E = \pm\sin k_x.$$

The first solution is specific to the model (it is an artifect of the special form of hopping chosen here). The second solution

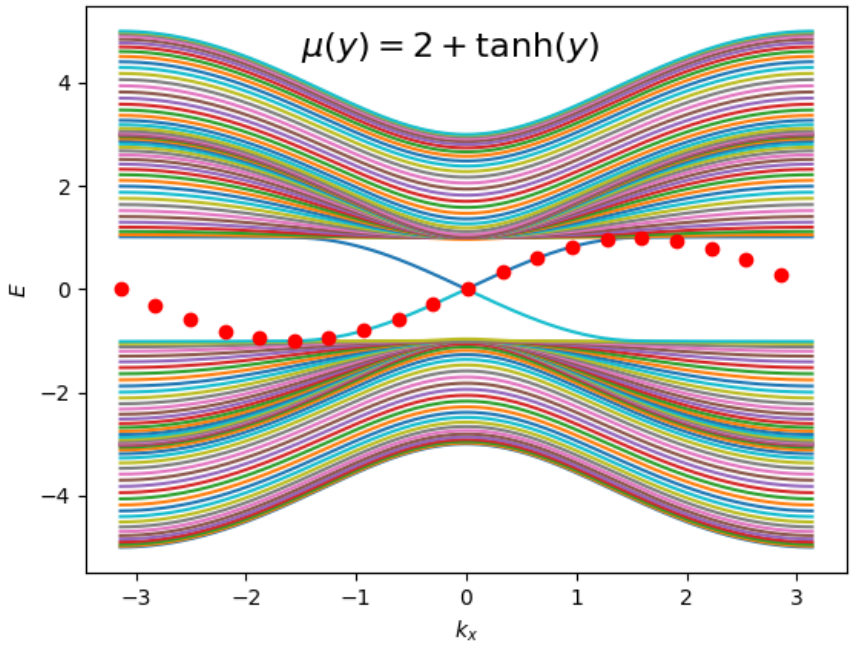



Figure 3: Energy dispersion of the chern insulator (37) with $\mu(y) = 2 + \tanh(y)$. Numerics is done using Eq. (45). Red dots are function $E = \sin(k_x)$, which fits very well with numerics.

We can compare these result with numerics; see Fig. 3. it turns out the correct dispursion is

$$e = \pm \sin k_x.$$

Notice that edge boundary corresponds to one edge state, and one dispersion. The two dispersion lines in the gap in Fig. 3 come from different boundaries.

This is known as the transfer matrix technique (or, really, a recursive sequence) to solve the edge states. The relevant papers are (Hatsugai, 1993) `https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.71.3697`, `https://journals.aps.org/prb/abstract/10.1103/PhysRevB.48.11851` which is a well known one.

Now we be a bit more careful about the convention when putting the QWZ model on a bar. Again assume $k_x$ is good quantum number so we have looking at

$$H = \sum_j c_j^\dagger A c_j + c_j^\dagger B c_{j+1} + c_{j+1}^\dagger B^\dagger c_j, \tag{40}$$

where $c_j^\dagger = (c_{jA}^\dagger, c_{jB}^\dagger)$. Using the usual Fourier convention, which is

$$c_k^\dagger = \frac{1}{\sqrt{L}} \sum_j e^{ikj} c_j^\dagger, \qquad c_k = \frac{1}{\sqrt{L}} \sum_j e^{-ikj} c_j, \qquad c_j^\dagger = \frac{1}{\sqrt{L}} \sum_k e^{-ikj} c_k^\dagger, \qquad c_j = \frac{1}{\sqrt{L}} e^{ikj} c_k, \tag{41}$$

we have

$$H = \sum_k c_{k_y}^\dagger A c_{k_y} + c_{k_y}^\dagger (B e^{ik_y} + B^\dagger e^{-ik_y}) c_{k_y}, \tag{42}$$

compare with

$$H(\vec{k}) = \begin{pmatrix} \mu - \cos k_x - \cos k_y & \sin k_x - i \sin k_y \\ \sin k_x + i \sin k_y & -\mu + \cos k_x + \cos k_y \end{pmatrix} = \begin{pmatrix} \mu - \cos k_x & \sin k_x \\ \sin k_x & -\mu + \cos k_x \end{pmatrix} + \begin{pmatrix} -\left(\frac{e^{ik_y} - e^{-ik_y}}{2}\right) & -i\left(\frac{e^{ik_y} - e^{-ik_y}}{2i}\right) \\ i\left(\frac{e^{ik_y} + e^{-ik_y}}{2i}\right) & \left(\frac{e^{ik_y} + e^{-ik_y}}{2}\right) \end{pmatrix},$$

we see that $A = \begin{pmatrix} \mu - \cos k_x & \sin k_x \\ \sin k_x & -\mu + \cos k_x \end{pmatrix}$, $B = \frac{1}{2} \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$.

Had we used $H(\vec{k}) = \begin{pmatrix} \cos k_x + \cos k_y - m & \sin k_x - i \sin k_y \\ \sin k_x + i \sin k_y & \mu - \cos k_x - \cos k_y \end{pmatrix}$ we then would have $A = \begin{pmatrix} \cos k_x - m & \sin k_x \\ \sin k_x & \mu - \cos k_x \end{pmatrix}$, and $B = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. Since this is what's in the python code now we adopt this convention. Putting on a finite slab and take, say, $k_x = 0.01$ and $m = 0.5$. The negative branch of the edge dispersion (i.e. $E = -\sin k_x$) is filled at $k_x = 0.01$, and by checking the eigenstate we find that it is localized on the small $j$ sites side. What this means is that the small $j$ site edges carries a dispersion $E \sim -k_x$ which has $v_F < 0$, meaning that it is flowing towards the $x < 0$ side. So, we see that there is an edge particle density current flowing in a clockwise fashion, which means that the electric charge current flows in the counterclockwise fashion. We know that for $m = 0.5$ the filled band has Chern number $C = +1$. Therefore, the right-hand rule here is that, if the filled band has a Chern number $C = +1$, then it will contribute to a charge current that follows according to the right-hand rule (i.e. counterclockwise). Futhermore, this should be consistent with a positive Hall conductivity $\sigma_H$: here we are assuming perfect antisymmetry so $\sigma_{xy} = -\sigma_{yx} = C = +1$ and $\sigma_H \equiv \frac{1}{2}(\sigma_{xy} - \sigma_{yx}) = \sigma_{xy} = C = +1$. According to Wikipedia, Hall coefficient $R_H = E_y/(j_x B_z)$: assuming a positive field $B_z$ along $+z$ direction and a positive field $E_x$ along $+x$ direction. If charge carriers are positive (negative): then classically positive (negative) charges will flow from the small $x$ (large $x$) side and accumulate on the small $y$ (also small $y$!!) edge, creating a $E_y$ that has direction $+y$ ($-y$), so $E_y > 0$ ($E_y < 0$). Note we have started with $B_z > 0$ and $E_y > 0$, so by definition $R_H > 0$ ($R_H < 0$) corresponds to positive (negative) charge carriers. We see that Hall coefficient is capable of distiguishing the signature of the charge carriers. Now the Hall resistance is defined as $R_{xy} = V_H/I_x$. Note that $I_x$ is positive since we already required $E_x > 0$, So $V_H$ determines the sign of $R_{xy}$ or $\sigma_{xy}$. It turns out that if the charge carrier is positive (negative) then the positive (negative) charges will accumulate on the small $y$ (also small $y$!!) edge so that the $E_y > 0$ ($E_y < 0$) so that $V_H = V(\text{large } y) - V(\text{small } y)$ is negative (positive), so that $\sigma_{xy}$ is negative (positive). Finally, we have seen that according to this usual convention, $\sigma_{xy} > 0$ means electronic carrier.

## 14.3 Weyl Semimetal: analytical solution

Model:

$$H(\vec{k}) = (\cos k_x + \cos k_y + \cos k_z - \cos k_0 - 2)\sigma_1 + \sin k_y \sigma_2 + \sin k_z \sigma_3, \tag{43}$$

or

$$H(\vec{k}) = \begin{pmatrix} \sin k_z & \cos k_x + \cos k_y + \cos k_z - \cos k_0 - 2 - i \sin k_y \\ \cos k_x + \cos k_y + \cos k_z - \cos k_0 - 2 + i \sin k_y & -\sin k_z \end{pmatrix},$$

energy

$$E = \pm\sqrt{(\cos k_x + \cos k_y + \cos k_z - \cos k_0 - 2)^2 + \sin k_y^2 + \sin k_z^2}. \tag{44}$$

band gap closes at $(k_x, k_y, k_z) = (\pm k_0, 0, 0)$.

Again put on a finite bar: $x$, $y$ directions are infinite but $z$ direction is finite $N$, but $N$ very large. We then must have $\cos k_z \to \frac{1}{2}(\delta_{z,1} + \delta_{z,-1})$ and $i \sin k_z \to \frac{1}{2}(\delta_{z,1} - \delta_{z,-1})$. The Hamiltonian again looks like

$$H_{k_0}(k_x, k_y) = \begin{pmatrix} A & B & & & & \\ B^\dagger & A & B & & & \\ & B^\dagger & A & B & & \\ & & B^\dagger & A & B & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{45}$$

where the empty elements are zero, and

$$A = \begin{pmatrix} 0 & \cos k_x + \cos k_y - \cos k_0 - 2 - i \sin k_y \\ \cos k_x + \cos k_y - \cos k_0 - 2 + i \sin k_y & 0 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix}$$

Following the previous line of reasoning for chern insulators, in order to have edge states we must have

$$|A + B\lambda| = |B^\dagger + A\lambda + B\lambda^2| = 0,$$

note we are only solving for zero energy modes (unlike the chern insulator case, where the bulk is gapped and any states with energy $< 1$ come from boundary). These gives

$$\begin{aligned} \cos k_0(\lambda + 2\cos k_x + 2\cos k_y - 4) - \cos^2 k_0 + \frac{1}{2}\left(4\lambda - 2\cos k_x(\lambda + 2\cos k_y - 4) + \sin^2 k_x - \cos^2 k_x - 2(\lambda - 4)\cos k_y - 11\right) = 0, \\ \frac{1}{2}\left(4\lambda^2 - 13\lambda + \lambda \sin^2 k_x - \lambda \cos^2 k_x - 2\cos k_x\left(\lambda^2 - 4\lambda + 2\lambda\cos k_y + 1\right) - 2\left(\lambda^2 - 4\lambda + 1\right)\cos k_y + 4\right) \\ -\lambda\cos^2 k_0 + \cos k_0\left(\lambda^2 - 4\lambda + 2\lambda\cos k_x + 2\lambda\cos k_y + 1\right) = 0, \end{aligned} \tag{46}$$

the funtion looks complicated. But we can presume that we know that there is no solution when $k_y \neq 0$. Therefore we set $k_y = 0$, and solve for $\lambda$, we get

$$\lambda = 1 + \cos k_0 - \cos k_y.$$

We also know that in order for it to be boundary states $|\lambda| < 1$, therefore we must have $\cos k_0 < \cos k_x$, therefore $-k_0 < k_x < k_0$. This is the known fermi arc.

A very good introductory paper for Weyl Semimetal is `https://www.annualreviews.org/doi/full/10.1146/annurev-conmatp` by Hassan.

The essence in understanding Fermi arc: for the example here, the Weyl nodes are located at $k_x = \pm k_0$. The fermi arc will appear when we attach boundary to the $z$ direction (then do the change $k_z \to z$). We must then *think $k_x$ as, not momentum, but some external parameter, at each of which we will have a 2D insulator.* The Weyl semimetal is then a superposition of these 2D insulators. Then the families of 2D insulators, parameterized by $k_x$, can be understood as going through phase transitions where gap closes and reopnens, and the middle phase is Chern phase.

# 15 Hopf Insulator – summary

## 15.1 remarks about two bands (i.e. the original Hopf insulator)

Define

$$\vec{N} = (N_1, N_2, N_3, N_4), \qquad \mathbf{z} = (N_1 + iN_2, N_3 + iN_4)^T, \tag{47a}$$

$$\boldsymbol{d} = \mathbf{z}^\dagger \boldsymbol{\sigma} \mathbf{z} = (d_1, d_2, d_3), \qquad H = \boldsymbol{d} \cdot \boldsymbol{\sigma}, \qquad \boldsymbol{n} = \frac{\boldsymbol{d}}{|\boldsymbol{d}|}, \tag{47b}$$

$$(d_1, d_2, d_3) = (2(N_1N_3 + N_2N_4), 2(N_1N_4 - N_2N_3), N_1^2 + N_2^2 - N_3^2 - N_4^2). \tag{47c}$$

For such a two band Hamiltonian, we write quite generally $H = \boldsymbol{d} \cdot \boldsymbol{\sigma} = d\, U\sigma^z U^\dagger$. Denoted $d = |\boldsymbol{d}|$, we have $d = N^2$.

For the unitary matrix $U$, We know the eigenvectors have a phase ambiguity, and as a consequence, there are many different forms of the unitary matrix $U$. Write $\boldsymbol{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, Then we know the general form of $U$ is

$$U = e^{-i\frac{\alpha}{2}\boldsymbol{p}\cdot\boldsymbol{\sigma}}, \qquad \text{s.t.} \qquad U^\dagger \boldsymbol{n}\cdot\boldsymbol{\sigma} U = e^{i\frac{\alpha}{2}\boldsymbol{p}\cdot\boldsymbol{\sigma}} \boldsymbol{n}\cdot\boldsymbol{\sigma} e^{-i\frac{\alpha}{2}\boldsymbol{p}\cdot\boldsymbol{\sigma}} = \sigma^z.$$

There are many choice of $\boldsymbol{p}$ and $\alpha$, corresponding to the phase ambiguity of eigenstates $U \to U\Theta$, where $\Theta = \mathrm{diag}(e^{i\theta_1}, e^{i\theta_2})$. But in terms of simplicity, there is perhaps a preferred choice, which is $\boldsymbol{p} = \boldsymbol{n}_\perp \equiv \frac{1}{\sqrt{n_1^2+n_2^2}}(-n_2, n_1, 0) = (-\sin\phi, \cos\phi) \perp (n_1, n_2, 0)$, which gives $\alpha = \frac{\theta}{2}$. Then we have

$$\begin{aligned}
U = e^{-i\frac{\theta}{2}\boldsymbol{n}_\perp\cdot\sigma} &= \cos\frac{\theta}{2} - i\boldsymbol{n}_\perp\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2}e^{i\phi} & \cos\frac{\theta}{2} \end{pmatrix} \\
&= \frac{1}{2\cos\frac{\theta}{2}} \begin{pmatrix} 1+\cos\theta & -\sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & 1+\cos\theta \end{pmatrix} \\
&= \frac{1}{\sqrt{(d_3+d)^2 + d_1^2 + d_2^2}} \begin{pmatrix} d+d_3 & -d_1+id_2 \\ d_1+id_2 & d+d_3 \end{pmatrix} \\
&= \frac{1}{\sqrt{2d(d+d_3)}}((d+d_3) + id_2\sigma^1 - id_1\sigma^2) \\
&= \frac{1}{|\vec{N}|} \begin{pmatrix} \sqrt{N_1^2+N_2^2} & -\frac{N_1+iN_2}{\sqrt{N_1^2+N_2^2}}(N_3 - iN_4) \\ \frac{N_1-iN_2}{\sqrt{N_1^2+N_2^2}}(N_3+iN_4) & \sqrt{N_1^2+N_2^2} \end{pmatrix},
\end{aligned} \tag{48}$$

Choosing $e^{i\theta_1} = \frac{N_1+iN_2}{\sqrt{N_1^2+N_2^2}}$ and $e^{i\theta_2} = \frac{N_1-iN_2}{\sqrt{N_1^2+N_2^2}}$, we have the new $U$

$$U \to U\Theta = \frac{1}{|\vec{N}|} \begin{pmatrix} N_1+iN_2 & -N_3+iN_4 \\ N_3+iN_4 & N_1-iN_2 \end{pmatrix} = \frac{1}{|\vec{N}|}(N_1 + iN_4\sigma^1 - iN_3\sigma^2 + iN_2\sigma^3). \tag{49}$$

We now seek a singularity-free expression of $U$ purely in terms of $\boldsymbol{d}$. First of all, we write

$$U = \frac{1}{d}\sqrt{\frac{d}{2(d+d_3)}} \begin{pmatrix} d+d_3 & -d_1+id_2 \\ d_1+id_2 & d+d_3 \end{pmatrix} = \frac{1}{d} \begin{pmatrix} \sqrt{\frac{d(d+d_3)}{2}} & \sqrt{\frac{d(d-d_3)}{2}}\frac{-d_1+id_2}{\sqrt{d_1^2+d_2^2}} \\ \sqrt{\frac{d(d-d_3)}{2}}\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}} & \sqrt{\frac{d(d+d_3)}{2}} \end{pmatrix},$$

where $d_1+id_2=\mathbf{z}^\dagger(\sigma^1+i\sigma^2)\mathbf{z}=(N_1-iN_2)(N_3+iN_4)$,

For a generic $U$ of the above form $U=\sigma^0u^0+i\sum_{i=1,2,3}\sigma^iu^i$, we look at the central quantity

$$I=\epsilon^{\mu\nu\rho}\mathrm{Tr}[U^\dagger\partial_\mu U\ U^\dagger\partial_\nu U\ U^\dagger\partial_\rho U]$$

each term is the product of six Pauli matrices including the identity $\sigma^0$. To have nonzero trace, the product must equal identity. Then note that the index of Pauli matrices are the same as that of the component index of $\vec{N}$; $\epsilon^{\mu\nu\rho}$ requires the latter index to be all distinct so the only nonzero terms must come from distinct Pauli matrices in $\partial_\mu U$, $\partial_\nu U$ and $\partial_\rho U$. Then, in order to have nonzero trace, the three $U^\dagger$'s must contribute the same three Pauli matrices, and this means that the only nonzero terms are of the form $\mathrm{Tr}[u^\alpha\sigma^\alpha\partial_\mu u^{\alpha'}\sigma^{\alpha'}u^\beta\sigma^\beta\partial_\nu u^{\beta'}\sigma^{\beta'}u^\gamma\sigma^\gamma\partial_\rho u^{\gamma'}\sigma^{\gamma'}]$ where $\{\alpha,\beta,\gamma\}=\{\alpha',\beta',\gamma'\}$, and this term becomes $u^\alpha u^\beta u^\gamma\partial_\mu u^\alpha\partial_\nu u^\beta\partial_\rho u^\gamma$. It turns out (easily checked in Mathematica) when all these terms are summed (with proper signs), there is a common factor $\sum_\alpha(u^\alpha)^2$ that can be factored out, which of course equals identity. The final result is (checked by Mathematica)

$$I=\epsilon^{\mu\nu\rho}\mathrm{Tr}[U^\dagger\partial_\mu U\ U^\dagger\partial_\nu U\ U^\dagger\partial_\rho U]=2\epsilon^{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho}u^\alpha\partial_\mu u^\beta\partial_\nu u^\gamma\partial_\rho u^\delta,$$

Note there are in total $2\times3!\times4!=288$ terms, and only 24 different types (so every 12 terms are the same). Substituting $U\to U\Theta$ in the above expression we get the result $2\epsilon^{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho}N^\alpha\partial_\mu N^\beta\partial_\nu N^\gamma\partial_\rho N^\delta$. Normalization: We have

$$\begin{aligned}\frac{1}{24\pi^2}\int I dxdydz&=\frac{1}{24\pi^2}\int\epsilon^{\mu\nu\rho}\mathrm{Tr}[U^\dagger\partial_\mu U\ U^\dagger\partial_\nu U\ U^\dagger\partial_\rho U]dxdydz\\&=\frac{2}{24\pi^2}\int\epsilon^{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho}u^\alpha\partial_\mu u^\beta\partial_\nu u^\gamma\partial_\rho u^\delta dxdydz\\&=\frac{12}{24\pi^2}\int\epsilon^{\alpha\beta\gamma\delta}u^\alpha\partial_x u^\beta\partial_y u^\gamma\partial_z u^\delta dxdydz\\&=\frac{1}{2\pi^2}\int\epsilon^{\alpha\beta\gamma\delta}u^\alpha\partial_x u^\beta\partial_y u^\gamma\partial_z u^\delta dxdydz\in\mathbb{Z},\end{aligned}\tag{50}$$

note that $S_3=2\pi^2$ is the area of the 3-sphere.

The above expression also suggests that if we use the naive version of $U$ defined in Eq. (48), then the above expression vanishes. In order to see this in a more clear manner, let's denote $\widetilde{U}=U\Theta$, and calculate $\widetilde{I}=\epsilon^{\mu\nu\rho}\mathrm{Tr}[\widetilde{U}^\dagger\partial_\mu\widetilde{U}\ \widetilde{U}^\dagger\partial_\nu\widetilde{U}\ \widetilde{U}^\dagger\partial_\rho\widetilde{U}]$. Since $\widetilde{U}^\dagger\partial\widetilde{U}=\Theta^\dagger U^\dagger[(\partial U)\Theta+U\partial\Theta]=\Theta^\dagger U^\dagger\partial_U\Theta+\Theta^\dagger\partial\Theta=\Theta^\dagger(U^\dagger\partial U+i\partial\theta)\Theta$, where we defined $\theta=\mathrm{diag}(\theta_1,\theta_2)$. Then we have $\widetilde{I}=\epsilon^{\mu\nu\rho}\mathrm{Tr}[(U^\dagger\partial_\mu U+i\partial_\mu\theta)(U^\dagger\partial_\nu U+i\partial_\nu\theta)(U^\dagger\partial_\rho U+i\partial_\rho\theta)]=I+3i\epsilon^{\mu\nu\rho}\mathrm{Tr}[U^\dagger\partial_\mu U\ U^\dagger\partial_\nu U\ \partial_\rho\theta]$, the last term vanishes after integrating by parts over the Brillouin zone. So there is an apparent contradiction: $\int I$ is gauge invariant, but why using $U$ in Eq. (48) and $U\Theta$ in Eq. (49) give different results? And which one is the correct one?

The problem is that the naive $U$ as in Eq. (48) has singularity: when $N_1=N_2=0$, $U$ is not well defined, while $\widetilde{U}=U\Theta$ is always well defined. When using the formula, we should always make sure $U$ is well defined.

Generally, when the first Chern number vanishes, there is always a way (at least in theory) to find a well defined $U$ (for any band number $n$). The way out is to use the Hamiltonian: denote $U=(\boldsymbol{u}_1,...,\boldsymbol{u}_n)$ so that $\boldsymbol{u}_j^\dagger H\boldsymbol{u}_i=\lambda_i\delta_{ij}$, then $(\partial\boldsymbol{u}_j^\dagger)\lambda_i\boldsymbol{u}_i+\boldsymbol{u}_j^\dagger(\partial H)\boldsymbol{u}_i+\boldsymbol{u}_j^\dagger\lambda_j\partial\boldsymbol{u}_i=\partial\lambda_i\delta_{ij}$, and using $(\partial\boldsymbol{u}_j^\dagger)\boldsymbol{u}_i=-\boldsymbol{u}_j^\dagger\partial\boldsymbol{u}_i$ we have

$$\boldsymbol{u}_j^\dagger(\partial H)\boldsymbol{u}_i+\boldsymbol{u}_j^\dagger(\lambda_j-\lambda_i)\partial\boldsymbol{u}_i=\partial\lambda_i\delta_{ij},$$

therefore

$$\boldsymbol{u}_j^\dagger\partial\boldsymbol{u}_i=\frac{\boldsymbol{u}_j^\dagger(\partial H)\boldsymbol{u}_i}{\lambda_i-\lambda_j},\qquad j\neq i.$$

Therefore

$$\partial\boldsymbol{u}_i=\sum_j\boldsymbol{u}_j\boldsymbol{u}_j^\dagger\partial\boldsymbol{u}_i=\boldsymbol{u}_i\boldsymbol{u}_i^\dagger\partial\boldsymbol{u}_i+\sum_{j\neq i}\boldsymbol{u}_j\boldsymbol{u}_j^\dagger\partial\boldsymbol{u}_i=\boldsymbol{u}_i\boldsymbol{u}_i^\dagger\partial\boldsymbol{u}_i+\sum_{j\neq i}\boldsymbol{u}_j\frac{\boldsymbol{u}_j^\dagger(\partial H)\boldsymbol{u}_i}{\lambda_i-\lambda_j},$$

$$[U^\dagger\partial U]_{ji}=\boldsymbol{u}_j^\dagger\partial\boldsymbol{u}_i=\delta_{ji}\boldsymbol{u}_i^\dagger\partial\boldsymbol{u}_i+\delta_{j\neq i}\frac{\boldsymbol{u}_j^\dagger(\partial H)\boldsymbol{u}_i}{\lambda_i-\lambda_j},$$

Therefore

$$U^\dagger\partial_\mu U=-iA^\mu+B^\mu,$$

where $A$ is the diagonal matrix consisting of the Berry connection of each band, $A_{ij}=i\delta_{ij}\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i$, and $B$ is the matrix with vanishing diagonal entries and nondiagonal elements $B_{ij}^\mu=\frac{\boldsymbol{u}_i^\dagger(\partial_\mu H)\boldsymbol{u}_j}{\lambda_j-\lambda_i}$, so that

$$(B^\mu)^\dagger=-B^\mu.$$

Then,

$$I = \epsilon^{\mu\nu\rho}\mathrm{Tr}[(-iA^\mu + B^\mu)(-iA^\nu + B^\nu)(-iA^\rho + B^\rho)],$$

the $A^3$ and $A^2$ terms are symmetric with respect to $\mu\nu\rho$ so they vanish. So we are left with

$$I = \epsilon^{\mu\nu\rho}\mathrm{Tr}[B^\mu B^\nu B^\rho] - 3i\epsilon^{\mu\nu\rho}\mathrm{Tr}[A^\mu B^\nu B^\rho] = \epsilon^{\mu\nu\rho}\mathrm{Tr}[B^\mu B^\nu B^\rho] - \frac{3i}{2}\epsilon^{\mu\nu\rho}\mathrm{Tr}[A^\mu[B^\nu, B^\rho]]. \tag{51}$$

We tested in Mathematica to verify that the $B^\mu B^\nu B^\rho$ term does not vanish for $n > 2$. Since $B$ involves $\boldsymbol{u}_j^\dagger$ and $\boldsymbol{u}_i$ with $i \neq j$, its form still changes with the phase ambiguity. However, if we write out explicitly

$$\epsilon^{\mu\nu\rho}B^\mu_{ij}B^\nu_{jk}B^\nu_{ki} = \epsilon^{\mu\nu\rho}\frac{\boldsymbol{u}_i^\dagger(\partial_\mu H)\boldsymbol{u}_j\boldsymbol{u}_j^\dagger(\partial_\nu H)\boldsymbol{u}_k\boldsymbol{u}_k^\dagger(\partial_\rho H)\boldsymbol{u}_i}{(\lambda_j-\lambda_i)(\lambda_k-\lambda_j)(\lambda_i-\lambda_k)},$$

note that the sum over $i, j, k$ must exclude any terms with $i = j$ and so on, we see that the expression is actually gauge invariant.

For $n = 2$, we cannnot satisfy $i \neq j, j \neq k, k \neq i$ so the $BBB$ term is zero.

We now look at the $ABB$ term. This term is

$$\epsilon^{\mu\nu\rho}\sum_i\sum_{j\neq i}A^\mu_{ii}B^\nu_{ij}B^\rho_{ji} = \epsilon^{\mu\nu\rho}\sum_i\sum_{j\neq i}\frac{i\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i\boldsymbol{u}_i^\dagger(\partial_\nu H)\boldsymbol{u}_j\boldsymbol{u}_j^\dagger(\partial_\rho H)\boldsymbol{u}_i}{-(\lambda_j-\lambda_i)^2},$$

Not that due to the part $\partial_\mu\boldsymbol{u}_i$ (i.e. $-i$ times the Berry connection) we must use the singularity-free $\boldsymbol{u}_i$ that comes from $U\Theta$, rather than $U$. This will then give the correct Hopf number. We clarify again that the expression is gauge invariant, in the sense that all the singularity-free $\boldsymbol{u}_i$ will give the same Hopf number.

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu = \partial_\mu(i\boldsymbol{u}_i^\dagger\partial_\nu\boldsymbol{u}_i) - \partial_\nu(i\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i) = i\partial_\mu\boldsymbol{u}_i^\dagger\partial_\nu\boldsymbol{u}_i - i\partial_\nu\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i = 2\mathrm{Im}(\partial_\nu\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i),$$

$$\partial_\mu\boldsymbol{u}_i^\dagger\partial_\nu\boldsymbol{u}_i = \partial_\mu\boldsymbol{u}_i^\dagger\boldsymbol{u}_i\boldsymbol{u}_i^\dagger\partial_\nu\boldsymbol{u}_i + \sum_{j\neq i}\frac{\boldsymbol{u}_i^\dagger(\partial_\mu H)\boldsymbol{u}_j\boldsymbol{u}_j^\dagger(\partial_\nu H)\boldsymbol{u}_i}{-(\lambda_i-\lambda_j)^2},$$

note the first term on the right is symemtric with respect to exchanging $\mu$ and $\nu$, therefore

$$\epsilon^{\mu\nu\rho}F^i_{\nu\rho} = i\epsilon^{\mu\nu\rho}\sum_{j\neq i}\frac{\boldsymbol{u}_i^\dagger(\partial_\nu H)\boldsymbol{u}_j\boldsymbol{u}_j^\dagger(\partial_\rho H)\boldsymbol{u}_i}{-(\lambda_i-\lambda_j)^2},$$

For two band system, using

$$\boldsymbol{u}_1\boldsymbol{u}_1^\dagger = \frac{1}{d}\begin{pmatrix} \frac{d+d_3}{2} & \frac{d_1-id_2}{2} \\ \frac{d_1+id_2}{2} & \frac{d-d_3}{2}\end{pmatrix}, \qquad \boldsymbol{u}_2\boldsymbol{u}_2^\dagger = \frac{1}{d}\begin{pmatrix} \frac{d-d_3}{2} & \frac{-d_1+id_2}{2} \\ \frac{-d_1-id_2}{2} & \frac{d+d_3}{2}\end{pmatrix},$$

Then

$$\begin{aligned}\epsilon^{\mu\nu\rho}F^1_{\nu\rho} &= i\frac{\epsilon^{\mu\nu\rho}}{-4d^4}\mathrm{Tr}\left[\begin{pmatrix} \partial_\nu d_3 & \partial_\nu(d_1-id_2) \\ \partial_\nu(d_1+id_2) & -\partial_\nu d_3\end{pmatrix}\begin{pmatrix} \frac{d-d_3}{2} & \frac{-d_1+id_2}{2} \\ \frac{-d_1-id_2}{2} & \frac{d+d_3}{2}\end{pmatrix}\cdot\right. \\ &\qquad\left.\begin{pmatrix} \partial_\rho d_3 & \partial_\rho(d_1-id_2) \\ \partial_\rho(d_1+id_2) & -\partial_\rho d_3\end{pmatrix}\begin{pmatrix} \frac{d+d_3}{2} & \frac{d_1-id_2}{2} \\ \frac{d_1+id_2}{2} & \frac{d-d_3}{2}\end{pmatrix}\right] \\ &= \epsilon^{\mu\nu\rho}\frac{i}{4d^4}\ 2id\ (d_1\partial_\nu d_2\partial_\rho d_3 - d_1\partial_\rho d_2\partial_\nu d_3 + \text{the rest four terms}) \\ &= -\frac{1}{2d^3}\epsilon^{\mu\nu\rho}\epsilon^{\alpha\beta\gamma}d_\alpha\partial_\nu d_\beta\partial_\rho d_\gamma,\end{aligned} \tag{52}$$

The original Hopf insulator paper uses the gauge $\partial_\mu A_\mu = 0$. Let's check: if $\boldsymbol{u}_i = \frac{1}{|\vec{N}|}(N_1 + iN_2, N_3 + iN_4)^T$, $\partial_\mu\boldsymbol{u}_i = -\frac{N_\alpha\partial_\mu N_\alpha}{|\vec{N}|^3}(N_1 + iN_2, N_3 + iN_4)^T + \frac{1}{|\vec{N}|}(\partial_\mu N_1 + i\partial_\mu N_2, \partial_\mu N_3 + i\partial_\mu N_4)^T$, So

$$\begin{aligned}\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i &= -\frac{N_\alpha\partial_\mu N_\alpha}{|\vec{N}|^2} + \frac{(N_1-iN_2)(\partial_\mu N_1 + i\partial_\mu N_2) + (N_3-iN_4)(\partial_\mu N_3 + i\partial_\mu N_4)}{|\vec{N}|^2} \\ &= \frac{i(N_1\partial_\mu N_2 - N_2\partial_\mu N_1) + i(N_3\partial_\mu N_4 - N_4\partial_\mu N_3)}{|\vec{N}|^2},\end{aligned} \tag{53}$$

since

$$\partial_\mu(\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i)=-2N_\alpha\partial_\mu N_\alpha\frac{i(N_1\partial_\mu N_2-N_2\partial_\mu N_1)+i(N_3\partial_\mu N_4-N_4\partial_\mu N_3)}{|\vec{N}|^4}+\frac{i(N_1\partial_\mu^2 N_2-N_2\partial_\mu^2 N_1)+i(N_3\partial_\mu^2 N_4-N_4\partial_\mu^2 N_3)}{|\vec{N}|^2},$$

under $\boldsymbol{u}_i\to\boldsymbol{u}_ie^{i\phi}$ we have $A_\mu^i=i\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i-\partial_\mu\phi$, $\partial_\mu A_\mu^i=i\partial_\mu\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i+i\boldsymbol{u}_i^\dagger\partial_\mu^2\boldsymbol{u}_i-\partial_\mu^2\phi$.

$$\begin{aligned}F_{\mu\nu}=-\frac{1}{|\vec{N}|^2}&\left[(\partial_\mu N_1\partial_\nu N_2-\partial_\nu N_1\partial_\mu N_2)(-d+d_3)+(\partial_\mu N_1\partial_\nu N_4-\partial_\nu N_1\partial_\mu N_4-\partial_\mu N_2\partial_\nu N_3+\partial_\nu N_2\partial_\mu N_3)d_1\right.\\&\left.+(\partial_\mu N_1\partial_\nu N_3-\partial_\nu N_1\partial_\mu N_3+\partial_\mu N_2\partial_\nu N_4-\partial_\nu N_2\partial_\mu N_4)(-d_2)+(\partial_\mu N_3\partial_\nu N_4-\partial_\nu N_3\partial_\mu N_4)(-d-d_3)\right],\end{aligned}\tag{54}$$

$$\int\epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho}dk_xdk_ydk_z=\int\epsilon^{\mu\nu\rho}i\partial_\mu\phi F_{\nu\rho}dk_xdk_ydk_z=2\int(\phi F_{xy})\Big|_{k_z=-\pi}^{k_z=\pi}dk_xdk_y+2\int(\phi F_{yz})\Big|_{k_x=-\pi}^{k_x=\pi}dk_ydk_z+2\int(\phi F_{zx})\Big|_{-k_y=-\pi}^{k_y=\pi}dxdz$$

set $e^{i\phi}=\frac{\sin k_x+i\sin k_y}{\sqrt{\sin^2 k_x+\sin^2 k_y}}$,

$$\int\epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho}dk_xdk_ydk_z=2\int(\phi(k_x=\pi)-\phi(k_x=-\pi))F_{yz}dk_ydk_z+2\int(\phi(k_y=\pi)-\phi(k_y=-\pi))F_{xz}dk_xdk_z,$$

it turns out $\phi(k_x=\pi)-\phi(k_x=-\pi)=\pi\delta_{k_y=0}-\pi\delta_{k_y=\pm\pi}$.

We see from above where the problem lies: the Berry curvature $F$ is invariant under the gauge transformation, while the Berry connection is. Under the gauge transformation $\boldsymbol{u}_1\to\boldsymbol{u}_1e^{i\phi}$, we get an extra term $\epsilon^{\mu\nu\rho}\int\partial_\mu\phi F_{\nu\rho}d^3k$. The question is whether this term is nonzero. If it is always zero for any $\phi$, this means the Abelian Chern-Simons is gauge invariant; otherwise it is not gauge invariant. From our analysis above, we think it is not gauge invariant, because $\int\partial_\mu\phi F_{\nu\rho}=\int\partial_\mu(\phi F_{\nu\rho})=\int(\phi F_{\nu\rho})|_{k_\mu=-\pi}^{k_\mu=\pi}$ (note that $\epsilon^{\mu\nu\rho}\partial_\mu F_{\nu\rho}=0$). But Vanderbilt's book says it is gauge invariant. So where is this term is his derivation? It is the first term of his Eq. (C.15): his $W_\mu$ is just $i\partial_\mu\phi$, and he used integrating by parts the other way: $\int\epsilon^{\mu\nu\rho}\partial_\mu\phi F_{\nu\rho}=\int\epsilon^{\mu\nu\rho}\partial_\mu\phi\partial_\nu A_\rho=\int\epsilon^{\mu\nu\rho}\partial_\nu((\partial_\mu\phi)A_\rho)$, (note that the $\epsilon^{\mu\nu\rho}(\partial_\nu\partial_\mu\phi)A_\rho$ term vanishes). Note there is a very important footnote on Page 322 of his book: "In eliminating the surface term in the integration by parts, we have implicitly assumed a smooth gauge,such that $A_\mu$ is smooth and periodic in the 3D Brillouin zone." So we see that, indeed, we must use a smooth gauge to start with!

Question: what is a smooth gauge? a smooth gauge is is vector field $\boldsymbol{A}$ such that 1. always have finite strength, $|\boldsymbol{A}|<\infty$, and 2. when a vortex appears the center of the vortex has vanishing amplitude $|\boldsymbol{A}|=0$.

So what can cause singularity to happen? One example: the derivative of $e^{i\theta_{\boldsymbol{k}}}$ is singular. This then means that the part $\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}}$ is singular since $d_1=\sin k_x\sim k_x$ and $d_2=\sin k_y\sim k_y$.

Now restricting to two band model, let's see if there is a surgery that we can perform to get the correct result. Remember that when using the form $\boldsymbol{u}_1=\frac{1}{\sqrt{2d}}(\sqrt{d+d_3},\sqrt{d-d_3}\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}})^T$, the integrand of the Hopf integral actually vanishes (which is true away from the singularities), so the only contribution must come from the singularities. By inspection, the singularity is only at $d_1=d_2=0$, which are the poles. So there must be a way to extract the Hopf number only at these poles.

$\boldsymbol{u}_1=\frac{1}{\sqrt{2d}}(\sqrt{d+d_3},\frac{d_1+id_2}{\sqrt{d+d_3}})^T$, $\partial_\mu\sqrt{d+d_3}=\frac{\partial_\mu d+\partial_\mu d_3}{2\sqrt{d+d_3}}$, $\partial_\mu(\frac{1}{\sqrt{d+d_3}})=-\frac{\partial_\mu d+\partial_\mu d_3}{2(d+d_3)^{3/2}}$,

$$\partial_\mu\boldsymbol{u}_1=-\frac{\partial_\mu d}{2d}\boldsymbol{u}_1+\frac{1}{\sqrt{2d}}\left(\frac{\partial_\mu(d+d_3)}{2\sqrt{d+d_3}},-(d_1+id_2)\frac{\partial_\mu d+\partial_\mu d_3}{2(d+d_3)^{3/2}}+\frac{\partial_\mu(d_1+id_2)}{\sqrt{d+d_3}}\right)^T,$$

Now we only care about the possibly singular part, so we write

$$\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1=-\frac{\partial_\mu d}{2d}+\frac{1}{4d}\left(\partial_\mu(d+d_3)-\frac{(d_1^2+d_2^2)\partial_\mu(d+d_3)}{(d+d_3)^2}+\frac{2(d_1-id_2)\partial_\mu(d_1+id_2)}{d+d_3}\right),$$

Now, note that all the derivatives are nonsingular, and the singularity only comes from zero denominators. So we further simply

$$
\begin{aligned}
\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1 &= -\frac{\partial_\mu d}{2d}+\frac{1}{4d}\partial_\mu(d+d_3)+\frac{1}{4d}\left(-\frac{(d_1^2+d_2^2)\partial_\mu(d+d_3)}{(d+d_3)^2}+\frac{2(d_1-id_2)\partial_\mu(d_1+id_2)}{d+d_3}\right)\\
&=\frac{1}{4d}\partial_\mu(-d+d_3)+\frac{1}{4d}\frac{-(d-d_3)\partial_\mu(d+d_3)+2(d_1-id_2)\partial_\mu(d_1+id_2)}{d+d_3}\\
&=\frac{1}{4d}\partial_\mu(-d+d_3)+\frac{1}{4d}\frac{-d\partial_\mu d+d_3\partial_\mu d_3+(d_3\partial_\mu d-d\partial_\mu d_3)+2d\partial_\mu d-2d_3\partial_\mu d_3+2i(d_1\partial_\mu d_2-d_2\partial_\mu d_1)}{d+d_3}\\
&=\frac{1}{4d}\partial_\mu(-d+d_3)+\frac{1}{4d}\frac{d\partial_\mu d-d_3\partial_\mu d_3+(d_3\partial_\mu d-d\partial_\mu d_3)+2i(d_1\partial_\mu d_2-d_2\partial_\mu d_1)}{d+d_3}\\
&=\frac{1}{4d}\partial_\mu(-d+d_3)+\frac{1}{4d}\partial_\mu(d-d_3)+\frac{i}{2d}\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{d+d_3}\\
&=\frac{i}{2d}\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{d+d_3},
\end{aligned}
\tag{55}
$$

Note that $d_1\partial_\mu d_2-d_2\partial_\mu d_1=4(N_3^2+N_4^2)(N_2\partial_\mu N_1-N_1\partial_\mu N_2)+4(N_1^2+N_2^2)(N_3\partial_\mu N_4-N_4\partial_\mu N_3)$, therefore

$$
\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1=\underbrace{\frac{i}{N_1^2+N_2^2}(N_2\partial_\mu N_1-N_1\partial_\mu N_2)}_{-iA_\mu^{\rm red}}+\underbrace{\frac{i}{|\vec{N}|^2}(N_1\partial_\mu N_2-N_2\partial_\mu N_1+N_3\partial_\mu N_4-N_4\partial_\mu N_3)}_{-iA_\mu^{\rm good}},
$$

compare with the expression under the good gauge, Eq. (53), we see that now the bad gauge contributed one more term which is $\frac{i}{N_1^2+N_2^2}(N_2\partial_\mu N_1-N_1\partial_\mu N_2)$. This term does not enter Berry curvature, as has been verified in Mathematica.

We verified that

$$
\epsilon^{\mu\nu\rho}A_\mu^{\rm red}F_{\nu\rho}+\epsilon^{\mu\nu\rho}A_\mu^{\rm good}F_{\nu\rho}=0,\qquad \partial_\nu A_\rho^{\rm red}-\partial_\rho A_\nu^{\rm red}=0,\qquad F_{\nu\rho}=\partial_\nu A_\rho^{\rm good}-\partial_\rho A_\nu^{\rm good},
$$
$$
\epsilon^{\mu\nu\rho}A_\mu^{\rm good}F_{\nu\rho}\sim\epsilon^{\alpha\beta\gamma\delta}\epsilon^{\mu\nu\rho}N_\alpha\partial_\mu N_\beta\partial_\nu N_\gamma\partial_\rho N_\delta.
$$

Since $\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i=-\partial_\mu\boldsymbol{u}_i^\dagger\ \boldsymbol{u}_i=-(\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i)^\dagger$, $\boldsymbol{u}_i^\dagger\partial_\mu\boldsymbol{u}_i$ is purely imaginary, which is a side check of the above expression. The above expression is equivalent to regular terms $+\,i\frac{d-d_3}{2d}\partial_\mu\left(\arctan\frac{d_2}{d_1}\right)$ when except for the point $d=-d_3$. We note that in our gauge, the singualrity can only happen at $d=-d_3$, not $d=d_3$. So,

$$
A_\mu^1=-\frac{1}{2d}\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{d+d_3}.
$$

using Eq. (52) which is

$$
\epsilon^{\mu\nu\rho}F_{\nu\rho}^1=-\frac{1}{2d^3}\epsilon^{\mu\nu\rho}\epsilon^{\alpha\beta\gamma}d_\alpha\partial_\nu d_\beta\partial_\rho d_\gamma
$$

we get

$$
\begin{aligned}
\epsilon^{\mu\nu\rho}A_\mu^1F_{\nu\rho}^1&=\frac{\epsilon^{\mu\nu\rho}}{4d^4}\frac{(2d_1d_2\partial_\mu d_2\partial_\nu d_3\partial_\rho d_1-2d_2d_1\partial_\mu d_1\partial_\nu d_2\partial_\rho d_3)}{d+d_3}\\
&=\frac{\epsilon^{\mu\nu\rho}}{2d^4}\frac{d_1-d_2}{d+d_3}\partial_\mu d_1\partial_\nu d_2\partial_\rho d_3)\\
&=
\end{aligned}
\tag{56}
$$

Now we use another gauge and write

$$
\widetilde{U}=\frac{1}{d}\sqrt{\frac{d}{2(d-d_3)}}\begin{pmatrix}d_1-id_2 & -d+d_3\\ d-d_3 & d_1+id_2\end{pmatrix}=\frac{1}{d}\begin{pmatrix}\sqrt{\frac{d(d+d_3)}{2}}\frac{d_1-id_2}{\sqrt{d_1^2+d_2^2}} & -\sqrt{\frac{d(d-d_3)}{2}}\\ \sqrt{\frac{d(d-d_3)}{2}} & \sqrt{\frac{d(d+d_3)}{2}}\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}}\end{pmatrix},
$$

which defines $\widetilde{U}=(\widetilde{\boldsymbol{u}}_1,\widetilde{\boldsymbol{u}}_2)$, where $\widetilde{\boldsymbol{u}}_1=\frac{1}{\sqrt{2d}}(\frac{d_1-id_2}{\sqrt{d-d_3}},\sqrt{d-d_3})^T$ corresponds to eigenvalue $+d$. Recall that

$$
U=\frac{1}{d}\begin{pmatrix}\sqrt{\frac{d(d+d_3)}{2}} & \sqrt{\frac{d(d-d_3)}{2}}\frac{-d_1+id_2}{\sqrt{d_1^2+d_2^2}}\\ \sqrt{\frac{d(d-d_3)}{2}}\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}} & \sqrt{\frac{d(d+d_3)}{2}}\end{pmatrix},
$$

which defines $U=(\boldsymbol{u}_1,\boldsymbol{u}_2)$ with $\boldsymbol{u}_1=\frac{1}{\sqrt{2d}}(\sqrt{d+d_3},\frac{d_1+id_2}{\sqrt{d+d_3}})^T$ corresponds to eigenvalue $+d$, so we have

$$\widetilde{U}=U\ \mathrm{diag}\left(\frac{d_1-id_2}{\sqrt{d_1^2+d_2^2}},\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}}\right),$$

i.e. $\widetilde{\boldsymbol{u}}_1=\boldsymbol{u}_1\frac{d_1-id_2}{\sqrt{d_1^2+d_2^2}}$ and $\widetilde{\boldsymbol{u}}_2=\boldsymbol{u}_2\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}}$.

which defines $\widetilde{U}=(\widetilde{\boldsymbol{u}}_1,\widetilde{\boldsymbol{u}}_2)$, where $\widetilde{\boldsymbol{u}}_1$ corresponds to eigenvalue $+d$. Then, we have

$$\partial_\mu\widetilde{\boldsymbol{u}}_1=-\frac{\partial_\mu d}{2d}\widetilde{\boldsymbol{u}}_1+\frac{1}{\sqrt{2d}}\left(\frac{\partial_\mu(d_1-id_2)}{\sqrt{d-d_3}}-\frac{\partial_\mu(d-d_3)}{2(d-d_3)^{3/2}}(d_1-id_2),\frac{\partial_\mu(d-d_3)}{2\sqrt{d-d_3}}\right),$$

so that

$$\begin{aligned}\widetilde{\boldsymbol{u}}_1^\dagger\partial_\mu\widetilde{\boldsymbol{u}}_1&=-\frac{\partial_\mu d}{2d}+\frac{(d_1+id_2)\partial_\mu(d_1-id_2)}{2d(d-d_3)}-\frac{1}{4d}\frac{(d_1^2+d_2^2)\partial_\mu(d-d_3)}{(d-d_3)^2}+\frac{1}{4d}\partial_\mu(d-d_3)\\&=-\frac{1}{4d}\partial_\mu(d+d_3)+\frac{2d\partial_\mu d-2d_3\partial_\mu d_3+2i(d_2\partial_\mu d_1-d_1\partial_\mu d_2)}{4d(d-d_3)}-\frac{1}{4d}\frac{(d+d_3)\partial_\mu(d-d_3)}{d-d_3}\\&=i\frac{d_2\partial_\mu d_1-d_1\partial_\mu d_2}{2d(d-d_3)}=-i\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{2d(d-d_3)},\end{aligned}\tag{57}$$

again, using $d_1\partial_\mu d_2-d_2\partial_\mu d_1=4(N_3^2+N_4^2)(N_2\partial_\mu N_1-N_1\partial_\mu N_2)+4(N_1^2+N_2^2)(N_3\partial_\mu N_4-N_4\partial_\mu N_3)$ we have

$$\widetilde{\boldsymbol{u}}_1^\dagger\partial_\mu\widetilde{\boldsymbol{u}}_1=\frac{i}{N_3^2+N_4^2}(N_3\partial_\mu N_4-N_4\partial_\mu N_3)+\frac{i}{|\vec{N}|^2}\left(N_1\partial_\mu N_2-N_2\partial_\mu N_1+N_3\partial_\mu N_4-N_4\partial_\mu N_3\right),$$

Recall that $\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1=\frac{i}{2d}\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{d+d_3}$, it is easy to verify that

$$\widetilde{\boldsymbol{u}}_1^\dagger\partial_\mu\widetilde{\boldsymbol{u}}_1=\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1+\frac{d_1+id_2}{\sqrt{d_1^2+d_2^2}}\partial_\mu\left(\frac{d_1-id_2}{\sqrt{d_1^2+d_2^2}}\right),\tag{58}$$

as should be. It is proper to call $A^N=i\boldsymbol{u}_1^\dagger\partial_\mu\boldsymbol{u}_1=-\frac{1}{2d}\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{d+d_3}$ and $A^S=i\widetilde{\boldsymbol{u}}_1^\dagger\partial_\mu\widetilde{\boldsymbol{u}}_1=\frac{d_1\partial_\mu d_2-d_2\partial_\mu d_1}{2d(d-d_3)}$, we see that on the equator they are opposite to each other: $A^N|_{d_3=0}=-A^S|_{d_3=0}$.

Now we further have

$$\begin{aligned}A_\mu^N&=\frac{d_3-d}{2d}\partial_\mu\arctan\frac{d_2}{d_1}=\frac{(d_3+d)\partial_\mu\arctan\frac{d_2}{d_1}}{2d}-\partial_\mu\arctan\frac{d_2}{d_1},\\A_\mu^S&=\frac{d_3+d}{2d}\partial_\mu\arctan\frac{d_2}{d_1}=\frac{(d_3-d)\partial_\mu\arctan\frac{d_2}{d_1}}{2d}+\partial_\mu\arctan\frac{d_2}{d_1},\end{aligned}$$

which is simply

$$A_\mu^S=A_\mu^N+\partial_\mu\arctan\frac{d_2}{d_1},\tag{59}$$

in agreement with Eq. (58).

At the level of $\boldsymbol{d}$, one cannot find a smooth gauge that is regular on the entire two-sphere. This can be either due to nonzero Chern number (obstruction to finding a smooth gauge) or artifact (bad parameterization), which cannot be distinguished on the level of $\boldsymbol{d}$.

## 15.2 Generalized Hopf insulator

A generalized Hopf insulator is protected by the symmetry

$$J^{-1}\mathcal{H}(\boldsymbol{k})J=-\mathcal{H}^*(\boldsymbol{k}),\quad J=i\sigma^2\otimes 1_{n\times n},\tag{60}$$

suppose $\mathcal{H}=U\Lambda U^\dagger$ is a $2n$-by-$2n$ Hamiltonian. Then it is classified by $\pi_3(Sp(n)/U(n))=\mathbb{Z}_2$, or

$$n_h=\frac{1}{24\pi^2}\int\mathrm{Tr}(U^\dagger dU)^3=\frac{1}{24\pi^2}\int d^3k\epsilon^{\mu\nu\lambda}\mathrm{Tr}[U^\dagger\partial_\mu UU^\dagger\partial_\nu UU^\dagger\partial_\lambda U]\in\mathbb{Z}_2\quad(n>1),\tag{61}$$

The symmetry (60) demands that $U$ is of the form $U=(V,JV^*)$, and that

$$U^\dagger dU=\begin{pmatrix}A & B\\ -B^\dagger & A^*\end{pmatrix},\tag{62}$$

where $A$ is an $n$-by-$n$ matrix and $B$ is related to $A$ by $A = V^\dagger dV$, $B = V^\dagger J dV^*$. We then have

$$\mathrm{Tr}(U^\dagger dU)^3 = \mathrm{Tr}(-3AdA - 2A^3) + \mathrm{Tr}(-3A^* dA^* - 2A^{*3}) = -3\mathrm{CS}[A] - 3\mathrm{CS}[A^*] = -6\mathrm{CS}[A], \tag{63}$$

where we used the fact that the Chern-Simons integral is always real. Note that now we have

$$n_h = -\frac{1}{4\pi^2}\mathrm{CS}[A], \tag{64}$$

which is a direct generalization of the 2-band Hopf insulator For time reversal invariant system, e.g. 3D TI, we know that the Chern-Simons term is quantized, i.e. integral $\frac{1}{4\pi^2}\int d^3\boldsymbol{k}\mathrm{Tr}[ada + \frac{2}{3}a^3] \in \mathbb{Z}$. Here of course we do not have time reversal symmetry, but our Chern-Simons term is also quantized to integer, which instead comes from the constraints for $V$, which are $V^\dagger V = 1$ and $V^\dagger J V^* = 0$.

In the 3D TI case, the time reversal symmetry $\mathcal{T}^2 = -1$ requires that the number of filled bands is even due to Kramer's theorem, and as a consequence $a$ must be an even-dimensional matrix. But in the case of generalized Hopf insulator, the symmetry (60) does not have this constraint, and $a$ can be a matrix of odd and even size alike. This just reflects the generality of the response term, which exists for any non-interacting model with arbitrary number of bands.

In Ludwig's paper, the generalized Hopf insulator sits in the class $(\epsilon_W, \eta_W) = (-1, -1)$.

### 15.3 New throught and papers

The Hamiltonian has 10 generators (see `https://math.stackexchange.com/questions/1293679/a-question-and-a-conjecture`

Marcel Franz paper on Witten effect in TI: `https://arxiv.org/pdf/1001.3179.pdf`

Proof that the axion integral equals the Fu-Kane formula: `https://arxiv.org/pdf/0910.5954.pdf`

`https://arxiv.org/pdf/2009.08466.pdf`

`https://arxiv.org/abs/1809.02853`

`https://arxiv.org/pdf/2206.10636v2.pdf`

## 16 Notes on the tenfold way

### 16.1 SPT phase and free fermion SPT

Symmetry protected topological (SPT) phase

- are short-range entangled, $G$-symmetric phases (whose classification depends on the dimension, and degrees of freedom being boson or fersmion, etc.);
- upon breaking the symmetry $G$ all reduces to the same trivial, product state;
- are "invertible phases": can be trivialized by stacking other $G$-SPTs.

Free fermion SPT, there are way too many unitary symmetries. So for simplicity and generality, only consider anti-unitary symmetries (tenfold way).

- There are only two anti-unitary symmetries which can be given the meaning of TR and PH/CC. Motivation for HP/CC: charge $U(1)$ broken down to $\mathbb{Z}_2$. On the 1st quantuzed Hamiltonian matrix we have, up to unitary transformation,
$$\mathcal{T}\colon H^* \sim H, \qquad \mathcal{C}\colon \quad H^* \sim -H, \tag{65}$$
- we call the composition of TR and PH/CC the chiral symmtery:
$$H \sim -H, \tag{66}$$
and when chiral symmetry is present we can always write $H$ in the form
$$H \sim \begin{pmatrix} 0 & h \\ h^\dagger & 0 \end{pmatrix}. \tag{67}$$
- here $\sim$ means up to unitary matrix $H$. For example, $H^* = U^\dagger H U$, so $H = U^T H^* U^* = U^T U^\dagger H U U^*$. Turns out $UU^* = \pm 1$ (Kramers theorem). Understanding: this in the $U(1)$ case is classified by $\mathcal{H}^2(\mathbb{Z}_2^{\mathcal{T}}, U(1)_{\mathcal{T}}) = \mathbb{Z}_2$.
- There is a corresponding Kramer's theorem for the anti-unitary CC. There is no such theorem for chiral. Therefore, there are in total ten classes:

| Class | A | AI | AII | AIII | BDI |
|---|---|---|---|---|---|
| $\mathcal{T}$ | 0 | + | − | 0 | + |
| $\mathcal{C}$ | 0 | 0 | 0 | 0 | + |
| $\mathcal{S}=\mathcal{TC}$ | 0 | 0 | 0 | + | + |
| $\mathcal{H}(\boldsymbol{k})\in$ | $G_{m,m+n}(\mathbb{C})$ | $G_{m,m+n}(\mathbb{C})$ | $G_{2m,2(m+n)}(\mathbb{C})$ | $U(m)$ | $U(m)$ |
| Subject to | – | $\mathcal{H}^*(\boldsymbol{k})=\mathcal{H}(-\boldsymbol{k})$ | $\sigma_y\mathcal{H}^*(\boldsymbol{k})\sigma_y=\mathcal{H}(-\boldsymbol{k})$ | – | $h(\boldsymbol{k})^*=h(-\boldsymbol{k})$ |
| Class | CII | D | C | DIII | CI |
| $\mathcal{T}$ | − | 0 | 0 | − | + |
| $\mathcal{C}$ | − | + | − | + | − |
| $\mathcal{S}=\mathcal{TC}$ | + | 0 | 0 | + | + |
| $\mathcal{H}(\boldsymbol{k})\in$ | $U(2m)$ | $G_{m,2m}(\mathbb{C})$ | $G_{m,2m}(\mathbb{C})$ | $U(2m)$ | $U(m)$ |
| Subject to | $\sigma_y h^*(\boldsymbol{k})\sigma_y=h(-\boldsymbol{k})$ | $\tau_x\mathcal{H}(\boldsymbol{k})^*\tau_x=-\mathcal{H}(-\boldsymbol{k})$ | $\tau_y\mathcal{H}(\boldsymbol{k})^*\tau_y=-\mathcal{H}(-\boldsymbol{k})$ | $h(\boldsymbol{k})^T=-h(-\boldsymbol{k})$ | $h(\boldsymbol{k})^T=h(-\boldsymbol{k})$ |

## 16.2 Classification of free fermion SPT

We follow Ludwig's review paper `https://arxiv.org/pdf/1512.08882.pdf`. A historical survey:

- The classification calcation first appears in the paper `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.78.195125`: this paper first establishes that for each dimension there are five nontrivial classes out of the ten symmetry classes. Using the argument of Anderson localization.
- Both Ludwig's `https://arxiv.org/abs/0912.2157` and Kitaev's `https://arxiv.org/abs/0901.2686` make the important claim that

  Any tenfold way class in any dimension has a massive Dirac Hamiltonian representative in the same class. (68)

  this claim is based on the experience in PhysRevB.78.195125 and proved through enumeration in 0912.2157. Kitaev on the other hand gives the circumstances for this assumption to be true and give the band topological classification using $K$-theory.
- Again built on the Dirac representative assumption, Joel's paper `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.85.045104` enumerates the quantum anomaly aspects of the Dirac Hamiltonians as a physical diagnose of the classification (i.e. classification using topological response).

In this sense, Ludwig summarizes that there are three ways for the classification:
(1) boundary theory of NL$\sigma$M for Anderson localization (2) bulk theory (homotopy, $K$-theory, band topology) (3) bulk-boundary correspondence using the Dirac fermion representative, i.e. anomaly (by Moore et al.)

## 16.3 Boudnary theory of anderson localization

We will omit this part. Relevant reference is `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.78.195125`, which cited `https://arxiv.org/pdf/cond-mat/0006360.pdf`. The basica idea is to study the NL$\sigma$M at the boundary theory and examine its symmetry and see what topological terms one can write down. Note that this is a classical theory. It turns out there are only two situations:

- a $\mathbb{Z}_2$ term iff $\pi_{d-1}(G/H)=\mathbb{Z}_2$;
- a WZW term iff $\pi_d(G/H)=\mathbb{Z}$.

where $d$ is the bulk dimension. $H$ is the Hamiltonian and $G$ is identified with the Fermion bilinears in the replica space. For Anderson localization, NL$\sigma$M, and the topological terms, See Altland&Simons.

## 16.4 Bulk classication

Spectral flattening: given a (1st quantized) Hamiltonian matrix in momentum space

$$\mathcal{H}(\boldsymbol{k})=U_{\boldsymbol{k}}\mathcal{E}(\boldsymbol{k})U_{\boldsymbol{k}}^\dagger\to U_{\boldsymbol{k}}\begin{pmatrix}1_m & \\ & -1_n\end{pmatrix}U_{\boldsymbol{k}}:=\mathcal{Q}(\boldsymbol{k}), \tag{69}$$

Underlying assumption: $\mathcal{H}(\boldsymbol{k})$ and $\mathcal{Q}(\boldsymbol{k})$ are homotopy equivalent.

Then: for each $\boldsymbol{k}$, $\mathcal{Q}$ in 1-to-1 correspondence with $U\in\frac{U(m+n)}{U(m)\times U(n)}$.

To summarize: after spectral flattening, $m$ ones and $n$ minus ones, the Hamiltonian is fully determined by the eigenvector matrix $U$. Call it the target space. Since $U(m)\times U(n)$ rotation does not change the flattened Hamiltonian, we see the topological information is fully stored in

$$f\colon BZ\to \frac{U(m+n)}{U(m)\times U(n)}\equiv G_{m,m+n}(\mathbb{C}). \tag{70}$$

- Step 1: fix band number $m$ and $n$. It is clear that topologically (i.e. up to smooth diformation) different Hamiltonian are characterized by homotoly types of $f$, which we denote $[BZ, G_{m,m+n}]$. If you are ok with $BZ^d=T^d\sim S^d$ (the spherical cow approximation), then calculating the homotopy $\pi_d(G_{m,m+n})$ gives the classification.

- Step 2: allow $m$ and $n$ to change. The classification should not change. So it seems the classification is given by classes of

$$[BZ, \underbrace{\cup_{0<m\le n} G_{m,m+n}}_{C_0(1)}]. \tag{71}$$

- Step 3: but the above overclassifies things. For example: $\left(\begin{smallmatrix} X & 0\\ 0 & -X\end{smallmatrix}\right)$ should be trivial! So need a coarser equivalence relation. This (weirdly) turns out to be a mathematically easier problem which is established: K-theory. It classifies the homotopy classes

$$K^0_{\mathbb{C}}(\overline{BZ}):=\Big[\overline{BZ}, \underbrace{\cup_{k\in\mathbb{Z}}\lim_{s\to\infty}\frac{U(2s)}{(U(s+k)\times U(s-k))}}_{C_0}\Big], \tag{72}$$

  where $\overline{BZ}$ is the topological space of $BZ$ seen by the target space: it identifies $\boldsymbol{k}$ and $-\boldsymbol{k}$ for all the eight real classes. For example, for the two complex classes we would use directly $BZ=T^d$.

- Step 4: adding symmetries. TR or PH/CC or Chiral imposes additional constraint on $U$, so the target space is more complicated. E.g. for class AII with TR$=-1$,

$$H_{\text{flattened}}(\boldsymbol{k})^*=Q_{\text{flattened}}(-\boldsymbol{k}),$$

  For fixed $m$ and $n$, this is a very hard problem.

Now we calculate the following two cases:

- Free-fermion Hamiltonian (i.e. Dirac): since we need to compactify infinity point ("spherical cow" classification), the correct mathematical object for the classification is the relative $K$-group

$$\widetilde{K}^{0,q}_{\mathbb{R}}(\bar{S}^d):=K^{0,q}_{\mathbb{R}}(\bar{B}^d,\partial\bar{B}^d)\cong\pi_0(R_{q-d}). \tag{73}$$

- Band insulator: this is the above case, and we have

$$K^0_{\mathbb{C}}(\overline{BZ})=\bigoplus_{s=0}^{d}C^s_d\pi_0(R_{q-s}). \tag{74}$$

  the $s=d$ term, $\pi_0(R_{q-d})$, gives the "spherical cow" classification in free fermion, and justifies the assumption (68).

Example: for the usual 3D TI (class AII, with $(\mathcal{T},\mathcal{C},\mathcal{S}=(-,0,0))$, we have $K^0_{\mathbb{C}}(\overline{BZ})=\mathbb{Z}_2\oplus 3\mathbb{Z}_2\oplus 0\oplus\mathbb{Z}$, where the first $\mathbb{Z}_2$ coincide with the continuous fermion result.

General frame work: Twisted equivariant $K$-theory. See Freed, Moore, Twisted Equivariant Matter (2013) `https://link.springer.com/content/pdf/10.1007/s00023-013-0236-x.pdf`

Idea: General ideal of twisted equivariant $K$-theory see Definition 7.33; application to band insulator (and arbitary symmetry) see Theorem 10.15; for band insulator with antiunitary symmetry (i.e. tenfold way), see Hypothesis 10.24 and Corollary 10.25. For fundations of band theory, see App. D.

Equivalence relation: $X'\sim X''$ if $X'\oplus Y\approx X''\oplus Y$;

Difference class $d(A,B)$: $(A',B')\sim(A'',B'')$ if $A'\oplus B''\sim A''\oplus B'$.

Therefore there is a map $(A,B)\mapsto k$.

$K$-theory (following Spin Geometry, P58): $X$ a compact space. $V(X)$: set of all isomorphism clsses of complex vector bundles over $X$ as a semigroup. $F(X)$: free abelian groups generated by $V(X)$; $E(X)$: subgroup of $F(X)$ generated by $[V]+[W]-([V]\oplus[W])$, where $+$ and $\oplus$ are from $F(X)$ and $V[X]$. The $K$-group of $X$ is defined as $K(X)=F(X)/E(X)$.

Categorical definition: If $G$ any abelian group and $f\colon V(X)\to G$ any semi-group homomorphism. Then there is a unique homomorphism $\tilde{f}\colon K(X)\to G$ s.t. $\tilde{f}\alpha=f$.

Table 3: The tenfold way table. The superscripts denote the anomaly type and topological terms in the response theory. $\mathcal{F}$ means gauge ($U(1)$ or $SU(2)$) anomaly, $\mathcal{R}$ means pure gravitational anomaly, and m means mixed anomaly.

| Cartan | Symmetry | | | Spatial dimension $d$ | | | | | | | | Note |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T | C | S | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| A | 0 | 0 | 0 | $\mathbb{Z}^{\mathcal{F},\mathrm{CS}}$ | 0 | $\mathbb{Z}^{\mathcal{F},\mathrm{CS}}$ | 0 | $\mathbb{Z}^{\mathcal{F},\mathrm{CS}}$ | 0 | $\mathbb{Z}^{\mathcal{F},\mathrm{CS}}$ | 0 | Complex class; QH |
| AIII | 0 | 0 | 1 | 0 | $\mathbb{Z}^{\mathcal{F},\theta}$ | 0 | $\mathbb{Z}^{\mathcal{F},\theta}$ | 0 | $\mathbb{Z}^{\mathcal{F},\theta}$ | 0 | $\mathbb{Z}^{\mathcal{F},\theta}$ | |
| AI | +1 | 0 | 0 | $\mathbb{Z}^{\mathrm{m,CS}}$ | 0 | 0 | 0 | $2\mathbb{Z}^{\mathrm{m,CS}}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | |
| BDI | +1 | +1 | 1 | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathrm{m},\theta}$ | 0 | 0 | 0 | $2\mathbb{Z}^{\mathrm{m},\theta}$ | 0 | $\mathbb{Z}_2$ | |
| D | 0 | +1 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathcal{R},\mathrm{CS}}$ | 0 | 0 | 0 | $2\mathbb{Z}^{\mathcal{R},\mathrm{CS}}$ | 0 | |
| DIII | −1 | +1 | 1 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathcal{R},\theta}$ | 0 | 0 | 0 | $2\mathbb{Z}^{\mathcal{R},\theta}$ | |
| AII | −1 | 0 | 0 | $2\mathbb{Z}^{\mathrm{m,CS}}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathrm{m,CS}}$ | 0 | 0 | 0 | TRI TI |
| CII | −1 | −1 | 1 | 0 | $2\mathbb{Z}^{\mathrm{m},\theta}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathrm{m},\theta}$ | 0 | 0 | |
| C | 0 | −1 | 0 | 0 | 0 | $2\mathbb{Z}^{\mathcal{R},\mathrm{CS}}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathcal{R},\mathrm{CS}}$ | 0 | |
| CI | +1 | −1 | 1 | 0 | 0 | 0 | $2\mathbb{Z}^{\mathcal{R},\theta}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}^{\mathcal{R},\theta}$ | |

## 16.5 Quantum anomalies

Original reference is the paper by Ryu, Moore and Ludwig, `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.85.045104`.

Perturbative anomaly and global anomaly. Take the representative massive Dirac Hamiltonian.

"gauging" the symmetry: now a standard technique for SPT. Physically makes sense – response theory as a probe.

- has U(1) symmetry? yes: couple to a $U(1)$ gauge field;
- has SU(2) symmetry? yes: couple to $SU(2)$ field.
- when no continuous symmetry? then couple to background gravitational field.

Integrating out fermions, get effective action for the gauge field.

$$e^{-W_{\mathrm{eff}}[A_\mu]} = \int \mathcal{D}[\bar{\psi},\psi] e^{-S[\bar{\psi},\psi;A_\mu]}. \tag{75}$$

where

$$W_{\mathrm{eff}} = i\int \Omega_{d+1},$$

and

$$\Omega_D = ch(\mathcal{F}), \quad \text{or } \hat{A}(\mathcal{R}), \quad \text{or } ch(\mathcal{F})\hat{A}(\mathcal{R}), \tag{76}$$

where $ch(\mathcal{F}) = r + \frac{i}{2\pi}\mathrm{tr}\mathcal{F} + \frac{1}{2!}\left(\frac{i}{2\pi}\right)^2 \mathrm{tr}\mathcal{F}^2 + \cdots$, and $\hat{A}(\mathcal{R}) = 1 + \frac{1}{(2\pi)^2}\frac{1}{12}\mathrm{tr}\mathcal{R}^2 + \frac{1}{(2\pi)^2}\left[\frac{1}{288}\mathrm{tr}\mathcal{R}^2 + \frac{1}{360}\mathrm{trR}^4\right] + \cdots$

So we have:

- In even spacetime dimension: a single massive Dirac fermion has a gauge U(1) chiral/axial anomaly.
- In odd spacetime dimension:

  $\theta$ term: bulk does not have anomaly but boundary gives Chern–Simons term which is anomalous. This gives the first descent relation:

  $$\Omega_{D,\theta} = d\Omega_{D-1,\mathrm{CS}}, \quad \text{for even spacetime dimension } D$$

- In odd spacetime dimension: Chern–simons term. This term is not gauge invariant, implying the bulk has anomaly; must be cancelled by anomaly on the boundary. Both boundary and bulk are anomalous. This gives the second descent relation:

  $$\delta_v \Omega_{D-1,\mathrm{CS}} = d\Omega_{D-2}.$$

Important characteristic classes:

- $H^{1,2,\ldots,n}(BO_n,\mathbb{Z}_2)$: Stiefel-Whitney class
- $H^{2,4,\ldots,n}(BU_n,\mathbb{Z})$: Chern class
- $H^{4,8,\ldots,2n}(BO_n,\Lambda)$: the rational Pontrjagin class with $\Lambda$ an integral domain containing $\frac{1}{2}$ ( e.g. $\mathbb{Q}$). the $\hat{A}$-genus belongs here.

- $H^{4,8,...,2n}(BO_n, \mathbb{Z})$: the integral Pontrjagin class. Related to the Chern class.
- $\chi \in H^n(BSP_{2n}, \Lambda)$, the universal Euler class. Here $\Lambda$ is any ontegral domain containing $\frac{1}{2}$. Note we have $H^*(BSO_{2n}, \Lambda) = \Lambda[p_1, ..., p_n, \chi]/\langle \chi^2 - p_n \rangle$.

## 16.6 Djigraaf-Witten theory

Throughout: $M$: oriented 3-manifold. $G$: compact Lie group. $E$: principal $G$ bundle. Goal: find TQFT.

Motivation: Q: why compact Lie group? A: any Lie group satisfies

$$1 \to G_0 \to G \to \Gamma \to 1,$$

$$1 \to \pi_1(G) \to \widetilde{G} \to G \to 1,$$

$G$ compact $\Rightarrow \Gamma$ a finite group. (Not ethat $\Gamma$ is also a Lie group!) Lie algebra is identified with the simply connected part of $G$, $G_0$. Both $G_0$ and $\widetilde{G}$ are simply connected. Compact Lie group has two extreme cases: finite group and connected, simply connected group. The TQFT for these two cases can be easily obtained which gives the same classification. So main result is that classification of TQFT is the same for all compact Lie group.

Statement 1: If $E$ is trivial: the TQFT (i.e. the action) can be defined as the usual CS in terms of $A$ (a Lie algebra valued one form).

Example: $G$ is connected and simply connected: then $E$ is trivial, TQFT is the usual CS.

If $E$ is nontrivial: requirement: TQFT needs to sum over all possible bundles.

Fact: any 3-manifold $M = \partial B$ for some 4-manifold $B$.

Statement 2: If can choose $B$ s.t. $E$ extends over $B$, then TQFT can be defined by a $\theta_B$ term on $B$. In this case, a flat connection $A$ that extends as a flat connection to $B$ gives $S_E(A) = 0$.

Statement 3: If cannot choose $B$ s.t. $E$ extends over $B$, can still have

$$S_E(A) - S_E(A') = \theta_{B:=M\times I}[F]. \tag{77}$$

here $A$ and $A'$ are two connections on $E$.

Statement 4: for finite group case, every principal $G$ bundle $E$ has a unique flat connection $A_{E,flat}$ ($A_{E,flat}$ for $E$ is in one-to-one correspondence with $\lambda$: $\pi_1(M) \to G$), giving $S_E(A_{E,flat}) = 0$. Therefore one can identifiy $S_E(A)$ with the $\theta$ term on RHS of Eq. (77). Physically, requirements are (i) and (ii) on Page 3, which gives the classification. So the final claim is that the classificadtion for TQFT of a finite group is $H^3(BG, R/Z) \cong H^4(BG, Z)$.

## 16.7 Relation between the spectral projector and the tenfold way classification

The concept of spectral projector should now be familiar: it is simply the space that the unitary matrix that diagonalizes the Hamiltonian, $U(\boldsymbol{k})$, lives in. It is also called the (momentum-space) target space (note there can be several notion of target spaces so one really has to be careful). For example, for class A, we know that the $U_{\boldsymbol{k}} \in U(N_+ + N_-)/U(N_+) \times U(N_-) = G_{N_+,N_+ + N_-}(\mathbb{C})$, so that $U$ defines a map $\mathrm{BZ}^d \to G_{N_+,N_+ + N_-}(\mathbb{C})$. On the other hand, in class AIII, the chiral class, $U$ can always be brought to the chiral form $\begin{pmatrix} 0 & q(\boldsymbol{k}) \\ q^\dagger(\boldsymbol{k}) & 0 \end{pmatrix}$ with $q(\boldsymbol{k}) \in U(N)$ without additional constraints, and that $U$ defines a map $\mathrm{BZ}^d \to U(N)$. Since there is no additional constraints on the projector, the above maps completely characterizies the classification of these classes (class A and class AIII), meaning that the relevant classification is just $\pi_d(G_{N_+,N_+ + N_-}(\mathbb{C})) = \begin{cases} \mathbb{Z}, & d \text{ even} \\ 0, & d \text{ odd} \end{cases}$ and $\pi_d(U(N)) = \begin{cases} \mathbb{Z}, & 0 \text{ even} \\ \mathbb{Z}, & d \text{ odd} \end{cases}$. This is also called topological band theory.

One then turn to the other eight classes with additional discrete symmetries. Due to the additional symmetries, the target space (i.e. space of spectral projectors) is no longer a simple established Lie group, and distinct maps that are not deformable to each other can no longer be characterized by homotopy groups (instead, one must resort to a K-theory approach, which we postpone to later sections).

The target space (i.e. the space of projectors) can be found in the 2nd column of Table III of the PRB paper by Schnyder, Ryu, Furusaki and Ludwig (2008), `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.78.195125`, or in Ludwig's review paper, Table I (`https://arxiv.org/pdf/1512.08882.pdf`).

The spectral projectors, afterall, contains the full information of band insualtors. So one may still wonder whether there is a (fairly simple) way to obtain the tenfold way classification (actually, the other eight classes with additional discrete symmetries) directly from the spectral projectors. I am not aware of that (other than the K-theory approach). The PRB paper above obtained the classification by counting the number of surface Dirac cones of each of the eight classes upon adding perturbations (i.e. a surface description), so I assume there is not a simple way (i.e. no simple bulk description in terms of the spectral projectors) to achieve that.

A partial summary: whenever the spectral projector belongs to a simple/established Lie group without additional constraints imposed by discrete symmetry, the classification is directly given by the homotopy group of that Lie group. This is why our generalized Hopf insulator works, because the target space is $Sp(n)/U(n)$, without additional discrete symmetry. [See also Table C2 of the tenfold way paper `https://iopscience.iop.org/article/10.1088/1367-2630/12/6/065010`, or the arXiv version `https://arxiv.org/pdf/0912.2157.pdf`; all four classes belongs to this case.]

## 16.8 Can we obtain the tenfold way classification directly from the spectral projector?

As mentioned in the last subsection, only in the complex classes A and AIII is the classification directly obtained from the homotopy of the spectral projector. The rest eight classes has a spectral projector that is a Lie group with more structure defined by the additional discrete symmetry, and the homotopy invariant is very hard to compute, if not impossible. but the spectral projector, after all, contains all the information of the ten classes, and there should be a way of obtaining the classification directly from them. So the question is: how to obtain the tenfold classification (we really only need the classification of the eight real classes) directly from the spectral projector?

This, first of all, cannot be just a homotopy theory, as just explained. It turns out that such a answer is provided by Kitaev in his famous paper `https://arxiv.org/abs/0901.2686`, using K-theory. In Kitaev's original formulation, he used the K-theory that gives the classification of the classifying space of the spectral projector in real space. The real spectral projector is related to the momentum space projectors $Q(\boldsymbol{k})$ or $q(\boldsymbol{k})$ in a certain way that is vaguely mentioned in Andreas Ludwig's review paper `https://arxiv.org/pdf/1512.08882.pdf` (In Table II of Andreas Ludwig's paper, he looked at the space of the momentum space projectors at TRIM, and defined the classifying space). I'm yet to understand the detail of it.

To summarize, yes, there is a way to obtain the tenfold way classification directly from the spectral projector, provided by Kitaev using K-theory. Such a classification recovers the homotopy classification for class A and AIII (the two complex classes), but for the other eight real classes, one must correctly resort to K-theory of the spectral projectors, rather than homotopy theory of the spectrak projectors. So in this sense, the classification using spectral projectors is K-theoretic in nature, rather than homotopy-theoretic in nature.

This the first of the three classification methods, as mentioned in Ludwig's review paper `https://arxiv.org/pdf/1512.08882.pdf`.

## 16.9 Is there, then, a homotopy theory for the tenfold way table?

The last subsection concludes that the homotopy of spectral projectors can only be employed to obtain the classification of class A and class AIII. So the next general question is, whether any entry of the tenfold way table can be expressed as some homotopy invariant.

The answer is positive. The classification is simply (see Table 4):

$$\text{Classification for class } X \text{ is } \begin{cases} \pi_d(G/H)=\mathbb{Z}, & \text{or} \\ \pi_{d-1}(G/H)=\mathbb{Z}_2, & \text{if } \pi_d(G/H) \text{ vanishes, or} \\ 0, & \text{if both above vanishes.} \end{cases} \tag{78}$$

here, importantly, $G/H$ is the target space of the NL$\sigma$M (some real space target space), which is not the space of the spectral projectors. This is clearly remarked in the footnote "‡" of the tenfold way paper.

As am example, we know that $\pi_d(O(N))=\pi_d(SO(N))=\begin{cases} 0, & \text{if } d=2,4,5,6(mod 8), \\ \mathbb{Z}_2, & \text{if } d=0,1(mod 8), \\ \mathbb{Z}, & \text{if } d=3,7(mod 8). \end{cases}$ From this we can immediately obtain the class DIII results, for which $G/H \in O(N)$.

To understand why the classification is given by the homotopy invariant given in Eq. (78), one of course has to go back to the Theory of NL$\sigma$M for Anderson localization, which we omit here. We stress again that for a given class (and dimension), the target space that $G/H$ lives in, and that target space for the spectral projectors $Q(\boldsymbol{k})$ (or $q(\boldsymbol{k})$ for chiral classes) are very different things, and are not the same space.

This the second of the three classification methods, as mentioned in Ludwig's review paper `https://arxiv.org/pdf/1512.08882.pdf`.

## 16.10 A few more comments about the dimensional reduction

For previous notes, see the WordReference blog and other .tex files.

There are many dimensional reduction sequence one can write. Sec. 2 of Ludwig's paper talks about the connection between A$(d=2n+2) \to$ AIII$(d=2n+1)$.

Table 4: The tenfold way table.

| Cartan | Symmetry | | | Projector | Spatial dimension $d$ | | | | | | | | Note |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T | C | S | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| A | 0 | 0 | 0 | $\frac{U(N_+ + N_-)}{U(N_+)\times U(N_-)}$ | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | 0 | Complex class; QH |
| AIII | 0 | 0 | 1 | chiral $U(N)$ | 0 | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | 0 | $\mathbb{Z}$ | |
| AI | +1 | 0 | 0 | PRB78 195125 | $\mathbb{Z}$ | 0 | 0 | 0 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | |
| BDI | +1 | +1 | 1 | PRB78 195125 | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | 0 | 0 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | |
| D | 0 | +1 | 0 | PRB78 195125 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | 0 | 0 | $2\mathbb{Z}$ | 0 | |
| DIII | −1 | +1 | 1 | PRB78 195125 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | 0 | 0 | $2\mathbb{Z}$ | |
| AII | −1 | 0 | 0 | PRB78 195125 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | 0 | 0 | TRI TI |
| CII | −1 | −1 | 1 | PRB78 195125 | 0 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | 0 | |
| C | 0 | −1 | 0 | PRB78 195125 | 0 | 0 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | 0 | |
| CI | +1 | −1 | 1 | PRB78 195125 | 0 | 0 | 0 | $2\mathbb{Z}$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ | |

- For class A, this is the complex class without any symmetry. the projector is $G_{N_+,N_++N_-}(\mathbb{C}) = \frac{U(N_++N_-)}{U(N_+)\times U(N_-)}$, and the topological band theory (i.e. map from the $d$-dimensional BZ to the projector target space) is simply $\pi_d(G_{N_+,N_++N_-}(\mathbb{C}))$ which is $\mathbb{Z}$ in even $d$ and 0 in odd $d$. Such a homotopy invariant can equivalently be computed using Chern number (the $n$-th Chern number for $d = 2n$ dimension). This class describes (integer) quantum Hall effect in $d = 2$ and $d = 4$.

- For class AIII which is a chiral class (and only have chiral symmetry), the projector can always be brought to the chiral form $Q(k) = \begin{pmatrix} 0 & q(k) \\ q^\dagger(k) & 0 \end{pmatrix}$ with $q(k) \in U(N)$, and the topological band theory classification is $\pi_d(U(N))$ which is $\mathbb{Z}$ in odd $d$ but 0 in even $d$. Such a homotopy invariant can equivalently be computed using winding number, which is an explicit integral formula over the odd-dimensional BZ. This completes the analysis of the AIII class; but if we want to go further, we can show that the response term (i.e. the Chern-Simons form) for the AIII class actually equals half of the winding number, i.e. the Chern-Simons term is quantized. In other words, the Chern-Simons value can tell the parity of the class, but not enough to tell the class exactly.

- a non-zero Chern number (on the even-dimensional BZ) means that wavefunction cannot be defined globally (on that BZ). One then defines two patches of the wavefunction, and we know that the Chern number can be expressed as a winding number of the transition function. The relation with the dimensional reduction A→AIII is that that the transition function is given by the off-diagonal block of the projector of the AIII class, and this is how the A($d = 2n + 2$) and AIII($d = 2n + 1$) are related, with the same classification $\mathbb{Z}$.

- These classes both have U(1) symmetry, and that's why we can talk about their response theory (to an external U(1) gauge field $A$). Note that for $d = 2n$, the response theory has a term $\int d^{2n}kch[f]\int d^{2n+1}xCS[A]$, while for $d = 2n - 1$, the response theory has a term $\nu_{2n-1}\int d^{2n}xch_n[F]$.

Xiaoliang's original paper talks about $A(d = 4) \to AII(d = 3) \to AII(d = 2)$.

## 16.11 How many integral invariants are there after all?

Questions. Reading Shinsei and Ludwig's paper, it seems to imply that: Chern number/winding is exactly the index that labels the homotopy class; and that whenenver this is the class, we have an integral formula that gives this index, and hence we have the integral formula to compute the homotopy class. The integral formula (?) is nothing but computing the degree of map, so there is some relation (not specified clearly) among homotopy class, (co)homology class, Chern class, Chern number/winding number, and degree of map. In the ten-fold way classification it seems the relation is a "benign" one. But it seems a not-benign-one can happen, see `https://math.stackexchange.com/questions/887396/is-the-homotopy-class-given-by-the-degree`.

Partial answer:

- Note that the Chern class is defined for complex vector bundles. This means the target space is at most the grassmannian $\mathrm{Gr}_k(\mathbb{C}^n) = \frac{U(n)}{U(k)\times U(n-k)}$, and that the Chern number is defined only on maps $f\colon \mathcal{M} \to \mathrm{Gr}_k(\mathbb{C}^n)$ with $\mathcal{M}$ any compact manifold of even dimensions. That is to say, Chern number (and the integral formula) can only be defined for homotopy $\pi_d(\mathrm{G}_k(\mathbb{C}^n))$; any other homotopy types cannot be defined via the Chern class. Note that this exactly corresponds to class A.

- Now we know how the integral formula for the Chern number is defined (via the Chern class). Next we turn to the winding number. So why is the formula $\propto \int_{\mathrm{BZ}^{2n-1}} \mathrm{Tr}[(q^{-1}dq)^{2n-1}]$ characterizing the homotopy class $\pi_{2n-1}(U(N))$ with $q \in U(N)$? in the case of $n = 2$, this has the intepretation of finding the winding of $\pi_3(S^2)$, since and $U(N)$ contains several copies of $S^2 \cong SU(2)$. But what about $n > 2$? To see this, remember how homotopy of $U(N)$ is calculated: the easiest way is through the long exact sequence for $U(N-1) \to U(N) \to S^{2N-1}$ (note that the homotopy long exact sequence becomes isomorphism since $S_d(S^{2N+1}) = 0$ for very large $N > d$), so that for large enough $N$, the homotopy of $U(N)$ "stabilizes" (i.e. does not change with $N$ anymore), so indeed for large $N$, $U(N)$ contains also spheres $d$ with $d$ odd. So we see that the winding number formula is always essentially computing $\pi_d(S^d)$, which is really the winding number $S^d \to S^d$ (also the degree of map for $S^d \to S^d$). This eliminates our doubts for the winding number.

- To summarize, we now see where the two distinct integral formula (that gives integer numbers) come from: the even dimensional one is for Grassmannian and comes from Chern class, while the odd dimensional one is for $U(N)$ (or its subgroup, such as $O(N)$) and comes essentially from the winding number of $S^d \to S^d$. All the integral formulas in the theory of TI eventually comes from these two (even the Chern-Simons term or the Hopf invariant too).

## 16.12 Response and anomaly for the ten classes

The paper to read is the one by Ryu, Moore and Ludwig, `https://journals.aps.org/prb/abstract/10.1103/PhysRevB.85.045104` (the arXiv version: `https://arxiv.org/abs/1010.0936`). This is the first paper that discusses the relation between quantum anomaly and tenfold way. The question answered in this paper is: what is the response theory and corresponding anomaly for the tenfold way?

This the third of the three classification methods, as mentioned in Ludwig's review paper `https://arxiv.org/pdf/1512.08882.pdf`.

Chern-Simons quantization for one massive Dirac Hamiltonian: according to Fradkin and Leon, one massive Dirac cone coupled to U(1) gauge field gives the term $e^{iS_{\mathrm{CS}}[A]}$, with

$$S_{\mathrm{CS,1DF}}[A] = -\frac{1}{16\pi}\frac{m}{|m|}\int d^3x \epsilon_{\mu\nu\lambda}A^\mu F^{\nu\lambda} = -\frac{1}{8\pi}\frac{m}{|m|}\int d^3x \epsilon_{\mu\nu\lambda}A^\mu \partial_\nu A_\lambda, \tag{79}$$

where $m$ is mass note the universal definition $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = 2\partial_{[\mu}A_{\nu]}$. Note here we have set $e = \hbar = 1$. Putting them back, we actually have

$$S_{\mathrm{CS,1DF}}[A] = -\frac{1}{16\pi}\frac{m}{|m|}\frac{e^2}{\hbar}\int d^3x \epsilon_{\mu\nu\lambda}A^\mu F^{\nu\lambda} = -\frac{1}{8}\frac{m}{|m|}\frac{e^2}{h}\int d^3x \epsilon_{\mu\nu\lambda}A^\mu F^{\nu\lambda},$$

for one massive Dirac, we have $\sigma_H = \frac{1}{2}\frac{e^2}{h}\frac{m}{|m|}$, so that

$$S_{\mathrm{CS,1DF}}[A] = -\frac{\sigma_H}{4}\int d^3x \epsilon_{\mu\nu\lambda}A^\mu F^{\nu\lambda}. \tag{80}$$

Using exterior differential forms, we usually write $F = dA$, which really means that $F = \frac{1}{2!}F^{\mu\nu}dx_\mu dx_\nu$ (the general rule for $r$-form has the prefactor $\frac{1}{r!}$), and $A = A_\mu dx_\mu$, so $dA = \partial_\nu A_\mu dx_\nu dx_\mu = \frac{1}{2}(\partial_\nu A_\mu - \partial_\mu A_\nu)dx_\nu dx_\mu$, which agrees with the usualy definition. So we concisely write

$$S_{\mathrm{SC,1DF}}[A] = -\frac{1}{8\pi}\frac{m}{|m|}\int A\wedge dA = -\frac{1}{8\pi}\frac{m}{|m|}\int AdA = -\frac{1}{8\pi}\frac{m}{|m|}\int A\wedge F. \tag{81}$$

The actin for $\nu = 1/m$ ($m$ is integer) Laughlin FQH state is $S = \frac{m}{4\pi}\int d^3x \epsilon^{\mu\nu\lambda}A_\mu\partial_\nu A_\lambda$. Wen showed using gauge and coordinate invariance of the CS term that the boundary theory is $S = \frac{m}{4\pi}\int dxdt(\partial_t\phi\partial_x\phi - v(\partial_x\phi)^2)$, where $v$ is the edge mode velocity viewed in the reference frame that has $A_0 = 0$.

Edge of $\mu = 1/m$ Langhlin FQH: (see Fradkin Eq. (15.17)) on the edge satisfies $h(x) = \frac{\rho(x)}{n_0} = \frac{\rho(x)}{\frac{\nu}{2\pi\ell_0^2}} = \frac{\rho(x)}{\frac{\nu}{2\pi\frac{\hbar c}{eB}}} = \frac{\rho(x)hc}{\nu eB}$, this gives the electrostatic energy to be $H = \int dx\pi\hbar\frac{v}{\nu}\rho(x)^2$, where $v = E/H$ is the drift velocity.

The chiral spin liquid mean field Hamiltonian has a Chern number of 1, and counting spin degeneracy gives 2. So the CS term has $m = 2$ and is the same as the $\nu = 1/m = 1/2$ Laughlin state.

Some comment in Fradkin's notation: Eq. (7.136) writes the CS term as $\frac{\theta}{4\pi}\epsilon_{\mu\nu\lambda}A^\mu F^{\nu\lambda}$, with $\theta = 2\pi S$ and Eq. (10.78) writes $\frac{\theta}{4}\epsilon_{\mu\nu\rho}A^\mu F^{\nu\rho}$ with $\theta = \frac{2S+1}{4\pi}(\mathrm{sgn}(m_+) + \mathrm{sgn}(m_-))$. The former uses $h = 1$ and the latter uses $\hbar = 1$. To avoid confusion, we will always explicitly write $e^2/\hbar$ in the following.

## 16.13 Anyons and Chern-Simons

Recap of Chern-Simons: the response term for one massive Dirac fermion is $e^{iS[A]}$, with

$$\begin{aligned} S_{1DF}[A] &= -\frac{1}{16\pi}\frac{e^2}{\hbar}\frac{m}{|m|}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho} = -\frac{1}{8\pi}\frac{e^2}{\hbar}\frac{m}{|m|}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu \partial_\nu A_\rho \\ &= -\frac{1}{8\pi}\frac{e^2}{\hbar}\frac{m}{|m|}\int A\wedge dA = -\frac{1}{8\pi}\frac{e^2}{\hbar}\frac{m}{|m|}\int A\wedge F, \end{aligned} \tag{82}$$

where note that the $F = dA = \frac{1}{2!}F^{\mu\nu}dx_\mu \wedge dx_\nu$. On the other hand, note that the following is quantized

$$\frac{1}{16\pi^2}\frac{e^2}{\hbar}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho} = \frac{1}{8\pi^2}\frac{e^2}{\hbar}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu \partial_\nu A_\rho = \frac{1}{8\pi^2}\frac{e^2}{\hbar}\int A\wedge \underbrace{dA}_{F} = \frac{1}{8\pi^2}\frac{e^2}{\hbar}\int A\wedge F \in \mathbb{Z}, \tag{83}$$

so that for the above case of one massive Dirac fermion, the partition function is $e^{i\pi \mathrm{Sgn}(m)}$, which is half quantized. This is the definition of parity anomaly. Note that more generally, for a tight-binding model, we have

$$\begin{aligned} S_{\text{t-b}}[A] &= -\frac{C}{8\pi}\frac{e^2}{\hbar}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho} = -\frac{C}{4\pi}\frac{e^2}{\hbar}\int d^3x\, \epsilon^{\mu\nu\rho}A_\mu \partial_\nu A_\rho \\ &= -\frac{C}{4\pi}\frac{e^2}{\hbar}\int A\wedge dA = -\frac{C}{4\pi}\frac{e^2}{\hbar}\int A\wedge F = -\frac{\sigma_{\text{H}}}{2}\int A\wedge F, \end{aligned} \tag{84}$$

with $C$ the Chern number, and $\sigma_{\text{H}} = \frac{1}{2}(\sigma_{xy}-\sigma_{yx}) = \frac{e^2}{h}C = \frac{e^2}{2\pi\hbar}C$. Note that for one massive Dirac we get $C = \frac{1}{2}\mathrm{sgn}(m)$, which is the anomaly. Note that Fradkin's text book defined $e = h = 1$ everywhere and put the $2\pi$ on the denominator of $\hbar$ explicitly in the numerical coefficient.

A word about the phase: we have heard that the CS term always appears as a phase in the partition function, i.e. if we define $S_{\text{t-b}}$ as above, which is real, then it appears in the partition function as $e^{iS_{\text{t-b}}[A]}$. A careful account for the phase nature can be found in Leon's notes 217B, page 59.

In usual literature, one finds the notation $\mathcal{L} = \frac{C}{4\pi}A\partial A$ to write the CS term, where $A\partial A$ stands for $\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho$. Note this is the coefficient for the majority ways of writing the CS term in Eq. (82), so no confusion should be caused. Also, in this notation, one has $e = \hbar = 1$.

The Hopf term in momentum space: there the gauge field $\boldsymbol{A}(\boldsymbol{k}) = i\langle \boldsymbol{k}|\partial_{\boldsymbol{k}}|\boldsymbol{k}\rangle$ which fixes the overall constant. In this case, the Hopf invariant is defined as

$$\frac{1}{8\pi^2}\int_{\text{BZ}} d^3\boldsymbol{k}\epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho} = \frac{1}{2\pi^2}\int_{\text{BZ}} d^3\boldsymbol{k}\epsilon^{\alpha\beta\gamma\delta}u^\alpha \partial_{k_x}u^\beta \partial_{k_y}u^\gamma \partial_{k_z}u^\delta \in \mathbb{Z}, \tag{85}$$

compare with the Hopf insulator notation: there $n_h = \int d^3k A_x^{MRW}F_{yz}^{MRW} + A_y^{MRW}F_{zx}^{MRW} + A_z^{MRW}F_{xy}^{MRW} = \frac{1}{4\pi^2}\int d^3k A_x F_{yz} + A_y F_{zx} + A_z F_{xy} = \frac{1}{8\pi^2}\int d^3k \epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho}$, so agree with the above normalization. Note that compared with the real space normalization, it is as if we were setting $e = 1$ and $\hbar = 1/2$.

More generally, if we have

$$S_{\theta\text{-anyon}}[A] = \frac{c}{4\pi}\frac{e^2}{\hbar}\int A\wedge F, \tag{86}$$

then the statistical angle is (note in Fradkin's notation this is $\delta$ and he defined $\theta$ as something else)

$$\theta_F = \frac{1}{2\pi}C,$$

so the statistic angle is

$$\theta = \frac{1}{2\theta_F} = \frac{\pi}{C}.$$

for QAHE, $C = 1$, so $\theta = \pi$ which is Fermion; for KL CSL, $C = (2S+1) = 2$ so $\theta = \frac{\pi}{2}$, the semion.

## 16.14 Chern-Simons counting

Following convention of David Tong, Subir Sachdev's book, and Steven Simon's book. Below we set $e = \hbar = 1$ as do must books (except for Fradkin's – see comments above)

Subir: a quasiparticle excitation is labeled by a vector $\boldsymbol{\ell}$, which is a set of integers representing its charges under the gauge fields $\boldsymbol{a}$. In the context of FQHE, the quasiparticle also carries the external $U(1)$ charge. Hence the complete action is

$$\mathcal{L} = \frac{1}{4\pi}\boldsymbol{a}\cdot K\cdot \wedge d\boldsymbol{a} + \frac{1}{2\pi}\boldsymbol{t}\cdot A\wedge d\boldsymbol{a} - j^\mu \ell^I a_\mu^I,$$

(Subir here is regarding the action above as the action for a single quasiparticle excitation labeled by $\boldsymbol{\ell}$, i.e. a single quasiparticle that carries charges under all external gauge fields. We can also instead regard this action as that for the elementary quasiparticles – each carry only a single charge with $\boldsymbol{\ell} = (0, ..., 0, 1, 0, ..., 0)$. This is just a minor change of point of view.) where the vectorized form is with respect to the gauge index $I$, i.e. $\boldsymbol{a}^I = a^I$. (The full component expression is $a^I_\mu$.) Here $\boldsymbol{t}$ is a set of integers; the $U(1)$ charge of the quasiparticle is

$$Q = \boldsymbol{\ell} \cdot K^{-1} \boldsymbol{t}.$$

In the absence of quasiparticles (i.e. when $\boldsymbol{\ell} = 0$) then we can integrate out the internal gauge fields to get

$$S = (\boldsymbol{t} K^{-1} \boldsymbol{t}) \frac{1}{4\pi} \int A \wedge dA.$$

compare with Eq. (84) we see that

$$\sigma_{\rm H} = \frac{\boldsymbol{t} K^{-1} \boldsymbol{t}}{2\pi} \xrightarrow{1=e=\hbar} \boldsymbol{t} K^{-1} \boldsymbol{t} \frac{e^2}{h}.$$

If there is no external field, i.e. $A = 0$, then we can solve for the equation of motion

$$j^\mu \ell_I = \frac{1}{2\pi} K_{IJ} \epsilon^{\mu\nu\rho} \partial_\nu a^J_\rho,$$

which is a proper generalization of

$$j^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu a_\rho.$$

(Note that instead of writing $j^\mu \ell_I = \frac{1}{2\pi} K_{IJ} \epsilon^{\mu\nu\rho} \partial_\nu a^J_\rho$ Steven Simons writes $j^\mu_I = \frac{1}{2\pi} K_{IJ} \epsilon^{\mu\nu\rho} \partial_\nu a^J_\rho$, and interpretes it as the current for the $I$-th quasiparticle. As we mentioned above, this is simply looking at a emergent charge charge vector of $\ell_I = (0, ..., 0, 1_I, 0, ..., 0)$.)

Following Steven Simon (but notation-wise we follow Subir): we have

$$j^0 \ell_I = \frac{K_{IJ}}{2\pi} (\partial_1 a^J_2 - \partial_2 a^J_1) = \frac{K_{IJ}}{2\pi} b^J \quad \Rightarrow \quad b^J = 2\pi [K^{-1}]^{IJ} j^0 \ell_I,$$

which says that the flux $b^J$ bounds to each of the charges $j^0 \ell_I$.

As Steven Simon comments: different fluxes (with index $I$) bound to different charges (with index $J$) but a particular charge only seens a particular flux (with the same index $J$). The saddle point action is then

$$\mathcal{L} = -\frac{1}{2} j^\alpha \ell_I a^I_\alpha.$$

The phase for wrapping a particle of type $\boldsymbol{\ell}$ around a particle of type $\boldsymbol{\ell'}$ is given by

$$\vartheta_{\ell,\ell'} = 2\pi (\boldsymbol{\ell} K^{-1} \boldsymbol{\ell'}),$$

(this is related to the $R$ matrix (which is the braiding matrix) by $e^{i\vartheta_{\ell,\ell'}} = R_{\ell,\ell'} R_{\ell',\ell}$.) and when talking about self statistics one means two identical particles $\ell$ and $\ell$ switching position (and this is really the $R$ matrix $e^{i\Theta_\ell} = R_{\ell,\ell}$). The angle is

$$e^{i\Theta_\ell} = e^{i\pi \boldsymbol{\ell} \cdot K^{-1} \cdot \boldsymbol{\ell}}.$$

Fradkin: $\mathcal{L}[\psi, A_1 + A_2] + k_1 \mathrm{CS}[A_1] + k_2 \mathrm{CS}[A_2] = \mathcal{L}[\psi, A] + k_1 \mathrm{CS}[A - A_2] + k_2 \mathrm{CS}[A_2] = \mathcal{L}[\psi, A] + k_1 \mathrm{CS}[A] + (k_1 + k_2) \mathrm{CS}[A_2] - 2k_1 A_1 dA_2$, integrating out $A_2$ we get $\mathcal{L}[\psi, A] + k_1 \mathrm{CS}[A] - (k_1 + k_2) \frac{k_1^2}{(k_1+k_2)^2} \mathrm{CS}[A]$, giving rise to a term $\left(k_1 - \frac{k_1^2}{k_1+k_2}\right) \mathrm{CS}[A] = -\frac{1}{\frac{1}{k_1} + \frac{1}{k_2}} \mathrm{CS}[A]$,

Subir's explanation of the parton method for FQHE is illustrative. Note that there the cases considered are the $\nu = 1/3$ Laughlin, and the Jain sequence. Other fillings would require some other choices of partons.

The parton $c = \psi_1 \psi_2 \psi_3$ enlargest the Hilbert space and the original one is recovered by $\psi_1^\dagger \psi_1 = \psi_2^\dagger \psi_2 = \psi_3^\dagger \psi_3$. Importantly, this enforces the parton density $\rho_1 = \rho_2 = \rho_3 = \rho$ where the last one is the electron density. The parton decomposition also gives

$$\mathcal{L} = \sum_i \frac{1}{2m_i} \psi_i^\dagger (\nabla - i b_i)^2 \psi_i,$$

(to be honest I don't fully understand how this is obtained – Subir mentioned this is obtained using Hubbard–Stratonovich but I don't see it), with $b_3 = -b_1 - b_2 + eB$, where $B$ is the external field. Now, importantly, we have

$$\rho = B\nu,$$

(where strictly we have $\rho = \nu B/\Phi_0$ with $\Phi_0 = h/e$ but we omit it since it is a constant) and furthermore, the action (lagrangian) $\mathcal{L}$ says we also have

$$\rho_i = B_i \nu_i,$$

and since $\rho_i = \rho$ for $i = 1, 2, 3$ we see the situation is exactly like $V = IR$ connected in parallel, with $(V, I, R) \mapsto (\rho, B, \nu)$. So it should not be surprising to see

$$\frac{1}{\nu} = \sum_i \frac{1}{\nu_i},$$

Up to now it is still general. From now on, we'd like to put the partons in integral filling so that $\nu_i \in \mathbb{Z}$. One can then ask what the maximal set of fillings $\nu$ this can give rise to. These are the Laughlin state at $\nu = 1/3$ and the Jain states with $\nu = N/2N + 1$.

Quite similarly we can consider $p$ partons, with the same constraints that $\psi_i^\dagger \psi_i = \psi_j^\dagger \psi_j$ for $i \neq j$. All things carry through and we have

$$\frac{1}{\nu} = \sum_{i=1}^{p} \frac{1}{\nu_i},$$

The Jain sequence is then putting $\nu_1 = \nu_2 = \cdots = \nu_{p-1} = 1$ and $\nu_p = N$, giving $\frac{1}{\nu} = p - 1 + \frac{1}{N} = \frac{pN-N+1}{N}$ so that $\nu = \frac{N}{pN+1}$ (we have shifted $p \to p + 1$).

The Moore–Read state describes the physics at $\nu = 1/2$, which can also be understood from the three-parton method by taking $N \to \infty$ (so that $B_3 = 0$).

## 16.15 Chapter 10 of Fradkin

- 10.2 A (say square lattice) $\pi$-flux-per-unit-cell tight-binding model gives rise to two bands with Chern numbers $C = \pm 1$. This can be viewed as contributed from two massive Dirac fermions giving

  $$C = \pm \frac{1}{2} \left( \mathrm{sgn}(m_1) + \mathrm{sgn}(m_2) \right) = \pm 1$$

  with $m_1 > 0$ and $m_2 > 0$. The effective action we arrive at is $\mathcal{S}[\psi_1, \psi_2]$ (here $\psi_{1,2}$ are the two massive Dirac fermions).

- 10.3 The phase fluctuations in 10.2 can be modeled by a (emergent, internal) gauge field. [The amplitude fluctuations can be manually supressed by formally considering a large $N$-flavor theory.] The effective action we arrive at is $\mathcal{S}[\psi_1, \psi_2, a]$ [$a$ is the (emergent, internal) gauge field].

- 10.4 Carrying out the integral of the Dirac fermions $\psi_1$ and $\psi_2$, we get

  $$S[a] = -\frac{C}{8\pi} \frac{e^2}{\hbar} \int d^3x \epsilon^{\mu\nu\rho} a_\mu f_{\nu\rho} \equiv -\frac{C}{4\pi} \frac{e^2}{\hbar} \int a \wedge da,$$

  where we have used the standard definition that $f = da = \frac{1}{2!} \sum_{\mu,\nu=1}^{3} f^{\mu\nu} dx_\mu \wedge dx_\nu$ and that $f^{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a - \mu$. Note that if the gauge field $a$ above were an external probe field (denoted as $A$), then it would give the Hall conducctivity $\sigma_{\mathrm{H}} \equiv \frac{1}{2}(\sigma_{xy} - \sigma_{yx}) \xrightarrow{\sigma_{xy} = -\sigma_{yx}} \sigma_{xy} = \frac{e^2}{\hbar} C = \frac{e^2}{2\pi\hbar} C$.

- Summary: from here we see that a $C = 1$ chern band contributes a Chern simons term in the action $S[a] = -\frac{1}{4\pi} \int ada$ (here we follow the usual convention to set $e = \hbar = 1$ and omit the wedge symbol). Using the $K$ matrix formalism (see e.g. Steven's book or the comments somewhere above), we see that the statistical angle is $\vartheta = \pi$. [Recall that the statistical angle for a single type of anyon $a$ is defined as the angle acquired in the partition function when switching the position of two (identical) $a$'s, i.e. half of the angle acquired when cycling one $a$ around the other $a$.] Importantly, as Fradkin mentioned, this is the statistical angle in additional to the intrinsic fermion/boson statistics.

- 10.5.

-

Fundamental fact: the phase (in the partition function) accumulated by encycling an (elementary) electric charge $e$ around a flux quantum $\Phi_0 = h/e$ is $2\pi$. This is a fact since it can be checked: this phase is just $\frac{1}{\hbar} \cdot e \cdot (h/e) = 2\pi$.

The flux is always in the units of $h/e$. It's just that in the unit where one sets $\hbar = e = 1$, the flux becomes in the units of $2\pi\hbar/e = 2\pi$. Thus a flux quantum (in the usual sense) is identified with a $2\pi$ flux. A magnetic monopole, when created, modifies the flux by $2\pi$, hence we say that a magnetic monopole carries flux quantum $2\pi$. A vison is a $\pi$ flux. (see also Subir's chapter 14 on extended XY model in 2+1D.)

(General fact about Dirac quantization condition: $q_e q_m = 2\pi$.)

From the fundamental fact above, we examine the composite particle $b$ of an electron binds one flux quantum $2\pi$. We then see that when $b$ encycles $b$ once, a phase of $2\pi$ is accumulated. This means that exchanging $b$ results in a phase of $\pi$ — and this phase is on top of the fermionic nature of electron — hence we claim that $b$ accumulates a phase of $2\pi \equiv 0$ when exchanged with another $b$, hence a boson. Similarly, we see that a particle with charge $e$ but itself is a boson, when binds a $2\pi$ flux, becomes a fermion.

This can also be rationalized using the $e, m, \varepsilon$ particles in toric code, but we really have to make more rules. The usual rule would be assuming $m$ (naturally the vison) to carry $\pi$ flux. This is a boson. $e$ is also a boson so assume it carries charge $2e$, but this is incompatible with the mutual statistics that when $e$ encycles $m$ we get a $-1$. So we must assume $e$ carries charge $e$ (and is a boson).

According to Steven Simon's book: $e$ is particle bound to 1 unit of eletric charge; $m$ is particle bound to $\pi$ flux; $\varepsilon$ is particle bound to 1 unit of electric charge and $\pi$ flux. And the $K$ matrix is $\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$. Then using the definition of $\vartheta_{\ell,\ell'} = 2\pi(\boldsymbol{\ell} K^{-1} \boldsymbol{\ell'})$ we see that encycling $e$ around $m$ (or vice versa) gives a phase $\pi$; encycling $e$ (or $m$) around $\varepsilon$ gives a phase $\pi$.

## 16.16 Eilenberg-Mac Lane space

Hatcher: A path-connected space whose fundamental group is isomorphic to a given group $G$ which has a contractible universal covering space is called a $K(G,1)$ space.

Rotman: for amy $G$, there exists the Eilenberg-Mac Lane space $K(G,1)$ is path-connected, aspherical (i.e. its $n$th homotopy groups vanish for all $n > 1$) and whose fundamental group $\pi_1(K(G,1)) \cong G$. The classifying space of $G$ is the universal covering space of $K(G,1)$; the space $BG$ is acyclic (i.e. $H_0(BG) \cong \mathbb{Z}$ and 0 for all $n \geq 1$), $G$ acts properly on $BG$, and $BG/G \approx K(G,1)$.

Example: $S^1$ is a $K(\mathbb{Z},1)$.

The cohomology of an abstract group $G$ coincides with the cohomology of a certain topological space: $H^n(K(G,1),A) \cong H^n(G,A)$.

`https://mcgreevy.physics.ucsd.edu/w21/final-papers/2021W-239-Lu-Dachuan.pdf`

## 16.17 Class AII With interactions?

Chong, Potter and Senthil, `https://www.science.org/doi/10.1126/science.1243326`

## 16.18 Dirac fermion in a Landau level

Consider a spatial 2D plane with a boundary. Suppose the boundary is at $y = 0$, perpendicular to the $y$ axis. So that throughout we will have translation invariance along the $x$ direction, meaning that $k_x$ is a good quantum number. Assume magnetic field is $\boldsymbol{B} = B(0,0,1)$. To respect the translation symmetry in the $y$ direction, we use the gauge (instead of $\boldsymbol{A} = (0, Bx, 0)$)

$$\boldsymbol{A} = (-yB, 0, 0),$$

so that $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} = (\partial_y A_z - \partial_z A_y, \partial_z A_x - \partial_x A_z, \partial_x A_y - \partial_y A_x) = (0,0,B)$. The Dirac Hamiltonian we are considering is, following the graphene review (RMP 81 109),

$$H_0 = v_F(\boldsymbol{\sigma} \cdot (-i\boldsymbol{\nabla} + e\boldsymbol{A}),$$

using $\psi(x,y) = e^{ik_x x}\phi(y)$ [4],

$$v_F \begin{pmatrix} 0 & \partial_y - k_x + Bey \\ -\partial_y - k_x + Bey & 0 \end{pmatrix} \phi(y) = E\phi(y).$$

Define

$$\xi = \frac{y}{\ell_B} - \ell_B k_x, \quad \omega_c = \sqrt{2}\frac{v_F}{\ell_B}, \qquad \psi_N(\xi) = \frac{1}{2^{N/2}\sqrt{N!}} e^{-\frac{\xi^2}{2}} H_N(\xi)$$

the solution is

$$E_{\pm N} = \pm\omega_c\sqrt{N}, \qquad \psi_{N,k_x}(x,y) = e^{ik_x x}\phi_N(\xi), \qquad \phi_{N,\pm}(\xi) = \begin{pmatrix} \psi_{N-1}(\xi) \\ \pm\psi_N(\xi) \end{pmatrix}; \qquad E_0 = 0, \qquad \phi_0 = \begin{pmatrix} 0 \\ \psi_0(\xi) \end{pmatrix}.$$

[4] instead of

$$v_F \begin{pmatrix} 0 & -i\partial_x - ik_y - iBex \\ -i\partial_x ik_y + iBex & 0 \end{pmatrix} \phi(x) = E\phi(x).$$

Now include a potential to the Hamiltonian

$$H = H_0 + H_{\rm DW}, \qquad H_{\rm DW} = \sigma^z m(y), \qquad m(y) = m_0 \tanh(y/y_0),$$

need to compute the matrix element

$$M_{n,k_x;n',k_x'} = \int dxdy \psi^*_{n,k_x} H_{\rm DW} \psi_{n',k_x'},$$

which actually imposes $k_x = k_x'$. We then write $M_{n,k_x;n',k_x'} = \delta(k_x - k_x') M_{n,n',k_x}$, where

$$M_{n,n',k_x} = \frac{m_0}{2^{n/2}\sqrt{n!}2^{n'/2}\sqrt{n'!}} \int dy e^{-\frac{1}{2\ell_B^2}(y-\ell_B^2 k_x)^2} H_n(\frac{y}{\ell_B} - \ell_B k_x) \tanh\left(\frac{y}{y_0}\right) e^{-\frac{1}{2\ell_B^2}(y-\ell_B^2 k_x)^2} H_{n'}(\frac{y}{\ell_B} - \ell_B k_x),$$

The essence is to calculate the integral

$$\int dy \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} H_n(\frac{y-\mu}{\sqrt{2}\sigma}) \tanh y H_{n'}(\frac{y-\mu}{\sqrt{2}\sigma}),$$

$$(\epsilon - m)\phi_a = L\phi_b,$$
$$(\epsilon + m)\phi_b = L^\dagger \phi_a, \quad L((\epsilon+m)\phi_b) = v_x(p_x m)\phi_b + (\epsilon+m)L\phi_b,$$

so

$$(\epsilon^2 - m^2)\phi_a = LL^\dagger \phi_a - v_x(p_x m)\frac{1}{\epsilon+m} L^\dagger \phi_a,$$

so that

$$\mathcal{L}\phi_a = 0,$$
$$\mathcal{L} = \epsilon^2 - m^2 + \frac{v_x(p_x m)L^\dagger}{\epsilon+m} - LL^\dagger.$$

when $m = \mu x/x_0$, we have $p_x m = -i\hbar(\mu/x_0)$, so we have

$$\mathcal{L} = \epsilon^2 - m^2 + \frac{v_x(-i\hbar\mu/x_0)L^\dagger}{\epsilon+m} - LL^\dagger,$$

## 16.19 Tenfold way – talk

SPT phase:

- distinct, $G$-symmetric phases; dimension; boson/fersmion;
- breaking $G$ then phase becomes trivial
- "Invertible phase": trivialized by stacking another $G$-SPT

Free fermion SPT, given anti-unitary symmetries (tenfold way).

- There are only two anti-unitary symmetries which can be given the meaning of TR and PH/CC. On the 1st quantuzed Hamiltonian matrix we have

  $$H^* \sim H, \quad H^* \sim -H,$$

  we call the composition of TR and PH/CC the chiral symmtery:

  $$H \sim -H.$$

  here $\sim$ means up to unitary matrix $H$. For exmaple, $H^* = U^\dagger H U$, so $H = U^T H^* U^* = U^T U^\dagger H U U^*$. Turns out $UU^* = \pm 1$.

- Physical motivation: TR and charge $U(1)$; latter can break down to $\mathbb{Z}_2$ which is PH/CC.

**Method 1: band topology**

after spectral flattening, $m$ ones and $n$ minus ones, the Hamiltonian is fully determined by the eigenvector matrix $U$. Call it the target space. Since $U(m) \times U(n)$ rotation does not change the flattened Hamiltonian, we see the topological information is fully stored in

$$f\colon BZ \to U(m+n)/U(m) \times U(n) \equiv G_{m,m+n}(\mathbb{C}).$$

Step 1: fix band number $m$ and $n$. It is clear that topologically (i.e. up to smooth diformation) different Hamiltonian are characterized by homotoly types of $f$, which we denote $[BZ, G_{m,m+n}]$. If you are ok with $BZ^d = T^d \sim S^d$ (the spherical cow approximation), then calculating the homotopy $\pi_d(G_{m,m+n})$ gives the classification.

Step 2: allow $m$ and $n$ to change. The classification should not change. This is indeed the case.

Step 3: adding symmetries. TR or PH/CC or Chiral imposes additional constraint on $U$, so the target space is more complicated. E.g. for class AII with TR$= -1$,

$$H_{\text{flattened}}(\boldsymbol{k})^* = Q_{\text{flattened}}(-\boldsymbol{k}),$$

For fixed $m$ and $n$, this is a very hard problem.

Step 4: turns out the physically relevant question is allowing to change $m$ and $n$. This (weirdly) turns out to be a mathematically easier problem which is established: K-theory. It classifies the homotopy classes

$$K^0_{\mathbb{C}}(\overline{BZ}) := \Big[\overline{BZ}, \underbrace{\cup_{k\in\mathbb{Z}} \lim_{s\infty} U(2s)/(U(s+k)\times U(s-k))}_{C_0}\Big],$$

where $\overline{BZ}$ is the topological space of $BZ$ seen by the target space: it identifies $\boldsymbol{k}$ and $-\boldsymbol{k}$ for all the eight real classes. For example, for the two complex classes we would use directly $BZ = T^d$.

Detailed computation tool: $K^{0,q}_{\mathbb{R}}(\bar{B}^d, \partial\bar{B}^d) = \widetilde{K}^{0,q}_{\mathbb{R}}(\bar{S}^d) \cong \pi_0(R_{q-d})$, $\widetilde{K}^{p,q}_{\mathbb{R}}(X) \cong \widetilde{K}^0_{\boldsymbol{R}}(S^r X)$, which roughly speaking allows to use the suspension of the base space to change the target space. This is still for sphere. Then for torus, relate to the $K$-homology of the real-space torus by the Baum-Connes isomorphism to get

$$[\pi, R_q] = \pi_0(R_{q-d}) \oplus \bigoplus_{s=0}^{d-1} C^s_d \pi_0(R_{q-s}),$$

the first term gives the spherical cow classification.

**Method 2: Anderson localization on the boundary and NL$\sigma$M**

**Method 3: Quantum anomaly**

Andreas Ludwig: `https://arxiv.org/pdf/1512.08882.pdf`

Kitaev: `https://arxiv.org/abs/0901.2686`

Three classifications: (1) boundary (NL$\sigma$M, Anderson localization) (2) bulk (homotopy, $K$-theory, band topology) (3) bulk-boundary correspondence, i.e. Anomaly (by Moore et al.)

# 17 FQHE

`https://www.damtp.cam.ac.uk/user/tong/qhe/qhe.pdf`

$\Phi_0 = \frac{h}{e} = 2\pi\frac{\hbar}{e} = 2\pi\ell_B^2 B$, $\ell_B = \sqrt{\frac{\hbar}{eB}}$.

Electron density: $n = \nu \cdot \frac{B}{\Phi_0}$.

One electron is has in itself one vortex

For $\nu = 1/m$ FQHE:

- Quasi-hole is defined via adding the factor $\prod_i (z_i - \eta)$ to the Laughlin wave function.
- A quasi-hole has charge $+e/m$ (where an electron has charge $-e$, here we defined $e = |e|$). Outline of argument: introducing $m$ quasi-holes amounts to introducing one electron; parameter vs dynamical variable means quasi-hole vs actual electron.
- $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = e^{i\pi\alpha}\psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$ after an anti-clockwise exchange
- AB phase: a test particle with charge $e^*$ moves along a path enclosing magnetic flux of $\Phi$ (but no other changes enclosed) picks up a phase in the wave function $e^{i\gamma}$ with $\gamma = \frac{e^*\Phi}{\hbar}$. This is also the Berry phase.
- $\alpha = 1/m$.

- $n$ quasi-holes as a single object: then $\alpha = n^2/m$.

Composite fermions for states Laughlin states $\nu = 1/m$ with $m$ odd with $\nu^* = 1$ and $\nu = 1, 2/3, 3/5, 4/7...$ ($\nu^* = 2, 3, 4, ...$):

- A composite fermion is an electron bound to $m-1$ further vortices, and the whole thing is a fermion when $m$ is odd.
- The magnetic fields experienced by electrons $B$ and composite fermions $B^*$ differ: $B^* = B - (m-1)n\Phi_0$, $n = \nu^* B^*/\Phi_0 = \nu B/\Phi_0$. $n$ is the electron density, which is the same as composite fermion density.
- The FQHE for electrons can be thought of as an IQHE for composite fermions

Composite fermion for the $\nu = 1/2$ fermionic FQHE: $\nu^* =$?

- One electron bound to two vortices

The Half-filled Landau level: biggest thing to explain: no Hall plateaux at $\nu = 1/2$.
Explanation from composite fermion: composite fermion sees no magnetic field, $B^* = 0$.

## 17.1 Benoit Etienne's talk at les Houches

$H = \hbar\omega_c(a^\dagger a + 1/2)$, where $\omega_C = |qB|/m$. LLL: $\psi(x,y) = f(z)e^{-|z|^2/4\ell_B^2}$ where $f$ is any holomorphic function. We defined $\ell_B = \sqrt{\frac{\hbar}{qB}}$ the magnetic length, each particle occupies $2\pi\ell_B^2$. For a Riemann surface with genus $g$, $N_{\text{orbital}} = \frac{\mathcal{A}}{2\pi\ell_B^2} + 1 - g$. With metric: then $F = B\sqrt{g}dx \wedge dy$ with $B$ constant, and $H = -\frac{1}{2m}\Delta_B = \frac{1}{2m}\nabla^*\nabla$

The Laughlin model:
Everything in the LLL. First, consider $\nu = 1/2$ (bosonic): $H = \sum_{i\leq j} \Pi_{LLL}\delta^{(2)}(r_i - r_j)\Pi_{LLL}$.
Fact: zero energy states: $\psi(z_1, ..., z_N) = F(z_1, ..., z_N)\prod_{i<j}(z_i - z_j)^m e^{-\sum_i |z_i|^2/(4\ell_B^2)}$ for $\nu = 1/m$.
Confining potential: $V(x,y) = \frac{1}{2}K(x^2+y^2)$, then $\Pi_{LLL}(\sum_i V_i)\Pi_{LLL} = L_z$.
$\Psi_\lambda(z_1, ..., z_N) = P_\lambda(z_1, ..., z_N)\prod_{i<j}(z_i - z_j)^m e^{-\sum |z_i|^2/(4\ell_B^2)}$ (now we are doing any Laughlin state $\nu = 1/m$)
$\lambda = (\lambda_1, ..., \lambda_p)$, $\lambda_1 \geq \lambda_2 \cdots \leq 1$, $P_\lambda = (\sum_i z_i^{\lambda_1}) \cdots (\sum_i z_i^{\lambda_p})$,
$E_\lambda = E_0 + K\ell_B^2|\lambda|$, $\rho = \frac{1}{m}\frac{1}{2\pi\rho_B^2}$
Quasihole: $\psi_{w_1,...,w_p}(z_1, ..., z_N) = \prod_{i,j}(z_i - w_j)^{a_j}\prod_{i<j} z_{ij}^m$, where $a_j = 0, 1, ..., m-1$. electric charge of quasihole $w_i$ is $a_i/m$.
Laughlin on the torus:
$\psi_a(z_1, ..., z_N) = \vartheta[^{\frac{a}{m}}_0](m\frac{\sum z_i}{L_1}|m\tau) \times \prod_{i<j}\theta_1(\frac{z_{ij}}{L_1}|\tau)^m$.
CFT encoding of Laughlin states:
chiral: scalar field $\hat{\phi}(z) = \hat{\phi}_0 - i\hat{a}_0 \log z + i\sum_{m\neq 0}\frac{\hat{a}_n}{m}z^{-n}$, where $[a_n, a_m] = m\delta_{n+m,0}$, $a_{-n}$ for $n > 0$ is the creation operator,i.e. $a_n = a_{-n}^\dagger$. $[\phi_0, a_n] = i$. $a_n$'s describe a function of decoupled harmonic oscillators, and $\phi_0$ is the zero mode; $\phi_0$ is the "position" and $a_n$ is the "momentum". We have $a_0^\dagger = a_0$
Define $L_0 = \frac{a_0^2}{2} + \sum_{n>0} a_{-n}a_n$ is a Hamiltonian for $1+1$d CFT; it is the transfer matrix for 2D stat mech models)
Hilbert space: vacuum are all states of the form

$$a_n|0\rangle$$

for $n \geq 0$ s.t.

$$a_0|0\rangle = 0$$

(i.e. charge neutral).
$|\lambda; 0\rangle = a_{-\lambda_1} \cdots a_{-\lambda_p}|0\rangle$, where $\lambda$ labels the partition. "Grading of the Hilbert space":

$$L_0|\lambda, 0\rangle = |\lambda||\lambda, 0\rangle,$$

$$L_0|\lambda, \alpha\rangle = \left(\frac{\alpha^2}{2} + |\lambda|\right)|\lambda, \alpha\rangle.$$

G.S.: $|0\rangle$ with $h = 0$; $a_{-1}|0\rangle$ for $h = 1$, $a_{-1}^2|0\rangle$ and $a_{-2}|0\rangle$ for $h = 2$, and so on.
Charged sectors (excitations):

$$a_n|\alpha\rangle = 0, \quad n \geq 0$$

s.t.

$$a_0|\lambda, \alpha\rangle = \alpha|\lambda, \alpha\rangle.$$

with $\alpha \in \frac{1}{\sqrt{m}}\mathbb{Z}$.

Vertex operator $V_\alpha(z) =: e^{-i\alpha\phi(z)} :\equiv e^{\alpha\sum_{n\geq 0}\frac{a_{-n}}{n}z^n} e^{-\alpha\sum_{n\geq 0}\frac{a_n}{n}z^{-n}} e^{i\alpha\phi_0} z^{\alpha a_0}$, the "coherent state" operator.

Electron operator: $V(z) = V_{\sqrt{m}}(z)$.

$$\langle 0| \underbrace{e^{-iN\sqrt{m}\phi_0}}_{=O:\text{(the background charge)}} V(z_N)\cdots V(z_1)|0\rangle = \prod_{i<j}(z_{ij})^m,$$

$$\langle \lambda; 0|e^{-iN\sqrt{m}\phi_0} V(z_N)\cdots V(z_1)|0\rangle = P_\lambda \prod_{i<j}(z_{ij})^m.$$

define $V(z) = V_{\sqrt{m}}(z)$, then we have OPE:

$$: e^{i\alpha\varphi(z_1)} :: e^{i\beta\varphi(z_2)} = (z_1 - z_2)^{\alpha\beta} : e^{i\alpha\varphi(z_1)+i\beta\varphi(z_2)} :$$

Confining potential: $V(r) = \frac{K}{2}r^2$, eigenstates can be labeled by all the homogeneous polynomial: $P_\lambda(z_1, ..., z_N)\prod_{i<j}(z_{ij})^m$ having energy $k\ell_B^2\left(|\lambda| + m\frac{\Delta N^2}{2}\right)$. (i.e. excitation energy) $\Delta N$ is the extra number of electrons one puts on top of ground state, which increases the charge (measured by $\alpha$).

(Essentially: $[\Pi x\Pi, \Pi y\Pi] = i\ell_B^2$)

On cylinder: tracing over topological sector: define $F_\alpha$ to be all states with charge $\alpha$, which is the kernel: $\mathrm{Ker}(a_0 - \alpha)$.

recall that $\alpha \in \frac{1}{\sqrt{m}}\mathbb{Z}$ (compactification radius), and $Q := \frac{\alpha}{\sqrt{m}}$ measures the electric charge, and is the charge operator.

$\mathcal{H}_a$ is defined as all states with $Q = \frac{a}{m}$ mod 1.

On torus:

$$\mathrm{Tr}_{\mathcal{H}_a}(e^{i2\pi\tau L_0 - \sqrt{m}N\phi_0} V(z_N)\cdots V(z_1)),$$

$a_n$ modes of $J = i\partial\phi$.

Morally speaking: CFT conformal blocks = all the Laughlin states.

OPE: anyon (with $\alpha = a/\sqrt{m}$) fuse with electron (with $\beta = \sqrt{m}$) is

$$: e^{i\frac{a}{\sqrt{m}}\varphi(w)} :: e^{i\sqrt{m}\varphi(z)} := ...$$

$$\langle \cdots V_{\frac{a_2}{\sqrt{m}}}(w_2) V_{\frac{a_1}{\sqrt{m}}}(w_1) V(z_n)...V(z_1)|0\rangle = \prod z_{ij}^m \prod (z_i - w_j)^{a_j} \prod (w_i - w_j)^{\frac{a_i a_j}{m}}.$$

Matrix product state (Zaletel and Mong, '12)

Cylinder, parametrized by $\mathbb{R}x\frac{\mathbb{R}}{L\mathbb{Z}}$. Gauge field $A = B(0, x)$.

Define $\psi_k(x, y) = e^{ik_y} f_k(x)$, where $f \in \frac{2\pi}{L}\mathbb{Z}$,

LLL WFs: $\psi_k(x, y) = e^{ik_y} e^{-\frac{(x-\ell_B^2 k)^2}{2\ell_B^2}} = e^{ik(x+iy)} e^{\frac{k^2}{2}} e^{-x^2/2}$ (which is a holomorphic function (we do not care about $e^{-x^2/2}$).

particles on the cylinder occupy $\frac{2\pi\ell_B^2}{L} \times L$. The "fat cylinder" refers to when $L \gg \ell_B$.

Put confining potential: then

$\langle 0, \lambda|OV(z_1)\cdots V(z_N)|0, N\rangle$, where $\lambda$ labels modes which used to be generate without putting in the confining potential.

Write $V(z) = \sum_k V_{-k} e^{kz}$,

Then $\langle 0, \lambda|OV(z_1)\cdots V(z_N)|0, N\rangle = \sum_{k_1,...,k_N} \langle 0, \lambda|OV_{-k_1}\cdots V_{-k_N}|0, k\rangle \underbrace{A\left(\prod_j e^{k_j z_j}\right)}_{\text{Slater determinant}}$,

We have $A^{[0]}[k] = 1$, $A^{[1]}[k] = e^{\frac{1}{2}k^2} V_{-k}$, ..., $A^{[m]}[k] = \frac{1}{\sqrt{m!}}(e^{\frac{1}{2}k^2} V_{-k})^m$.

$k = \frac{2\pi}{L}j$. Fact: $e^{\frac{1}{2}k^2} V_{-k} = U^{-j} V_0 U^j$, where $U = e^{-\left(\frac{2\pi}{L}\right)^2 L_0} e^{-i\frac{1}{\sqrt{m}}\varphi_0}$.

Lecture: generalized symmetries and Dualities by Lootens

(Let's say symmetry here only mean global internal symmetry)

TQFT (Turaev-Viro invariants in $2+1$D ) in the bulk, with one gapped boundary having symmetry $G$, while the other boundary a trivial theory (call it $T$).

other names for the same thing: "SYMTFT". "Topological Holography", "Sandwich construction", "Strange correlators", "Shadow world" (that goes back to Moore and Reshetikhin)

In 3D: TQFT = Turaev-Viro state sum

$\langle \Psi(\mathcal{M})|\Omega\rangle Z = \mathrm{Tr}(e^{-\beta H})$ in $2+0$D or $\mathrm{Tr}(e^{-i\tau H})$ in $1+1$D.

First, let's talk about $\langle\Psi(\mathcal{M})|$. Toric code in $2+1$D:

Symmetries: closed loops of $\prod Z$ (or $\prod X$) on the primal (dual) lattice. ("1-form" symmetry: acting on a codimension-1 manifold)

4GS's on the torus: $|0\rangle$, $\prod_{horizontal} Z|0\rangle$, $\prod_{vertical} Z|0\rangle$, and $\prod_{horizontal} Z \prod_{vertical} Z|0\rangle$.

define the GHZ tensors $T_{ijkl} = \delta_{i=j=k=l=0 \text{ or } 1}$, and $P_{ijkl} = \sum_{i'j'k'l'} (H \otimes H \otimes H \otimes H)_{ii',jj',kk',ll'}$, where $H$ is the $2 \times 2$ Hadamard matrix, with $HX = ZH$.

1) define reference state $|0^{\otimes N}\rangle$, then $2P_v = (1 + \prod_v XXXX) = \prod_v P'P'P'P'$ where $P'$ is the Hadamard-type tensor with only three legs ]2 0 Reference state $|+^{\otimes N}\rangle$: $2P_p = 1 + \prod_p ZZZZ = \prod_v T'T'T'T'$, where $T'$ is written as ....

Ising: $Z_1 = \langle TC_1|\Omega\rangle$, where $|\Omega\rangle = |\beta\rangle^{\otimes N}$, $|\beta\rangle = (e^{\beta}, e^{-\beta})^T$. This gives the partition function of the classicla Ising model at inverse temperature $\beta$.

$Z_2 = \langle TC_2|\Omega\rangle$

Vincent Pasquier:

$Q$-state Potts model: $Z = \sum_c e^{\sum_{i\sim j} \beta\delta_{\sigma_i=\sigma_j}}$. When $\beta\infty\infty$: $Q$ degeneracy ($Z = Q$). $\beta\infty 0$: each lattice site is independent, completely disordered state.

Second way to compute $Z$: Cluster picture: all the spins on the cluster have equal spins.

$Z = \sum_{Cluster} v^{\#\text{ of links}} Q^{\#\text{ of Clusters}}$, where $v = e^{\beta} - 1$.

Third way to compute $Z$: "dual picture", the "media lattice", map to the six-vertex model: assign orientation of the blue lines, so that when summing over two orientations labeld by $q$ and $q^{-1}$, wth $q + q^{-1} = \sqrt{Q}$.

Six possible configs for the blue lines: out of all possible eight ways of labeling the arrows for a 4-leg vertex, we only keep the 2-in 2-out configurations, with weights $1, xq, 1 + xq, 1, xq^{-1}, 1 - xq$, redefine as $1, a, b, b', c, c'$ (?) with $\frac{a^2+b^2-c^2}{2\sqrt{ab}} = Q = q + q^{-1}$.

Another way is to introduce transfer matrix:

$Z = \text{tr}[(T_1T_2)^{\#\text{ rows}}]$.

define $e_{2i} = \sqrt{Q}\delta_{\sigma_i=\sigma_{i+1}}$.

Define $e_{2i+1} = \frac{1}{\sqrt{Q}} A$, where $A$ is a $Q$-by-$Q$ matrix with 1 everywhere.

$T_1 = (1 + xe_1)(1 + xe_3)\cdots$, $T_2 = (1 + ye_2)(1 + ye_4)\cdots$, where vertical weight is $e^{\beta} = 1 + \sqrt{Q}/x$, and horizontal weight $e^{\beta} = 1 + y\sqrt{Q}$. We must have $xy = 1$ in order to make the model consistent.

The point of introducing $T_{1,2}$ is that they satisfy the TL algebra: $e_i^2 = Qe_i$, $e_ie_{i+1}e_i = e_i$, $e_{i+1}e_ie_{i+1} = e_{i+1}$.

(Best rep of TL algebra is loop model)

Key results: $[T(z), T(w)] = 0$. Lax operator $L$; $R : (z)L(w) = L(w)L(z)R$.

# 18 Loop models

Basic concepts and characterization:

- Loop number $|\mathcal{C}|$: $e^{-\alpha|\mathcal{C}|} = (e^{-\alpha})^{|\mathcal{C}|}$, with $e^{-\alpha} \equiv n$ called number fugacity;
- Loop length $\ell$: $e^{-\beta x} = (e^{-\beta})^x$ with $e^{-\beta} \equiv x$ called loop length fugacity;
- Is crossing

`https://arxiv.org/abs/0806.3484` Temperley–Lieb algebra, the Potts model, the Jones polynomial, and SU(2) Chern–Simons gauge theory

Various presentations of the TL algebra can be used to define lattice statistical–mechanical models.

# 19 Modern theory of invertible phases

## 19.1 Kitaev's proposal of gapped invertible phases

Gapped invertible systems form an $\Omega$ spectrum in homotopy theory.

Follow `https://journals.aps.org/prb/pdf/10.1103/PhysRevB.108.125147`:

Assumption 1: the existence of a classifying space of invertible gapped phases in $d$ spatial dimensions, called $E_d$ (a topological space). [Here classifying space refers to the proposal that the path components correspond to invertible phases of systems. (System: one parameter point; phase: phases of matter.)]

Example: $E_0$: a $0D$ gapped bosonic system without any symmetry over $X$ (the parameter space): the ground state assemble into a line bundle over $X$, whose first Chern class in $H^2(X, \mathbb{Z})$ is a complete phase invariant. So the space $E_0$ can be chosen to be the infinite complex projective space (i.e. a $K(\mathbb{Z}, 2)$).

Assumption 2: $E_d$ forms an $\Omega$ spectrum. That is to say that $E_d \cong \Omega E_{d+1}$.

Generalized cohomology theory of gapped invertible phases: the grounds $E^d(X)$ arise as homotopy classes of maps, in particular: $E^d(X) = [X, E_d]$, the (set of) homotopy classes of continuous maps from $X$ to $E_d$.

# 20 Doodles for talks

$$\overline{S_n(N)} = \frac{1}{1-n} \lim_{k\to 0} \frac{d}{dk} \overline{(\mathrm{Tr}(\check{\rho}^n))^k (\mathrm{Tr}(\check{\rho}))^{N-nk}} \tag{87}$$

$$\mathcal{O} = S_n = \frac{1}{1-n} \log \mathrm{tr}[\rho_t^n]$$

$$\langle \mathcal{O} \rangle_{\text{Circuit}} = \sum_{\text{outcomes}} \overline{(\mathrm{tr}\check{\rho}_t)\mathcal{O}} = (2^{m_{\text{measured}}})^t \cdot \overline{(\mathrm{tr}\check{\rho}_t)\mathcal{O}}$$

$$\langle e^{-k(n-1)S_n} \rangle_{\text{Circuit}} = (2^{m_{\text{measured}}})^t \cdot \overline{(\mathrm{tr}\check{\rho})^{1-nk} (\mathrm{tr}\check{\rho}^n)^k}$$

$$\overline{S_1}$$

$$T_{\sigma,\rho} = q^{N-|\sigma^{-1}\rho|}$$

$$\overline{S_n} \equiv \frac{1}{1-n} \overline{\ln \mathrm{tr}(\rho_t^n)} = f(x)$$

$$= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

$$A_{\text{LH}} \subseteq \ker[\tilde{i} \colon \mathcal{H}^3(G_{\text{space}}, \mathbb{Z}_2) \to \mathcal{H}^3(G_{\text{space}}, U(1)_\rho)]$$

$$\langle S_1 \rangle = -\ln x + 1 - \gamma + \frac{11}{24}x^2 - \frac{1739}{2880}x^4 + \frac{329489}{181440}x^6 - \frac{83530439}{9676800}x^8 + \cdots$$

$$\langle S_1 \rangle = -\ln x + 1 - \gamma + \frac{5}{24}x^2 - \frac{239}{2880}x^4 + \frac{3679}{36288}x^6 - \frac{2423279}{9676800}x^8 + \cdots$$

$$A_{\text{LH}} \cong \ker[\tilde{i} \colon \mathcal{H}^2(G_{\text{space}}, \mathbb{Z}_2) \to \mathcal{H}^2(G_{\text{space}}, U(1)_\rho)]$$

$$Z_1 = e^{i\pi \int_{\mathcal{M}_5} A_i^2 \wedge (A_i + A_m) \wedge \omega}$$

$$Z_2 = e^{i\pi \int_{\mathcal{M}_5} B_{xy} \wedge A_i \wedge \omega}$$

$$w_i \colon$$

$$Sq^i \colon$$

$$\mathcal{SQ}^i \colon$$

$$\mathcal{H}^{d+1}(G_1 \times G_2, U(1)) = \sum_{i=0}^{d} \mathcal{H}^{d-i}(G_2, \mathcal{H}^{i+1}(G_1, U(1))) = \sum_{i=0}^{d+1} (G_1 \leftrightarrow G_2)$$

$$\mathcal{H}^{d+1}(G_{\text{int}}, U(1))$$

$$H_G^{d+1}(EG, U(1)^\rho) \cong \mathcal{H}^{d+1}(G, U(1)^\rho).$$

$$\overline{e^{-k(n-1)S_n}} = \overline{(\mathrm{Tr}\rho^n)^k \, (\mathrm{Tr}\rho)^{N-nk}}$$

$$\left(T^{t+1}\right)_{1^{N-nk}n^k,1^N} \propto \sum_{\ell=0}^{\infty} H_{N,n,k,\ell} \frac{x^\ell}{\ell!}$$

$$\langle 1^{N-nk}n^k|T^{t+1}|1^N\rangle \propto \sum_{\ell=0}^{\infty} H_{N,n,k,\ell} \frac{x^\ell}{\ell!}$$

$$= \frac{1}{N!}\sum_{\lambda\vdash N} d^\lambda \chi^\lambda(\sigma) e^{\frac{x}{2}\left(\sum_j \lambda_j^2 - (\lambda_j^\top)^2\right)}$$

$$\sum_\ell H_{\sigma,\ell}\frac{x^{2\ell}}{(2\ell)!}$$

$$H_{\sigma,\ell}$$

$$S_g = \left|\frac{\mathrm{Hom}(\pi_1(M_g)\to G)}{G}\right|$$

$$\prod_{i=1}^{g}[a_i,b_i] = e$$

$$a_i, b_i \in G$$

$$I(t) = \int_{G^n} H(t, f(h), e)dh$$

$$H(t,x,y) = \frac{1}{|G|}\sum_{\lambda\in \mathrm{Irr}(G)} d_\lambda \chi_\lambda(xy^{-1})e^{-tp(\lambda)}$$

$$(a_1,b_1,...,a_g,b_g) \sim (ga_1g^{-1}, gb_1g^{-1}, ..., ga_gg^{-1}, gb_gg^{-1})$$

$$H_{N,k,n,1} = \frac{(n+1)(N-1)N}{4} + \left(N - \frac{(n+1)(5n+6)}{12}\right)\frac{n(n-1)k}{2} + \frac{n^2(n-1)^2k^2}{4}$$

$$Cl((12))$$

$$S_{g,n} = \frac{|G|^{2g+n-1}}{\prod_{j=1}^n |Cl(c_j)|}\sum_{\lambda\in\mathrm{Irr}(G)}\frac{\prod_{j=1}^n \chi_\lambda(c_j)}{d_\lambda^{2g+n-2}}$$

$$S_g = \begin{cases} (2^{2g}-1)\frac{n^{2g-2}+1}{2} + \frac{n^{2g}+1}{2}, & n \text{ odd;} \\ 2(2^{2g}-1)(n^{2g-2}+2^{2g-2}) + \frac{n^{2g}+2^{2g}}{2}, & n \text{ even.} \end{cases}$$

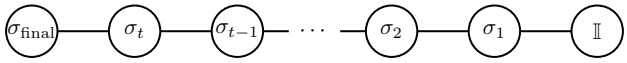
$$\prod_{i=1}^{g}[x_i,y_i]\prod_{i=1}^{n} z_i = e$$

$$x_i, y_i \in G$$

$$z_i \in Cl(c_i)$$

$$\tau_1\tau_2\cdots\tau_\ell = \sigma$$

$$\tau_i \in \mathrm{Sym}(N), \quad |\tau_i| = 1$$



Protocol I

Protocol II

$$x := \frac{t}{q}$$

$$\frac{1}{q} = \frac{1}{q_{\text{un.}}} - \frac{1}{q_{\text{tot}}}$$

$$\overline{S_n} = -\ln x + \text{const.} + o(x^2)$$

$$\begin{aligned} \overline{S_1} &= -\ln x + 1 - \gamma + \frac{11}{24}x^2 - \frac{1739}{2880}x^4 \cdots \\ \overline{S_1} &= -\ln x + 1 - \gamma + \frac{5}{24}x^2 - \frac{239}{2880}x^4 + \cdots \end{aligned} \tag{88}$$

$$\rho_0 \sim \mathbb{I}$$

$$\rho_{t-1} \to \rho_t \sim M_t^\dagger \rho_{t-1} M_t,$$

$$\rho_t = M_t^\dagger M_t,$$

$$\sigma \in \text{Sym}(N)$$

$$M_t$$

$$m_i :$$

$$\overline{e^{-k(n-1)S_n}} \propto \lim_{N \to 0 \text{ or } 1} (T^{t+1})_{1^N,\sigma},$$

$$\sigma = 1^{N-nk} n^k \in \text{Sym}(N)$$

$$H = \sum_{\langle ij \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j$$

$$H^{d+2}(G_{\text{spatial}} \times SO(3), U(1)^{\text{or}}) = H^d(G_{\text{spatial}}, H^2(SO(3), U(1))) \times \cdots$$

$$\lambda \in H^2(G_{\text{spatial}}, \mathbb{Z}_2), \quad \eta \in H^2(SO(3), U(1)) = \mathbb{Z}_2,$$

$$\text{No. } 227\colon Fd\bar{3}m = \langle T_{1,2,3}, C_{2z}, C_{2x}, C_3, M_{xy}, P \rangle$$
$$\text{No. } 216\colon F\bar{4}3m = \langle T_{1,2,3}, C_{2z}, C_{2x}, C_3, M_{xy} \rangle$$
$$\text{No. } 196\colon F23 = \langle T_{1,2,3}, C_{2z}, C_{2x}, C_3 \rangle$$
$$\text{No. } 22\colon F222 = \langle T_{1,2,3}, C_{2z}, C_{2x} \rangle$$
$$\text{No. } 5\colon C2 = \langle T_{1,2,3}, C_{2z} \rangle$$

$$C2 \subset F222 \subset F23 \subset F\bar{4}3m \subset Fd\bar{3}m$$
$$C2 \subset P3_221 \subset R32 \subset R\bar{3}m \subset Fd\bar{3}m$$

From Mike Zaletel, TopoQuantum16Conf:

$$[\Lambda]_{227,\mathbb{Z}_2} \in \mathbb{Z}_2^4$$

$$[\Lambda]_{216,\mathbb{Z}_2} \in \mathbb{Z}_2^4$$
$$[\Lambda]_{196,\mathbb{Z}_2} \in \mathbb{Z}_2^4$$
$$[\Lambda]_{22,\mathbb{Z}_2} \in \mathbb{Z}_2^4$$
$$[\Lambda]_{5,\mathbb{Z}_2} \in \mathbb{Z}_2^2$$

Main results:
1. Topological partition function corresponding to the LSM anomalies:

$$A_{Fd\bar{3}m} = \langle \beta\iota, \iota^3 + \tau, \iota^2(\sigma+\iota), \sigma^2(\sigma+\iota), \psi \rangle, \tag{89a}$$
$$A_{F\bar{4}3m} = \langle \gamma, \tau, \omega, \sigma^3, \psi \rangle, \tag{89b}$$
$$A_{F23} = \langle \gamma, \tau, \omega, \psi \rangle, \tag{89c}$$
$$A_{F222} = \langle \gamma, \tau, \chi_{12}\rho(\rho+\rho'), \rho\rho'(\rho+\rho') \rangle, \tag{89d}$$
$$A_{C2} = \langle \gamma, \beta'\rho \rangle. \tag{89e}$$

2. Topological invariants:
3. filling constraints:

$$n_{216} = 2\mathbb{Z}, \qquad n_{227} = 4\mathbb{Z}. \tag{90}$$

Application:
Symmetry breaking $G_{227} \to G_{216}$.
Problems:
1. In 2D: lattice homotopy matches exactly with the SPT argument. In 3D it seems not (cases of 196-230):
2.

| Wallpaper group | $\mathcal{H}^2(G, \mathbb{Z}^{\mathrm{or}})$ |
|---|---|
| 1,4,5 | $\mathbb{Z}$ |
| 2,6 | $\mathbb{Z} \times \mathbb{Z}_2^3$ |
| 3,8 | $\mathbb{Z} \times \mathbb{Z}_2$ |
| 7,9 | $\mathbb{Z} \times \mathbb{Z}_2^2$ |
| 10,11 | $\mathbb{Z} \times \mathbb{Z}_2 \times \mathbb{Z}_4$ |
| 12 | $\mathbb{Z} \times \mathbb{Z}_4$ |
| 13,14 | $\mathbb{Z} \times \mathbb{Z}_3^2$ |
| 15 | $\mathbb{Z} \times \mathbb{Z}_3$ |
| 16,17 | $\mathbb{Z} \times \mathbb{Z}_2 \times \mathbb{Z}_3$ |

$$\sigma^3: \quad \Omega(M,M,M) = -1, \tag{91a}$$
$$\iota^3: \quad \Omega(P,P,P) = -1, \tag{91b}$$
$$\sigma^3, \sigma^2\iota, \sigma\iota^2, \iota^3: \quad \Omega(MP,MP,MP) = -1, \tag{91c}$$
$$\sigma^3, \sigma\iota^2: \quad \Omega(M,MP,MP)\Omega(MP,M,MP)\Omega(MP,MP,M) = -1, \tag{91d}$$
$$\sigma\delta: \quad \Omega(M,C_2,C_2)\Omega(C_2,M,C_2)\Omega(C_2,C_2,M) = -1, \tag{91e}$$
$$\psi: \quad \Omega(C_2,C_2',C_2')\Omega(C_2',C_2,C_2')\Omega(C_2',C_2',C_2) = -1, \tag{91f}$$
$$\iota^3, \beta\iota: \quad \Omega(T_1T_2T_3P, T_1T_2T_3P, T_1T_2T_3P) = -1, \tag{91g}$$
$$\begin{aligned}\beta\sigma: \quad & \Omega(M,T_3,T_1)\Omega(M,T_1T_2^{-1}T_3,T_1)\Omega(M,T_1,T_3)\Omega(M,T_1,T_1T_2^{-1}T_3)\cdot \\ & \Omega(T_3,T_3M,T_1)\Omega(T_3,T_1,T_3M)\Omega(T_3,T_1M,T_3)\Omega(T_3,T_1M,T_1T_2^{-1}T_3)\cdot \\ & \Omega(T_1T_2^{-1}T_3, T_1^{-1}T_2T_3M, T_1)\Omega(T_1T_2^{-1}T_3, T_2M, T_3)\cdot \\ & \Omega(T_1T_2^{-1}T_3, T_2M, T_3)\Omega(T_1T_2^{-1}T_3, T_1, T_1^{-1}T_2T_3M)\cdot \\ & \Omega(T_1,M,T_3)\Omega(T_1,M,T_1T_2^{-1}T_3)\Omega(T_1,T_3,T_3M)\Omega(T_1,T_1T_2^{-1}T_3,T_1^{-1}T_2T_3M) = -1,\end{aligned} \tag{91h}$$
$$\begin{aligned}\tau: \quad & \Omega(T_1T_2T_3^{-1}, T_1, T_1^{-1}T_2^{-1}C_2'P)\Omega(T_1, T_1T_2T_3^{-1}, T_1^{-1}T_2^{-1}C_2'P)\cdot \\ & \Omega(T_1T_2T_3^{-1}, T_2^{-1}C_2'P, T_1)\Omega(T_2^{-1}C_2'P, T_1T_2T_3^{-1}, T_1)\cdot \\ & \Omega(T_2^{-1}C_2'P, T_1, T_1T_2T_3^{-1})\Omega(T_1T_2T_3^{-1}, T_1^{-1}T_2^{-1}C_2'P, T_1T_2T_3^{-1})\cdot \\ & \Omega(T_1, T_1^{-1}T_2^{-1}C_2'P, T_1T_2T_3^{-1}) = -1,\end{aligned} \tag{91i}$$

$$\lambda(P,P,P) = -1,$$
$$\lambda(T_1T_2T_3P, T_1T_2T_3P, T_1T_2T_3P) = -1$$
$$\lambda(C_2, C_2', C_2')\lambda(C_2', C_2, C_2')\lambda(C_2', C_2', C_2) = -1$$
$$\lambda(\widetilde{C}_2, \widetilde{C}_2', \widetilde{C}_2')\lambda(\widetilde{C}_2', \widetilde{C}_2, \widetilde{C}_2')\lambda(\widetilde{C}_2', \widetilde{C}_2', \widetilde{C}_2) = -1$$
$$\widetilde{C}_2 = T_3C_2, \quad \widetilde{C}_2' = T_2C_2'$$

$$C_2 \to T_3C_2, C_2' \to T_2C_2'$$
$$C_2 \to T_3^2C_2, C_2' \to T_2^2C_2'$$
$$C_2 \to T_3^3C_2, C_2' \to T_2^3C_2'$$

$$Z_{\mathrm{UV}} = e^{i\pi \int_{\mathcal{M}_5} \lambda[A_{\mathrm{spatial}}] \cup \omega[A_{\mathrm{spin}}]}$$
$$\in H^3(G_{\mathrm{spatial}}, H^2(SO(3), U(1))),$$

$$\lambda \in H^3(G_{\mathrm{spatial}}, \mathbb{Z}_2), \quad \omega \in H^2(SO(3), U(1))) = \mathbb{Z}_2,$$

$$A_{\mathrm{LH}} \cong \ker[\tilde{i}\colon H^3(G, \mathbb{Z}_2) \to H^3(G, U(1)^{\mathrm{or}})]$$

$$A_{\mathrm{LH}} \cong H^3(G, \mathbb{Z}^{\mathrm{or}}) \otimes \mathbb{Z}_2$$

$$\lambda(g_1, g_2, g_3)\colon G \times G \times G \to \mathbb{Z}_2$$

$\lambda \in$

| Wyckoff | Little group | | Coordinates | LSM anomaly class $\lambda$ | | | | Topological invariant |
|---|---|---|---|---|---|---|---|---|
| | Intl. | Schönflies | | $A_i^2(A_m+A_i)+A_iB$ | $A_iB$ | $C_{n\gamma}+C_{s\psi}$ | $C_{n\gamma}$ | |
| 16d | $\bar{3}$m | $D_{3d}$ | $(1/2,1/2,1/2)$ | 1 | 0 | 0 | 0 | $\phi_1[\lambda] = \lambda(I,I,I)$ |
| 16d | $\bar{3}$m | $D_{3d}$ | $(0,0,0)$ | 0 | 1 | 0 | 0 | $\phi_2[\lambda] = \lambda(T_1T_2T_3I, T_1T_2T_3I, T_1T_2T_3I)$ |
| 8b | $\bar{4}$3m | $T_d$ | $(3/8,3/8,3/8)$ | 0 | 0 | 1 | 0 | $\phi_3[\lambda] = \lambda(C_2', C_2, C_2)$ |
| 8a | $\bar{4}$3m | $T_d$ | $(1/8,1/8,1/8)$ | 0 | 0 | 0 | 1 | $\phi_4[\lambda] = \prod_{\mathrm{cyc}} \lambda(T_2C_2', T_3C_2, T_3C_2)$ |

$$[\lambda] \in H^3(G_{\text{space group}}, H^2(SO(3), U(1)))$$

$$\lambda \in H^3(G_{\text{space group}}, \mathbb{Z}_2)$$

$$\lambda(g_1, g_2, g_3) \in H^2(SO(3), U(1)) = \mathbb{Z}_2 = \{+, -\}$$

$$\lambda(g_1, g_2, g_3) = \pm \in \mathbb{Z}_2$$

$$\tilde{i} = T_1T_2T_3\, i$$
$$\widetilde{C}_2 = T_3C_2$$
$$\widetilde{C}_2' = T_2C_2'$$

$$|\uparrow\rangle$$
$$|\downarrow\rangle$$

$$|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$$

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$$

$$S = \frac{1}{2}$$

$$\text{“}-\text{”} \in \mathbb{Z}_2$$

$$\text{“}+\text{”} \in \mathbb{Z}_2$$

$$\begin{gathered} \mathbb{Z}_2^4 \\ \mathbb{Z}_2^4 \\ \mathbb{Z}_2^4 \\ \mathbb{Z}_2^4 \\ \mathbb{Z}_2^2 \end{gathered} \tag{92}$$

$$\begin{aligned} &\lambda \in \langle B_{xy} A_i, A_i^2 (A_i + A_\sigma), C_\tau, C_\psi \rangle \\ &\lambda \in \langle A_x A_y A_z, C_\omega, C_\tau, C_\psi \rangle \\ &\lambda \in \langle A_x A_y A_z, C_\omega, C_\tau, C_\psi \rangle \\ &\lambda \in \langle A_x A_y A_z, C_\omega, (A_x + A_y) A_\rho (A_\rho + A_{A_{\rho'}}), A_\rho A_{A_{\rho'}} (A_\rho + A_{A_{\rho'}}) \rangle \\ &\lambda \in \langle A_x A_y A_z, B_{xy} A_\rho \rangle \end{aligned} \tag{93}$$

$$\begin{aligned} \mathbb{Z}_2^5 &\cong \langle \text{LSM's}, A_\sigma^2 (A_i + A_\sigma) \rangle \\ \mathbb{Z}_2^5 &\cong \langle \text{LSM's}, A_\sigma^3 \rangle \\ \mathbb{Z}_2^4 &\cong \langle \text{LSM's} \rangle \\ \mathbb{Z}_2^4 &\cong \langle \text{LSM's} \rangle \\ \mathbb{Z}_2^2 &\cong \langle \text{LSM's} \rangle \end{aligned} \tag{94}$$

$$\begin{aligned} \text{No. 227:} \quad & F d \bar{3} m \\ \text{No. 216:} \quad & F \bar{4} 3 m \\ \text{No. 196:} \quad & F 23 \\ \text{No. 22:} \quad & F 222 \\ \text{No. 5:} \quad & C2 \end{aligned} \tag{95}$$

| Wallpaper group | $\mathcal{H}^2(G, \mathbb{Z}^{\mathrm{or}})$ |
|---|---|
| 1,4,5 | $\mathbb{Z}$ |
| 2,6 | $\mathbb{Z} \times \mathbb{Z}_2^3$ |
| 3,8 | $\mathbb{Z} \times \mathbb{Z}_2$ |
| 7,9 | $\mathbb{Z} \times \mathbb{Z}_2^2$ |
| 10,11 | $\mathbb{Z} \times \mathbb{Z}_2 \times \mathbb{Z}_4$ |
| 12 | $\mathbb{Z} \times \mathbb{Z}_4$ |
| 13,14 | $\mathbb{Z} \times \mathbb{Z}_3^2$ |
| 15 | $\mathbb{Z} \times \mathbb{Z}_3$ |
| 16,17 | $\mathbb{Z} \times \mathbb{Z}_2 \times \mathbb{Z}_3$ |

$$\cong$$

$\mathcal{H}^5(G_{\mathrm{UV}}, U(1)) = \mathcal{H}^6(G_{\mathrm{UV}}, \mathbb{Z}) \leftarrow \mathcal{H}^3(G_{\mathrm{UV}}, \mathbb{Z}) \cup \mathcal{H}^3(G_{\mathrm{UV}}, \mathbb{Z}) = \mathcal{H}^2(G_{\mathrm{UV}}, U(1)) \cup \mathcal{H}^2(G_{\mathrm{UV}}, U(1))$
$\mathcal{H}^5(G_{\mathrm{UV}}, U(1)) \leftarrow \mathcal{H}^5(G_{\mathrm{IR}}, U(1))$
$\mathcal{H}^5(G_{\mathrm{UV}}, U(1)) \leftarrow \mathcal{H}^2(G_{\mathrm{UV}}, \mathbb{Z}_2) \cup \mathcal{H}^3(G_{\mathrm{UV}}, \mathbb{Z}_3)$

$$SQ^1 \lambda = 0, \quad SQ^2 \lambda = 0$$

$$\begin{aligned} SQ^1 &= (Sq^1 + w_1 \cup), \\ SQ^2 &= (Sq^2 + w_1 \cup Sq^1 + \widetilde{w}_2 \cup), \end{aligned} \tag{96}$$

Table 5: Point group LSM.

| Schönflies | Abstract | Mod-2 Coh. Ring | $H^3(PG, Z^{\mathrm{or}}) \otimes \mathbb{Z}_2$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|---|---|---|
| $C_1$ | Trivial | Trivial | Trivial | 0 | 0 | 0 |
| $C_i$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2[A_\iota]$ | $\langle A_\iota^3 \rangle$ | $A_\iota$ | $A_\iota^2$ | $A_\iota^3$ |
| $C_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2[A_\rho]$ | Trivial | 0 | 0 | 0 |
| $C_s$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2[A_{\sigma'}]$ | $\langle A_{\sigma'}^3 \rangle$ | $A_{\sigma'}$ | 0 | 0 |
| | | $\mathbb{Z}_2[A_\rho, A_{\sigma'}]$ | $\langle A_\rho^2 A_{\sigma'}, A_{\sigma'}^3 \rangle$ | $A_{\sigma'}$ | $A_\rho^2$ | $A_\rho^2 A_{\sigma'}$ |
| $C_{2h}$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_2[A_{\sigma'}, \iota]$ | $\langle A_\iota^2(A_\iota + A_{\sigma'}), A_\sigma^2(A_\iota + A_{\sigma'}) \rangle$ | $A_\iota + A_{\sigma'}$ | $A_\iota^2$ | $A_\iota^2(A_\iota + A_{\sigma'})$ |
| | | $\mathbb{Z}_2[A_\rho, A_\iota]$ | $\langle A_\iota A_\rho^2, A_\iota^3 \rangle$ | $A_\iota$ | $A_\rho^2 + A_\iota^2$ | $A_\iota(A_\rho^2 + A_\iota^2)$ |
| $D_2$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_2[A_\rho, A_{\rho'}]$ | $\langle A_\rho A_{\rho'}(A_\rho + A_{\rho'}) \rangle$ | 0 | $A_\rho^2 + A_\rho A_{\rho'} + A_{\rho'}^2$ | $A_\rho A_{\rho'}(A_\rho + A_{\rho'})$ |
| $C_{2v}$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_2[A_\rho, A_\sigma]$ | $\langle A_\rho^2 A_\sigma, A_\sigma^3 \rangle$ | $\sigma$ | $A_\rho(A_\rho + A_\sigma)$ | 0 |
| $D_{2h}$ | $\mathbb{Z}_2^3$ | $\mathbb{Z}_2[A_\rho, A_{\rho'}, A_{\sigma'}]$ | $\langle A_\rho A_{\rho'}(A_\rho + A_{\rho'} + A_{\sigma'}), A_\rho^2 A_{\sigma'}, A_{\rho'}^2 A_{\sigma'}, A_{\sigma'}^3 \rangle$ | $A_{\sigma'}$ | $A_\rho^2 + A_\rho A_{\rho'} + A_{\rho'}^2 + A_{\rho'} A_{\sigma'}$ | $A_\rho(A_\rho + A_{\rho'})(A_{\rho'} + A_{\sigma'})$ |
| $C_4$ | $\mathbb{Z}_4$ | $\mathbb{Z}_2[A_{\bar{\rho}}, B_\delta]/(A_{\bar{\rho}}^2)$ | Trivial | 0 | 0 | 0 |
| $S_4$ | $\mathbb{Z}_4$ | $\mathbb{Z}_2[A_{\bar{\iota}}, B_\delta]/(A_{\bar{\iota}}^2)$ | $\langle A_{\bar{\iota}} B_\delta \rangle$ | $A_{\bar{\iota}}$ | $B_\delta$ | $A_{\bar{\iota}} B_\delta$ |
| $C_{4h}$ | $\mathbb{Z}_4 \times \mathbb{Z}_2$ | $\mathbb{Z}_2[A_{\bar{\rho}}, A_{\sigma'}, B_\delta]/(A_{\bar{\rho}}^2)$ | $\langle B_\delta A_{\sigma'}, A_{\sigma'}^3 \rangle$ | $A_{\sigma'}$ | $B_\delta$ | $A_{\sigma'} B_\delta$ |
| $D_4$ | $Dih_4$ | $\mathbb{Z}_2[A_{\bar{\rho}}, A_{\rho'}, B_\delta]/(A_{\bar{\rho}}(A_{\bar{\rho}} + A_{\rho'}))$ | $\langle A_{\bar{\rho}}^3 + B_\delta A_{\rho'} \rangle$ | 0 | $A_{\rho'}^2 + B_\delta$ | $A_{\rho'} B_\delta$ |
| $C_{4v}$ | $Dih_4$ | $\mathbb{Z}_2[A_{\bar{\rho}}, A_\sigma, B_\delta]/(A_{\bar{\rho}}(A_{\bar{\rho}} + A_\sigma)$ | $\langle A_{\bar{\rho}}^3, A_\sigma^3 \rangle$ | $A_\sigma$ | $B_\delta$ | 0 |
| $D_{2d}$ | $Dih_4$ | $\mathbb{Z}_2[A_{\bar{\iota}}, A_{\rho'}, B_\delta]/(A_{\bar{\iota}}(A_{\bar{\iota}} + A_{\rho'}))$ | $\langle A_{\bar{\iota}}^3, B_\delta(A_{\bar{\iota}} + A_{\rho'}) \rangle$ | $A_{\bar{\iota}}$ | $(A_{\rho'} + A_{\bar{\iota}})A_{\rho'} + B_\delta$ | $(A_{\rho'} + A_{\bar{\iota}})B_\delta$ |
| $D_{4h}$ | $Dih_4 \times \mathbb{Z}_2$ | $\mathbb{Z}_2[A_{\bar{\rho}}, A_{\rho'}, B_\delta, A_{\sigma'}]/(A_{\bar{\rho}}(A_{\bar{\rho}} + A_{\rho'}))$ | $\langle A_{\bar{\rho}}^3 + (A_{\rho'} + A_{\sigma'})B_\delta, A_{\bar{\rho}}^2 A_{\sigma'}, A_{\rho'}^2 A_{\sigma'}, A_{\sigma'}^3 \rangle$ | $A_{\sigma'}$ | $A_{\rho'}^2 + B_\delta$ | $(A_{\sigma'} + A_{\rho'})B_\delta$ |
| $C_3$ | $\mathbb{Z}_3$ | *id.* $C_1$ | *id.* $C_1$ | *id.* $C_1$ | *id.* $C_1$ | *id.* $C_1$ |
| $S_6$ | $\mathbb{Z}_3 \times \mathbb{Z}_2$ | *id.* $C_i$ | *id.* $C_i$ | *id.* $C_i$ | *id.* $C_i$ | *id.* $C_i$ |
| $D_3$ | $Dih_3$ | *id.* $C_2'$ | *id.* $C_2'$ | *id.* $C_2'$ | *id.* $C_2'$ | *id.* $C_2'$ |
| $C_{3v}$ | $Dih_3$ | *id.* $C_s'$ | *id.* $C_s'$ | *id.* $C_s'$ | *id.* $C_s'$ | *id.* $C_s'$ |
| $D_{3d}$ | $Dih_3 \times \mathbb{Z}_2$ | *id.* $C_{2h}'$ | *id.* $C_{2h}'$ | *id.* $C_{2h}'$ | *id.* $C_{2h}'$ | *id.* $C_{2h}'$ |
| $C_6$ | $\mathbb{Z}_3 \times \mathbb{Z}_2$ | *id.* $C_3$ | *id.* $C_3$ | *id.* $C_3$ | *id.* $C_3$ | *id.* $C_3$ |
| $C_{3h}$ | $\mathbb{Z}_3 \times \mathbb{Z}_2$ | *id.* $C_s$ | *id.* $C_s$ | *id.* $C_s$ | *id.* $C_s$ | *id.* $C_s$ |
| $C_{6h}$ | $\mathbb{Z}_3 \times \mathbb{Z}_2^2$ | *id.* $C_{2h}$ | *id.* $C_{2h}$ | *id.* $C_{2h}$ | *id.* $C_{2h}$ | *id.* $C_{2h}$ |
| $D_6$ | $Dih_3 \times \mathbb{Z}_2$ | *id.* $D_2$ | *id.* $D_2$ | *id.* $D_2$ | *id.* $D_2$ | *id.* $D_2$ |
| $C_{6v}$ | $Dih_3 \times \mathbb{Z}_2$ | *id.* $C_{2v}$ | *id.* $C_{2v}$ | *id.* $C_{2v}$ | *id.* $C_{2v}$ | *id.* $C_{2v}$ |
| $D_{3h}$ | $Dih_3 \times \mathbb{Z}_2$ | *id.* $C_{2v}'$ | *id.* $C_{2v}'$ | *id.* $C_{2v}'$ | *id.* $C_{2v}'$ | *id.* $C_{2v}'$ |
| $D_{6h}$ | $Dih_3 \times \mathbb{Z}_2^2$ | *id.* $D_{2h}$ | *id.* $D_{2h}$ | *id.* $D_{2h}$ | *id.* $D_{2h}$ | *id.* $D_{2h}$ |
| $T$ | $A_4$ | $\mathbb{Z}_2[A_\rho, A_{\rho'}]^{\mathbb{Z}_3}$ | $\langle A_\rho A_{\rho'}(A_\rho + A_{\rho'}) \rangle$ | 0 | $B_\delta$ | $A_\rho A_{\rho'}(A_\rho + A_{\rho'})$ |
| $T_h$ | $A_4 \times \mathbb{Z}_2$ | $\mathbb{Z}_2[A_\rho, A_{\rho'}]^{\mathbb{Z}_3} \otimes \mathbb{Z}_2[A_\iota]$ | $\langle A_\rho A_{\rho'}(A_\rho + A_{\rho'}), A_\iota^3 \rangle$ | $A_\iota$ | $A_\iota^2 + B_\delta$ | $A_\rho A_{\rho'}(A_\rho + A_{\rho'}) + A_\iota^3$ |
| $O$ | $S_4$ | $\mathbb{Z}_2[A_{\rho'}, B_\delta, C_\psi]/(A_{\rho'} C_\psi)$ | $\langle A_{\rho'} B_\delta + C_\psi \rangle$ | 0 | $A_{\rho'}^2 + B_\delta$ | $A_{\rho'} B_\delta + C_\psi$ |
| $T_d$ | $S_4$ | $\mathbb{Z}_2[A_\sigma, B_\delta, C_\psi]/(A_\sigma C_\psi)$ | $\langle A_\sigma^3, C_\psi \rangle$ | $\sigma$ | $B_\delta$ | $C_\psi$ |
| $O_h$ | $S_4 \times \mathbb{Z}_2$ | $\mathbb{Z}_2[A_\sigma, B_\delta, C_\psi]/(A_\sigma C_\psi) \otimes \mathbb{Z}_2[A_\iota]$ | $\langle A_\iota^2(A_\iota + A_\sigma), A_\sigma^2(A_\iota + A_\sigma), A_\iota B_\delta + C_\psi \rangle$ | $A_\iota + A_\sigma$ | $A_\iota^2 + B_\delta$ | $A_\iota^2(A_\iota + A_\sigma) + \iota B_\delta + C_\psi$ |
| | | $\mathbb{Z}_2[A_\sigma, A_{\sigma'}, B'_\delta, C'_\psi]/((A_\sigma + A_{\sigma'})C'_\psi)$ | $\langle (A_\sigma + A_{\sigma'})A_{\sigma'}^2, (A_\sigma + A_{\sigma'})A_\sigma^2, C'_\psi \rangle$ | $A_{\sigma'} + A_\sigma$ | $A_{\sigma'} A_\sigma + B'_\delta$ | $C'_\psi$ |

$$\widetilde{w}_2 = w_2 + w_1 \cup w_1$$

$$w_1 \in \mathcal{H}^1(G_{\mathrm{space}}, \mathbb{Z}_2)\text{: First Stiefel–Whitney class}$$

$$w_2 \in \mathcal{H}^2(G_{\mathrm{space}}, \mathbb{Z}_2)\text{: Second Stiefel–Whitney class}$$

$$0 = |S_{12} = 0\rangle |S_{34} = 0\rangle$$

$$0 = |S_{13} = 0\rangle |S_{24} = 0\rangle$$

$$\mathcal{H}^5(G, U(1))$$

$R_{xi} = A_x^i, \quad R_{yi} = A_y^i$

$$\mathcal{H}^5(G_{\mathrm{IR}}, U(1)) \to \mathcal{H}^5(G_{\mathrm{UV}}, U(1))$$

$$Z_{\mathrm{UV}} = e^{i\pi \int_{\mathcal{M}_4} A_x \cup A_y \cup \omega[A_{\mathrm{spin}}]}$$

$$Z_{\mathrm{UV}} = e^{i\pi \int_{\mathcal{M}_5} A_x \cup A_y \cup A_z \cup \omega[A_{\mathrm{spin}}]}$$

$$j \wedge *A$$

$$*dA = j_{\mathrm{top}}$$

$$dA \wedge *dA = *d\tilde{A} \wedge d\tilde{A}$$

$$d\tilde{A} = *dA$$

$$A \to A + \lambda$$

$$\tilde{A} \to \tilde{A} + f \colon f \text{ is a } d-2 \text{ form}$$
$$d\lambda = 0, \quad df = 0$$

$\lambda$ is a 1-form, $f$ is a $(d-2)$-form

$$A \wedge B + j_A \wedge *A + j_B \wedge *B$$

$$\delta \quad \partial \quad \omega \quad \omega' \quad \omega'' \quad \omega''' \quad \varepsilon \quad \varepsilon' \quad \varepsilon'' \quad \varepsilon''' \quad \eta \quad \eta' \quad \eta'' \quad \eta''' \quad D = \partial + \delta$$

$$p = 0 \qquad p = 1 \qquad p = 2 \qquad p = 3 \qquad p = 4$$
$$q = 0 \qquad q = 1 \qquad q = 2 \qquad q = 3$$

$$E_0^{p,q} = \mathcal{C}^p(G_{\text{pt}}, C_{d-q}(\mathbb{T}^3, M))$$

$$E_0^{p,q} = \mathcal{C}^p(G_{\text{pt}}, C_{d-q}(\mathbb{T}^2, M))$$

$B\mathbb{Z}^2 = \mathbb{T}^3$

$B\mathbb{Z}^3 = \mathbb{T}^3$

$$\mathcal{H}^3(G, M) \cong H_0^G(\mathbb{R}^3, M) = H_0^{G_{\text{pt}}}(\mathbb{T}^3, M)$$

$$\begin{aligned} \delta\omega &= \partial\omega' \\ \delta\omega' &= \partial\omega'' \\ \delta\omega'' &= \partial\omega''' \end{aligned} \tag{97}$$

Isotropic SS:

vs

LHS SS:

$$\begin{aligned} \partial\omega &= 0 \\ \delta\omega &= 0 \end{aligned} \tag{98}$$

$$\Rightarrow D\omega = 0$$

$$\begin{aligned} \partial\omega &= 0 \\ \varepsilon &:= \delta\omega \\ \varepsilon &\overset{?}{=} \partial\omega' \\ \text{s.t. } \delta\omega' &= 0 \end{aligned} \tag{99}$$

$$\Rightarrow D(\omega - \omega') = 0$$

$$\begin{aligned} \partial\omega &= 0 \\ \varepsilon &:= \delta\omega \\ \varepsilon &\overset{\checkmark}{=} \partial\omega' \\ \varepsilon' &:= \delta\omega' \\ \varepsilon' &\overset{?}{=} \partial\omega'' \\ \text{s.t. } \delta\omega'' &= 0 \end{aligned} \tag{100}$$

$$\Rightarrow D(\omega - \omega' + \omega'') = 0$$

$$
\begin{aligned}
\partial\omega &= 0 \\
\varepsilon &:= \delta\omega \\
\varepsilon &\overset{\checkmark}{=} \partial\omega' \\
\varepsilon' &:= \delta\omega' \\
\varepsilon' &\overset{\checkmark}{=} \partial\omega'' \\
\varepsilon'' &:= \delta\omega'' \\
\varepsilon'' &\overset{?}{=} \partial\omega''' \\
\text{s.t. } \delta\omega''' &= 0
\end{aligned}
\tag{101}
$$

$$\Rightarrow D(\omega - \omega' + \omega'' - \omega''') = 0, \omega - \omega' + \omega'' - \omega''' \in H(G).$$

$$
\begin{aligned}
\delta\omega &= 0 \\
\omega &\overset{?}{=} \partial\eta \\
\text{s.t. } \delta\eta &= 0 \\
D\eta &= \omega
\end{aligned}
\tag{102}
$$

$$
\begin{aligned}
\delta\omega &= 0 \\
\omega &\overset{\checkmark}{=} \partial\eta \\
\omega' &:= \delta\eta \\
\omega' &\overset{?}{=} \partial\eta' \\
\text{s.t. } \delta\eta' &= 0 \\
D(\eta - \eta') &= \omega
\end{aligned}
\tag{103}
$$

$$
\begin{aligned}
\delta\omega &= 0 \\
\omega &\overset{\checkmark}{=} \partial\eta \\
\omega' &:= \delta\eta \\
\omega' &\overset{\checkmark}{=} \partial\eta' \\
\omega'' &:= \delta\eta' \\
\omega'' &\overset{?}{=} \partial\eta'' \\
\text{s.t. } \delta\eta'' &= 0 \\
D(\eta - \eta' + \eta'') &= \omega
\end{aligned}
\tag{104}
$$

$$
\begin{aligned}
\delta\omega &= 0 \\
\omega &\overset{\checkmark}{=} \partial\eta \\
\omega' &:= \delta\eta \\
\omega' &\overset{\checkmark}{=} \partial\eta' \\
\omega'' &:= \delta\eta' \\
\omega'' &\overset{\checkmark}{=} \partial\eta'' \\
\omega''' &:= \delta\eta'' \\
\omega &\in H(G) \\
(\omega, \omega', \omega'', \omega''')
\end{aligned}
\tag{105}
$$

$$Fd\bar{3}m \to I4_1/amd \to I\bar{4}2d \to I\bar{4} \to I2$$

Table 6: Topological invariants applied to the 3-cocycles of $H^3(F222, \mathbb{Z}_2)$.

| Cohomology class | LSM invariants | | | | Non-LSM invariants | | | | | | Triggers LSM? |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\phi_4$ | $\phi_5$ | $\phi_6$ | $\phi_7$ | $\phi_8$ | $\phi_9$ | $\phi_{10}$ | |
| $C_\gamma$ | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $C_\tau$ | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $A_{x+y}A_c^2$ | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | ✓ |
| $A_{x+y}A_cA_{c'}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | No |
| $A_{x+y}A_{c'}^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | No |
| $A_{x+z}A_c^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | No |
| $A_c^3$ | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | No |
| $A_c^2A_{c'}$ | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | ✓ |
| $A_cA_{c'}^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | No |
| $A_{c'}^3$ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | No |

$$Fd\bar{3}m \to F\bar{4}3m \to F23 \to F222 \to C2$$
$$Fd\bar{3}m \to R\bar{3}m \to R32 \to P3_221 \to C2$$

$$I2 \subset I\bar{4} \subset I\bar{4}2d \subset I4_1/amd \subset Fd\bar{3}m$$

$$C2 \subset F222 \subset F23 \subset F\bar{4}3m \subset Fd\bar{3}m$$

$$C2 \subset P3_221 \subset R32 \subset R\bar{3}m \subset Fd\bar{3}m$$

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - eA_\mu J^\mu$$

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{\theta e^2}{32\pi^2}\epsilon^{\mu\nu\lambda\rho}F_{\mu\nu}F_{\lambda\rho}$$

Setting $\hbar = 1$: (for prefactors, see `https://www.diva-portal.org/smash/get/diva2:1215729/FULLTEXT01.pdf` and xiaoliang's paper `https://arxiv.org/pdf/0802.3537.pdf`)

$$j = \sigma E$$

$$j_x = \sigma_{\mathrm{H}}E_y, \delta\rho = \sigma_{\mathrm{H}}B_z$$

$$\mathcal{L}_{2+1\mathrm{D}} = \frac{1}{4\pi}\varepsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho - eA_\mu J^\mu$$

$$\mathcal{L}_{3+1\mathrm{D}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{\theta e^2}{32\pi^2}\epsilon^{\mu\nu\lambda\rho}F_{\mu\nu}F_{\lambda\rho} - eA_\mu J^\mu$$

$$\mathcal{L} = -\frac{1}{4g_0^2}f_{\mu\nu}f^{\mu\nu} + \frac{\theta}{32\pi^2}\epsilon^{\mu\nu\lambda\rho}f_{\mu\nu}f_{\lambda\rho}$$

$$\mathcal{Q}_e = \frac{\theta}{2\pi}\mathcal{Q}_m$$

$$B_\delta(g_1, g_2) = c_1c_2 + c_1c_2' + c_1'c_2' + (c_2 + c_1c_2' + c_2c_2')m_1$$

$$H = \sum_{\langle ij\rangle} J_z S_i^z S_j^z + J_{xy}(S_i^x S_j^x + S_i^y S_j^y)$$
$$+ J_{\pm\pm}\sum_{\langle ij\rangle}(\gamma_{ij}S_i^+ S_j^+ + h.c.)$$
$$+ J_{\pm z}\sum_{\langle ij\rangle}[(i\gamma_{ij}^* S_i^z S_j^+ + h.c.) + (i \leftrightarrow j)],$$

$$H = \sum_{\langle ij \rangle} J_z S_i^z S_j^z + J_\pm (S_i^+ S_j^- + S_i^- S_j^+)$$

$$H_{\text{eff}} = J_z \sum_{i^*} q_{i^*}^2 + U \sum_i B_i^2 - K \sum_p \cos\left(\Sigma_{i \in p} A_i\right)$$

$$S_i^\pm \sim e^{\pm i A_i}, \quad S_i^z \sim B_i$$

$$H^2(\mathbb{Z}^2, \mathbb{Z}_2) = H^2(\pi_1(G/H), \mathbb{Z}_2) = \mathbb{Z}_2$$

$$T_c$$

$$H = J_1 \sum_{\langle ij \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j + J_2 \sum_{\langle\langle ij \rangle\rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j$$

$$H = J \sum_{\langle ij \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j$$

$$H = \sum_{\langle ij \rangle} J_z S_i^z S_j^z + J_{xx} (S_i^x S_j^x + S_i^y S_j^y)$$

$$\theta = \pi$$

Table 7: Topological invariants applied to the 3-cocycles of $H^3(Fd\bar{3}m, \mathbb{Z}_2)$.

| Cohomology class $\lambda$ | LSM invariants | | | | Non-LSM invariants | | | | | Triggers LSM? |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\phi_4$ | $\phi_5$ | $\phi_6$ | $\phi_7$ | $\phi_8$ | $\phi_9$ | |
| $A_i^2(A_m + A_i) + A_i B_{xy+xz+yz}$ | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $A_i B_{xy+xz+yz}$ | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $C_{n\gamma} + C_{s\psi}$ | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $C_{n\gamma}$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ✓ |
| $A_m(A_m^2 + A_i^2)$ | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | No |
| $A_m A_i^2$ | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | No |
| $A_m A_i(A_m + A_i)$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | No |
| $A_m B_{cc'}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | No |
| $A A_m B_{xy+xz+yz}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | No |

$$H_{\text{T.C.}} = -K_p \sum_p (XXXX)_p - K_s \sum_s (ZZZZ)_s$$

$$\boldsymbol{M} \neq 0$$

$$\boldsymbol{M} = \boldsymbol{N}_a - \boldsymbol{N}_b = 0,$$

$$\boldsymbol{N}_{a,b} \neq 0$$

$S = 1/2$

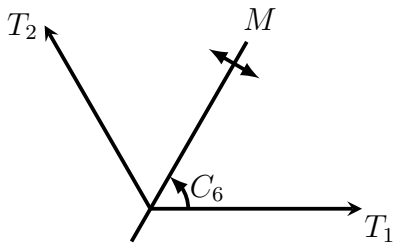
$$S = 1$$

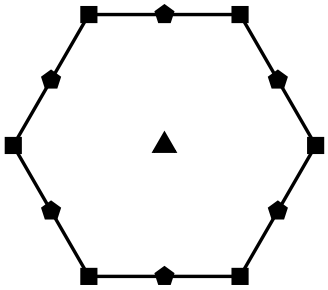
$$H = J \sum_{\langle ij \rangle} S_i^z S_j^z$$

$$J > 0$$

$$J < 0$$

$$H = \sum_i S_i^z S_{i+1}^z$$



$a$: ▲
$b$: ■
$c$: ⬟

$$H = \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1},\ S = 1/2$$

$$H = \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1},\ S = 1$$

$$a_\mu$$

$$\boldsymbol{S} = \frac{1}{2} f^\dagger \boldsymbol{\sigma} f,$$

$$f = \begin{pmatrix} f_\uparrow \\ f_\downarrow \end{pmatrix}$$

$f \to e^{i\theta} f$

$$H_f = \sum_{i,j} t_{ij} f_i^\dagger f_j + h.c. \to |\Psi_f\rangle$$

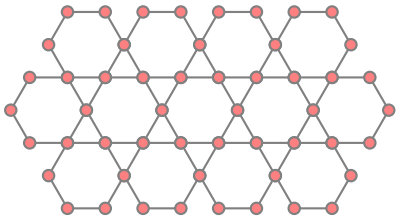
$$|\Psi_{\text{trial}}\rangle = P_{\text{G}} |\Psi_f\rangle$$

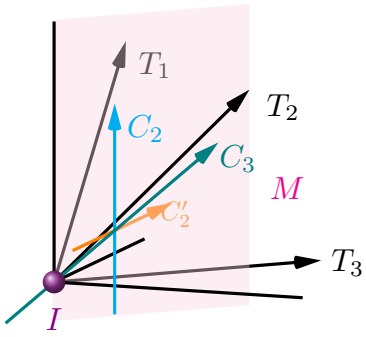
$$Z_{\text{TI}} = e^{i\frac{\theta}{32\pi^2}\int_{\mathcal{M}_4} F\wedge F},\ F = dA$$

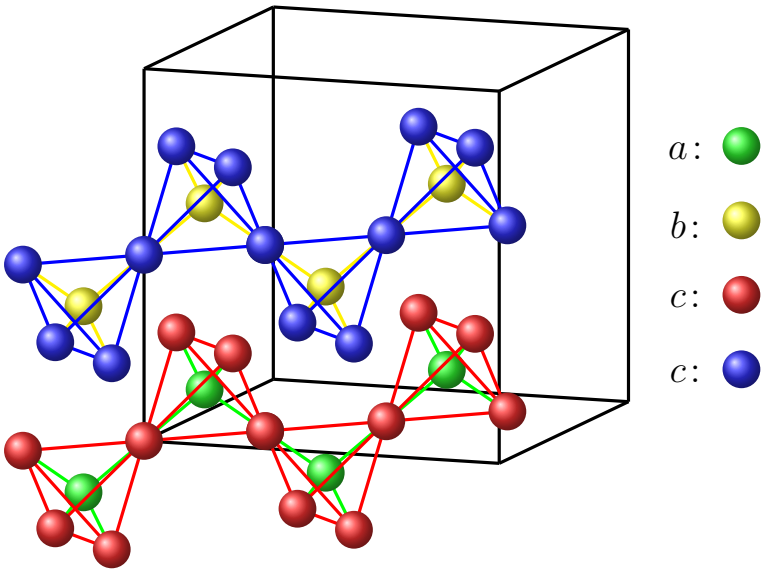
$$t^* \sim 2^N$$

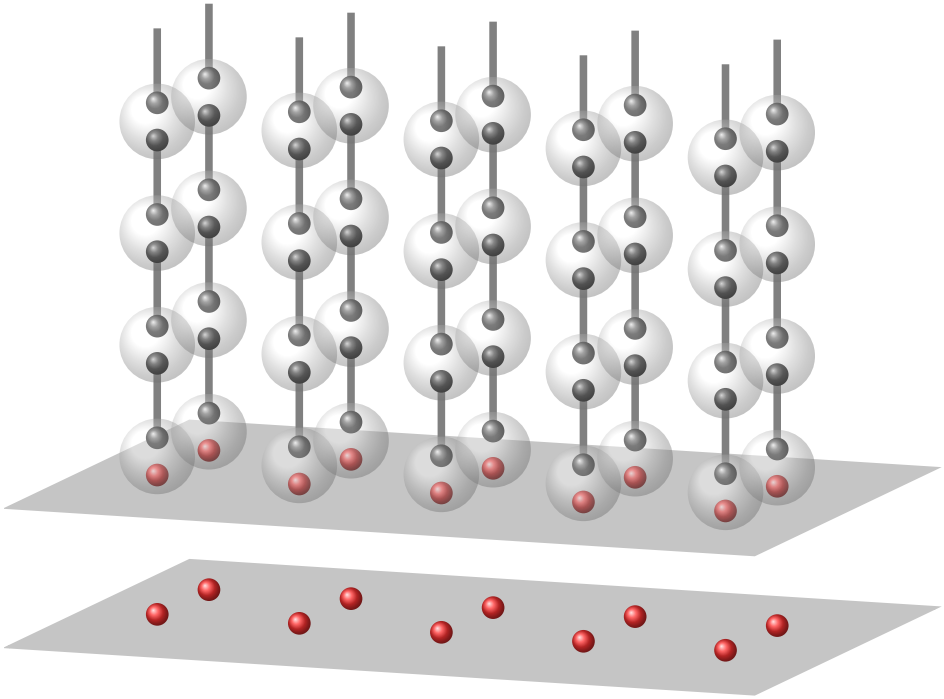
$$S_{\text{von Neumann}} = f(x), \quad x \equiv t/2^N.$$

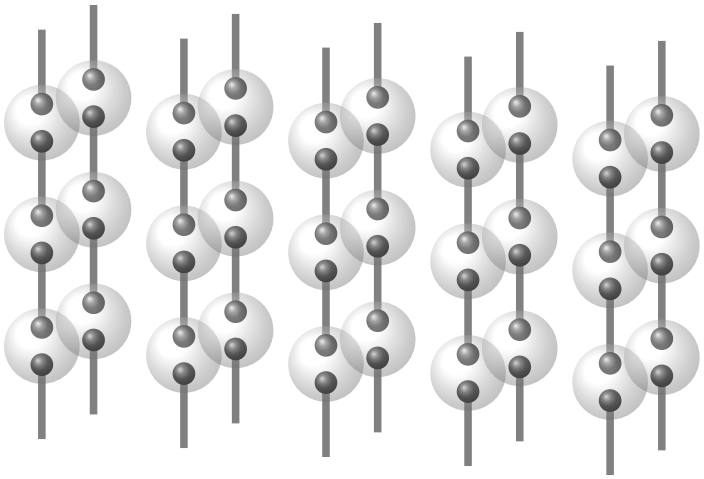
$$S_{\text{von Neumann}} = -\ln x + \text{Const.} + \frac{5}{24}x^2 + \cdots$$




$T_1$
$T_2$
$C_2$
$C_3$
$M$
$C_2'$
$T_3$
$I$


$a$:
$b$:
$c$:
$c$:





$$\beta_2 \colon H^2(G,\mathbb{Z}_2) \to H^3(G,\mathbb{Z}^{or})$$

$$\beta_2 \colon H^3(G,\mathbb{Z}_2) \to H^4(G,\mathbb{Z}^{or})$$

$$\mathcal{SQ}^2 \colon H^3(G,\mathbb{Z}_2) \to H^5(G,\mathbb{Z}_2)$$

$$\mathcal{SQ}^2 := Sq^2 + w_1 \cup Sq^1 + (w_2 + w_1^2) \cup$$

$$Z_{\text{1D}} = e^{i\pi \int_{\mathcal{M}_3} A_x \cup \omega_2^{\text{spin}}}$$

$$Z_{\text{2D}} = e^{i\pi \int_{\mathcal{M}_4} A_x \cup A_y \cup \omega_2^{\text{spin}}}$$

$$Z = e^{i\pi \int_{\mathcal{M}_{d+2}} \lambda \cup \omega_2^{\text{spin}}}$$

$$\omega_2^{\text{SO}(3)} \in H^2(SO(3),U(1)) = \mathbb{Z}_2$$
$$\lambda \in H^d(G,\mathbb{Z}_2)$$

$$\begin{aligned} &H^{d+2}(G \times SO(3), U(1)^{\text{or}}) \\ &= H^d(G, H^2(SO(3),U(1))) \oplus \cdots \\ &= H^d(G,\mathbb{Z}_2) \oplus \cdots \end{aligned} \tag{106}$$

$$Z_{\text{UV}} = \exp\left( i\pi \int_{\mathcal{M}_5} A_x A_y A_z \cup \omega_2^{\text{SO}(3)} \right)$$

$$G_1 \times G_2$$

$$\rho_1 \otimes \rho_2$$

$$H^n(G_1 \times G_2, U(1))$$
$$= \prod_{p+q=n} H^p(G_1, H^q(G_2,U(1)))$$

$$H = J_1 \sum_{\langle ij \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j + J_2 \sum_{\langle\langle ij \rangle\rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j$$

$$\lambda := \frac{M}{N}$$

$$x := \frac{t}{2^q}$$

$$\overline{S_1} = M\left(\ln 2 - \frac{1-\lambda}{\lambda}\ln(1-\lambda) - 1\right)$$

$$\overline{S_1} = -\ln x + \text{const.} + \frac{5}{24}x^2 - \frac{239}{2880}x^4 + \cdots$$

$$C_V \sim aT^{\frac{3}{2}} + b\frac{T^{\frac{3}{2}}}{\ln T} + \cdots$$