Non-invertible symmetries for qubits

Shu-Heng Shao; notes by Bruno Le Floch

July 2024

Contents

1 Basic aspects of non-invertible symmetries

References:

- TASI lectures by Shu-Heng Shao from 2023.
- Seiberg, Shu-Heng Shao `https://arxiv.org/abs/2307.02534`
- Seiberg, Seifnashri `https://arxiv.org/abs/2401.12281`
- Seifnashri, Shu-Heng Shao `https://arxiv.org/abs/2404.01369`

Ordinary symmetries follow the paradigm of Wigner theory: symmetries are generated by unitary or anti-unitary operators, which by definition have an inverse. In recent years it has become increasingly clear that symmetries should be generalized as soon as we go beyond quantum mechanics, going to quantum field theory or lattice systems. Many symmetries are implemented by operators which are conserved but do not necessarily have inverses.

1.1 Examples

The $1+1$ dimensional Ising CFT. (conformal field theory)[1]. Known for decades. This is arguably the simplest conformal field theory of all: it is the first minimal model, it has the smallest central charge $c = 1/2$ among all non-trivial unitary conformal field theory. This theory has a $\mathbb{Z}_2$ symmetry, with a symmetry operator η that obeys $\eta^2 = 1$. In typical lattice realization, η flips the spins from $\downarrow$ to $\uparrow$. At the CFT point, we have an extra operator $\mathcal{D}$, which obeys the algebra

$$\mathcal{D}\eta = \eta\mathcal{D} = \mathcal{D}, \qquad \mathcal{D}^2 = 1 + \eta, \qquad \mathcal{D}^\dagger = \mathcal{D}. \tag{1}$$

From the first equation we see immediately that $\mathcal{D}$ cannot have an inverse: otherwise multiplying on the left by this inverse would tell us that η equals 1. We will eventually rederive this result in these lectures. Nice example for bootstrappers, or string theorists.

Non-invertible chiral symmetry. (Known since 2022.) Non-invertible chiral symmetry in $3+1$ dimensional QED with ABJ anomaly, to be precise a $U(1)$ gauge theory with charged chiral fermions. Classically the axial $U(1)$ acts as $\Psi \to e^{i\alpha\gamma_5/2}\Psi$ with $\alpha \in [0, 2\pi)$. This was discussed in Clay Córdova's lectures. Nice example for particle physicists.

Quantum Ising lattice model. The simplest example, suitable for an advanced undergraduate, is the non-invertible symmetry in the quantum Ising lattice model in $1+1$ dimensions, to be precise space will be discrete while time will be continuous. It is also called the transverse field Ising model.

The $\mathrm{Rep}(G)$ **construction.** For those who know what $\mathrm{Rep}(G)$, it is a very nice construction in 2d. But its generalization to higher dimensions gives a $(d-2)$-form non-invertible symmetry, which is more tricky to understand than non-invertible 0-form symmetries.

1.2 Ising lattice model

1.2.1 The model

This is a spin chain with sites labelled by $j \in \{1, \dots, L\}$ (periodic). The Hilbert space at each site is a qubit $\mathcal{H}_j = \mathbb{C}^2$. The whole Hilbert space is

$$\mathcal{H} = \bigotimes_{j=1}^{L} \mathcal{H}_j, \qquad \dim \mathcal{H} = 2^L.$$

We define operators X_j, Z_j acting on $\mathcal{H}_j$ with

$$X_j Z_j = -Z_j X_j, \qquad Z_j^2 = X_j^2 = 1.$$

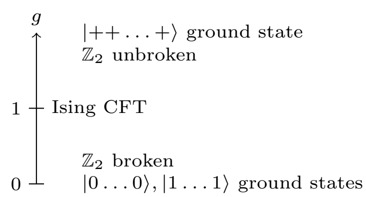
[1] In these lectures, by CFT we really mean the continuum theory, not critical lattice realizations.

There is a basis $|0\rangle$, $|1\rangle$ with $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$ and another basis with $X|\pm\rangle = \pm|\pm\rangle$. Then we define the Hamiltonian (with coupling g)

$$H = -g \sum_{j=1}^{L} X_j - g^{-1} \sum_{j=1} Z_j Z_{j+1}.$$

This is the leading-order Hamiltonian if we impost the following usual symmetries:

- the $\mathbb{Z}_2$ on-site symmetry $\eta = \prod_{j=1}^{L} X_j$, with $\eta^2 = 1$;
- lattice translation T, which obeys $T X_j T^{-1} = X_{j+1}$ and $T Z_j T^{-1} = Z_{j+1}$.



1.2.2 What is an analogue of $\mathcal{D}$?

Question: given what people know about the Ising CFT, does $\mathcal{D}$ exist at $g = 1$ (and finite L)? Conditions are

It acts within the Hilbert space $\mathcal{H}$. In some related contexts this condition is sometimes relaxed, which is why we are explicit about it here.

It must commute with the Hamiltonian H. Spacetime symmetries like Lorentz boosts can fail to commute with the Hamiltonian, but here we are not interested in such symmetries. Another situation is models where the Hamiltonian depends on time, in which case there may be subtleties.

It becomes $\mathcal{D}$ in the $L \to +\infty$ limit. The Ising CFT, and the Ising lattice model, are integrable, so there are actually infinitely many symmetries; we have to be careful when identifying which symmetry we managed to find. We cannot impose the algebra (1) because it will turn out to be deformed, as we will see. One should take into account locality. For instance, the projector $\Pi_{|E\rangle}$ onto a given energy eigenstate definitely commutes with the Hamiltonian, but it is not a good symmetry because after rotating time and space (in Euclidean signature) the corresponding "defect" does not give a Hilbert space of integer dimension.

1.2.3 Kramers–Wannier transformation

At $g = 1$ there is a well-known transformation that does **roughly**

$$X_j \mapsto Z_j Z_{j+1}, \qquad Z_j Z_{j+1} \mapsto X_{j+1}. \tag{2}$$

This looks very much like a symmetry! Wigner's paradigm says there should exist U such that $UX_jU^{-1} = Z_jZ_{j+1}$. Now

$$U\eta U^{-1} = U\prod_{j=1}^{L} X_j U^{-1} = \prod_{j=1}^{L} Z_j Z_{j+1} = 1 \qquad \text{if } U \text{ existed}$$

because each Z_j appears exactly twice and $Z_j^2 = 1$ (and these operators commute with one another). Then multiplying by U^{-1} and U on the left and right would give $\eta = 1$. There are several ways to make this work.

- Make the Hilbert space smaller: impose a global constraint on the Hilber space, namely reduce to states on which $\eta = 1$. Then the Hilbert space no longer has the structure of a tensor product.
- Make the Hilbert space larger: see later?
- Allow $\mathcal{D}$ to project out some states.

Lesson: in physics we should not write arrows, because it can hide some subtleties. We should try to write equalities. If the lecturer writes arrows later in the lecture, we should stop him.

(Question from the audience: what if we work with an open chain instead of a closed chain? It turns out that the Kramers–Wannier transformation does not exist on an open chain because its square is a lattice translation, which does not exist on a (finite) open chain.[2])

1.2.4 Definition and properties of D

Definition of D We define (the normalization is important but complicated to explain[3])

$$D = \sqrt{2}e^{-2\pi i L/8}U_{\text{KW}}\frac{1+\eta}{2}, \qquad U_{\text{KW}} = \left(\prod_{j=1}^{L-1}\frac{1+iX_j}{\sqrt{2}}\frac{1+iZ_jZ_{j+1}}{\sqrt{2}}\right)\frac{1+iX_L}{\sqrt{2}}.$$

Note that each of the fractions in U_{KW} are unitary, for instance $\frac{1+iX_j}{\sqrt{2}} = e^{i\pi X_j/4}$. The operator U_{KW} is a "sequential quantum circuit", meaning roughly that the number of factors is proportional to the number of sites, and the operators act

[2]What about an infinite open chain?

[3]E.g., the factor of 1/8 is related to the critical dimension of superstring theory. In light-cone gauge there are $10 - 2 = 8$ transverse directions. We will not need this connection in these lectures.

on nearby sites. The last factor, $(1+\eta)/2$, is a projector onto the $\eta = 1$ subspace. The operator U_{KW} is not translationally-invariant, see the exercises, but D is, thanks to the projection. The projector has a kernel, so D is a *non-invertible matrix*. Let us check whether it commutes with the Hamiltonian H.

Question: why the $\sqrt{2}$ normalization in D? *Answer:* to match the normalization in the continuum, where there is a good notion of what is the correct normalization and we wouldn't like factors of $1/2$ in the upcoming fusion rule $D^2 = 1 + \eta$.

Commutation with the Hamiltonian One can calculate

$$U_{\mathrm{KW}} X_j U_{\mathrm{KW}}^{-1} = \begin{cases} Z_j Z_{j+1}, & j \neq L, \\ \eta Z_L Z_1, & j = L \end{cases}$$

where we recall $\eta = \prod_{j=1}^{L} X_j$. Thus, U_{KW} clearly is not translationally-invariant. But once we impose the projector, we get the rigorous version of the arrows of (2): for all j,

$$DX_j = Z_j Z_{j+1} D, \qquad DZ_j Z_{j+1} = X_{j+1} D.$$

Thus,

$$DH = HD \text{ for } g = 1,$$

and more generally $DH = H|_{g \to g^{-1}} D$. (Audience question: what about DZ_j? Answer, Z_j is $\mathbb{Z}_2$-odd so DZ_j cannot be written as $\mathcal{O}D$ since it has to project onto the $\eta = -1$ sector.)

How we evaded Wigner. Wigner's theorem assumed that the operator preserves the norm of the state, to preserve probabilities. Our operator does not preserve the norm.

Question from Slava: is this operator useful for finite L? We will get to that, see the section "what is it good for?"

1.3 The algebra of D

The algebra of D and its friends Roughly speaking, D is a $1/2$ lattice translation.

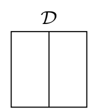
$$\eta^2 = 1, \qquad T^L = 1, \qquad \eta T = T\eta,$$
$$TD = DT, \qquad \eta D = D\eta = D, \qquad T^{-1}D = DT^{-1} = D^\dagger, \qquad D^2 = (1+\eta)T.$$

Interestingly, $D^\dagger$ is not D but it differs by a lattice translation. We also have a translation in D^2. We don't quite get the continuuum algebra. But in the continuum limit this translation does not matter.

Remark by Giulio: $DD^\dagger = D^\dagger D = 1 + \eta$ is not deformed. But there is still T in the $D^\dagger = T^{-1}D$.

Numerous remarks and questions. *Remark by a student:* if we take a symmetry and multiply by some projector, or if we take the projector $(1+\eta)/2$ itself, then that's also some symmetries in the sense that it commutes with the Hamiltonian. *Answer:* such cases would be fully constructed from unitary symmetry operators, so it is not teaching us anything new. In contrast, the one we have here really does not exist as an invertible symmetry.

Question: what is special about D and locality? *Answer:* in the continuum $\mathcal{D}$ can be used to twist space and has to give a nice Hilbert space interpretation to a torus partition function



$$= \mathrm{Tr}_{\mathcal{H}_\mathcal{D}}\, e^{-iHt}.$$

Follow-up question: this is not an interpretation about the operator acting on the Hilbert space. *Answer in the Hamiltonian lattice language:* every symmetry should give you two objects, an operator acting on the Hilbert space, and a defect, see early in Max Metlitski's lectures. *Follow-up question:* there is no way to characterize such good operators by some locality property etc? *Answer:* probably not because there are many local operators that are not mapped to local operators.

Follow-up by someone else: even the continuum story is not so good because it needs Lorentz invariance. *Answer:* good point.

Question by Yifan Wang: you said there are two pieces of information, and the second tells you you know how to act on only part of the space, so it is a kind of locality property.

1.3.1 Derivation of the algebra

A model of Majorana fermions Write each chiral fermion in terms of a pair of Majoranas so that we end up with $2L$ Majorana fermions χ_l, $l = 1, \dots, 2L$, with

$$\{\chi_l, \chi_{l'}\} = 2\delta_{l,l'}.$$

We take the Hamiltonian

$$H_\pm = i \sum_{l=1}^{2L-1} \chi_{l+1}\chi_l \pm i\chi_1\chi 2L.$$

This differs from the models Max Metlitski mentioned in his lectures: we have terms for every pair of contiguous Majoranas, not every other one. Usual unitary symmetries of the Hamiltonian:

- Fermion parity $(-1)^F = i^L \chi_1\chi_2 \dots \chi_{2L}$, which obeys $(-1)^F(-1)^F = 1$ adn $(-1)^F\chi_l(-1)^F = -\chi_l$.
- Majorana translations T_+ with $T_+\chi_l T_+^{-1} = \chi_{l+1}$ and T_- with $T_-\chi_l T_-^{-1} = \chi_{l+1}$ for $l \neq 2L$ and $T_-\chi_{2L}T_-^{-1} = -\chi_1$.

Here $T_\pm$ is a symmetry of $H_\pm$, respectively.

Commutation of translation and fermion parity. For H_+ we compute

$$T_+(-1)^F T_+^{-1} = T_+ i^L \chi_1\chi_2 \dots \chi_{2L} T_+^{-1} = i^L \chi_2\chi_3 \dots \chi_{2L}\chi_1 = -(-1)^F$$

since we need to move χ_1 through an odd number of χ_j, $j = 2, \dots, 2L$. People in the literature have interpreted this as an LSM-type anomaly. For H_- we get $T_-(-1)^F T_-^{-1} = (-1)^F$.

Continuum model. We consider $\mathcal{L} = i\psi_L(\partial_t - \partial_x)\psi_L + i\psi_R(\partial_t + \partial_x)\psi_R$. We have two fermion parities $(-1)^{F_L}$ and $(-1)^{F_R}$ flipping ψ_L and ψ_R, respectively. Then H_+ corresponds to Ramond–Ramond boundary conditions, for which we can show $(-1)^F(-1)^{F_L} = -(-1)^{F_L}(-1)^F$ while H_- corresponds to Neveu–Schwarz–Neveu–Schwarz boundary conditions for which there is no such sign.

Ising versus Majorana People often say Ising $\stackrel{?}{=}$ Majorana. What it means is that you can do a Jordan–Wigner transformation by pairing up the Majoranas,

$$\chi_{2j-1} = \left(\prod_{k=1}^{j-1} \sigma_k^x\right)\sigma_j^y,$$
$$\chi_{2j} = \left(\prod_{k=1}^{j-1} \sigma_k^x\right)\sigma_j^z,$$

where σ are Pauli matrices. We will see shortly why we use a different notation σ_k^x rather than X_k as before; there are some differences.

The important thing here is that a local operator is mapped to a non-local operator in terms of Paulis. In fact it corresponds to a gauging, as we will see next time. If you apply this transformation to $H_\pm$ you find

$$H_\pm = -\sum_{j=1}^{L} \sigma_j^x - \sum_{j=1}^{L-1} \sigma_j^z\sigma_{j+1}^z \pm (-1)^F \sigma_L^z\sigma_1^z. \tag{3}$$

Let us meditate on this Hamiltonian. This is almost the same as the critical Ising model. But there is one "mistake" involving $(-1)^F$, which is non-local in terms of the Pauli matrices:

$$(-1)^F = \prod_{j=1}^{L} \sigma_j^x.$$

The reason we want to use this map from Ising to Majorana is that it is supposed to help identify the symmetry. We need to pin down this one problem in the conversion from one model to the other.

One diagnostic of the problem is that $T_+^{2L} = 1$ on the Majorana chain, which is a $\mathbb{Z}_{2L}$ symmetry. But the Ising model only has $\mathbb{Z}_L$ symmetry, and translation does not have a square root.

As it turns out, the Ising model will be a gauging of the Majorana model by $(-1)^F$.

July 2. Some reminders from yesterday We have $(-1)^F = i^L \chi_1 \dots \chi_{2L}$. It obeys $(-1)^F \chi_\ell (-1)^F = -\chi_\ell$. We also have a Majorana translation $T_\pm$ with $T_\pm \chi_{2L} T_\pm^{-1} = \pm\chi_1$ and all other $T_\pm \chi_\ell T_\pm^{-1} = \chi_{\ell+1}$.

Anomaly (Majorana fermions). We have

$$T_+(-1)^F = -(-1)^F T_+, \quad T_-(-1)^F = (-1)^F T_-.$$

The minus sign for T_+ prevents us from diagonalizing both operators at the same time. It is a sort of anomaly.

Question: is T_- non-anomalous? *Answer:* yes and no: we should think of $T_\pm$ as one symmetry with different boundary conditions, in which case we say that this translation symmetry T has an anomaly because it has one for some choice of boundary condition.

1.3.2 Lattice bosonization

We consider two copies of $\mathcal{H}$, denoted $\mathcal{H}_\pm$, with $\dim \mathcal{H}_\pm = 2^L$. We set $\widetilde{\mathcal{H}} = \mathcal{H}_- \oplus \mathcal{H}_+$, which has dimension 2^{L+1}. (This is similar to how in string theory we begin by putting together the NSNS and RR sectors before performing the GSO projection; actually it is most similar to the type 0 case.) Then we take the Hamiltonian to be

$$\widetilde{H} = \begin{pmatrix} H_- & 0 \\ 0 & H_+ \end{pmatrix}, \qquad H_\pm = \sum_{\ell=1}^{2L-1} i\chi_{\ell+1}\chi_\ell \pm i\chi_1\chi_{2L}.$$

Then we perform the projection $(-1)^F = +1$ in $\mathcal{H}_-$ and $(-1)^F = -1$ in $\mathcal{H}_+$. This ensures that the last term in (3) has the same (minus) sign as all of its friends, and is explicitly periodic. Then we consider

$$\mathcal{H} = \widetilde{\mathcal{H}}|_{\text{proj}}.$$

The notation is no coincidence: this matches the Ising model Hilbert space. Then

$$X_j = \begin{pmatrix} \sigma_j^x & 0 \\ 0 & \sigma_j^x \end{pmatrix}, \qquad Z_j = \begin{pmatrix} 0 & \sigma_j^z \\ \sigma_j^z & 0 \end{pmatrix},$$

which explains the difference of notation.

The fate of Majorana translations. In the extended Hilbert space $\widetilde{\mathcal{H}}$ there is a nice Majorana translation

$$\widetilde{T} = \begin{pmatrix} T_- & 0 \\ 0 & T_+ \end{pmatrix}.$$

In the next step of projecting the Hilbert space, we see that since T_+ does not commute with $(-1)^F$, $\widetilde{T}$ does not act within the projected Hilbert space $\mathcal{H} = \widetilde{\mathcal{H}}|_{\text{proj}}$.

The square $\widetilde{T}^2$ commutes with $(-1)^F$ hence it does act on the projected Hilbert space. We can thus define $T = \widetilde{T}^2|_{\text{proj}}$, which is a $\mathbb{Z}_L$ translation as expected.

We are being too cruel throwing away T_-: after all there is nothing wrong with it, it commutes with $(-1)^F$. We shouldn't punish it for the sin of its brother. We consider

$$D = \sqrt{2}\begin{pmatrix} T_- & 0 \\ 0 & 0 \end{pmatrix}\Bigg|_{\text{proj}} \qquad \text{on } \mathcal{H}.$$

This precisely reproduces the non-invertible symmetry D (including the phase?). It is then clear what D^2 is: the Ising translation projected to one block, so $D^2 = (1+\eta)T$ as announced. It also makes $DT = TD$ obvious using that $\widetilde{T}$ is block diagonal.

This is part of a typical setting where *an invertible symmetry with an anomaly can become a non-invertible symmetry.*

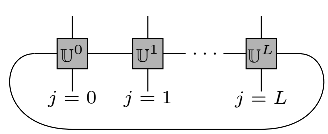
Question: here we managed to "purify" a non-invertible symmetry into an invertible symmetry on a larger Hilbert space; is this generalizable. *Answer:* not completely generally, what is happening here is related to the fact that Kramers–Wannier duality is a gauging operation. See recent paper by Tachikawa and Okada.

1.3.3 Alternative expression for D

Claim: after a non-trivial calculation,

$$D = \mathrm{Tr}_2\Big[\mathbb{U}^1\mathbb{U}^2\dots\mathbb{U}^L\Big], \qquad \mathbb{U}^j = \begin{pmatrix} |0\rangle_j\langle +|_j & |0\rangle_j\langle -|_j \\ |1\rangle_j\langle -|_j & |1\rangle_j\langle +|_j \end{pmatrix}$$

where $\mathbb{U}^j$ are operator-valued 2×2 matrices, we use the matrix product and the trace of 2×2 matrices. Schematic representation



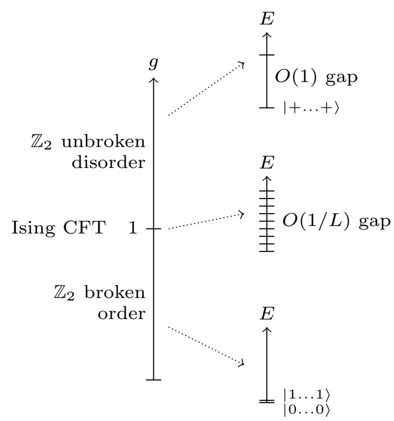
From this point of view rescaling the $\mathbb{U}^j$ by some factor like 7 would rescale D by 7^L, so it is not natural to try and remove the $\sqrt{2}$ overall factor. Somehow this MPO (matrix product operator) point of view knows about the preferred normalization of D.

Question: is the bond dimension related to the quantum dimension? *Answer:* I don't know because the quantum dimension is only defined in the continuum limit.

Question: is there a continuum interpretation of the horizontal legs in this picture? *Answer:* maybe related to a gauge field, not clear.

1.3.4 What is it good for?

Phase diagram of the Ising model.



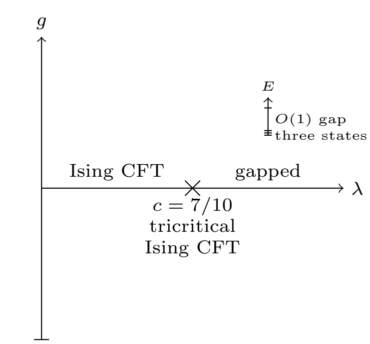
Self-dual deformation. There is a nice self-dual deformation of the self-dual Ising model:

$$H = -g\sum_{j=1}^{L} X_j - g^{-1}\sum_{j=1}^{L} Z_j Z_{j+1} + \frac{\lambda}{2}\sum_{j=1}^{L}\Big(X_{j-1}Z_jZ_{j+1} + Z_{j-1}Z_jX_{j+1}\Big).$$

This manifestly commutes with the Kramers–Wannier duality transformation.

	$\mathbb{Z}_2 : \eta$	D
g	✓	No
λ	✓	✓

Phase diagram (found through general arguments and confirmed by DMRG)



We can actually prove the threefold multiplicity in the gapped phase at large λ.

For instance, there is no relevant operator in the Ising CFT that preserves the non-invertible symmetry $\mathcal{D}$, which explains why the Ising CFT continues being the fixed point in a neighborhood of $\lambda = 0$.

2 Non-invertible symmetry protected topological (SPT) phases

Given a symmetry, can it be realized in a trivially gapped phase?

- Case 1: no. Suppose you manage to prove for a given system that the answer is no. Then it puts constraints on the phase diagram. For instance there can be anomalous invertible symmetries, or symmetries involving lattice translations (like in the LSM theorem, or the non-invertible Kramers–Wannier symmetry etc). For some community of people we would say that this is an anomaly.
- Case 2: yes. This is much easier to show, just identify one example. Then we can ask to classify SPTs. For instance it can be useful for on-site invertible symmetries (as in Max's lectures) or non-invertible SPTs.

2.1 Cluster model and $\mathbb{Z}_2 \times \mathbb{Z}_2$ SPTs

Cluster model. We consider a periodic chain with an even number $L \in 2\mathbb{Z}$ of qubits. Our first Hamiltonian has a product form $H_{\text{prod}} = -\sum_{j=1}^{L} X_j$. It has a single vacuum and an order-1 gap. Our second Hamiltonian is

$$H_{\text{cluster}} = -\sum_{j=1}^{L} Z_{j-1} X_j Z_{j+1},$$

which is actually a sum of independent[4] commuting Pauli operators. The spectrum is exactly the same as the one of H_{prod}. The ground state $|\text{cluster}\rangle$ is a bit more messy to write.

$$|\text{cluster}\rangle = V|{+}{+}\dots{+}\rangle, \qquad V = \prod_{j=1}^{L} \mathrm{CZ}_{j,j+1},$$

where $\mathrm{CZ}_{j_1,j_2} = \frac{1}{2}(1 + Z_{j_1} + Z_{j_2} - Z_{j_1} Z_{j_2})$ is invertible (it squares to 1). In fact more generally

$$V H_{\text{prod}} V^{-1} = H_{\text{cluster}}.$$

If we did not care about symmetries, then the difference between H_{prod} and H_{cluster} would be unimportant.

Symmetries. We have $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ symmetries (where e, o stand for even and odd):

$$\eta^e = \prod_{j \text{ even}} X_j, \quad \eta^o = \prod_{j \text{ odd}} X_j,$$

which commute and square to the identity. They are obviously symmetries of the product Hamiltonian. They are also symmetries of the cluster Hamiltonian

[4]Slava complained about this word missing.

because that Hamiltonian only involves pairs of Z whose indices have the same parity.

Claim 1. *The ground states* $|{+}{+}\ldots{+}\rangle$ *and* $|\text{cluster}\rangle$ *are distinct* $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ *SPTs. In continuum the difference of SPTs is captured by the invertible field theory* $\exp\left[i\pi \int A^e \cup A^o\right]$.

Comparing the two phases. The first way of detecting the difference between the two phases is the edge modes. Let us put them side by side on a closed periodic chain. From $j = 1$ to $j = L'$ we use the cluster Hamiltonian and from $j = L' + 1$ to $j = L$ we use the product Hamiltonian. We choose

$$H_{\text{interface}} = -Z_1 X_2 Z_3 - Z_2 X_3 Z_4 - \cdots - Z_{L'-2} X_{L'-1} Z_{L'} - X_{L'+1} - \cdots - X_L.$$

This preserves $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ symmetry. We assume L, L' even. Later L a multiple of 4.

We could have done a slightly different choice at the interface, like adding X_L' and X_1, but the present choice ensures that all terms continue commuting. All that we will say about the phases will not depend on details of the interface. We could also change the relative factor between the two parts of the Hamiltonian. For instance putting an infinite factor in front of $-X_{L'+1} - \cdots - X_L$ would pin down all the spins in that half of the spin chain, which would match the usual approach of considering an open spin chain.

Let us solve this Hamiltonian. The Hamiltonian has $L - 2$ independent commuting Pauli matrices, so there are $2^L/2^{L-2} = 4$ ground states. A ground state $|\psi\rangle$ obeys

$$\begin{aligned} Z_{j-1} X_j Z_{j+1} |\psi\rangle &= |\psi\rangle, \qquad j = 2, 3, \ldots, L' - 1, \\ X_j |\psi\rangle &= |\psi\rangle, \qquad j = L' + 1, \ldots, L. \end{aligned}$$

We have

$$\eta^e |\psi\rangle = Z_1 \prod_{j=1}^{L'/2-1} (Z_{2j-1} X_{2j} Z_{2j+1}) Z_{L'-1} X_{L'} X_{L'+2} \ldots X_L |\psi\rangle = Z_1 Z_{L'-1} X_{L'} |\psi\rangle.$$

We denote $\eta^e_{\mathrm{L}} = Z_1$ and $\eta^e_{\mathrm{R}} = Z_{L'-1} X_{L'}$. Likewise we have $\eta^o |\psi\rangle = \eta^o_{\mathrm{L}} \eta^o_{\mathrm{R}}$ with $\eta^o_{\mathrm{L}} = X_1 Z_2$ and $\eta^o_{\mathrm{R}} = Z_{L'}$. Then it turns out that we have

$$\eta^e_{\mathrm{L}} \eta^o_{\mathrm{L}} = -\eta^o_{\mathrm{L}} \eta^e_{\mathrm{L}}, \qquad \eta^e_{\mathrm{R}} \eta^o_{\mathrm{R}} = -\eta^o_{\mathrm{R}} \eta^e_{\mathrm{R}}.$$

Each of these algebras imply a two-fold degeneracy of the ground state. Together this explains the fourfold degeneracy of the ground state.

Another point of view is the finite-depth local unitary circuit (FDLU).

Kennedy–Tasaki transformation. Another approach is the Kennedy–Tasaki transformation, which gauges the $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ symmetry, with a twist. In the high-energy literature it is usually denoted as TST (composition of modular T and

S transformations). The transformation maps (in a sense that we will have to make precise)

$$X_j \to \widehat{X}_j, \qquad Z_{j-1}Z_{j+1} \to \widehat{Z}_{j-1}\widehat{X}_j\widehat{Z}_{j+1}.$$

Then

$$\begin{aligned}
H_{\text{prod}} &= -\sum_j X_j & &\to \widehat{H}_{\text{prod}} = -\sum_j \widehat{X}_j & &\widehat{\mathbb{Z}_2}\times\widehat{\mathbb{Z}_2} \text{ unbroken},\\
H_{\text{cluster}} &= -\sum_j Z_{j-1}X_jZ_{j+1} & &\to \widehat{H}_{\text{cluster}} = -\sum_j \widehat{Z}_{j-1}\widehat{Z}_{j+1} & &\widehat{\mathbb{Z}_2}\times\widehat{\mathbb{Z}_2} \text{ broken}.
\end{aligned}$$

This gauging operation gives an invertible field theory in the first case, and a non-invertible one (four ground states) in the second. This is consistent with the idea that the phases are only different when we track the symmetry, or when we gauge it. The $\mathbb{Z}_2 \times \mathbb{Z}_2$ before and after the Kennedy–Tasaki transformation is different (the second is the quantum symmetry generated by Wilson lines of the gauged $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$).

2.2 Non-invertible symmetry of the cluster model

We introduce an operator D (different from the Ising one) acting as

$$DX_j = Z_{j-1}Z_{j+1}D, \qquad DZ_{j-1}Z_{j+1} = X_jD.$$

It can be constructed as $D = T^{-1}D^eD^o$ where D^e and D^o are Kramers–Wannier transformations on the even and odd sites, and T is the lattice translation by one site. For instance, D^eD^o maps X_j to Z_jZ_{j+2} (for j even/odd this is done by D^e or D^o respectively), which is then "recentered" by the lattice translation.

We can find the algebra

$$D\eta^e = \eta^e D = \eta^o D = D\eta^o = D, \qquad D^2 = (1+\eta^e)(1+\eta^o).$$

Contrarily to the Kramers–Wannier case where we had a lattice translation in D^2, this does not happen here.

As it turns out, the fusion category underlying this set of operators is $\mathrm{Rep}(D_8)$, whose objects are the five irreducible representations

$$1, \eta^e, \eta^o, \eta^e\eta^o, D$$

of the dihedral group D_8. The fusion rules are simply decompositions of tensor products of representations into irreducible ones.[5]

[5] Caveat: from the presentation here, we could not determine whether the group is D_8 or Q_8, but one can compute the "F-symbols" on the lattice and distinguish these two cases. Alternatively there is a construction of the theory by gauging a D_8 group, so that the gauged theory automatically has $\mathrm{Rep}(D_8)$ symmetry.

Action on the cluster state. We have a new symmetry D, which extends the $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ symmetry to $\mathrm{Rep}(D_8)$. We can compute

$$D|\text{cluster}\rangle = 2|\text{cluster}\rangle, \qquad \eta^e|\text{cluster}\rangle = \eta^o|\text{cluster}\rangle = |\text{cluster}\rangle.$$

This means that $|\text{cluster}\rangle$ is invariant under $\mathrm{Rep}(D_8)$ hence is a $\mathrm{Rep}(D_8)$-SPT state. Said in a different way, H_{cluster} is a trivially-gapped $\mathrm{Rep}(D_8)$-SPT phase.

Naive question: is it a trivial $\mathrm{Rep}(D_8)$-SPT state? The notion of triviality is meaningless. Correct question: are there other $\mathrm{Rep}(D_8)$-SPT states?

Very important note: we cannot consider an open chain, because the open boundary conditions would break D. This can also be seen by noting that the product Hamiltonian H_{prod} is not invariant under D so it does not make sense to study $\mathrm{Rep}(D_8)$ for this Hamiltonian, or to put it on half of a closed chain like we did before.

In mathematics, it is known that there are 3 distinct $\mathrm{Rep}(D_8)$-SPT states.

On stacking SPT. For a group G, consider two G-SPT states $|\psi_i\rangle$, $i = 1, 2$ and consider $|\psi_1\rangle \otimes |\psi_2\rangle$. It has a $G \times G$ symmetry coming from G acting on each factor separately. Under the diagonal subgroup $G \subset G \times G$ we get a new G-SPT. Thus there is an addition structure among SPT. Not quite a group, it is a torsor over a group, specifically a torsor over $H^2(G, U(1))$.

For non-invertible symmetry there is no analogue of stacking, because there is no analogue of diagonal subgroup. Concretely for $\mathrm{Rep}(D_8)$, if you try to map $D \mapsto (D \otimes D)$ in $\mathrm{Rep}(D_8) \times \mathrm{Rep}(D_8)$ then it doesn't work because $(D \otimes D)^2 = (1+\eta^e)(1+\eta^o) \otimes (1+\eta^e)(1+\eta^o)$ has many more terms than $1 \otimes 1 + \eta^e \otimes \eta^e + \eta^o \otimes \eta^o + \eta^e\eta^o \otimes \eta^e\eta^o$. Thus, there is really no group structure.

The three SPTs. We consider $L = 0 \mod 4$. We denote the three SPTs as $|\text{cluster}\rangle$, $|\text{even}\rangle$, $|\text{odd}\rangle$. We have already seen the first. The $|\text{even}\rangle$ state is characterized by

$$\begin{aligned} Z_{2k}X_{2k+1}Z_{2k+2} &= -1, \\ Y_{2k-1}X_{2k}Y_{2k+1} &= 1. \end{aligned}$$

The $|\text{odd}\rangle$ state is characterized by

$$\begin{aligned} Z_{2k-1}X_{2k}Z_{2k+1} &= -1, \\ Y_{2k}X_{2k+1}Y_{2k+2} &= 1. \end{aligned}$$

It is a bit subtle to make a Hamiltonian with one of these as ground state and that is $\mathrm{Rep}(D_8)$-invariant. Adding up the $Z_{2k-1}X_{2k}Z_{2k+1}$ and $-Y_{2k}X_{2k+1}Y_{2k+2}$ is not enough to preserve the symmetry.[6]

[6] Concretely, $H = \sum_k Z_{2k-1}X_{2k}Z_{2k+1} - \sum_k Y_{2k}X_{2k+1}Y_{2k+2} + \sum_k Z_{2k-1}Z_{2k}X_{2k+1}Z_{2k+2}Z_{2k+3}$.

Putting phases side by side. We can prove that $|\text{cluster}\rangle$, $|\text{even}\rangle$, $|\text{odd}\rangle$ are the same $\mathbb{Z}_2^e \times \mathbb{Z}_2^o$ SPT. For instance when placing $|\text{cluster}\rangle$ and $|\text{odd}\rangle$ side by side, we find that ground states $|\psi\rangle$ are invariant under η^e, and that $\eta^o|\psi\rangle = \eta^o_{\text{L}}\eta^o_{\text{R}}|\psi\rangle$ with local factors on the left and right boundaries.

The non-invertible symmetry distinguishes these phases. We have

$$D|\psi\rangle = (-1)^{(L-L')/4}\big(D^{(1)}_{\text{L}} D^{(1)}_{\text{R}} + D^{(2)}_{\text{L}} D^{(2)}_{\text{R}}\big),$$

with the non-trivial algebra $\eta^o_{\text{L}} D^{(I)}_{\text{L}} = -D^{(I)}_{\text{L}} \eta^o_{\text{L}}$ and $\eta^o_{\text{R}} D^{(I)}_{\text{R}} = -D^{(I)}_{\text{R}} \eta^o_{\text{R}}$. This forces the ground state to be degenerate, meaning that there must be some modes at the boundary between the two halves of the spin chain.

Interestingly the $0+1$ dimensional quantum mechanics at each boundary has a projective action of an invertible group $\mathbb{Z}_2 \times \mathbb{Z}_2$, no longer anything non-invertible. This need more explanations.

Kennedy–Tasaki transformation. We have $\text{Rep}(D_8)$-SPTs. What happens under the Kennedy–Tasaki transformation? We know that we have $\widehat{\eta}^e$, $\widehat{\eta}^o$ from the $\mathbb{Z}_2 \times \mathbb{Z}_2$ Wilson lines (dual symmetry from gauging). On top of that we have a new operator $\widehat{V}$:

$$\begin{aligned}
\widehat{\eta}^e &= \prod_{j \text{ even}} \widehat{X}_j \\
\widehat{\eta}^o &= \prod_{j \text{ odd}} \widehat{X}_j \\
\widehat{V} &= \prod_{j} \widehat{\text{CZ}}_{j,j+1}
\end{aligned}$$

It turns out that all three $\widehat{\eta}^e, \widehat{\eta}^o, \widehat{V}$ all commute, it is a $(\mathbb{Z}_2)^3$ symmetry. This (invertible) symmetry has an anomaly classified by $H^3(\mathbb{Z}_2, U(1))$, called “type III anomaly” whose topological action is $e^{i\pi \int a\cup b\cup c}$ where a, b, c are background fields for these three symmetries.

We can see this in the other direction since gauging a discrete symmetry can always be undone by gauging the corresponding quantum symmetry. So H_{cluster} is obtained by gauging $\widehat{\mathbb{Z}}_2^e \times \widehat{\mathbb{Z}}_2^o$. This is a situation we have seen many times: starting from a theory with a mixed anomaly (here $(\mathbb{Z}_2)^3$), gauging part of the symmetry (here $(\mathbb{Z}_2)^2$) will give a non-invertible symmetry.

Symmetry breaking pattern. Consider all three $\text{Rep}(D_8)$-SPT states $|\text{cluster}\rangle$, $|\text{even}\rangle$, $|\text{odd}\rangle$. After the Kennedy–Tasaki transformation, we know (from the $\mathbb{Z}_2 \times \mathbb{Z}_2$ discussion) that $\widehat{\mathbb{Z}}_2^e \times \widehat{\mathbb{Z}}_2^o$ is broken. This does not tell us which part of the full $(\mathbb{Z}_2)^3$ is unbroken. There seems to be four options:

- $\widehat{\mathbb{Z}}_2^V$ could be unbroken, this is achieved by (the Kennedy–Tasaki image of) the cluster state;

- the diagonal subgroup of $\widehat{\mathbb{Z}}_2^o \times \widehat{\mathbb{Z}}_2^V$ could be unbroken, this is achieved by (the Kennedy–Tasaki image of) the even state;
- the diagonal subgroup of $\widehat{\mathbb{Z}}_2^e \times \widehat{\mathbb{Z}}_2^V$ could be unbroken, this is achieved by (the Kennedy–Tasaki image of) the odd state;
- the diagonal subgroup of the whole $(\mathbb{Z}_2)^3$ could be unbroken, but actually this generator (generator of the CZX symmetry) is subject to the Levin–Gu anomaly, so it cannot be unbroken.