Lectures on Generalized Symmetries

T. Daniel Brennan
University of Birmingham

These lecture notes are for the 2026 *Fudan Particle Physics Summer School. They are under-construction and may contain mistakes or be missing some references.*

Contents

1 Lecture 1 — Group-like Global Symmetries

Goal and philosophy. These lectures aim to give the tools needed to understand the new literature on "generalized symmetries."[1] The guiding philosophy is two-fold: to teach you the basics in depth (better equipps you for research) and to introduce you to some of the more advanced topics, to give you a starting point for your own investigation. The subject organizes into a hierarchy (Figure 1): group-like higher-form symmetries sit inside higher-group symmetries, which sit inside fully categorical symmetries; only the innermost layer is group-like. We will focus mostly on higher-form symmetries, and sketch the outer layers in the final lecture.

[1] For more details, one can see the reviews [5, 1, 2, 3, 4].

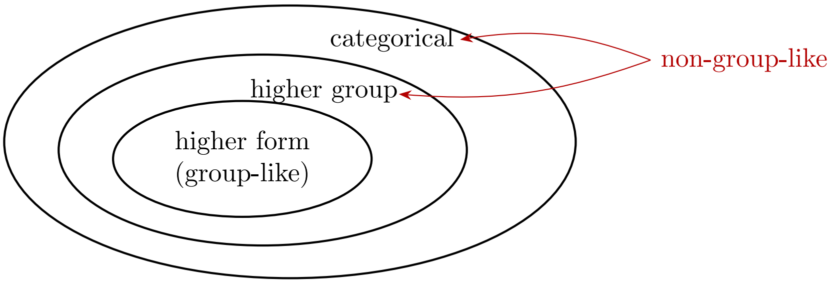

Figure 1: The hierarchy of generalized symmetries.

1.1 Ordinary symmetries, three ways

Consider a d-dimensional field theory with fields $\phi(x)$ acted on by a continuous Lie group G.

$$\text{Classical:} \qquad \phi(x) \to g \cdot \phi(x) \ , \quad g = e^{i\alpha} \tag{1}$$

is a symmetry of the action, and Noether's theorem gives $\partial_\mu j^\mu(x) = 0$, derived by imposing invariance of the action under coordinate-dependent transformations

$$\delta_{\alpha(x)} S = \int \alpha(x)\, \partial_\mu j^\mu(x)\, d^d x = 0 \ . \tag{2}$$

In terms of differential forms, the conservation equation can be written as

$$\partial_\mu j^\mu(x) = 0 \quad \Longleftrightarrow \quad d \star j(x) = 0 \quad , \quad \star j \in \Omega^{d-1}(M) \tag{3}$$

For a review on differential forms, see Appendix A). The resulting charge $Q = \oint_\Sigma * j$ is a constant of classical motion.

In the *quantum* theory, the current conservation becomes the operator equation

$$\partial_\mu j^\mu(x)\, \mathcal{O}_q(y) = q\, \delta^{(d)}(x-y)\, \mathcal{O}_q(y) \ . \tag{4}$$

In terms of the Hilbert-space, a spatial (Cauchy) slice Σ carries a Hilbert space $\mathcal{H}_\Sigma$; assuming a time direction that foliates spacetime into slices Σ_t,

$$\mathcal{H}_{\Sigma_t} = \mathcal{H}_t \ , \qquad U(\Delta t) : \mathcal{H}_t \to \mathcal{H}_{t+\Delta t} \ . \tag{5}$$

G-symmetry means that the group acts on the Hilbert space as

$$U_\alpha = e^{i\alpha Q_\Sigma} \ , \qquad U_\alpha U_\beta = U_{\alpha+\beta} \ , \tag{6}$$

so that $\mathcal{H}_\Sigma$ decomposes into G-representations (superselection sectors). Since $[Q_\Sigma, \hat{H}] = 0$, time evolution preserves the G-representation,

$$U_\alpha\, U(\Delta t) = U(\Delta t)\, U_\alpha \ , \tag{7}$$

which follows from Baker–Campbell–Hausdorff; pictorially, the symmetry operator is a sheet that can be pushed through the time slices (Figure 2).

Conservation buys more than commutation with the Hamiltonian[2] — it makes the operator deformable:

$$e^{i\alpha \oint_\Sigma \star j} \cdot e^{-i\alpha \oint_{\Sigma+\delta\Sigma} \star j} = e^{i\alpha\left(\oint_\Sigma - \oint_{\Sigma+\delta\Sigma}\right)\star j} = \exp\left[i\alpha \oint_{\delta\Sigma} d \star j\right] = 1 \ , \tag{8}$$

so $e^{i\alpha Q_\Sigma}$ is invariant under deformations of Σ.

[2]Note that topologically invariant implies commutativity with the stress tensor.

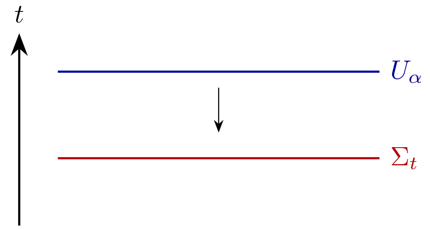

Figure 2: The symmetry operator U_α pushed through time slices.

1.2 Lorentzian vs. Euclidean spacetime

Spacetime is Lorentzian, but it is often useful to work in the Euclidean formulation. A state $|\psi\rangle \in \mathcal{H}_\Sigma$ does not care whether Σ bounds a Lorentzian or Euclidean manifold — quantization yields the same Hilbert space — so we may study $\mathcal{H}_\Sigma$ in Euclidean signature. In fact $|\psi\rangle$ can be *prepared* by the Euclidean path integral over a manifold M_d with operators inserted and $\partial M = \Sigma$ (Figure 3):

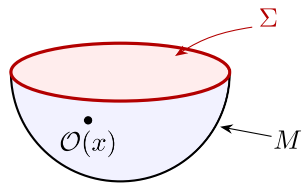

Figure 3: A state in $\mathcal{H}_\Sigma$, $|\mathcal{O}(x)\rangle$ can be prepared by the Euclidean path integral over M with $\partial M = \Sigma$ and with an insertion of the operator $\mathcal{O}(x)$ for $x \in M$.

Important: there is no preferred time direction, so the choice of Σ is arbitrary. For example, in 2d $\mathbb{R}^2$ one may quantize on lines of constant y *or* on concentric circles (radial quantization) — a choice that is important for CFT. Since the slice is arbitrary, so is the surface supporting the charge, and that is the modern starting point.

Punchline: in Euclidean QFT, $U_\alpha(\Sigma)$ is a *topological operator* for every $\Sigma \subset M$; if Σ is contractible, $U_\alpha(\Sigma) \to \mathbb{1}$ (squashing a small sphere flattens it onto two coincident sheets supporting $U_{-\alpha}U_\alpha \sim \mathbb{1}$; Figure 4).

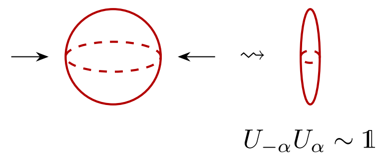

Figure 4: Topological operators at work. Left: a contractible U_α squashes onto two oppositely oriented coincident sheets, $U_{-\alpha}U_\alpha \sim \mathbb{1}$.

To find the action of U_α on local operators, sweep the sheet past an insertion (Figure 5): sandwiching $\mathcal{O}(x)$ between U_α and U_α^{-1} on neighbouring slices Σ_2, Σ_1 (Figure 5) gives

$$U_\alpha\, \mathcal{O}(x)\, U_\alpha^{-1} = e^{i\alpha\left(\oint_{\Sigma_2} \star j - \oint_{\Sigma_1} \star j\right)}\, \mathcal{O}(x) = e^{i\alpha \oint_{\delta\Sigma} d\star j}\, \mathcal{O}(x) = e^{i\alpha q}\, \mathcal{O}(x)\ . \tag{9}$$

1.3 Background gauge fields

An important tool for analyzing symmetries is coupling to background gauge fields.

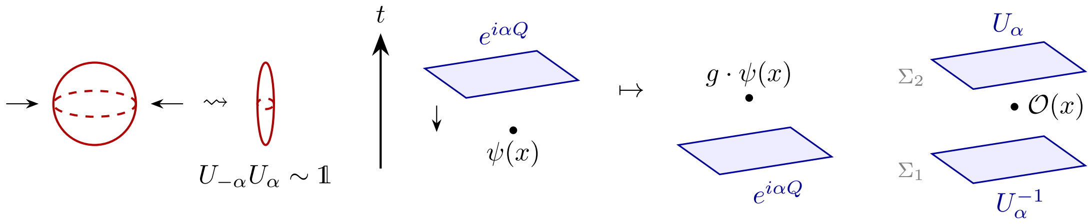

Figure 5: Topological operators at work. Left: a contractible U_α squashes onto two oppositely oriented coincident sheets, $U_{-\alpha}U_\alpha \sim \mathbb{1}$. Middle: sweeping the sheet down past the insertion turns $\psi(x)|0\rangle$ into $g \cdot \psi(x)|0\rangle$. Right: the sandwich computing (9).

Question: what does it mean to gauge a symmetry, and how is that related to coupling to a background gauge field? Consider $U(1)$ gauge theory,

$$Z = \int_{\mathcal{A}/\mathcal{G}} [dA]\, Z_{\text{matter}}[A] \ , \qquad Z_{\text{matter}}[A] = \int [d\Phi]\, e^{-S[\Phi,A]} \ , \tag{10}$$

where $\mathcal{A}/\mathcal{G}$ is the space of gauge fields modulo gauge equivalence. It is sometimes claimed that coupling to a background gauge field is the same as the $e^2 \to 0$ limit. This is incorrect: gauging is not just coupling to a background field, because gauging always imposes gauge invariance no matter what the gauge field does. In QED, for instance, $\psi\psi(x)$ is not a gauge-invariant local operator, yet it is a perfectly good operator in the theory where the gauge field is a background. The difference is whether the gauge symmetry is "real" or fake: in the $e^2 \to 0$ limit the symmetry is fake, while with a background gauge field the would-be gauge symmetry is a genuine global symmetry.

Concretely, for a continuous symmetry one couples the background gauge field to j (in vector notation $i\int A_\mu j^\mu d^dx$) so that the action is invariant under background gauge transformations,

$$S = \cdots + i\int A \wedge \star j \ , \qquad A \to A + d\lambda : \quad S \to S + i\int d\lambda \wedge \star j = i\int \lambda\,(d\star j) = 0 \ , \tag{11}$$

so that $Z[A]$ exists and symmetry preservation is equivalent to background gauge invariance. One can also recover the symmetry operator itself: inserting $U_\alpha(\Sigma)$ shifts $S \to S + i\alpha\oint_\Sigma \star j$, and using the Poincaré dual $\delta(\Sigma) \in \Omega^1(M)$ of the $(d-1)$-manifold Σ (the Thom class — morally a δ-function),

$$i\alpha\oint_\Sigma \star j = i\alpha\int_M \delta(\Sigma)\wedge\star j \ \sim\ i\int_M (\alpha\,\delta(\Sigma))\wedge\star j \ , \qquad A = \alpha\,\delta(\Sigma) \sim d\big(\alpha\,\Theta(\Sigma)\big) \ , \tag{12}$$

a singular flat background, with $\Theta(\Sigma)$ the global angular form.

Q: what is the utility of background gauge fields?

A: they detect anomalies and generalized symmetries.

Ex: For a massless Dirac fermion in 4d with $U(1)_v \times U(1)_a$ and partition function $Z[A_v, A_a]$,

$$Z[A_v + d\lambda_v,\ A_a + d\lambda_a] = Z[A_v, A_a] \times \exp\left\{ i\int \lambda_a \left(\frac{F_a \wedge F_a}{8\pi^2} + \frac{F_v \wedge F_v}{8\pi^2} \right) \right\} \ , \tag{13}$$

the familiar triangle diagram (Figure 6). As we will see, this technology is very useful.

2 Lecture 2 — Higher-Form Symmetries and Gauge Theory

Consider a totally antisymmetric conserved current $J^{[\mu_1\cdots\mu_{p+1}]}$:

$$\partial_\mu J^{[\mu\mu_1\cdots\mu_p]}(x) = 0 \quad \Longrightarrow \quad d\star J = 0 \ , \qquad \star J \in \Omega^{d-p-1}(M) \ , \tag{14}$$

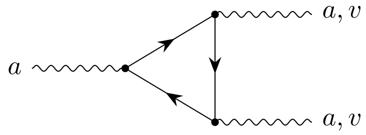

Figure 6: The anomaly triangle for the axial current.

and define $U_\alpha(\Sigma) = e^{i\alpha \oint_\Sigma \star J}$, where now Σ has codimension $p+1$. Exactly as before,

$$U_\alpha(\Sigma)\, U_{-\alpha}(\Sigma + \delta\Sigma) = e^{i\alpha\left(\oint_\Sigma - \oint_{\Sigma+\delta\Sigma}\right)\star J} = e^{i\alpha \int_{\delta\Sigma} d\star J} = 1 \ , \tag{15}$$

so $U_\alpha(\Sigma)$ is topological. Is this a symmetry? Yes — a *p-form symmetry*, acts on p-dimensional operators; a 0-form symmetry is an "ordinary" symmetry [6].

2.1 Example: 4d Maxwell theory

Take $S = \int \frac{1}{2e^2} F \wedge \star F$. The two identities

$$d\Big(\underbrace{\frac{F}{2\pi}}_{\star J_m}\Big) = 0 \ , \qquad d\Big(\underbrace{\star\frac{F}{e^2}}_{\star J_e}\Big) = 0 \tag{16}$$

give, since $F \in \Omega^2(M)$, two 1-form symmetries: $U(1)_e^{(1)} \times U(1)_m^{(1)}$. The magnetic current acts on 't Hooft (monopole) lines and the electric current on Wilson lines: $\oint_\Sigma \star J_m = \oint_\Sigma \frac{F}{2\pi}$ measures the magnetic flux through Σ, counting the monopoles it encloses, and unlinking the symmetry surface from a monopole line produces the phase $e^{i\alpha p}$ (Figure 7).

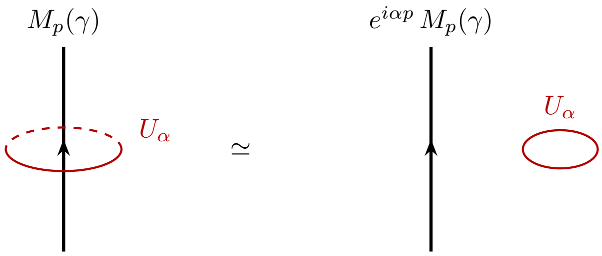

Figure 7: The 1-form symmetry operator acts on line operators by linking.

The electric current is nothing but the momentum conjugate to the gauge field, $\star J_e = \Pi_A$;[3] the operator $e^{i\alpha \oint_\Sigma \star F/e^2}$ shifts the dynamical gauge field, $\delta S = -\int \alpha\, \delta(\Sigma) \wedge \star \frac{F}{e^2}$, i.e. U_α acts by linking. Equivalently one shifts $A \to A + \alpha\, \Theta(\Sigma)$; moving $\Sigma \to \Sigma + \delta\Sigma$ gives $\Delta A = \alpha\, d\Theta(\delta\Sigma) = \alpha\, \delta(\delta\Sigma)$, which shifts Wilson loops $W = e^{in \oint A} \to e^{in\alpha} W$.

We can couple both symmetries to background 2-form gauge fields,

$$S = \int \frac{1}{2e^2}(F - B_e)^2 + i \int B_m \wedge \star J_m \ , \qquad \delta B_{e,m} = d\Lambda_{e,m} \ . \tag{17}$$

However, $\star J_m = \frac{F}{2\pi}$ shifts under δ_e, but improving $\star J_m \to (\star J_m - B_e)$ ruins invariance under δ_m:

$$\delta_m S = i \int d\Lambda_m \wedge \Big(\frac{F}{2\pi} - B_e\Big) = i \int \Lambda_m \wedge dB_e \ . \tag{18}$$

[3] Note that the choice of momentum is dependent on the choice of Cauchy slice. Here, the hodge dual makes it so that when you restrict $\star J_e\big|_\Sigma$ that it will be Π_A with respect to the choice of Σ.

Why: a mixed anomaly,

$$Z[B_e + d\Lambda_e,\ B_m + d\Lambda_m] = Z[B_e, B_m]\, e^{i\oint \Lambda_m \wedge dB_e}\ , \tag{19}$$

with 5d inflow $S_{5d} = i\int B_m \wedge dB_e \sim i\int B_e \wedge dB_m$, the two presentations related by the local counterterm $S_{c.t.} = i\int B_m \wedge B_e$.

2.2 Example: 4d periodic scalar

Take $S = \int \frac{d\phi\wedge\star d\phi}{2}$ with $\phi \sim \phi + 2\pi$. The equation of motion gives $d(\star d\phi) = 0$ and the Bianchi identity $d(\frac{d\phi}{2\pi}) = 0$, so

$$\star d\phi = \star j \in \Omega^3(M)\ \to\ \text{0-form symmetry}\ ,\quad d\phi = \star J \in \Omega^1(M)\ \Rightarrow\ \text{2-form symmetry}\ , \tag{20}$$

the latter acting on strings. A cosmic string $S(\Gamma)$ on a $2d$ surface Γ with transverse $\mathbb{R}^2$ denotes the disorder operator around which ϕ winds n times; $e^{i\alpha\oint\star J}$ measures this winding, acting on the string by linking exactly as in the monopole case.

Much like Maxwell theory, $U(1)^{(0)}$ shifts $\phi \to \phi + \alpha$ while $U(1)^{(2)}$ measures winding, and coupling to backgrounds again requires a choice:

$$S = \int \frac{(d\phi - A)^2}{2} + i\int d\phi \wedge C_3\ ,\qquad \delta C_3 = d\Lambda_2\ ,\quad \delta A = d\lambda\ ,\ \phi \to \phi + \lambda\ , \tag{21}$$

$$\delta S = i\int d\lambda \wedge C_3 \quad \overset{c.t.}{\Longleftrightarrow} \quad \delta S = i\int A \wedge d\Lambda_2\ . \tag{22}$$

Again there is an anomaly:

$$Z[A + d\lambda,\ C + d\Lambda] = Z[A, C] \times e^{i\int d\lambda\wedge C_3}\ ,\qquad \mathcal{A}_{5d} = i\int A_1 \wedge dC_3 \simeq i\int dA_1 \wedge C_3\ . \tag{23}$$

General pattern: a $U(1)^{(p)}$ "electric" symmetry (shifting the field) pairs with a $\tilde{U}(1)^{(d-p-2)}$ "magnetic" symmetry (acting on the winding defect).

The road map for the rest of the course runs through discrete symmetries (Figure **??**): the most interesting higher-form symmetries in physics are discrete, and they live in non-abelian gauge theories.

2.3 Interlude: discrete symmetry and discrete gauge theory

So far the discussion covered only continuous (really $U(1)$) symmetries, but the paradigm holds for general groups, including discrete groups. For a 0-form symmetry the operators $U_g(\Sigma)$, $g \in G$, fuse as $U_g(\Sigma)\cdot U_h(\Sigma') = U_{gh}(\Sigma)$ and act on charged local operators by $U_g(\Sigma)\cdot\mathcal{O}_R(x) = R(g)\cdot\mathcal{O}(x)\,U_g(\Sigma)$. For $p >$ 0-form symmetries, $U_g(\Gamma)U_h(\Gamma') \simeq U_h(\Gamma')U_g(\Gamma)$: operators of codimension greater than one can be deformed past each other, so higher-form symmetry groups are abelian. For the rest of the lecture we focus on $G^{(p)} = \mathbb{Z}_N^{(p)}$.

What is the background gauge field for a discrete symmetry? A $\mathbb{Z}_N$ gauge field is tricky: a "smooth" $\mathbb{Z}_N$ gauge field A means $e^{i\oint_\gamma A} = e^{2\pi i n/N}$, and smoothness forces the holonomy to depend only on the topological class of γ — there are no fluxes, in fact $dA = F = 0$. Geometrically, the gauge field describes an N-fold cover of spacetime, with charged matter transforms by a phase when shifting between sheets.

As physicists it is more useful to build $\mathbb{Z}_N$ gauge theory from what we know: since $\mathbb{Z}_N \subset U(1)$, we can Higgs $U(1) \to \mathbb{Z}_N$. Take 4d $U(1)$ gauge theory with a charge-N scalar (Abelian-Higgs Model)

$$S = \int \frac{1}{2e^2}F^2 + |D\Phi|^2 - V(\Phi)\ ,\qquad D\Phi = d\Phi - iNA\,\Phi\ , \tag{24}$$

so that $\alpha = \frac{2\pi k}{N}$ leaves Φ invariant and $\mathbb{Z}_N \subset U(1)$ survives. Note that the equation of motion $d\star\frac{F}{e^2} + N\star j = 0$ with $\oint \star j \in \mathbb{Z}$ makes the operator

$$e^{\frac{2\pi i n}{N}\left(\oint_\gamma \star\frac{F}{e^2} + N\int_\Gamma \star j\right)} \tag{25}$$

depend only on the topological class of γ: the electric symmetry is broken $U(1)^{(1)}_e \to \mathbb{Z}^{(1)}_N$. Taking $V(\Phi) = \lambda(|\Phi|^2 - v^2)^2$ condenses $\langle|\Phi|^2\rangle = v^2$ and spontaneously breaks $U(1) \rightsquigarrow \mathbb{Z}_N$. Writing $\Phi = (\rho+v)e^{i\phi}$ (ρ massive, ϕ eaten),

$$S = \int \frac{1}{2e^2}F^2 + v^2(d\phi - NA)^2 \ . \tag{26}$$

Something we can do to simplify is to dualize the field ϕ to a 2-form gauge field B_2 ($\delta B_2 = d\Lambda_1$, with $e^{i\oint B_2} \in U(1)$ and $\oint \frac{dB_2}{2\pi} \in \mathbb{Z}$). The way this is implemented is by making ϕ a $\mathbb{R}$-valued field (instead of a $U(1)$-valued field) and then using B_2 as a Lagrange multiplier to impose the periodicity on ϕ. This is accomplished by introducing the coupling to the action

$$S = ...2\pi i \int \frac{d\phi}{2\pi} \wedge \frac{dB_2}{2\pi} \tag{27}$$

We can now complete the square

$$S = \int \frac{1}{2e^2}F^2 + v^2(d\phi - NA)^2 + \frac{i}{2\pi}\int (d\phi - NA)\wedge dB_2 + \frac{iN}{2\pi}\int A\wedge dB \ . \tag{28}$$

and integrate out ϕ. Doing so imposes $A = \frac{d\phi}{N}$ so $F = 0$, leaving

$$S = \frac{iN}{2\pi}\int A\wedge dB \ \simeq \ \frac{iN}{2\pi}\int B\wedge F \tag{29}$$

which is also known as $\mathbb{Z}_N$ BF theory, the prototype $\mathbb{Z}_N$ gauge theory. In this theorythe gauge transformations are $\delta A = d\lambda$, $\delta B = d\Lambda_1$ and equations of motion are given by

$$\frac{N\,dA}{2\pi} = 0 \quad , \quad \frac{N\,dB}{2\pi} = 0 \tag{30}$$

We can see the $\mathbb{Z}_N$ discreteness manifestly from the partition function. Integrating over B imposes the B-equation of motion, and the path integral contains a sum over dB-flux,

$$Z = \int [dB][dA]\, e^{\frac{iN}{2\pi}\int A\wedge dB} \ \to \ \sum_{[dB]\in H^3}\int [dA]\, e^{\frac{iN}{2\pi}\int A\wedge[dB]} \ \to \ e^{iN\oint A} = 1 \ . \tag{31}$$

Alternatively, analyze BF theory directly as a TQFT through its operators $W_n = e^{in\oint_\gamma A}$ and $S_m(\Gamma) = e^{im\oint_\Gamma B}$:

$$\langle W_n(\gamma)\, S_m(\Gamma)\rangle = \int [dA\, dB]\ e^{\frac{iN}{2\pi}\int A\wedge dB + in\int A\wedge\delta(\gamma) + im\int B\wedge\delta(\Gamma)} \ . \tag{32}$$

Absorbing $A \mapsto A - \frac{2\pi m}{N}\Theta(\Gamma)$ and $B \mapsto B - \frac{2\pi n}{N}\Theta(\gamma)$,

$$\langle W_n(\gamma)\, S_m(\Gamma)\rangle = \langle \mathbb{1}\rangle \times e^{-\frac{2\pi i\, nm}{N}\int\Theta(\gamma)\wedge\delta(\Gamma)} = e^{-\frac{2\pi i\, nm}{N}\,\mathrm{Link}(\gamma,\Gamma)} \ . \tag{33}$$

Notice that this means that the surface operators $S_m(\Gamma)$ shift the gauge field A along the transverse space because it shifts the holonomy of A (measured by the expectation value of the Wilson line). Similarly, $W_n(\gamma)$ shifts the background gauge field B. Since both W_n, S_m are topological, they generate a symmetry:

$$\mathbb{Z}^{(1)}_N \times \mathbb{Z}^{(2)}_N \tag{34}$$

which are generated by $S_m(\Gamma)$ and $W_n(\gamma)$ respectively. As in the case of Maxwell theory, these symmetries are "dual" since the $\mathbb{Z}_N$ gauge theory of A_1, necessarily has B_2 which gives rise to the two symmetries above.

Also as in Maxwell theory, these symmetries have a mixed anomaly. We can again couple $\mathbb{Z}_N$ BF theory to background $\mathbb{Z}_N^{(1)} \times \mathbb{Z}_N^{(2)}$ background gauge fields as:[4]

$$S = \frac{iN}{2\pi}\int (dA_1 - b_2)\wedge B_2 - \frac{iN}{2\pi}\int A_1 \wedge c_3 \tag{35}$$

Again, since the action is not background gauge invariant

$$S = -\frac{iN}{2\pi}\int (\lambda_1 \wedge c_3 + b_2 \wedge \lambda_2)\ \mathrm{mod}_{2\pi\mathbb{Z}} \tag{36}$$

under the total background gauge transformations $\delta b_2 = d\lambda_1$ and $\delta c_3 = d\lambda_2$. Notice that this term cannot be cancelled by a local counter term. This anomaly is described by the anomaly 5-form:

$$\mathcal{A} = \frac{iN}{2\pi}\int b_2 \wedge c_3 \tag{37}$$

Lesson: Dual symmetries are $G^{(p)} \leftrightarrow \widetilde{G}^{(d-p-2)}$ for continuous symmetries and $G^{(p)} \leftrightarrow \widetilde{G}^{(d-p-1)}$ for discrete symmetries.

2.4 Non-abelian gauge theory

Now let us apply this to the study of non-abelian gauge theories. Consider 4d $SU(N)$ Yang–Mills theory. The gauge field transforms in the adjoint representation — much like the charge-N scalar above, not the minimal representation — so gluons cannot screen Wilson lines $W_R(\gamma)$ unless R appears in tensor powers $N^{\otimes nN}$, which is classified by how the center of the gauge group $Z(G)$ acts on the representation. This corresponds to a 1-form symmetry called *center symmetry*, $Z(G) = \mathbb{Z}_N^{(1)}$, the analog of the electric 1-form symmetry of Maxwell theory.

Conversely, the $PSU(N)$ theory has no center, $Z(PSU(N)) = 0$, but it has new monopole lines. Recall that a Dirac monopole is made up of (at least) two patches which on the northern/southern hemisphere take the form

$$A = \frac{1}{2}T(\mp 1 - \cos(\theta))d\phi \quad , \quad T \in Lie(G) = \mathfrak{g} \tag{38}$$

so that there is no singular Dirac string. The connection in the two patches must be related by a gauge transformation along the great circle where they overlap. These non-trivial gauge transformations are classified by element of $\pi_1(G)$. For $G = PSU(N)$, $\pi_1(PSU(N)) = \mathbb{Z}_N$ while $\pi_1(SU(N)) = 0$, and thus there are non-trivial (Dirac) monopole lines in $PSU(N)$ gauge theory that do not exist in $SU(N)$ gauge theory. [5] The two theories are related by gauging:
Q: what does it mean to gauge a discrete symmetry $\mathbb{Z}_N^{(p)}$?

In the $U(1)$ gauge theory, we summed over all possible gauge field configurations in the path integral. For $\mathbb{Z}_N$ gauge theory, summing over all background gauge fields is equivalent to summing over all insertions of symmetry operators — equivalently, makes the background gauge field dynamical (a sum over flat connections) — which projects out charged operators and produces a new dual symmetry $\tilde{\mathbb{Z}_N}^{(d-p-2)}$, generated by the "Wilson line" of the new gauge field.

[4]Here we are imagining that the background gauge fields b, c are flat $U(1)$ gauge fields that we have restricted to take values in $\mathbb{Z}_N$.

[5]Note: one can break $SU(N+1) \to PSU(N)$ by condensing $\langle |h|^2\rangle \neq 0$ where h transforms in the fundamental representation.

$$\begin{tikzcd} \boxed{SU(N)} \arrow[r, bend left, "\text{gauge } \mathbb{Z}_N^{(1)}", "\text{emerge } \widetilde{\mathbb{Z}_N}^{(1)}"'] & \boxed{PSU(N)} \arrow[l, bend left, "\text{gauge } \widetilde{\mathbb{Z}_N}^{(1)}", "\text{emerge } \mathbb{Z}_N^{(1)}"'] \end{tikzcd}$$

Figure 8: The gauging round trip between $SU(N)$ and $PSU(N)$.

Physically speaking, gauging the $\mathbb{Z}_N^{(1)}$ center of $SU(N)$ allows fractional flux in the gauge field, just as for $U(1) \to U(1)/\mathbb{Z}_N$ with $\mathbb{Z}_N^{(1)} \subset U(1)_e^{(1)}$:

$$F \to F + \underbrace{\frac{2\pi n}{N}\delta(\Gamma^2)}_{\text{fractional flux } w_2} \quad . \tag{39}$$

Consequence: a change in instanton-number quantization. For $U(1)$ on a spin manifold, $\int \frac{F\wedge F}{8\pi^2} \in \mathbb{Z}$; for $U(1)/\mathbb{Z}_N$,

$$I \;\to\; \int \frac{F\wedge F}{8\pi^2} + \frac{1}{N}\int \frac{F\wedge w_2}{2\pi} + \frac{1}{2N^2}\int w_2 \wedge w_2 \implies I \in \frac{1}{2N^2}\mathbb{Z} \; ; \tag{40}$$

and for $SU(N) \to PSU(N)$ [7, 8],

$$I \;\to\; I + \frac{N-1}{2N}\int w_2 \cup w_2 \qquad (w_2 \text{ summed over}) \; . \tag{41}$$

This will be important for axion physics, as we will see in Section 4.3.

3 Lecture 3 — Higher Symmetries

We are finally ready to discuss higher symmetries, guided by the correspondence [6]

$$\Big\{\text{topological operators}\Big\} \longleftrightarrow \Big\{\text{generalized symmetry}\Big\} \; . \tag{42}$$

The set of topological operators is generally controlled by the topological sector of the theory, and it is described by category theory — many different flavors of categories are used in many different contexts. There is no time for a full course, so we illustrate the idea of higher structure.

3.1 Categories in one page

A category $\mathcal{C}$ consists:

- of a set of objects $\mathrm{obj}(\mathcal{C})$
- a set of morphisms $\mathrm{Mor}(\mathcal{C})$ between objects, $F_{x,y} : X \to Y$ such that:
  - You can compose when composition makes sense

$$F_{x,y} : X \to Y \quad , \quad G_{y,z} : Y \to Z \quad \Longrightarrow \quad G_{y,z} \circ F_{x,y} : X \to Z \tag{43}$$

  - The composition is associative

$$H_{z,w} \circ (G_{y,z} \circ F_{x,y}) = (H_{z,w} \circ G_{y,z}) \circ F_{x,y} : X \to W \tag{44}$$

- There is an identity morphism for every object: $id_x : X \to X$

For example, the case obj$(\mathcal{C}) = \{*\}$ with all morphisms invertible is a group.

Categories can be extended by layers of structure: morphisms between morphisms (2-morphisms), morphisms between those (3-morphisms), and so on. A theory which has structure all the way up to n-morphisms is called an n-category.

In fact, a normal category is an n-category with the higher morphisms trivial, so any category is "trivially" an ∞-category.

Geometric/Physical Picture: take a theory with $\mathcal{C}$-symmetry and a single object assigned to the vacuum; then 1-morphisms are codimension-1 operators preserving the vacuum, 2-morphisms are codimension-2 operators that change the surface operators, 3-morphisms are codimension-3 operators, and so on.

3.1.1 Symmetry fractionalization

To illustrate, consider a 2-category with a single object and invertible 1-morphisms, but with $f_g \circ f_h \simeq f_{gh}$ only *up to a 2-morphism*: physically, two codimension-1 sheets g and h merge into gh along a codimension-2 seam — the 2-morphism — which can itself act on line operators much like a 1-form symmetry operator (Figure 9).

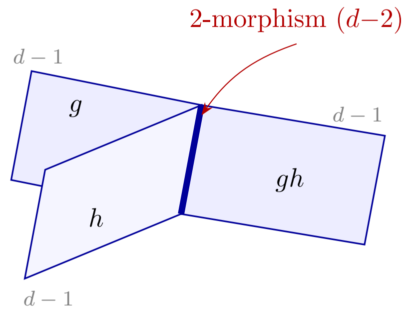

Figure 9: Two codimension-1 sheets g and h fusing into gh along a codimension-2 seam — the 2-morphism.

Junctions of the kind just drawn come in two physically important types.

- When the theory has both a 0-form and a 1-form symmetry, junctions can be dressed by 1-form symmetry operators, and this dressing encodes how the 0-form symmetry acts on local degrees of freedom [9, 10, 11].

  As an example, take 4d $SU(2)$ Yang–Mills with $2N_f$ fundamental fermions, an adjoint scalar and a Yukawa coupling, with flavor symmetry $G_F = SO(2N_f)/\mathbb{Z}_2$. Higgsing to electromagnetism leaves a $U(1)^{(1)}_m$ magnetic symmetry, with G_F preserved and acting trivially on local fields. But monopoles carry zero-mode degrees of freedom on their world-lines valued in spin$(2N_f)$, which requires junctions of G_F dressed by $U(1)^{(1)}_m$: the flavor symmetry is *fractionalized* through the magnetic 1-form symmetry.

- Even with no 1-form symmetry there can be non-trivial junctions [12].

  Take 4d $SU(N)$ QCD with N flavors, $G_F = \frac{SU(N)_L \times SU(N)_R}{\mathbb{Z}_N}$. The fundamental Wilson line links non-trivially with junctions of flavor symmetry defects (Figure 10): sliding a $g, h \to gh$ junction across the line, and comparing the two orders of action on the endpoint, shows that the endpoint quark transforms *projectively* under G_F. Because the junction is topological, the Wilson line's projective class can be tracked under RG flow.

Punchline: higher symmetry is opening Pandora's box — many things can happen, and the full implications are still being worked out.

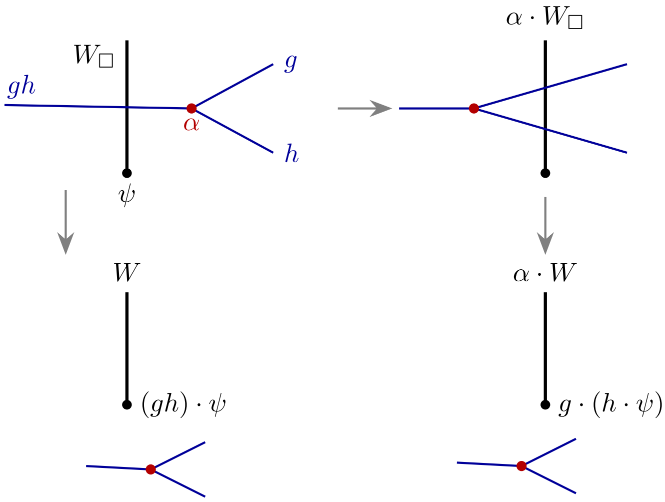

Figure 10: Sliding a $g, h \to gh$ junction past the Wilson-line endpoint in two orders: passing the junction through the line first (right column) gives $\alpha \cdot W$ ending on $g \cdot (h \cdot \psi)$, while sweeping the fused gh defect straight down (left column) gives W ending on $(gh) \cdot \psi$. Hence $(gh) \cdot \psi = \alpha\, g \cdot (h \cdot \psi)$: the endpoint transforms projectively under G_F.

3.2 Higher groups

An important example of higher symmetry is a *higher group*: an n-group is an n-category with a single object where all k-morphisms are invertible. Unfortunately, there is no unique definition of an n-group, so we take a physicist's approach.

Higher groups can in a sense be organized by the homotopy groups $\pi_k(\mathcal{C})$. A higher group decomposes into "component groups" — the k-morphisms preserving the $(k-1)$-identity — which geometrically are free codimension-k topological operators living inside the $(k-1)$-identity surface, inside the $(k-2)$-identity, and so on; the higher structure emerges only when looking at junctions, junctions of junctions, etc.

The *homotopy hypothesis* associates to any $\mathcal{C}$ a topological space X with, for finite symmetry, $\pi_k(\mathcal{C}) = \pi_k(X)$.

Classifying spaces.[6] To make the correspondence with spaces concrete we introduce the gadget BG, called the *classifying space* of G, which classifies the topological properties of G-gauge theory. It is built by constructing a non-contractible loop $\gamma_g \subset BG$ to every $g \in G$. A map $f : M \to BG$ then associates $g \in G$ to every loop in spacetime and defines a G-gauge field via the holonomy. In fact any principal G-bundle arises as the pullback $P = f^*(EG)$ of the universal bundle (Appendix A.3).

This is useful in two ways:

1. Mathematically, $H^*(BG, \mathbb{Z})$ classifies topological invariants — e.g. the instanton number sits in $H^4(BG, \mathbb{Z})$ for smooth Lie Groups — and is especially sensitive to discrete effects.
2. Symmetry-wise it works for discrete G: $\pi_1(BG) = G$ with $\pi_{k\geq 2}(BG) = 0$. The construction generalizes to higher-form symmetries (G abelian): B^kG has $\pi_k(B^kG) = G$, and a map $f : M \to B^kG$ identifies a G-gauge field B_{k+1} with $e^{i\oint B_{k+1}} \in G$.

Point: given an n-group $\mathcal{C}$, one builds $X_{\mathcal{C}}$ out of the blocks $B^{k+1}G_k$ where $\pi_{k+1}(\mathcal{C}) = G_k$ is the "k-form symmetry."

Postnikov towers and a 2-group example. Consider massless 4d QED with two flavors, $G_0 = SU(2)_L \times SU(2)_R$, with charges

[6] For further details, refer to [13, 14].

	$U(1)_g$	$SU(2)_L$	$SU(2)_R$
ψ_+	$+1$	$\square$	1
ψ_-	-1	1	$\square$

The flavor–flavor–gauge triangle (coefficient 1) is not a gauge anomaly, but it involves the gauge field:

$$\delta S = i\int \left(\frac{\mathrm{Tr}[\lambda_L F_L]}{8\pi^2} + \frac{\mathrm{Tr}[\lambda_R F_R]}{8\pi^2}\right) \wedge F_g \ . \tag{45}$$

The theory also has $G_1 = U(1)^{(1)}_m$ with $\star J = \frac{F}{2\pi}$, coupled as $S = \cdots + i\int B_2 \wedge \frac{F_g}{2\pi}$. The anomalous variation is cancelled by modifying

$$\delta B_2 = d\Lambda_1 - \frac{\mathrm{Tr}[\lambda_{L/R} F_{L/R}]}{4\pi} \ , \tag{46}$$

so the 0-form and 1-form symmetries fuse into a *2-group*. In the classifying-space picture the fibration

$$\begin{array}{ccc} B^2U(1) & \longrightarrow & X_{\mathcal{C}} \\ & & \downarrow \\ & & BG_0 \end{array} \tag{47}$$

does not split: a background flavor transformation (motion in BG_0) induces a transformation in $B^2U(1)$.

More generally, an n-group can be described by a *Postnikov tower*

$$\begin{array}{ccc} B^nG_{n-1} & \longrightarrow & X_{\mathcal{C}} \\ & & \downarrow \\ B^{n-1}G_{n-2} & \longrightarrow & X_{-1} \\ & & \downarrow \\ & & \vdots \\ & & \downarrow \\ B^2G_1 & \longrightarrow & X_{-n+2} \\ & & \downarrow \\ & & BG_0 \ , \end{array} \tag{48}$$

with backgrounds $\{A_1, B_2, \ldots, C_n\}$ transforming as

$$\delta A_1 = d\lambda \ , \qquad \delta B_2 = d\Lambda_1 + \omega_2(\lambda, A_1) \ , \qquad \delta C_3 = d\Lambda_2 + \omega_3(\lambda, \Lambda_1, B_2, A_1) \ , \ \ldots \tag{49}$$

where the ω_n are classified by cohomology classes called *Postnikov classes.*

Physical consequences. First, a restriction on symmetry breaking: since G_q transformations are turned on by $G_{p\le q}$ backgrounds, breaking G_q must necessarily include breaking $G_{p\le q}$. Second, a nested structure

$$G_n = G_n^{(0)} \supseteq G_n^{(1)} \supseteq \cdots \supseteq G_n^{(n-1)} = G_{n-1} \tag{50}$$

which an emergent G_n symmetry must respect along the RG flow (Figure 11, left):[7]

$$E_p \gtrsim E_q \qquad p > q \ . \tag{51}$$

[7]Note that 2-group global symmetries are actually not "compatible" with interacting $4d$ CFTs. Because of the constraints of conformal representation theory, you can show that interacting CFTs in $4d$ cannot have 2-group global symmetries – the 1-form symmetries decouple from the dynamics.

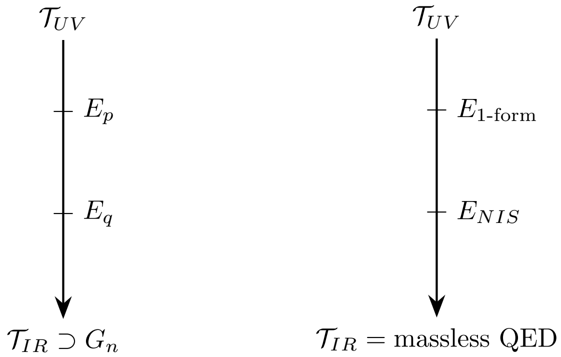

Figure 11: Left: emergent higher-group symmetry respects the nesting, $E_p \gtrsim E_q$ for $p > q$. Right: the corresponding scale ordering for the non-invertible chiral symmetry of Lecture 4.

4 Lecture 4: SymTFT, Non-Invertible Symmetries, and Axions

4.1 SymTFT

The organizing idea is: symmetry category → TQFT in one higher dimension.[8] The full symmetry category of a d-dimensional QFT can be physically realized by a TQFT on the slab $Y_{d+1} = M_d \times [0,1]_t$ — the *symmetry TQFT*, or SymTFT [16]. See Figure 12.

In this construction, the physical QFT lives on the boundary at $t = 0$: every symmetry operator U_g of the QFT corresponds to a topological operator $\widehat{U}_g$ of the bulk pushed against that boundary, while a charged operator $\mathcal{O}(\Gamma)$ lives at the end of a one-higher-dimensional topological operator L_γ stretched across the slab, $\partial\gamma = \Gamma_{t=0} \cup \overline{\Gamma}_{t=1}$. The coupling is consistent only if the action of the symmetry on charged operators is reproduced by the braiding and fusion of the purely topological bulk operators — the SymTFT is a physical realization of the symmetry category of the QFT.

Since we are working in Euclidean signature, we can choose to quantize the theory along the interval so that the two boundaries are boundary conditions on the space of fields, or transverse to the interval (i.e. at fixed time t) so that the boundary conditions are really just states in the SymTFT. With this duality in mind, we will refer to the physical boundary by the state $\langle QFT|$ and the other boundary by $|\mathfrak{B}\rangle$.

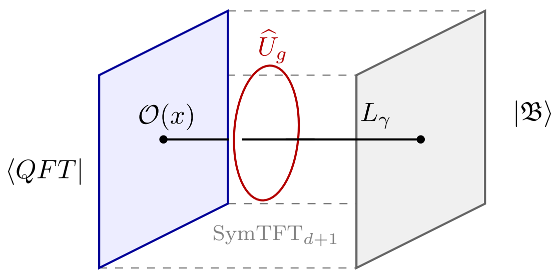

Figure 12: The SymTFT: the physical QFT lives on the left boundary $\langle QFT|$ and the topological "quiche" boundary condition sits on the right, $|\mathfrak{B}\rangle$. Symmetry operators of the QFT lift to bulk topological operators $\widehat{U}_g$, and charged operators sit at the ends of bulk topological lines L_γ, on which the $\widehat{U}_g$ act by linking.

The Quiche boundary. The construction requires a choice of *topological* boundary condition $\mathfrak{B}$ at $t = 1$ — the "quiche" — and this choice determines how the symmetry category is realized in the QFT: for a non-anomalous discrete symmetry G there are boundary conditions that couple the QFT to a background G-gauge field, and others that gauge G (or a subgroup). Quantizing

[8] This discussion follows and uses images from [15].

along M_d, with t as time, turns boundary conditions into states $|\mathfrak{B}\rangle$ of the SymTFT Hilbert space. The QFT boundary is subtler — it must encode the partition function of a generically non-topological theory — and in an orthonormal basis $|A\rangle$ it is the state

$$\langle QFT| = \sum_A Z_{QFT}[A]\,\langle A| \ . \tag{52}$$

The partition function is a *vector*: the QFT is a relative theory, and each pairing $\langle QFT|\mathfrak{B}\rangle$ produces an absolute d-dimensional theory.

Example: $\mathbb{Z}_N^{(0)}$. For a $\mathbb{Z}_N$ 0-form symmetry the SymTFT is the $\mathbb{Z}_N$ BF theory of the interlude (§2.3), now in $(d+1)$ dimensions,

$$S = \frac{iN}{2\pi}\int_{Y_{d+1}} a_1 \wedge db_{d-1} \ , \tag{53}$$

with topological operators

$$W_n(\gamma) = e^{in\oint_\gamma a_1} \ , \qquad U_m(\Sigma) = e^{im\oint_\Sigma b_{d-1}} \ , \tag{54}$$

generating $\mathbb{Z}_N^{(d-1)} \times \mathbb{Z}_N^{(1)}$ symmetries of the SymTFT, with

$$\langle W_n(\gamma)\,U_m(\Sigma)\rangle = e^{-\frac{2\pi i\,nm}{N}\,\mathrm{Link}(\gamma,\Sigma)} \tag{55}$$

by the linking computation of §2.3. Choose the orthonormal basis of Dirichlet boundary conditions for a_1: states $|A\rangle$ with $a_1\big|_\partial = \frac{2\pi}{N}A$ a flat $\mathbb{Z}_N$ gauge field. The Wilson lines W_n then take definite boundary values, while the U_m shift the boundary field,

$$e^{im\oint_\Sigma b_{d-1}}\,|A\rangle = \big|\,A + m\,\Theta(\Sigma)\,\big\rangle \ , \tag{56}$$

with $\Theta(\Sigma)$ the global angular form from Lecture 1. In this basis

$$\langle QFT| = \frac{1}{\sqrt{|H^1(M_d;\mathbb{Z}_N)|}}\sum_{A\in H^1(M_d;\mathbb{Z}_N)} Z_{QFT}[A]\,\langle A| \ , \tag{57}$$

so $\langle QFT|A\rangle = Z_{QFT}[A]$ is the partition function coupled to a fixed background. The same bulk admits Neumann boundary conditions for a_1 (Dirichlet for b_{d-1}),

$$|N_B\rangle = \frac{1}{\sqrt{|H^1(M_d;\mathbb{Z}_N)|}}\sum_{A\in H^1(M_d;\mathbb{Z}_N)} e^{\frac{2\pi i}{N}\int_{M_d} A\cup B}\,|A\rangle \ , \tag{58}$$

with B a fixed background for the dual $\mathbb{Z}_N^{(d-2)}$ symmetry, and then $\langle QFT|N_B\rangle \propto \sum_A Z_{QFT}[A]\,e^{\frac{2\pi i}{N}\int A\cup B}$ is the partition function of the *gauged* theory of Lecture 2. Coupling to backgrounds and gauging are simply different gapped boundary conditions of one bulk TQFT.

4.2 Categorical symmetry: the non-invertible chiral symmetry

Now we will discuss an important type of categorical symmetry for particle physics: the non-invertible chiral symmetry [17, 18].

Consider 4d massless QED

$$S = \int \left(\frac{1}{2g^2}F^2 + i\bar{\Psi}\not{D}\Psi\right) \ . \tag{59}$$

The electric $U(1)^{(1)}$ is broken by the charged matter; the magnetic $U(1)^{(1)}$ survives, $d\frac{F}{2\pi} = 0 \Rightarrow \star J_2 = \frac{F}{2\pi}$; and the axial symmetry $\Psi \to e^{i\alpha\gamma_5}\Psi$ with $\star j_5 = \bar{\Psi}\gamma\gamma_5\Psi$ suffers the ABJ anomaly

$$d\star j_5 = 2\,\frac{F\wedge F}{8\pi^2} \tag{60}$$

The naive defect $U_\alpha(\Sigma) = e^{i\alpha \oint_\Sigma \star j_5}$ fails to be topological,

$$\delta_\Sigma U_\alpha(\Sigma) = e^{2i\alpha \oint \frac{F\wedge F}{8\pi^2}} \ , \tag{61}$$

but for $\alpha = \frac{2\pi}{N}$ the symmetry can be "saved" by dressing with a fractional quantum Hall TQFT — morally $e^{\frac{i}{4\pi N}\int A\wedge dA}$, defined properly through the minimal TQFT $\mathcal{A}^{N,1}$ [19]:

$$\mathcal{D}_{1/N} = U_{\frac{2\pi}{N}}(\Sigma) \underbrace{\int [dv_1]\, e^{\frac{i}{4\pi}\int (N v_1\wedge dv_1 + 2v_1\wedge dA_1)}}_{\mathcal{A}^{N,1}(\Sigma)} \ , \qquad \mathcal{D}_{\frac{p}{N}} = U_{2\pi\frac{p}{N}}(\Sigma)\times \mathcal{A}^{N,p}(\Sigma) \tag{62}$$

is topological. The price is a lack of invertibility. If we write $v_\pm = \frac{1}{2}(v_1 \pm v_2)$:

$$\begin{aligned}\mathcal{D}_{1/N}\otimes\mathcal{D}_{-1/N} &= \int [dv_+ dv_-]\ \exp\left[\frac{iN}{4\pi}\int v_+\wedge dv_- + \frac{i}{2\pi}\int v_-\wedge F\right]\\ &= \sum_{v_-\in H^1(M;\mathbb{Z}_N)} e^{\frac{2\pi i}{N}\int v_-\wedge\frac{F}{2\pi}} \neq \mathbb{1} \ ,\end{aligned} \tag{63}$$

so in general

$$\mathcal{D}_q(\Sigma) = e^{2\pi i q\int \star j_5}\otimes \mathcal{A}^{N,p}(\Sigma) \ , \qquad q = \frac{p}{N}\in\mathbb{Q} \ , \qquad \mathcal{D}_q\otimes\mathcal{D}_{-q}\neq \mathbb{1} \ . \tag{64}$$

Equivalently, $\mathcal{D}_q$ is an interface between the theory $\mathcal{T}$ and its gauging $\mathcal{T}/\mathbb{Z}_N^{(1)}$ by the magnetic 1-form symmetry, with the FQH degrees of freedom on the wall (Figure 13) [20, 21]. The gauged action is

$$S[A_g,\Psi] + \frac{iN}{2\pi}\int da_1\wedge b_2 + \frac{i}{2\pi}\int b_2\wedge F + \frac{ikN}{4\pi}\int b_2\wedge b_2 \tag{65}$$

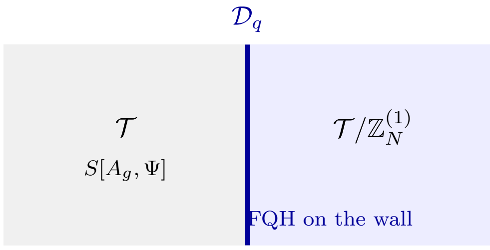

Figure 13: The non-invertible defect as an interface between the theory and its 1-form gauging.

As with higher groups, the payoff is a constraint on RG flows. The UV theory need *not* factorize as $U(1)^{(1)}\times\mathcal{C}_{NIS}$ (crossed out in the notes — the point being that the two structures emerge at separate scales), and the ordering of scales is fixed (Figure 11, right). For example, $SU(2) + N_f\,\Psi + \phi_{\text{adj}} \longrightarrow U(1) + N_f\,\Psi$ Dirac gives

$$E_{\text{1-form}} = v/g \ , \qquad E_{NIS} = v \quad (E_{NIS} \textit{ is instanton core size}) \tag{66}$$

4.3 Axion theories

Deforming the previous example by a Yukawa coupling lands on the axion theory, where both non-invertible structures of this lecture reappear — and constrain phenomenology.[9]

[8]This section follows the review [5], but the results come from the papers [22, 23].

Axion electromagnetism and the non-invertible shift symmetry. Take axion electrodynamics,

$$S = \int \frac{1}{2g^2} f \wedge *f + (\partial a)^2 + \frac{i}{8\pi^2 f_a} \int a\, f \wedge f \ , \qquad a \sim a + 2\pi f_a \ , \quad f = dA \ . \tag{67}$$

It is commonly said that the axion, being a periodic boson, has a continuous shift symmetry broken by the axion interaction: under $\delta a = \alpha f_a$,

$$\delta S = i\alpha \int \frac{f \wedge f}{8\pi^2} \ , \tag{68}$$

identical to the variation of a chiral symmetry with an ABJ anomaly. Given the discussion above, the shift symmetry is restored as a $\mathbb{Q}/\mathbb{Z}$ *non-invertible shift symmetry* in exactly the same way: with $*J_1 = *da$,

$$\mathcal{D}_q = e^{2\pi i q \int *J_1} \times \mathcal{A}^{N,p}(f, \Sigma_3) \ , \qquad q = \frac{p}{N} \in \mathbb{Q}/\mathbb{Z} \ . \tag{69}$$

This is obvious in hindsight: the axion is the pseudo-Goldstone mode of an ABJ-anomalous chiral symmetry, as in the KSVZ model

$$S_{\text{KSVZ}} = \int \frac{1}{2g^2} f \wedge *f + (\partial\phi)^2 + i\bar{\psi}\not{D}\psi - \lambda_\phi(|\phi|^2 - f_a^2)^2 + i\lambda_\psi\, \phi\psi\psi + \text{c.c.} \tag{70}$$

— ϕ condenses, the fermions get mass $\lambda_\psi f_a$, and the chiral phase becomes the axion. The non-invertible chiral symmetry of the UV KSVZ theory is matched by the non-invertible shift symmetry of the IR axion theory.

The non-invertible electric symmetry. The electric 1-form symmetry of free Maxwell theory, $*J_2^{(e)} = \frac{1}{g^2} * f$, is also broken by the axion coupling,

$$d * J_2^{(e)} = \frac{1}{4\pi^2 f_a}\, da \wedge f \ , \tag{71}$$

while the magnetic 1-form and winding 2-form currents remain conserved,

$$d * J_2^{(m)} = d\frac{f}{2\pi} = 0 \ , \qquad d * J_3 = d\frac{da}{2\pi f_a} = 0 \ , \tag{72}$$

and the violation is a product of conserved currents:

$$d * J_2^{(e)} = *J_3 \wedge *J_2^{(m)} \ . \tag{73}$$

(For the shift symmetry the same structure holds, $d * J_1 = \frac{1}{2} * J_2^{(m)} \wedge * J_2^{(m)}$, which is what permits the "saving.") The electric symmetry is then rescued as a non-invertible 1-form symmetry by dressing with a 2d TQFT,

$$\mathcal{D}_q^{(1)}(\Sigma_2) = e^{2\pi i q \int *J_2^{(e)}} \times \int [D\phi\, Dc_1] \exp\left\{\frac{i}{2\pi}\int \left(N\phi\, dc_1 + p\frac{a}{f_a}dc_1 + \phi f\right)\right\} \ , \qquad q = \frac{p}{N} \ , \tag{74}$$

with $\phi \sim \phi + 2\pi$ a periodic scalar and c_1 a dynamical $U(1)$ gauge field on the defect. Its fusion is explicitly non-invertible,

$$\mathcal{D}_{1/N}^{(1)} \times \mathcal{D}_{1/N}^{(1)\dagger} = \int [D\phi_\pm Dc_1^\pm] \exp\left\{\frac{i}{2\pi}\int \left(N\phi_+ dc_1^+ + N\phi_- dc_1^- + 2\frac{a}{f_a}dc_1^- + \phi_- f\right)\right\} \neq \mathbb{1} \ . \tag{75}$$

By comparison with the 0-form case: the non-invertible *shift* symmetry descends from the UV chiral symmetry, while the non-invertible *electric* symmetry is emergent along the RG flow from the KSVZ model.

Constraints on UV physics. These symmetries constrain UV completions: being good symmetries of the IR theory, they must be matched in the UV or be emergent. If emergent, the symmetries must emerge respecting the categorical nested structure.

For the non-invertible shift symmetry, the magnetic 1-form symmetry must emerge above the non-invertible shift symmetry. For the non-invertible electric symmetry, the magnetic and winding symmetries must emerge above the emergent classical electric symmetry. This leads to the inequality

$$\text{non-invertible:} \quad E_{\text{electric}} \lesssim \min\{E_{\text{winding}}, E_{\text{magnetic}}\} \ , \quad E_{\text{shift}} \lesssim E_{\text{magnetic}} \ . \tag{76}$$

In the KSVZ-type completion above (quartic λ_ϕ, Yukawa λ_ψ, $\langle|\phi|\rangle = v$),

$$E_{\text{electric}} = m_\psi = \lambda_\psi v \ , \qquad E_{\text{wind}} = \sqrt{\lambda_\phi}\, v \ \ (\text{strings stabilize}) \ , \qquad E_{\text{mag}} = E_{UV} \ , \tag{77}$$

$$\implies \qquad \boxed{\lambda_\psi \lesssim \sqrt{\lambda_\phi}} \tag{78}$$

This coincides with the boundary of EFT validity: for $\sqrt{\lambda_\phi} \ll \lambda_\psi$ the fermion contribution to the scalar potential is large and outside EFT control.

Axion–Yang–Mills and 3-Group Symmetry. Axion-Yang-Mills may also have a 3-group global symmetry. Take 4d $SU(N)$ axion–Yang–Mills,

$$S = \frac{1}{2}\int da \wedge *da + \frac{1}{g^2}\int \mathrm{Tr}[F \wedge *F] - \frac{i}{8\pi^2 f_a}\int a\, \mathrm{Tr}[F \wedge F] \ , \qquad a \sim a + 2\pi f_a \ . \tag{79}$$

The theory has a $\mathbb{Z}_N^{(1)}$ center symmetry and a $U(1)^{(2)}$ winding symmetry with $*J_3 = \frac{1}{2\pi f_a} da$. Coupling backgrounds B_2, C_3 and activating B_2 makes the instanton number fractional,

$$\frac{N}{8\pi^2}\int \mathrm{Tr}[(F - B_2) \wedge (F - B_2)] \in \mathbb{Z} \quad \implies \quad \delta_a S = -\frac{2\pi N(N-1)}{8\pi^2}\int B_2 \wedge B_2 \notin 2\pi\mathbb{Z} \ , \tag{80}$$

so axion periodicity — a gauge transformation — appears to break the 1-form symmetry. The cure is to promote C_3 to a 3-group background with invariant field strength

$$G_4 = dC_3 + \frac{N(N-1)}{4\pi} B_2 \wedge B_2 \ , \tag{81}$$

which requires the transformations

$$B_2 \mapsto B_2 + d\Lambda_1 \ , \qquad C_3 \mapsto C_3 + d\Lambda_2 - \frac{N(N-1)}{2\pi}\Lambda_1 \wedge B_2 - \frac{N(N-1)}{4\pi}\Lambda_1 \wedge d\Lambda_1 \ : \tag{82}$$

a 3-group symmetry (here with trivial 0-form part). As in Lecture 3, an emergent 3-group must respect its nesting, $E_{2\text{-form}} \gtrsim E_{1\text{-form}}$. In the standard UV completion ($SU(N)$ with fermions $\psi_\pm$ and a neutral scalar φ, potential $V = \lambda_\varphi(|\varphi|^2 - f_a^2)^2$, Yukawa λ_ψ), the winding symmetry emerges when axion strings stabilize, $E_{2\text{-form}} \sim \sqrt{\lambda_\varphi}\, f_a$, and the center symmetry emerges when the fundamental fermions are integrated out, $E_{1\text{-form}} = m_\psi = \lambda_\psi f_a$. The 3-group bound then gives

$$\sqrt{\lambda_\varphi} \gtrsim \lambda_\psi \ , \tag{83}$$

again coinciding with the EFT-validity bound — the abelian hierarchy above is its non-invertible refinement.

A Differential Forms and Gauge Theory

This follows the appendix in [5]. For a more complete review see [14].

A.1 Crash course on differential forms

Differential forms describe geometric properties of manifolds and bundles in a coordinate-independent, index-free language. They are labeled by a degree n; the set of n-forms on X is $\Omega^n(X)$. The 1-forms,

$$\omega_1 = \sum_i f_i(x)\, dx_i \in \Omega^1(X)\ , \tag{84}$$

are the "duals" of vector fields via the pairing $\langle \vec{v}, \omega \rangle = \sum_i v^i(x) f_i(x)$; with a metric,

$$\vec{v} \longmapsto v_1 = \sum_{ij} g_{ij} v^j(x)\, dx_i\ . \tag{85}$$

Higher forms are built with the *wedge product*, the antisymmetric product $dx_i \wedge dx_j = -dx_j \wedge dx_i$. The antisymmetrization is what makes n-forms fundamental n-volumes: for instance $d\mathrm{Vol} = d\vec{x}_1 \cdot (d\vec{x}_2 \times d\vec{x}_3) = dx_1 \wedge dx_2 \wedge dx_3$, and under a shear $(x, y, z) \mapsto (x,\ y + f(x),\ z + g(x))$ the volume form is invariant — the antisymmetry projects out the redundancies of non-orthogonal coordinates. A generic n-form is

$$\omega_n = \sum_{i_1,\ldots,i_n} f_{i_1 \ldots i_n}(x)\, dx_{i_1} \wedge \cdots \wedge dx_{i_n}\ , \tag{86}$$

and for a p-form and a q-form, $\alpha_p \wedge \beta_q = (-1)^{pq} \beta_q \wedge \alpha_p$.

Forms are naturally integrable: an n-form integrates over a closed n-manifold,

$$\oint_{\Sigma_n} \omega_n = \sum_{i_1,\ldots,i_n} \oint_{\Sigma_n} f_{i_1 \ldots i_n}(x)\, dx_{i_1} \cdots dx_{i_n}\ , \tag{87}$$

where the pullback keeps only components proportional to the volume element of Σ_n. One can also integrate an n-form along a p-manifold with $n > p$, obtaining an $(n{-}p)$-form, by decomposing ω_n against $d\mathrm{Vol}_p(\Sigma_p)$ and integrating the volume factor.

The *exterior derivative* generalizes the partial derivative,

$$d\omega_n = \sum_{i, i_1,\ldots,i_n} \frac{\partial f_{i_1 \ldots i_n}(x)}{\partial x_i}\, dx_i \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_n}\ , \tag{88}$$

and squares to zero, $d^2 = 0$, since partial derivatives commute while the wedge antisymmetrizes. These properties hold on curved manifolds — the machinery uses only partial derivatives; the "magic" is in the exterior algebra. The product rule reads

$$d(\alpha_p \wedge \beta_q) = d\alpha_p \wedge \beta_q + (-1)^p\, \alpha_p \wedge d\beta_q\ . \tag{89}$$

The *Stokes theorem* states that for an oriented $(n{+}1)$-manifold Σ with boundary,

$$\int_\Sigma d\omega_n = \oint_{\partial\Sigma} \omega_n\ , \tag{90}$$

which together with (89) gives integration by parts,

$$\int_M d\alpha_p \wedge \beta_q = \int_{\partial M} \alpha_p \wedge \beta_q + (-1)^p \int_M \alpha_p \wedge d\beta_q\ . \tag{91}$$

The *Hodge star* on a d-dimensional Riemannian manifold is a map $* : \Omega^n(M) \to \Omega^{d-n}(M)$ depending on orientation and metric,

$$*\big(dx^{i_1} \wedge \cdots \wedge dx^{i_n}\big) = \frac{\sqrt{\det g}}{(d-n)!}\, g^{i_1 j_1} \cdots g^{i_n j_n} \epsilon_{j_1 \ldots j_d}\, dx^{j_{n+1}} \wedge \cdots \wedge dx^{j_d}\ , \tag{92}$$

with $**\,\omega_n = (-1)^{n(d-n)} s\,\omega_n$, $s = +1$ Euclidean and $s = -1$ Lorentzian. Divergences become form operations: for the 1-form v_1 dual to $\vec{v}$ as in (85), $\vec{\nabla} \cdot \vec{v} = *\,d * v_1$ (exercise).

Finally, *de Rham cohomology*: on topologically non-trivial manifolds there are closed forms ($d\omega_n = 0$) that are not exact ($\omega_n \neq d\alpha_{n-1}$), classified by

$$H^n_{dR}(M) = Z^n(M)/C^{n-1}(M)\ , \qquad H^n_{dR}(M) = H^n(M;\mathbb{R})\ . \tag{93}$$

We will also use forms with *integral periods* $\Omega^n_{\mathbb{Z}}(M)$ — integrating to integers on all n-cycles — with integral cohomology $H^n(M;\mathbb{Z}) = Z^n_{\mathbb{Z}}(M)/C^{n-1}_{\mathbb{Z}}(M)$.

A.2 Gauge theory

Gauge theory is the theory of principal G-bundles with connection (G a Lie group, including discrete groups such as $\mathbb{Z}_N$): fibrations $G \to P \xrightarrow{\pi} X$ over spacetime. Over a "nice" cover $\{\mathcal{U}_\alpha\}$ of X the bundle trivializes, $\pi^{-1}(\mathcal{U}_\alpha) \cong G \times \mathcal{U}_\alpha$, with patches glued by transition functions

$$g_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \to G\ , \qquad (x, p_\alpha) = (x,\, g_{\alpha\beta}(x) \cdot p_\beta)\ , \tag{94}$$

subject to the triple-intersection condition

$$g_{\alpha\beta}(x) \cdot g_{\beta\gamma}(x) \cdot g_{\gamma\delta}(x) = \mathbb{1} \qquad x \in \mathcal{U}_{\alpha\beta\gamma} \tag{95}$$

— this is what will differ for "twisted" bundles. A connection compares fibers at different points: parallel transport along the base induces translations along the fiber, described locally by a $\mathfrak{g}$-valued 1-form A glued across patches by

$$A_\beta = g^{-1}_{\alpha\beta} A_\alpha g_{\alpha\beta} + g^{-1}_{\alpha\beta} d g_{\alpha\beta}\ , \tag{96}$$

with covariant derivative and curvature (field strength)

$$D_A = d + A\wedge\ , \qquad F = dA + A \wedge A\ . \tag{97}$$

A.3 Characteristic classes

For fixed G, any principal G-bundle over X is pulled back from a universal bundle over the *classifying space* BG (constructed as EG/G for contractible EG with free G-action; $P = f^*(EG)$ for $f : X \to BG$). Consequently all characteristic classes are pullbacks of cohomology classes of BG: the cohomology of BG classifies the possible topology of G-bundles. For simply connected G the classes are gauge-invariant polynomials in F (Chern classes),

$$\frac{\mathrm{Tr}[F \wedge \cdots \wedge F]}{n!(2\pi)^n} \in H^{2n}(BG;\mathbb{Z})\ ; \tag{98}$$

for $U(1)$, $H^{2n}(BU(1);\mathbb{Z}) = \mathbb{Z}\big[\frac{F\wedge\cdots\wedge F}{n!(2\pi)^n}\big]$; and for $\mathbb{Z}_N$, $H^{2n}(B\mathbb{Z}_N;\mathbb{Z}) = \mathbb{Z}_N$ (the "field strength" is the Bockstein $\beta_N(A)$). For non-simply-connected G such as $SO(N)$ or $PSU(N)$, the universal coefficient theorem gives

$$H_1(G;\mathbb{Z}) \cong H^2(BG;\mathbb{Z})\ : \tag{99}$$

G-gauge theory has discrete 2-form fluxes corresponding to $\pi_1(G)$ — the 1-form magnetic flux sectors.

A.4 Twisted bundles and center symmetry

Center symmetry of a G-bundle with center $Z = Z(G)$ acts by shifting the gauge field by a Z-gauge field; mathematically this is the *twisting* of the bundle, classified by $w_2 \in H^2(BG; Z)$. For $Z = \mathbb{Z}_N$ and matter uncharged under $\mathbb{Z}_N \subset G$, the triple-overlap condition can be twisted to

$$g_{\alpha\beta} \circ g_{\beta\gamma} \circ g_{\gamma\alpha} = \omega \in \mathbb{Z}_N \tag{100}$$

without inducing branch cuts, shifting the curvature by the discrete flux $F_G \mapsto F_G + \frac{1}{N} w_2 \mathbb{1}_G$. Equivalently one lifts to a G^c-bundle — a principal $\frac{G\times U(1)}{\mathbb{Z}_N}$-bundle with overlap condition $g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = (z, z^{-1}) \sim (1,1)$, $z \in \mathbb{Z}_N \subset U(1)$ — a twisted G-bundle and a twisted $U(1)$-bundle combining into an untwisted total bundle. The curvature obeys $F_{G^c} = \omega_{U(1)} + F_G$ with integral total class, so the two factors carry opposite fractional fluxes. For center symmetry one replaces $U(1)$ by $\mathbb{Z}_N$ ($\frac{G\times\mathbb{Z}_N}{\mathbb{Z}_N} \cong G$): no new degrees of freedom, just fractional center flux, imposed in the path integral by a Lagrange multiplier.

Example: $G = SU(N)$. Twisting by a $\mathbb{Z}_N$-valued b_2 means lifting to

$$G^c = \frac{SU(N)\times U(1)}{\mathbb{Z}_N} \cong U(N)\,, \qquad F_{SU(N)} = F_{U(N)} - F_{U(1)}\,, \qquad \oint \frac{F_{U(1)}}{2\pi} = \oint \frac{b_2}{2\pi} \operatorname{mod}_{\mathbb{Z}} \in \frac{1}{N}\mathbb{Z}\,. \tag{101}$$

Imposing the constraint via a Lagrange multiplier,

$$S = \cdots + \frac{i}{2\pi}\int \varphi_2 \wedge \mathrm{Tr}[F_{U(N)} - B_2 \mathbb{1}_N] \quad \Longrightarrow \quad \mathrm{Tr}[F_{U(N)}] = N B_2\,, \tag{102}$$

projects $U(N) \to SU(N)$ twisted by b_2. The instanton number then becomes fractional:

$$\int \frac{\mathrm{Tr}[F_{SU(N)} \wedge F_{SU(N)}]}{8\pi^2} = N(N-1)\int \frac{B_2 \wedge B_2}{8\pi^2} \operatorname{mod}_{\mathbb{Z}} = \frac{N-1}{2N}\int w_2 \wedge w_2 \operatorname{mod}_{\mathbb{Z}}\,, \tag{103}$$

with $B_2 = \frac{2\pi}{N} w_2$ and $\oint w_2 \in \mathbb{Z}$. This is the computation behind the fractional instanton numbers of Lecture 2 and the 3-group structure of axion–Yang–Mills.

References

[1] J. McGreevy, "Generalized Symmetries in Condensed Matter," *Ann. Rev. Condensed Matter Phys.* **14** (2023) 57–82, arXiv:2204.03045 [cond-mat.str-el].

[2] S.-H. Shao, "What's Done Cannot Be Undone: TASI Lectures on Non-Invertible Symmetries," in *Theoretical Advanced Study Institute in Elementary Particle Physics 2023: Aspects of Symmetry*. 8, 2023. arXiv:2308.00747 [hep-th].

[3] L. Bhardwaj, L. E. Bottini, L. Fraser-Taliente, L. Gladden, D. S. W. Gould, A. Platschorre, and H. Tillim, "Lectures on generalized symmetries," *Phys. Rept.* **1051** (2024) 1–87, arXiv:2307.07547 [hep-th].

[4] D. Costa *et al.*, "Simons Lectures on Categorical Symmetries," 11, 2024. arXiv:2411.09082 [math-ph].

[5] T. D. Brennan and S. Hong, "Introduction to Generalized Global Symmetries in QFT and Particle Physics," arXiv:2306.00912 [hep-ph].

[6] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, "Generalized Global Symmetries," *JHEP* **02** (2015) 172, arXiv:1412.5148 [hep-th].

[7] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, "Theta, Time Reversal, and Temperature," *JHEP* **05** (2017) 091, arXiv:1703.00501 [hep-th].

[8] E. Witten, "Supersymmetric index in four-dimensional gauge theories," *Adv. Theor. Math. Phys.* **5** (2002) 841–907, arXiv:hep-th/0006010.

[9] M. Barkeshli, P. Bonderson, M. Cheng, and Z. Wang, "Symmetry Fractionalization, Defects, and Gauging of Topological Phases," *Phys. Rev. B* **100** no. 11, (2019) 115147, arXiv:1410.4540 [cond-mat.str-el].

[10] D. G. Delmastro, J. Gomis, P.-S. Hsin, and Z. Komargodski, "Anomalies and symmetry fractionalization," *SciPost Phys.* **15** no. 3, (2023) 079, arXiv:2206.15118 [hep-th].

[11] T. D. Brennan, C. Cordova, and T. T. Dumitrescu, "Line Defect Quantum Numbers & Anomalies," arXiv:2206.15401 [hep-th].

[12] T. D. Brennan, T. Jacobson, and K. Roumpedakis, "Consequences of symmetry fractionalization without 1-form global symmetries," *JHEP* **11** (2025) 153, arXiv:2504.08036 [hep-th].

[13] A. Hatcher, *Algebraic Topology*. Cambridge University Press, Cambridge, 2002.

[14] M. Nakahara, *Geometry, Topology and Physics*. CRC Press, Boca Raton, 2nd ed., 2003.

[15] T. D. Brennan and K. Intriligator, "Generalized Families of QFTs," arXiv:2602.09105 [hep-th].

[16] D. S. Freed, G. W. Moore, and C. Teleman, "Topological symmetry in quantum field theory," arXiv:2209.07471 [hep-th].

[17] Y. Choi, H. T. Lam, and S.-H. Shao, "Noninvertible Global Symmetries in the Standard Model," *Phys. Rev. Lett.* **129** no. 16, (2022) 161601, arXiv:2205.05086 [hep-th].

[18] C. Cordova and K. Ohmori, "Noninvertible Chiral Symmetry and Exponential Hierarchies," *Phys. Rev. X* **13** no. 1, (2023) 011034, arXiv:2205.06243 [hep-th].

[19] P.-S. Hsin, H. T. Lam, and N. Seiberg, "Comments on One-Form Global Symmetries and Their Gauging in 3d and 4d," *SciPost Phys.* **6** no. 3, (2019) 039, arXiv:1812.04716 [hep-th].

[20] Y. Choi, C. Cordova, P.-S. Hsin, H. T. Lam, and S.-H. Shao, "Non-invertible Condensation, Duality, and Triality Defects in 3+1 Dimensions," *Commun. Math. Phys.* **402** no. 1, (2023) 489–542, arXiv:2204.09025 [hep-th].

[21] Y. Choi, C. Cordova, P.-S. Hsin, H. T. Lam, and S.-H. Shao, "Noninvertible duality defects in 3+1 dimensions," *Phys. Rev. D* **105** no. 12, (2022) 125016, arXiv:2111.01139 [hep-th].

[22] T. D. Brennan and C. Cordova, "Axions, higher-groups, and emergent symmetry," *JHEP* **02** (2022) 145, arXiv:2011.09600 [hep-th].

[23] Y. Choi, H. T. Lam, and S.-H. Shao, "Non-invertible Gauss law and axions," *JHEP* **09** (2023) 067, arXiv:2212.04499 [hep-th].