University of Amsterdam Institute for Theoretical Physics

Master Thesis

Generalised symmetries, anomalies, and other tales

Stathis Vitouladitis
12239127

Supervisor: Dr Diego Hofman
Second Examiner: Prof. Jan de Boer

A thesis submitted in partial fulfilment of the requirements
for the degree of
MSc Physics and Astronomy: track Theoretical Physics

July 2020

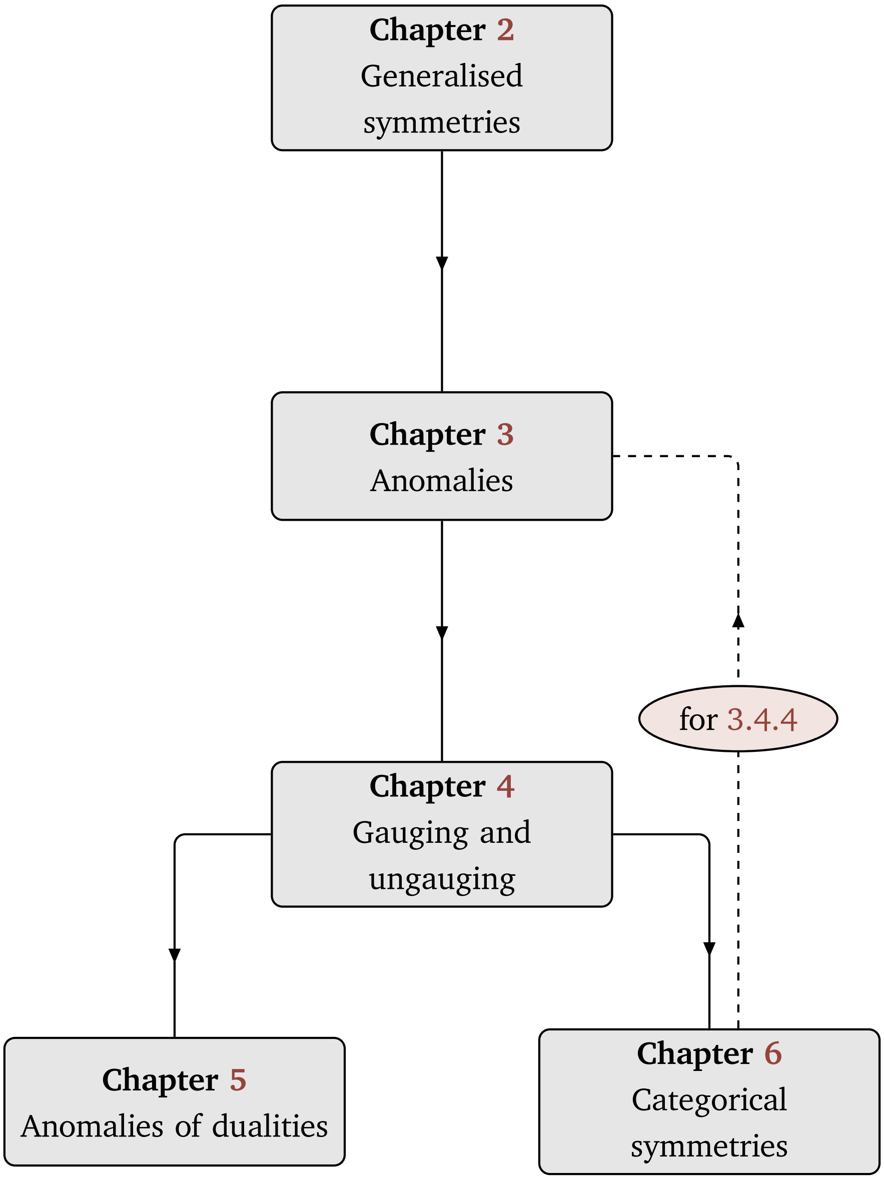
Abstract

This thesis focuses mainly on two aspects of generalised symmetries. These symmetries (also known as p-form symmetries) are captured by a codimension-$(p+1)$ topological operator which acts on p-dimensional operators. They are proper symmetries of the theory, obeying modified versions of the usual laws that symmetries obey: Spontaneous symmetry breaking, Mermin–Wagner and Goldstone theorem. More importantly, they can have 't Hooft anomalies. In the modern understanding, anomalies of d-dimensional QFTs are captured by Symmetry Protected Topological (SPT) phases in $(d+1)$ dimensions. An anomalous QFT should be matched by a non-trivial SPT phase, and since anomalies are robust under renormalisation group flow, this puts restrictions on the allowed low energy physics. Therefore higher-form anomalies provide a further set of tools to constrain and study the IR phases of quantum field theories. One aspect of this thesis concerns "lower"-form, i.e. (-1)-form symmetries and their anomalies. These are effectively a study of dualities recast in terms of generalised symmetries. In this spirit, *we identify a new mixed anomaly in quantum mechanics, between modular-invariance/T-duality and conformal symmetry* and we study its implications. The second part of this thesis consists of a further generalisation of symmetries. In particular, relaxing the condition that symmetries form a group and taking the topological nature of charge operators as the defining characteristic of a symmetry, we study non-invertible symmetries. In two dimensions it is well known that (not necessarily invertible) topological operators form a fusion category and explain symmetries such as those of of RCFTs. We propose a generalisation in order to describe non-invertible (both ordinary and higher-form) symmetries in any dimension. We argue that a p-form (not necessarily invertible) symmetry in d dimensions is given by a $(p+1)$*-fusion* $(d-p-1)$*-category* and we discuss examples and possible applications.

Contents

Organisation of this thesis

The organisation of this thesis and the interdependence of the chapter is outlined in the following diagram. Solid arrows indicate that the chapter at the origin is necessary to understand the targeted chapter. In contrast, dashed arrows indicate some dependence of the target-chapter on the origin-chapter, whereby revisiting the target-chapter after reading the origin-chapter should shed some light on possibly unclear points in the discussion there.



1

Introduction

1.1 Pseudoaxiomatic definition of quantum field theory

If we were to understand quantum field theory axiomatically à la Atiyah–Segal [1, 2], we could come up with the following definition. A relativistic, Euclidean quantum field theory, Q, in d dimensions, with some structure $\bullet$ (such as orientation, spin structure, metric, a map to another space X, a principal G-bundle with connections, etc.) is a functor from a suitable bordism category to the category of vector spaces. In particular, if the structure $\bullet$ includes *the smooth structure*, the resulting QFT is called a TQFT — or rather an *extended* TQFT — and the (yet-to-come) definition reduces to the standard Atiyah–Segal axioms.

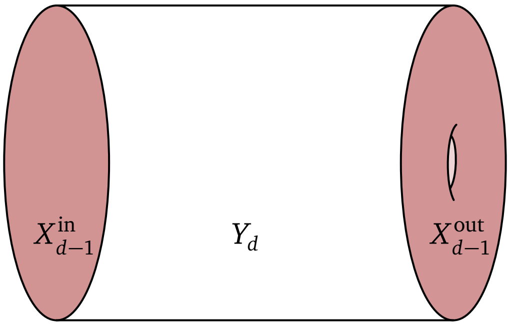
We say that the QFT is performing the following assignments:

X_d	a d-dimensional closed manifold with structure $\bullet$
$\rightsquigarrow$	
$\mathcal{Z}_Q(X_d)$	the partition function (a $\mathbb{C}$-number)
X_{d-1}	a $(d-1)$-dimensional closed manifold with structure $\bullet$
$\rightsquigarrow$	
$\mathcal{H}_Q(X_{d-1})$	the space of states (a $\mathbb{C}$-vector space)
X_{d-2}	a $(d-2)$-dimensional closed manifold with structure $\bullet$
$\rightsquigarrow$	
$\mathcal{C}_Q(X_{d-2})$	data to factorise the space of states across an entangling surface X_{d-2} $\big($a category$\big)$
$\vdots$	$\vdots$

For our purposes — and the majority of physical/non-exotic constructions — the first two assignments are enough, and we will safely forget the whole tower after them. The reason why such an axiomatic definition is not groundbreaking for QFTs as it is for TQFTs is that in TQFTs the structure is enough to pin down the functor exactly and thus to solve the theory completely, while in general QFTs this is not the case. An interesting, yet very premature comment is in place here. It might prove useful to consider a general QFT as a TQFT where we remove T. Then we can work in reverse to restore it. If we add a symmetry to the QFT, it has some topological nature (cf. chapter 2) hidden in its symmetry operators. Then we can add more and more symmetries of all dimensions to render the theory entirely topological. Reversing the logic, we can use a TQFT as the "free theory" off of which we might construct a perturbation expansion away from topologicality to try to approximate generic QFTs. Let us now get back to the main discussion.

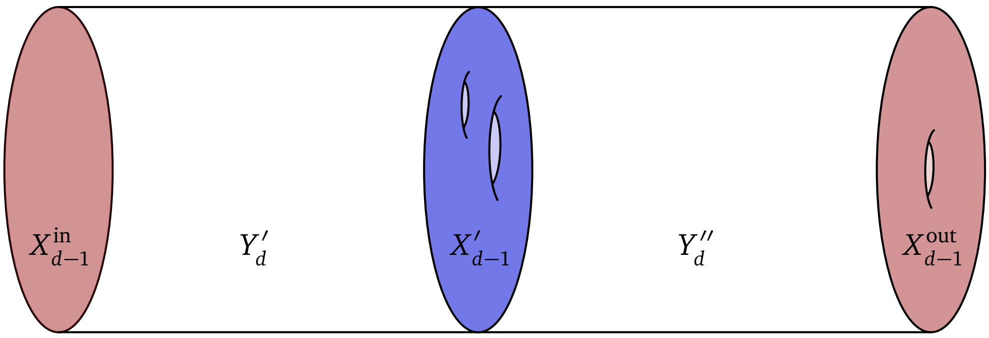
Albeit this axiomatisation does not buy us much in terms of performing calculations, it does buy us something; a way to build generic properties that all QFTs should possess. We can see these, by rephrasing the above assignments as follows. Given manifolds X^{in}_{d-1} and X^{out}_{d-1} such that there exists a Y_d, with $\partial Y_d = X^{\text{in}}_{d-1} \sqcup \overline{X^{\text{out}}_{d-1}}$, we can view $\mathcal{Z}_Q(Y_d)$ as a linear map:

$$\mathcal{Z}_Q(Y_d) : \mathcal{H}_Q\big(X^{\text{in}}_{d-1}\big) \longrightarrow \mathcal{H}_Q\big(X^{\text{out}}_{d-1}\big)$$



and we should think of $\mathcal{Z}_Q$ as the transition amplitude between an in- and an out-state. It also has to satisfy correct composition, i.e.,

$$\mathcal{Z}_Q(Y_d) = \mathcal{Z}_Q\big(Y'_d\big)\, \mathcal{Z}_Q\big(Y''_d\big).$$



Furthermore, we require that

$$\begin{gathered}\mathcal{H}_Q\big(X_{d-1} \sqcup X'_{d-1}\big) = \mathcal{H}_Q(X_{d-1}) \otimes \mathcal{H}_Q\big(X'_{d-1}\big)\\ \mathcal{H}_Q(\varnothing) = \mathbb{C}.\end{gathered}$$

Then a closed Y_d has $X_{d-1}^{\text{in}} = X_{d-1}^{\text{out}} = \varnothing$ and so the transition amplitude,

$$\mathcal{Z}_Q(Y_d) : \mathbb{C} \longrightarrow \mathbb{C}, \tag{1.1.1}$$

is the partition function.

When the structure $\bullet$ contains the metric we can write $Y_d = X_{d-1} \times [0, \beta]$ and define

$$\begin{aligned} \mathcal{Z}_Q(Y_d) &: \mathcal{H}_Q(X_{d-1}) \longrightarrow \mathcal{H}_Q(X_{d-1}) \\ &\;\;!! \\ &\mathrm{e}^{-\beta H(X_{d-1})}, \end{aligned}$$

where $H(X_{d-1})$ is the Hamiltonian.

For two theories Q and Q' with the same dimension and structure we can define $Q \times Q'$ and $Q + Q'$ by demanding

$$\begin{aligned} \mathcal{H}_{Q\times Q'}(X_{d-1}) &= \mathcal{H}_Q(X_{d-1}) \otimes \mathcal{H}_{Q'}(X_{d-1}) \quad \text{and} \\ \mathcal{H}_{Q+Q'}(X_{d-1}) &= \mathcal{H}_Q(X_{d-1}) \oplus \mathcal{H}_{Q'}(X_{d-1}). \end{aligned}$$

Let us give a few examples of classes of QFTs in the above language.

Trivial QFT

The simplest such example is that of a trivial QFT. It is specified By

$$\begin{aligned} &\mathcal{H}_{\text{triv}}(X_{d-1}) = \mathbb{C} \\ &\mathcal{Z}_{\text{triv}}(Y_d) : \{\mathcal{H}_{\text{triv}}(X_{d-1}^{\text{in}}) = \mathbb{C}\} \xrightarrow{\text{id}} \{\mathcal{H}_{\text{triv}}(X_{d-1}^{\text{out}}) = \mathbb{C}\}. \end{aligned}$$

It clearly defines the identity element of the operation $Q \times Q'$.

Invertible QFT

The second simplest example is to have

$$\mathcal{H}_Q(X_d) \cong \mathbb{C},$$

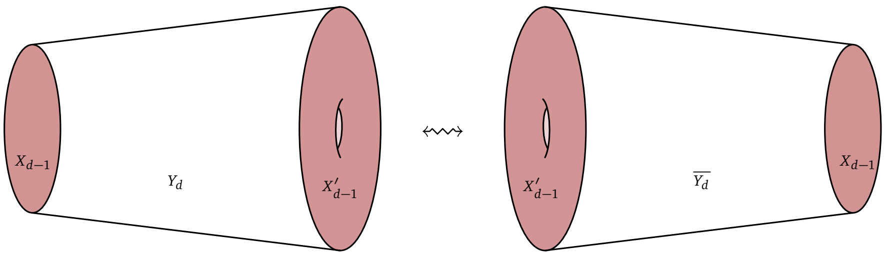
but not canonically isomorphic. Then we can define the inverse theory, Q^{-1}, by

$$\mathcal{H}_{Q^{-1}}(X_{d-1}) := \mathcal{H}_Q(X_d)^\vee,$$

so that $Q \times Q^{-1} = \text{triv}$. The set of invertible QFTs forms an Abelian group under $Q \times Q'$. Invertible QFTs are conjectured to be the low energy limits of Symmetry Protected Topological phases. As we will see in section 3.2 the latter capture the anomalies of anomalous QFTs in one dimension lower.

1.1.1 Unitarity of Euclidean QFTs

Before we discuss unitarity of Euclidean QFTs let us take a detour in the world of Lorentzian QFTs. There we want the operators that correspond to observables/conserved charges, such as the Hamiltonian, to be Hermitian. This constraint is transferred to the transition amplitudes, by the demand that the evolution is unitary. In Euclidean QFT this translates to demanding that the QFT has reflection positivity.

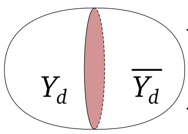


In particular, the reflection $\overline{Y_d}$, as defined above, by exchanging the in- and out-boundaries and reversing the orientation defines a conjugate partition function

$$\mathcal{Z}_Q(\overline{Y_d}) : \mathcal{H}_Q(X'_{d-1}) \longrightarrow \mathcal{H}_Q(X'_{d-1}). \tag{1.1.2}$$

The statement of unitarity in a Euclidean QFT is then that $\mathcal{Z}_Q(Y_d)\mathcal{Z}_Q(\overline{Y_d})$ is positive definite. More precisely, since the partition function on a manifold with boundaries defines a vector rather than a number, a more concrete statement [3] is that if Y_d has no incoming boundary, then

$$\mathcal{Z}_Q(\Delta Y_d) := \mathcal{Z}_Q(Y_d)\mathcal{Z}_Q(\overline{Y_d}) = \mathcal{Z}_Q\left(\quad \right) > 0. \tag{1.1.3}$$



1.2 Low energy behaviour of quantum field theories

In the Wilson approach to renormalisation, one can take their favourite quantum field theory and flow with it all the way to the infrared, integrating out degrees of freedom heavier than the cutoff scale at each point during the RG flow. A critical question that one can ask is what is the theory describing the low energy degrees of freedom if there are any. For that, it is useful to zoom out a little and consider what *types* of theories can there be in the IR. For this, we generally distinguish theories between:

- **gapless:** there is a massless excitation in the spectrum. Generally, gapless theories describe either free or interacting conformal theories (CFTs). In some parts of the literature, they make a distinction between trivial and non-trivial fixed points with regards to the IR behaviour, regarding trivial fixed points as too simple to be called a CFT and reserving the term CFT only for non-trivial fixed points. The difference is mainly semantic; therefore, for our purposes, gaplessness will be used interchangeably with conformality.
- **gapped:** the lightest excitation in the spectrum has mass $m > 0$. We further distinguish gapped theories between:
    - **trivially gapped:** there are no degrees of freedom in the IR theory, or
    - **TQFTs:** there are degrees of freedom, that are topological in nature. They show up as degenerate vacua.

For concreteness let us give some examples theories falling into each category.

- **gapless:**
    - **free:** An example of such a theory is one that spontaneously breaks a continuous global symmetry along the RG flow. The spectrum includes then massless Goldstone modes. Once we flow to the deep IR, what we are left with is only the Goldstones, and maybe also massless degrees of freedom whose masslessness is protected by some (unbroken) symmetry. A specific example is an abelian superfluid, whose low energy effective action is just

$$S_{\mathrm{IR}}[\varphi] = \int P\Big(\sqrt{-\mathrm{d}\varphi \wedge \star \mathrm{d}\varphi}\Big),$$

where φ is the phase of the original field $\Phi = re^{i\varphi}$, i.e. the Goldstone mode of a spontaneously broken U(1) theory.

Another class of examples is a theory which has some protected massless degrees of freedom and some massive degrees of freedom, at some intermediate RG point. Then, while flowing to the IR we are instructed to throw away the massive degrees of freedom, leaving us only with the massless ones. Concretely, consider QED. The theory consists of the Maxwell sector, with a gauge field, a, and the matter sector with the electron ψ. At energies below the electron mass $E \ll m_\psi$, we cannot excite any matter so we are left with the massless Maxwell sector,

$$S_{\mathrm{IR}}[a] = \frac{1}{2}\int \mathrm{d}a \wedge \star \mathrm{d}a.$$

Even if the electron was massless, since it is not protected by symmetry, the theory flows logarithmically to the free Maxwell theory. Under a new light, this class of examples falls into the above class of examples by considering spontaneous breaking of higher-form symmetries. Then we can consider the photon of the Maxwell theory as the Goldstone mode of a spontaneously broken $\mathrm{U}(1)^{[1]}$ symmetry [4,5] (see also section 2.4).

- **interacting:** This type of theories can emerge, when one hits a fixed point, before reaching the extreme IR. The theory then becomes conformal, and the RG flow cannot continue (unless of course one explicitly deforms it). One such example provides the Banks-Zaks fixed point. For this the setup that we will consider is Yang–Mills theory, with coupling constant g, with N colours and n fermions. The two-loop beta function of the theory is given by

$$\beta(g) = -\beta_0 \frac{g^3}{16\pi^2} + \beta_1 \frac{g^5}{(16\pi^2)^2} + \mathcal{O}(g^7),$$

where $\beta_0 = \frac{11}{3} N - \frac{2}{3} n$ and $\beta_1 = \mathcal{O}(N^2, Nn)$. Choosing $N, n \gg 1$, in such a way that $\beta_0 = \mathcal{O}(1) > 0$, and hence implying that $\beta_1 = \mathcal{O}(N^2) > 0$, we can rewrite the beta function as

$$\beta(g) = -\frac{g^3}{16\pi^2}\left(1 - \frac{\beta_1}{\beta_0}\frac{g^2}{16\pi^2} + \cdots\right).$$

We see that the theory possesses a non-trivial fixed point at $g = g_\star$, such that $\dfrac{g^2}{16\pi^2} = \mathcal{O}\big(\frac{1}{N^2}\big)$. This behaviour can also happen at finite N and n, when

$$n_{\mathrm{crit}} < n < \tfrac{11}{2} N.$$

Here n_{crit} is some lower critical value of n, such that the obtained beta function is trustworthy and the regime where the Banks–Zaks fixed point can be reached has been dubbed "conformal window".

- **gapped:**
  - **trivially gapped:** Here lie most theories that do not break a symmetry. The most trivial trivially gapped theory is that of a free massive scalar

$$S[\phi] = \frac{1}{2}\int \left(\mathrm{d}\phi \wedge \star\mathrm{d}\phi + m^2\phi \wedge \star\phi\right).$$

    There really is nothing left once we flow to the deep IR; the particles are too heavy to excite and the theory is empty.

    A little less trivial example is Yang–Mills theory in four dimensions:

$$S[a] = -\frac{1}{2g^2}\int \mathrm{Tr}\left(f \wedge \star f\right), \qquad f = \mathrm{d}a + a \wedge a.$$

    The beta function is negative $\beta(g) \sim -g^3 + \mathcal{O}\left(g^5\right)$, so the theory is asymptotically free at high energies, where the coupling goes to zero while at low energies, when the coupling grows, it becomes non-perturbative at the scale

$$\Lambda_{\mathrm{IR}} = \Lambda_{\mathrm{UV}} \exp\left(-\frac{\#}{g_{\mathrm{UV}}^2}\right).$$

    The lightest excitation is then of order Λ_{IR} and hence the theory is trivially gapped.
  - **TQFTs:** Famously an example of such a theory is Chern–Simons theory, with action

$$S[a] = \int_{X_3} \mathrm{Tr}\left(a \wedge \mathrm{d}a + \frac{2}{3} a \wedge a \wedge a\right). \tag{1.2.1}$$

    In Chern–Simons theory there are no local excitations, but one can excite topological configurations. For example, it captures the physics of the quantum Hall effect, it allows one to move between vacua of four-dimensional Yang–Mills theories and more.

2 Generalised symmetries

2.1 Symmetries

Consider a d-dimensional Euclidean QFT, with a continuous global symmetry[1], given by a simply connected Lie group, G. Ward identities — the quantum upgrade of Noether's theorem — assert that there must be a current J_μ operator, which is conserved on-shell, satisfying $\partial^\mu J_\mu(x) = 0$, as an operator equation, i.e.

$$\left\langle \partial^\mu J_\mu(x)\, \mathcal{O}(y) \right\rangle \sim \delta(x-y)\mathcal{O}(y),$$

for any local operator insertion, $\mathcal{O}(y)$. One might equivalently write, in adult terms, that there exists a conserved one-form $J_{[1]} := J_\mu(x)\, \mathrm{d}x^\mu$, satisfying

$$\mathrm{d} \star J_{[1]} = 0, \tag{2.1.1}$$

[1] We will drop the adjective global, before the word symmetry from now on, except for cases where it is really necessary.

as an operator equation.

From $J_\mu(x)$ one would normally construct a conserved charge by integrating its zeroth component along a Cauchy slice,

$$Q = \int_{\text{space}} \mathrm{d}^{d-1}x\, J_0(x).$$

In the coordinate-free language one would better write

$$Q(M_{d-1}) := \int_{X_{d-1}} \star J_{[1]}, \tag{2.1.2}$$

where now X_{d-1} is *any* manifold of codimension one. Note that since we are in Euclidean signature we can pick whichever codimension-one manifold we want, and we are guaranteed a conserved charge. If we want to do Lorentzian quantum field theory, we should pick a spacelike manifold to integrate the current on. The next reasonable move from here is to exponentiate the charge, in order to define an operator acting on the Hilbert space, hereby called *symmetry operetor* or *charger*:

$$U_{\mathrm{e}^{i\alpha}}(X_{d-1}) := \exp\big(i\alpha \cdot Q(X_{d-1})\big) = \exp\left(i\alpha \cdot \int_{X_{d-1}} \star J_{[1]}\right). \tag{2.1.3}$$

Here α is an element of the Lie algebra, $\mathfrak{g}$, of G and the dot product is taken with respect to the generators of the algebra.

Note that $U_{\mathrm{e}^{i\alpha}}(X_{d-1})$ is topological, in the sense that it is insensitive to deformations of X_{d-1}. This can be seen as follows. Suppose X'_{d-1} is a deformation of X_{d-1}, written as $X'_{d-1} = X_{d-1} + \partial Y_d$. Then, at the purely classical level

$$\begin{aligned} U_{\mathrm{e}^{i\alpha}}\big(X'_{d-1}\big) = U_{\mathrm{e}^{i\alpha}}(X_{d-1} + \partial Y_d) = \exp\left(i\alpha \cdot \int_{X_{d-1}+\partial Y_d} \star J_{[1]}\right) = \\ = \exp\left(i\alpha \cdot \left[\int_{X_{d-1}} \star J_{[1]} + \int_{Y_d} \mathrm{d} \star J_{[1]}\right]\right) = U_{\mathrm{e}^{i\alpha}}(X_{d-1}), \end{aligned}$$

due to the conservation of $J_{[1]}$. However at the quantum level (2.1.1) holds only when the topological change does not cross any operator charged under the symmetry. Therefore we define the topological property at the quantum level, as *deformations of X_{d-1} that do not cross any local operators*. With this definition of topological movements, the symmetry operators are by construction topological at the full quantum level.

Zooming out, the existence of a continuous symmetry, G is fully equivalent to the existence of a set of topological operators,

$$\Big\{U_g(X_{d-1}) \,\Big|\, g \in G\Big\},$$

supported on a codimension-1 manifold.

This observation leads to define the existence of a symmetry, based on a group G — which is not constrained to be a simply-connected Lie group, or even a continuous group whatsoever — by the existence of a set of codimension-1 topological operators

$$\left\{U_g(X_{d-1}) \mid g \in G\right\}. \tag{2.1.4}$$

The upshot of this definition is that it puts discrete and continuous symmetries on the same footing, i.e. does not rely on the existence of a current, and does not rely on a Lagrangian description of the theory.

How do the chragers charge?

These symmetry operators act on the local operators of the theory, by surrounding them, and we write

$$U_g(X_{d-1})\,\mathcal{O}(x) = g \cdot \mathcal{O}(x), \tag{2.1.5}$$

Schematically we can see the action of a symmetry operator as in figure 2.1. More specif-

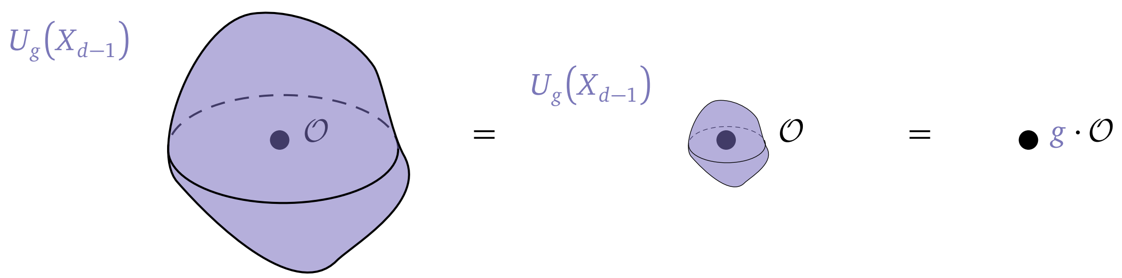


Figure 2.1: Action of a symmetry operator on a local operator

ically, when X_{d-1} is $\mathbf{S}^{d-1}$, we have

$$U_g(X_{d-1})\,\mathcal{O}_i(x) = (g \cdot \mathcal{O})_i(x) = R_i^{\ j}(g)\mathcal{O}_j(x), \tag{2.1.6}$$

where $R_i^{\ j}(g)$ is the representation of the group element g, of $\mathcal{O}$. For example, in the case of a U(1) symmetry, then $g = \mathrm{e}^{i\theta}$, and $R_i^{\ j}(\mathrm{e}^{i\theta}) = \mathrm{e}^{i\theta q(\mathcal{O})}\delta_i^{\ j}$, where $q(\mathcal{O})$ is the charge of $\mathcal{O}$.

If as a special case we consider X_{d-1} to be all of space, at some fixed time, then (2.1.6) implies the equal time commutation relation:

$$U_g(X_{d-1})\mathcal{O}_i(x) = R_i^{\ j}(g)\mathcal{O}_j(x)U_g(X_{d-1}). \tag{2.1.7}$$

Here x is exactly on some spacetime point of X_{d-1}, giving rise to the non-commutativity.

Group structure

The way to define $U_g(X_{d-1})$, both for discrete and continuous symmetry groups, is to cut spacetime along X_{d-1} and insert a symmetry transformation in the complete set of states for the Hilbert space associated to $U_g(X_{d-1})$. Given two such symmetry operators associated to different elements g and g', they satisfy

$$U_{g'}(X_{d-1}) \otimes U_g(X_{d-1}) = U_{g\,g'}(X_{d-1}), \tag{2.1.8}$$

i.e. the action of the fusion of two topological operators is the action of the topological operator associated to the product of group elements. In other words, the operator that corresponds to the element $g\,g'$ comes of the fusion of U_g with $U_{g'}$. The symmetry operators also satisfy an associativity relation, meaning that when one is interested in the fusion of three symmetry operators, it does not matter if they fuse the outer ones first and then the product of the fusion with the innermost operator, or the other way around. Associativity can in some cases be relaxed and it gives rise to the physics of 2-groups [6–9] (cf. chapter 4). Similarly, these operators are invertible, thus closing the group structure of the symmetry. Also similarly, invertibility can be relaxed, and the physics of non-invertible symmetry operators is the subject of much recent research [10–15]. The mathematics underlying this sort of operators are captured by (higher-) categories and is the focus of chapter 6.

2.2 Higher-form symmetries

A natural generalisation of the above discussion is to consider topologcial operators supported on a codimension-$(p+1)$ manifold, where $p > 1$. The objects that are naturally charged under these operators are then p-dimensional extended operators — i.e. operators, $\mathcal{W}(\Sigma_p)$ supported on p-dimensional manifolds — and the action of the symmetry operators is again by surrounding, as sketched in figure 2.2 for the case $p = 1$ (where we imagine that both operators extend in all other dimensions that we cannot draw).

In the case of a continuous symmetry the symmetry operators are constructed by exponentiating a conserved charge, as before, but this time the charge is supported on a codimension-$(p+1)$ manifold and hence the conserved current associated to it is a $(p+1)$-

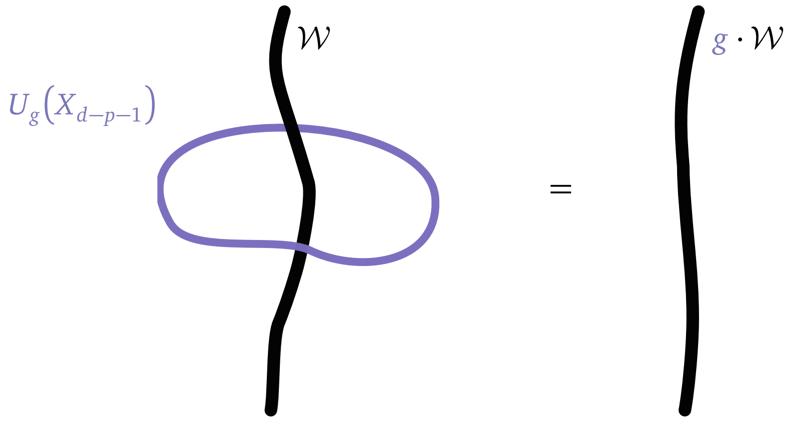


Figure 2.2: Action of codimension-2 symmetry operator on a line operator

form:

$$Q(X_{d-p-1}) = \int_{X_{d-p-1}} \star J_{[p+1]} \tag{2.2.1}$$

$$\mathrm{d} \star J_{[p+1]} = 0. \tag{2.2.2}$$

Due to the higher-form current — as well as because at the level of a Lagrangian these symmetries couple to $(p+1)$-form connections (cf. section 2.3) — these symmetries are dubbed [16, 17] *higher-form* or *p-form symmetries*. Setting $p = 0$ the discussion reduces to the discussion of conventional symmetries, discussed above, hence we will hereby refer to conventional symmetries as 0-form symmetries, whenever we need to distinguish them.

Similarly as with 0-form symmetries, higher-form symmetry operators fuse according to a group law

$$U_g(X_{d-p-1}) \otimes U_{g'}(X_{d-p-1}) = U_{gg'}(X_{d-p-1}) \tag{2.2.3}$$

Notice however that in the case $p > 1$ there are enough codimensions that the operators can slide across each other without touching. This implies directly that *higher-form symmetries are abelian*, since the the product of two operators $U_g(X_{d-p-1})U_{g'}(X_{d-p-1})$ can be deformed to $U_{g'}(X_{d-p-1})U_g(X_{d-p-1})$ as shown in figure 2.3

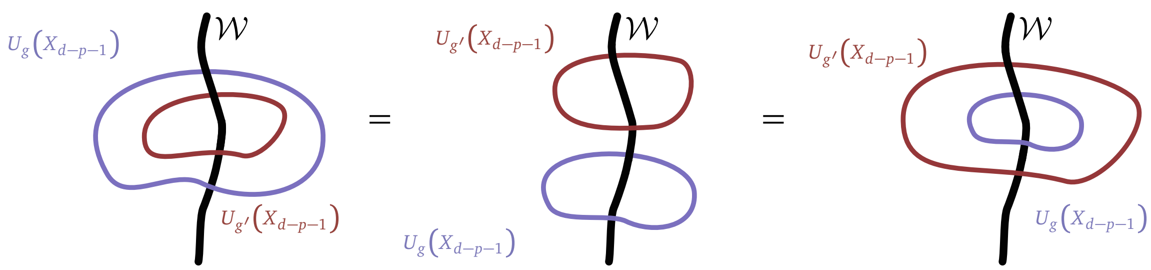


Figure 2.3: Higher-form symmetries are are abelian

2.2.1 Lower-form symmetries

Instead of cranking up the form-degree of the conserved current we could consider a zero-form, i.e. a scalar current, corresponding to setting $p=-1$ in the context of p-form symmetries. Associated to such a current there is a charge supported on a codimension-zero manifold, i.e. on a volume in spacetime:

$$Q(\Sigma_d)=\int_{\Sigma_d\subset X_d}\star J_{[0]} \tag{2.2.4}$$

$$\mathrm{d}\star J_{[0]}=0. \tag{2.2.5}$$

In the more abstract universal presentation through topological symmetry operators, the symmetry operators of such a (-1)-form symmetry are supported on a d-dimensional manifold. Therefore, what is charged under (-1)-form symmetries are objects whose excitations are zero-dimensional, i.e. instantons.

Interestingly, (-1)-form symmetries need not have group structure; it suffices that they are just sets. One can arrive at this conclusion through the more mathematical description of higher-form symmetries, through Eilenberg-MacLane spaces. For a p-form symmetry associated to a group or set, G, the mathematical structure that one uses to construct the field data is the Eilenberg-MacLane space $\mathrm{B}^{p+1}G:=K(G,p+1)$, whose defining characteristic is that

$$\pi_n\left(\mathrm{B}^{p+1}G\right)=G\delta_{p+1,n},$$

i.e. only the $(p+1)$-st homotopy group is non-trivial and equal to G. For $p=0$, only the fundamental group matters. The fundamental group is firstly a group, while secondly, it can be non-abelian. For $p>1$ we only find higher-homotopy groups, which are necessarily abelian — giving another view to why higher-form symmetries are abelian. However, for $p=-1$, the relevant structure is $\pi_0(G)$, which need not even be a group. From a physical viewpoint, the group structure of zero- and higher-form symmetries arises from the ability to fuse topological operators, while the non-abelian structure of zero-form symmetries arises from the existence or not of contact terms when one exchanges the order in which symmetry operators act. For (-1)-form symmetries, topological operators fill up a whole part of spacetime so one cannot fuse them without having to cope with an ambiguous volume-worth of contact terms.

Albeit (-1)-form symmetries look obscure, their background gauge fields are simply (space-dependent) coupling constants. To see this, consider an action such as

$$S[\phi;\lambda]=S_0[\phi]+\int_{X_d}\lambda\mathcal{O}. \tag{2.2.6}$$

Since $\mathcal{O}$ is a d-form, it is naturally closed. Therefore one can define a zero-form conserved current as

$$J_{[0]} := \star\mathcal{O}. \tag{2.2.7}$$

Then, making λ spacetime dependent amounts to background gauging the theory, with λ as the gauge field. To understand the action of a (-1)-form symmetry one can put the theory on a twisted sector. There the background defects for a p-form symmetry become of codimension-$(p+2)$ and hence for (-1)-form symmetries they are codimension-1; they are just domain walls across which the value of λ jumps.

2.2.2 Dimensional Reduction of higher-form symmetries

Upon dimensionally reducing a d-dimensional theory with a p-form symmetry, $G^{[p]}$, on a circle we are lead to a $G^{[p]}$ and a $G^{[p-1]}$ in the $(d-1)$-dimensional theory. To illustrate this, notice that the charged p-dimensional operators of the d-dimensional theory, can either wrap or not wrap the circle. Those that wrap the circle will become $(p-1)$-dimensional operators in the $(d-1)$-dimensional theory, whereas those that don't wrap the circle will remain p-dimensional in the new theory. Their charges will be inferred by the original theory. In even more generality, in a similar fashion, upon a dimensional reduction on some $\mathbf{T}^k$, to land on a $(d-k)$ dimensional theory, the total symmetry group will be

$$G^{[p]} \longmapsto G^{[p]} \times G^{[p-1]} \times \cdots \times G^{[p-\min\{k,p\}]}, \tag{2.2.8}$$

where each factor arises from the number of non-trivial cycles the operators in the dimensionally reduced theory wrap.

2.3 Turning on Background Fields

An equivalent way to study the symmetries of a theory is to couple the theory to flat background gauge fields, A, and consider the partition function, $\mathcal{Z}[A]$, as a functional of the background gauge fields. In theories with a continuous p-form symmetry, i.e. with a conserved current $(p+1)$-form, $J_{[p+1]}$ this coupling amounts to including in the action the term

$$i\int_{X_d} A_{[p+1]} \wedge \star J_{[p+1]} \tag{2.3.1}$$

plus possible corrections that are quadratic in A to restore gauge invariance. We see that the gauge field is a $(p+1)$-form connection taking values in G. In general, since it is a background field, it suffices to restrict ourselves to flat background connections. Therefore in what follows we will draw our background gauge fields from $\mathrm{H}^{p+1}(X_d, G)$. We use the cohomology group to explicitly throw pure gauge configurations out, since of course

$$\int_{X_d} \mathrm{d}\Lambda_{[p]} \wedge \star J_{[p+1]} = (-1)^{p+1} \int_{X_d} \Lambda_{[p]} \wedge \mathrm{d} \star J_{[p+1]} = 0. \tag{2.3.2}$$

In the case of discrete symmetry, most of the discussion follows from above, namely background gauge fields are flat $(p+1)$-form connections, drawn from $\mathrm{H}^{p+1}(X_d, G)$. The difference is that now we do not have a conserved current, so the way to couple them to the theory is usually through some invariant $(d-p-1)$-form (call it $\omega_{[d-p-1]}$) that our theory possesses

$$i \int A_{[p+1]} \smile \omega_{[d-p-1]}. \tag{2.3.3}$$

Here $\smile$, the cup product, is the analogue of $\wedge$ that is correct for forms valued in cohomology. A comment is in place here. If we have a pairing based on $\wedge$, it descends to a pairing based on $\smile$, by just restricting to cohomology classes. This does not work the other way around, i.e. a pairing with $\smile$ does not extend uniquely to a pairing with $\wedge$. The correct way to think even about continuous symmetries is through a cup product. $\int A \wedge \star J$ should be in fact written as $\int [A] \smile [\star J]$. Of course, we will not clutter the notation with confusing classes and cups, and we will stick to the standard notation with wedge product, bearing in mind this subtlety.

Coming back to the main discussion, we should think of the coupling of background fields both in the discrete and in the continuous case, through Poincaré duality. Specifically, Poincaré asserts that

$$\int_{B_{p+1}} A_{[p+1]} = \int A_{[p+1]} \smile \hat{B}_{[d-p-1]}, \tag{2.3.4}$$

where $B_{[p+1]} \in \mathrm{H}_{p+1}(X_d, G)$ and $A_{[p+1]} \in \mathrm{H}^{p+1}(X_d, G)$. Then $\hat{B}_{[d-p-1]} \in \mathrm{H}^{d-p-1}(X_d, G)$ is the Poincaré dual of $B_{[p+1]}$. Drawing again an analogy with the continuous case, where

$$\int_{X_d} A_{[p+1]} \wedge \star J_{[p+1]} = \int_{\hat{A}_{[d-p-1]}} \star J, \tag{2.3.5}$$

we should define the partition function in the presence of a background gauge field, $A_{[p+1]}$, universally as

$$\mathcal{Z}\big[A_{[p+1]}\big] := \Big\langle U_g\big(\hat{A}_{d-p-1}\big)\Big\rangle. \tag{2.3.6}$$

In the above expression, $\hat{A}_{[d-p-1]}$ is the Poincaré dual of $A_{[p+1]}$, g is an arbitrary reference group element and $\langle\cdots\rangle$ is taken with respect to the original theory.

In the following paragraphs, we will see various examples of higher-form symmetries and how to study them by coupling to background fields.

2.3.1 Maxwell theory

The first example of a theory with higher-form symmetries is certainly Maxwell theory. Consider the Maxwell action in d dimensions, of a $\mathrm{U}(1)$ dynamical gauge field, $a \in \Omega^1(X_d, \mathrm{U}(1))$:

$$S[a] = \frac{1}{2e^2}\int_{X_d} \mathrm{d}a \wedge \star\, \mathrm{d}a = \frac{1}{2e^2}\int_{X_d} f \wedge \star f. \tag{2.3.7}$$

The classical equations of motion and the Bianchi identities, on the absence of sources give

$$\begin{aligned} \mathrm{d}\star f &= 0 \\ \mathrm{d}\ f &= 0, \end{aligned}$$

allowing us to identify a f as a 2-form conserved current giving rise to an "electric" one-form global symmetry $\mathrm{U}(1)^{[1]}_{\mathrm{e}}$ and $\star f$ as a $(d-2)$-form current, corresponding to a "magnetic" $(d-3)$-form global symmetry, $\mathrm{U}(1)^{[d-3]}_{\mathrm{m}}$.

For the electric part, the one-form symmetry is the invariance under shifts $a-\omega$ where ω is a one-form flat conenction. We can promote this symmetry to (background) gauge symmetry by allowing ω to have some curvature, as long as we compensate for that with a 2-form background gauge field $B \in \mathrm{H}^2(X_d, \mathrm{U}(1))$, whose gauge transformations $B \longmapsto B + \mathrm{d}\omega$ cancel the unwanted curvature. Therefore the action with the electric one-form symmetry turned on is

$$S[a,B,C] = \frac{1}{2e^2}\int_{X_d} (\mathrm{d}a - B)\wedge \star(\mathrm{d}a - B) \tag{2.3.8}$$

The magnetic $(d-3)$-form symmetry does not have such a description in terms of local fields, therefore the way to couple a background field, $C \in \mathrm{H}^{d-2}(X_d, \mathrm{U}(1))$ for it is through minimal coupling $C \wedge f$. Notice however that with B already turned on, this coupling becomes $C \wedge (\mathrm{d}a - B)$. This extra B in the coupling is the starting kick of a mixed anomaly in Maxwell theory, that we will revisit in section 3.3.1. In total the action with backgrounds for both symmetries turned on reads

$$S[a,B,C] = \frac{1}{2e^2} \int_{X_d} (\mathrm{d}a - B) \wedge \star(\mathrm{d}a - B) + \frac{i}{2\pi} \int_{X_d} C \wedge (\mathrm{d}a - B). \tag{2.3.9}$$

2.3.2 Chern–Simons theory

Another example of higher-form symmetries that we will frequently invoke is three dimensional Chern–Simons theory. A Chern–Simons theory comes with a gauge group, $\mathcal{G}$ and a level $k \in \mathbb{Z}$, written usually together as $\mathcal{G}_k$. Mathematically, it also comes with a $\mathcal{G}$-bundle. When the latter is the trivial bundle, we can globally define a connection, a, on it and the theory is defined by the action

$$S_{\mathrm{CS},\mathcal{G}_k}[a] := \frac{k}{4\pi} \int_{X_3} \mathrm{Tr}_{\mathcal{G}} \underbrace{\left(a \wedge \mathrm{d}a + \frac{2}{3} a \wedge a \wedge a \right)}_{\substack{\shortparallel \\ \mathrm{CS}[a]}}. \tag{2.3.10}$$

When the bundle is non-trivial though we cannot define a connection globally and so we have to consider the Chern–Simons theory on a four-dimensional manifold, Y_4 such that $\partial Y_4 = X_3$, by defining

$$S_{\mathrm{CS},\mathcal{G}_k}[a] := \frac{k}{4\pi} \int_{Y_4} \mathrm{Tr}_{\mathcal{G}} \left(f \wedge f \right), \tag{2.3.11}$$

where f is the curvature of a chosen on patches, obeying compatability conditions on overlaps. When $k \in \mathbb{Z}$ the action (2.3.11) is uniquely defined, modulo 1, independently of the choice of Y_4 and the extension of a and the $\mathcal{G}$-bundle.

However, any $\mathcal{G}$-bundle on a three-dimensional manifold is necessarily trivial when $\mathcal{G}$ is a connected, simply connected compact Lie group, and since we will mainly be interested in the cases of $\mathrm{U}(1)$ and $\mathrm{SU}(N)$, we can use the action in the form (2.3.10) and we can safely forget about (2.3.11).

In what follows, we examine the global symmetries of the Chern–Simons theory.

$\mathrm{U}(1)_k$ Chern–Simons

Consider $\mathrm{U}(1)_k$ Chern–Simons theory,

$$S[a] = \frac{k}{4\pi}\int_{X_3} a \wedge \mathrm{d}a\,. \tag{2.3.12}$$

It has a $\mathbb{Z}_k^{[1]}$ symmetry[2], which can be seen by measuring the charges of the Wilson line operators of the theory,

$$W(\gamma) := \exp\left(in\oint_\gamma a\right), \qquad n \in \mathbb{Z}_k. \tag{2.3.13}$$

The fact that $n \sim n+k$ can be seen from the fact that (2.3.12) is invariant under the shift $a \longmapsto a + \frac{1}{k}C$, with C a properly quantised flat $\mathrm{U}(1)$ connection. Alternatively if we compactify (2.3.12) on a circle, we find a theory with a Lagrangian

$$\frac{k}{4\pi}a_3\ \mathrm{d}a\,,$$

which has a $\mathbb{Z}_k^{[0]}$ symmetry.

Now if we want to turn on a background gauge field for the $\mathbb{Z}_k^{[1]}$ symmetry all we have to do is add a 2-form background gauge field B, as

$$S[a,B] = \frac{k}{4\pi}\int a\wedge \mathrm{d}a + \frac{1}{2\pi}\int a\wedge B. \tag{2.3.14}$$

Note that it looks as if B is $\mathrm{U}(1)$ valued; however, let us demonstrate that this in not the case. Suppose that $B = k\widetilde{B}$, where $\widetilde{B}$ is a properly quantised background field. We can write $\widetilde{B}$ locally as a curvature

$$\widetilde{B} \approx \mathrm{d}C\,. \tag{2.3.15}$$

Upon a change of variables in the path integral, $a \longmapsto a - C$ the action takes the form

$$\begin{aligned} S[a,B] \longmapsto S[a,C] &= \frac{k}{4\pi}\int (a-C)\wedge \mathrm{d}(a-C) + \frac{1}{2\pi}\int (a-C)\wedge \mathrm{d}(kC) = \\ &= \underbrace{\frac{k}{4\pi}\int a\wedge \mathrm{d}a}_{S[a]} - \underbrace{\frac{k}{4\pi}\int C\wedge \mathrm{d}C}_{\text{counterterm}}, \end{aligned} \tag{2.3.16}$$

which makes evident that background fields of the form $B = k\widetilde{B}$, do not affect the original action. Hence B is $\mathbb{Z}_k$-valued, up to counterterms.

[2]this symmetry will be proven to be anomalous, but let us not worry about it for now

$\mathrm{SU}(N)_k$ Chern–Simons

Non-abelian $\mathrm{SU}(N)_k$ Chern–Simons theory has a $\mathbb{Z}_N^{[1]} = Z\big(\mathrm{SU}(N)\big)^{[1]}$ symmetry generated by the Wilson lines

$$W(\gamma) = \mathrm{Tr}_{k/N}\, \wp \exp\left(im \oint a \right), \tag{2.3.17}$$

in the $j = k/N$ representation. A background field B for this symmetry can be turned on via

$$S[a,B] = \frac{k}{4\pi} \int \mathrm{CS}[a] + \frac{1}{2\pi} \int w_2 \smile B, \tag{2.3.18}$$

where w_2 is the second Stiefel–Whitney class, measuring precisely the obstruction of lifting a $\mathrm{PSU}(N) = \mathrm{SU}(N)/\mathbb{Z}_N$ bundle to an $\mathrm{SU}(N)$ bundle.

2.3.3 Yang–Mills theory and the importance of bundle topology

If one wants to explore the global symmetries of Yang–Mills theory, the choice of the gauge group is of crucial importance. On a very formal level, we will see that if $\mathcal{E}$ is an extension by the centre $Z(\mathcal{G})$, of $\mathcal{G}$,

$$1 \longrightarrow Z(\mathcal{G}) \longrightarrow \mathcal{E} \longrightarrow \mathcal{G} \longrightarrow 1, \tag{2.3.19}$$

then a Yang–Mills theory with gauge group $\mathcal{E}$ will have $Z(\mathcal{G})$ as a one-form global symmetry, whereas a Yang–Mills theory with gauge group $\mathcal{G}$ will not, by virtue of the different bundles the two theories can accommodate.

To unpack the takaway message hidden in the above cryptic paragraph, let us consider four-dimensional SU(2) Yang–Mills theory:

$$S[a] = \frac{1}{2g^2} \int_{X_4} \mathrm{Tr}\big(f \wedge \star f\big) \tag{2.3.20}$$

It has a $\mathbb{Z}_2^{[1]} = Z\big(\mathrm{SU}(2)\big)^{[1]}$ symmetry, akin to the centre symmetry of the $\mathrm{SU}(N)_k$ Chern–Simons theory discussed above. A background gauge field for the centre symmetry should be a flat 2-form field $B \in \mathrm{H}^2\big(X_4, \mathbb{Z}_2\big)$. The SU(2) bundles are labelled by the instanton number of the theory,

$$\frac{1}{8\pi^2} \int_{X_4} \mathrm{Tr}\big(f \wedge f\big) =: I \in \mathbb{Z}. \tag{2.3.21}$$

The full path integral of the theory reads then:

$$\mathcal{Z}_{\mathrm{SU}(2)} = \underbrace{\sum_{I\in\mathbb{Z}} \underbrace{\int \mathrm{D}a}_{\substack{\text{connection on}\\ \text{a fixed bundle}}}}_{\text{sum over bundle topologies}} \mathrm{e}^{-S[a]}. \tag{2.3.22}$$

However, for a theory with gauge group SO(3) instead, there is a new invariant, the second Stiefel–Whintey class, $w_2 \in \mathrm{H}^2(X_4, \mathbb{Z}_2)$, which describes a $\mathbb{Z}_2$-valued magnetic flux, on a two-surface,

$$\int_{\Sigma_2} w_2 \in \mathbb{Z}_2. \tag{2.3.23}$$

For a very concrete example consider $a = a^{\mathfrak{a}} \dfrac{\sigma_{\mathfrak{a}}}{2}$, with $\mathfrak{a} \in \{1,2,3\}$, be the Lie algebra indices and let only $a^3 \neq 0$. Then the magnetic flux is

$$\int_{\Sigma_2} w_2 = \int_{\Sigma_2} \frac{\mathrm{d}a^3}{2\pi} \bmod 2. \tag{2.3.24}$$

We see that the SO(3) theory has more bundles labelled by the magnetic fluxes. The full SO(3) path integral is

$$\mathcal{Z}_{\mathrm{SO}(3)} = \underbrace{\sum_{w_2\in\mathrm{H}^2\left(X_4,\mathbb{Z}_2\right)} \underbrace{\sum_{I\in\mathbb{Z}} \underbrace{\int \mathrm{D}a}_{\substack{\text{connection on}\\ \text{a fixed bundle}}}}_{\text{sum over bundle topologies}}}_{\text{sum over magnetic fluxes}} \mathrm{e}^{-S[a]}. \tag{2.3.25}$$

We see then that all the bundles of SU(2) are special SO(3) bundles for which $w_2 = 0 \in \mathrm{H}^2(X_4, \mathbb{Z}_2)$.

For $\mathrm{SU}(N)$ versus $\mathrm{PSU}(N) := \mathrm{SU}(N)/Z\big(\mathrm{SU}(N)\big) = \mathrm{SU}(N)/\mathbb{Z}_N$, in four dimensions, the story is exactly similar; the $\mathrm{PSU}(N)$ gauge theory has the generalised second Stiefel–Whitney class as an invariant measuring the $\mathbb{Z}_N$ magnetic flux, $w_2 \in \mathrm{H}^2\big(X_4, \mathbb{Z}_N\big)$, which means that

$$\mathcal{Z}_{\mathrm{PSU}(N)} = \sum_{w_2\in\mathrm{H}^2\left(X_4,\mathbb{Z}_N\right)} \sum_{I\in\mathbb{Z}} \int \mathrm{D}a\, \mathrm{e}^{-S[a]}. \tag{2.3.26}$$

$$\mathcal{Z}_{\mathrm{SU}(N)} = \sum_{I\in\mathbb{Z}} \int \mathrm{D}a\, \mathrm{e}^{-S[a]}. \tag{2.3.27}$$

On the other hand, we have already mentioned that SU(N) Yang–Mills theory has a $\mathbb{Z}_N^{[1]}$ symmetry. This means that we can turn on a background gauge field $B \in \mathrm{H}^2(X_4, \mathbb{Z}_N)$ for this one-form symmetry. Since turning on background fields amounts to twisting the bundles, we can turn the background field on by turning on the Stiefel–Whitney class $w_2 \in \mathrm{H}^2(X_4, \mathbb{Z}_N)$ but constraining it to $w_2 = B$, i.e. Therefore we have:

$$\mathcal{Z}_{\mathrm{SU}(N)}[B] = \sum_{\substack{I\in\mathbb{Z}\\ w_2=B}} \int \mathrm{D}a\ \mathrm{e}^{-S[a]} \equiv \sum_{I\in\mathbb{Z}} \int \mathrm{D}a\ \mathrm{e}^{-S[a]} \exp\left(i \int w_2 \smile B \right). \tag{2.3.28}$$

Gauging[3] the one-form symmetry amounts to summing over all background fields and thus

$$\mathcal{Z}_{\mathrm{SU}(N)/\!/\mathbb{Z}_N^{[1]}} = \sum_{B\in\mathrm{H}^2(X_4,\mathbb{Z}_N)} \mathcal{Z}_{\mathrm{SU}(N)}[B] = \sum_{B\in\mathrm{H}^2(X_4,\mathbb{Z}_N)} \sum_{\substack{I\in\mathbb{Z}\\ w_2=B}} \int \mathrm{D}a\ \mathrm{e}^{-S[a]} = \mathcal{Z}_{\mathrm{PSU}(N)}. \tag{2.3.29}$$

The story can be made more general. A gauge theory of a simply connected gauge group $\mathcal{G}$ is again characterised by the instanton number and a generalised magnetic flux invariant, $w_2 \in \mathrm{H}^2(X_4, Z(\mathcal{G}))$[4]. It has a non-anomalous one-form symmetry $\Gamma^{[1]} \subset Z(\mathcal{G})^{[1]}$ (if there are no matter fields transforming under $\Gamma^{[1]}$) and by gauging it one obtains:

$$\mathcal{Z}_{\mathcal{G}/\!/\Gamma^{[1]}} = \mathcal{Z}_{\mathcal{G}/\Gamma}. \tag{2.3.30}$$

2.4 Higher-form symmetry breaking

We have seen by now that higher-form symmetries are much like zero-form symmetries and should not be thought of as exotic or non-standard. To strengthen this view, we ought to mention the existence of theorems regarding the spontaneous breaking of higher-form symmetries which are modified versions of the Coleman–Mermin–Wagner theorem and the Goldstone theorem. In this section, we will briefly review the content of these theorems.

[3] we denote gauging by // to remind us of orbifolding for reasons that will become apparent in chapter 4

[4] it can be seen by remembering that $Z(\mathcal{G}) = \frac{\Lambda_{\mathrm{weight}}}{\Lambda_{\mathrm{root}}}$

2.4.1 Higher Goldstone theorem

A modification of Goldstone theorem for higher-form symmetries appeared in [4] and in [5] using different approaches, but the claim was in both cases the following.

A spontaneously broken continuous p-form symmetry implies the existence of a massless excitation in the spectrum, a higher-form Goldstone mode. The Goldstones of the spontaneously broken symmetry are p-form gauge fields.

More precisely, in the symmetry-broken phase, the conserved current, $J_{[p+1]}$, is realised nonlinearly in terms of a p-form gauge field $a_{[p]}$ as $J_{[p+1]} = v^2 \mathrm{d}a_{[p]}$, with action

$$S[a] = -v^2 \int_{X_d} \mathrm{d}a_{[p]} \wedge \star \mathrm{d}a_{[p]}. \tag{2.4.1}$$

Applied to ordinary Maxwell theory, higher Goldstone theorem has the remarkable consequence that the real-world photon's masslessness is protected by the spontaneous breaking of the electric one-form symmetry.

2.4.2 Higher Coleman–Mermin–Wagner theorem

The Mermin–Wagner theorem for higher-form symmetries [17] states the following.

A continuous p-form symmetry in d dimensions cannot be spontaneously broken if $d \leqslant p+2$. Similarly, a discrete p-form symmetry in d dimensions cannot be spontaneously broken if $d \leqslant p+1$.

A sketch of a proof of this theorem goes as follows. Consider a theory with a p-form symmetry, $G^{[p]}$ in d dimensions and compactify on various p-dimensional tori $\mathbf{T}^p$, in such a way that each time there is at least one $G^{[0]}$ factor. Then one can use the standard Coleman–Mermin–Wagner argument to claim that $G^{[0]}$ is unbroken if $d-p \leqslant 2$ ($d-p \leqslant 1$ for discrete symmetries) and thus the original $G^{[p]}$ is unbroken if $d \leqslant p+2$ ($d \leqslant p+1$ for discrete symmetries).

3 Anomalies

The anomaly,
anomaly, anomaly...
Oh, oh, oh

— Princess Nokia, Anomaly

3.1 ’t Hooft anomalies

Consider a theory with a global symmetry, G and couple it to a G-background gauge field, A. Consider a general gauge transformation $A \longmapsto A^{\lambda}$. Naïvely we expect the partition function to be invariant under this transformation. However, in some cases it might transform as

$$\mathcal{Z}[A^\lambda] = \exp\left(i\int_{X_d} \alpha[A,\lambda]\right)\mathcal{Z}[A]. \tag{3.1.1}$$

We see that the partition function has a phase ambiguity, which is nevertheless controllable. Turning off the gauge field the ambiguity vanishes, so since the gauge field is not dynamical, we are not really at fear. The ambiguity does, however, mean concretely that we *cannot* make the gauge field dynamical, i.e. we cannot gauge the theory. This controllable ambiguity goes by the name *'t Hooft anomaly*.

Before we declare the theory anomalous, though, notice that we can always add local counterterms, by changing the regularisation scheme. Different regularisation schemes give rise to a similar shift in the partition function. Therefore we might be able to exploit the counterterms to absorb the above phase. When we fail to do so, we say that the theory is indeed anomalous, and we define the anomaly to be those shifts in the partition function modulo identification by counterterms.

The last sentence hints greatly at cohomology. We can reformulate the argument in terms of free energies, where the cohomology appears more naturally. Following Komargodski [18], the argument goes as follows. In a quantum field theory one can compute the free energy in the presence of a background field, $\mathcal{W}[A] := -\log \mathcal{Z}[A]$, only up to local counterterms,

$$\Omega[A] := \int \omega[A],$$

identifying

$$\mathcal{W}[A] \sim \mathcal{W}[A] + \Omega[A]. \tag{3.1.2}$$

Then, we say that the theory is 't Hooft anomalous, when under a gauge transformation $A \longmapsto A^\lambda$,

$$\mathcal{W}[A^\lambda] - \mathcal{W}[A] = \alpha[A,\lambda], \tag{3.1.3}$$

modulo identifications by (3.1.2). This clearly shows that picking out anomalies is a problem in cohomology; (3.1.3) is the closure condition and (3.1.2) is the exactness identification. The relation between anomalies and cohomology will be made concrete in the next section, where we will unpack what anomalies mean in quantum mechanics and it will be generalised in section 3.2 for anomalies of QFTs in arbitrary dimensions.

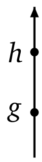
3.1.1 Projective representations

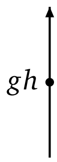
Consider a very simple system; a $(0+1)$-dimensional QFT — also known as quantum mechanics — with a symmetry G. For simplicity, let us take the Hamiltonian of the system to be $H=0$, and the Hilbert space, $\mathcal{H}$ to be finite-dimensional. In QFT terms we could refer to it as a $(0+1)$-dimensional TQFT.

In quantum mechanics two states that differ by a phase are equivalent, i.e. $\mathcal{H} \ni |\psi\rangle \sim e^{i\gamma}|\psi\rangle$. This implies that if the symmetry group is in some representation ρ, we can allow for representations that obey the rule:

$$\rho(g)\rho(h) = \rho(gh)\, \exp\big(i\gamma(g,h)\big),. \tag{3.1.4}$$

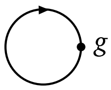
Such representations are known as *projective representations*, and we will show that they are a particularly simple case where an anomaly can arise. In the spirit of turning on background gauge fields, or equivalently putting symmetry defects, we can rephrase equation (3.1.4) in pictures as



$$= \exp\big(i\gamma(g,h)\big)$$


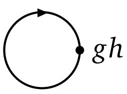
(3.1.5)

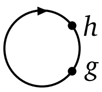
The partition function, at the presence of a defect is:

$$\mathcal{Z}[g] := \mathrm{Tr}_{\mathcal{H}}\big(\rho(g)e^{-H}\big) = \mathrm{Tr}_{\mathcal{H}}\,\rho(g) =$$


(3.1.6)

Therefore we see that

$$\mathcal{Z}[gh] =$$


$$= \exp\big(i\gamma(g,h)\big)$$


$$= \exp\big(i\gamma(g,h)\big)\, \mathcal{Z}[g,h],$$
(3.1.7)

so the partition function suffers by a controllable phase ambiguity, i.e. the theory is anomalous.

If we now take three such defects we have that

$$\begin{aligned}
&\rho(g)\rho(h)\rho(k) = \rho(gh)\rho(k)\, \exp\big(i\gamma(g,h)\big) = \rho(ghk)\, \exp\big(i\gamma(g,h)\big)\, \exp\big(i\gamma(gh,k)\big)\\
&\qquad\qquad |||\\
&\rho(g)\rho(h)\rho(k) = \rho(g)\rho(hk)\, \exp\big(i\gamma(h,k)\big) = \rho(ghk)\, \exp\big(i\gamma(h,k)\big)\, \exp\big(i\gamma(g,hk)\big).
\end{aligned}$$

Therefore we need the phases γ to satisfy

$$\gamma(g,h) + \gamma(gh,k) = \gamma(g,hk) + \gamma(h,k) \quad \mod 2\pi. \tag{3.1.8}$$

i.e. γ is a U(1) valued 2-cocycle of G. However we are allowed to redefine $\rho(g)$ by a phase, $\widetilde{\rho}(g) := \rho(g)\exp\big(i\beta(g)\big)$, and $\widetilde{\rho}(g)$ describes the same physics as $\rho(g)$. This leads to

$$\begin{aligned} \widetilde{\rho}(g)\widetilde{\rho}(h) &= \widetilde{\rho}(g,h)\exp\big(i\widetilde{\gamma}(g,h)\big), \quad \text{with} \\ \widetilde{\gamma}(g,h) &= \gamma(g,h) + \beta(g) + \beta(h) - \beta(gh). \end{aligned} \tag{3.1.9}$$

We recognise (3.1.9) as shifting γ by a coboundary, bearing in mind the definition of discrete differentials. This gives the algebraic definition of the second group cohomology group, also known to mathematicians as *Borel*, or *Borel–Moore* group cohomology ($\mathscr{H}^{\bullet}$). The Borel cohomology is equivalent to the usual, topological, cohomology ($\mathrm{H}^{\bullet}$) of the classifying space of the group in question.

Therefore we are led to the following claim:

Projective representations/anomalies in $(0+1)$ dimensions are classified by γ, satisfying (3.1.8) modulo identification by (3.1.9). In other words they are classified by

$$[\gamma] \in \mathscr{H}^2\big(G, \mathrm{U}(1)\big) \equiv \mathrm{H}^2\big(BG, \mathrm{U}(1)\big).$$

Here, let us clarify a potentially confusing point. The group structure of the symmetry, namely the fusion algebra of the topological operators instructs us to take $U_g U_h \overset{!}{=} U_{gh}$. However, the fusion algebra is allowed to hold projectively on the states, as explained above. Moreover, since the ambiguity is multiplicative, it cancels at the operator level, as the operators of the theory transform as $\mathcal{O} \longmapsto g\mathcal{O}g^{-1}$. Altogether we have that in the presence of a non-zero $[\gamma]$, all operators are in faithful representations of the group whereas the states are in projective representations (with the same $[\gamma]$).

The whole intricate structure of anomalies in QFT stems from this simple fact about quantum mechanics, decorated by the complexity that higher dimensions bring in the game. We will see that all known anomalies of quantum field theories are classified as well, using group cohomology, or some refined, generalised cohomology.

Another indicator of what is about to become a general motif, is to adopt a bulk point-of-view. We can think of quantum mechanics as living on the boundary of a $(1+1)$-dimensional theory. There, the merging of the defects looks like figure 3.1. The bulk

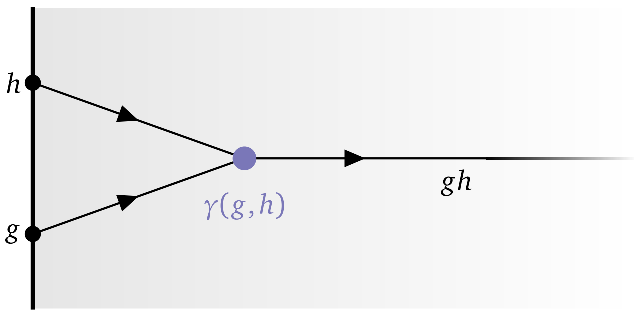
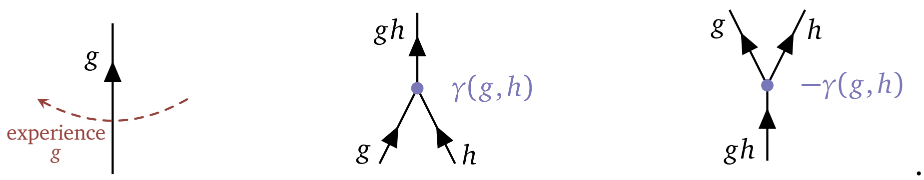
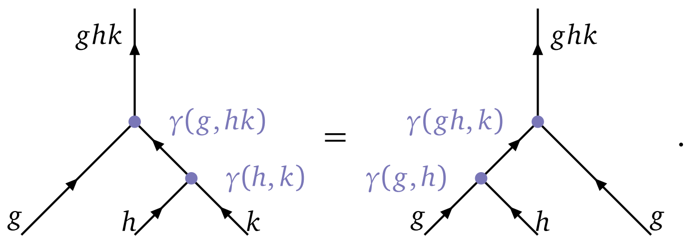


Figure 3.1: Merging of defects in one-higher dimension

theory describes an ungauged Dijkgraaf–Witten theory [19] in $(1 + 1)$ dimensions. We can describe it by the following rules:

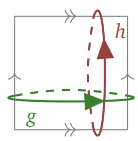
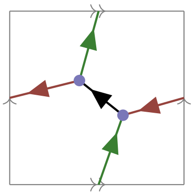


Then the relation (3.1.8) takes on the following geometric interpretation:



(3.1.10)

If we take the two dimensional manifold to be a torus, we can caluclate various partition functions. For example:

$$\mathcal{Z}_{\mathbb{T}^2}[g,h] := \mathcal{Z}\Bigg[$$  $$\Bigg] =$$  $$= e^{i\gamma(g,h)}\, e^{-i\gamma(h,g)} =$$

$$= \exp\big(i[\gamma(g,h)-\gamma(h,g)]\big) =: \exp(i\mathrm{SPT}(g,h)). \tag{3.1.11}$$

The above picture generalises. When the bulk theory is placed on a closed manifold it is an invertible field theory. Namely, there exists a $\widetilde{\mathcal{Z}}_{Y_2}[g,h]$), such that $\mathcal{Z}_{Y_2}[g,h]\,\widetilde{\mathcal{Z}}_{Y_2}[g,h] = 1$. Moreover, when the manifold Y_2 is cut open, the theory has edge modes which can be exactly cancelled placing the anomalous quantum mechanics we started with, with appropriate insertions on the boundary. Such bulk theories are known as symmetry protected topological (SPT) orders/phases and it is another general theme in the study of anomalies, as well as the topic of section 3.2.

Before going there though, let us work through some examples of potential anomalies in quantum mechanics.

- $\boldsymbol{G = \mathbb{Z}_2}$**:** This one is annoyingly simple. In particular $\mathrm{H}^2\big(\mathrm{B}\mathbb{Z}_2, \mathrm{U}(1)\big) = 0$, so it cannot accomodate projective representations.

- $\boldsymbol{G = \mathbb{Z}_2 \times \mathbb{Z}_2}$**:** Here we have $\mathrm{H}^2\big(\mathrm{B}(\mathbb{Z}_2 \times \mathbb{Z}_2), \mathrm{U}(1)\big) = \mathbb{Z}_2$ so it can accomodate a unique projective representation. If we label the generators of the first $\mathbb{Z}_2$ as $\{0, g\}$ and of the second as $\{0, h\}$ we get that the unique projective representation is given by

$$\rho(g) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \rho(h) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \rho(gh) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.1.12}$$

  we can easily read off from here that $\gamma(g, h) = \pi \bmod 2\pi$. and the rest $\gamma(\bullet, \circ)$ are trivial. For the bulk theory, the partition function that corresponds to (3.1.11), evaluates to -1. More generally we could turn on background fields for each factor of the group $A \in \mathrm{H}^1\Big(N_2, \overset{g}{\mathbb{Z}_2}\Big)$, $B \in \mathrm{H}^1\Big(N_2, \overset{h}{\mathbb{Z}_2}\Big)$ and write an SPT action as

$$\mathrm{SPT}[a, b] = \pi i \int_{N_2} A \smile B. \tag{3.1.13}$$

  Such an anomaly can arise from a variety of starting points.

  - A particle on a circle, with Lagrangian

  $$L = \frac{1}{2}\dot{q}^2 + \frac{i\theta}{2\pi}\dot{q},$$

    with $q \sim q + 2\pi$ has a symmetry $\mathbb{Z}_2 \times \mathbb{Z}_2$, with such an anomaly at $\theta = \pi$. The anomaly implies that all energy levels are degenerate. Of course one can explicitly solve this problem and find that $E_n = \frac{1}{2}\big(n - \frac{\theta}{2\pi}\big)^2$, verifying that energy levels cross. However the result holds for an arbitrary deformation $L \longmapsto L + V(\phi)$, such that $V(\phi)$ is $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetric at $\theta = \pi$.

  - It describes the phenomenon of discrete torsion on orbifolds [20].

  - It describes how a D-brane with discrete torsion carries a projective representation of the discrete torsion phase [21].

3.2 Anomalies as Symmetry Protected Topological phases

In this section, we will see how and why we can think of anomalies of an arbitrary QFT as SPT phases in one higher dimension, generalising the quantum mechanical paradigm of the previous passage.

We can describe the anomalies of a d-dimensional QFT, via a classical local *invertible action* in $(d+1)$ dimensions, that is the low-energy QFT of a *symmetry protected topological phase* (SPT). SPT phases originate in condensed matter and are short-range entangled gapped phases of matter, meaning that they can all be deformed to the same trivial product state if we break the symmetry. For our purposes we will use the phrases "SPT phase" and "invertible field theory" interchangeably, as we are only interested in the low-energy description.

To understand anomalies from the bulk point of view, consider the following setup. Given X_d, the spacetime manifold, we can choose a $(d+1)$ dimensional manifold, Y_{d+1} such that $\partial Y_{d+1} = X_d$, on which there is a local lagrangian $(d+1)$-form, $\mathcal{A}[A]$, such that

$$\exp\left(2\pi i \int_{Y_{d+1}} \mathcal{A}\left[A^{\lambda}\right] - 2\pi i \int_{Y_{d+1}} \mathcal{A}[A]\right) = \exp\left(2\pi i \int_{X_d} \alpha[A,\lambda]\right). \tag{3.2.1}$$

We can then define the *partition function of the anomaly theory* as

$$\mathcal{Z}_{\mathrm{SPT}}[A] := \exp\left(2\pi i \int_{Y_{d+1}} \mathcal{A}[A]\right). \tag{3.2.2}$$

Extending the partition function of the original theory into the classical bulk as

$$\widetilde{\mathcal{Z}}[A] := \mathcal{Z}[A]\mathcal{Z}_{\mathrm{SPT}}[A], \tag{3.2.3}$$

where $\mathcal{Z}$ is calculated on Y_{d+1}, we have that

$$\widetilde{\mathcal{Z}}\left[A^{\lambda}\right] = \widetilde{\mathcal{Z}}[A], \tag{3.2.4}$$

so now the theory is gauge invariant at the cost of specifying the data in the bulk.

3.3 Mixed anomalies

If the theory has a symmetry group given by $\Gamma = G \times H$ (or more generally $1 \longrightarrow G \longrightarrow \Gamma \longrightarrow H \longrightarrow 1$) there can be a mixed anomaly between G and H, showing up as follows. When one turns on background gauge fields for both G and H and performs a background gauge transformation on the background G-gauge field, there appears a non-trivial phase in front of the partition function, involving the background H-gauge field:

$$\mathcal{Z}[A_G, A_H + \mathrm{d}\lambda_H] = \exp\left(i \int \alpha[A_G, \lambda_H] \right) \mathcal{Z}[A_G, A_H], \tag{3.3.1}$$

where A_G and A_H are the background gauge fields for G and H respectively and λ_H is a gauge shift of the H-background field.

For continuous symmetries, the mixed anomaly has another presentation, far more common than the transformation of the partition function and this is through the anomalously conserved currents, J_G, J_H. The mixed anomaly appears as

$$\begin{aligned} \mathrm{d} \star J_G &= 0 \\ \mathrm{d} \star J_H &= f[A_G] \end{aligned} \qquad \text{or} \qquad \begin{aligned} \mathrm{d} \star J_G &= g[A_H] \\ \mathrm{d} \star J_H &= 0, \end{aligned}$$

where f and g are arbitrary functions of the gauge fields and their derivatives, such that $f[0] = g[0] = 0$.

3.3.1 Mixed anomaly in Maxwell theory

A particularly interesting example with a mixed anomaly is Maxwell theory, in d dimensions. Consider the Maxwell action of a U(1) dynamical gauge field, $a \in \Omega^1(X_d, \mathrm{U}(1))$:

$$S[a] = \frac{1}{2e^2} \int_{X_d} \mathrm{d}a \wedge \star \mathrm{d}a = \frac{1}{2e^2} \int_{X_d} f \wedge \star f. \tag{3.3.2}$$

The classical equations of motion and the Bianchi identities, on the absence of sources give

$$\begin{aligned} \mathrm{d} \star f &= 0 \\ \mathrm{d} \; f &= 0, \end{aligned}$$

allowing us to identify a f as a 2-form conserved current giving rise to an "electric" one-form global symmetry $\mathrm{U}(1)_{\mathrm{e}}^{[1]}$ and $\star f$ as a $(d-2)$-form current, corresponding to a "magnetic" $(d-3)$-form global symmetry, $\mathrm{U}(1)_{\mathrm{m}}^{[d-3]}$. Let us now turn on background fields B and C for $\mathrm{U}(1)_{\mathrm{e}}^{[1]}$ and $\mathrm{U}(1)_{\mathrm{m}}^{[d-3]}$ respectively:

$$S[a,B,C]=\frac{1}{2e^2}\int_{X_d}(\mathrm{d}a-B)\wedge\star(\mathrm{d}a-B)+\frac{i}{2\pi}\int_{X_d}C\wedge(\mathrm{d}a-B). \tag{3.3.3}$$

Under a background gauge transformation:

$$a\longmapsto a+\eta,\qquad B\longmapsto B+\mathrm{d}\eta\,,\qquad C\longmapsto C+\mathrm{d}\chi\,, \tag{3.3.4}$$

the action transforms as

$$S[a,B,C]\longmapsto S[a,B,C]+\frac{i}{2\pi}\int_{X_d}\mathrm{d}\chi\wedge(\mathrm{d}a-B)=S[a,B,C]-\frac{i}{2\pi}\int_{X_d}\mathrm{d}\chi\wedge B. \tag{3.3.5}$$

This is precisely our anomaly; At the level of the partition function it shows up as

$$\mathcal{Z}[B^\eta,C^\chi]=\exp\left(-\frac{i}{2\pi}\int_{X_d}\mathrm{d}\chi\wedge B\right)\mathcal{Z}[B,C], \tag{3.3.6}$$

whereas at the level of the currents it shows up as

$$\begin{aligned}\mathrm{d}\star f&=0\\ \mathrm{d}\ f&=\mathrm{d}B.\end{aligned}$$

Alternatively, we can introduce a classical $(d+1)$-dimensional bulk, Y_{d+1} and place an SPT phase,

$$\mathcal{Z}_{\mathrm{SPT}}[B,C]:=\exp\left(\frac{i}{2\pi}\int_{Y_{d+1}}C\wedge\mathrm{d}B\right). \tag{3.3.7}$$

The SPT phase transforms under a background gauge transformation as

$$\begin{aligned}\mathcal{Z}_{\mathrm{SPT}}[B^\eta,C^\chi]&=\exp\left(\frac{i}{2\pi}\int_{Y_{d+1}}(C+\mathrm{d}\chi)\wedge\big(\mathrm{d}(B+\mathrm{d}\eta)\big)\right)=\\ &=\exp\left(\frac{i}{2\pi}\int_{Y_{d+1}}\mathrm{d}\chi\wedge\mathrm{d}B\right)\mathcal{Z}_{\mathrm{SPT}}[B,C],\end{aligned} \tag{3.3.8}$$

thus cancelling the anomaly on the boundary, since

$$\mathcal{Z}[B^\eta,C^\chi]\mathcal{Z}_{\mathrm{SPT}}[B^\eta,C^\chi]=\mathcal{Z}[B,C]\mathcal{Z}_{\mathrm{SPT}}[B,C]. \tag{3.3.9}$$

We can take it one step further. Since the choice of Y_{d+1} is arbitrary, up to the fact that $\partial Y_{d+1} \stackrel{!}{=} X_d$, we can choose a different bulk, Y'_{d+1} and place the same SPT action. Now if there exists W_{d+2} such that $\partial W_{d+1} = Y_{d+1} \sqcup \overline{Y'}_{d+1}$ we can naturally rewrite the difference of the two SPT actions as

$$\frac{i}{2\pi}\int_{Y_{d+1}\sqcup\overline{Y'}_{d+1}} C \wedge \mathrm{d}B = 2\pi i \int_{W_{d+2}} \frac{\mathrm{d}B}{2\pi} \wedge \frac{\mathrm{d}C}{2\pi}. \tag{3.3.10}$$

Defining a $(d+2)$ form, called the anomaly polynomial:

$$I_{d+2} := \frac{\mathrm{d}B}{2\pi} \wedge \frac{\mathrm{d}C}{2\pi}. \tag{3.3.11}$$

3.3.2 General higher-form mixed anomaly

Now suppose that we have a theory in d dimensions, with continuous p- and q-form symmetries, with a mixed anomaly between them. Since higher-form symmetries are always abelian, each factor is a U(1) and assuming no non-trivial mixing of the symmetries (cf. section 4.3) the whole symmetry is $\mathrm{U}(1)^{[p]} \times \mathrm{U}(1)^{[q]}$. The story is mimicking the case of the Maxwell theory, with different form degrees. Let us turn on background gauge fields $B_{[p+1]}$, $C_{[q+1]}$ for $\mathrm{U}(1)^{[p]}$ and $\mathrm{U}(1)^{[q]}$ respectively.

If $d = p+q+2$, then the SPT phase that captures the anomaly is given by

$$\mathcal{Z}_{\mathrm{SPT}}[B,C] = \exp\left(\frac{i}{2\pi}\int_{Y_{d+1}} C_{[q+1]} \wedge \mathrm{d}B_{[p+1]}\right) \quad \text{or} \tag{3.3.12}$$

$$= \exp\left(\frac{i}{2\pi}\int_{Y_{d+1}} \mathrm{d}C_{[q+1]} \wedge B_{[p+1]}\right). \tag{3.3.13}$$

3.4 Classification of Symmetry Protected Topological phases

As of now, we have seen that a quantum field theory that has a symmetry might be anomalous, so we must be extra careful before drawing conclusions based on the symmetry. The natural question that arises here is *when can a symmetry be anomalous?* And the follow-up question is *how many distinct anomalies can it have?*

To answer these questions, we can turn to the presentation of anomalies through Symmetry Protected Topological phases, and understand their classification. We can then infer from those a classification of anomalies.

The main goal of this section is to give a review of the classification of SPT phases, together with new directions and ongoing yet-incomplete pursuits, in order to preclude or allow and understand certain anomalies in the chapters to follow.

3.4.1 Mild SPT phases and group cohomology

For the time being, we will restrict ourselves to the classification of SPT phases/anomalies associated to a zero-form symmetry group, and we will see how some of the results generalise (as well as how some others do not generalise) to higher-form SPT phases/anomalies in a later subsection.

In the preliminary example of anomalies in quantum mechanics, in section 3.1.1, we saw that anomalies of a symmetry group $G^{[0]}$ were classified by the group cohomology groups $\mathcal{H}^2(G, \mathrm{U}(1)) \cong \mathrm{H}^2(\mathrm{B}G, \mathrm{U}(1))$. We also saw that the two-dimensional SPT phase corresponding to the quantum mechanical anomaly was a 2d ungauged Dijkgraaf–Witten theory and this was essentially the link between the SPT phase and group cohomology in this baby example.

Cranking up dimensions, the argument presented in [22] and [23] was that even in higher dimensions, bosonic SPT phases correspond to ungauged Dijkgraaf–Witten theories of higher dimensions. Namely consider a bosonic SPT phase in $(d+1)$ dimensions, protected by a discrete symmetry group, G. On a branched triangulation[5] of the manifold, X_d where the theory is defined, we can write its partition function as

$$\mathcal{Z}_{\mathrm{SPT}}(X_d) = \frac{1}{|G|^{|\triangle_0|}} \sum_{\{g\}} \exp\big(-S[\{g\}]\big), \tag{3.4.1}$$

with

$$\exp\big(-S[\{g\}]\big) = \prod_{\triangle_{d+1}} \triangle_{d+1}^{s_{i_0 i_2 \cdots i_{d+1}}} \big(g_{i_0} g_{i_2} \cdots g_{i_{d+1}}\big). \tag{3.4.2}$$

In the above expressions $|G|$ is the number of elements in G and $|\triangle_0|$ is the number of

[5] i.e. a triangulation together with a choice of orientation for each edge of an n-simplex, such that there is no oriented loop on any triangle.

vertices on the triangulation. $\triangle_{d+1}$ defines a $(d+1)$-simplex,

$$\triangle_{d+1} : \underbrace{G \times G \times \cdots \times G}_{d+2 \text{ times}} \longrightarrow \mathrm{U}(1)$$

satisfying

$$\triangle_{d+1}(gg_0, gg_1, \ldots, gg_{d+1}) = \triangle_{d+1}(g_0, g_1, \ldots, g_{d+1}) \quad \text{and} \tag{3.4.3}$$

$$\prod_{k=0}^{d+1} \triangle_{d+1}^{(-1)^k}\big(g_0, \cdots, g_{k-1}, g_{k+1, \cdots, g_{d+1}}\big) = 1. \tag{3.4.4}$$

Also, $s_{i_0 \cdots i_{d+1}} = \pm 1$ depending on the orientation of the simplex. We should think of $\triangle_{d+1}$ as the $(d+1)$-dimensional generalisation of $\mathrm{e}^{i\gamma(g,h)}$ that we had in section 3.1.1; a $(d+1)$-cocycle. Then we if we combine (3.4.3) and (3.4.4) we get the $(d+1)$-dimensional generalisation of (3.1.8) written in multiplicative notation. We can further see that making the transformation

$$\triangle_{d+1} \longmapsto \triangle_{d+1} \prod \widetilde{\triangle}_d^{s_{\ldots}},$$

with $\widetilde{\triangle}_d$ a d-cocycle, applied on appropriate group elements, $\exp\big(-S[\{g_i\}]\big)$ stays intact. This is the $(d+1)$-dimensional version of (3.1.9). It is precisely the algebraic condition for the $(d+1)$-st Borel group cohomology group of G. Therefore we get the claim:

Mild bosonic SPT phases in $(d+1)$ dimensions/'t Hooft anomalies of bosonic QFTs in d dimensions with symmetry G are classified by

$$\mathscr{H}^{d+1}\big(G, \mathrm{U}(1)\big) \cong \mathrm{H}^{d+1}\big(\mathrm{B}G, \mathrm{U}(1)\big). \tag{3.4.5}$$

At this point, we would like to comment on a few facts. Firstly this procedure applies to what we call *mild* bosonic SPT phases, i.e. most SPT phases of bosonic systems. It fails to classify fermionic SPTs because, from the quantum field theory point of view, we need an extra ingredient in (3.4.2); a choice of spin structure. If we choose a spin structure and proceed in the same way, we will end up with a group supercohomology that classifies fermionic SPT phases [24]. More importantly, it fails to classify certain SPT phases involving time-reversal, as noted in [25–27] — these SPT phases we will call *wild*, and we will classify using cobordisms.

Secondly, (3.4.2) is precisely a $(d+1)$ dimensional ungauged Dijkgraaf–Witten model and so gauging G amounts to summing over inequivalent configurations, i.e. over $\mathrm{H}^{d+1}(\mathrm{B}G, \mathrm{U}(1))$ elements. This is exactly the Dijkgraaf–Witten procedure cranked up in dimensions. A third comment regards the generalisation to continuous groups. The above argument was formulated for discrete groups, but it is directly generalised to be applicable also for continuous groups, by replacing $\sum_{\{g_i\}}$ by a path integral $\int \mathrm{D}g$ and the product over simplices by a continuous product.

3.4.2 Wild SPT phases and cobordisms

Since the bag of generalised cohomology theories is open, we might explore other possibilities to classify SPT phases beyond group cohomology. The first such instance of classification was initiated by Kapustin in [28] for bosonic SPT phases and soon after was modified to include fermionic SPT phases [29].

Here we will proceed using more field-theoretic techniques, rather than condensed-matter theoretic, following [30]. The work of Freed [31] permits us to do so, by showing that the low energy effective description of a condensed matter system with short-range entanglement is, in fact, a unitary invertible topological field theory. Such condensed matter systems are precisely those that we call SPT orders [32]. Therefore, the classification of SPT phases amounts to the classification of unitary invertible field theories — modulo certain equivalences which we will point when the time is right.

Let us now define the relevant terms, to arrive to the main claim. As in chapter 1, denote by $\bullet$ the appropriate structure for the systems we would like to classify. For example, $\bullet$ consists of a spin structure and a G background gauge field for systems with fermion number symmetry and an internal symmetry G. We define the $\bullet$-bordism group $\Omega_d^\bullet$ in dimension d to be

$$\Omega_d^\bullet := \left\{ \begin{matrix} d\text{-dimensional manifolds} \\ \text{with structure } \bullet \end{matrix} \right\} \Big/ \sim, \tag{3.4.6}$$

where the equivalence relation is defined so that $X_d \sim X'_d$ if and only if there exists a Y_{d+1} with the structure $\bullet$, such that ∂Y_{d+1} consists of X_d as incoming boundary and X'_d as outgoing boundary. Define then the *cobordism group* of d dimensional $\bullet$-manifolds as the Pontryagin dual to $\Omega_d^\bullet$:

$$\Omega_\bullet^d := \mathrm{Hom}\big(\Omega_d^\bullet, \mathrm{U}(1)\big). \tag{3.4.7}$$

Since $\bullet$ consists of various components we might want to change part of it, but not all. In particular we will be interested in changing the internal symmetry groups, so it will be practical to adopt the notation accordingly. We will denote as $\Omega_d^\bullet(\mathrm{B}G)$, the bordism group of $\bullet$-manifolds, equipped with a map to $\mathrm{B}G$. As an example, a structure $\tilde{\bullet}$ consisting of a spin structure and a $\mathbb{Z}_2$ gauge field can be seen as having a structure $\tilde{\bullet} = (\bullet, \mathrm{B}G) = (\text{spin structure}, \mathrm{B}\mathbb{Z}_2)$. Therefore, $\Omega_d^{\tilde{\bullet}} = \Omega_d^{\mathrm{Spin}}(\mathrm{B}\mathbb{Z}_2)$. Furthermore if pt denotes a point we have that $\Omega_d^\bullet = \Omega_d^\bullet(\mathrm{pt})$. The cobordism group is then,

$$\Omega_\bullet^d(\mathrm{B}G) := \mathrm{Hom}\big(\Omega_d^\bullet(\mathrm{B}G), \mathrm{U}(1)\big). \tag{3.4.8}$$

The claim of [30] is that

Wild SPT phases in $(d+1)$ dimensions/'t Hooft anomalies in d dimensions with symmetry G are classified by

$$\Omega_{\bullet}^{d+1}(\mathrm{B}G) = \mathrm{Hom}\big(\Omega_{d}^{\bullet}(\mathrm{B}G), \mathrm{U}(1)\big). \tag{3.4.9}$$

The main points towards the proof of (3.4.9) are the following. Any bordism $X_d \sim X'_d$, can be decomposed into elementary procedures of removing $\mathbf{S}^k \times \mathbf{D}^{d-k}$ from X_d and pasting $\mathbf{D}^{k+1} \times \mathbf{S}^{d-k-1}$ back to it, using a Morse function [33]. Here $\mathbf{D}^p$ is a p-dimensional ball, with boundary $\partial\mathbf{D}^p = \mathbf{S}^{p-1}$. The next main point is to use unitarity, and get that

$$\mathcal{Z}\big(\Delta\big(\mathbf{D}^d \times \mathbf{S}^{d-1}\big)\big) := \mathcal{Z}\big(\mathbf{D}^d \times \mathbf{S}^{d-1}\big)\mathcal{Z}\Big(\overline{\mathbf{S}^{d-1} \times \mathbf{D}^d}\Big) = \mathcal{Z}\big(\mathbf{S}^d\big) > 0,$$

where ΔM is the double of M, produced by gluing M to $\overline{M}$ along their common boundary, where $\overline{M}$ denotes orientation reversal. In fact, taking $\mathbf{S}^d$ as $\partial\mathbf{D}^{d+1}$ we get that $\mathcal{Z}\big(\mathbf{S}^d\big) = 1$[6]. Then the theorem of [30] identifies elements of $\Omega_{\bullet}^{d+1}(\mathrm{B}G)$ with unitary functors $\mathcal{Z}(Y_{d+1}) : \mathcal{H}(X_d) \longrightarrow \mathcal{H}(X'_d)$, where $\mathcal{H}(X_d)$ and $\mathcal{H}(X'_d)$ are one-dimensional, i.e. with unitary invertible field theories.

Certain comments are again in place here. Firstly, the assumption of unitarity is crucial for the cobordism classification of SPT phases. Unitarity arises naturally when one deals with Lorentzian field theory but it is something that has to be imposed in Euclidean theories. If one relaxes unitarity there can be counterexamples; namely non-unitary invertible theories that are not bordism invariants. One such example is presented in [34]. It is a "massive bc ghost system" in $d = 4n+1$ dimensions. Its action is given by

$$S[b,c] = \int_{X_d} b \wedge (\star\mathrm{d} + \mathrm{d}\star)c + m b \wedge \star c, \tag{3.4.10}$$

where b and c are sections of $\mathcal{S} \otimes \mathcal{S}^*$, where $\mathcal{S}$ is the spin bundle over X_d and $\mathcal{S}^*$ its dual [7]. Regularising the theory with a Pauli–Villars mass M and later taking $m = -M$ and $M \longrightarrow \infty$ we obtain

$$\begin{aligned}\mathcal{Z}(X_d) = \frac{\det(\star\mathrm{d} + \mathrm{d}\star + m)}{\det(\star\mathrm{d} + \mathrm{d}\star + M)} = \exp\big(-2\pi i \eta(\star\mathrm{d} + \mathrm{d}\star)\big)\Big|_{\text{acting on even forms}} = \\ = (-1)^{\sum_{\ell} \dim \mathrm{H}^{2\ell}(X_d)}.\end{aligned} \tag{3.4.11}$$

[6] automatically for odd d; choosing an appropriate Euler term for even d

[7] note that the theory is bosonic, despite b and c being fermionic, since $\mathcal{S} \otimes \mathcal{S}^* \cong \bigoplus_{i=0}^{n/2} \bigwedge^{2i} \mathrm{T}^* X_d$, i.e. we do not need a spin structure to define the theory

In particular, we get that $\mathcal{Z}(\mathbf{S}^d) = -1 < 0$, which is inconsistent with $\mathbf{S}^d$ being a bordism invariant.

A second comment regards a refinement of $\Omega_\bullet^{d+1}$, to classify more generic anomalies. In particular, $\Omega_\bullet^{d+1}$ classifies theories whose partition function depend only on the topological data of background fields. Therefore it is suitable for finite groups and for global anomalies, but it does not detect perturbative anomalies. We can account for the latter by considering the Anderson dual of $\Omega^\bullet_{d+1}$: $\mho_\bullet^{d+1}$. The latter classifies the deformation classes of SPT phases, which can depend continuously on the background fields. Formally the Anderson dual is obtained by the universal coefficient theorem for some generalised cohomology; namely for normal (co)homology we have

$$0 \longrightarrow \mathrm{Ext}(\mathrm{H}_d(X,\mathbb{Z}),\mathbb{Z}) \longrightarrow \mathrm{H}^{d+1}(X,\mathbb{Z}) \longrightarrow \mathrm{Hom}(\mathrm{H}_{d+1}(X,\mathbb{Z}),\mathbb{Z}),$$

while for a generalised (co)homology, such as (co)bordisms we take

$$0 \longrightarrow \mathrm{Ext}\big(\Omega^\bullet_d(X,\mathbb{Z}),\mathbb{Z}\big) \longrightarrow \mho_\bullet^{d+1}(X,\mathbb{Z}) \longrightarrow \mathrm{Hom}\big(\Omega^\bullet_{d+1}(X,\mathbb{Z}),\mathbb{Z}\big).$$

Notice that albeit $\Omega_\bullet^{d+1} \neq \mho_\bullet^{d+1}$, their torsion part coincides. This was first noted by Freed and Hopkins in [3] in the context of extended invertible theories but it also applies to non-extended theories. As a final comment we would like to mention that for $d < 3$, $\mathrm{H}_{d+1} \cong \Omega^{\mathrm{O}}_{d+1}$, therefore in dimensions less than three the group cohomology classification is complete.

3.4.3 Towards higher-SPT phases

These rich results dealt with ordinary, zero-form symmetries. However there can of course be anomalous higher-form symmetries. Coming back to subsection 3.4.1, we can modify equations (3.4.1, 3.4.2) so that

$$\mathcal{Z}_{\mathrm{SPT}}(X_d) = \frac{1}{|G|^{|\triangle_0|}} \sum_{k \in C^0(X_d,G)} \prod_{\triangle_{d+1}} \triangle_{d+1}(k), \tag{3.4.12}$$

where k as a zero-cochain encodes now the information about the group elements and the orientation of the triangulation. This form is now immediately generalisable to higher-form symmetries [35]. For a p-form symmetry, the degrees of freedom that live on the vertices of the triangulations are not local anymore, but rather p-extended. Therefore they should be described by p-form connections,

$$k_{[p]} \in \widetilde{C}^p(X_d,G) := \frac{C^p(X_d,G)}{\mathrm{d}\widetilde{C}^{p-1}(X_d,G)}, \tag{3.4.13}$$

where $\widetilde{C}^1(X_d, G) \cong C^1(X_d, G)$. Then the partition function of an SPT phase protected by a p-form symmetry can be written as

$$\mathcal{Z}_{\mathrm{SPT}}(X_d) = \frac{1}{|G|^{|\triangle_{0\to p}|}} \sum_{k_{[p]}} \prod_{\triangle_{d+1}} \triangle_{d+1}\big(k_{[p]}\big). \tag{3.4.14}$$

In the above expression $\left|\triangle_{0\to p}\right| := \sum_{k=0}^{p} (-1)^k \left|\triangle_{p-k}\right|$. In this language, the correct cocycle conditions modulo identifications that $\triangle_{d+1}(k_{[p]})$ satisfy give rise to $\mathrm{H}^{d+1}(\mathrm{B}^{p+1}G, \mathrm{U}(1))$. Therefore the equivalent of claim (3.4.5) is

> Mild bosonic SPT phases in $(d+1)$ dimensions/'t Hooft anomalies of bosonic QFTs in d dimensions with p-form symmetry $G^{[p]}$ are classified by
>
> $$\mathrm{H}^{d+1}\big(\mathrm{B}^{p+1}G, \mathrm{U}(1)\big). \tag{3.4.15}$$

In the above expression, $\mathrm{B}^{p+1}G = K(G, p+1)$, is an Eilenberg–MacLane space, obtained by taking the classifying space $p+1$ times, as reviewed in Appendix B.

A minor technical comment regarding the above classification is that we needed to shift to the geometric description with cochains instead of group elements. For the case of zero-form symmetries, we have that $\mathscr{H}^{d+1}(G, \mathrm{U}(1)) \cong \mathrm{H}^{d+1}(\mathrm{B}G, \mathrm{U}(1))$ so an algebraic description with group elements gives rise to the Borel cohomology which can be reduced to the topological cohomology simply by taking a classifying space. The topological cohomology could be independently obtained by the geometric description with cochains. For higher-form symmetries we can turn the topological cohomology to a Borel cohomology by removing a classifying space, namely $\mathrm{H}^{d+1}\big(\mathrm{B}^{p+1}G, \mathrm{U}(1)\big) \cong \mathscr{H}^{d+1}(\mathrm{B}^pG, \mathrm{U}(1))$, but this would amount to a (totally unnatural) mixed geometroalgebraic description. In principle we should be able to keep removing classifying spaces until we arrive to a purely algebraic description, $\mathrm{H}^{d+1}\big(\mathrm{B}^{p+1}G, \mathrm{U}(1)\big) \cong \mathscr{H}^{d+1}(\mathrm{B}^pG, \mathrm{U}(1)) \cong \cdots \cong \mathscr{H}^{d+1}(G, \mathrm{U}(1))$. It is not known however what type of cohomology would $\mathscr{H}$ be. This question is interesting in algebraic topology, but it is uninteresting from a physical viewpoint.

To classify deformation classes of wild higher-SPT phases, that fall outside the group cohomology classification, Wan and Wang proposed [36] an extension of the work of Freed and Hopkins [3] namely:

Deformation classes of wild SPT phases in $(d+1)$ dimensions/'t Hooft anomalies of d-dimensional QFTs, with a p-form symemtry $G^{[p]}$ are classified by

$$\mho_{\bullet}^{d+1}\left(\mathrm{B}^{p+1}G^{[p]}\right). \tag{3.4.16}$$

In fact, the claim in [36] includes also higher-groups i.e. mixtures of higher-form symmetries of different degrees (see also chapter 4). However, their claim has only been checked for certain one-form symmetries and few 2-group symmetries. This is partly because the way to calculate cobordism groups one has to run multiple spectral sequences that do not collapse very quickly and partly because higher classifying spaces are generally not known even for quite common groups. It would be worth looking into obtaining a classification using the Pontryagin dual, instead of the Anderson dual similar to (3.4.9) for higher-symmetries.

3.4.4 Categorical approach

In the modern understanding, both topological order and symmetries correspond to categories. Chapter 6 is dedicated to studying symmetries from a categorical perspective and in Appendix B we lay out the basic notions of (higher-) category theory, so for the rest of this subsection we will not define here all the notions we borrow from category theory. Seen through the categorical lense, classifying topological phases with a symmetry arises naturally in a categorical setting. Such a classification of SPTs appeared recently in [37].

The main points towards this classification is the boundary-bulk relation for topological orders [38, 39] and the condensation completion for higher categories [40]. Let us focus on zero-form symmetries for the time being. The idea behind the boundary-bulk relation for topological orders is that one can recover the fusion d-category of bulk codimension-1 defects, $\mathbf{C}$, from the fusion d-category of boundary codimension-1 defects, $\mathbf{B}$, through a two-step procedure:

$$\mathbf{C} = \Sigma\mathfrak{Z}_1(\mathbf{B}). \tag{3.4.17}$$

Here $\mathfrak{Z}_1$ denotes the Drinfeld centre, while Σ is the condensation completion. To obtain the Drinfeld centre, notice that if one brings the trivial defect, id, (i.e. the absence of a defect) from the bulk to the boundary, the boundary does not change. Thus all codimension-2 defects, that correspond to the delooping of $\mathbf{C}$: $\Omega\mathbf{C} \cong \mathrm{Hom}(\mathsf{id}, \mathsf{id})$ can be safely brought to the boundary. Then there exists naturally a central[8] monoidal functor

$$
\begin{aligned}
F : \Omega\mathbf{C} &\longrightarrow \mathbf{B} \\
X &\longmapsto F(X) := X \otimes \mathrm{id},
\end{aligned}
$$

where $\otimes$ is the fusion of two defects. Hence one defines the Drinfeld centre as

$$
\mathfrak{Z}_1(\mathbf{B}) := \left\{ \begin{array}{c} \text{maximal braided monoidal} \\ d\text{-category such that there} \\ \text{exists a central monoidal functor} \\ F : \mathfrak{Z}_1(\mathbf{B}) \longrightarrow \mathbf{B} \end{array} \right\}.
$$

The condensation completion is morally the d-categorical analogue of the linear algebra procedure to construct a vector space from a set of basis vectors, by taking all different linear combinations. Descriptively one defines the condensation completion as

$$
\Sigma\mathbf{B} := \left\{ \begin{array}{c} \text{include all (d+1)-dimensional defects} \\ \text{that can be obtained via the} \\ \text{condensation of lower-dimensional defects} \end{array} \right\}
$$

Mathematically it is described in [40] as the Karoubi completion of the one-point delooping, $\mathrm{B}_{\mathrm{pt}}\mathbf{B}$ of $\mathbf{B}$:

$$
\Sigma\mathbf{B} := \mathrm{Kar}\big(\mathrm{B}_{\mathrm{pt}}\mathbf{B}, \mathrm{pt}\big).
$$

The one-point delooping is the $(d+1)$-category which has only one object, pt, and has the objects of $\mathbf{B}$ as 1-morphisms, 1-morphisms of $\mathbf{B}$ as 2-morphisms and so on.

To understand a topological phase with a symmetry G, one can focus on the charges, i.e. (one-dimensional) excitations of G, given by $\mathbf{Rep}G$, the fusion category of the representations of G. The symmetry charges can condense on a line or a plane, or a higher-dimensional surface to give rise to higher dimensional defects, that we can obtain through condensation completion:

$$
\begin{aligned}
&\text{2d defects:} && \Sigma\mathbf{Rep}G =: \mathbf{Rep}_2(G) \\
&\text{3d defects:} && \Sigma^2\mathbf{Rep}G =: \mathbf{Rep}_3(G) \\
&\quad\vdots && \\
&d\text{d defects:} && \Sigma^{d-1}\mathbf{Rep}G =: \mathbf{Rep}_d(G).
\end{aligned}
$$

Therefore, given a $(d+1)$-dimensional anomaly-free topological phase[9], given by a functor $C : \mathbf{1}_{d+2} \longrightarrow \mathbf{1}_{d+2}$ and let $\mathbf{C} = \mathrm{Hom}(C, C)$ be the category of all d-dimensional defects in C.

[8] i.e. for all $X \in \Omega\mathbf{C}$ and $Y \in \mathbf{B}$: $F(X) \otimes Y \cong Y \otimes F(X)$

[9] note that the topological phase being anomaly free does not imply that the theory on its boundary is anomaly free

The fact that C has a symmetry G means that $\mathbf{C}$ is a fusion d-category over $\mathbf{Rep}_d(G)$. Namely, the symmetry charges and their higher-dimensional condensation descendants are included in $\mathbf{C}$, i.e. there exists a natural embedding

$$\mathbf{Rep}_d(G) \hookrightarrow \mathbf{C}. \tag{3.4.18}$$

Furthermore, the codimension-2 excitations in the bulk $\mathbf{1}_{d+2}$ of $\mathbf{C}$ must be a subcategory of the Drinfeld centre, $\mathfrak{Z}_1(\mathbf{C})$, i.e. there exists a natural embedding

$$\mathbf{Rep}_d(G) \hookrightarrow \mathfrak{Z}_1(\mathbf{C}). \tag{3.4.19}$$

Finally, requiring compatibility of the Drinfeld centre with $\mathbf{Rep}G$ we demand that

$$\begin{tikzcd} & \mathbf{Rep}_d(G) \arrow[dl, hook'] \arrow[dr, hook] & \\ \mathfrak{Z}_1(\mathbf{C}) \arrow[rr] & & \mathbf{C} \end{tikzcd} \tag{3.4.20}$$

commutes.

In the above, $\mathbf{Rep}_d(G) \hookrightarrow \mathfrak{Z}_1(\mathbf{C})$ fixes the bulk of $\mathbf{C}$. For a trivial phase its bulk should be the trivial bulk $\mathbf{Rep}_d(G) \hookrightarrow \mathfrak{Z}_1\big(\mathbf{Rep}_d(G)\big)$. Moreover, since $\mathbf{C}$ are the defects of an anomaly-free phase, its bulk should be the trivial bulk, namely there is a braided equivalence

$$\mathfrak{Z}_1\big(\mathbf{Rep}_d(G)\big) \xrightarrow{\sim} \mathbf{C}, \tag{3.4.21}$$

such that

$$\begin{tikzcd} & \mathbf{Rep}_d(G) \arrow[dl, hook'] \arrow[dr, hook] & \\ \mathfrak{Z}_1\big(\mathbf{Rep}_d(G)\big) \arrow[rr, "\sim"] & & \mathfrak{Z}_1(\mathbf{C}) \end{tikzcd} \tag{3.4.22}$$

commutes.

The main result of [37] is that those categories, $\mathbf{C}$ with a braided equivalence (3.4.21) such that (3.4.22) holds, classify of $(d+1)$-dimensional G-Symmetry Enriched Topological (SET) phases up to invertible phases without symmetry. As a corollary to that, we get the classification of G-SPTs as follows. If we take $\mathrm{Hom}(C,C) \equiv \mathbf{C} = \mathbf{Rep}_d(G)$, then C is an invertible phase, i.e. it describes an SPT phase. Therefore we have the following claim.

G-SPT phases in $(d+1)$-dimensions are classified by the autoequivalences

$$\mathfrak{Z}_2(\mathbf{Rep}_d(G)) \xrightarrow{\sim} \mathfrak{Z}_2(\mathbf{Rep}_d(G)),$$

that preserve the embedding

$$\mathbf{Rep}_d(G) \hookrightarrow \mathfrak{Z}_1\big(\mathbf{Rep}_d(G)\big).$$

These are denoted by $\mathrm{Aut}\big(\mathfrak{Z}_1(\mathbf{Rep}_d(G)), \mathbf{Rep}_d(G)\big)$. In other words

$$\left\{\begin{matrix}(d+1)\text{-dimensional} \\ G\text{-SPT phases}\end{matrix}\right\} = \mathrm{Aut}\big(\mathfrak{Z}_1(\mathbf{Rep}_d(G)), \mathbf{Rep}_d(G)\big). \tag{3.4.23}$$

This is an extension of the group cohomology classification of SPT phases, since one can show that

$$\mathrm{Aut}(\mathfrak{Z}_1(\mathbf{Rep}(G)), \mathbf{Rep}(G)) = \mathrm{H}^2(\mathrm{B}G, \mathrm{U}(1)) \tag{3.4.24}$$

$$\mathrm{Aut}\big(\mathfrak{Z}_1(\mathbf{Rep}_2(G)), \mathbf{Rep}_2(G)\big) = \mathrm{H}^3(\mathrm{B}G, \mathrm{U}(1)) \tag{3.4.25}$$

$$\mathrm{Aut}\big(\mathfrak{Z}_1(\mathbf{Rep}_d(G)), \mathbf{Rep}_d(G)\big) \supset \mathrm{H}^{d+1}(\mathrm{B}G, \mathrm{U}(1)), \qquad d \geqslant 3. \tag{3.4.26}$$

4 Gauging and ungauging

Gauging is not a physical process, but rather a process performed by physicists

— Eric Sharpe, [41]

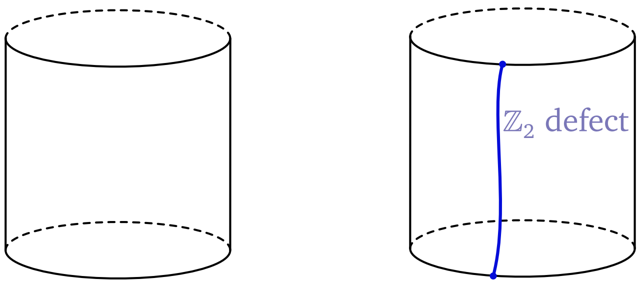
If one is given a QFT, with some global symmetry, they might be tempted to gauge it. A global symmetry can have objects charged under it, while a gauge redundancy cannot. What happens to those objects upon gauging? The goal of this chapter is to show that the information about them stays intact — at least when the symmetry group is a finite abelian group. Another danger lingers when one wants to gauge an anomalous symmetry. We know of course that it cannot be gauged, but a non-anomalous subgroup of it may still be gaugable. These issues are better viewed through identifying gauging with orbifolding.

4.1 Gauging as orbifolding in 2d

Let us review the well-understood [42] gauging $\equiv$ orbifolding story in two dimensions, in order to generalise to arbitrary dimensions. We will start with the most straightforward example of a $\mathbb{Z}_2$ symmetry and afterwards we will gauge a general finite abelian group.

4.1.1 $\mathbb{Z}_2$ example

We will start with a very concrete example, a two-dimensional theory, Q, acted by a *non-anomalous* $\mathbb{Z}_2$, on $\Sigma_2 = \mathbb{R} \times \mathbf{S}^1$. For the Hilbert space $\mathcal{H}_{\mathrm{Q}}$, we can consider either the untwisted sector $(+)$, or the twisted sector $(-)$. These can be viewed as if we implemented a $\mathbb{Z}_2$ codimension-1 defect on Σ_2.



Then the Hilbert space of the two sectors is correspondingly

$$\mathcal{H}_{\mathrm{Q}}^{+} = \mathcal{H}_{\mathrm{Q}}^{++} \oplus \mathcal{H}_{\mathrm{Q}}^{+-} \tag{4.1.1}$$

$$\mathcal{H}_{\mathrm{Q}}^{-} = \mathcal{H}_{\mathrm{Q}}^{-+} \oplus \mathcal{H}_{\mathrm{Q}}^{--}, \tag{4.1.2}$$

where $\pm$ is the charge of the states in the corresponding Hilbert space under $\mathbb{Z}_2$. Now since $\mathbb{Z}_2$ is non-anomalous we can orbifold by it. The Hilbert space of the orbifolded theory $\mathrm{Q}/\!/\mathbb{Z}_2$ is obtained by keeping only the $\mathbb{Z}_2$-even (or only the $\mathbb{Z}_2$-odd) parts of both the twisted and the untwisted sectors. Doing that we find

$$\mathcal{H}_{\mathrm{Q}/\!/\mathbb{Z}_2}^{+} = \mathcal{H}_{\mathrm{Q}}^{++} \oplus \mathcal{H}_{\mathrm{Q}}^{-+} \tag{4.1.3}$$

$$\mathcal{H}_{\mathrm{Q}/\!/\mathbb{Z}_2}^{-} = \mathcal{H}_{\mathrm{Q}}^{--} \oplus \mathcal{H}_{\mathrm{Q}}^{+-}, \tag{4.1.4}$$

Therefore the new theory has a global $\mathbb{Z}_2$. Similarly, orbifolding $\mathrm{Q}/\!/\mathbb{Z}_2$ by $\mathbb{Z}_2$, we get back

Q. We write:

$$\begin{tikzcd}
\mathbb{Z}_2 \rhd \mathrm{Q} \arrow[d, bend left, "/\!/\mathbb{Z}_2"] \\
\mathbb{Z}_2 \rhd \mathrm{Q}/\!/\mathbb{Z}_2 \arrow[u, bend left, "/\!/\mathbb{Z}_2"]
\end{tikzcd}.$$

This is equivalent to gauging the symmetry. At the level of the partition function we have the following statement. If $A \in \mathrm{H}^1(\Sigma_2, \mathbb{Z}_2)$ is a $\mathbb{Z}_2$-background gauge fieldand $B \in \mathrm{H}^1(\Sigma_2, \mathbb{Z}_2)$ a $\mathbb{Z}_2$-background gauge field, then

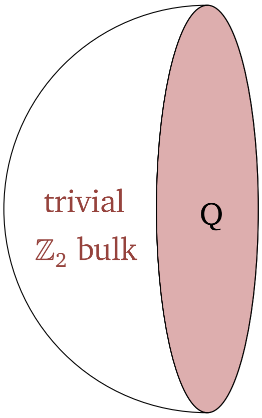
$$\mathcal{Z}_{\mathrm{Q}/\!/\mathbb{Z}_2}[\Sigma_2, B] \sim \sum_A \exp\left(\pi i \int_{\Sigma_2} A \smile B\right) \mathcal{Z}_{\mathrm{Q}}[\Sigma_2, A] \tag{4.1.5}$$

$$\begin{aligned}
\mathcal{Z}_{\mathrm{Q}/\!/\mathbb{Z}_2/\!/\mathbb{Z}_2}[\Sigma_2, A] &\sim \sum_B \exp\left(\pi i \int_{\Sigma_2} B \smile A\right) \mathcal{Z}_{\mathrm{Q}/\!/\mathbb{Z}_2}[\Sigma_2, B] \\
&\sim \sum_B \exp\left(\pi i \int_{\Sigma_2} B \smile A\right) \sum_{A'} \exp\left(\pi i \int_{\Sigma_2} A' \smile B\right) \mathcal{Z}_{\mathrm{Q}}[\Sigma_2, A'] \\
&= \mathcal{Z}_{\mathrm{Q}}[\Sigma_2, A].
\end{aligned} \tag{4.1.6}$$

Yet another way to see it is to attach the theory to a classical 3d bulk. $\mathbb{Z}_2$ is not anomalous, so we do not *have to* attach it to a bulk to define it properly. We *can*, however, attach it to a bulk; the trivial bulk, i.e. the trivial SPT phase. We then proceed to gauge the $\mathbb{Z}_2$ symmetry of the full system. Notice that since the ungauged bulk theory is a completely trivial theory, by gauging it, we must get at most a topological theory. In fact the gauged theory is equivalent to a $\mathrm{U}(1)^2$ Chern Simons theory with level matrix $K^{IJ} = \left(\begin{smallmatrix} 0 & 2 \\ 2 & 0 \end{smallmatrix}\right)$:

$$S[a] = \frac{1}{4\pi} K^{IJ} \int a_I \, \mathrm{d}a_J = \frac{1}{2\pi} \int a_1 \, \mathrm{d}a_2 \tag{4.1.7}$$

The Chern–Simons theory has $\mathbb{Z}_2$ Wilson lines for a_1 and a_2. For the rest of this section, we will not focus on the Chern–Simons theory, but we use the discrete gauge theory presentation to describe the bulk theory.



The 3d $\mathbb{Z}_2$ theory has a Wilson line $L_e(a)$ and a 't Hooft line $L_m(b)$, where a and b are cycles in the bulk manifold and e and m are the charges of the lines.These lines, since they're $\mathbb{Z}_2$ charged obey commutation relations as $L_e(a)L_m(b) = L_m(b)L_e(a)\ (-1)^{\int_{\Sigma_2'} a\smile b}$, where Σ_2' is some "timeslice" where both a and b have support on.

Similarly one could consider $Q/\!/\mathbb{Z}_2$ with a trivial $\mathbb{Z}_2$ bulk. Now the $\mathbb{Z}_2$ theory has Wilson and 't Hooft lines again with $e = m$ and $m = e$. Since the charged lines do not commute we can at most diagonalise one, so we can pick states $|e,A\rangle$ or $|m,B\rangle$ such that

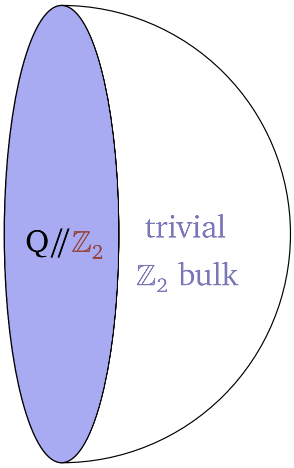
$$L_e(a)\,|e,A\rangle = \exp\left(\pi i \int a\smile A\right)|e,A\rangle \quad \textbf{or}$$
$$L_m(b)\,|m,B\rangle = \exp\left(\pi i \int b\smile B\right)|m,B\rangle\,.$$

Notice that

$$\langle e,A|m,B\rangle = \exp\left(\pi i \int A\smile B\right) \tag{4.1.8}$$

If we write that $\mathcal{Z}_{Q}[A] = \langle e,A|0\rangle$ and similarly $\mathcal{Z}_{Q/\!/\mathbb{Z}_2} = \langle 0|m,B\rangle$, we have that

$$\begin{aligned}\mathcal{Z}_{Q/\!/\mathbb{Z}_2}[B] &= \langle 0|m,B\rangle = \langle 0|\sum_A |e,A\rangle\,\langle e,A|m,B\rangle = \\ &= \sum_A \langle e,A|m,B\rangle\,\langle 0|e,A\rangle = \sum_A \langle e,A|m,B\rangle\,\langle e,A|0\rangle \\ &= \sum_A \exp\left(\pi i \int A\smile B\right)\mathcal{Z}_{Q}[A],\end{aligned} \tag{4.1.9}$$



thus recovering again, that $Q/\!/\mathbb{Z}_2$ is indeed the $\mathbb{Z}_2$-gauged Q.

We could have instead done the following

$$\mathcal{Z}_{Q/\!/\mathbb{Z}_2}[B] = \langle 0|m,B\rangle = \langle 0|m,B\rangle \sum_{A'} \big\langle e,A\big|e,A'\big\rangle = \langle 0|\sum_{A'} |m,B\rangle\langle e,A|\,\big|e,A'\big\rangle, \tag{4.1.10}$$

seeing that gauging amounts to inserting a doubly open bulk which accommodates a pure $\mathbb{Z}_2$ gauge theory with $\mathbb{Z}_2$-Dirichlet boundary conditions from the one side and $\mathbb{Z}_2$-Dirichlet from the other side and glueing $|e,A'\rangle$ to that, summing over the connections.

4.1.2 General finite group

The situation is similar for an arbitrary abelian finite group, G:

$$\begin{tikzcd} G \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/G"] \\ \hat{G} \triangleright \mathrm{Q}/\!/G \arrow[u, bend left=60, "/\!/\hat{G}"] \end{tikzcd}$$

Here $\hat{G}$ is the Pontryagin dual group, related to G as follows. One can define the set of unitary irreducible representations of G

$$\hat{G} := \left\{\chi : G \longrightarrow \mathrm{U}(1) \mid \chi \text{ is an irreducible representation}\right\} \equiv \mathrm{Hom}(G, \mathrm{U}(1)). \qquad (4.1.11)$$

Since G is abelian, all the irreducible representations are automatically one-dimensional, which makes the product of two irreducible representations a new irreducible representation, making $\hat{G}$ into a group $\hat{G}$, isomorphic to G, called the Pontryagin dual group.

At the partition function level we have:

$$\mathcal{Z}_{\mathrm{Q}/\!/G}[B] = \sum_{A} \mathrm{e}^{i\langle A,B\rangle} \mathcal{Z}_{\mathrm{Q}}[A], \qquad (4.1.12)$$

where $A \in \mathrm{H}^1(\Sigma_2, G)$, $B \in \mathrm{H}^1\left(\Sigma_2, \hat{G}\right)$ and

$$\mathrm{e}^{i\langle \bullet,\circ\rangle} : \mathrm{H}^1(\Sigma_2, G) \times \mathrm{H}^1\left(\Sigma_2, \hat{G}\right) \longrightarrow \mathrm{H}^2(\Sigma_2, \mathrm{U}(1)) \cong \mathrm{U}(1). \qquad (4.1.13)$$

This implies that

$$\begin{aligned} \mathcal{Z}_{\mathrm{Q}/\!/G/\!/\hat{G}}[A] &= \sum_{B}\sum_{A'} \mathrm{e}^{i\langle B,A\rangle} \mathrm{e}^{i\left\langle A',B\right\rangle} \mathcal{Z}_{\mathrm{Q}}[A'] = \\ &= \sum_{B}\sum_{A'} \mathrm{e}^{i\left\langle B,A-A'\right\rangle} \mathcal{Z}_{\mathrm{Q}}[A'] = \sum_{A'} \delta_{A,A'} \mathcal{Z}_{\mathrm{Q}}[A'] = \\ &= \mathcal{Z}_{\mathrm{Q}}[A]. \end{aligned} \qquad (4.1.14)$$

4.2 General Dimensions, higher-form symmetries

To generalise the above argument, let us consider a d-dimensional QFT, Q, on a manifold X_d, with a finite non-anomalous p-form global symmetry $G^{[p]}$ (note that higher-form symmetries are automatically abelian). Turn on a background $(p+1)$-form gauge field $\left[A_{[p+1]}\right] \in \mathrm{H}^{p+1}\left(X_d, G^{[p]}\right)$. Since the partition function is gauge invariant it can be seen as a proper function:

$$\begin{aligned} \mathcal{Z}_{\mathrm{Q}} : \mathrm{H}^{p+1}\left(X_d, G^{[p]}\right) &\longrightarrow \mathbb{C} \\ \left[A_{[p+1]}\right] &\longmapsto \mathcal{Z}_{\mathrm{Q}}\left[A_{[p+1]}\right]. \end{aligned}$$

Gauging is again orbifolding, i.e. summing over the twisted sectors, or equivalently over (equivalence classes of) flat $A_{[p+1]}$, with some weight. Assigning equal weight to all gauge fields we have:

$$\mathcal{Z}_{\mathrm{Q}/\!/G^{[p]}}[B] = \sum_{A_{[p+1]}} \mathrm{e}^{i\left\langle A_{[p+1]}, B\right\rangle} \mathcal{Z}_{\mathrm{Q}}\left[A_{[p+1]}\right],$$

However now the pairing $\langle \bullet, \circ \rangle$ should be of the form

$$\left\langle A_{[p+1]}, B\right\rangle \sim \int_{X_d} A_{[p+1]} \smile B. \tag{4.2.1}$$

which means that B should be a $(d-p-1)$-form field and hence a background field for a $(d-p-2)$-form symmetry, $\hat{G}^{[d-p-2]} = \mathrm{Hom}\left(G^{[p]}, \mathrm{U}(1)\right)$, giving finally

$$\begin{tikzcd} G^{[p]} \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/G^{[p]}"] \\ \hat{G}^{[d-p-2]} \triangleright \mathrm{Q}/\!/G^{[p]} \arrow[u, bend left=60, "/\!/\hat{G}^{[d-p-2]}"] \end{tikzcd}$$

with

$$\mathcal{Z}_{\mathrm{Q}/\!/G^{[p]}}[B] = \sum_{A_{[p+1]}} \mathrm{e}^{i\left\langle A_{[p+1]}, B_{[d-p-1]}\right\rangle} \mathcal{Z}_{\mathrm{Q}}\left[A_{[p+1]}\right] \tag{4.2.2}$$

$$\mathcal{Z}_{\mathrm{Q}/\!/G^{[p]}/\!/\hat{G}^{[d-p-2]}}\left[A_{[p+1]}\right] = \mathcal{Z}_{\mathrm{Q}}\left[A_{[p+1]}\right]. \tag{4.2.3}$$

4.3 Don't gauge the whole thing

There are two main cases where we are cannot gauge the whole symmetry group. If the symmetry group is anomalous, we have seen (cf. chapter 3) that the partition function is not background gauge invariant and therefore we are not allowed to make the gauge fields dynamical, i.e. fully gauge the symmetry. Another such case arises when the symmetry group is not abelian. In that case, although we are technically allowed to gauge it, we do not know what the fate of the charged objects is and how to regauge the gauged theory. This is because non-abelian symmetries, can have higher-dimensional irreducible representations, therefore multiplying two irreducible representations is not necessarily an irreducible representation and hence the set $\mathrm{Rep}(G) = \mathrm{Hom}(G, \mathrm{U}(1))$ no longer combines into a group. There is a remedy to that, but it will come in chapter 6.

In any case, we may be able to find a finite abelian non-anomalous subgroup of the whole symmetry group and gauge that. This procedure raises some questions. What is the global symmetry of the partially gauged theory? Can the gauging be undone? These questions were partially answered in [43] and will be reviewed below, together with attempts towards a fuller understanding.

4.3.1 Without an anomaly

Consider a d-dimensional QFT, Q, on X_d, with a (zero-form) symmetry group Γ and assume there is a finite, abelian, non-anomalous subgroup $\mathcal{A}$. Suppose that Γ is not anomalous, and $\mathcal{A}$ enters Γ through a direct product,

$$\Gamma = \mathcal{A} \times \Gamma/\mathcal{A} =: \mathcal{A} \times G. \tag{4.3.1}$$

To gauge $\mathcal{A}$ we proceed exactly as in the previous section. Essentially one can think of this case as being the previous case with an extra symmetry that comes along for the ride. We have then

$$\begin{tikzcd} \mathcal{A} \times G \,\triangleright\, \mathrm{Q} \arrow[d, bend left=60, "/\!/\mathcal{A}"] \\ \hat{\mathcal{A}}^{[d-p-2]} \times G \,\triangleright\, \mathrm{Q}/\!/\mathcal{A} \arrow[u, bend left=60, "/\!/\hat{\mathcal{A}}^{[d-2]}"] \end{tikzcd} .$$

This case was rather uninteresting, so consider a more interesting case, when $\mathcal{A}$ enters Γ through a nontrivial group extension. In general if $\mathcal{A}$ enters Γ through a short exact sequence,

$$1 \longrightarrow \mathcal{A} \longrightarrow \Gamma \longrightarrow G \longrightarrow 1, \tag{4.3.2}$$

where $\mathcal{A}$ is a normal subgroup of Γ and $G = \Gamma/A$. Γ is said to be an extension of G by $\mathcal{A}$ and the possible extensions are classified by classes of the second group cohomology

$\varepsilon \in \mathrm{H}^2(G, \mathcal{A})$. The extension is defined as follows: $\Gamma = \mathcal{A} \times_\varepsilon G$ is identified with $\mathcal{A} \times G$ as a set, but the group operation is defined as

$$(0, g) * (0, h) := (\varepsilon(g, h), gh).$$

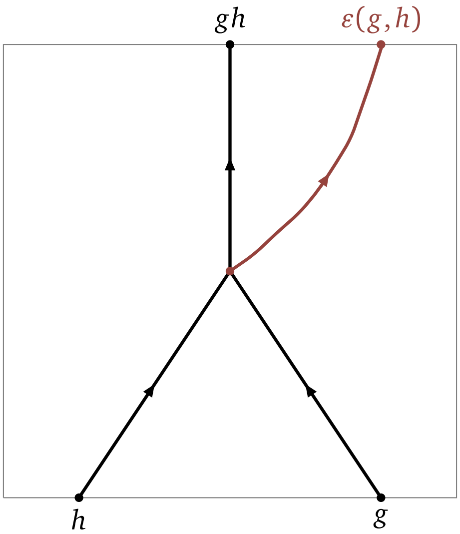


Figure 4.1: Domain walls for $\mathcal{A}$ created by the fusion of G-domain walls.

In terms of background gauge fields, if $\varepsilon = 0$, a background gauge field for $\mathcal{A}$, is $A \in \mathrm{H}^1(X_d, \mathcal{A})$. Now if $\varepsilon \neq 0$, group elements of $\mathcal{A}$ can be created solely form group elements of G, i.e. domain walls for $\mathcal{A}$ can have boundaries, as seen in figure 4.1. This implies that A need not be a cocycle but instead it must compensate the contribution of the boundary. If then B is a background gauge field for G, $A \in C^1(X_d, \mathcal{A})$ must satisfy

$$dA = \varepsilon(B) \in \mathrm{H}^2(X_d, \mathcal{A}). \tag{4.3.3}$$

This is precisely the Green–Schwarz effect of string theory [44]. Here $\varepsilon(B)$ is defined as follows. B determines a map $f_B : X_d \longrightarrow \mathrm{B}G$, which we can use to pull ε back to X_d, so $\varepsilon(B) = {f_B}^* \varepsilon \in \mathrm{H}^2(X_d, \mathcal{A})$.

By gauging $\mathcal{A}$ the gauged theory will certainly have $\hat{\mathcal{A}}^{[d-2]}$ as a symmetry. However, when passing a $\hat{\mathcal{A}}^{[d-2]}$ defect, i.e. a topological line operator, $\mathcal{W}$, across a codimension-2 region where the domain wall $\varepsilon(g, h) \in \mathcal{A}$ starts, the line picks up an extra phase $\mathcal{W}(\varepsilon(g, h))$. When gauging $\mathcal{A}$, such domain walls are summed over and hence invisible, therefore we pick up this extra phase when passing through the intersection of three G-domain walls. This implies that the $\mathcal{A}$-gauged theory has a $\hat{\mathcal{A}}^{[d-2]} \times G$ global symmetry with a mixed anomaly, given by the SPT phase:

$$\mathcal{Z}_{\mathrm{SPT}}\left[A_{[d-1]}, B\right] = \exp\left(i \int_{Y_{d+1}} A_{[d-1]} \smile \varepsilon(B) \right), \tag{4.3.4}$$

with $\partial Y_{d+1} = X_d$, where both $A_{[d-1]}$ and B are extended into the bulk of the SPT phase — and hence we now use $f_B : Y_{d+1} \longrightarrow \mathrm{B}G$ to define $\varepsilon(B)$.

Thus we have

$$\begin{tikzcd}
\left[\begin{array}{c} 1 \longrightarrow \mathcal{A} \longrightarrow \Gamma \longrightarrow G \longrightarrow 1 \\ \varepsilon \in \mathrm{H}^2(G,\mathcal{A}) \end{array}\right] \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/\mathcal{A}"] \\
\left[\mathrm{SPT}\big[A_{[d-1]}, B\big] = \overset{\widehat{\mathcal{A}}^{[d-2]} \times G}{\displaystyle\int_{Y_{d+1}} A_{[d-1]} \smile \varepsilon(B)}\right] \triangleright \mathrm{Q}/\!/\mathcal{A} \arrow[u, bend left=60, dashed, "/\!/\widehat{\mathcal{A}}^{[d-2]}"]
\end{tikzcd}.$$

The question is now whether we can follow the dotted arrow back, gauging $\widehat{\mathcal{A}}^{[d-2]}$. We turn to this question in the subsection below.

4.3.2 With an anomaly

Consider first the case when the extension class is zero, i.e. that our theory, Q, has the symmetry $\Gamma = \mathcal{A} \times G$, with an anomaly

$$\mathcal{Z}_{\mathrm{SPT}}[A, B] = \exp\left(i \int_{Y_{d+1}} A \smile \widetilde{\varepsilon}(B) \right), \tag{4.3.5}$$

with $A \in \mathrm{H}^1(Y_{d+1}, \mathcal{A})$, $B \in \mathrm{H}^d(Y_{d+1}, G)$, and $\widetilde{\varepsilon} \in \mathrm{H}^d(G, \widehat{\mathcal{A}})$, where $\partial Y_{d+1} = X_d$.

One can gauge $\mathcal{A}$, after turning the background G-gauge field off. However, upon gauging $\mathcal{A}$, we see that $\widetilde{\varepsilon}$ is still there, mixing $\widehat{\mathcal{A}}^{[d-2]}$ and G non-trivially, exactly inverting the logic of the previous case. Therefore the gauged theory has a symmetry given by an extension of G by $\widehat{\mathcal{A}}^{[d-2]}$, controlled by $\widetilde{\varepsilon} \in \mathrm{H}^d(G, \widehat{\mathcal{A}}^{[d-2]})$.

$$1 \longrightarrow \widehat{\mathcal{A}}^{[d-2]} \longrightarrow \Gamma^{[d-2,0]} \longrightarrow G \longrightarrow 1. \tag{4.3.6}$$

Here, $\Gamma^{[d-2,0]}$ is a mixture of a zero-form symmetry and a $(d-2)$-form symmetry, in the form of a $(d-1)$-group. In particular to turn on background gauge fields for this symmetry we need

$$B \in \mathrm{H}^1(X_d, G) \quad \& \quad \widehat{A}_{[d-1]} \in C^{d-1}\left(X_d, \widehat{\mathcal{A}}^{[d-2]}\right), \text{ such that } \quad \mathrm{d}\widehat{A}_{[d-1]} = \widetilde{\varepsilon}(B). \tag{4.3.7}$$

In this construction we have then that:

$$\begin{tikzcd}
\left[\begin{array}{c} \mathcal{A}\times G \\ \mathrm{SPT}[A,B]=\displaystyle\int_{Y_{d+1}} A\smile\tilde{\varepsilon}(B) \end{array}\right] \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/\mathcal{A}"] \\
\left[\begin{array}{c} 1\longrightarrow \hat{\mathcal{A}}^{[d-2]} \longrightarrow \Gamma^{[d-2,0]} \longrightarrow G \longrightarrow 1 \\ \tilde{\varepsilon}\in \mathrm{H}^d\big(G,\hat{\mathcal{A}}^{[d-2]}\big) \end{array}\right] \triangleright \mathrm{Q}/\!/\mathcal{A} \arrow[u, bend left=60, dashed, "/\!/\hat{\mathcal{A}}^{[d-2]}"]
\end{tikzcd}$$

We still have not answered the question of whether we can follow the dashed arrow in the previous construction, but now we have a new dashed arrow which we do not know if we can follow. Eventually the answer will come in subsection 4.4.

However before going to the general case for generic higher-form symmetries, let us consider the case where everything is turned on. Γ is a non trivial extension of G by $\mathcal{A}$,

$$\begin{aligned} 1\longrightarrow \mathcal{A}\longrightarrow \Gamma\longrightarrow G\longrightarrow 1, &\quad \text{controlled by} \quad \varepsilon\in \mathrm{H}^2(G,\mathcal{A}), \\ \text{with anomaly} \quad \mathrm{SPT}\in \mathrm{H}^{d+1}(\Gamma,\mathrm{U}(1)), &\quad \text{controlled by} \quad \tilde{\varepsilon}\in \mathrm{H}^d(G,\hat{\mathcal{A}}). \end{aligned}$$

Again as above we have the Green–Schwarz condition $\mathrm{d}A=\varepsilon(B)$. However now if we take $\tilde{\varepsilon}\in \mathrm{H}^d\big(G,\mathrm{H}^1(\mathcal{A},\mathrm{U}(1))\big)=\mathrm{H}^d\big(G,\hat{\mathcal{A}}\big)$ we cannot use $A\smile\tilde{\varepsilon}(B)$ to define an element of $\mathrm{H}^{d+1}\big(\Gamma,\mathrm{U}(1)\big)$, since

$$\begin{aligned} \mathrm{d}(A\smile\tilde{\varepsilon}(B)) &= \mathrm{d}(A\smile {f_B}^*\tilde{\varepsilon}) = \mathrm{d}A\smile {f_B}^*\tilde{\varepsilon} - A\smile \mathrm{d}({f_B}^*\tilde{\varepsilon}) = \\ &= {f_B}^*\varepsilon\smile {f_B}^*\tilde{\varepsilon} - A\smile \cancelto{0}{{f_B}^*(\mathrm{d}\tilde{\varepsilon})} = {f_B}^*(\varepsilon\smile\tilde{\varepsilon}). \end{aligned} \tag{4.3.8}$$

To make $\mathrm{d}(A\smile\tilde{\varepsilon}(B))=0$ for all B, thus defining a proper cohomology class element, we need to require that

$$\varepsilon\smile\tilde{\varepsilon}=0\in \mathrm{H}^{d+2}(G,\mathrm{U}(1)). \tag{4.3.9}$$

For that to hold it suffices to take $\varepsilon\smile\tilde{\varepsilon}$ to be a boundary, i.e. pick an $\omega\in C^{d+1}(G,\mathrm{U}(1))$, such that

$$\mathrm{d}\omega=\varepsilon\smile\tilde{\varepsilon}\in C^{d+1}(G,\mathrm{U}(1)), \tag{4.3.10}$$

corresponding indeed to the trivial element in the cohomology. Now, we have in our disposal an element of $\mathrm{H}^{d+1}(\Gamma,U(1))$, namely

$$\mathrm{SPT}(\tilde{\varepsilon},\omega):=\int_{Y_{d+1}}\Big(A\smile\tilde{\varepsilon}(B)-\omega(B)\Big). \tag{4.3.11}$$

Now, proceeding to gauge $\mathcal{A}$, we obtain exactly like previously a theory whose symmetry data are

$$
\begin{aligned}
&1 \longrightarrow \hat{\mathcal{A}}^{[d-2]} \longrightarrow \widetilde{\Gamma}^{[d-2,0]} \longrightarrow G \longrightarrow 1 \\
&\text{controlled by} \quad \widetilde{\varepsilon} \in \mathrm{H}^d(G, \hat{\mathcal{A}}), \\
\text{with anomaly} \quad &\widetilde{\mathrm{SPT}} \in \mathrm{H}^{d+1}(\widetilde{\Gamma}, \mathrm{U}(1)) \quad \text{controlled by} \quad \varepsilon \in \mathrm{H}^2(G, \mathcal{A}),
\end{aligned}
$$

where

$$
\widetilde{\mathrm{SPT}}(\varepsilon, \omega) := \int_{Y_{d+1}} \left(\hat{A} \smile \varepsilon(B) - \omega(B)\right). \tag{4.3.12}
$$

Said differently we have the following gauging-ungauging pattern

$$
\begin{tikzcd}
\left[\begin{array}{c} 1 \longrightarrow \mathcal{A} \longrightarrow \Gamma \longrightarrow G \longrightarrow 1 \\ \varepsilon \in \mathrm{H}^2(G, \mathcal{A}) \\ \mathrm{SPT}(\widetilde{\varepsilon}, \omega) = \displaystyle\int_{Y_{d+1}} \left(A \smile \widetilde{\varepsilon}(B) - \omega(B)\right) \end{array}\right] \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/\mathcal{A}"] \\
\left[\begin{array}{c} 1 \longrightarrow \hat{\mathcal{A}}^{[d-2]} \longrightarrow \widetilde{\Gamma}^{[d-2,0]} \longrightarrow G \longrightarrow 1 \\ \widetilde{\varepsilon} \in \mathrm{H}^d\left(G, \hat{\mathcal{A}}^{[d-2]}\right) \\ \widetilde{\mathrm{SPT}}(\varepsilon, \omega) = \displaystyle\int_{Y_{d+1}} \left(\hat{A} \smile \varepsilon(B) - \omega(B)\right) \end{array}\right] \triangleright \mathrm{Q}/\!/\mathcal{A} \arrow[u, bend left=60, dashed, "/\!/\hat{\mathcal{A}}^{[d-2]}"]
\end{tikzcd}
$$

Now we have a third dashed arrow which we do not know if we can trust, so it is urgent to address this issue.

4.4 Higher-form subgroup gaugings

The generalisation of the classifying space $\mathrm{B}G$, out of which we make our cohomology groups is Eilenberg–Mac Lane spaces,

$$
\mathrm{B}^{p+1}G := K(G, p+1). \tag{4.4.1}
$$

In this subsection, we will mostly repeat the above arguments for higher-form cases. Consider elements $\varepsilon \in \mathrm{H}^{q+2}\left(\mathrm{B}^{p+1}G, \mathcal{A}\right)$ and $\widetilde{\varepsilon} \in \mathrm{H}^{d-q}\left(\mathrm{B}^{p+1}G, \hat{\mathcal{A}}\right)$, such that

$$
\varepsilon \smile \widetilde{\varepsilon} = 0 \in \mathrm{H}^{d+2}\left(\mathrm{B}^{p+1}G, \mathrm{U}(1)\right). \tag{4.4.2}
$$

Choose a cochain $\omega \in C^{d+1}(\mathrm{B}^{p+1}G, \mathrm{U}(1))$ to trivialise the above cup, via

$$\mathrm{d}\omega = \varepsilon \smile \widetilde{\varepsilon}. \tag{4.4.3}$$

We will construct an extension of the p-form group $G^{[p]}$ by the q-form group $\mathcal{A}^{[q]}$.

$$1 \longrightarrow \mathcal{A}^{[q]} \longrightarrow \Gamma^{[q,p]} \longrightarrow G^{[p]} \longrightarrow 1, \tag{4.4.4}$$

with the extension specified by ε. $\Gamma^{[q,p]}$ is defined through its background fields, by a pair $\left(A_{[q+1]}, B_{[p+1]}\right)$, where $A_{[q+1]} \in C^{q+1}(X_d, \mathcal{A})$ and $B_{[p+1]} \in \mathrm{H}^{p+1}(X_d, G)$, satisfying

$$\mathrm{d}A_{[q+1]} = \varepsilon\left(B_{[p+1]}\right). \tag{4.4.5}$$

Now $\varepsilon\left(B_{[p+1]}\right)$ is not striclty a pullback, but rather a more abstract cohomology operation that preserves the identities of a pullback, namely compatibility with the exterior derivate and the cup product

$$\begin{aligned} \varepsilon : \mathrm{H}^{p+1}(\bullet, G) &\longrightarrow \mathrm{H}^{q+2}(\bullet, \mathcal{A}), \\ B_{[p+1]} &\longmapsto \varepsilon\left(B_{[p+1]}\right). \end{aligned}$$

$\Gamma^{[q,p]}$ can be anomalous with the anomaly given by

$$\mathrm{SPT}(\widetilde{\varepsilon}, \omega) := \int_{Y_{d+1}} \left(A_{[q+1]} \smile \widetilde{\varepsilon}\left(B_{[p+1]}\right) - \omega\left(B_{[p+1]}\right)\right), \tag{4.4.6}$$

where $\widetilde{\varepsilon}$ is again a cohomology operation

$$\begin{aligned} \widetilde{\varepsilon} : \mathrm{H}^{p+1}(\bullet, G) &\longrightarrow \mathrm{H}^{d-q}(\bullet, \hat{\mathcal{A}}), \\ B_{[p+1]} &\longmapsto \varepsilon\left(B_{[p+1]}\right). \end{aligned}$$

Conversely we could imagine extending $G^{[p]}$ by $\hat{\mathcal{A}}^{[d-q-2]}$, as

$$1 \longrightarrow \hat{\mathcal{A}}^{[d-q-2]} \longrightarrow \widetilde{\Gamma}^{[d-q-2,p]} \longrightarrow G^{[p]} \longrightarrow 1, \tag{4.4.7}$$

with the extension controlled by $\widetilde{\varepsilon}$ and an anomaly

$$\widetilde{\mathrm{SPT}}(\varepsilon, \omega) = \int_{Y_{d+1}} \left(\hat{A}_{[d-q-1]} \smile \varepsilon\left(B_{[p+1]}\right) - \omega\left(B_{[p+1]}\right)\right), \tag{4.4.8}$$

where we have here the 'dual' Green–Schwarz relation $\mathrm{d}\hat{A}_{[d-q-1]} = \widetilde{\varepsilon}\left(B_{[p+1]}\right)$.

According to the gauging construction of subsection 4.3.1, the two extensions are exchanged by gauging, i.e. we have the gauging-ungauging pattern

$$
\begin{tikzcd}
\left[\begin{array}{c} 1 \longrightarrow \mathcal{A}^{[q]} \longrightarrow \Gamma^{[q,p]} \longrightarrow G^{[p]} \longrightarrow 1 \\ \varepsilon \in \mathrm{H}^{q+2}(G,\mathcal{A}) \\ \mathrm{SPT}(\widetilde{\varepsilon},\omega) = \displaystyle\int_{Y_{d+1}} \Big(A_{[q+1]} \smile \widetilde{\varepsilon}\big(B_{[p+1]}\big) - \omega\big(B_{[p+1]}\big)\Big) \end{array}\right] \triangleright \mathrm{Q} \arrow[d, bend left=60, "/\!/\mathcal{A}"] \\
\left[\begin{array}{c} 1 \longrightarrow \hat{\mathcal{A}}^{[d-q-2]} \longrightarrow \widetilde{\Gamma}^{[d-q-2,p]} \longrightarrow G^{[p]} \longrightarrow 1 \\ \widetilde{\varepsilon} \in \mathrm{H}^{d-q}\Big(G,\hat{\mathcal{A}}^{[d-q-2]}\Big) \\ \widetilde{\mathrm{SPT}}(\varepsilon,\omega) = \displaystyle\int_{Y_{d+1}} \Big(\hat{\mathcal{A}}^{[d-q-1]} \smile \varepsilon\big(B_{[p+1]}\big) - \omega\big(B_{[p+1]}\big)\Big) \end{array}\right] \triangleright \mathrm{Q}/\!/\mathcal{A} \arrow[u, bend left=60, "/\!/\hat{\mathcal{A}}"]
\end{tikzcd}
\tag{4.4.9}
$$

From here we see that all the dashed arrows above are very well justified, by taking the $q = p = 0$ case, so we might replace them with normal arrows.

4.4.1 Higher-group symmetries

In this subsection we will elaborate a bit on the structure of the symmetry

$$
1 \longrightarrow \mathcal{A}^{[q]} \longrightarrow \Gamma^{[q,p]} \longrightarrow G^{[p]} \longrightarrow 1, \qquad \varepsilon \in \mathrm{H}^{q+2}(\mathrm{B}^{p+1}G,\mathcal{A}).
$$

for the $p = 0$ case. The structure is known to mathematicians and lately to physicists as well; $\Gamma^{[q,0]}$ is a $(q+1)$-group.

$$
1 \longrightarrow \mathcal{A}^{[q]} \longrightarrow \Gamma^{[q,0]} \longrightarrow G \longrightarrow 1, \qquad \varepsilon \in \mathrm{H}^{q+2}(\mathrm{B}G,\mathcal{A}).
$$

It can be described as follows. The topological operators for G are codimension-1 defects. Generically codimension-1 defects will intersect. When $(q+2)$ such G-defects intersect they give rise to a codimension-$(q+1)$ defect $\varepsilon(g_1, g_2, \ldots, g_{q+2})$ which acts as a sourced topological operator for $\mathcal{A}$. The case $q = 0$ was the starting point of this quest, while when $q = 1$ it gives rise to the physics of 2-groups, studied by [6, 7]. In the case of 2-groups, ε is known as a Postnikov class and is usually denoted by β. Higher-groups arise naturally in category theory, where a $(q+1)$ group is a group-like $(q+1)$-category, where q-associativity is weakened down to isomorphism. Those $q > 2$ cases are straightforward to study, in the physics language, by modifying [6,7]'s arguments, albeit not well understood in mathematics, for problems linked with the understanding of higher categories.

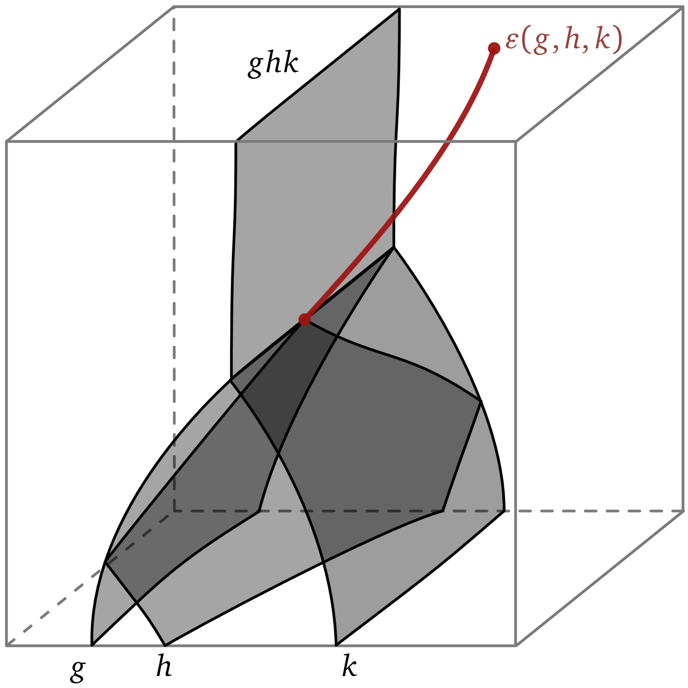


Figure 4.2: Three codimension-1 G-defects create a codimension-2 $\mathcal{A}^{[1]}$ defect, forming a 2-group $\Gamma^{[1,0]}$-defect

For an even more general setup, we must consider the case where G acts on $\mathcal{A}$, which is easy to imagine, as a codimension-$(q+1)$ $\mathcal{A}$-defect, a can pierce through a codimension-1 G-defect, g and transform into $a' := \rho_g a$. This defines a group homomorphism:

$$\rho : G \longrightarrow \mathrm{Aut}\left(\mathcal{A}^{[q]}\right). \tag{4.4.10}$$

In total a $(q+1)$-group symmetry can be defined as

$$\Gamma^{[q,0]} := (G, \mathcal{A}^{[q]}, \rho, \varepsilon), \tag{4.4.11}$$

where G is a group, $\mathcal{A}^{[q]}$ is an abelian q-form group, ρ is the group homomorphism defined above and $\varepsilon \in \mathrm{H}^{q+2}_{\rho}(\mathrm{B}G, \mathcal{A})$, H^{q+2}_{ρ} being the ρ-twisted (topological) cohomology. It can be equivalently defined via its background gauge fields, $B \in \mathrm{H}^1(X_d, G)$, $A_{[q+1]} \in C^{q+1}(X_d, \mathcal{A})$ that obey the following relation:

$$\mathrm{d}_{\rho} A_{[q+1]} = \varepsilon(B). \tag{4.4.12}$$

For simplicity let us focus again on the case where ρ is trivial, i.e. there is no action of G on $\mathcal{A}^{[q]}$. If the $(q+1)$-group is continuous, namely $\mathcal{A}^{[1]} \cong G \cong \mathrm{U}(1)$, we can turn on U(1) background gauge fields, $A_{[q+1]}$ and $B_{[1]}$, with transforming as

$$B_{[1]} \longmapsto B^{\lambda}_{[1]} := B_{[1]} + \mathrm{d}\lambda \tag{4.4.13}$$

$$A_{[q+1]} \longmapsto A^{\Lambda,\lambda}_{[q+1]} := A_{[q+1]} + \mathrm{d}\Lambda_{[q]} - \lambda\zeta\big(B_{[1]},\lambda\big), \tag{4.4.14}$$

where $\zeta\big(B_{[1]},\lambda\big)$ is determined by the equation,

$$\mathrm{d}\big(\lambda\zeta\big(B_{[1]},\lambda\big)\big) \stackrel{!}{=} \varepsilon\Big(B^{\lambda}_{[1]}\Big) - \varepsilon\big(B_{[1]}\big). = \varepsilon(\mathrm{d}\lambda). \tag{4.4.15}$$

Notice that (4.4.15) has always solutions because a gauge transformation of of $B_{[1]}$ corresponds to a homotopy of $A : X_d \longrightarrow \mathrm{B}G$ which does not change the cohomology class of $\varepsilon\big(B_{[1]}\big)$, and hence the RHS of (4.4.15) vanishes in the cohomology. Essentially, ζ is determined by descent from the pullback.

Then the theory has conserved currents, which we can find by adding to the Lagrangian the coupling terms

$$\int A_{[q+1]} \wedge \star j_{[q+1]} + B_{[1]} \wedge \star j_{[1]}. \tag{4.4.16}$$

We then find

$$\mathrm{d} \star j_{[q+1]} = 0 \tag{4.4.17}$$

$$\mathrm{d} \star j_{[1]} = \zeta\big(B_{[1]},\lambda\big) \wedge \star j_{[q+1]}, \tag{4.4.18}$$

which may look anomalous upon a first glance, but in reality they are not. It is merely a consequence of the Green–Schwarz mechanism controlling the background gauge fields. In some cases we might write a useful form of ζ in terms of its arguments and the Postnikov class. Following a result from [22],

$$\mathrm{H}^{q+1}\big(\mathrm{BU}(1),\mathrm{U}(1)\big) = \begin{cases} \mathrm{U}(1) & q = -1, \\ 1 & q \in 2\mathbb{N}+1, \\ \mathbb{Z} & q \in 2\mathbb{N}, \end{cases} \tag{4.4.19}$$

we find that for example in the case of a 2-group it is $\zeta\big(B_{[1]},\lambda\big) = \frac{\kappa}{2\pi}\lambda \mathrm{d}B_{[0]}$, where $\kappa \in \mathbb{Z}$ are labeling the cohomology class $\mathrm{H}^3(\mathrm{BU}(1),\mathrm{U}(1)) = \mathbb{Z}$.

Higher-groups can have anomalies but with a slight modification: pure G-anomalies are not counted by $\mathrm{H}^{d+1}(\mathrm{B}G,\mathrm{U}(1))$, but rather by

$$\frac{\mathrm{H}^{d+1}(\mathrm{B}G,\mathrm{U}(1))}{\mathrm{H}^{d-q-2}_{\rho}\Big(\mathrm{B}G,\hat{\mathcal{A}}^{[d-q-2]}\Big) \smile \varepsilon}. \tag{4.4.20}$$

4.5 More general anomaly?

So far we achieved gauging and ungauging of $\mathcal{A}$ in $1 \longrightarrow \mathcal{A} \longrightarrow \Gamma \longrightarrow G \longrightarrow 1$, when the anomaly of Γ has an anomaly with a very particular form, determined by $\widetilde{\varepsilon} \in \mathrm{H}^d\big(G, \hat{\mathcal{A}}\big)$. However, a general anomaly of Γ is determined by classes of $\mathrm{H}^{d+1}\big(\Gamma, \mathrm{U}(1)\big)$. Such a group cohomology can be studied using $\mathrm{H}^p\big(G, \mathrm{H}^q(\mathcal{A}, \mathrm{U}(1))\big)$ with $p + q = d + 1$ and using the Lyndon–Hochschild–Serre spectral sequence (LHSss) (B.2.5).

So far we have only used

$$\mathrm{H}^{d+1}\big(G, \mathrm{H}^0(\mathcal{A}, \mathrm{U}(1))\big) = \mathrm{H}^{d+1}(G, \mathrm{U}(1)) \quad \text{and} \tag{4.5.1}$$

$$\mathrm{H}^d\big(G, \mathrm{H}^1(\mathcal{A}, \mathrm{U}(1))\big) = \mathrm{H}^d\big(G, \mathrm{Hom}(\mathcal{A}, \mathrm{U}(1))\big) = \mathrm{H}^d\big(G, \hat{\mathcal{A}}\big). \tag{4.5.2}$$

A natural continuation would be to consider anomalies that collapse deeper in the LHSss, i.e. anomalies of the form

$$\mathrm{H}^{d+1-k}\big(G, \mathrm{H}^k(\mathcal{A}, \mathrm{U}(1))\big). \tag{4.5.3}$$

This is however rather cumbersome because, already in two dimensions we shall abandon the notion of a symmetry group and consider symmetry categories instead [10].

Consider for example the simplest such case, where $\Gamma = \mathcal{A} \times G$ and it has an anomaly determined by $\xi \in \mathrm{H}^{d-1}\big(G, \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1))\big)$, through

$$\mathrm{SPT}[A, B] = \int_{Y_{d+1}} A \smile A \smile \xi(B). \tag{4.5.4}$$

In two dimensions, to gauge $\mathcal{A}$ we must decorate the domain walls with projective representations to cancel the mixed flavour–gauge anomaly. Thus the topological operator of the gauged theory $\mathrm{Q}/\!/\mathcal{A}$ are

$$(\rho, g) \in \big(\mathrm{Rep}_{\xi(g)}\mathcal{A}, G\big), \tag{4.5.5}$$

where $\mathrm{Rep}_{\xi(g)}\mathcal{A}$ is the set of projective representations of $\mathcal{A}$ specified by $\xi(g) \in \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1))$. Given such operators we can fuse them via

$$(\rho, g) \otimes (\sigma, h) = (\rho \otimes \sigma, gh), \tag{4.5.6}$$

which is consistent since the phase of the projective representation $\rho \otimes \sigma$ is $\xi(gh) = \xi(g)\xi(h)$. This rule defines in turn a fusion category.

In $d > 2$, if we start with a theory with symmetry $\Gamma = G \times \mathcal{A}$ with an anomaly specified by $\xi \in \mathrm{H}^{d-1}\big(G, \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1))\big)$ $(d-1)$ G-domain walls fuse to make a new domain wall and determine an element $\xi(g_1, g_2, \ldots, g_{d-1}) \in \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1))$. Now we decorate this new element by a projective representation of $\mathcal{A}$ with phase given by ξ. In $d = 3$ this construction gives rise to non-invertible 2-categories. The general framework of categorical symmetries is developed in chapter 6.

5 Anomalies of dualities

In this chapter we will focus on anomalies of dualities. Under the light of [45] and [46,47] it is possible to recast (self-)dualities as (-1)-form symmetries and explore their anomalies. We will review the general formalism for detecting whether a duality group is anomalous and we will apply it to specific examples in QFT and in quantum mechanics. We will also see that the framework developed in the previous chapter will be useful for certain examples.

5.1 How to detect an anomaly of a duality group?

The general procedure to detect whether a (self-)duality group is anomalous goes as follows [45]. Let $\widehat{\mathcal{M}}$ be the space of background data — background gauge fields, background metrics, coupling constants, etc. — and G be the duality group we want to examine. The partition function is generally a function on $\widehat{\mathcal{M}}$. However, it is not always a function on $\mathcal{M} := \widehat{\mathcal{M}}/G$. It can be that G is anomalous, so the partition function is merely a section of a line bundle over $\mathcal{M}$.

Consider a complex line bundle over $\widehat{\mathcal{M}}$. For simplicity we will only focus on the trivial complex line bundle, $\mathbb{C} \times \widehat{\mathcal{M}}$, however all the discussion is generalisable to nontrivial bundles using local coordinate charts. Let $(\ell, m) \in \mathbb{C} \times \widehat{\mathcal{M}}$ be a coordinate system and $g \in G$, acting as

$$g : (\ell, m) \longmapsto (\varphi_m(g)\ell,\ g \cdot m). \tag{5.1.1}$$

For the action to be consistent we demand that $\varphi_m(g) \in \mathbb{C}^\times$, satisfies

$$\varphi_m(f \circ g) = \varphi_{g\cdot m}(f)\varphi_m(g). \tag{5.1.2}$$

Then we might construct the equivariant, up to fixed points, bundle

$$\mathcal{L} := \frac{\mathbb{C} \times \widehat{\mathcal{M}}}{G}$$

Then if G is anomalous, the partition function is a section of $\mathcal{L}$, determined by $\varphi_m(g)$:

$$\mathcal{Z}[g \cdot m] = \varphi_m(g)\mathcal{Z}[m]. \tag{5.1.3}$$

Therefore a way to test if a non-trivial transformation of a partition function yields indeed an anomaly is to act with the potentially anomalous group on the background data, pick up the transformation of the partition function after each group element and test whether they satisfy (5.1.2), thus constituting a section of the equivariant line bundle on $\widehat{\mathcal{M}}$.

5.2 Anomaly of the electric–magnetic duality

Consider Maxwell theory in four dimensions with a θ-angle turned on:

$$S[a] = \frac{1}{g^2}\int_{X_4} f \wedge \star f + \frac{i\theta}{8\pi^2}\int_{X_4} f \wedge f, \tag{5.2.1}$$

where a is a one-form gauge field and $f = \mathrm{d}a$. On spin manifolds (5.2.1) has the duality group $\mathrm{SL}(2,\mathbb{Z})$. In order to see this, let us combine the coupling constants into

$$\tau := \frac{4\pi i}{g^2} + \frac{\theta}{2\pi}. \tag{5.2.2}$$

Then the group $\mathrm{SL}(2,\mathbb{Z})$ acts on τ, regarded as a zero-form background field. In particular, the generators of $\mathrm{SL}(2,\mathbb{Z})$ act on τ as

$$\mathsf{T}:\tau \longmapsto \tau+1 \tag{5.2.3}$$

$$\mathsf{S}:\tau \longmapsto \frac{1}{\tau}. \tag{5.2.4}$$

We know due to Witten [48], that the partition function, on a closed, oriented, spin four-manifold

$$\mathcal{Z}[\tau,\overline{\tau}] := \int \mathrm{D}a$$

$$exp(-S[a;\tau,\overline{\tau}]),$$

transforms as follows:

$$\mathcal{Z}[\mathsf{T}\cdot(\tau,\overline{\tau})] = \mathcal{Z}[\tau+1,\overline{\tau}+1] = \mathcal{Z}[\tau,\overline{\tau}] \tag{5.2.5}$$

$$\mathcal{Z}[\mathsf{S}\cdot(\tau,\overline{\tau})] = \mathcal{Z}\left[-\frac{1}{\tau},-\frac{1}{\overline{\tau}}\right] = \tau^{(\chi+\sigma)/4}\overline{\tau}^{(\chi-\sigma)/4}\mathcal{Z}[\tau,\overline{\tau}], \tag{5.2.6}$$

where

$$\chi = \frac{1}{32\pi^2}\int_{X_4} \epsilon^{abcd}R_{ab}\wedge R_{cd} \quad \text{and} \tag{5.2.7}$$

$$\sigma = \frac{1}{24\pi^2}\int_{X_4} R^a{}_b\wedge R^b{}_a = \int_{X_4}\frac{p_1(X_4)}{3} \tag{5.2.8}$$

are the Euler characteristic and the signature of the four-manifold respectively and $R^a{}_b$ is the curvature two-form $R^a{}_b = \mathrm{d}\omega^a{}_b + \omega^a{}_c\wedge\omega^c{}_b$. As we will see in a bit we will interpret (5.2.5,5.2.6) as a mixed anomaly between the duality group $\mathrm{SL}(2,\mathbb{Z})$ and gravity.

Before doing that, though, we have the freedom to add local counterterms to simplify the discussion. Since only the signature and the Euler characteristic appear, the counterterm Lagrangian should have the form

$$L_{\mathrm{ct}} = \frac{1}{32\pi^2}f(\tau,\overline{\tau})\,\epsilon^{abcd}R_{ab}\wedge R_{cd} + \frac{1}{24\pi^2}g(\tau,\overline{\tau})\,R^a{}_b\wedge R^b{}_a. \tag{5.2.9}$$

Following [45] we choose $f(\tau,\overline{\tau}) = \mathrm{Re}\log\eta(\tau)$ and $g(\tau,\overline{\tau}) = i\,\mathrm{Im}\log\eta(\tau)$, where $\eta(\tau)$ is the Dedekind η-function. With L_{ct} the partition function gets modified to

$$\mathcal{Z}'[\tau,\overline{\tau}] := \eta(\tau)^{-(\chi+\sigma)/2}\eta(\overline{\tau})^{-(\chi-\sigma)/2}\mathcal{Z}[\tau,\overline{\tau}]. \tag{5.2.10}$$

Through the modular properties of the Dedekind η-function:

$$\eta(\tau+1) = \mathrm{e}^{i\pi/12} \quad \text{and} \quad \eta\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\,\eta(\tau),$$

we find that the new partition function transforms as

$$\mathcal{Z}'[\mathsf{T}\cdot(\tau,\overline{\tau})]=\mathcal{Z}'[\tau+1,\overline{\tau}+1]=\mathrm{e}^{-i\pi\sigma/3}\mathcal{Z}'[\tau,\overline{\tau}] \tag{5.2.11}$$

$$\mathcal{Z}'[\mathsf{S}\cdot(\tau,\overline{\tau})]=\mathcal{Z}'\left[-\frac{1}{\tau},-\frac{1}{\overline{\tau}}\right]=\mathcal{Z}'[\tau,\overline{\tau}]. \tag{5.2.12}$$

It is evident that the above transformation constitutes an anomaly from the fact that we weren't able to absorb the transformation of the partition function in a counterterm. However, let us also prove it using the general algorithm for duality anomalies described above. The parameter τ takes values in the upper half plane $\mathbb{H}:=\left\{z\in\mathbb{C}\;\middle|\;\mathrm{Im}(z)>0\right\}$. The duality group in question is $G=\mathrm{SL}(2,\mathbb{Z})$ and the space of background data is $\widehat{\mathcal{M}}=\{\tau\}=\mathbb{H}$. We then consider the line bundle $\mathcal{L}$ over $\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$. We first take the space $\mathbb{C}\times\widehat{\mathcal{M}}$ with coordinates $(\ell,m)=(\ell,\tau)$ and we specify an action of $\mathrm{SL}(2,\mathbb{Z})$ as

$$g\cdot(\ell,\tau)=(\varphi_\tau(g),g\cdot\tau), \tag{5.2.13}$$

where $\mathsf{S}\cdot\tau=-\frac{1}{\tau}$ and $\mathsf{T}\cdot\tau=\tau+1$.

In our case we can choose $\varphi_\tau(g)$ to be the one-dimensional representations of $\mathrm{SL}(2,\mathbb{Z})$, that are τ-independent. Namely, recall that $\mathrm{SL}(2,\mathbb{Z})$ can be expressed through its generators, $\mathsf{T},\ \mathsf{S}$ as

$$\mathrm{SL}(2,\mathbb{Z})=\left\langle\mathsf{S},\mathsf{T}\mid\mathsf{S}^2=\left(\mathsf{T}^{-1}\mathsf{S}\right)^3,\mathsf{S}^4=1\right\rangle. \tag{5.2.14}$$

Then a one-dimensional representation must obey $\rho(\mathsf{S})=\rho(\mathsf{T})^3$ and $\rho(\mathsf{T})^12=1$. Hence, a general one-dimensional representation of $\mathrm{SL}(2,\mathbb{Z})$ is given by:

$$\rho_n(\mathsf{T})=\exp\left(-\frac{2\pi in}{12}\right),\qquad\rho_n(\mathsf{S})=\exp\left(-\frac{2\pi in}{4}\right),\qquad n\in\mathbb{Z}_{12}. \tag{5.2.15}$$

Taking $\varphi_\tau(g)=\rho_n(g)$ specifies the section and reproduces the anomaly. This is because the signature σ that appears in (5.2.11) is a multiple of 16 in spin manifolds, and hence $\exp\left(-\frac{\pi i\sigma}{3}\right)=\exp\left(-\frac{\pi i\sigma}{12}\right)$. So we take n to be $\sigma/2 \bmod 12$. As an aside, the bundle with $n=1$ is called Hodge bundle, Λ, in the literature, while bundles corresponding to ρ_n for general n are $\Lambda^{\otimes n}$ with $\Lambda^{\otimes 12}$ being trivial.

We can also compute the SPT phase in five dimensions that saturates the anomaly. The signature ω is expressible through the first Pontryagin class of the manifold, as

$$\sigma=\int_{X_4}\frac{p_1(X_4)}{3}. \tag{5.2.16}$$

Then if $a\in\mathrm{H}^1(\mathrm{BSL}(2,\mathbb{Z}),\mathrm{U}(1))=\mathbb{Z}_{12}$, the SPT phase

$$\mathcal{Z}_{\mathrm{SPT}}=\exp\left(-\frac{\pi i}{12}\int_{Y_5}a\wedge\frac{p_1(Y_5)}{3}\right), \tag{5.2.17}$$

saturates the anomaly (5.2.11), where $\partial Y_5 = X_4$ and the Pontryaging class is computed in the bulk manifold.

5.3 Theta angle anomaly in quantum mechanics

A famous example of a (-1)-form symmetry afflicted with an anomaly regards a particle on a circle, with a θ angle turned on:

$$S[\phi;\theta] := \frac{1}{2}\int_{\mathbf{S}^1} \mathrm{d}\phi \wedge \star \mathrm{d}\phi + \frac{i}{2\pi}\int_{\mathbf{S}^1} \theta \, \mathrm{d}\phi. \tag{5.3.1}$$

This action has two symmetries: it has a U(1) shift symmetry $\phi \longmapsto \phi + \alpha$ for constant α, coming from the fact that ϕ lives on a circle. We can turn on a background gauge field, $A = A_\tau \, \mathrm{d}\tau$, for U(1), by promoting the exterior derivative to a covariant derivative, i.e. replacing $\mathrm{d}\phi$ by $(\mathrm{d}\phi - A)$. Then the action is invariant under the local symmetry

$$\begin{aligned} \phi &\longmapsto \phi + \alpha(\tau) \\ A &\longmapsto A + \mathrm{d}\alpha(\tau). \end{aligned}$$

The action has also a (-1)-form symmetry, under which $\theta \sim \theta + 2\pi$. This symmetry comes from the fact that $\int \mathrm{d}\phi \in 2\pi\mathbb{Z}$ and thus

$$\exp\left(\frac{i\theta}{2\pi}\int \mathrm{d}\phi\right) = \exp\left(\frac{i(\theta+2\pi)}{2\pi}\int \mathrm{d}\phi\right).$$

For this symmetry θ should be regarded as the 0-form background gauge field. As a group this (-1)-form symmetry is $\mathbb{Z}^{[-1]}$.

As was first shown in [49] and later rephrased in the language of (-1)-form symmetry anomalies in [46, 47], the two symmetries have a mixed anomaly manifesting itself as

$$\mathcal{Z}[\theta+2\pi, A] = \exp\left(i\int_{\mathbf{S}^1} A\right)\mathcal{Z}[\theta, A], \tag{5.3.2}$$

or through the inflow action

$$\mathcal{Z}_{\mathrm{SPT}}[\theta, A] = \exp\left(i\int_{Y_2} \theta \, \mathrm{d}A\right), \qquad \partial Y_2 = \mathbf{S}^1. \tag{5.3.3}$$

Since this is one of the simplest instances of an anomalous (-1)-form symmetry we can test the algorithm for detecting such anomalies in this example. We shall regard θ taking values in $\mathbb{R}$ rather than $\mathbf{S}^1$, so that the action of the group takes us to different values of θ that incidentally are identified in the partition function. Then the space of background data is

$$\widehat{\mathcal{M}} = \{\theta\} \times \{A\} = \mathbb{R} \times \widetilde{\Omega}^1\big(\mathbf{S}^1, \mathrm{U}(1)\big), \tag{5.3.4}$$

where $\widetilde{\Omega}^1\big(\mathbf{S}^1, \mathrm{U}(1)\big)$ is the space of U(1) connections on $\mathbf{S}^1$, i.e. $\widetilde{\Omega}^1\big(\mathbf{S}^1, \mathrm{U}(1)\big) = \Omega^1\big(\mathbf{S}^1, \mathrm{U}(1)\big)/\Omega^0\big(\mathbf{S}^1, \mathrm{U}(1)\big)$. Next we take the space $\mathbb{C}\times\widehat{\mathcal{M}}$ with coordinates $(\ell, m) = (\ell, \theta, A)$ as instructed by the algorithm and we denote the group elements as $n \in \mathbb{Z}^{[-1]}$ and the multiplication of two group elements by addition: $n_1 \circ n_2 := n_1 + n_2$, since $\mathbb{Z}^{[-1]}$ is an abelian group. The action of $\mathbb{Z}^{[-1]}$ on $\mathbb{C} \times \widehat{\mathcal{M}}$ is then:

$$n \cdot (\ell,\ \theta,\ A) = (\varphi_{\theta,A}(n)\ell,\ \theta + 2\pi n,\ A), \tag{5.3.5}$$

while the consistency relation (5.1.2) is

$$\varphi_{\theta,A}(n_2 + n_1) = \varphi_{\theta+2\pi n_1,A}(n_2)\varphi_{\theta,A}(n_1). \tag{5.3.6}$$

Choosing

$$\varphi_{\theta,A}(n) := \exp\left(in \int_{S^1} A\right), \tag{5.3.7}$$

is consistent with (5.3.6) and reproduces the anomaly (5.3.2). Thus indeed the partition function $\mathcal{Z}[\theta, A]$ is a section of the equivariant line bundle

$$\mathcal{L} = \frac{\mathbb{C} \times \mathbb{R} \times \widetilde{\Omega}^1\big(\mathbf{S}^1, \mathrm{U}(1)\big)}{\mathbb{Z}^{[-1]}},$$

specified by the above choice of $\varphi_{\theta,A}(n)$. Clearly, since $\mathbb{Z}^{[-1]}$ doesn't act on A the problem lies in taking the $\mathbb{R}/\mathbb{Z}^{[-1]}$ quotient, so we should in fact view the budle $\mathcal{L}$ as a bunlde over $\mathbf{S}^1 = \mathbb{R}/\mathbb{Z}^{[-1]}$. This tells us that θ cannot take values in $\mathbf{S}^1 = \mathbb{R}/\mathbb{Z}^{[-1]}$ but rather in a line bundle over it.

5.4 Duality anomalies in quantum mechanics

5.4.1 Motivation: two-dimensional compact boson

The usual way to approach the $(d+1)$-dimensional compact boson is to view it as a function $\phi : \mathcal{M}_d \times \mathbf{S}^1_\beta \longrightarrow \mathbf{S}^1_R$, where $\mathcal{M}_d$ is some compact d-dimensional manifold. Under that light, the two-dimensional compact boson is

$$X : \mathbf{S}^1_L \times \mathbf{S}^1_\beta \longrightarrow \mathbf{S}^1_R. \tag{5.4.1}$$

Hence the field X satisfies $X \sim X + 2\pi R$.

Now, the scale of the "spatial"[10] manifold helps us create two dimensionless constants: $\frac{\beta}{L}$ and $\frac{R}{L}$. If we further redefine the field to $\phi := \frac{X}{R}$ so that $\phi \sim \phi + 2\pi$, the compact boson action becomes

$$S[\phi] = \frac{1}{2\pi} \int_{\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}} \frac{\widetilde{R}^2}{2} \, \mathrm{d}\phi \wedge \star \mathrm{d}\phi, \tag{5.4.2}$$

where $\widetilde{R} := \frac{R}{L}$ and $\widetilde{\beta} := \frac{\beta}{L}$ are the dimensionless target space scale and inverse temperature respectively. The partition function is a function of the dimensionless quantities, $\mathcal{Z}[\widetilde{\beta}, \widetilde{R}]$ and is invariant upon inverting the temperature or the target space radius, so

$$\begin{tikzcd}
\mathcal{Z}[\widetilde{\beta}, \widetilde{R}] \arrow[r, equal, "\text{T-duality}"] \arrow[d, equal, "\text{Modular invariance}"] & \mathcal{Z}[\widetilde{\beta}, \widetilde{R}^{-1}] \arrow[d, equal] \\
\mathcal{Z}[\widetilde{\beta}^{-1}, \widetilde{R}] \arrow[r, equal] & \mathcal{Z}[\widetilde{\beta}^{-1}, \widetilde{R}^{-1}].
\end{tikzcd} \tag{5.4.3, 5.4.4}$$

Here we can regard R as a coupling constant (it's like a mass term) and β as a background field, since it's tied with the metric of the base space which we take as a background field.

Let us review the path integral derivation of the T-duality. Consider $B \in \Omega^1\left(\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}\right)$, which we should think of, morally as a background gauge field for the $\mathrm{U}(1)$ shift symmetry: $\phi(\tau) \longmapsto \phi(\tau) + \chi$ of (5.4.2). Take an action, pulled out of a hat for now:

$$\check{S}[B, \phi] := \frac{1}{2\pi} \int_{\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}} \frac{1}{2\widetilde{R}^2} B \wedge \star B + \frac{i}{2\pi} \int_{\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}} B \wedge \mathrm{d}\phi \,. \tag{5.4.5}$$

Since B is an auxiliary field, we can path integrate over it simply by inserting the equations of motion. These are $B = i\widetilde{R}^2 \star \mathrm{d}\phi$. Therefore,

$$S_{\text{eff}}[\phi] = -\log \int \mathrm{D}B \exp\left(-\check{S}[B, \phi]\right) = \frac{1}{2\pi} \int_{\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}} \frac{\widetilde{R}^2}{2} \mathrm{d}\phi \wedge \star \mathrm{d}\phi. \tag{5.4.6}$$

On the other hand, doing the path integral for ϕ first, fixes B to be flat. Note, that in general B is not exact. We can make it exact however, by gauge fixing, so that it's pure gauge, $B = \mathrm{d}\varphi$. Therefore,

$$S_{\text{eff}}[\varphi] = -\log \int \mathrm{D}\phi \exp\left(-\check{S}[B, \phi]\right) \overset{B = \mathrm{d}\varphi}{=\!=\!=} \frac{1}{2\pi} \int_{\mathbf{S}^1 \times \mathbf{S}^1_{\widetilde{\beta}}} \frac{1}{2\widetilde{R}^2} \mathrm{d}\varphi \wedge \star \mathrm{d}\varphi. \tag{5.4.7}$$

[10] it doesn't make sense to speak about spatial manifolds since the theory is defined in Euclidean signature, but by spatial here we denote the piece which is not the thermal circle

Integrating (5.4.6) over ϕ is the same as integrating (5.4.7) over φ since both are the full path integral of (5.4.5), therefore this is a proof of (5.4.3).

5.4.2 One-dimensional compact boson

Consider now the one dimensional compact boson,

$$S[X]=\frac{1}{2}\int_{S^1_\beta}\mathrm{d}\tau\,(\partial X)^2,\qquad X\sim X+2\pi R. \tag{5.4.8}$$

In one dimension, the only parameters we have at our disposal are the size of the thermal circle, β, and the size of the target space circle, R. However, in one dimension the compact boson field has dimensions $[X]=-\frac{1}{2}$, thus also R has dimensions $[R]=-\frac{1}{2}$. Therefore the only dimensionless parameter that we can create combining β and R is (the number 2 is there just for future convenience)

$$\lambda^2:=\frac{2R^2}{\beta}. \tag{5.4.9}$$

For convenience, we strip off the dimensions from X by defining $\phi:=\frac{X}{R}$ like previously and also from τ by redefining $\tau ::= \beta\tau$, so the action takes the form

$$S[\phi]=\frac{1}{2}\int_{S^1}\mathrm{d}\tau\,\frac{R^2}{\beta}(\partial\phi)^2=\frac{1}{4}\int_{S^1}\mathrm{d}\tau\,\lambda^2(\partial\phi)^2. \tag{5.4.10}$$

From the discussion above it follows that neither T-duality nor modular invariance makes sense on its own, because the theory is not allowed to know about dimensionful parameters, so we can only explore what happens under changes of λ^2. In other words, the target space and the base space are mixed up.

There is a mixed duality, however, relating high temperatures to low temperatures, or *equivalently*, large target space circles to small target space circles. We can see the duality as follows. First, observe that there are two ways to obtain the partition fuction. Firstly, it is a very simple quantum mechanical model, so we can solve it exactly. Its spectrum is simply $E_n=\frac{n^2}{2R^2}$. The the partition function is then

$$\mathcal{Z}_{\mathrm{QM}}[\lambda]=\sum_{n\in\mathbb{Z}}\exp\left(-\beta\frac{n^2}{2R^2}\right)=\sum_{n\in\mathbb{Z}}\exp\left(-\frac{n^2}{\lambda^2}\right). \tag{5.4.11}$$

On the other hand, as a quadratic quantum field theory we can perform the path integral explicitly, so

$$\begin{aligned}\mathcal{Z}[\lambda] &= \int_{\text{periodic}} \mathrm{D}\phi \, \exp\left(-\frac{1}{2}\int_0^1 \mathrm{d}\tau \;\lambda^2(\partial\phi)^2\right) = \\ &= \sum_{m\in\mathbb{Z}} \int_{\phi(1)-\phi(0)=2\pi m} \exp\left(-\frac{1}{2}\int_0^1 \mathrm{d}\tau \;\lambda^2(\partial\phi)^2\right) = \\ &= \sum_{m\in\mathbb{Z}} \exp\left(-\lambda^2 m^2\right) = \mathcal{Z}_{\mathrm{QM}}[\lambda^{-1}],\end{aligned} \tag{5.4.12}$$

where the last path integral can be performed by decomposing $\phi(\tau)$ in Fourier modes. The two results can be mapped to one another by Poisson resummation. Performing the Poisson resummaion formula carefully we find that

$$\mathcal{Z}\left[\lambda^{-1}\right] = \lambda^{-1/4}\, \mathcal{Z}[\lambda]. \tag{5.4.13}$$

This is interpreted as a duality which mixes up the modular transformations in one dimension, $\beta \longrightarrow \frac{1}{\beta}$, and target space duality, $R \longrightarrow \frac{1}{R}$. Nevertheless it is not an exact duality. It has is an ambiguity controlled by λ, which we will interpret as a duality anomaly. The main goals in the following sections is to prove that it is indeed an anomaly, to understand what the origin of the anomaly is and to give an possible implications of this anomaly.

5.4.3 The partition function is not a function

A test to verify that the theory is anomalous is to show that its partition function is not a genuine function on the space of coupling constants, but rather a section of a line bundle over that space. Here we only have one coupling constant, λ or equivalently, σ, defined by $\mathrm{e}^{i\sigma} = \lambda^2$. The space of coupling constants is $\widehat{\mathcal{M}} := \{\sigma\} = i\mathbb{R}$. The relevant duality group is $\mathbb{Z}_2$. We will show that $\mathcal{Z}[\sigma]$ is not a function on $\widehat{\mathcal{M}}/\mathbb{Z}_2$, but a section of some line bundle, which we will construct. $\widehat{\mathcal{M}}$ is trivial as a space, thus the only complex line bundle we can take over it is the trivial bunlde, $\mathbb{C}\times\widehat{\mathcal{M}}$. Take then $(\ell,\sigma)\in\mathbb{C}\times\widehat{\mathcal{M}}$, be coordinates in that space, along with an action of $\mathbb{Z}_2$ as

$$\mathbb{Z}_2 = \langle 1, \bullet \mid \bullet^2 = 1\rangle \ni g : (\ell,\sigma) \longmapsto (\varphi_\sigma(g)\ell,\; g\cdot\sigma), \tag{5.4.14}$$

where $g\cdot\sigma$ is: $1\cdot\sigma = \sigma$ and and $\bullet\cdot\sigma = -\sigma$. To be a consistent action, $\varphi_\sigma(g)\in\mathbb{C}\setminus\{0\}$ must satisfy the following relation:

$$\varphi_\sigma(f\,g) = \varphi_{g\cdot\sigma}(f)\varphi_\sigma(g). \tag{5.4.15}$$

The partition function is then a section of the (honest) equivariant[11] line bundle $\mathscr{M} := \dfrac{\mathbb{C} \times \widehat{\mathscr{M}}}{\mathbb{Z}_2}$ specified by

$$\mathcal{Z}[g \cdot \sigma] = \varphi_\sigma(g)\mathcal{Z}[\sigma]. \tag{5.4.16}$$

From (5.1.2), follows that $\varphi_\sigma(1) = 1$ and $\varphi_{-\sigma}(\bullet) = [\varphi_\sigma(\bullet)]^{-1}$. Choosing $\varphi_\sigma(\bullet) = \exp\left(-\frac{i\sigma}{8}\right)$, or equivalently

$$\varphi_\sigma(g) = \exp\left(-\frac{i\sigma}{16} + \frac{ig \cdot \sigma}{16}\right), \qquad g \in \mathbb{Z}_2 \tag{5.4.17}$$

satisfies the relation (5.1.2) and determines the section.

5.4.4 On the search for an SPT

As we have seen in 3.2, every 't Hooft anomaly must be captured by an SPT phase in one higher dimension. In two dimensions the group cohomology classification is complete. For the group cohomology of $\mathbb{Z}_2$ we have the following.

If the $\mathbb{Z}_2$ doesn't act on U(1),

$$\mathrm{H}^2(\mathrm{B}\mathbb{Z}_2, \mathrm{U}(1)) \cong \mathscr{H}^2(\mathbb{Z}_2, \mathrm{U}(1)) = 0. \tag{5.4.18}$$

Thus every SPT phase must be trivial, which is cannot match the fact that there is an anomaly.

If the $\mathbb{Z}_2$ acts on the $U(1)$, then the group cohomology classes are the different group extensions, E:

$$1 \longrightarrow \mathbb{Z}_2 \longrightarrow E \xrightarrow{\pi} \mathrm{U}(1) \longrightarrow 1. \tag{5.4.19}$$

These are either the trivial group extension $E = \mathrm{U}(1) \times \mathbb{Z}_2$, with π being just the projection on the first factor, or the non-trivial extension, $E = U(1)$, with $\pi(x) = x^2$; in other words the Möbius strip. Hence

$$\mathrm{H}^2_\pi(\mathrm{B}\mathbb{Z}_2, \mathrm{U}(1)) \cong \mathscr{H}^2_\pi(\mathbb{Z}_2, \mathrm{U}(1)) = \mathbb{Z}_2. \tag{5.4.20}$$

Symmetries of this sort are spacetime symmetries. However in one dimension the only such symmetry is time-reversal and it is not the case at hand.

Thus this anomaly is not a pure anomaly. Luckily, the theory we are considering has a conformal symmetry which we can exploit to identify the anomaly.

[11] up to an abuse in semantics regarding fixed points

5.4.5 Pinning down the anomaly

We want to think about this duality anomaly as a mixed anomaly between the conformal group in one dimension, $\mathrm{SL}(2,\mathbb{R})$, and the (-1)-form symmetry, $\mathbb{Z}_2^{[-1]}$ corresponding to the duality.

For this let us consider a different system with symmetry Γ given by a non-trivial extension of $\mathrm{SL}(2,\mathbb{R})$, as

$$1 \longrightarrow \mathbb{Z}_2 \longrightarrow \Gamma \longrightarrow \mathrm{SL}(2,\mathbb{R}) \longrightarrow 1. \tag{5.4.21}$$

It is known that the unique non-trivial $\mathbb{Z}_2$ central extension of $\mathrm{SL}(2,\mathbb{R})$ is given by the Weil representation, i.e. the metaplectic group of order 2 over the reals, $\mathrm{Mp}(2,\mathbb{R})$. Therefore, we can essentially say that

$$\mathrm{H}^2\big(\mathrm{SL}(2,\mathbb{R}),\mathbb{Z}_2\big) = \mathbb{Z}_2. \tag{5.4.22}$$

This means that if we have a theory with $\mathrm{Mp}(2,\mathbb{R})$ symmetry, where $0 \neq \varepsilon \in \mathrm{H}^2\big(\mathrm{SL}(2,\mathbb{R}),\mathbb{Z}_2\big)$ we can gauge a $\mathbb{Z}_2$ subgroup of it and land on a theory with symmetry

$$\mathbb{Z}_2^{[-1]} \times \mathrm{SL}(2,\mathbb{R}), \tag{5.4.23}$$

with a mixed anomaly

$$\mathcal{Z}_{\mathrm{SPT}}[\hat{A},B] = \exp\left(i\int_{Y_2} \hat{A} \smile \varepsilon(B)\right), \tag{5.4.24}$$

where $\hat{A}$ is a zero-form $\mathbb{Z}_2^{[-1]}$-background gauge field, and B is a one-form $\mathrm{SL}(2,\mathbb{R})$-valued background gauge field for the conformal symmetry.

A priori it is not certain that this is our anomaly. In general the anomaly can be hidden anywhere inside

$$\mathrm{H}^{d+1-k}\big(\mathrm{SL}(2,\mathbb{R}),\mathrm{H}^k(\mathbb{Z}_2,\mathrm{U}(1))\big), \qquad k \in \{0,\dots,d+1\}. \tag{5.4.25}$$

Here, $d=1$, so we have very few steps to check. We have:

- $\boldsymbol{k=0}$: $\mathrm{H}^2(\mathrm{SL}(2,\mathbb{R}),\mathbb{Z}_2) = \mathbb{Z}_2$. This corresponds to (5.4.24)

- $\boldsymbol{k=1}$: $\mathrm{H}^1\big(\mathrm{SL}(2,\mathbb{R}),\mathrm{H}^1(\mathbb{Z}_2,\mathrm{U}(1))\big) = \mathrm{H}^1(\mathrm{SL}(2,\mathbb{R}),\mathbb{Z}_2) = \mathrm{Hom}(\mathrm{SL}(2,\mathbb{R}),\mathbb{Z}_2) = 0$. The last step comes from the fact that $\mathrm{SL}(2,\mathbb{R})$ equals its own commutator subgroup, $[\mathrm{SL}(2,\mathbb{R}),\mathrm{SL}(2,\mathbb{R})]$, and hence its abelianisation is trivial, which in turn implies that the set of group homomorphisms of $\mathrm{SL}(2,\mathbb{R})$ to an abelian group is trivial. Physically it means that there cannot be an anomaly of the form $\hat{A}\smile\hat{A}\smile\zeta(B)$.
- $\boldsymbol{k=2}$: $\mathrm{H}^0\big(\mathrm{SL}(2,\mathbb{R}),\mathrm{H}^2(\mathbb{Z}_2,\mathrm{U}(1))\big) = \mathrm{H}^0\big(\mathrm{SL}(2,\mathbb{R}),0\big) = 0$. So there cannot be a third order anomaly as well.

To summarise, the anomaly in question is indeed a mixed anomaly between the $\mathbb{Z}_2^{[-1]}$ duality and conformal symmetry, given by (5.4.24).

5.5 Berry phase as an anomaly

Let us consider the standard quantum mechanical scenario that gives rise to the Berry phase. Consider a system given by a Hamiltonian, $H\big(\lambda(t)\big)$, where $\lambda(t)$ is a set of *slowly varying* parameters, with a periodicity condition, $\lambda(t)=\lambda(t+T)$, for some T.

Pick an instantaneous eigenbasis, $|\psi_n(\lambda)\rangle$, of this Hamiltonian. The overlap $\langle\psi_n(\lambda)|\Psi(t)\rangle$, where $|\Psi(t)\rangle$ is a solution to the Schrödinger equation for H, transform after a period as

$$\langle\psi_n(\lambda)|\Psi(t+T)\rangle = \exp\left(i\oint_S A_n\right)\langle\psi_n(\lambda)|\Psi(t)\rangle\,, \tag{5.5.1}$$

with

$$A_n := \langle\psi_n(\lambda)|\partial_\lambda|\psi_n(\lambda)\rangle\cdot \mathrm{d}\lambda. \tag{5.5.2}$$

up to what we would call a local counterterm.

We may reformulate the equations a bit. If we write $\langle\psi_n(\lambda)|\Psi(t)\rangle =: \mathcal{Z}[A_n,t]$, which we may do as A_n entails all the information that we need, (5.5.1) becomes

$$\mathcal{Z}[A_n,t+T] = \exp\left(i\oint_S A_n\right)\mathcal{Z}[A_n,t]. \tag{5.5.3}$$

This now looks exactly like a mixed anomaly between a $t\sim t+T$ (-1)-form symmetry and a 0-form $\mathrm{U}(1)$ symmetry. In fact (5.5.3) can be captured by the 2d SPT (with X such that $\partial X = S$):

$$\mathcal{Z}_{\mathrm{SPT}}[A_n,t] = \exp\left(i\int_X \frac{t}{T}\,F_n\right). \tag{5.5.4}$$

Very recently Kapustin and Spodyneiko [50] studied higher Berry curvatures, where the curvature is now a $(d+1)$ form, $F_{[d+1]}$ valued on the parameter space of a d-dimensional QFT. It can be thought as a mixed anomaly between (-1)-form symmetry $t \sim t + T$ and $\mathrm{U}(1)^{[d-1]}$.

Hsin, Kapustin and Thorngren [51] classified the possible Berry phases in a QFT, with parameter space M as a family of invertible theories parametrised by M classified by

$$\mho_{\bullet}^{d+1}(M). \tag{5.5.5}$$

This viewpoint strengthens the relation of Berry phase with anomalies. Compared to (3.4.16) we see that Berry phases can indeed be thought of as a (-1)-form anomaly.

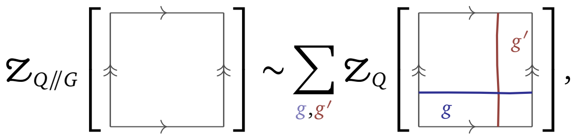
6 Categorical symmetries

To motivate this chapter, let us go back to the concept of gauging and ungauging (chapter 4). There, the most general case involved a theory Q, with a finite abelian group, G. To that, we found that upon gauging, the symmetry comes back, this time as the Pontryagin dual $\hat{G}$. We saw further that upon gauging $\hat{G}$, we find the original theory, effectively ungauging the symmetry. This concept can be morally extended to the case of non-finite abelian groups. The most general non-finite abelian group can be written as products of two basic such groups: $\mathrm{U}(1)$ and $\mathbb{Z}$, and any finite-abelian group we would like. Therefore for our purposes, it suffices to study these two groups. In fact, these groups are both a limit $\mathbb{Z}_{k\to\infty}$, in a particular way, creating either a compact or a non-compact group. Furthermore, they are Pontryagin dual to each other. The only caveat to fully extending the gauging/ungauging pattern to non-finite abelian groups is that $\mathrm{U}(1)$ can accommodate non-flat bundles and those bundles cannot be described by the naïve sum over defects procedure we presented above. One possible solution would be to include δ-function curvature on the triple intersections of defects and then proceed to sum over them, taking into account the curvature. Equivalently this can be captured by summing over gauge fields with a fixed holonomy. It is still not fully understood how to overcome this problem, but it is believed that most of the results of chapter 4 carry through. We will not focus on this problem here.

Instead, we will focus on the next natural question one should ask; whether we can extend this pattern to non-abelian groups. For simplicity, let us focus on finite non-abelian groups so that we tackle one problem at a time. In this case, when we consider a non-abelian finite group G, not all of its representations are one-dimensional. This implies the well-known fact that the product of two irreducible representations is not an irreducible representation. Therefore we cannot pass a group structure to the representations of G. We instead have a representation ring, which in turn forms a representation category $\mathbf{Rep}(G)$. We can, therefore, try to consider $\mathbf{Rep}(G)$ as the symmetry of $Q/\!/G$. By the Tannaka reconstruction theorem, $\mathbf{Rep}(G)$ entails enough information to recover G, so we could in principle gauge $\mathbf{Rep}(G)$ in such a way that $Q/\!/G/\!/\mathbf{Rep}(G) \cong Q$. Furthermore, since $\mathbf{Rep}(G)$ is obviously not the same mathematical structure as G, we are tempted to ask whether we can find a structure such that G and $\mathbf{Rep}(G)$ are special cases of and if the answer is positive we will conjecture that this is the most general symmetry structure a QFT can have.

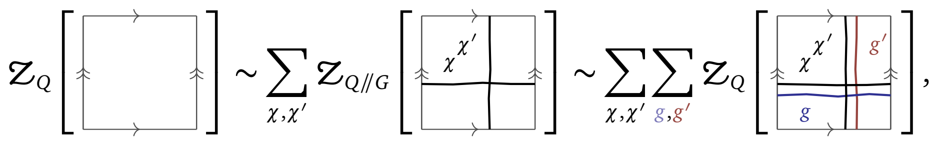
6.1 Categorical symmetry in two dimensions

To elaborate on the introductory paragraph, we will focus on the two-dimensional case, where the result is known in the literature [10]. Consider for simplicity a theory, Q, with a finite abelian (non-anomalous) symmetry, G, defined on a torus. The partition function of the gauged theory is



(6.1.1)

since gauging is summing over the background defects. The background defects of the gauged theory are the various characters of the various original group elements, which are the elements of the Pontryagin dual symmetry group and hence regauging means that (since $Q/\!/G/\!/\hat{G} \equiv Q$)

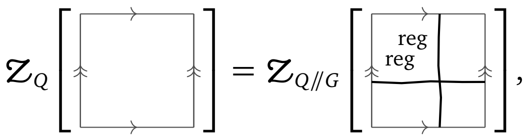


(6.1.2)

The last diagram produces $\chi'(g)$ and $\chi(g')$ at the meeting points and the summation over χ enforces $g' = 1$, while summation over χ' enforces $g = 1$. There is however a distinguished representation, called the regular representation, the character of which obeys (or rather is defined through) the relation

$$\sum_{\chi\in\widehat{G}} \chi(g) =: \chi_{\mathrm{reg}}(g). \tag{6.1.3}$$

In this representation there is a basis e_g such that $\rho_{\mathrm{reg}}(g)e_h = e_{gh}$, and hence we see that $\chi_{\mathrm{reg}}(g) = \delta_{g,1}$. This allows us to formally "pass the sum inside the argument", writing

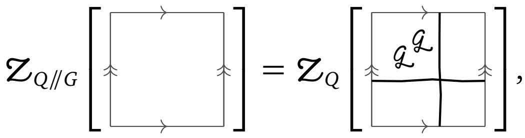


(6.1.4)

where now

$$\Big|^{\mathrm{reg}} = \sum_{R\ \mathrm{irreps}} \dim R\ \Big|^{R} . \tag{6.1.5}$$

Now we can extend the latter to the case where the group is non-abelian, as



(6.1.6)

with

$$\Big|^{\mathcal{L}} = \sum_{g} \Big|^{g} . \tag{6.1.7}$$

In the last generalisation, the non-trivial step is that we are considering the lines $\mathcal{L}$, not as domain walls implementing group elements, but as domain walls implementing objects of a *unitary fusion category*.

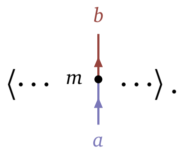
A unitary fusion category, henceforth referred to as a symmetry category **Sym**, consists of the following data:

Objects: The objects are the topological line operators generating the symmetry. For any topological line operator labelled by an object $a \in \mathrm{Obj}(\mathbf{Sym})$, inserted along some oriented curve γ, we can compute the partition function at the presence of a and denote it as

$$\langle \cdots a(\gamma) \cdots \rangle := \langle \cdots \ \big\uparrow_{a} \ \cdots \rangle , \tag{6.1.8}$$

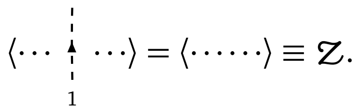
where the dots denote other operators, inserted away from γ.

Morphisms: The morphisms of **Sym** correspond to topological possibly-line-changing local operators. For any two objects $a, b \in \text{Obj}(\textbf{Sym})$ we can insert a line-changing operator $m : a \longrightarrow b$ ($m \in \text{Hom}(a, b)$) to construct the following partition function



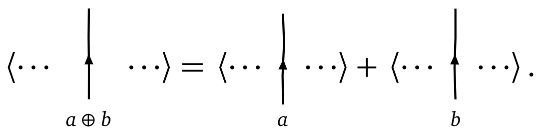
(6.1.9)

Trivial line: **Sym** has an object **1** which is the trivial line:



(6.1.10)

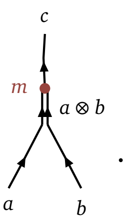
Additive structure: For every two objects $a, b \in \text{Obj}(\textbf{Sym})$ there exists an object $a \oplus b \in \text{Obj}(\textbf{Sym})$, which implements the partition function:



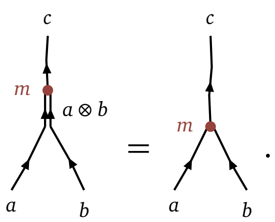
(6.1.11)

In the following we might write $a \oplus a$ as $2a$ to abbreviate the notation.

Tensor structure: For every two objects $a, b \in \text{Obj}(\textbf{Sym})$ there exists an object $a \otimes b \in \text{Obj}(\textbf{Sym})$. This corresponds to having a and b running parallel to each other, as one line operator. We can use the morphism $m \in \text{Hom}(a \otimes b, c)$ to fuse $a \otimes b$ and c together:



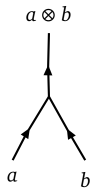
Since all these operators are topological we can move the morphism exactly at the point were a and b started to run parallelly, so we have



Finally after that we can reinterpret what $a \otimes b$ means. It means that we fuse a with b using the identity morphism

$$\begin{aligned} \text{id} : \textbf{Sym} \otimes \textbf{Sym} &\longrightarrow \textbf{Sym} \\ a \otimes b &\longmapsto (a \otimes b), \end{aligned}$$

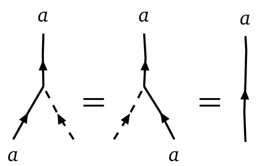
so we will draw



To this tensor structure the object $1(\gamma)$ is the multiplicative unit, and for that we will sometimes call it the unit object. It satisfies

$$a \otimes 1 \cong 1 \otimes a \cong a. \tag{6.1.12}$$

Furthermore we can always find an equivalent category such that this isomorphism is a strict equality. Therefore we will assume that we have already done so. Diagrammatically we write



(6.1.13)

Finiteness: Symmetry categories correspond to finite symmetries of the physical theory. The mathematical requirement is that every object in **Sym** is *semisimple*. In particular, simple objects, $a \in$ **Sym** are objects for which $\dim \mathrm{Hom}(a,a) = 1$. To be more precise, this means that for simple objects, the only morphism that exists is the identity morphism. This makes $\mathrm{Hom}(a,a)$ isomorphic to $\mathbb{C}$ as an algebra. For all categorical symmetries, we will assume that the identity object, 1, is simple. In other words, this tells us that there are no non-trivial topological local operators that do not change simple objects. Semisimplicity requires that all objects, $x \in$ **Sym** can be written as a finite sum:

$$x = \bigoplus_{a \text{ simple}} N_a a, \tag{6.1.14}$$

where $N_a \in \mathbb{Z}_{\geqslant 0}$. To impose finiteness to the above sum, we assume that there is a finite number of isomorphism classes of simple objects.

Associative structure: In a symmetry category there is a distinguished collection of isomorphisms, that are called *associators*:

$$\alpha_{a,b,c} : (a \otimes b) \otimes c \longrightarrow a \otimes (b \otimes c). \tag{6.1.15}$$

In the literature of RCFTs and three-dimensional TQFTs, these are known as fusion matrices F or as 6j symbols. The associators have to satisfy the pentagon identity, i.e.

$$
\begin{tikzcd}
 & \big((a\otimes b)\otimes c\big)\otimes d \arrow[dl, "\alpha_{a,b,c}\otimes 1_d"'] \arrow[dr, "\alpha_{a\otimes b,c,d}"] & \\
\big(a\otimes(b\otimes c)\big)\otimes d \arrow[d, "\alpha_{a,b\otimes c,d}"] & & (a\otimes b)\otimes(c\otimes d) \arrow[d, "\alpha_{a,b,c\otimes d}"] \\
a\otimes\big((b\otimes c)\otimes d\big) \arrow[rr, "1_a\otimes\alpha_{a,c,d}"] & & a\otimes\big(b\otimes(c\otimes d)\big)\ .
\end{tikzcd}
\tag{6.1.16}
$$

Dual structure: For every object $a \in \mathbf{Sym}$, there exists a dual object, $a^* \in \mathbf{Sym}$, such that

$$
\langle\cdots \underset{a}{\big\uparrow} \cdots\rangle = \langle\cdots a(\gamma)\cdots\rangle \overset{!}{=} \langle\cdots a^*(\overline{\gamma})\cdots\rangle = \langle\cdots \underset{a^*}{\big\downarrow} \cdots\rangle\,, \tag{6.1.17}
$$

where $\overline{\gamma}$ is the same path as γ but with the opposite orientation. We require that $(a^*)^* \cong a$. Furthermore, taking duals reverses the order of tensoring:

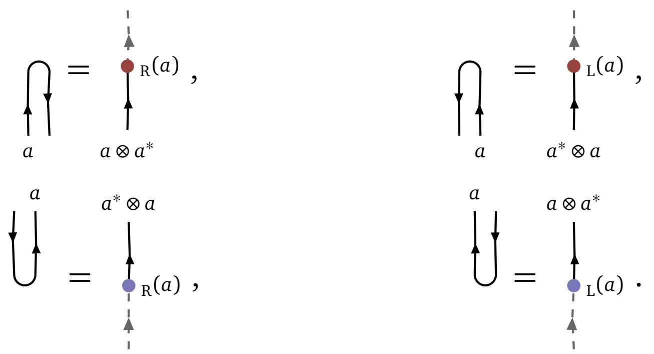
$$
(a\otimes b)^* = b^*\otimes a^*. \tag{6.1.18}
$$

We also demand that there exist evaluation and coevaluation morphisms

$$
\color{brown}{\bullet}_{\mathrm{R}}(a) : a\otimes a^* \longrightarrow 1, \qquad \color{brown}{\bullet}_{\mathrm{L}}(a) : a^*\otimes a \longrightarrow 1, \tag{6.1.19}
$$

$$
\color{blue}{\bullet}_{\mathrm{R}}(a) : 1 \longrightarrow a^*\otimes a, \qquad \color{blue}{\bullet}_{\mathrm{L}}(a) : 1 \longrightarrow a\otimes a^*. \tag{6.1.20}
$$

These maps correspond to folding the operator a:



(6.1.21)

(6.1.22)

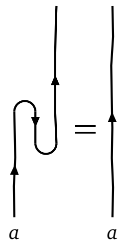
Note that $\bullet_{\mathrm{R}}(a)$ is not necessarily equal to $\bullet_{\mathrm{L}}(a^*)$ but they satisfy the following relations

$$
\bullet_{\mathrm{R}}(a) = \bullet_{\mathrm{L}}(a^*)\circ\big(p(a)\otimes 1\big) = \bullet_{\mathrm{L}}(a^*)\circ\big(1\otimes p(a^*)^{-1}\big),
$$

where $p(a) : a \longrightarrow a$, is known as a *pivotal structure*. Similar relations apply for ${\color{blue}\bullet}_{\mathrm{R}}(a)$ and ${\color{blue}\bullet}_{\mathrm{L}}(a^*)$. Furthermore, ${\color{red}\bullet}$ and ${\color{blue}\bullet}$, should obey the consistency condition:

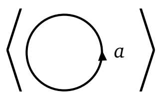
$$({\color{red}\bullet}_{\mathrm{R}}(a) \otimes 1) \circ \left(\alpha_{a,a^*,a}\right)^{-1} \circ (1 \otimes {\color{blue}\bullet}_{\mathrm{R}}(a)) = 1, \tag{6.1.23}$$

ensuring that the following equation is true:

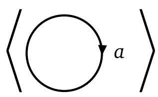


$$. \tag{6.1.24}$$

Evaluation and coevaluation morphisms allow us to define the dimension of a line operator. A subtlety is that a priori there exist clockwise and counter-clockwise dimensions:



$$=: \dim_{\mathrm{cc}} a \tag{6.1.25}$$



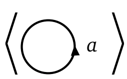
$$=: \dim_{\mathrm{c}} a, \tag{6.1.26}$$

where the closed loops are defined by the morphisms $1 \xrightarrow{{\color{blue}\bullet}_{\mathrm{R}}(a)} a^* \otimes a \xrightarrow{{\color{red}\bullet}_{\mathrm{L}}(a)} 1$ and $1 \xrightarrow{{\color{blue}\bullet}_{\mathrm{L}}(a)} a \otimes a^* \xrightarrow{{\color{red}\bullet}_{\mathrm{R}}(a)} 1$. Since changing the orientation of the line amounts to renaming a to a^* we find that

$$\dim_{\mathrm{cc}} a = \dim_{\mathrm{c}} a^* \tag{6.1.27}$$

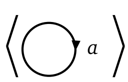
$$\dim_{\mathrm{c}} a = \dim_{\mathrm{cc}} a^*. \tag{6.1.28}$$

In certain cases, for example when the theory is placed on a sphere we can further see that $\dim_{\mathrm{cc}} a = \dim_{\mathrm{c}} a$. In particular, take the only insertion to be a small counter-clockwise loop of a around the north pole of the sphere. Then,



$$= (\dim_{\mathrm{cc}} a)\, \mathcal{Z}_{S^2}, \tag{6.1.29}$$

where $\mathcal{Z}_{S^2}$ is the bare partition function of the theory on the sphere. Since a is topological and there are no other insertions we can move it so that it is now a small clockwise loop around the south pole, in which case we find,



$$= (\dim_{\mathrm{c}} a)\, \mathcal{Z}_{S^2}. \tag{6.1.30}$$

Since this is the same computation, equating the two descriptions we find

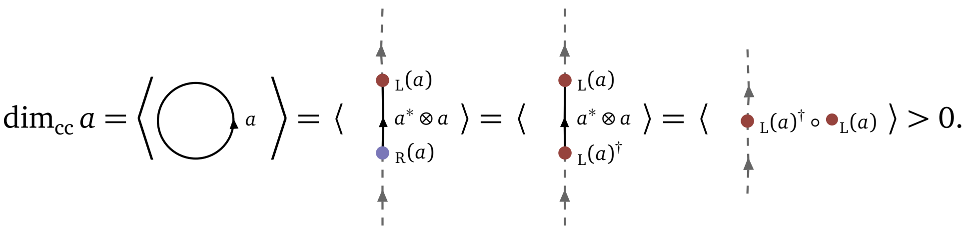
$$\dim_{\mathrm{cc}} a = \dim_{\mathrm{c}} a =: \dim a. \tag{6.1.31}$$

For theories on generic two-dimensional manifolds this does not hold in general; for example a counter-clockwise loop wrapping a non-trivial cycle of a torus will always be counter-clockwise. If we demand nevertheless that (6.1.31) holds, the symmetry category is called a *spherical fusion category*.

Unitary structure: The last demand of a symmetry category is the existence of a unitary structure, i.e. an anti-linear map $\mathrm{Hom}(a,b) \ni m \longmapsto m^\dagger \in \mathrm{Hom}(b,a)$, for which it holds that $m^\dagger \circ m \in \mathrm{Hom}(a,a)$ is semi-positive definite[12]. The (co)evaluation morphisms transform under this map as

$$\bullet_{\mathrm{R}}(a) = \bullet_{\mathrm{L}}(a)^\dagger, \qquad \bullet_{\mathrm{L}}(a) = \bullet_{\mathrm{R}}(a)^\dagger. \tag{6.1.32}$$

As the name implies, the unitarity structure is tied to unitarity of QFTs, or rather reflection-positivity since we are working in Euclidean signature. This makes evident that all the topological lines have non-negative dimension:

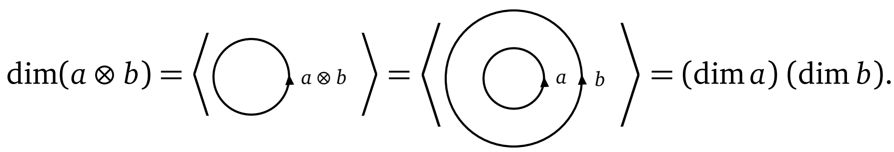


(6.1.33)

The non-negativity is clear from the semi-positive definitness of any $m^\dagger \circ m$. Moreover here we have that $\bullet_{\mathrm{L}}(a)^\dagger \circ \bullet_{\mathrm{L}}(a) \in \mathrm{Hom}(1,1)$, therefore since the identity object is simple we find that $\dim_{\mathrm{cc}} a$ is strictly positive. Similarly it holds that $\dim_{\mathrm{c}} a > 0$ for all $a \in \mathrm{Obj}(\mathbf{Sym})$.

In fact one can show through an intermediate argument using the Perron–Frobenius theorem from linear algebra, and the fusion ring (cf. a few paragraphs later) of the symmetry category, that in fact $\dim_{\mathrm{cc}} a = \dim_{\mathrm{c}} a =: \dim a$ and hence unitarity implies sphericity. Another consequence of that is that $\dim a = \dim a^*$. Finally the dimensions of direct sums of objects and tensor products of objects are given as follows:

$$\dim(a \oplus b) = \dim a + \dim b \tag{6.1.34}$$



(6.1.35)

Albeit this long list of axioms looks complicated, it is merely a formalisation in a categorical setting of the behaviour of topological operators in a unitary two-dimensional QFT.

[12] Due to semisimplicity and finiteness $\mathrm{Hom}(a,a)$ can be thought of as a direct sum of a martix algebra. Then we require $m^\dagger \circ m$ to have non-negative eigenvalues.

The construction of the above requirements is akin to the Moore–Seiberg [52] categorification of symmetries of two-dimensional RCFTs, and three-dimensional TQFTs. There the categorification yielded what is known as a unitary modular tensor category (UMTC), which is, in fact, a unitary fusion category, with extra structure.

Symmetry functors In traditional group theory we have group homomorphisms $G_1 \longrightarrow G_2$ between groups G_1 and G_2. Using these we can establish criteria for when the two groups are equivalent (isomorphic). Similarly, for symmetry categories, we have *symmetry functors*

$$F : \mathbf{Sym}_1 \longrightarrow \mathbf{Sym}_2, \tag{6.1.36}$$

accompanied by morphisms, that map all the data and structures of the symmetry category $\mathbf{Sym}_1$ to data and structures of $\mathbf{Sym}_2$. One such is a set of isomorphisms

$$\epsilon_F(a,b) \in \mathrm{Hom}(F(a) \otimes F(b), F(a \otimes b)), \tag{6.1.37}$$

which specify how the tensor structure of $\mathbf{Sym}_2$ corresponds to the tensor structure of $\mathbf{Sym}_1$. Two symmetry functors, F, G are equivalent whenever there is a set of isomorphisms

$$\eta(a) \in \mathrm{Hom}(F(a), G(a)),$$

such that

$$\eta(a \otimes b)\, \epsilon_F(a,b) = \epsilon_G(a,b)\, [\eta(a) \otimes \eta(b)]. \tag{6.1.38}$$

When a symmetry functor $F : \mathbf{Sym}_1 \longrightarrow \mathbf{Sym}_2$ has an inverse, $F^{-1} : \mathbf{Sym}_2 \longrightarrow \mathbf{Sym}_1$, then $\mathbf{Sym}_1$ and $\mathbf{Sym}_2$ are equivalent as symmetry categories.

Product symmetry For two symmetry categories $\mathbf{Sym}_1$ and $\mathbf{Sym}_2$ we can define the product category $\mathbf{Sym}_1 \boxtimes \mathbf{Sym}_2$, whose simple objects are $a_1 \boxtimes a_2$, where a_1 and a_2 are simple objects of $\mathbf{Sym}_1$ and $\mathbf{Sym}_2$ respectively. This product is called Deligne's product.

Fusion ring If we are interested in global features of the symmetry category, that do not depend on the extra structures that we have postulated, we can work with isomorphism classes; for each simple object a, we denote its isomorphism class as $[a]$. Then, if $a,\ b$ are simple objects, $a \otimes b$ is a semisimple object and as we have seen before it has a decomposition in simple objects as

$$a \otimes b = \bigoplus_c N_{ab}^c c, \qquad N_{ab}^c \in \mathbb{Z}_{\geqslant 0}. \tag{6.1.39}$$

If we define the product of isomorphism classes as

$$[a][b] := \sum_c N_{ab}^c [c], \tag{6.1.40}$$

this promotes the non-negative linear combinations of classes of simple objects to an algebra over $\mathbb{Z}_{\geqslant 0}$, which is known as the fusion ring $R(\mathbf{Sym})$ of the symmetry category. It is clear that having access to the fusion ring coefficients, we can determine the dimensions of the objects. For RCFTs, i.e. when the symmetry category is enriched with extra structure, this is achieved by the Verlinde formula [53], but a Verlinde-like formula for general unitary fusion categories is not known. Finally, the dimension of the symmetry category is defined as

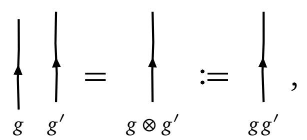
$$\dim \mathbf{Sym} := \sum_{[a]} (\dim a)^2. \tag{6.1.41}$$

6.2 Examples in two-dimensions

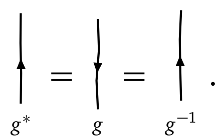
In this section, we will apply the formalism presented above to explore some instances of categorical symmetry in two-dimensional theories.

6.2.1 Group symmetry as categorical symmetry

As a first example let us simply recast the familiar case of a group symmetry, based on a finite group G, in terms of the new categorical notation. The simple objects are group elements $g, g', \dots \in G$ for which



(6.2.1)



(6.2.2)

Taking all the associators α_{g_1,g_2,g_3} to be trivial for all $g_i \in G$, all the other data, such as $\bullet_{\mathrm{R,L}}$ and $\bullet_{\mathrm{R,L}}$ are determined by α and thus we form a fusion category where all the simple objects are invertible, that we denote by $\mathbf{Sym}(G)$[13]. Since for any $g \in G$ there is some

[13]In other places in the literature, such as [13] it is denoted as Vec_G

$n \in \mathbb{Z}_{\geqslant 0}$ such that $g^n = 1$, we have that

$$1 = \dim(1) = \dim(g^n) = (\dim g)^n.$$

Furthermore since $\dim g > 0$, this implies that $\dim g = 1$. Due to the simplicity of the objects, the total dimension of the symmetry category is

$$\dim \mathbf{Sym}(G) = |G|, \tag{6.2.3}$$

the order of the group.

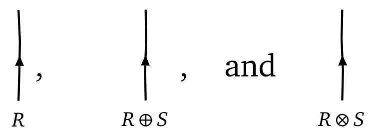
Above we made the choice to take all α_{g_1,g_2,g_3} to be trivial. If we relax this choice we see that due to the pentagon identity (6.1.16) α is a 3-cocycle on G. In addition, if any of g_1, g_2, g_3 is the identity α should vanish, which makes α a normalised cocycle. Therefore it takes values in U(1). Again the rest of the data are determined by the choice of α. Therefore this structure also defines a unitary fusion category that we denote as $\mathbf{Sym}(G,\alpha)$[14]. Whenever we have two symmetry categories, $\mathbf{Sym}(G,\alpha_1)$ and $\mathbf{Sym}(G,\alpha_2)$ where $\alpha_1 - \alpha_2 = d\epsilon$ for some 2-cocycle ϵ, we can show that the two categories are equivalent, in the sense of the discussion on symmetry functors. Therefore we can safely mod out by this sort of coboundaries in the definition of $\mathbf{Sym}(G,\alpha)$ and take $\alpha \in \mathrm{H}^3\big(G, \mathrm{U}(1)\big)$. This is the most general fusion category whose simple objects are invertible. Therefore we can claim that

A symmetry category whose simple lines are all invertible is of the form $\mathbf{Sym}(G,\alpha)$ for some group G and $\alpha \in \mathrm{H}^3\big(G, \mathrm{U}(1)\big)$.

Notice, that in two dimensions, the group cohomology group $\mathrm{H}^3\big(G, \mathrm{U}(1)\big)$ classifies the anomalies of a QFT with symmetry G, therefore in the categorical language one has the remarkable statement that *a symmetry category knows about its anomalies*.

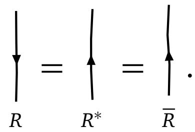
6.2.2 Category of representations

The next important example to discuss is the category of representations of a finite group G. We will denote this category as $\mathbf{Rep}(G)$. Its objects are representations of G, so we can think of



[14]which is again denoted in the respective places in the literature as Vec_G^α

as G-Wilson lines in the representation R, $R \oplus S$, and $R \otimes S$ respectively. The dual line is a line in the complex conjugate representation:



(6.2.4)

The simple objects of this category are the irreducible representations of G, while the morphisms, $\mathrm{Hom}(R,S)$, are the intertwiners. In this category we can choose the associators to be trivial, meaning that $R \otimes (S \otimes T) \cong (R \otimes S) \otimes T$ for all representations. Finally the dimensions of the representations equal their group theoretic dimensions, therefore the total dimension of the symmetry category is

$$\dim \mathbf{Rep}(G) = \sum_{R \text{ irreps}} (\dim R)^2 = |G|. \tag{6.2.5}$$

Later we will see how gauging a theory with a $\mathbf{Sym}(G)$ symmetry yields a $\mathbf{Rep}(G)$ symmetry and vice versa.

As we have seen before (for the first time in subsection 4.1.2), when G is an abelian group, $\mathbf{Rep}(G)$ forms itself a group, therefore we can identify $\mathbf{Rep}(G) = \mathbf{Sym}(\hat{G})$. Therefore, for distinct groups $G \neq H$ we get that $\mathbf{Rep}(G) \neq \mathbf{Rep}(H)$ which we can also infer from the fact that their fusion rings, are distinct $R(G) \neq R(H)$. However for non-abelian groups we might have different groups that yield the same fusion ring. For example, if we take the dihedral group with eight elements: D_8 and the group of eight quaternions, $\{\pm 1, \pm i, \pm j, \pm k\}$, Q_8:

$$\begin{aligned} D_8 &:= \left\langle x, y \,\middle|\, x^4 = y^2 = (xy)^2 = 1 \right\rangle \\ Q_8 &:= \left\langle x, y \,\middle|\, x^2 = y^2 = (xy)^2, x^4 = 1 \right\rangle, \end{aligned}$$

it is known that they both have four one-dimensional irreducible representations

$$1, \quad a, \quad b, \quad a \otimes b =: ab$$

and one two-dimensional irreducible representation, m, for which

$$m \otimes m = 1 \oplus a \oplus b \oplus ab.$$

Therefore the two symmetry categories have the same fusion ring $R(D_8) = R(Q_8)$. However, as symmetry categories $\mathbf{Rep}(D_8)$ and $\mathbf{Rep}(Q_8)$ they are different, because they differ in the associator of the two-dimensional representation $\alpha_{m,m,m}$.

6.2.3 Symmetry of rational CFTs

A much more physical example has to do with the symmetry of rational conformal field theories. Of course, this example is well-known since the work of Moore and Seiberg [52], but it is nice to understand it from the point of view of symmetry categories, following also [54].

A 2d CFT, enjoys a Virasoro symmetry. More precisely, in terms of the symmetry algebra it has a holomorphic and an anti-holomorphic Virasoro algebra, Vir and $\overline{\text{Vir}}$ respectively, with generators L_n and $\overline{L}_n$ respectively, obeying

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{c}{12}n(n^2-1)\delta_{n+m,0} \tag{6.2.6}$$

$$\left[\overline{L}_n, \overline{L}_m\right] = (n-m)\overline{L}_{n+m} + \frac{\overline{c}}{12}n(n^2-1)\delta_{n+m,0} \tag{6.2.7}$$

$$\left[L_n, \overline{L}_m\right] = 0. \tag{6.2.8}$$

$$\tag{6.2.9}$$

In the language of CFTs, an operator a is topological if

$$[L_n, a] = \left[\overline{L}_n, a\right] = 0, \quad \text{for all } n \in \mathbb{Z}. \tag{6.2.10}$$

If it obeys the weaker condition that

$$\left[L_n - \overline{L}_{-n}, a\right] = 0, \quad \text{for all } n \in \mathbb{Z}, \tag{6.2.11}$$

the operator is called conformal. All topological operators are conformal, but not vice versa.

In a CFT, the operators close an algebra, the operator product expansion (OPE). Since the OPE of two holomorphic operators is holomorphic, the holomorphic operators of the theory close a subalgebra, which is called the chiral algebra $\mathfrak{a}$. Similarly, the anti-holomorphic operators give rise to the anti-chiral algebra $\overline{\mathfrak{a}}$. Moreover, since every CFT contains at least two holomorphic operators; the unit $\mathbf{1}$ and the stress tensor $T(z)$, whose modes are L_n, the Virasoro algebra is contained in the chiral algebra, Vir $\subset \mathfrak{a}$. Similarly, $\overline{\text{Vir}} \subset \overline{\mathfrak{a}}$

Now, we suppose that $\mathfrak{a}$ has a finite number of unitary representations V_i (and similarly $\overline{\mathfrak{a}}$, with $\overline{V}_i$), in which case $\mathfrak{a}$ and $\overline{\mathfrak{a}}$ are known as rational chiral algebras. Then we can construct a rational CFT by taking a diagonal combination of $\mathfrak{a}$ and $\overline{\mathfrak{a}}$. Namely, the Hilbert space of the resulting theory is

$$\mathcal{H}_{\text{RCFT}} = \bigoplus_i V_i \otimes \overline{V}_i. \tag{6.2.12}$$

Using a chiral vertex operator a corresponding to a representation of $\mathfrak{a}$ we can construct a topological line operator of the RCFT [52]. However, since the RCFT is the diagonal combination of $\mathfrak{a}$ and $\overline{\mathfrak{a}}$, chiral and anti-chiral operators give rise to the same topological operators. The fusion algebra they satisfy is, of course, governed by the Moore–Seiberg data, forming a unitary modular tensor category. This can be thought of as the fusion symmetry category of the RCFT, by forgetting the braiding. In particular, the irreducible representations of $\mathfrak{a}$ obey the Verlinde algebra

$$V_i \otimes V_j = \bigoplus_k N_{ij}^k V_k. \tag{6.2.13}$$

Since the topological line operators are determined exactly by the irreducible representations, we will be careless with distinguishing between the algebra of the representations and the fusion algebra of the topological operators. The resulting symmetry category will be very similar to the representation categories discussed above.

As a concrete example, consider an $\mathrm{SU}(2)_k$ Wess–Zumino–Witten model. Its chiral algebra is $\widehat{\mathfrak{su}}(2)_k$ and has $k+1$ irreducible representations

$$V_i, \qquad i \in \{0, 1/2, \cdots, k/2\}. \tag{6.2.14}$$

The fusion rule for two representations is

$$V_i \otimes V_j = V_{|i-j|} \oplus V_{|i-j|+1} \oplus \cdots \oplus V_m, \tag{6.2.15}$$

where $m = \min\{i+j, k-(i+j)\}$. We will denote the symmetry category based on $\widehat{\mathfrak{su}}(2)_k$ as $\mathbf{Rep}(\widehat{\mathfrak{su}}(2)_k)$.

SU(2)$_1$: The $\mathrm{SU}(2)_1$ model has two irreducible representations and the corresponding topological operators are

$$V_0 = \mathbf{1}, \quad V_{1/2} =: x, \tag{6.2.16}$$

with fusion rule

$$x \otimes x = \mathbf{1}. \tag{6.2.17}$$

So this is a grouplike symmetry, i.e. $\mathbf{Rep}(\widehat{\mathfrak{su}}(2)_k) = \mathbf{Sym}(\mathbb{Z}_2, \alpha)$, where $\alpha \in \mathrm{H}^3\big(\mathbb{Z}_2, \mathrm{U}(1)\big) = \mathbb{Z}_2$. In fact, α is determined in terms of the quantum 6j symbol with four x's. Here it can be found to be non-trivial. Hence we find that

$$\mathbf{Rep}(\widehat{\mathfrak{su}}(2)_k) = \mathbf{Sym}(\mathbb{Z}_2, \alpha), \tag{6.2.18}$$

is just a $\mathbb{Z}_2$ group with a non-trivial anomaly.

SU(2)$_2$: Here, there are three topological operators, $V_0 = \mathbf{1}$, $V_{1/2} =: \sigma$, $V_1 =: x$. The fusion algebra is

$$x \otimes x = \mathbf{1} \tag{6.2.19}$$

$$\sigma \otimes \sigma = x \oplus \mathbf{1}, \tag{6.2.20}$$

so this is the simplest non-grouplike symmetry in a physical model. We find that it also contains a $\mathbb{Z}_2$ subgroup, generated by x and $\mathbf{1}$, for which the anomaly is trivial this time.

SU(2)$_3$: Finally the SU(2)$_3$ model has four topological operators, $V_0 = \mathbf{1}$, $V_{1/2} =: a$, $V_1 =: b$ $V_{3/2} =: x$, with fusion rules

$$x \otimes x = \mathbf{1} \tag{6.2.21}$$

$$a \otimes x = x \otimes a = b \tag{6.2.22}$$

$$b \otimes x = x \otimes b = a \tag{6.2.23}$$

$$a \otimes a = \mathbf{1} \oplus b \tag{6.2.24}$$

$$b \otimes b = \mathbf{1} \oplus b. \tag{6.2.25}$$

Here we see that except for a $\mathbf{Sym}(\mathbb{Z}_2, \alpha)$ subgroup symmetry generated by $\mathbf{1}$ and x we also have a subcategory symmetry generated by $\mathbf{1}$ and b. The latter goes sometimes by the name Fibonacci category in the literature.

6.3 Categorical symmetry in general dimensions

To recapitulate the main idea of the previous section, Bhardwaj and Tachikawa [10] formalised the properties of topological line operators two dimensions and showed the topological moves are precisely the string diagrams of a unitary fusion category. In higher dimensions, string diagrams are not enough. A codimension-k topological operator should correspond, if anything, to a k-brane diagram[15].

Taking the brane diagrams seriously, we can identify a zero-form symmetry in d dimensions, by a *unitary fusion* $(d-1)$*-category*[16]. For three dimensions a similar construction has been done in the context of defect TQFTs [55]. There the authors needed that the theory be strictly topological because they were interested in defining and calculating the partition function as the functor $\mathcal{Z} : \mathbf{Bord}_n^{\mathrm{def}}(\mathbb{D}) \longrightarrow \mathbf{Vec}$). However, for our purposes, we just need that the symmetry defects be topological, so their results apply to a generic theory, as far as the symmetry defects are of concern. In this setup, a zero-form symmetry category in d dimensions contains the following data:

[15] A k-brane in our conventions extends in k dimensions, not in $k+1$ which is the case in string theory.

[16] A technical remark is that whenever we speak about higher categories in this thesis we refer to *weak* higher-categories

Objects: The objects are the codimension-1 topological operators of the theory.

1-morphisms: The morphisms between objects, are again topological operators that change an object to a different one. The difference is that this time they are supported on the boundaries of the objects, so they are codimension-2 topological operators.

2-morphisms: Here it is where the situation changes from the two-dimensional case. Because in two-dimensions the morphisms where local operators so there was nothing to be changed to. However, in higher dimensions, codimension-2 operators can change to other codimension-2 operators with the help of codimension-3 operators. In other words, there should be 2-morphisms that are the codimension-3 topological operators.

$\vdots$

k-morphisms: The pattern continues and we identify k-morphisms with topological codimension-$(k+1)$ operators.

$\vdots$

$(d-1)$-morphisms: The last type of operators are the local = codimension-$(d-1+1)$ operators. These are, of course, the $(d-1)$-morphisms.

Rest of the data: the rest of the data can be inferred by the two-dimensional case. Namely, the category symmetry should have additivity ($\oplus$), tensor structure ($\otimes$), associators, dualisability of objects, unitary structure and should be finite.

Condensation completion: The last datum that is necessary to define a Symmetry category is the notion of condensation completion [40]. This is a higher-categorical analogue of the Karoubi completion for 1-categories, and it is necessary, in order to ensure that all objects have duals.

We denote the symmetry $(d-1)$-category as $\mathbf{Sym}_{d-1}$. An important remark is that it differs from the usual notion of zero-form symmetries in that it contains not only codimension-1 topological operators but also topological operators of higher codimension, so long as they are supported themselves on topological operators of one-lower-codimension. From this general consideration, we can recover familiar examples of symmetry as special cases. Groups are the symmetry category for which all simple objects are invertible and all $(k \geqslant 2)$-morphisms are trivial. Groups with anomalies are again a symmetry category with all simple objects being invertible and all $(k \geqslant 2)$-morphisms trivial, for which, however, the associators are non-trivial.

Another important class of symmetries that we identify are 2-groups. 2-groups can be identified as a symmetry $(d-1)$-category for which all simple objects are invertible, but there are non-trivial 1-morphisms. For example, in a 2-group, instead of demanding that $g \otimes g^{-1} = \mathbf{1}$, there is a 1-isomorphism $\iota_{\mathrm{R}} : g \otimes g^{-1} \xrightarrow{\sim} \mathbf{1}$. Similarly, $\iota_{\mathrm{L}} : g^{-1} \otimes g \xrightarrow{\sim} \mathbf{1}$, $\ell_g : \mathbf{1} \otimes g \xrightarrow{\sim} g$ and $r_g : g \otimes \mathbf{1} \xrightarrow{\sim} g$. Finally also in the case of a non-anomalous 2-group, the associativity is that $\alpha_{g,h,k} : g \otimes (h \otimes k) \xrightarrow{\sim} (g \otimes h) \otimes k$, while the case of an anomalous 2-group, i.e. with non-trivial associators already built-in is more involved and is discussed in [6]. Similarly, we can extend to 3-groups, by weakening the relations between 1-morphisms to isomorphism, implemented by 2-morphisms, and so on.

Note that albeit fusion $(d-1)$-categories include topological operators of higher codimension, these are not enough to describe higher-form symmetries. To demonstrate this, notice that the natural guess in terms of a fusion $(d-1)$-category for a one-form symmetry would be to take the symmetry category to only have a single object; the object $\mathbf{1}$. This is the transparent defect. On it and upon tensoring many such transparent defects one could place codimension-2 operators/1-morphisms

$$U_i(X_{d-2}) : \mathbf{1}_{d-1} \longrightarrow \mathbf{1}_{d-1}, \tag{6.3.1}$$

and use these to recover the group algebra or the fusion algebra of the one-form symmetry. However, one of the restrictions of fusion categories is the simplicity of the identity. This imposes that any 1-morphism in $\mathrm{Hom}(\mathbf{1}_{d-1}, \mathbf{1}_{d-1})$ is proportional to the identify morphism. One way out of this problem is to drop the demand for simplicity of the identity. This turns the symmetry category into a braided fusion category.

The aforementioned process is closely related to the notions of looping and delooping in the category theory literature. In particular, given a fusion n-category, $\mathbf{C}_n$, one can construct a braided fusion $(n-1)$-category,

$$\Omega \mathbf{C}_n := \mathrm{Hom}(\mathbf{1}_{n-1}, \mathbf{1}_{n-1}), \tag{6.3.2}$$

where $\mathbf{1}_{n-1}$ is the codimension-1 trivial object in $\mathbf{C}_n$. $\Omega \mathbf{C}_n$ is called the looping of $\mathbf{C}_n$. More generally, as it will come in handy later on, given a k-fusion n-category $\mathbf{C}_n^{[k]}$, its looping

$$(\Omega \mathbf{C})_{n-1}^{[k+1]} := \Omega \mathbf{C}_n^{[k]} = \mathrm{Hom}(\mathbf{1}_{n-1}, \mathbf{1}_{n-1}) \tag{6.3.3}$$

is a $(k+1)$-fusion $(n-1)$-category. Conversely, given a k-fusion n-category $\mathbf{D}_n^{[k]}$, one can define the one-point delooping, as the $(n+1)$-category $\Sigma_* \mathbf{D}_n^{[k]}$ with only one object, $*$ and with 1-morphisms given by the objects of $\mathbf{D}_n^{[k]}$, 2-morphisms given by the 1-morphisms of $\mathbf{D}_n^{[k]}$ and so on. This is a $(k-1)$-fusion $(n-1)$-category.

The notions of looping and delooping give us a way to identify what the appropriate notion of higher-form symmetries should be. In particular taking a fusion zero-form symmetry category $\mathbf{Sym}_{d-1} \equiv \mathbf{Sym}^{[1]}_{d-1}$ and applying a looping, we essentially project to the purely one-form part of it, which makes it a braided fusion $(d-2)$-category,

$$(\Omega\mathbf{Sym})^{[2]}_{d-2} = \Omega\mathbf{Sym}^{[1]}_{d-1}. \tag{6.3.4}$$

Furthermore, given a braided fusion $(d-2)$-category $\mathbf{C}^{[2]}_{d-2}$, we can deloop around an object $\mathbf{1}_{d-1}$, to make a fusion $(d-1)$-category, or in other words, a symmetry category:

$$\Sigma_{\mathbf{1}_{d-1}}\mathbf{C}^{[2]}_{d-2} = \left(\Sigma_{\mathbf{1}_{d-1}}\mathbf{C}\right)^{[1]}_{d-1} = \widetilde{\mathbf{Sym}}_{d-1}, \tag{6.3.5}$$

describing the embedding of a one-form symmetry in a zero-form symmetry category. Therefore we can say that *a one-form categorical symmetry in d dimensions is given by a braided fusion $(d-2)$-category*, $\mathbf{Sym}^{[2]}_{d-2}$.

Taking it one step further for a two-form symmetry, we can start from a one-form symmetry and loop once to land on a 3-fusion $(d-3)$-category. We can see that all 3-fusion $(d-3)$-categories can deloop twice to fusion $(d-1)$-categories. Hence we identify *a two-form symmetry in d dimensions with a 3-fusion $(d-3)$-category* $\mathbf{Sym}^{[3]}_{d-3}$. Continuing the pattern we can give a general definition of higher-form categorical symmetries

A d-dimensional quantum field theory can have some p-form symmetry, implemented by a $(p+1)$-fusion $(d-p-1)$-category $\mathbf{Sym}^{[p+1]}_{d-p-1}$.

A note on terminology is in place. A 1-fusion category is simply called fusion. A 2-fusion category is called braided fusion. A 3-fusion category goes sometimes by the name syllep-tic fusion, while any higher-fusion category does not have a specific name.

A different way to wrap one's head around the above definition of symmetries is to first think about a one-form symmetry in three dimensions. Usually, it is said that such higher-form symmetries are abelian. Indeed, locally, this is the case, as shown in chapter 2. However, globally the symmetry operators might come linked as shown in figure 6.1. If we think of the symmetry operators configurations as string diagrams, these are precisely the string diagrams for a *braided fusion category*. Similarly, for a one-form symmetry in four dimensions, its symmetry operators are surfaces (2-surfaces). Two and three surfaces can link in four dimensions [56, 57] according to figure 6.2. Having surfaces makes the symmetry into a fusion 2-category, and this extra linking gives a braiding structure, making it a *braided fusion 2-category* in total.

Another sanity check is to look at the periodic table of k-fusion n-categories [58–60] (table 6.1). Wee see several things there. First of all, we see that any $(p+1)$-fusion (-1)-category

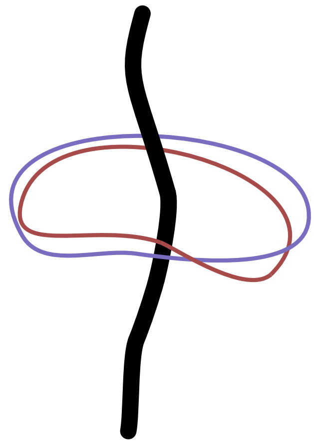


Figure 6.1: one-form symmetry is *almost* abelian in three dimensions

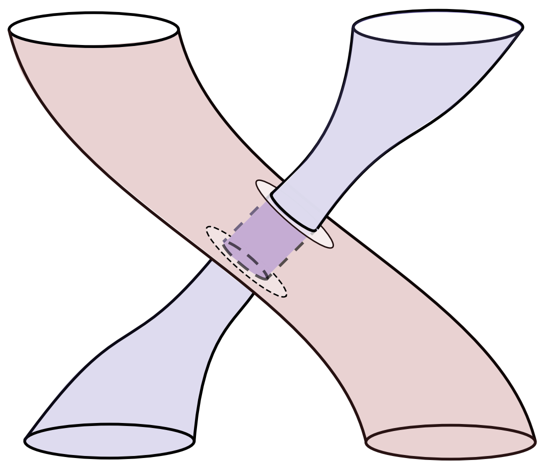


Figure 6.2: Surfaces braiding in four dimensions. The figure should be seen as the trace that two loops leave behind as they pass through each-other

is trivial, which implies that a d-form symmetry is trivial in d dimensions. Secondly, a (-1)-form symmetry in d dimensions is simply a d-category without any extra structure, as expected (cf. subsection 2.2.1). Finally, due to the stabilisation hypothesis [58], we get that for $k \geqslant n+2$, the symmetry category is maximally abelian. In other words for sufficiently large p, p-form symmetries are trully abelian, as we had expected. For low $p \geqslant 1$, the abelianity is hindered by global features.

Very recently, Johnson–Freyd [61] axiomatised the extended operators in topological orders and found that in a d-dimensional topological order[17], the codimension-($\geqslant p+1$) lie in a $(p+1)$-fusion $(d-p-1)$-category. Considering that the low-energy theory that corre-

[17] Here by d-dimensional we mean d space-time dimensions, while in [61] the conventions are such that nd means $(n+1)$ space-time dimensions. Therefore the result on [61] differs by what we mention here, but they can be mapped to one another if we make the correct substitutions of the dimensions

$(p+1) = k$ \ $(d-p-1) = n$	-1	0	1	2	3	$\cdots$
0	truth value	set	category	2-category	3-category	$\cdots$
1	trivial	monoid	fusion category	fusion 2-category	fusion 3-category	$\cdots$
2	trivial	abelian monoid	braided fusion category	braided fusion 2-category	braided fusion 3-category	$\cdots$
3	trivial	abelian monoid	abelian fusion category	sylleptic fusion 2-category	sylleptic fusion 3-category	$\cdots$
4	trivial	abelian monoid	abelian fusion category	abelian fusion 2-category	4-fusion 3-category	$\cdots$
5	trivial	abelian monoid	abelian fusion category	abelian fusion 2-category	abelian fusion 3-category	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Table 6.1: Periodic table of k-fusion n-categories

sponds to a topological order is a TQFT, the low-energy TQFT has precisely the categorical symmetry that we found above.

Another recent exploration in topological order [13], revealed that a topological order can have a so-called algebraic higher symmetry, which is given by a local fusion higher-category. Moreover, the authors of [13] took a bulk point-of-view for the topological orders they considered, thus essentially considering SPT phases for anomalous categorical symmetries. The implications of both these results for field theory and the way they can be incorporated in our classification are fascinating and certainly a theme that we will pursue in future research.

6.4 Examples of higher-dimensional categorical symmetries

Since we have already discussed how to write a group symmetry and a higher-group symmetry as a categorical symmetry, by mimicking the two-dimensional case and using the definition of a higher-group, let us dive into examples of the primary theme of interest; non-invertible categorical symmetries.

6.4.1 Non-invertible one-form symmetry in three dimensions

In three dimensions, a one-form symmetry is implemented by topological line operators. In a TQFT these lines are the worldlines of anyons. Invertible lines represent abelian anyons, while the non-invertible ones signify non-abelian anyons. They are famously described by modular tensor categories (MTC) [52]. Similarly to the discussion in two dimensions, forgetting the extra structure of the MTC and demanding finiteness, we regard these lines as given by a braided fusion category. Note that contrary to the two-dimensional case, we now need the braiding structure.

Two lines, a and b can fuse according to the fusion rule

$$a \otimes b = \bigoplus_c N_{ab}^c c. \tag{6.4.1}$$

The fusion is rule is abelian, i.e. $N_{ab}^c = N_{ba}^c$, which reflects the abelian nature of higher-form symmetries. Two lines can also braid, and this defines an action of the topological lines as follows. We surround the line a with a loop of b and since b is topological, we shrink b (see figure 6.3). Doing this, we get back a, up to a braiding factor, B_{ba}. Notice that B_{ba} is not the same as B_{ab}, i.e. the braiding action is not abelian. This refines the abelian nature to "*almost* abelian". Note also that taking $b = \mathbf{1}$ to be the trivial defect, renders $B_{\mathbf{1}a} = 1$. Taking a to be trivial renders $B_{b\mathbf{1}} = \dim(b)$. In a TQFT, the fusion coefficients N_{ab}^c can be expressed in terms of the modular S-matrix, through the Verlinde formula [53] and similarly for the braiding coefficients B_{ab}:

$$N_{ab}^c = \sum_d \frac{S_{ad} S_{bd} S_{dc}^*}{S_{\mathbf{1}d}^*}, \qquad B_{ab} = \frac{S_{ab}^*}{S_{\mathbf{1}b}^*}. \tag{6.4.2}$$

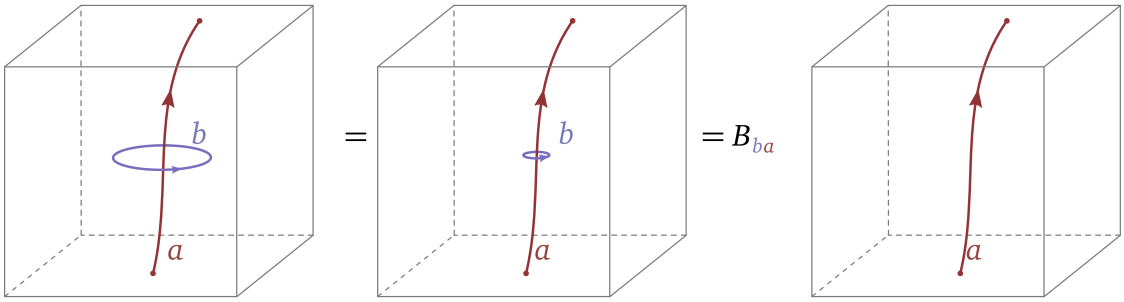


Figure 6.3: Action a topological line on another through braiding

Finite gauge theories

A way to construct such non-invertible symmetries is by using non-abelian finite gauge theory. The discussion in this passage follows [12]. The one-form group symmetry of the gauge theory is, in general, a subcategory of the full one-form symmetry category. In particular, the invertible operators will appear in the fusion channel of other, possibly non-invertible operators.

To see this, consider a gauge theory of a finite group, G. A class of topological operators that we can construct are the Wilson lines. They are labelled by a representation of G, $\mathcal{W}_\rho$. However, the fusion rule of two Wilson lines is governed by the representation ring of G, which, as discussed previously, is not a group if G is non-abelian. Wilson lines labelled by a one-dimensional representation are invertible, and they are responsible for the magnetic one-form symmetry of the gauge theory. Besides, there are Gukov–Witten operators[18]. They are labelled by a conjugacy class $[g]$ of G, $\mathcal{V}_{[g]}$. In general, we do not expect them to be invertible. The invertible ones are responsible for the electric one-form symmetry. Finally, since in three dimensions both Wilson and Gukov–Witten operators are lines, they can be fused to create an operator labelled both by a conjugacy class and by a representation of the stabiliser subgroup Stab$[g]$, of an element in the conjugacy class $[g]$. A generic topological line operator will be denoted as $[g]_\rho$, where

$$[1]_\rho = \mathcal{W}_\rho \tag{6.4.3}$$

$$[g]_{\mathbf{1}} = \mathcal{V}_{[g]}, \tag{6.4.4}$$

where $[1]$ denotes the conjugacy class of the identity element, $\mathbf{1}$ is the trivial representation and $\rho \in \mathrm{Rep}(\mathrm{Stab}[g])$.

[18]These are like the standard Gukov–Witten surfaces in four dimensions [62], but one dimension lower, thus lines instead of surfaces.

S_3 gauge theory

As an example, we will construct the one-form symmetry category of the simplest finite non-abelian gauge theory; the S_3 gauge theory. In the symmetric group, S_3, there are three conjugacy classes, $[(1)] =: [1]$, $[(12)] =: [2]$, and $[(123)] =: [3]$, corresponding to cycles of size 1, 2 and 3 respectively. A note on the notation of S_3 is that the an element (abc) sends a to b, b to c and c to a, while an element (ab) sends a to b and leaves the third number intact. Finally (1) is the identity element.

The stabiliser subgroups are as follows. Of course $\text{Stab}[1] = S_3$, $\text{Stab}[2] = \mathbb{Z}_2$, since (ab) is stabilised by the identity and (ab). Finally, $\text{Stab}[3] = \mathbb{Z}_3$, since (abc) is stabilised by the identity, (abc) and (acb). The Wilson lines are

$$[1]_{\mathbf{1}} =: \mathbf{1} \text{ (the trivial object)}, \qquad [1]_{\mathbf{sgn}} =: \mathcal{W}_{\mathbf{sgn}}, \quad \text{and} \quad [1]_{\mathbf{2}} = \mathcal{W}_{\mathbf{2}}, \tag{6.4.5}$$

corresponding to the trivial, the sign and the two-dimensional representation of S_3 respectively. For the more generic lines, we have

$$[2]_{\mathbf{1}} =: \tau_+, \quad \text{and} \quad [2]_{\mathbf{sgn}} =: \tau_-, \tag{6.4.6}$$

corresponding to the trivial and the sign representation respectively, of $\mathbb{Z}_2 = \text{Stab}[2]$. Finally we also have

$$[3]_a =: \theta_a, \qquad a \in \{1,2,3\}, \tag{6.4.7}$$

for the three irreducible representations of $\mathbb{Z}_3 = \text{Stab}[3]$.

The modular S-matrix for S_3 is given by

$$S = \frac{1}{6}\left(\begin{array}{c|cccccccc} & \mathbf{1} & \mathcal{W}_{\mathbf{sgn}} & \mathcal{W}_{\mathbf{2}} & \tau_+ & \tau_- & \theta_1 & \theta_2 & \theta_3 \\ \hline \mathbf{1} & 1 & 1 & 2 & 3 & 3 & 2 & 2 & 2 \\ \mathcal{W}_{\mathbf{sgn}} & 1 & 1 & 2 & -3 & -3 & 2 & 2 & 2 \\ \mathcal{W}_{\mathbf{2}} & 2 & 2 & 4 & 0 & 0 & -2 & -2 & -2 \\ \tau_+ & 3 & -3 & 0 & 3 & -3 & 0 & 0 & 0 \\ \tau_- & 3 & -3 & 0 & -3 & 3 & 0 & 0 & 0 \\ \theta_1 & 2 & 2 & -2 & 0 & 0 & 4 & -2 & -2 \\ \theta_2 & 2 & 2 & -2 & 0 & 0 & -2 & -2 & 4 \\ \theta_3 & 2 & 2 & -2 & 0 & 0 & -2 & 4 & -2 \end{array}\right). \tag{6.4.8}$$

From here we can read of the fusion and the braiding coefficients through (6.4.2). The various topological lines fuse as follows:

$$\begin{aligned}
\mathcal{W}_{\mathbf{sgn}} \otimes \mathcal{W}_{\mathbf{sgn}} &= \mathbf{1}, \\
\mathcal{W}_{\mathbf{sgn}} \otimes \mathcal{W}_{\mathbf{2}} &= \mathcal{W}_{\mathbf{2}}, \\
\mathcal{W}_{\mathbf{2}} \otimes \mathcal{W}_{\mathbf{2}} &= \mathbf{1} \oplus \mathcal{W}_{\mathbf{sgn}} \oplus \mathcal{W}_{\mathbf{2}}, \\
\mathcal{W}_{\mathbf{sgn}} \otimes \tau_{\pm} &= \tau_{\mp}, \\
\tau_{+} \otimes \tau_{-} &= \mathcal{W}_{\mathbf{sgn}} \oplus \mathcal{W}_{\mathbf{2}} \oplus \bigoplus_{a} \theta_a \\
\tau_{\pm} \otimes \tau_{\pm} &= \mathbf{1} \oplus \mathcal{W}_{\mathbf{2}} \oplus \bigoplus_{a} \theta_a, \\
\mathcal{W}_{\mathbf{2}} \otimes \tau_{\pm} = \theta_a \otimes \tau_{\pm} &= \tau_{+} \oplus \tau_{-}, \\
\theta_a \otimes \mathcal{W}_{\mathbf{sgn}} &= \theta_a, \\
\theta_a \otimes \theta_a &= \mathbf{1} \oplus \mathcal{W}_{\mathbf{sgn}} \oplus \theta_a, \\
\theta_a \otimes \theta_{b \neq a} &= \mathcal{W}_{\mathbf{2}} \oplus \theta_{c \neq a,b}, \\
\theta_a \otimes \mathcal{W}_{\mathbf{2}} &= \bigoplus_{b \neq a} \theta_b.
\end{aligned} \tag{6.4.9}$$

We see that $\{\mathbf{1}, \mathcal{W}_{\mathbf{sgn}}\}$ generate a subgroup, which is precisely the $\mathbb{Z}_2$ magnetic one-form symmetry. The braiding structure can also be read off from the S-matrix. For example, $B_{\theta_1, \mathcal{W}_{\mathrm{sgn}}} = 2$ and $B_{\mathcal{W}_2, \theta_1} = -1$.

6.4.2 A non-invertible 2-category

A way to produce a theory with a non-invertible two category follows from an extension of the discussion of section 4.5 and [43]. There we saw that in two dimensions, gauging a non-anomalous subgroup, $\mathcal{A}$ of a symmetry $\mathcal{A} \times G$ with a mixed anomaly in $\mathrm{H}^{3-k}\big(G, \mathrm{H}^k(\mathcal{A}, \mathrm{U}(1))\big)$, for $k \geqslant 2$, results in a symmetry given by a fusion category, similar to those discussed in section 6.1. Let us now perform the analogous computation in three dimensions.

Start with a theory with a symmetry $\mathcal{A} \times G$, where $\mathcal{A}$ is an abelian group, with an anomaly determined by $\xi \in \mathrm{H}^2\big(G, \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1))\big)$ through the SPT action

$$\mathrm{SPT}[A, B] = \int_{Y_4} A \smile A \smile \xi(B). \tag{6.4.10}$$

A is a background field for $\mathcal{A}$ and B a background field for G. The element $\xi(B)$ characterises the anomaly as follows: Placing a theory on a surface $\Sigma_{d-1} = \Sigma_2$, with a G flux given by B turned on and dimensionally reduce to a one-dimensional theory that we take as a theory with symmetry $\mathcal{A}$. Then the anomaly of the $\mathcal{A}$-symmetric one dimensional theory is given by

$$\int_{\Sigma_2} \xi(B) \in \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1)). \tag{6.4.11}$$

This viewpoint makes evident that we should see $\xi(B)$ as a one-dimensional locus, where $(d-1) = 2$ G-defects, g and g' fuse to create a defect labelled by $g'' = gg'$, thus determining an element

$$\xi(g, g') \in \mathrm{H}^2(\mathcal{A}, \mathrm{U}(1)) \tag{6.4.12}$$

Here we demand that this one-dimensional locus carries a projective representation of $\mathcal{A}$ whose projective phase is $\xi(g, g')$. The set of topological operators that appear in the theory are surfaces implementing $\mathcal{A}$ together with their projective representations, manifesting themselves as one-dimensional 1-morphisms on the $\mathcal{A}$-surfaces. In total this defines the data for a fusion 2-category [55]. Since projective representations of abelian groups are in general not one-dimensional this is a non-invertible 2-category, describing a zero-form symmetry in three dimensions.

7 Discussion and outlook

7.1 Summary

Symmetries are one of the few tools that we have at our disposal to study universal features of quantum field theories and their dynamics. It is natural to try to push the boundaries of what we can think of as symmetry, in order to gain more handles to understand quantum field theory universally. The modern understanding of ordinary global symmetries is that they are described by topological operators, supported on a codimension-1 manifold. These operators implement the symmetry action by surrounding the local operators of the theory. This description of symmetries has led to the emergence and understanding of higher-form symmetries [16, 17] as a new, very powerful organising principle. These correspond to topological operators of higher codimension. It is standard by now, that higher-form symmetries are as important as ordinary symmetries and therefore we should immediately look out for them whenever we are assessing the symmetries of a theory. This is also the starting point of this, thesis, reviewed in chapter 2. A different generalisation of symmetries relies on the existence of non-invertible topological operators. In particular, it is known for some time that there are topological operators that are not necessarily invertible, for example, the topological lines of RCFTs [52], the Kramers-Wannier duality defects in the Ising CFT [63], etc. Mathematically it was understood in [10] that in two dimensions not-necessarily-invertible topological operators form a unitary fusion category. One aspect of this thesis (chapter 6) dealt with understanding the mathematical structure of not-necessarily-invertible operators in all dimensions. We found that these correspond in fact to unitary fusion $(d-1)$-categories in a d-dimensional quantum field theory. Moreover, we examined not-necessarily-invertible topological operators of higher codimension — namely higher-form non-invertible symmetries — and found that their description necessitates the existence of braiding (and higher braiding) morphisms. In particular we proposed that a p-form symmetry in d-dimensions is described by a unitary $(p+1)$-fusion $(d-p-1)$-category.

A different universally applicable tool for quantum field theories is 't Hooft anomalies, i.e. background gauge non-invariance. The importance of 't Hooft anomalies lies in constraining the renormalisation group flows. In particular, a theory with a symmetry afflicted with an anomaly cannot have a trivially gapped vacuum. The introduction of higher-form symmetries has served as the start-point for a plethora of new 't Hooft anomalies, leading in turn to new constraints. A notable example is the remarkable work of Gaiotto, Kapustin, Komargodski, and Seiberg [49], in which the low-energy phases of four-dimensional Yang–Mills theory at were severely constrained due to the existence of a mixed anomaly between the $\mathbb{Z}_N^{[N]}$ one-form centre symmetry and time-reversal at $\theta = \pi$. More recently, there has been an interest in anomalies of duality groups [45] as well as in anomalies in the space of couplings [46, 47]. These two are highly correlated since in both cases the background gauge fields are zero-form gauge fields. In fact, we can describe both, in a somewhat notation-abusing way, as anomalies of (-1)-form symmetries. In this thesis, after reviewing anomalies of ordinary and higher-form symmetries in chapter 3, we discuss duality anomalies and identify a new mixed duality anomaly in quantum mechanics in chapter 5. The anomaly we identify concerns a particle on a ring of radius R, at inverse temperature β. This theory has the duality $\beta/R^2 \longmapsto R^2/\beta$, which should be thought of as a mixture between T-duality and modular invariance. We proved, however, that this duality is anomalous and moreover that the anomaly is mixed, involving the conformal symmetry of the model.

Central to this work has been the idea of gauging and ungauging finite symmetries, that we review in chapter 4. Traditionally the idea goes back to Vafa [42], where in two dimensions gauging an abelian finite symmetry group G results in the gauged theory having a Pontryagin dual global symmetry $\hat{G}$, dubbed quantum symmetry. In higher dimensions, such quantum symmetries were found only after the introduction of higher-form symmetries, since in d dimensions the global symmetry of a theory with a gauged symmetry G is a $(d-2)$-form symmetry $\hat{G}^{[d-2]}$. More important for this thesis is gauging a subgroup of a symmetry, as explained in [43]. There it was found that gauging an abelian subgroup G of a symmetry given by group extension $1 \longrightarrow G \longrightarrow \Gamma \longrightarrow K \longrightarrow 1$ controlled by $\varepsilon \in \mathrm{H}^2(K, G)$, results in a theory with global symmetry $\hat{G}^{[d-2]} \times K$ together with a mixed anomaly controlled by the same ε. Here this result was applied to pin down exactly what the mixed duality anomaly that we found is. Furthermore, the whole idea of gauging a finite symmetry — this time a non-abelian finite symmetry — served as a starting point in [10] and therefore as a starting point for the whole field of categorical symmetries that we explore in chapter 6.

7.2 Outlook

During this thesis, many interesting questions, with possibly non-trivial answers, arose. In this final section, we will lay some of them out, together with possible routes towards them.

Towards discrete Goldstone theorem Very early on in this thesis, when we first introduced higher-form symmetries, we found a modification of Goldstone theorem, valid for higher-form symmetries. Since in the modern understanding, both discrete and continuous symmetries have been put on the same footing, for the most part, it is natural to ask whether there is an equivalent of Goldstone theorem for discrete symmetries.

In [64], Wen studied condensed matter systems with a spontaneously broken finite higher-form symmetry and found that a spontaneous symmetry broken state of a finite higher-form symmetry always corresponds to a topologically ordered state. However, in two dimensions gauging the symmetry of a Symmetry Protected Topological (SPT) order, results in genuine topological order [65]. For non-trivial SPT orders, the resulting topological order has protected gapless edge modes. There is a caveat, however; that not all topological orders correspond to spontaneous higher-form symmetry breaking. For example, they can arise from explicit symmetry breaking, or they can be non-abelian, thus precluding the involvement of higher-form symmetries. The question is whether the topological orders of [65], are precisely the topological orders that correspond to spontaneous higher-form symmetry breaking, implying that it protects the gapless edge modes, thus giving a discrete-Goldstone theorem analogue. This question deserves further research, and we would like to revisit it in future works.

2-groups, higher-groups, and sticky symmetries We saw in chapter 4, that symmetries of the form $1 \longrightarrow G^{[1]} \longrightarrow \Gamma^{[1,0]} \longrightarrow K^{[0]} \longrightarrow 1$ correspond to so-called 2-groups, that have been getting much attention recently [6, 7, 9]. Similarly, symmetries of the form $1 \longrightarrow G^{[p]} \longrightarrow \Gamma^{[p,0]} \longrightarrow K^{[0]} \longrightarrow 1$, could be described as a $(p+1)$-group, that corresponds to a codimension-$(p+1)$ topological operator appearing at the fusion of $p+2$ codimension-1 operators. We expect that this kind of mixing appears naturally in dimensions $d \geqslant p+2$. It would be interesting to see whether this sort of structure has any implications for usual quantum field theories. More generally, a symmetry such as $1 \longrightarrow G^{[p]} \longrightarrow \Gamma^{[p,q]} \longrightarrow K^{[q]} \longrightarrow 1$, would give rise to a codimension-$(p+1)$ topological operator at the fusion of $(p+2)$ codimension-$(q+1)$ topological operators. Unlike the original higher-group case, this should happen less frequently as higher-codimensional operators are less likely to intersect. We, therefore, expect that it will not be as important a structure as higher-groups. It is still interesting to understand the physics of this type of symmetry, in any case. Finally, a special case of the previous scenario that we expect to be more important is the case of $p=0$, and $q=1$, namely $1 \longrightarrow G^{[0]} \longrightarrow \Gamma^{[0,1]} \longrightarrow K^{[1]} \longrightarrow 1$.

This case describes codimension-2 defects intersecting at a point. They generically do not intersect. However, they braid among each other. Hence, a natural way we can get two codimension-2 defects to touch is when they try to unbraid them, pulling them apart. In the case of $\Gamma^{[0,1]}$, when they are pulled apart, they give rise to a sourced codimension-1 defect of G, given by $\varepsilon(g,h)$, as shown in figure 7.1. We refer to this symmetry as sticky symmetry. In $d>2$ dimensions codimension-2 defects can always braid, thus can always

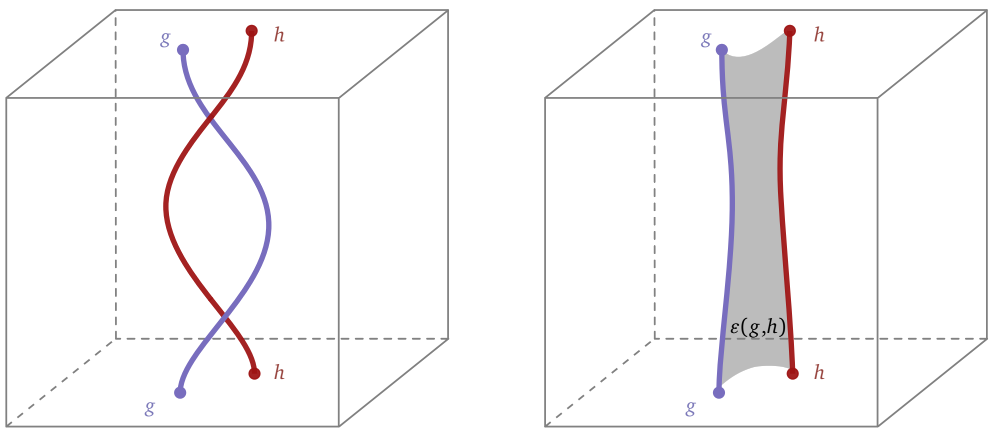


Figure 7.1: Two codimension-2 $K^{[1]}$-defects $(g,\ h)$, untangled, creating a codimension-1 G-defect $(\varepsilon(g,h))$.

give rise to such a sticky symmetry. One example of stickiness can be realised by Chern–

Simons theories [43]. In particular, taking two $\mathrm{U}(1)_{2k}$ Chern–Simons theories, with their $\mathbb{Z}_{2k}^{[1]}$ global symmetry identified, can fit in an extension $1 \longrightarrow \mathbb{Z}_k^{[1]} \longrightarrow \mathbb{Z}_{2k}^{[1]} \longrightarrow \mathbb{Z}_k^{[1]} \longrightarrow 1$, in which $\mathbb{Z}_k^{[1]}$ is non-anomalous and can be gauged. The gauged theory has $1 \longrightarrow \mathbb{Z}_k^{[0]} \longrightarrow \Gamma^{[0,1]} \longrightarrow \mathbb{Z}_k^{[1]} \longrightarrow 1$, with both a non-trivial extension and a mixed anomaly. The full story of sticky symmetries is yet unknown and we would like to understand it in more detail in future works. Moreover it is unclear where it fits in the categorical classification that we gave in chapter 6. Finally, cranking up codimensions gets even trickier, because there is no non-trivial braiding between codimension-($\geqslant 2$) objects. We expect that syllepses take the role of braiding in codimension-2, while higher yet-unkown analogues in higher-codimension.

Duality anomalies and quantum mechanics The story explored in section 5.4, is still not finished. First of all, it still remains to uncover what exactly the relevant background fields are and what are the connections between them. In particular, the $\mathrm{SL}(2,\mathbb{R})$ gauge field can be written similarly as a Levi-Civita connection in one dimension. Namely, we can write $B = \theta^{-1}\mathrm{d}\theta$, where θ transforms under $\mathrm{SL}(2,\mathbb{R})$ like a derivative. In this way, if θ is further equal to $\partial_\tau \varphi$, where φ is a scalar field, invariant under $\mathrm{SL}(2,\mathbb{R})$, then the low-energy dynamics of the system, are controlled by a Schwarzian, in a similar fashion as in [66]. It is natural to wonder, then how does this story relate to $\mathrm{AdS}_2/\mathrm{CFT}_1$ [67, 68] and JT gravity. Finally, it would be interesting to understand whether this or a similar anomaly constrains other systems of quantum mechanics with conformal symmetry as well.

Anomalies of categorical symmetries Something that is discernibly absent from our study of categorical symmetries is an understanding of anomalies and anomaly inflow for categorical symmetries. Although in two dimensions, we argued that anomalies manifest themselves as the associator morphisms in the symmetry category, the connection with categorical SPT phases is still unclear. One attempt towards anomaly inflow in two-dimensional symmetry categories is presented in [11], however, their analysis relies heavily on the construction of three-dimensional TQFTs and hence it is not easy to generalise to higher dimensions. On the other hand, generalising the associator argument of [10], we might postulate that a categorical symmetry $\mathbf{Sym}_{d-p-1}^{[p+1]}$ in d dimensions, with anomaly $\mathcal{A}$, corresponds to a TQFT in $(d+1)$-dimensions, with symmetry $\mathbf{Sym}_{d-p}^{[d-p]}$, whose $(d-p)$-morphisms are given by $\mathcal{A}$. This sentence is far from precise, however, and requires much further analysis. Other notions that appear related are the so-called braided-automorphisms, that appear in the categorical classification of topological orders [37], and the fibre functor that was used heavily in [11]. We would like to explore how these three seemingly unrelated notions connect to each other to give us a full understanding of categorical anomalies. Finally, it is of crucial importance to understand the constraints that these anomalies provide to the RG flows.

Categorical symmetries and the symmetries of quantum gravity Something that seems important is the presence of categorical symmetries in quantum gravity. On the one hand, it is conjectured [69, 70] and proven in the context of AdS/CFT [71] that quantum gravity cannot admit any global symmetries. Based on that, it was recently conjectured than non-invertible symmetries should too be excluded [12]. However, the proof of [71], for group-like symmetries does not go through in the case of non-invertible topological operators. One of the reasons for that is that the splittability requirement, that was crucial for the proof does not hold for non-invertible operators. It might be the case that an extended notion of splittability might be applicable for non-invertible symmetries too, but the existing arguments do not preclude categorical symmetries.

Furthermore, formally speaking, a theory of quantum gravity is a theory where Lorentz symmetry $SO(d,1)$ has been gauged. Modifying the arguments presented in chapters 4 and 6, we should in principle expect that quantum gravity should have a categorical symmetry $\mathbf{Rep}(SO(d,1))$. Of course, many questions arise here. How is this consistent with the "no global symmetries" argument and black hole physics? Does this symmetry exist or break somehow? Can we use this symmetry to understand gravity as an effective theory?

Acknowledgements

First of all, I am very grateful to Diego for his guidance and the many intriguing discussions we had throughout this project, as well as for encouraging me to explore whatever caught my interest and make this thesis all about having fun with physics.

Moreover, I want to thank Vladimir, always welcoming all sorts of questions and having at least one reference that would prove helpful (although sometimes written in Russian). I also wish to thank Tarek, Marcos, and Greg for interesting discussions, as well as for helping out with my PhD applications. Furthermore, I would like to acknowledge helpful email correspondence with Yuji Tachikawa, about a part of this thesis and fruitful discussions with Clay Córdova, during the CERN winter school.

Last, but by any means not least I want to thank my friends in Amsterdam: Archishna, for nice discussions, late night working company in the master-room, and for taking care of my plants whenever I was not in Amsterdam, including these past three-weeks-that-became-five-months; Diego, for being the baking aunt that everyone deserves, bringing us sweets he would bake all the time and for many nice physics discussions; Keivan, for endless discussions about physics, music, and life in general, and for a great friendship altogether.

I am also very grateful to the Alexander S. Onassis foundation for awarding me a scholarship towards my master's studies.

A

Notation

Symbol	Meaning
$\bullet := \circ$	$\bullet$ is defined as $\circ$
$\bullet =: \circ$	$\circ$ is defined as $\bullet$
$\bullet ::= \circ$	$\bullet$ (already defined) is redefined as $\circ$
$\bullet =:: \circ$	$\circ$ (already defined) is redefined as $\bullet$
$\bullet \equiv \circ$	$\bullet$ is equivalent to $\circ$
$\bullet \cong \circ$	$\bullet$ is isomorphic to $\circ$
$G^{[p]}$	p-form symmetry group
$G := G^{[0]}$	short for 0-form symmetry group
$A_{[q]}$	q-form
$\mathrm{CS}[A]$	the Chern-Simons form: A, $\left(A \wedge \mathrm{d}A + \frac{2}{3} A \wedge A \wedge A\right)$, ... depending on the dimensions
$\mathscr{H}^*$	Borel–Moore/Borel (group) cohomology group/ring
H^*	topological/ordinary (group) cohomology group/ring
$\mathcal{H}$	Hilbert space
$\Omega_*^\bullet$	bordism group/ring of manifolds with structure $\bullet$

Symbol	Meaning
$\Omega^*_\bullet$	cobordism group/ring of manifolds with structure $\bullet$
	$\left(\Omega^d_\bullet := \mathrm{Hom}\left(\Omega^\bullet_d, \mathrm{U}(1)\right)\right)$
$\mathcal{Z}$	partition function
$\mho^*_\bullet$	Anderson dual cobordism group/ring of manifolds with structure $\bullet$
C, **Bord**, **Fus**	examples of categories. Categories will be denoted by one or a string of upright boldface letters.

B Mathematical Facts

B.1 Group cohomology and (Co)bordisms

It holds generally that

$$\mathrm{H}^k(\mathbb{Z}, M) \cong \mathrm{H}^k_{\text{singular}}\big(\mathbf{S}^1, M\big) \cong \mathrm{H}^k_{\text{dR}}\big(\mathbf{S}^1, M\big) = \begin{cases} M, & k \in \{0,1\}, \\ 0 & \text{otherwise}, \end{cases} \tag{B.1.1}$$

where M is any $\mathbb{Z}$-module, which allows us to compute the cohomology of the integers. Another important equation is the following:

$$\mathrm{H}^k(\mathrm{B}G, \mathrm{U}(1)) = \mathscr{H}^k(G, \mathrm{U}(1)) = \mathscr{H}^k(G, \mathrm{B}\mathbb{Z}) = \mathscr{H}^{k+1}(G, \mathbb{Z}) = \mathrm{H}^{k+1}(\mathrm{B}G, \mathbb{Z}). \tag{B.1.2}$$

This allows us to reduce cohomology with $\mathrm{U}(1)$ coefficients to cohomology with $\mathbb{Z}$ coefficients, that is generally simpler. There is a generalisation for higher classifying spaces stating that

$$\mathrm{H}^k\big(\mathrm{B}^{p+1}G, \mathrm{U}(1)\big) = \mathrm{H}^{k+1}\big(\mathrm{B}^{p+1}G, \mathbb{Z}\big) = \mathrm{H}^{k+1}(K(G, p+1), \mathbb{Z}), \tag{B.1.3}$$

where $K(G,n)$ is an Eilenberg-McLane space, i.e. a topological group defined by the condition

$$\pi_m\big(K(G,n)\big) = \delta_{m,n} G. \tag{B.1.4}$$

In the main text we made frequent use of (-1)-form symmetry. For this we will need $K(G,0)$. This is defined as the group G itself, with the discrete topology. If we are only interested in finite groups, they can only have the discrete topology, therefore $K(G_{\text{finite}},0) \cong G_{\text{finite}}$. And hence

$$\mathrm{H}^k\big(\mathrm{B}^0 G_{\text{finite}}, \mathrm{U}(1)\big) = H^k(K(G_{\text{finite}},0),\mathbb{Z}) = H^{k+1}(G_{\text{finite}},\mathbb{Z}). \tag{B.1.5}$$

B.2 Spectral sequences

A spectral sequence is a tool to compute (generalised) cohomologies of chain complexes. Essentially one breaks the computation into pieces, which are 'easier' chain complexes, called pages, $E_0, E_1, \dots, E_\infty$, with differentials $\mathrm{d}_0, \mathrm{d}_1, \dots, \mathrm{d}_\infty$, such that E_{p+1} is the cohomology of E_p. Each 'next page' is a better approximation for the cohomology group, $H^\bullet$ one wishes to study. If the spectral sequence converges, the last page E_∞ is in principle equivalent to $H^\bullet$, although in real world there might be reconstruction problems.

To put into maths, say that we want to compute some cohomology group; call it H^d. The spectral sequence creates a (infinitely long) book – a sequence of pages consisting of tables: $E_r^{p\cdot q}$, with $p+q=d$, such that there is a differential

$$\mathrm{d}_r : E_r^{p,q} \longrightarrow E_r^{p-r+1,q+r}, \tag{B.2.1}$$

with $\mathrm{d}_r^2 = 0$ and

$$E_{r+1}^{p,q} := \frac{\ker \mathrm{d}_r}{\mathrm{im} \mathrm{d}_r}. \tag{B.2.2}$$

Then $E_\infty^{p,q}$ is related to $H^{d=p+q}$ as follows:

$$H^d =: F^d \supset F^{d-1} \supset \cdots \supset F^1 \supset F^0, \tag{B.2.3}$$

where

$$E_\infty^{d,0} = \frac{F^d}{F^{d-1}}, \quad E_\infty^{d-1,1} = \frac{F^{d-1}}{F^{d-2}}, \cdots \quad , \quad E_\infty^{1,d-1} = \frac{F^1}{F^0}, \quad E_\infty^{0,d} =: F^0. \tag{B.2.4}$$

B.2.1 Lyndon–Hochschild–Serre spectral sequence

Given a group extension $1 \longrightarrow \mathcal{A} \longrightarrow \Gamma \longrightarrow G \longrightarrow 1$ and a Γ-module, M, there is a spectral sequence

$$E_2^{p,q} \cong \mathrm{H}^p\big(G, \mathrm{H}^q(\mathcal{A}, M)\big) \Longrightarrow \mathrm{H}^{p+q}(\Gamma, M). \tag{B.2.5}$$

B.2.2 Atiyah–Hirzenbruch spectral sequence

Given a fibration $F \longrightarrow E \longrightarrow B$ and a generalised cohomology theory $\mathcal{H}^\bullet$, there is a spectral sequence

$$E_2^{p,q} \cong \mathrm{H}^p(B, \mathcal{H}^q(F)) \Longrightarrow \mathcal{H}^{p+q}(E). \tag{B.2.6}$$

The Atiyah–Hirzenbruch spectral sequence can be used to compute SPT phases that are beyond group cohomology. For example if they are classified by $\mho_\bullet^d$, one can take the fibre F to be trivial, i.e. a point, $F = \mathsf{pt}$ so one simply has $\mathsf{pt} \longrightarrow E \longrightarrow B \Longleftrightarrow E \cong B$ and can compute

$$\mathrm{H}^p\big(B, \mho_\bullet^q(\mathsf{pt})\big) \Longrightarrow \mho_\bullet^{p+q}(B). \tag{B.2.7}$$

B.3 Spin structure vs spin manifolds

A theory, on a manifold X, can accommodate fermions/spinors iff there exists a spinor bundle over X, of which spinors will be smooth sections. Such a bundle can exist if and only if the manifold that it is defined in has *spin structure*. Spin structures exist if and only if the second Stiefel-Whitney class vanishes:

$$H^2(M, \mathbb{Z}_2) \ni w_2(M) = 0. \tag{B.3.1}$$

A manifold is called *spin* if and only if $w_2(\mathrm{T}M) = 0$. Said differently, the Stiefel–Whitney class measure the obstruction of lifting a $\mathrm{PSU}(N)$ bundle to an $\mathrm{SU}(N)$ bundle. This is the reason why the second Stiefel Whitney class is of uttermost importance for anomalies in gauge theories.

B.4 Basics of category theory

In this section, we lay out the basic definitions of category and higher-category theory so that the thesis is as self-contained as possible. However, this is far from a complete exposition. A longer list of essential notions is in [72], while a much more complete and reliable exposition to higher-categories is [73].

Definition B.4.1 *A category $\boldsymbol{C}$ consists of*

1. *A collection of objects:* $\mathrm{Obj}(\boldsymbol{C}) \ni a, b, \ldots$.
2. *For any* $a, b \in \mathrm{Obj}(\boldsymbol{C})$*, a set of morphisms* $\mathrm{Hom}(a \longrightarrow b)$*, denoted by arrows:*
$$a \xrightarrow{f} b. \tag{B.4.1}$$
3. *A rule to compose arrows:*
$$\begin{tikzcd} a \arrow[r, "f"] \arrow[dr, "gf"'] & b \arrow[d, "g"] \\ & c \end{tikzcd} . \tag{B.4.2}$$
4. *An identity morphism:* $\begin{tikzcd} a \arrow[loop right, "1_a"] \end{tikzcd}$, *for every object* $a \in \mathrm{Obj}(\boldsymbol{C})$.

Morphisms should be associative, i.e. $(fg)h = f(gh) = fgh$ *and the identity should compose trivially for any object* $\bullet \in \mathrm{Obj}(\boldsymbol{C})$*, i.e.* $f\,1_\bullet = f = 1_\bullet f$.

Definition B.4.2 *A morphism* $f : a \longrightarrow b$ *is said to be an isomorphism if there exists an inverse* $f^{-1} : b \longrightarrow a$ *such that* $f f^{-1} = 1_b$ *and* $f^{-1} f = 1_a$.

Definition B.4.3 *A functor* $F : \boldsymbol{C} \longrightarrow \boldsymbol{D}$ *is a map between categories, meaning that it maps* $\mathrm{Obj}(\boldsymbol{C}) \overset{F}{\longmapsto} \mathrm{Obj}(\boldsymbol{D})$ *and* $\mathrm{Hom}(a \longrightarrow b) \overset{F}{\longmapsto} \mathrm{Hom}(F(a) \longrightarrow F(b))$, $\forall a, b \in \mathrm{Obj}(\boldsymbol{C})$.

Definition B.4.4 *A 2-category, ${}_2\boldsymbol{C}$ consists of everything that a category consists of, together with 2-morphisms, i.e. maps between morphisms:*

$$\begin{tikzcd} a \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\Downarrow\mu"] & b \end{tikzcd} . \tag{B.4.3}$$

I.e. a 2-category contains

1. *Objects,*
2. *1-morphisms between objects, and*
3. *2-morphisms between 1-morphisms.*

More generally we can define an n-category, ${}_n\boldsymbol{C}$, which consists of

1. *Objects,*
2. *1-morphisms between objects,*
3. *2-morphisms between 1-morphisms,*

$\vdots$

n. *$(n-1)$-morphisms between $(n-2)$-morphisms.*

Definition B.4.5 *A category $\boldsymbol{C}$ is called monoidal if there is a notion of tensor product $\otimes : \boldsymbol{C} \times \boldsymbol{C} \longrightarrow \boldsymbol{C}$. This brings together friends: a new morphism, called associator: $\alpha_{a,b,c} : (a \otimes b) \otimes c \overset{\sim}{\longmapsto} a \otimes (b \otimes c)$, and a special object $\bullet \in \mathrm{Obj}(\boldsymbol{C})$.The associator comes* *with a constraint, namely, the pentagon identity:*

$$\begin{tikzcd}
 & \big((a\otimes b)\otimes c\big)\otimes d \arrow[dl, "\alpha_{a,b,c}\otimes 1_d"'] \arrow[dr, "\alpha_{a\otimes b,c,d}"] & \\
\big(a\otimes(b\otimes c)\big)\otimes d \arrow[d, "\alpha_{a,b\otimes c,d}"] & & (a\otimes b)\otimes(c\otimes d) \arrow[d, "\alpha_{a,b,c\otimes d}"] \\
a\otimes\big((b\otimes c)\otimes d\big) \arrow[rr, "1_a\otimes\alpha_{a,c,d}"] & & a\otimes\big(b\otimes(c\otimes d)\big)
\end{tikzcd}, \tag{B.4.4}$$

while $\bullet$ *acts as the identity of tensoring:* $a\otimes\bullet = a = \bullet\otimes a$, $\forall a\in \mathrm{Obj}(C)$.

In physics, everybody wants tensor categories. The definition of those varies in the literature. Some define them as monoidal categories [10], while others are more strict [74] and define them as $\mathbb{C}$-linear semisimple rigid monoidal categories are called tensor categories. These will be the most important.

Definition B.4.6 *A braiding on a monoidal category is a natural isomorphism:* $\beta_{a,b} : a\otimes b \overset{\sim}{\longmapsto} b\otimes a$, *such that*

$$\begin{tikzcd}
 & a\otimes(b\otimes c) \arrow[r, "\beta_{a,b\otimes c}"] & (b\otimes c)\otimes a \arrow[dr, "\alpha_{b,c,a}"] & \\
(a\otimes b)\otimes c \arrow[ur, "\alpha_{a,b,c}"] \arrow[dr, "\beta_{a,b}\otimes 1_c"'] & & & b\otimes(c\otimes a) \\
 & (b\otimes a)\otimes c \arrow[r, "\alpha_{b,a,c}"'] & b\otimes(a\otimes c) \arrow[ur, "1_b\otimes\beta_{a,c}"'] &
\end{tikzcd} \tag{B.4.5}$$

and

$$
\begin{tikzcd}
 & (a \otimes b) \otimes c \arrow[r, "\beta_{a \otimes b, c}"] & c \otimes (a \otimes b) \arrow[dr, "\alpha^{-1}{}_{c,a,b}"] & \\
a \otimes (b \otimes c) \arrow[ur, "\alpha^{-1}{}_{a,b,c}"] \arrow[dr, "1_a \otimes \beta_{b,c}"'] & & & (c \otimes a) \otimes b \\
 & a \otimes (c \otimes b) \arrow[r, "\alpha^{-1}{}_{a,c,b}"'] & (a \otimes c) \otimes b \arrow[ur, "\beta_{a,c} \otimes 1_b"'] &
\end{tikzcd}
\tag{B.4.6}
$$

commute.

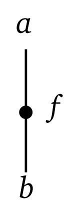
A braiding is essentially a step towards complete abelianisation/symmetrisation; this means that we can exchange $a \otimes b$ with $b \otimes a$ but only up to isomorphism. If we also demand that $\beta_{a,b}\beta_{a,b} = \bullet$ then we land on a symmetric monoidal category.

We can start thinking about higher categories and define a tensor product structure on a 2-category, to get a monoidal 2-category, which we can then make a little bit more commutative, by adding a braiding to get a braided monoidal 2-category and now we have an extra step before full symmetrisation, namely the syllepsis, which brings us to a sylleptic monoidal 2-category. Finally, if we symmetrise more, we get to the symmetric monoidal category. We see that in a 1-category, there are three steps to complete symmetry: tensoring, braiding, symmetry; in a 2-category, there are three steps: tensoring, braiding, syllepsing, symmetry. If we took one step more, we would fall into the trivial (2-)category. In general, it is conjectured, that, starting from an n-category, we fall into the trivial n-category after $n+2$ steps; this is known as the stabilisation hypothesis [73]

String diagrams

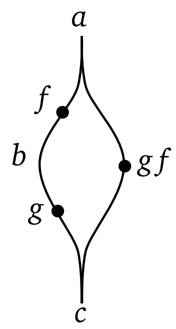
There is a Poincaré dual way to thing about categories. Say we are in a normal category. Instead of thinking of the objects as the points and the morphisms as arrows between them, we can think of the points as lines, starting from the top and the function as a point (or a black box) that changes them to a different line.

$$a \xrightarrow{f} b \qquad \Longleftrightarrow$$



Now commutative diagrams become string equations; we do not win too much yet. We do win, however, when we define a tensor product. Then commutative diagrams can be fully seen within a string. For example

$$\begin{tikzcd} a \arrow[r, "f"] \arrow[dr, "gf"'] & b \arrow[d, "g"] \\ & c \end{tikzcd} \qquad \Longleftrightarrow$$



Here the tensor product structure was needed to split a into $a \otimes a$ and combine $c \otimes c$ to c. Every diagram we will consider from now on is to be thought within a monoidal (n-)category.

Now if we want to define a braiding, we need three dimensions, to allow lines to pass behind other lines. Had we wanted to take one more step and go symmetric, we would extend to one more dimension, so all the cables would untangle. We attempted to draw a sketch, but we ran out of dimensions.

If we try to think about monoidal 2-categories, we have objects, morphisms and 2-morphisms. we have to be able to go as low as codimension-3. So a monoidal 2-category can be represented similarly by a "membrane" diagram in three dimensions, where sheets (codim-1) are the objects, lines connecting sheets (codim-2) are the 1-morphisms and points connecting sheet-conecting-lines (codim-3) are the 2-morphisms

B

Bibliography

[1] M. Atiyah, *Topological quantum field theories*, *Inst. Hautes Etudes Sci. Publ. Math.* **68** (1989) 175. (cited on page 02)

[2] G. Segal, *The definition of conformal field theory*, in *Symposium on Topology, Geometry and Quantum Field Theory (Segalfest)*, pp. 421–575, 6, 2002. (cited on page 02)

[3] D. S. Freed and M. J. Hopkins, *Reflection positivity and invertible topological phases*, 1604.06527. (cited on pages 06 , 39, and 40)

[4] D. M. Hofman and N. Iqbal, *Goldstone modes and photonization for higher form symmetries*, *SciPost Phys.* **6** (2019) 006 [1802.09512]. (cited on pages 08 and 24)

[5] E. Lake, *Higher-form symmetries and spontaneous symmetry breaking*, 1802.07747. (cited on pages 08 and 24)

[6] F. Benini, C. Córdova and P.-S. Hsin, *On 2-group global symmetries and their anomalies*, *JHEP* **03** (2019) 118 [1803.09336]. (cited on pages 13, 57, 58, 91, and 103)

[7] C. Córdova, T. T. Dumitrescu and K. Intriligator, *Exploring 2-group global symmetries*, *JHEP* **02** (2019) 184 [1802.04790]. (cited on pages 13, 57, 58, and 103)

[8] E. Sharpe, *Notes on generalized global symmetries in QFT*, *Fortsch. Phys.* **63** (2015) 659 [1508.04770]. (cited on page 13)

[9] C. Córdova, T. T. Dumitrescu and K. Intriligator, "To appear.". (cited on pages 13 and 103)

[10] L. Bhardwaj and Y. Tachikawa, *On finite symmetries and their gauging in two dimensions*, *JHEP* **03** (2018) 189 [1704.02330]. (cited on pages 13, 60, 76, 89, 101, 102, 104, and 114)

[11] R. Thorngren and Y. Wang, *Fusion category symmetry I: anomaly in-flow and gapped phases*, 1912.02817. (cited on pages 13 and 104)

[12] T. Rudelius and S.-H. Shao, *Topological operators and completeness of spectrum in discrete gauge theories*, 2006.10052. (cited on pages 13, 96, and 105)

[13] L. Kong, T. Lan, X.-G. Wen, Z.-H. Zhang and H. Zheng, *Algebraic higher symmetry and categorical symmetry – a holographic and entanglement view of symmetry*, 2005.14178. (cited on pages 13, 84, and 94)

[14] W. Ji and X.-G. Wen, *Categorical symmetry and non-invertible anomaly in symmetry-breaking and topological phase transitions*, 1912.13492. (cited on page 13)

[15] Z. Komargodski, K. Ohmori, K. Roumpedakis and S. Seifnashri, "To appear.". (cited on page 13)

[16] A. Kapustin and N. Seiberg, *Coupling a QFT to a TQFT and duality*, *JHEP* **04** (2014) 001 [1401.0740]. (cited on pages 14 and 100)

[17] D. Gaiotto, A. Kapustin, N. Seiberg and B. Willett, *Generalized global symmetries*, *JHEP* **02** (2015) 172 [1412.5148]. (cited on pages 14, 24, and 100)

[18] Z. Komargodski, "Dynamics in four dimensions and domain walls 1." 2018, youtube link. (cited on page 26)

[19] R. Dijkgraaf and E. Witten, *Topological gauge theories and group cohomology*, *Commun. Math. Phys.* **129** (1990) 393. (cited on page 29)

[20] C. Vafa, *Modular invariance and discrete torsion on orbifolds*, *Nucl. Phys. B* **273** (1986) 592. (cited on page 30)

[21] M. R. Douglas, *D-branes and discrete torsion*, hep-th/9807235. (cited on page 30)

[22] X. Chen, Z.-C. Gu, Z.-X. Liu and X.-G. Wen, *Symmetry protected topological orders and the group cohomology of their symmetry group*, *Phys. Rev. B* **87** (2013) 155114 [1106.4772]. (cited on pages 35 and 59)

[23] X. Chen, Z.-C. Gu, Z.-X. Liu and X.-G. Wen, *Symmetry-protected topological orders in interacting bosonic systems*, *Science* **338** (2012) 1604 [1301.0861]. (cited on page 35)

[24] Z.-C. Gu and X.-G. Wen, *Symmetry-protected topological orders for interacting fermions: Fermionic topological nonlinear σ models and a special group supercohomology theory*, *Phys. Rev. B* **90** (2014) 115141 [1201.2648]. (cited on page 36)

[25] A. Vishwanath and T. Senthil, *Physics of three dimensional bosonic topological insulators: surface deconfined criticality and quantized magnetoelectric effect*, *Phys. Rev. X* **3** (2013) 011016 [1209.3058]. (cited on page 36)

[26] C. Wang and T. Senthil, *Boson topological insulators: A window into highly entangled quantum phases*, *Phys. Rev. B* **87** (2013) 235122 [1302.6234]. (cited on page 36)

[27] F. Burnell, X. Chen, L. Fidkowski and A. Vishwanath, *Exactly soluble model of a three-dimensional symmetry-protected topological phase of bosons with surface topological order*, *Phys. Rev. B* **90** (2014) 245122 [1302.7072]. (cited on page 36)

[28] A. Kapustin, *Symmetry protected topological phases, anomalies, and cobordisms: beyond group cohomology*, 1403.1467. (cited on page 37)

[29] A. Kapustin, R. Thorngren, A. Turzillo and Z. Wang, *Fermionic symmetry protected topological phases and cobordisms*, *JHEP* **12** (2015) 052 [1406.7329]. (cited on page 37)

[30] K. Yonekura, *On the cobordism classification of symmetry protected topological phases*, *Commun. Math. Phys.* **368** (2019) 1121 [1803.10796]. (cited on pages 37 and 38)

[31] D. S. Freed, *Short-range entanglement and invertible field theories*, 1406.7278. (cited on page 37)

[32] Z.-C. Gu and X.-G. Wen, *Tensor-entanglement-filtering renormalization approach and symmetry protected topological order*, *Phys. Rev. B* **80** (2009) 155131 [0903.1069]. (cited on page 37)

[33] J. Milnor, L. Siebenmann and J. Sondow, *Lectures on the H-Cobordism Theorem*. Princeton University Press, 1965. (cited on page 38)

[34] C.-T. Hsieh, Y. Tachikawa and K. Yonekura, *Anomaly inflow and p-form gauge theories*, 2003.11550. (cited on page 38)

[35] C. Delcamp and A. Tiwari, *From gauge to higher gauge models of topological phases*, *JHEP* **10** (2018) 049 [1802.10104]. (cited on page 39)

[36] Z. Wan and J. Wang, *Higher anomalies, higher symmetries, and cobordisms I: classification of higher-symmetry-protected topological states and their boundary fermionic/bosonic anomalies via a generalized cobordism theory*, *Ann. Math. Sci. Appl.* **4** (2019) 107 [1812.11967]. (cited on pages 40 and 41)

[37] L. Kong, T. Lan, X.-G. Wen, Z.-H. Zhang and H. Zheng, *Classification of topological phases with finite internal symmetries in all dimensions*, 2003.08898. (cited on pages 41, 43, and 104)

[38] L. Kong, X.-G. Wen and H. Zheng, *Boundary-bulk relation for topological orders as the functor mapping higher categories to their centers*, *arXiv* (2015) [1502.01690]. (cited on page 41)

[39] L. Kong, X.-G. Wen and H. Zheng, *Boundary-bulk relation in topological orders*, *Nucl. Phys. B* **922** (2017) 62 [1702.00673]. (cited on page 41)

[40] D. Gaiotto and T. Johnson-Freyd, *Condensations in higher categories*, 1905.09566. (cited on pages 41, 42, and 90)

[41] E. Sharpe, *Discrete torsion, quotient stacks, and string orbifolds*, math/0110156. (cited on page 45)

[42] C. Vafa, *Quantum Symmetries of String Vacua*, *Mod. Phys. Lett. A* **4** (1989) 1615. (cited on pages 46 and 101)

[43] Y. Tachikawa, *On gauging finite subgroups*, *SciPost Phys.* **8** (2020) 015 [1712.09542]. (cited on pages 51, 98, 102, and 104)

[44] M. B. Green and J. H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, *Phys. Lett. B* **149** (1984) 117. (cited on page 52)

[45] N. Seiberg, Y. Tachikawa and K. Yonekura, *Anomalies of duality groups and extended conformal manifolds*, *PTEP* **2018** (2018) 073B04 [1803.07366]. (cited on pages 62, 64, and 101)

[46] C. Córdova, D. S. Freed, H. T. Lam and N. Seiberg, *Anomalies in the space of coupling constants and their dynamical applications I*, *SciPost Phys.* **8** (2020) 001 [1905.09315]. (cited on pages 62, 66, and 101)

[47] C. Córdova, D. S. Freed, H. T. Lam and N. Seiberg, *Anomalies in the space of coupling constants and their dynamical applications II*, *SciPost Phys.* **8** (2020) 002 [1905.13361]. (cited on pages 62, 66, and 101)

[48] E. Witten, *On S duality in abelian gauge theory*, *Selecta Math.* **1** (1995) 383 [hep-th/9505186]. (cited on page 64)

[49] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, *Theta, time reversal, and temperature*, *JHEP* **05** (2017) 091 [1703.00501]. (cited on pages 66 and 101)

[50] A. Kapustin and L. Spodyneiko, *Higher-dimensional generalizations of the berry curvature*, *Phys. Rev. B* **101** (2020) 235130 [2001.03454]. (cited on page 74)

[51] P.-S. Hsin, A. Kapustin and R. Thorngren, *Berry phase in quantum field theory: diabolical points and boundary phenomena*, 2004.10758. (cited on page 74)

[52] G. W. Moore and N. Seiberg, *Classical and quantum conformal field theory*, *Commun. Math. Phys.* **123** (1989) 177. (cited on pages 83, 86, 87, 95, and 101)

[53] E. P. Verlinde, *Fusion rules and modular transformations in 2d conformal field theory*, *Nucl. Phys. B* **300** (1988) 360. (cited on pages 84 and 95)

[54] J. Frohlich, J. Fuchs, I. Runkel and C. Schweigert, *Defect lines, dualities, and generalised orbifolds*, in *16th International Congress on Mathematical Physics*, 9, 2009, DOI [0909.5013]. (cited on page 87)

[55] N. Carqueville, C. Meusburger and G. Schaumann, *3-dimensional defect TQFTs and their tricategories*, *Adv. Math.* **364** (2020) 107024 [1603.01171]. (cited on pages 89 and 99)

[56] C. Wang and M. Levin, *Braiding statistics of loop excitations in three dimensions*, *Phys. Rev. Lett.* **113** (2014) 080403 [1403.7437]. (cited on page 92)

[57] S. Jiang, A. Mesaros and Y. Ran, *Generalized modular transformations in* $(3+1)d$ *topologically ordered phases and triple linking invariant of loop braiding*, *Physical Review X* **4** (2014) [1404.1062]. (cited on page 92)

[58] J. Baez and J. Dolan, *Higher dimensional algebra and topological quantum field theory*, *J. Math. Phys.* **36** (1995) 6073 [q-alg/9503002]. (cited on pages 92 and 93)

[59] E. Cheng and N. Gurski, *The periodic table of n-categories for low dimensions I: degenerate categories and degenerate bicategories*, *arXiv e-prints* (2007) [0708.1178]. (cited on page 92)

[60] E. Cheng and N. Gurski, *The periodic table of n-categories for low dimensions II: degenerate tricategories*, *arXiv e-prints* (2007) [0706.2307]. (cited on page 92)

[61] T. Johnson-Freyd, *On the classification of topological orders*, 2003.06663. (cited on page 93)

[62] S. Gukov and E. Witten, *Rigid surface operators*, *Adv. Theor. Math. Phys.* **14** (2010) 87 [0804.1561]. (cited on page 96)

[63] J. Frohlich, J. Fuchs, I. Runkel and C. Schweigert, *Kramers-Wannier duality from conformal defects*, *Phys. Rev. Lett.* **93** (2004) 070601 [cond-mat/0404051]. (cited on page 101)

[64] X.-G. Wen, *Emergent anomalous higher symmetries from topological order and from dynamical electromagnetic field in condensed matter systems*, *Phys. Rev. B* **99** (2019) 205139 [1812.02517]. (cited on page 102)

[65] M. Levin and Z.-C. Gu, *Braiding statistics approach to symmetry-protected topological phases*, *Phys. Rev. B* **86** (2012) 115109 [1202.3120]. (cited on page 102)

[66] D. Anninos, D. A. Galante and D. M. Hofman, *De sitter horizons & holographic liquids*, *JHEP* **07** (2019) 038 [1811.08153]. (cited on page 104)

[67] K. Alexei, "A simple model of quantum holography." 2015, http://online.kitp.ucsb.edu/online/entangled15/. (cited on page 104)

[68] J. Maldacena and D. Stanford, *Remarks on the sachdev-ye-kitaev model*, *Phys. Rev. D* **94** (2016) 106002 [1604.07818]. (cited on page 104)

[69] R. Kallosh, A. D. Linde, D. A. Linde and L. Susskind, *Gravity and global symmetries*, *Phys. Rev. D* **52** (1995) 912 [hep-th/9502069]. (cited on page 105)

[70] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys. Rev. D* **83** (2011) 084019 [1011.5120]. (cited on page 105)

[71] D. Harlow and H. Ooguri, *Symmetries in quantum field theory and quantum gravity*, 1810.05338. (cited on page 105)

[72] J. C. Baez, *Some definitions everyone should know*, 2010, http://math.ucr.edu/home/baez/qg-winter2001/definitions.pdf. (cited on page 112)

[73] J. C. Baez and M. Shulman, *Lectures on n-categories and cohomology*, math/0608420. (cited on pages 112 and 115)

[74] D. J. Binder and S. Rychkov, *Deligne categories in lattice models and quantum field theory, or making sense of $O(N)$ symmetry with non-integer N*, *JHEP* **04** (2020) 117 [1911.07895]. (cited on page 114)