Classifications of Quantum Field Theories

by

Lakshya Bhardwaj

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2018

© Lakshya Bhardwaj 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Amihay Hanany, Imperial College London
Supervisor(s):	Davide Gaiotto, Perimeter Institute Robert Myers, Perimeter Institute
Internal Member:	Robert Mann, University of Waterloo
Internal-External Member:	Ben Webster, University of Waterloo
Other Member(s):	Kevin Costello, Perimeter Institute

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 2 of this thesis consists of material from the paper [1], co-authored with Michele Del Zotto, Jonathan J. Heckman, David R. Morrison, Tom Rudelius and Cumrun Vafa.

Chapter 3 of this thesis consists of material from the paper [2], co-authored with David R. Morrison, Yuji Tachikawa and Alessandro Tomasiello.

Chapter 4 of this thesis consists of material from the paper [3], co-authored with Davide Gaiotto and Anton Kapustin.

Chapter 5 of this thesis consists of material from the paper [4].

Chapter 6 of this thesis consists of material from the paper [5], co-authored with Yuji Tachikawa.

Abstract

We discuss classifications of UV complete supersymmetric theories in six dimensions, and (spin-)topological field theories admitting a finite global symmetry and possibly time-reversal symmetry in three dimensions. We also discuss a generalization of finite global symmetries and their gauging in two dimensions.

First, we start with LSTs which are UV complete non-local 6D theories decoupled from gravity in which there is an intrinsic string scale. We present a systematic approach to the construction of supersymmetric LSTs via the geometric phases of F-theory. Our central result is that all LSTs with more than one tensor multiplet are obtained by a mild extension of 6D superconformal field theories (SCFTs) in which the theory is supplemented by an additional, non-dynamical tensor multiplet, analogous to adding an affine node to an ADE quiver, resulting in a negative semidefinite Dirac pairing. We also show that all 6D SCFTs naturally embed in an LST. Motivated by physical considerations, we show that in geometries where we can verify the presence of two elliptic fibrations, exchanging the roles of these fibrations amounts to T-duality in the 6D theory compactified on a circle.

Second, we study the interpretation of $O7_+$-planes in F-theory, mainly in the context of the six dimensional models. In particular, we study how to assign gauge algebras and matter contents to seven-branes and their intersections, and the implication of anomaly cancellation in our construction, generalizing earlier analyses without any $O7_+$-planes. By including $O7_+$-planes we can realize 6d superconformal field theories hitherto unobtainable in F-theory, such as those with hypermultiplets in the symmetric representation of special unitary gauge algebra. We also examine a couple of compact models. These reproduce some famous perturbative models, and in some cases enhance their gauge symmetries non-perturbatively.

Third, we argue that it is possible to describe fermionic phases of matter and spin-topological field theories in 2+1d in terms of bosonic "shadow" theories, which are obtained from the original theory by "gauging fermionic parity". The fermionic/spin theories are recovered from their shadow by a process of fermionic anyon condensation: gauging a one-form symmetry generated by quasi-particles with fermionic statistics. We apply the formalism to theories which admit gapped boundary conditions. We obtain Turaev-Viro-like and Levin-Wen-like constructions of fermionic phases of matter. We describe the group structure of fermionic SPT phases protected by the product of fermion parity and internal symmetry G. The quaternion group makes a surprise appearance.

Fourth, we generalize two facts about oriented 3d TFTs to the unoriented case. On one hand, it is known that oriented 3d TFTs having a topological boundary condition admit a state-sum construction known as the Turaev-Viro construction. This is related to the string-net construction of fermionic phases of matter. We show how Turaev-Viro construction can be generalized to unoriented 3d TFTs. On the other hand, it is known that the "fermionic" versions of oriented TFTs, known as Spin-TFTs, can be constructed in terms of "shadow" TFTs which are ordinary oriented TFTs with an anomalous $\mathbb{Z}_2$ 1-form symmetry. We generalize this correspondence to Pin^+-TFTs by showing that they can be constructed in terms of ordinary unoriented TFTs with anomalous $\mathbb{Z}_2$ 1-form symmetry having a mixed anomaly with time-reversal symmetry. The corresponding Pin^+-TFT does not have any anomaly for time-reversal symmetry however and

hence it can be unambiguously defined on a non-orientable manifold. In case a Pin^+-TFT admits a topological boundary condition, one can combine the above two statements to obtain a Turaev-Viro-like construction of Pin^+-TFTs. As an application of these ideas, we construct a large class of Pin^+-SPT phases.

Finally, we recall that it is well-known that if we gauge a $\mathbb{Z}_n$ symmetry in two dimensions, a dual $\mathbb{Z}_n$ symmetry appears, such that re-gauging this dual $\mathbb{Z}_n$ symmetry leads back to the original theory. We describe how this can be generalized to non-Abelian groups, by enlarging the concept of symmetries from those defined by groups to those defined by unitary fusion categories. We will see that this generalization is also useful when studying what happens when a non-anomalous subgroup of an anomalous finite group is gauged: for example, the gauged theory can have non-Abelian group symmetry even when the original symmetry is an Abelian group. We then discuss the axiomatization of two-dimensional topological quantum field theories whose symmetry is given by a category. We see explicitly that the gauged version is a topological quantum field theory with a new symmetry given by a dual category.

Acknowledgements

I would like to thank my supervisor Davide Gaiotto for always pushing me to become an independent researcher and for teaching me the cutting-edge viewpoint on quantum field theory.

A special thanks goes to Yuji Tachikawa who was always welcoming for discussions and guidance.

Next, I would like to thank all of my collaborators who contributed their experience and expertise to our joint research projects. These are Michele Del Zotto, Davide Gaiotto, Jonathan Heckman, Anton Kapustin, Dave Morrison, Tom Rudelius, Yuji Tachikawa, Alessandro Tomasiello and Cumrun Vafa.

I would also like to thank Rob Myers for serving as my co-supervisor, and Kevin Costello and Rob Mann for serving on my advisory committee.

Finally, I am indebted to my parents for their love and support.

Dedication

This thesis is dedicated to Stephen Hawking and Joseph Polchinski - two brilliant physicists who unfortunately passed away recently.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this thesis, our operational definition of a quantum field theory is any relativistic quantum mechanical theory in the absence of dynamical gravity. Thus, it is a matter of extreme importance to understand quantum field theories. Even before trying to understand QFTs, one might ask: What is the space of QFTs? It turns out that even this zeroth order question is an extremely challenging one to answer.

Recently, a lot of progress was made on this question on two fronts. In the rest of this introductory chapter, we will provide an overview of these two developments which should be accessible to non-experts. The other chapters of the thesis will record the technical work performed by the author as a part of these developments. Each subsequent chapter will be accompanied by its own introductory section addressed to the experts.

1.1 6d supersymmetric QFTs

On one front, people made progress by restricting their attention to six-dimensional supersymmetric UV complete QFTs. UV completeness of a QFT means that the QFT describes physical processes at all energy scales. It is typically assumed because it is extremely restrictive. The criterion of UV completeness drastically cuts down the space of QFTs one needs to explore. This was a lesson learned from the famous renormalization program of the last century. Supersymmetry is assumed in order to provide a strong computational control on the quantum fluctuations in the theories under consideration because supersymmetry forces quantum fluctuations of bosonic degrees of freedom to cancel against the quantum fluctuations of fermionic degrees of freedom.

To simplify the problem further, we could imagine considering only those theories in which physical phenomena happening at different energy scales are the same. In other words, we would like to consider theories in which the physics is not a function of the energy scale at which we probe the system. Such QFTs are called conformal field theories (CFTs). Such theories occupy a very small subspace of the space of QFTs. Flipping this statement around, we can now think of QFTs as deformations of CFTs. Thus the problem of understand QFTs essentially boils down to first understanding the space of CFTs and then understanding the space of deformations of CFTs.

Now, it is a general expectation that the space of QFTs gets smaller as one increases spacetime dimension. The reasoning is that starting from a higher dimensional QFT, one can in principle obtain myriad of lower dimensional QFTs by compactifying some of the space directions on a compact manifold of very small size. However, it is known that the maximal spacetime dimension in which a SCFT can exist is six. So one obtains the most amount of restrictions when one looks for six-dimensional SCFTs.

One very successful strategy for constructing 6d SCFTs is via compactifications of F-theory. F-theory is a name used to describe a certain class of non-perturbative vacua of type IIB string theory in which the axion and dilaton obtain space dependent vevs. The vacua can be described by means of an elliptic fibration over physical space, thus making the effective spacetime dimension twelve.

To construct 6d SCFTs, one compactifies F-theory on a non-compact elliptically fibered Calabi-Yau threefold with the intersection pairing of 2-cycles in the base being positive definite. Thus the classification of 6d SCFTs boils down to classification of such Calabi-Yaus. This classification was performed in [9, 10] by assuming additional input from anomaly cancellation in the resulting 6d QFT. The authors hoped that the actual list of such Calabi-Yaus would exactly match the list obtained using anomaly cancellation without any mismatch.

Now it turns out that there exists a class of supersymmetric UV complete QFTs in six dimension whose low-energy physics is described by 6d SCFTs. Unlike 6d SCFTs, these theories are non-local because they do not distinguish between a circular dimension of radius R and a circular dimension of radius $1/R$. In particular, if one tries to compactify these theories on a small circle, one obtains again a six dimensional theory rather than a five dimensional theory. Since this property is shared by string theory as well, these theories are called supersymmetric little string theories (or LSTs).

One can construct 6d LSTs by compactifying F-theory on a non-compact elliptically fibered Calabi-Yau threefold with the intersection pairing of 2-cycles in the base being positive semi-definite (but not positive definite). This classification was performed in [1] and is described in Chapter 2 below.

A twist in the above story came when it was realized that upon comparing the list of theories obtained via F-theory with the list of theories obtained via other methods of classification [11], there were some theories missing in the F-theoretic classification. The situation was particularly disturbing because the missing theories could be constructed via type IIA string theory which is supposed to be dual to F-theory. Thus if F-theory could not reproduce these theories, it would be a big conundrum.

It was soon realized that the above F-theoretic classifications assumed that one is dealing with F-theory compactifications which do not involve frozen singularities, or $O7_+$ planes in type IIB language. Incorporating frozen singularities into the machinery of F-theory turns out not be straightforward and many of the nice properties enjoyed by F-theory compactifications without frozen singularities is not shared by compactifications with frozen singularities. The incorporation of frozen singularities is the subject of Chapter 3 where it is also shown how to reproduce the missing theories via such compactifications.

1.2 3d topological field theories

On the other front, people restricted their attention to topological field theories (or TFTs) in two and three dimensions. TFTs are structurally even simpler than CFTs because they enjoy much more spacetime symmetry. The data of CFTs is insensitive to the size of the spacetime manifold but the data of TFTs is insensitive to any topological deformation of the spacetime manifold.

This seems to imply that TFTs should be trivial because there could be no spacetime dependence in the expectation values of local observables. However, most of the interesting data of TFTs is captured in the expectation values of non-local observables which are defined over a non-zero dimensional locus in spacetime. These observables are also referred to as defects. Defects in TFTs capture properties of spacetime manifold as they can wrap around non-trivial cycles in spacetime. Even more refined data is captured by junctions between defects which can be thought of as defects living inside defects. Of course, one can obtain even finer information by considering defects living inside defects living inside defects, and so on. Given this nested structure of non-local observables, it pays to restrict to low-enough dimensions where the nesting structure could be handled more easily.

The properties of line defects in TFTs are captured by the data of a fusion category possibly enriched by other structures on the category. In this language, the line defects are recognized as objects of teh category and local operators living at the junction of two line defects are recognized as morphisms between the corresponding objects in the category. Two line defects can be brought together and converted into a single line defect via topological deformations. This makes the category into a fusion category.

The extra structure needed to describe line defects in a 3d TFT is that of braiding which captures properties of lines as they wind around each other in three dimensional spacetime. The correct mathematical structure turns out to be that of a modular tensor category. However, Turaev and Viro [12] presented a construction of 3d TFTs based on the data of a spherical fusion category which does not admit a braiding structure.

This puzzle was resolved in [3] where it was argued that the fusion category of Turaev and Viro is actually describing lines on a topological boundary of the 3d TFT. Since the boundary is two dimensional, the lines living on it do not require a braiding structure and hence it makes sense. The Turaev-Viro prescription for obtaining partition function of the 3d TFT was also obtained using abstract principles of TFTs. Thus, one converts the problem of classififying 3d TFTs to the problem of classifying spherical fusion categories. One can restrict to those 3d TFTs which admit an inverse TFT in the sense that taking a direct product of the two TFTs gives rise to a trivial TFT. In this case, it is easy to classify the corresponding spherical fusion categories. In [3], a prescription was also given to obtain the data of a 3d spin-TFT from the data of a 3d TFT with an anomalous $\mathbb{Z}_2$ symmetry and vice-versa. A spin-TFT is a TFT which requires a choice of spin structure on the spacetime manifold to be well defined. Thus spin-TFTs can be thought of as TFTs with fermionic degrees of freedom. Using all of the above, one can classify invertible 3d spin-TFTs via spherical fusion categories. All of this is the content of Chapter 4.

The story of last paragraph holds for 3d TFTs which are defined only on oriented spacetime manifolds. This story can be enriched by generalizing to 3d TFTs which can be defined even of unoriented spacetime manifolds. Such unoriented TFTs are important because these TFTs

admit time-reversal symmetry. This symmetry can be thought to be the consequence of the existence of an orientation reversing codimension one defect in the TFT. The lines on a topological boundary in the presence of an orientation reversing defect were studied in [4] and the resulting mathematical structure was identified as that of a "twisted" spherical fusion category. The generalization of Turaev-Viro prescription in terms of data of a twisted spherical fusion category was also given. On an unoriented spacetime, there can be two types of TFTs with fermionic degrees of freedom. They are called Pin^+-TFTs and Pin^--TFTs and they are well-defined after choosing a Pin^+-structure or a Pin^--structure on the spacetime respectively. In [4], a correspondence was established between 3d Pin^+-TFTs and 3d unoriented TFTs with an anomalous $\mathbb{Z}_2$ symmetry and was used to provide a partial classification of invertible 3d Pin^+-TFTs. All of this is the subject of Chapter 5.

1.3 Finite symmetries in two dimensions

In Chapter 6, we go slightly outside the main theme of this thesis. Instead of talking about classification of QFTs, we will talk about correspondence between QFTs. It was shown a long time ago [13] that if one gauges a finite, abelian internal global symmetry group G of a two-dimensional QFT, the resulting 2d QFT has a finite, abelian internal global symmetry group $\widehat{G}$ which is the Pontryagin dual of G. Furthermore, if one gauge $\widehat{G}$ in the resulting 2d QFT, one obtains the original 2d QFT that one started with.

Chapter 6 asks what happens if G is a finite, non-abelian group. Gauging G produces Wilson line defects labeled by representations of G which form a fusion category known as the representation category of G. If G is abelian, the data of representation category is equivalent to the data of $\widehat{G}$. However, if G is non-abelian, there is no such notion of $\widehat{G}$. Is there a notion of "gauging the representation category of the gauged 2d QFT" to obtain the original 2d QFT with non-abelian global symmetry G?

This question was answered in affirmative in [5]. In fact, a much more general statement was argued. In general one can consider the fusion category $\mathcal{C}$ formed by topological line defects in a 2d QFT $\mathfrak{T}$. A notion of gauging $\mathcal{C}$ is provided by the choice of a special object known as Frobenius algebra object A in $\mathcal{C}$. This object satisfies some nice properties. Once gauges $\mathfrak{T}$ by inserting a fine-enough network of A lines in the 2d spacetime. In the gauged theory $\widehat{\mathfrak{T}}$, the network of A lines becomes invisible and the fusion category $\widehat{\mathcal{C}}$ formed by topological lines is the category of A-bimodules in $\mathcal{C}$. It can be shown that $\widehat{\mathcal{C}}$ admits a Frobenius algebra object $\widehat{A}$ which is dual to A in the sense that the category of $\widehat{A}$-bimodules in $\widehat{\mathcal{C}}$ can be identified with $\mathcal{C}$. Lifting this up, it can be shown that gauging $\widehat{\mathfrak{T}}$ by $\widehat{A}$ gives back the theory $\mathfrak{T}$ with associated fusion category $\mathcal{C}$.

Chapter 2

F-theory and the Classification of Little Strings

2.1 Introduction

One of the concrete outcomes from the post-duality era of string theory is the wealth of insights it provides into strongly coupled quantum systems. In the context of string compactification, this has been used to argue, for example, for the existence of novel interacting conformal field theories in spacetime dimensions $D > 4$. In a suitable gravity-decoupling limit, the non-local ingredients of a theory of extended objects such as strings are instead captured by a quantum field theory with a local stress energy tensor.

String theory also predicts the existence of novel non-local theories. Our focus in this work will be on 6D theories known as little string theories (LSTs).[1] For a partial list of LST constructions, see e.g. [14, 6, 15–19]. In these systems, 6D gravity is decoupled, but an intrinsic string scale M_{string} remains. At energies far below M_{string}, we have an effective theory which is well-approximated by the standard rules of quantum field theory with a high scale cutoff. However, this local characterization breaks down as we reach the scale M_{string}. The UV completion, however, is not a quantum field theory.[2]

The mere existence of little string theories leads to a tractable setting for studying many of the essential features of string theory –such as the presence of extended objects– but with fewer complications (such as coupling to quantum gravity). It also raises important conceptual questions connected with the UV completion of low energy quantum field theory. For example, in known constructions these theories exhibit T-duality upon toroidal compactification [16,17,21] and a Hagedorn density of states [22], properties which are typical of closed string theories with tension set by M^2_{string}.

Several families of LSTs have been engineered in the context of superstring theory by using various combinations of branes probing geometric singularities. The main idea in many of these constructions is to take a gravity-decoupling limit where the 6D Planck scale $M_{pl} \to \infty$ and

[1] We leave open the question of whether non-supersymmetric little string theories exist.

[2] In fact, all known properties of LSTs are compatible with the axioms for quasilocal quantum field theories [20].

the string coupling $g_s \to 0$, but with an effective string scale M_{string} held fixed. Even so, an overarching picture of how to construct (and study) LSTs has remained somewhat elusive.

Our aim in this chapter will be to demystify various aspects of LSTs, and in particular to give a systematic approach to realizing LSTs arising in F-theory[3]. To do this, we will use both a bottom up characterization of little string theories on the tensor branch (i.e. where all effective strings have picked up a tension), as well as a formulation in terms of compactifications of F-theory. To demonstrate UV completeness of the resulting models we will indeed need to use the F-theory characterization.

Recall that in F-theory, we have a non-compact base B of complex dimension two, which is supplemented by an elliptic fibration to reach a non-compact Calabi-Yau threefold. In the resolved phase, the intersection pairing of the base coincides with the Dirac pairing for two-form potentials of the theory on its tensor branch. For an SCFT, we demand that the Dirac pairing is negative definite. For an LST, we instead require that this pairing is negative semidefinite, i.e., we allow for a non-trivial null space.

F-theory also imposes the condition that we can supplement this base by an appropriate elliptic fibration to reach a non-compact Calabi-Yau threefold. In field theory terms, this is usually enforced by the condition that all gauge theoretic anomalies are cancelled on the tensor branch of the theory. For 6D gauge theories which complete to LSTs, this condition was discussed in detail in reference [26]. Even when no gauge theory interpretation is available, this means that in the theory on the tensor branch, some linear combination of tensor multiplets is non-dynamical, and instead defines a dimensionful parameter (effectively a UV cutoff) for the 6D effective field theory.

In F-theory terms, classifying LSTs thus amounts to determining all possible elliptic Calabi-Yau threefolds which support a base B with negative semidefinite intersection pairing. One of our results is that all LSTs are given by a small extension of 6D SCFTs, i.e. they can always be obtained by adding just one more curve to the base of an SCFT so that the resulting base has an intersection pairing with a null direction. Hence, much as in the case of Lie algebras, all LSTs arise from an affine extension of SCFTs. See figure 2.1.1 for a depiction of this process.

In fact, the related classification of 6D SCFTs has already been successfully carried out. See e.g. the partial list of references [9, 27–30, 10, 11]. What this means is that we can freely borrow this structure to establish a classification of LSTs. Much as in reference [10], we establish a similar "atomic classification" of how LSTs are built up from smaller constitutent elements. We find that the base of an F-theory geometry is organized according to a single spine of "nodes" which are decorated by possible radicals, i.e. links which attach to these nodes. As opposed to the case of SCFTs, however, the topology of an LST can be either a tree or a loop.

Using this characterization of LSTs, we also show that all 6D SCFTs can be embedded in some LST by including additional curves and seven-branes:

$$6\text{D SCFTs} \to 6\text{D LSTs}. \tag{2.1.1}$$

[3]More precisely, we focus on the case of geometric phases of F-theory, ignoring the (small) list of possible models with "frozen" singularities (see e.g. [23–25]). We return to this point later in section 2.10 when we discuss the mismatches between field theory motivated LST constructions and their possible lifts to string constructions

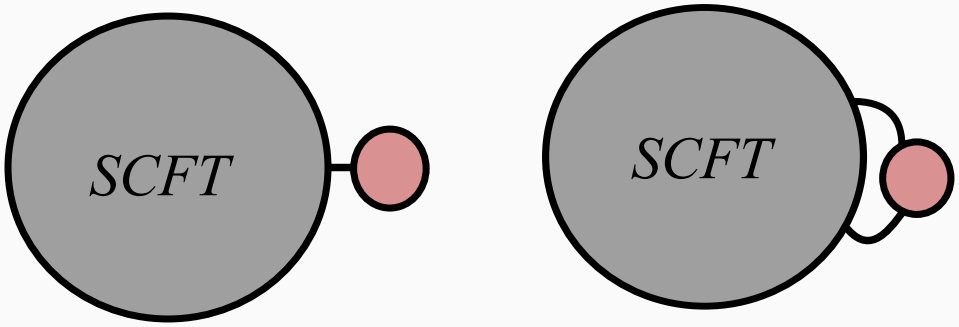


Figure 2.1.1: Depiction of how to construct the base of an F-theory model for an LST. All LST bases are obtained by adding one additional curve to the base for a 6D SCFT. This additional curve can intersect either one or two curves of the SCFT base. Much as in the study of Lie algebras, LSTs should be viewed as an "affine extension" of SCFTs.

Deformations in both Kähler and complex structure moduli for the LST then take us back to the original SCFT. It is curious to note that although many 6D SCFTs cannot be coupled to 6D supergravity, they *can* always be embedded in another theory with an intrinsic length scale.

A hallmark of all known LSTs is T-duality, that is, by compactifying on a small circle,[4] we reach another 6D LST compactified on a circle of large radius. This motivates a physical conjecture that all LSTs exhibit such a T-duality. In geometries where we can verify the presence of two elliptic fibrations, we find that exchanging the roles of these fibrations amounts to T-duality in the 6D theory compactified on a circle. In some cases, we find that T-duality takes us to the *same* LST. For a recent application of this double elliptic fibration structure in the study of the correspondence between instantons and monopoles via compactifications of little string theory, see reference [31].

The rest of this chapter is organized as follows. In section 2.2 we state necessary bottom up conditions to realize a LST. This includes the core condition that the Dirac pairing for an LST is a negative semidefinite matrix. After establishing some of the conditions this enforces, we then turn in section 2.3 to the rules for constructing LSTs in F-theory. We also explain the (small) differences between the rules for constructing LSTs versus SCFTs. Section 2.4 gives some examples of known constructions of LSTs, and their embedding in F-theory. In section 2.5 we show how decoupling a tensor multiplet to reach an SCFT leads to strong constraints on possible F-theory models. In section 2.6 we present an atomic classification of bases, and in section 2.7 we turn to the classification of possible elliptic fibrations over a given base. In section 2.8 we demonstrate that every 6D SCFT constructed in F-theory can be embedded into at least one 6D LST constructed in F-theory. In section 2.9 we show how T-duality of the LST shows up as the existence of a double elliptic fibration structure, and the exchange in the roles of the elliptic fibers. As a consequence, we show that LSTs can acquire discrete gauge symmetries for

[4] That is, small when compared with the effective string scale.

particular values of their moduli. In section 2.10 we discuss the small mismatch with possible LST constructions suggested by field theory, and their potential embedding in a non-geometric phase of an F-theory model. Section 2.11 contains our conclusions, and some additional technical material is deferred to a set of Appendices.

2.2 LSTs from the Bottom Up

In this section we state some of the conditions necessary to realize a supersymmetric little string theory.

We consider 6D supersymmetric theories which admit a tensor branch (which can be zero dimensional, as will be the case for many LSTs), that is, we will have a theory with some dynamical tensor multiplets, and vacua parameterized (at low energies) by vevs of scalars in these tensor multiplets. We will tune the vevs of the dynamical scalars to zero to reach a point of strong coupling. Our aim will be to seek out theories in which this region of strong coupling is not described by an SCFT, but rather, by an LST. In addition to dynamical tensor multiplets, we will allow the possibility of non-dynamical tensor multiplets which set mass scales for the 6D supersymmetric theory.

Recall that in a theory with T tensor multiplets, we have scalars S^I and their bosonic superpartners $B_{\mu\nu}^{-,I}$, with anti-self-dual field strengths. The vevs of the S^I govern, for example, the tension of the effective strings which couple to these two-form potentials. In a theory with gravity, one must also include an additional two-form potential $B_{\mu\nu}^{+}$ coming from the graviton multiplet. Given this collection of two-form potentials, we get a lattice of string charges Λ_{string}, and a Dirac pairing:[5]

$$\Lambda_{string} \times \Lambda_{string} \to \mathbb{Z}, \tag{2.2.1}$$

in which we allow for the possibility that there may be a mull space for this pairing. It is convenient to describe the pairing in terms of a matrix A in which all signs have been reversed. Thus, we can write the signature of A as (p, q, r) for q self-dual field strengths, p anti-self-dual field strengths, and r the dimension of the null space.

Now, in a 6D theory with q self-dual field strengths and p anti-self-dual field strengths, the signature of A is $(p, q, 0)$. For a 6D supergravity theory with T tensor multiplets, the signature is $(T, 1, 0)$. In fact, even more is true in a 6D theory of gravity: diffeomorphism invariance enforces the condition found in [32] that $\det A = -1$.

Now, since we are interested in supersymmetric theories decoupled from gravity we arrive at the *necessary* condition that the signature of A is $(p, 0, r)$. In this special case, each of our two-form potentials has a real scalar superpartner, which we denote as S^I. The kinetic term for these scalars is:

$$\mathcal{L}_{eff} \supset A_{IJ} \partial S^I \partial S^J. \tag{2.2.2}$$

Observe that if A has a zero eigenvector, some linear combinations of the scalars will have a trivial kinetic term. When this occurs, these tensor multiplets define parameters of the effective theory on the tensor branch (i.e. they are non-dynamical fields).

[5] Here we ignore possible torsional contributions to the pairing.

This leaves us with two general possibilities. Either A is positive definite (i.e. $A > 0$), or it is positive semidefinite (i.e. $A \geq 0$). Recall, however, that to reach a 6D SCFT, a necessary condition is that we must also demand that $A > 0$. We summarize the various possibilities for self-consistent 6D theories:

	6D SUGRA	6D LST	6D SCFT
signature	$(T, 1, 0)$	$(p, 0, r)$	$(T, 0, 0)$
$\det A$:	$\det A = -1$	$\det A = 0$	$\det A > 0$

(2.2.3)

For now, we have simply indicated an LST as any theory where $\det A = 0$.

As already mentioned, when $\det A = 0$, some linear combinations of the scalar fields for tensor multiplets will have trivial kinetic term. This means that they are better viewed as defining dimensionful parameters. For example, in the case of a 6D theory with a single gauge group factor and no dynamical tensor multiplets, this parameter is just the overall value $S_{null} = 1/g_{YM}^2$, with g_{YM} the Yang-Mills coupling of a gauge theory. Indeed, this Yang-Mills theory contains solitonic solutions which we can identify with strings:

$$F = - *_4 F, \tag{2.2.4}$$

that is, we dualize in the four directions transverse to an effective string. More generally, we can expect A to contain some general null space, and with each null direction, a non-dynamical tensor multiplet of parameters:

$$\overrightarrow{v}_{null} \equiv N_1 \overrightarrow{v}^1 + ... + N_T \overrightarrow{v}^T \text{ such that } A \cdot \overrightarrow{v}_{null} = 0. \tag{2.2.5}$$

for the two-form potential, and:

$$S_{null} = N_1 S^1 + ... + N_T S^T \tag{2.2.6}$$

for the corresponding linear combination of scalars. Since they specify dimensionful parameters, we get an associated mass scale, which we refer to as M_{string}:

$$S_{null} = M_{string}^2. \tag{2.2.7}$$

Returning to our example from 6D gauge theory, the tension of the solitonic string in equation (2.2.4) is just $1/g_{YM}^2 = M_{string}^2$. At energies above M_{string}, our effective field theory is no longer valid, and we must provide a UV completion.

On general grounds, $A \geq 0$ could have many null directions. However, in the case where where we have a single interacting theory, i.e. when A is simple, there are further strong restrictions. As explained in reference [33], when $A \geq 0$ is simple, all of its minors are positive definite: $A_{minor} > 0$. Consequently, there is precisely *one* zero eigenvalue, and the eigenvector is a positive linear combination of basis vectors. Consequently, there is only one dimensionful parameter M_{string}. This also means that if we delete any tensor multiplet, we reach a positive definite intersection pairing, and consequently, a 6D SCFT. What we have just learned is that if we work in the subspace orthogonal to the ray swept out by S_{null}, then the remaining scalars can all be collapsed to the origin of moduli space. When we do this, we reach the LST limit.

We shall refer to this property of the matrix A as the "tensor-decoupling criterion" for an LST. As we show in subsequent sections, the fact that decoupling any tensor multiplet takes us to an SCFT imposes sharp restrictions.

Even so, our discussion has up to now focussed on some necessary conditions to reach a UV complete theory different from a 6D SCFT. In references [26, 11] the specific case of 6D supersymmetric gauge theories was considered, and closely related consistency conditions for UV completing to an LST were presented. Here, we see the same consistency condition $A \geq 0$ appearing for *any* effective theory with (possibly non-dynamical) tensor multiplets.

Indeed, simply specifying the tensor multiplet content provides an incomplete characterization of the tensor branch. In addition to this, we will also have vector multiplets and hypermultiplets. For theories with only eight real supercharges, anomaly cancellation often imposes tight consistency conditions.

There is, however, an important difference in the way anomaly cancellation operates in a 6D SCFT compared with a 6D LST. The crucial point is that because A has a zero eigenvalue, there is a non-dynamical tensor multiplet which does not participate in the Green-Schwarz mechanism. In other words, on the tensor branch of an LST with T tensor multiplets, at most only $T - 1$ participate. This is not particularly worrisome since as explained in reference [32] and further explored in reference [34], there is in general a difference between the tensor multiplets which participate in anomaly cancellation and those which appear in the tensor branch of a general 6D theory.

Though we have given a number of *necessary* conditions that any putative LST must satisfy, to truly demonstrate their existence we must pass beyond effective field theory, embedding these theories in a UV complete framework such as string theory. We therefore now turn to the F-theory realization of little string theories.

2.3 LSTs from F-theory

In this section we spell out the geometric conditions necessary to realize LSTs in F-theory. Recall that in a little string theory, we are dealing with a 6D theory which contains strings with finite tension. As such, they are an intermediate case between the case of a 6D superconformal field theory (which only contains tensionless strings), and the full string theory (i.e., one in which gravity is dynamical).

Any supersymmetric F-theory compactification to six dimensions is defined by an elliptically fibered Calabi-Yau threefold $X \to B$. Here, X is the total space and B is the base. The elliptic fibration can be described by a local Weierstrass model

$$y^2 = x^3 + fx + g, \tag{2.3.1}$$

where f and g are local functions on B, that globally are sections respectively of $\mathcal{O}_B(-4K_B)$ and $\mathcal{O}_B(-6K_B)$, K_B being the canonical class of B. The discriminant of the elliptic fibration is:

$$\Delta \equiv 4f^3 + 27g^2 \tag{2.3.2}$$

which globally is a section of $\mathcal{O}_B(-12K_B)$. The discriminant locus $\Delta = 0$ is a divisor, and its irreducible components tell us the locations of degenerations of elliptic fibers. Such singularities determine monodromies for the complex structure parameter τ of the elliptic fiber, which is interpreted in type IIB string theory as the axio-dilaton field. In type IIB language, the discriminant locus signals the location of seven-branes in the F-theory model.

In F-theory, decoupling gravity means we will always be dealing with a non-compact base B. When all curves of B are of finite non-zero size, we get a 6D effective theory with a lattice of string charges:

$$\Lambda_{string} = H_2^{cpct}(B, \mathbb{Z}). \tag{2.3.3}$$

The intersection form defines a canonical pairing:

$$A_{intersect} : \Lambda_{string} \times \Lambda_{string} \to \mathbb{Z}, \tag{2.3.4}$$

which we identify with the Dirac pairing:

$$A_{Dirac} = A_{intersect}. \tag{2.3.5}$$

We also introduce the "adjacency matrix"

$$A_{adjacency} = -A_{Dirac}. \tag{2.3.6}$$

To streamline the notation, we shall simply denote the adjacency matrix as A. The two-form potentials of the 6D theory arise from reduction of the four-form potential of type IIB string theory. Additionally, the volumes of the various compact two-cycles translate to the real scalars of tensor multiplets:

$$S^I \propto \mathrm{Vol}(\Sigma_I). \tag{2.3.7}$$

In the F-theory model, the appearance of a null vector for $A_{intersect}$ means that some of these moduli are not dynamical in the 6D effective field theory. Rather, they define dimensionful parameters / mass scales.

To define an F-theory model, we need to ensure that there is an elliptic Calabi-Yau X in which B is the base. A necessary condition for realizing the existence of an elliptic model is that the collection of curves entering in a base B are obtained by gluing together the "non-Higgsable clusters" (NHCs) of reference [35] via $\mathbb{P}^1$'s of self-intersection -1.

Recall that the non-Higgsable clusters are given by collections of up to three $\mathbb{P}^1$'s in which the minimal singular fiber type is dictated by the self-intersection number of the $\mathbb{P}^1$. The self-intersection number, and associated gauge symmetry and matter content are as follows:

Self-intersection	-3	-4	-5	-6	-7	-8
Gauge Theory	$\mathfrak{su}_3$	$\mathfrak{so}_8$	$\mathfrak{f}_4$	$\mathfrak{e}_6$	$\mathfrak{e}_7 \oplus \frac{1}{2}56$	$\mathfrak{e}_7$

(2.3.8)

Self-intersection	-9	-10	-11	-12
Gauge Theory	$\mathfrak{e}_8 \oplus$ 3 inst	$\mathfrak{e}_8 \oplus$ 2 inst	$\mathfrak{e}_8 \oplus$ 1 inst	$\mathfrak{e}_8$

(2.3.9)

Self-intersection	$-3, -2$	$-2, -3, -2$	$-3, -2, -2$
Gauge Theory	$\mathfrak{g}_2 \times \mathfrak{su}_2 \oplus \frac{1}{2}(7+1,2)$	$\mathfrak{su}_2 \times \mathfrak{so}_7 \times \mathfrak{su}_2 \oplus \frac{1}{2}(2,8,1) \oplus \frac{1}{2}(1,8,2)$	$\mathfrak{g}_2 \times \mathfrak{sp}_1 \oplus \frac{1}{2}(7,2) \oplus \frac{1}{2}(1,2)$

(2.3.10)

in addition, we can also consider a single -1 curve, and configurations of -2 curves arranged either in an ADE Dynkin diagram, or its affine extension (in the case of little string theories). The local rules for building up an F-theory base compatible with these NHCs amount to a local gauging condition on the flavor symmetries of a -1 curve: We scan over product subalgebras of the $\mathfrak{e}_8$ flavor symmetry which are also represented by the minimal fiber types of the NHCs. When they exist, we get to "glue" these NHCs together via a -1 curve.

For a general elliptic Calabi-Yau threefold, the curves appearing in a given gluing configuration can lead to rather intricate intersection patterns. For example, two curves may intersect more than once, and may therefore form either a closed loop, or an intersection with some tangency. Additionally, we may have three curves all meeting at the same point, as in the case of the type IV Kodaira fiber. Finally, a single -1 curve may in general intersect more then just two curves. The possible ways to locally glue together such NHCs has also been worked out explicitly in reference [35] (see also [36]). The main idea, however, is that since the -1 curve theory defines a 6D SCFT with E_8 flavor symmetry, we must perform a gluing compatible with gauging some product subalgebra of the Lie algebra $\mathfrak{e}_8$.

What this means in general is that the adjacency matrix provides only a partial characterization of intersecting curves in the base of a geometry. To handle these different possibilities, we therefore introduce the following notation:

$$\text{Normal Intersection: } a, b \text{ or } ab \tag{2.3.11}$$

$$\text{Tangent Intersection: } a||b \tag{2.3.12}$$

$$\text{Triple Intersection: } a \overset{b}{\bigtriangledown} c \tag{2.3.13}$$

$$\text{Looplike configuration: } //a_1 ... a_k//. \tag{2.3.14}$$

Now, decoupling gravity to reach an SCFT or an LST leads to significant restrictions on the possible ways to glue together NHCs. In the case of a 6D SCFT, contractibility of all curves in the base means first, that all of the compact curves are $\mathbb{P}^1$'s, and further, that a -1 curve can intersect at most two other curves. Additionally, all off-diagonal entries of the intersection pairing are either zero or one. In the case of LSTs, however, the curves of the base could include a T^2, and a -1 curve can potentially intersect more than two curves. Additionally, there is also the possibility that the off-diagonal entries of the adjacency matrix may be different than just zero or one.

Again, we stress that the intersection pairing provides only *partial* information. For example, a curve of self-intersection zero could refer either to a $\mathbb{P}^1$, or to a T^2. In the case of a T^2 of self-intersection zero, the normal bundle need not be trivial, but could be a torsion line bundle instead. Additionally, an off-diagonal entry in the adjacency matrix which is two may either refer to a pair of curves which intersect twice, or to a single intersection of higher tangency. The case of tangent intersections violates the condition of normal crossing (which is known to hold for SCFTs [9] but fails for LSTs). An additional type of normal crossing violation appears when we blow down a -1 curve meeting more than two curves. In Appendix A.2 we determine the types of matter localized when there are violations of normal crossing.

2.3.1 Geometry of the Gravity-Decoupling Limit

We now discuss how to obtain limits of F-theory compactifications in which gravity is decoupled, following a program initiated in [37], worked out in detail in [38] (see also [39]), and extended to the case of 6D SCFTs in [40]. For this purpose, we consider F-theory from the perspective of the type IIB string, with the volume of the base B of the F-theory compactification providing a Planck scale for the compactified theory. We will see that the quest for decoupled gravity leads to the same condition on semidefiniteness of the intersection matrix of the compact curves, and moreover we will see how to ensure that the F-theory base B in such cases has a metric of the appropriate kind.

The Case of Compact Base

We begin with the case in which the F-theory base B is a compact surface, and suppose we have a sequence of metrics (specified by their Kähler forms ω_i) which decouple gravity in the limit $i \to \infty$. In particular, the volume must go to infinity: $\lim_{i\to\infty} \mathrm{vol}(\omega_i) = \infty$.

To investigate the geometry of this family of metrics, we temporarily rescale them and consider the Kähler forms

$$\widetilde{\omega}_i := \frac{\omega_i}{\sqrt{\mathrm{vol}(\omega_i)}}. \tag{2.3.15}$$

The rescaled metrics all have volume 1, and since the closure of the set of volume 1 Kähler classes on B is compact, there must be a convergenct subsequence of Kähler classes $[\widetilde{\omega}_{i_j}]$ whose limit

$$[\widetilde{\omega}_\infty] = \lim_{j\to\infty} [\widetilde{\omega}_{i_j}] \tag{2.3.16}$$

lies in the closure of the Kähler cone. If the original sequence was chosen generically, the limit of the rescaled sequence will be an interior point of the Kähler cone, and in this case all areas and volumes grow uniformly as we take the limit of the original sequence ω_i. Gravity decouples, but all other physical quantities measured by areas and volumes approach either zero or infinity, leaving us with a trivial theory.

However, if the rescaled limit (2.3.16) lies on the boundary of the Kähler cone, more interesting things can happen. In favorable circumstances, such as those present in Mori's cone theorem [41] and its generalizations [42], we can form another complex space $\overline{B}$ out of B by identifying pairs of points p and q whenever they are both contained in a curve C whose area vanishes in the limit. There is a holomorphic map $\pi : B \to \overline{B}$ for which all such curves of zero limiting area are contained in fibers $\pi^{-1}(t)$, $t \in \overline{B}$.

As already pointed out in [38], there are two qualitatively different cases: $\overline{B}$ might be a surface or it might be a curve. (It is not possible for $\overline{B}$ to be a point since there are some curves $C \subset B$ whose area does not vanish in the limit.) If $\overline{B}$ is a surface, then the map $\pi : B \to \overline{B}$ contracts some curves to points, and may create singularities in $\overline{B}$. It is widely believed, and has been mathematically proven under certain hypotheses [43, 44], that the limiting metric $\widetilde{\omega}_\infty$ can be interpreted as a metric $\omega_{\overline{B}}$ on the smooth part of $\overline{B}$.

On the other hand, if $\overline{B}$ is a curve, so that π has curves Σ_t, $t \in \overline{B}$ as fibers, then we again expect the limiting metric $\widetilde{\omega}_\infty$ to be induced by a metric on $\overline{B}$, although there are fewer mathematical theorems covering this case. (See [45] for one known theorem of this kind.)

In general, we do not expect the curves contracted by π to necessarily have zero area in the gravity-decoupling limit. This can be achieved by starting with a reference Kähler form ω_0 on B as well as a (possibly degenerate) Kähler form $\omega_{\overline{B}}$ on $\overline{B}$, and constructing a limit of the form

$$\lim_{t\to\infty} \left(\omega_0 + t\pi^*(\omega_{\overline{B}})\right). \tag{2.3.17}$$

In the case of an SCFT, we wish all curves contracted by π to be at zero area in the limit, so in that case we should omit ω_0 and simply scale up $\pi^*(\omega_{\overline{B}})$).

The Case of Non-Compact Base

Our discussion of the compact bases makes it clear that the decoupling limit only depends on the metric in a (finite volume) neighborhood of a collection of curves on the original F-theory base B, together with a rescaling which takes that neighborhood to infinite volume and smooths out its features in the process. This analysis can be applied to an arbitrary base, compact or non-compact.

If the collection of curves is disconnected, the corresponding points on $\overline{B}$ to which the collection is mapped will be moved infinitely far apart during the rescaling process, thus leading to several decoupled quantum theories. So to study a single theory, it suffices to consider a connected configuration. To reiterate the two cases we have found:

1. We may have a connected collection of curves Σ_j which can be simultaneously contracted to a singular point on a space $\overline{B}$. (The contractibility implies that the intersection matrix is negative definite.) When the metric on $\overline{B}$ is rescaled, gravity is decoupled giving a 6D SCFT (in which every curve in the collection remains at zero area).

   Alternatively, we can combine this rescaled metric with another reference metric which provides finite area to each Σ_j. This produces a quantum field theory in the Coulomb branch of the SCFT.

2. Or we may have a connected collection of curves Σ_j which are all contained in a single fiber of a map $\pi : B \to \overline{B}$, and include all components of that fiber. (This implies that the intersection matrix is negative semidefinite, with a one-dimensional zero eigenspace.) We combine a reference metric on B that sets the areas of the individual Σ_j's with a metric on $\overline{B}$ which is rescaled to decouple gravity, yielding a little string theory (with the string provided by a D3-brane wrapping the entire fiber). The overall area of the fibers of π sets the string scale, and the possible areas of the Σ_j's map out the moduli space of the theory. Gravity is decoupled, and we find an LST.

Note that in the second case, there are two distinct possibilities for the fibers of the map π: the general fiber can be a curve of genus 0 or a curve of genus 1. In the case of genus 1, it is possible for the central fiber to have a nontrivial multiplicity, that is, the fiber can take the form $m\Sigma$ for some $m > 1$.

2.4 Examples of LSTs

In the previous section we gave the general rules for constructing LSTs. Our plan in this section will be to show how the F-theory realization allows us to recover well-known examples of LSTs previously encountered in the literature.

To this end, we begin by first showing how LSTs with sixteen supercharges arise in F-theory constructions. After this, we turn to known constructions of LSTs with eight supercharges (i.e. minimal supersymmetry). This will also serve to illustrate how F-theory provides a single coherent framework for realizing LSTs.

2.4.1 Theories with Sixteen Supercharges

To set the stage, we begin with little string theories with sixteen supercharges. In this case, we have two possibilities given by $\mathcal{N} = (2,0)$ supersymmetry or $\mathcal{N} = (1,1)$ supersymmetry. Note that only the former is possible in the context of 6D SCFTs.

One way to generate examples of $\mathcal{N} = (2,0)$ LSTs is to take k M5-branes filling $\mathbb{R}^{5,1}$ and probing the geometry $S^1_\perp \times \mathbb{C}^2$ with $S^1_\perp$ a transverse circle of radius R. To reach the gravity-decoupling limit for an LST we simultaneously send the radius $R \to 0$ and $M_{pl} \to \infty$ whilst holding fixed the effective string scale. In this case, it is the effective tension of an M2-brane wrapped over the circle which we need to keep fixed. Performing a reduction along this circle, we indeed reach type IIA string theory with k NS5-branes. By a similar token, we can also consider IIB string theory with k NS5-branes. This realizes LSTs with $\mathcal{N} = (1,1)$ supersymmetry.

T-dualizing the k NS5-branes of type IIA, we obtain type IIB string theory on the local geometry given by a configuration of -2 curves arranged in the affine $\widehat{A}_{k-1}$ Dynkin diagram. Similarly, we also get an LST by taking type IIA string theory on the same geometry.

Consider next the F-theory realization of these little string theories. First of all, we reach the aforementioned theories by working with F-theory models whose associated Calabi–Yau threefold takes the form $T^2 \times S$, in which $S \to C$ is an elliptically fibered (non-compact) Calabi-Yau surface. If we treat the T^2 factor as the elliptic fiber of F-theory, we get IIB on S, and if we treat the elliptic fiber of S as the elliptic fiber of F-theory, we get (after shrinking the T^2 factor to small size) F-theory with base $T^2 \times C$, which is dual to IIA on S. To refer to both cases, it will be helpful to label the auxiliary elliptic curve T^2 as T^2_F (for fiber) and the other elliptic curve as T^2_S (since it lies in the surface S).

Now, by allowing S to develop a singular elliptic fiber, we can realize the same local geometries obtained perturbatively. For example, the $\mathbb{C}^2/\mathbb{Z}_k$ lifts to a Kodaira fiber of type I_k. Resolving this local singularity, we find k compact cycles $\Sigma_I \simeq \mathbb{P}^1$'s which intersect according to the affine $\widehat{A}_{k-1}$ Dynkin diagram. In this case the null divisor class is:

$$[\Sigma_{null}] = [\Sigma_1] + ... + [\Sigma_k] \tag{2.4.1}$$

that is, it is the ordinary minimal imaginary root of $\widehat{A}_{k-1}$. By shrinking T^2_F to small size, this engineers in F-theory the $\mathcal{N} = (2,0)$ LST of k M5-branes, or of k NS5-branes in type IIA.

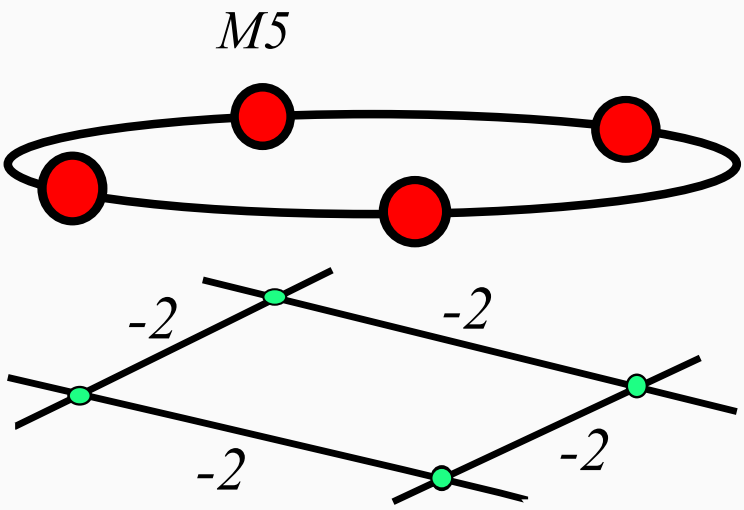


Figure 2.4.1: Depiction of the tensor branch of the $\mathcal{N} = (2,0)$ $\widehat{A}_3$ LST. In the top figure, we engineer this example using spacetime filling M5-branes probing the geometry $S^1_\perp \times \mathbb{C}^2$. In the dual F-theory realization, we have four -2 curves in the base, which are arranged as the affine $\widehat{A}_3$ Dynkin diagram. The Kähler class of each -2 curve in the F-theory realization corresponds in the M-theory realization to the relative separation between the M5-branes.

(See figure 2.4.1 for a depiction of the A-type $\mathcal{N} = (2,0)$ LSTs.) In the other case, one obtains F-theory on $T^2 \times C$ whose fibers have an I_k singularity along $T^2 \times \{0\}$. Then $[\Sigma_{null}]$ is precisely the class of the F-theory fiber T^2_S, and supersymmetry enhances to $\mathcal{N} = (1,1)$.

More generally, we can consider any of the degenerations of the elliptic fibration classified by Kodaira, i.e. the type $I_n, II, III, IV, I_n^*, II^*, III^*, IV^*$ fibers and produce a $\mathcal{N} = (2,0)$ model with that degeneration occuring as a curve configuration on the F-theory base, as well as a $\mathcal{N} = (1,1)$ model with that same degeneration occurring as the F-theory fiber over some T^2 in the base.

As a brief aside, a convenient way to realize examples of both the $\mathcal{N} = (1,1)$ and $\mathcal{N} = (2,0)$ theories is to consider F-theory on the Schoen Calabi-Yau threefold $dP_9 \times_{\mathbb{P}^1} dP_9$ [46]. Then, we can keep the elliptic fiber on one dP_9 factor generic, and allow the other to degenerate. Switching the roles of the two fibers then moves us from the IIA to IIB case. Note that although this strictly speaking only yields eight real supercharges (as we are on a Calabi-Yau threefold), in the rigid limit used to reach the little string theory, we expect a further enhancement to either $\mathcal{N} = (2,0)$ or $\mathcal{N} = (1,1)$ supersymmetry. The specific chirality of the supersymmetries depends on which elliptic curve we take to be in the base, and which to be in the fiber of the corresponding F-theory compactification.

Let us also address whether each of the different Kodaira fiber types leads us to a different little string theory. Indeed, some pairs of Kodaira fiber types lead to identical gauge symmetries in the effective field theory. To illustrate, consider the type IV Kodaira fiber, and compare it with

the type I_3 fiber. There is a complex structure deformation which moves the triple intersection appearing in the type IV case out to the more generic type I_3 case. This modulus, however, is decoupled from the 6D little string theory. The reason is that if we consider a further compactification on a circle, we reach a 5D gauge theory which is the same for both fiber types. The additional complex structure modulus from deforming IV to I_3 does not couple to any of the modes of the 5D theory. So, there does not appear to be any difference between these theories. In other words, we should classify all of the $\mathcal{N} = (2,0)$ little string theories in terms of affine ADE Dynkin diagrams rather than in terms of Kodaira fiber types.

For the $\mathcal{N} = (1,1)$ little string theories, the absence of a chiral structure actually leads to more possibilities. For example, if we consider M-theory on an ADE singularity compactified on a further circle, we have the option of twisting by an outer automorphism of the simply laced ADE Lie algebra [47]. In other words, for the $\mathcal{N} = (1,1)$ theories we have an ABCDEFG classification according to all of the simple Lie algebras.

To realize these LSTs in F-theory, we make an orbifold of the previous construction. Suppose that $S \to C$ is an elliptically fibered (non-compact) Calabi–Yau surface which has compatible actions of $\mathbb{Z}_m$ on the base and on the total space, such that the action on the total space preserves the holomorphic 2-form. Then $\mathbb{Z}_m$ acts on $T^2_F \times S$ with the action on T^2 being translation by a point of order m. The quotient $X := (T^2_F \times S)/\mathbb{Z}_m$ is then an elliptically fibered Calabi–Yau threefold (with two genus one fibrations as before).

The elliptic fibration $X \to (T^2_F \times C)/\mathbb{Z}_m$ leads to an F-theory model with $\mathcal{N} = (1,1)$ supersymmetry. Note that the base $(T^2_F \times C)/\mathbb{Z}_m$ of the F-theory fibration contains a curve Σ of genus 1 and self-intersection 0 such that $m\Sigma$ can be deformed into a one-parameter family although no smaller multiple can be deformed. Note also that if the action of $\mathbb{Z}_m$ on S preserves the section of the fibration $S \to C$, then $X \to (T^2_F \times C)/\mathbb{Z}_m$ also has a section and, as we will explain in section 2.7.1, $m \in \{2,3,4,6\}$ since every elliptic fibration with section has a Weierstrass model [48].

There is a second fibration $X \to (S/\mathbb{Z}_m)$ which is a genus one fibration without a section and leads to theories with $\mathcal{N} = (2,0)$ supersymmetry. We discuss additional details about this second fibration, as well as T-duality for these theories, in section 2.9.

It is instructive to study the structure of the moduli space of the LSTs with maximal supersymmetry. Recall that the tensor branch for a $\mathcal{N} = (2,0)$ SCFT of ADE type $\mathfrak{g}$ is given by:

$$\mathcal{M}_{(2,0)}[\mathfrak{g}] = \mathbb{R}^T_{\geq 0}/\mathcal{W}_{\mathfrak{g}}, \tag{2.4.2}$$

where in the above, T is the number of tensor multiplets and $\mathcal{W}_{\mathfrak{g}}$ is the Weyl group of the ADE Lie algebra $\mathfrak{g}$. That is, the moduli space is given by a Weyl chamber of the ADE Lie algebra and is therefore non-compact. In the present case of LSTs, we see that the condition that we have a string scale leads to one further constraint on this moduli space, effectively "compactifying" it to the compact Coxeter box for an affine root lattice [17].

Finally, one of the prominent features of these examples is the manifest appearance of two elliptic fibrations in the geometry. Indeed, in passing from the $\mathcal{N} = (2,0)$ theories to the $\mathcal{N} = (1,1)$ theories, we observe that we have simply switched the role of the two fibrations. In section 2.9, we return to this general phenomenon for how T-duality of LSTs is realized in F-theory.

2.4.2 Theories with Eight Supercharges

Several examples of LSTs with minimal, i.e. $(1,0)$ supersymmetry are realized by mild generalizations of the examples reviewed above.

To begin, let us consider again the case of k coincident M5-branes filling $\mathbb{R}^{5,1}$ and probing the geometry $S^1_\perp \times \mathbb{C}^2$. We arrive at a $(1,0)$ LST by instead taking a quotient of the $\mathbb{C}^2$ factor by a non-trivial discrete subgroup $\Gamma_G \subset SU(2)$ so that the geometry probed by the M5-brane is $\mathbb{C}^2/\Gamma_G$. The discrete subgroups admit an ADE classification, and the corresponding simple Lie group G_{ADE} specifies the gauge group factors on the tensor branch. We reach a 6D SCFT by decompactifying the $S^1_\perp$. In this limit, we have an emergent $G_L \times G_R$ flavor symmetry. From this perspective, the little string theory arises from gauging a diagonal G subgroup of the flavor symmetry. In the IIB realization of NS5-branes probing the affine geometry, applying S-duality takes us to a stack of D5-branes probing an ADE singularity. On its tensor branch, this leads to an affine quiver gauge theory [49].

The F-theory realization of these LSTs is simply an affine A-type Dynkin diagram of k curves of self-intersection -2 decorated with $I_n, I_n^*, IV^*, III^*, II^*$ fibers, respectively for $G = A_{n-1}, D_{n+4}, E_{6,7,8}$. We reach a 6D SCFT by decompactifying any of the -2 curves in the loop, and we recover a 6D SCFT with an emergent $G_L \times G_R$ flavor symmetry. From this perspective, the little string theory arises from gauging a diagonal G subgroup of the flavor symmetry. Note that for $G \neq A_n$, all these systems involve conformal matter in the sense of reference [28].

Another class of LSTs is given by taking k M5-branes in heterotic M-theory, i.e. M-theory on $S^1/\mathbb{Z}_2 \times \mathbb{C}^2$. In this case, we have two E_8 flavor symmetry factors; one for each endpoint of the interval $S^1/\mathbb{Z}_2$. In this case, the gravity-decoupling limit requires us to collapse the size of the interval to zero size (i.e. to reach perturbative heterotic strings), whilst still holding the effective string scale finite. (The ratios of the lengths of subintervals between the endpoints and the various M5-branes to the length of the total interval will remain finite in the gravity-decoupling limit and provide parameters for the tensor branch.) In perturbative heterotic string theory, we have k NS5-branes probing $\mathbb{C}^2$. A related example is provided by instead working with the $Spin(32)/\mathbb{Z}_2$ heterotic string in the presence of k NS5-branes. Indeed, once suitable Wilson line data has been specified, these two examples are T-dual to one another.

The F-theory realization of the theory of k M5-branes is given by a non-compact base with a configuration of curves:

$$[E_8]\underbrace{1,2,...,2,1}_{k}[E_8] \tag{2.4.3}$$

where we have indicated the flavor symmetry factors in square brackets. In this configuration, we reach the LST limit by holding fixed the volume of the null divisor (given by a sum over each divisor with multiplicity one), and collapse all other Kähler moduli to zero size. The construction of the T-dual characterization is somewhat more involved, and so we defer a full discussion to section 2.9 and Appendix A.6. See figure 2.4.2 for a depiction of the M-theory and F-theory realizations of this LST.

We can also combine the effects of different orbifold group actions. For example, we can consider k M5-branes filling $\mathbb{R}^{5,1}$ and probing the geometry $S^1/\mathbb{Z}_2 \times \mathbb{C}^2/\Gamma_G$. In F-theory terms,

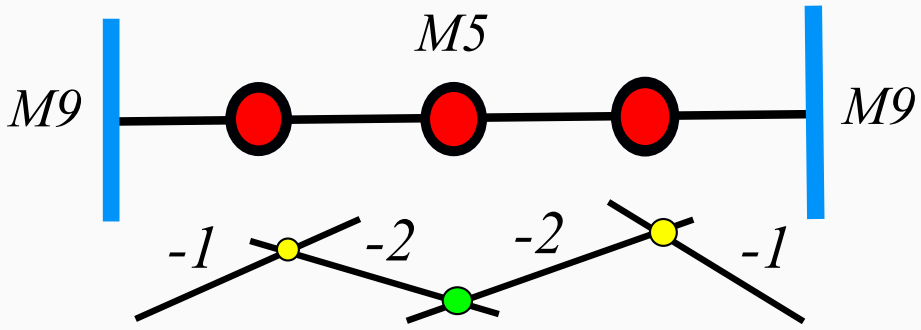


Figure 2.4.2: TOP: Depiction of the LST realized by k M5-branes in between the two Horava-Witten nine-brane walls of heterotic M-theory ($k = 3$ above). This leads to an LST with an $E_8 \times E_8$ flavor symmetry BOTTOM: The Corresponding F-theory base given by the configuration of curves $[E_8], 1, 2, ..., 2, 1, [E_8]$ for k total compact curves. In this realization, the E_8 flavor symmetry is localized on two non-compact 7-branes, one intersecting each -1 curve.

this is given by the geometry:

$$[E_8]\underbrace{\overset{\mathfrak{g}}{1}, \overset{\mathfrak{g}}{2}, ..., \overset{\mathfrak{g}}{2}, \overset{\mathfrak{g}}{1}}_{k}[E_8] \tag{2.4.4}$$

i.e. we decorate by a $\mathfrak{g}$-type ADE gauge symmetry over each curve of self-intersection -1 or -2. This geometry was studied in detail in reference [15]. Further blowups in the base are needed for all fibers to remain in Kodaira-Tate form. This leads to conformal matter between each simply laced gauge group factor [28].

Summarizing, we have seen in the above that the various LSTs which have been constructed via perturbative string theory and M-theory all have a natural embedding in the context of specific F-theory constructions. With this in mind, we now turn to a systematic construction of all 6D LSTs in F-theory.

2.5 Constraints from Tensor-Decoupling

As a first step towards the classification of LSTs, we now show how to classify possible bases using the "tensor-decoupling criterion," that is, the requirement that decoupling any tensor multiplet from an LST must take us to an SCFT. In geometric terms, deleting any curve of the base (with possible fiber enhancements along this curve) must take us back to an SCFT base (with possibly disconnected components). Since all SCFTs have the structure of a tree-like graph of intersecting curves [10], our task reduces to scanning over the list of *connected* SCFTs, and asking whether adding an additional curve (with possible fiber enhancements on this curve) will produce an LST. This inductive approach to classification will allow us to effectively constrain the overall structure of bases for LSTs.

In this section we show how the tensor-decoupling criterion constrains many candidate bases for LSTs. We first use this criterion to limit the possible graph topologies of curves in the base. Next, we give a general inductive rule for how to take an SCFT and verify whether it enhances to an LST. We shall refer to this as an inductive classification, since it implicitly accounts for all possible structures for LSTs. In section 2.6 we use these constraints to present a more explicit construction of possible bases for LSTs.

2.5.1 Graph Topologies for LSTs

For any compact curve Σ in the base which remains in the gravity-decoupling limit, the self-intersection Σ^2 must be $-n$ for $0 \leq n \leq 12$. Moreover, since having an F-theory model requires that $-4K$, $-6K$, and $-12K$ be effective divisors, if $K \cdot \Sigma + \Sigma^2 > 0$ (so that $K \cdot \Sigma > 0$), then $-4K$ would have multiplicity at least 4 along Σ, $-6K$ would have multiplicity at least 6 along Σ, and $-12K$ would have multiplicity at least 12 along Σ. Since this is not allowed in the Kodaira classification, we conclude that $2g - 2 = K \cdot \Sigma + \Sigma^2 \leq 0$, in other words, that Σ is either $\mathbb{P}^1$ or T^2. We now use the tensor-decoupling criterion to argue that the possible topologies of LST bases are limited to tree-like structures and appropriate degenerations of an elliptic curve.

Let us first show that a curve Σ of self-intersection zero (of topology $\mathbb{P}^1$ or T^2) can only appear in isolation, i.e. it cannot intersect any other curve. If it met another curve and we decoupled everything that this curve touches, we would be left with an SCFT base containing a curve of self-intersection zero, a contradiction. If Σ has genus 0, then the base takes the form $\mathbb{C} \times \Sigma$, while if Σ has genus 1, then the base takes the form $(\mathbb{C} \times \Sigma)/\mathbb{Z}_m$, with $\mathbb{Z}_m$ acting on Σ by a translation and on $\mathbb{C}$ by multiplication by a root of unity. Note that if either $g = 0$ or $m = 1$, the base is just a product. Hence, to get a six-dimensional theory we must wrap seven-branes over Σ, i.e. we *must* include a non-trivial fiber enhancement over this curve, unless $g = 1$ and $m > 1$. (We will see examples of this latter case in section 2.9.)

Consider next adjacency matrices in which the off-diagonal entries are different from zero or one. For example, this can occur when a -4 and -1 curve form a closed loop (i.e. intersect twice), or when the same curves intersect along a higher order tangency. Again, this possibility is severely limited because if this were to a occur in a configuration with three or more curves, we would contradict the tensor-decoupling criterion. By the same token, the value of all off-diagonal entries are bounded below by two:

$$-2 \leq A_{IJ} \leq 0 \text{ for } I \neq J, \tag{2.5.1}$$

and in the case where -2 appears, we are limited to just two curves. The only possibilities for a rank one LST base (i.e. with two curves) are therefore:

$$1,1 \text{ or } //2,2// \text{ or } //4,1// \text{ or } 2||2 \text{ or } 4||1. \tag{2.5.2}$$

In Appendix A.2 we analyze the possible fiber enhancements which can occur when the two curves meet along a tangency (i.e. do not respect normal crossing), as is the case in the last two configurations.

For all other LST bases, we see that all curves must be constructed from $\mathbb{P}^1$'s of self-intersection $-x$ for $1 \leq x \leq 12$, which all intersect with normal crossings, i.e. all off-diagonal entries of the adjacency matrix are either zero or one.

To further constrain the structure, we next observe that the base of any 6D SCFT is always tree-like [9]. This means that the graph associated to an LST adjacency matrix can admit at most one loop, and when it contains a loop, there can be no additional curves branching off. This is because the tensor-decoupling criterion would be violated by joining a loop of curves to anything else. We are therefore left with two general types of configurations:

- Tree-like LSTs
- Loop-like LSTs.

Note that some of the tree-like structures we shall encounter can also be viewed as loops, that is, as degenerations of an elliptic curve.

2.5.2 Inductive Classification

To proceed further, we now present an inductive strategy for constructing the base of any LST with three or more curves. The main idea is that we simply need to sweep over the list of SCFT bases and ask whether we can append an additional curve of self-intersection $-x$ to such a base. By the remarks on decoupling already noted, we see that this additional curve can intersect either one curve or two curves of an SCFT base. In the latter case, we obtain a loop-like configuration of curves in the base. The latter possibility can only occur for an SCFT base which consists of a single line of curves (i.e. no branches emanating off of the primary spine of the base). The main condition we need to check is that after adding this curve, we obtain a positive semidefinite adjacency matrix. In particular, the determinant must vanish. Implicit in this construction is that we only append an additional curve compatible with the gluing rules for bases.

Consider first the case of an LST with adjacency matrix A_{LST} which describes a tree-like base given by adding a single curve of self-intersection $-x$ to some SCFT with adjacency matrix A_{SCFT}:

$$A_{LST}^{tree} = \begin{pmatrix} y & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & & & & & & & \\ 0 & & & & A_{SCFT} & & & \\ \vdots & & & & & & & \vdots \\ 0 & & & & & & & \end{pmatrix} \tag{2.5.3}$$

Let $A_{SCFT'}$ be the principal submatrix for the set of nodes $\{3, 4, ..., N\}$. Evaluating the determinant of A_{LST}^{tree}, we obtain the condition:

$$0 = \det(A_{LST}^{tree}) = y \det(A_{SCFT}) - \det(A_{SCFT'}). \tag{2.5.4}$$

or:

$$y = \det(A_{SCFT'})/\det(A_{SCFT}). \tag{2.5.5}$$

SCFT	12	13	14	15	16	17	18	19	1(10)	1(11)	1(12)	22	23
det(A)	1	2	3	4	5	6	7	8	9	10	11	3	5
y	5	3	7/3	2	9/5	5/3	11/7	3/2	13/9	7/5	15/11	2	7/5

Table 2.5.1: Candidate loop-like rank two LSTs from adding an additional curve to a rank two SCFT.

Consider now the case of a loop-like LST. In this case, the only SCFTs we need consider are those constructed from a single line of curves (i.e. no trivalent vertices at all), and we can only add the additional curve to the leftmost and rightmost ends of a candidate SCFT. The adjacency matrix is then of the form:

$$A_{LST}^{loop} = \begin{pmatrix} y & -1 & 0 & 0 & \dots & 0 & 0 & -1 \\ -1 & & & & & & & \\ 0 & & & & & & & \\ \vdots & & & & A_{SCFT} & & & \vdots \\ 0 & & & & & & & \\ -1 & & & & & & & \end{pmatrix} \tag{2.5.6}$$

To have an LST we must have

$$0 = y \det(A_{SCFT}) - (A_{SCFT})_{(1,1)} - (A_{SCFT})_{(N-1,N-1)} + 2(-1)^{N+1}(A_{SCFT})_{(1,N-1)} \tag{2.5.7}$$

where we have denoted the $(i,j)^{th}$ minor of A_{SCFT} by an appropriate subscript. Solving for y, we obtain:

$$y = \frac{(A_{SCFT})_{(1,1)} + (A_{SCFT})_{(N-1,N-1)} - 2(-1)^{N+1}(A_{SCFT})_{(1,N-1)}}{\det(A_{SCFT})}. \tag{2.5.8}$$

The above algorithm allows us to systematically classify LSTs: From this structure, we see that the locations of where we can add an additional curve to an existing SCFT are quite constrained. Indeed, in order to *not* produce another SCFT, but instead an LST, we will typically only be able to add our extra curve at the end of a configuration of curves, or at the second to last curve. Otherwise, we could not reach an SCFT upon decoupling other curves in the base.

2.5.3 Low Rank Examples

To illustrate how the algorithm works in practice, we now give some low rank examples. In table 2.5.1 we list all of the rank two SCFT bases which we attempt to enhance to loop-like LST. Of the cases where y is an integer, some are further eliminated since the resulting base requires

further blowups.[6] The full list of rank two LST bases is then:

$$\begin{array}{c}\text{Three curve LST bases:}\\ 2\overset{2}{\bigtriangledown}2 \text{ and } 121 \text{ and } 212 \text{ and } //222//\end{array} \tag{2.5.9}$$

where the first entry denotes a triple intersection of -2 curves (that is, a type IV Kodaira degeneration), and $//x_1x_2...x_{n+1}//$ denotes a loop in which the two sides are identified.

2.6 Atomic Classification of Bases

In principle, the remarks of the previous section provide an implicit way to characterize all LSTs. Indeed, we simply need to sweep over the list of bases for SCFTs obtained in reference [10] and then determine whether there is any place to add one additional curve to reach an LST. The self-intersection of this new curve is constrained by the condition that the determinant of the adjacency matrix vanishes, and the location of where we add this curve is likewise constrained by the tensor-decoupling criterion.

In this section we use the atomic classification of 6D SCFTs presented in [10] to perform a corresponding atomic classification of bases for LSTs. We now use the explicit structure of 6D SCFTs found in [10] to further cut down the possibilities. It is helpful to view the bases as built out of smaller "atoms" and "radicals". In particular, we introduce the convention of a "node" referring to a single curve in which the minimal fiber type leads to a D or E-type gauge algebra. We refer to a "link" as any collection of curves which does not contain any D or E-type gauge algebras for the minimal fiber type. The results of [10] amount to a classification of all possible links, as well as all possible ways of attaching links to the nodes. Quite remarkably, the general structure of the resulting bases is quite constrained. For all 6D SCFTs, we can filter the theories according to the number of nodes in the graph. These nodes are always arranged along a single line joined by links:

$$S_{0,1}\overset{S_1}{g_1}L_{12}\overset{\mathbf{I}^{\oplus s}}{g_2}L_{2,3}g_3...g_{k-2}L_{k-2,k-1}\overset{\mathbf{I}^{\oplus t}}{g_{k-1}}L_{k-1,k}\overset{\mathbf{I}^{\oplus u}}{g_k}S_{k,k+1}, \tag{2.6.1}$$

here, the g_i's denote the nodes, the $L_{i,i+1}$'s denote interior links (since they join to two nodes) and the S's are side links as they can only join to one node. The notation $\mathbf{I}^{\oplus m}$ refers to decorating by m small instantons, these are further classified according to partitions of m (i.e. how many of the small instantons are coincident with one another). One of the key points is that for $k \geq 6$, there is no decoration on any of the interior nodes, i.e. for $3 \leq i \leq k-2$. This holds both for the types of links which can attach to these nodes (which are always the minimal ones forced by the resolution algorithm of reference [9]), as well as the possible fiber enhancements (there are none). When $k = 5$, it is possible to decorate the middle node g_3 by a single -1 curve. In reference [10], the explicit form of all such sequences of g's, as well as the possible side links and minimal links was classified. An additional important property is that all of the interior links blow down to a trivial endpoint, the blowdown of a single -1 curve.

[6] For example, configurations such as $//512//$ and $//313//$ require a further blowup. Doing this, we instead reach a four curve LST base, respectively given by $//6131//$ and $//4141//$.

Turning now to LSTs, we can ask whether we can add one more curve to the base quiver, resulting in yet another tree-like graph, or in a loop-like graph. By inspection, we can either add this additional curve to a side link, an interior link, or a base node.

Restrictions on Loop-like Graphs: In fact, a general loop-like graph which is an LST is tightly constrained by the tensor-decoupling criterion. The reason is that if we consider the resulting sequence of nodes, we must have a pattern of the form:

$$//g_1 L_{12} g_2 L_{2,3} g_3 ... g_{k-2} L_{k-2,k-1} g_{k-1} L_{k-1,k} g_k L_{k,1} // \tag{2.6.2}$$

where the notation "//" indicates that the left and the right of the base quiver are joined together to form a loop. Now, another important constraint from reference [10] is that the minimal fiber type on the nodes obeys a nested sequence of containment relations. But in a loop, no such ordering is possible. We therefore conclude that all of the nodes for a loop-like LST must be identical, and moreover, that the interior links must all be minimal. We therefore can specify all such loops simply by the type of node (i.e. a -4 curve, a -6 curve, a -8 curve or a -12 curve), and the number of such nodes.

For this reason, we now confine our attention to tree-like graphs, i.e. where we add an additional curve which intersects only one other curve in the base. The main restriction we now derive is that the resulting configuration of curves is basically the same as that of line (2.6.1). Indeed, we will simply need to impose further restrictions on the possible side links and sequences of nodes which can appear in an LST base.

Restrictions on Adding to Interior Links: Our first claim is that we can only possibly add an extra curve to an interior link in a base with two or fewer nodes. Indeed, suppose to the contrary. Then, we will encounter a configuration such as:

$$\overset{y}{...g_i L_{i,i+1} g_{i+1}...} \tag{2.6.3}$$

where y denotes our additional curve attached in some way to the link. The notation "..." denotes the fact that there is at least one more curve in the base. Now, since the interior link blows down to a single -1 curve, we will get a violation of normal crossing. This is problematic if we have one additional curve (as denoted by the "..."), since deleting that curve would produce a putative SCFT with a violation of normal crossing, a contradiction. By the same token, in a two node base, if any side links are attached to this node, then we cannot add anything to the interior link. This leaves us with the case of just:

$$\overset{y}{g_1 L_{1,2} g_2}. \tag{2.6.4}$$

In this case, it is helpful to simply enumerate once again all of the possible interior links, and ask whether we can attach an additional curve. This we do in Appendix A.4, finding that the options are severely limited. Summarizing, then, we find that we can attach an extra curve to an interior link only in the case where there are two nodes, and then only if these two nodes do not attach to any side links.

Restrictions on Adding to Nodes: Let us next turn to restrictions on adding an extra curve to a node in a base. If we add a curve to a node, we observe that this extra curve must have self-intersection -1. Note that the endpoint of the SCFT must therefore be trivial in these cases. We now ask which of the nodes of the base can support an additional -1 curve. Since we must be able to delete a single curve and reach a collection of SCFTs, we cannot place this -1 curve too far into the interior of the configuration. More precisely, we see that for $k \geq 7$ nodes, we are limited to adding a -1 curve to the first three, or last three nodes. In the specific case where we attach a -1 curve to the third interior node, we see that there cannot be any side links whatsoever. Otherwise, we would find a subconfiguration of curves which is not a 6D SCFT.

Restrictions on Adding to Side Links: Consider next restrictions on adding an extra curve to a side link. In the case of a small instanton link such as $1, 2..., 2$, we can append an additional -1 curve to the rightmost -2 curve, but then it can no longer function as a side link (via the tensor-decoupling criterion). In Appendices D and E we determine the full list of LSTs comprised of just adding one more curve to a side link. If we instead attempt to take an existing SCFT and add an additional curve to a side link to reach an LST, then we either produce a new side link (i.e. if the curve has self-intersection $-1, -2, -3$ or -5), or we produce a base quiver with one additional node (i.e. if the curve has self-intersection $-4, -6, -7, -8, -9, -10, -11, -12$). Phrased in this way, we see that the rules for which side links can join to an SCFT are slightly different, but cannot alter the overall topology of a base quiver from the case of an SCFT.

Summarizing, we see that unless we have precisely two nodes, and no side links, we cannot decorate any interior link. Moreover, we can only decorate the three leftmost and rightmost node in special circumstances. So in other words, the general structure of a tree-like LST base is essentially the same as that of a certain class of SCFTs. All that remains is for us to determine the possible sequences of nodes (with no decorations) which can generate an LST, and to also determine which of our side links can be attached to an SCFT such that the resulting configuration is an LST.

Overview of Appendices: This final point is addressed in a set of Appendices. In the appendices we collect a full list of the building blocks for constructing LSTs. The tensor-decoupling criterion prevents a direct gluing of smaller LSTs to reach another LST. Rather, we are always supplementing an SCFT to reach an LST. Along these lines, in Appendix A.3 we collect the list of bases which are comprised of a single spine of nodes with no further decoration from side links. In Appendix A.4 we collect the full list of bases in which no nodes appear. Borrowing from the terminology used for 6D SCFTs, these links are "noble" in the sense that they cannot attach to anything else in the base. Finally, in Appendix A.5 we give a list of LSTs given by attaching a single side link to a single node. Much as in the classification of 6D SCFTs, the further task of sweeping over all possible ways to decorate a base quiver by side links is left implicit (as dictated by the number of blowdowns induced by a given side link). All of these rules follow directly from reference [10].

This completes the classification of bases for LSTs. We now turn to the classification of elliptic fibrations over a given base.

2.7 Classifying Fibers

Holding fixed the choice of base, we now ask whether we can enhance the singularities over curves of the base whilst keeping all fibers in Kodaira-Tate form. As this is a purely local question (i.e. compatible with the matter enhancements over the neighboring curves), most of the rules for adding extra gauge groups / matter are fully specified by the rules spelled out in reference [10]. Rather than repeat this discussion, we refer the interested reader to these cases for further discussion of the "standard" fiber enhancement rules for curves which intersect with normal crossings.

There are, however, a few cases which cannot be understood using just the SCFT considerations of reference [10]. Indeed, we have already seen that a curve of self-intersection zero, an elliptic curve, tangent intersections and triple intersections of curves can all occur in the base of an LST. We have also seen, however, that all of these cases are comparatively "rare" in the sense that they do not attach to larger structures. Our plan in this section will therefore be to deal with all of these low rank examples. In Appendix A.1 we give general constraints from anomaly cancellation in F-theory models and in Appendix A.2 we present some additional technical material on the localization of matter in the case of tangent intersections such as the $4||1$ and $2||2$ configurations. Finally, compared with the case of 6D SCFTs, the available fiber enhancements over a given base are also comparatively rare. To illustrate this point, we give some examples in which the base is an affine Dynkin diagram of -2 curves. In these cases, the presence of the additional imaginary root (and the constraints from anomaly cancellation) typically dictate a small class of possible fiber enhancements.

2.7.1 Low Rank LSTs

In this subsection we give a complete characterization of fiber enhancements for low rank LSTs. To begin, we consider the case of the rank zero LSTs, i.e. those where the F-theory base consists of a single compact curve. In these cases, we *only* get a 6D theory once we wrap some 7-branes over the curve unless the normal bundle of the curve is a torsion line bundle. An interesting feature of this and related examples is that because the corresponding tensor multiplet is non-dynamical it cannot participate in the Green-Schwarz mechanism and we must cancel the anomaly using just the content of the gauge theory sector. We then turn to the other low rank examples where other violations of normal crossing appear. In all of these cases, the F-theory geometry provides a systematic tool for determining which of these structures can embed in a UV complete LST.

Rank Zero LSTs

In a rank zero LST, we have a single compact curve, which must necessarily have self-intersection zero. There are only a few inequivalent configurations consisting of a single curve with self-intersection zero:

$$\Sigma_{null} = P, I_0, I_1, II, {}_mI_0, {}_mI_1. \tag{2.7.1}$$

Here I_0 (resp. P) is shorthand for a base B consisting of a smooth torus (resp. two sphere) with trivial normal bundle $T^2 \times \mathbb{C}$ (respectively $\mathbb{P}^1 \times \mathbb{C}$), while I_1 (resp. II) is a curve with a node (resp. a cusp) singularity and trivial normal bundle. These configurations can give rise to LSTs *only if* 7-branes wrap Σ_{null}. Otherwise, we do not have a genuine 6D model. The variants ${}_mI_0$ and ${}_mI_1$ describe curves whose normal bundle is torsion of order $m > 1$; these can also support 6D theories for $m \in \{2, 3, 4, 6\}$ as discussed below. Observe also that if we apply the tensor-decoupling criterion in these cases, we find that the resulting 6D SCFT is empty, i.e. trivial.

As curves of self-intersection zero do not show up in 6D SCFTs, it is important to explicitly list the possible singular fiber types which can arise on each curve of line (2.7.1).[7]

We begin with the "multiple fiber phenomenon" – a genus one curve Σ whose normal bundle is torsion of order $m > 1$. The F-theory base B is a (rescaled) small neighborhood of Σ, and its canonical bundle $\mathcal{O}_B(K_B)$ must also be torsion of the same order by the adjunction formula. Now to construct a Weierstrass model, we need sections f and g of $\mathcal{O}_B(-4K_B)$ and $\mathcal{O}_B(-6K_B)$, respectively, but nontrivial torsion bundles do not have nonzero sections. Thus, in order to have a nonzero f, the order m of the torsion must divide 4, while to have a nonzero g, the order m must divide 6. There are thus three cases:

1. If $m = 2$, then both f and g may be nonzero.
2. If $m = 3$ or 6, then f must be zero but g may be nonzero.
3. If $m = 4$, then g must be zero but f may be nonzero.

For any other value of $m > 1$, Weierstrass models do not exist (since f and g are not both allowed to vanish identically).

Note that the fact that some quantities obtained from coefficients in a Weierstrass model are sections of torsion bundles also provides the possibility that those sections do not exist (if they are known to be nonzero). As described in Table 4 of [51], the criterion for deciding whether a given Kodaira type leads to a gauge algebra whose Dynkin diagram is simply laced or not simply laced reduces in almost every case to a question of whether a certain quantity has a square root.[8] If the desired square root is in fact a section of a 2-torsion bundle, then it cannot exist.

We can give explicit examples of this phenomenon which do not involve enhanced gauge symmetry, using the framework outlined in section 2.4.1. We start with $S = T^2_S \times \mathbb{C}$, where T^2_S admits an automorphism of order m which acts faithfully on the holomorphic 1-form. If we extend the action to include multiplication by an appropriate root of unity on $\mathbb{C}$, then the

[7] In what follows we focus on those cases where the fiber enhancement leads to a non-abelian gauge symmetry, i.e. a gauge theory description. In the cases where we have a type I_1 or type II fiber enhancement, the resulting 6D theory will consist of some number of weakly coupled free hypermultiplets, where the precise number depends on whether the base curve has non-trivial arithmetic and / or geometric genus. Much as in the case of 6D SCFTs, these cases can be covered through a mild extension of the analysis presented in reference [10]. See also [50] for additional information about these theories.

[8] In the remaining case, one must consider a more complicated cubic equation, but in the situation being described here, the question in that case boils down to the existence or non-existence of a cube root. If the bundle of which the desired cube root is a section is a 3-torsion bundle, the cube root cannot exist.

Base Curve	Matter Content
I_0	Any simple Lie algebra, $n_{Adj} = 1$
I_1,II	$\mathfrak{su}(N)$, $n_{\text{sym}} = 1$, $n_{\Lambda^2} = 1$
	$\mathfrak{su}(6)$, $n_f = 1$, $n_{\text{sym}} = 1$, $n_{\Lambda^3} = \frac{1}{2}$
P	$\mathfrak{su}(N)$, $N \geq 2$, $n_f = 16$, $n_{\Lambda^2} = 2$
	$\mathfrak{su}(6)$, $n_f = 17$, $n_{\Lambda^2} = 1$, $n_{\Lambda^3} = \frac{1}{2}$
	$\mathfrak{su}(6)$, $n_f = 18$, $n_{\Lambda^3} = 1$
	$\mathfrak{sp}(N)$, $N \geq 1$, $n_f = 16$, $n_{\Lambda^2} = 1$
	$\mathfrak{sp}(3)$, $n_f = 17\frac{1}{2}$, $n_{\Lambda^3} = \frac{1}{2}$
	$\mathfrak{so}(N)$, $N = 6, ..., 14$, $n_f = N - 4$, $n_s = \frac{64}{d_s}$
	$\mathfrak{g}_2$, $n_f = 10$
	$\mathfrak{f}_4$, $n_f = 5$
	$\mathfrak{e}_6$, $n_f = 6$
	$\mathfrak{e}_7$, $n_f = 4$
	$\mathfrak{e}_8$, $n_{\text{inst}} = 12$

Table 2.7.1: Rank zero LSTs. In the above, Adj refers to the adjoint representation, sym refers to a two-index symmetric representation, and Λ^n refers to an n-index anti-symmetric representation.

holomorphic 2-form on S is preserved. As is well-known, such automorphisms exist exactly for $m \in \{2, 3, 4, 6\}$. As explained in section 2.4.1, the quotient $(T_F^2 \times S)/\mathbb{Z}_m$ has an elliptic fibration $(T_F^2 \times S)/\mathbb{Z}_m \to (T_F^2 \times \mathbb{C})/\mathbb{Z}_m$ whose fibers over $T_F^2 \times \{0\}$ are all nonsingular elliptic curves, but with $\mathbb{Z}_m$ acting upon them as loops are traversed on T_F^2. This same geometry has a second fibration $(T_F^2 \times S)/\mathbb{Z}_m \to S/\mathbb{Z}_m$ with no section and some multiple fibers in codimension two, which will be further discussed in section 2.9.

Turning now to enhancements of fibers, from the relation between anomaly cancellation and enhanced singular fibers [52, 53, 51, 54] we find all possible gauge theories compatible with a given choice of base curve. This is summarized in table 2.7.1. We find that in general, such theories can support hypermultiplets in the adjoint representation (denoted Adj), the two-index symmetric representation (denoted sym) and the n-index anti-symmetric representation (denoted Λ^n).

The greatest novelty here relative to the case of 6D SCFTs are the theories with $n_{Adj} = 1$ or $\mathfrak{su}(N)$ gauge algebra, $n_{\text{sym}} = 1$, $n_{\Lambda^2} = 1$ or, in the special case of $\mathfrak{su}(6)$, $n_f = 1$, $n_{\text{sym}} = 1$,

$n_{\Lambda^3} = 1/2$. The first of these cases, with $n_{Adj} = 1$, corresponds simply to a smooth curve of genus 1 in the base. The cases with symmetric representations of $\mathfrak{su}(N)$, on the other hand, arise when the *base* curve is of Kodaira type I_1 (i.e. has a nodal singularity). As reviewed in Appendix A.2, the notion of genus is ambiguous for singular curves. A type I_1 curve has topological genus 0 but arithmetic genus 1, and as a result it must support a hypermultiplet in the two-index symmetric representation, rather than one in the adjoint representation.

For LSTs, these are the only examples in which a curve of (arithmetic) genus 1 shows up, and a curve of genus $g \geq 2$ is never allowed. Note also that a curve of genus 0 and self-intersection 0 cannot itself support an $\mathfrak{e}_8$ gauge algebra: it must be blown up at 12 points, resulting in an $\mathfrak{e}_8$ theory with 12 small instantons:

$$(12), 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 \tag{2.7.2}$$

Rank One LSTs

Consider next the case of rank one LSTs, i.e. those in which there are two compact curves in the base. As we have already remarked, in this and all higher rank LSTs, all the curves of the base will be $\mathbb{P}^1$'s, and moreover, they will have self-intersection $-x$ for $1 \leq x \leq 12$. Now, in the case of two curves, we can have various violations of normal crossing. For example, we can have curves which intersect along a tangency. This occurs in both the $4||1$ and $2||2$ configurations, the latter describing a type III Kodaira fiber. Observe that in both cases, there is a smoothing deformation which takes us from an order two tangency to a loop, i.e. we can deform to $//4, 1//$ and $//2, 2//$. In addition to these rank one LSTs, there is just one more configuration given by $1, 1$, which in some sense is the most "conventional" possibility (as all intersections respect normal crossing).

To this end, let us first discuss fiber enhancements for the $1, 1$ configuration. We shall then turn to the cases where there is either a violation of normal crossing or a loop configuration. Whenever curves with gauge algebras intersect, matter charged under each gauge algebra will pair up into a mixed representation of the gauge algebras. The mixed anomaly condition places strong constraints on which representations are allowed to pair up. The allowed set of mixed representations for two curves intersecting at a single point is given in section 6.2 of reference [10]. Consider for example, the $1, 1$ base. We have:

$$\begin{matrix} \mathfrak{g}_L & \mathfrak{g}_R \\ 1 & 1 \end{matrix} \tag{2.7.3}$$

with the following list of allowed gauge algebras:

- $\mathfrak{g}_L = \mathfrak{so}(M)$, $\mathfrak{g}_R = \mathfrak{sp}(N)$, $M = 7, ..., 12$, $M - 5 \geq N$, $4N + 16 \geq M$.
- $\mathfrak{g}_L = \mathfrak{so}(M)$, $\mathfrak{g}_R = \mathfrak{sp}(N)$, $M = 7$, $N \leq 6$.
- $\mathfrak{g}_L = \mathfrak{g}_2$, $\mathfrak{g}_R = \mathfrak{sp}(N)$, $N \leq 7$.
- $\mathfrak{g}_L = \mathfrak{sp}(M)$, $\mathfrak{g}_R = \mathfrak{sp}(N)$, $2M + 8 \geq 2N$, $2N + 8 \geq 2M$.

- $\mathfrak{g}_L = \mathfrak{sp}(M)$, $\mathfrak{g}_R = \mathfrak{su}(N)$, $2M + 8 \geq N$, $N + 8 + \delta_{N,3} + \delta_{N,6} \geq 2M$.
- $\mathfrak{g}_L = \mathfrak{su}(M)$, $\mathfrak{g}_R = \mathfrak{su}(N)$, $M + 8 + \delta_{M,3} + \delta_{M,6} \geq N$, $N + 8 + \delta_{N,3} + \delta_{N,6} \geq M$.
- $\mathfrak{g}_L = \mathfrak{f}_4$, $\mathfrak{e}_6$, $\mathfrak{e}_7$ or $\mathfrak{e}_8$, $\mathfrak{g}_R = \emptyset$.

Here, it is understood that $\mathfrak{sp}(0)$ is the same as an empty -1 curve, and $\mathfrak{e}_8$ on a -1 curve implies that 11 points on the -1 curve have been blown up.

Consider next the configuration $//2, 2//$, i.e. a loop of two -2 curves. In this case the only gauge algebra enhancement is given by

$$\overset{\mathfrak{su}(N)}{//} \overset{\mathfrak{su}(N)}{2} \quad \overset{}{2} \overset{}{//} \tag{2.7.4}$$

with a bifundamental localized at each intersection point.

When the -2 curves intersect tangentially, i.e. in the $2||2$ configuration, more options are available, as suggested by the effective field theory on the tensor branch. We find the following general possibilities for enhancements of the gauge algebra and matter:

- $\mathfrak{g}_a = \mathfrak{su}(N_a)$, $\mathfrak{g}_b = \mathfrak{so}(N_b)$, $R_a = \mathbf{N_a}$, $R_b = \mathbf{N_b}$. The only allowed possibilities are $N_a = 6$, $N_b = 12$ and $N_a = 7$, $N_b = 13$.
- $\mathfrak{g}_a = \mathfrak{su}(4)$, $\mathfrak{g}_b = \mathfrak{so}(7), \mathfrak{so}(8)$, $R_a = \mathbf{4}$, $R_b = \mathbf{8}$.
- $\mathfrak{g}_a = \mathfrak{su}(4)$, $\mathfrak{g}_b = \mathfrak{g}_2$, $R_a = \mathbf{4}$, $R_b = \mathbf{7}$.

Just as it was necessary to deform the singular Kodaira I_1 curve to get hypermultiplets in the symmetric representation of $\mathfrak{su}(N)$, so it is necessary to deform the I_2 base to the tangentially intersecting Kodaira type III curve configuration to get these matter pairings. In conclusion, for the Kodaira type III configuration $2||2$, we may enhance the gauge algebras as:

$$\overset{\mathfrak{so}(7)}{\underset{[Sp(1)]}{2}} || \overset{\mathfrak{su}(4)}{2}$$

$$\overset{\mathfrak{so}(8)}{\underset{[Sp(2)]}{2}} || \overset{\mathfrak{su}(4)}{2}$$

$$\overset{\mathfrak{g}_2}{2} || \overset{\mathfrak{su}(4)}{2}$$

$$\overset{\mathfrak{so}(12)}{\underset{[N_s=1]}{2}} || \overset{\mathfrak{su}(6)}{2}$$

$$\overset{\mathfrak{so}(13)}{\underset{[N_s=1/2]}{2}} || \overset{\mathfrak{su}(7)}{2}$$

Here, there is a mixed representation in the bifundamental of the two gauge algebras in the last three cases. These representations can only show up when the curves are tangent to each other.

Finally, we turn to the case of the bases $//4,1//$ and $4||1$. In both cases, the only enhancement possible is an SO-type algebra over the -4 curve and an Sp-type or SU-type algebra over the -1 curve. The matter content of the theory, however, depends on the type of intersection. Consider first the case of

$$\overset{\mathfrak{so}(2N+8)}{//\quad 4} \qquad \overset{\mathfrak{sp}(N)}{1\quad //} \tag{2.7.5}$$

with a half hypermultiplet localized at each intersection point. Indeed, we can reach this theory by starting from the 6D SCFT:

$$[\mathfrak{so}(2N+8)] \overset{\mathfrak{sp}(N)}{1} [\mathfrak{so}(2N+8)] \tag{2.7.6}$$

and gauging the diagonal subalgebra of the flavor symmetry.

In the case of the tangential intersection $4||1$, we again find novel configurations of matter which are missing from the case of normal crossing. The gauge algebras are the same as those of (2.7.5):

$$\overset{\mathfrak{so}(2N+8)}{4} \quad || \quad \overset{\mathfrak{sp}(N)}{1} \tag{2.7.7}$$

However, there is now a single full hypermultiplet in the bifundamental of the two gauge algebras rather than two half hypermultiplets.

There is another configuration possible in the case of $4||1$:

$$\overset{\mathfrak{su}(N)}{1} \quad || \quad \overset{\mathfrak{so}(N+8)}{4} \tag{2.7.8}$$

with a hypermultiplet in the bifundamental of the two gauge algebras, as well as a hypermultiplet in the two index anti-symmetric representation of the $\mathfrak{su}(n)$ factor, all of which are located at the collision point between the two branes.

Rank Two LSTs

We now turn to the case of rank two LSTs, i.e. those with three curves. Here, we can have no tangential intersections. In this case, the adjacency matrix again provides only partial information about the geometry of intersecting curves. In the case where we have normal crossings for all pairwise intersections, the rules of enhancing fiber enhancements follow from reference [10] and are also reviewed in Appendix A.1. There is, however, also the possibility of a Kodaira type IV configuration of -2 curves, i.e. $2 \overset{2}{\bigtriangledown} 2$. We shall therefore confine our attention to fiber decorations with this base.

To illustrate, suppose the gauge algebra localized on each of the three -2 curves is $\mathfrak{su}(2)_i, i = 1,2,3$. Anomaly cancellation dictates that in such a case there must be a single half-trifundamental $\frac{1}{2}(\mathbf{2},\mathbf{2},\mathbf{2})$ plus two fundamentals charged under each $\mathfrak{su}(2)$. We note in passing that the loop-like configuration $//222//$ also admits a similar enhancement in the gauge algebras, i.e. with $\mathfrak{su}(2)_i, i = 1,2,3$ gauge groups, but that the corresponding matter content is given by three bifundamentals $(\mathbf{2},\mathbf{2},\mathbf{1})$, $(\mathbf{2},\mathbf{1},\mathbf{2})$, $(\mathbf{1},\mathbf{2},\mathbf{2})$. These two configurations have the same anomaly polynomials.

But in contrast to the $//222//$ configuration, for $2\overset{2}{\bigtriangledown}2$, no other gauge algebra enhancements are possible. To see this, suppose we $\mathfrak{su}(N)$ factors on each -2 curve. We would then need N^2 fundamentals of each $\mathfrak{su}(N)$ to get a (N, N, N) representation. However, anomaly cancellation considerations constrain us to 2N such fundamentals. In other words, we are limited to $N \leq 2$.

2.7.2 Higher Rank LSTs

Turning next to the case of LSTs with at least four curves in the base, all of these local violations of normal crossing do not appear. Nevertheless, we encounter such violations when we attempt to blow down -1 curves which touch more than two curves. Even so, the local rules for fiber enhancements follow the same algorithm already spelled out in detail in reference [10]. In some cases, however, there can be additional restrictions compared with the fiber enhancements which are possible for 6D SCFTs. To illustrate, we primarily focus on some simple examples, i.e. the affine ADE bases, fiber decorations for the base $1, 2, ..., 2, 1$, and fiber decorations for the loop-like bases.

Affine ADE Bases

Let us next consider the case of fiber enhancements in which the base is given by an affine Dynkin diagram of -2 curves. If we assume that no further blowups are introduced in the base, we will be limited to just $\mathfrak{su}(N)$ gauge algebras over each curve. In the case of 6D SCFTs, it is typically possible to obtain a rich class of possible sequences of gauge group factors, because 6D anomaly cancellation can be satisfied by introducing an appropriate number of additional flavor symmetry factors. This in turn leads to a notion of a "ramp" in the increase in the ranks [28] (see also [18]). For an affine quiver, this is much more delicate, since all of these cases can be viewed as a degeneration of an elliptic curve. For example, in the case of an affine $\widehat{A}_k$ base (i.e. the I_k Kodaira type), anomaly cancellation tells us that all of the gauge algebra factors are the same $\mathfrak{su}(N)$. By a similar token, 6D anomaly cancellation tells us that the gauge algebra of any of these cases is $\mathfrak{su}(Nd_i)$, where d_i is the Dynkin label of the node in the affine graph, and $N \geq 1$ is an overall integer.[9] In F-theory language, we have fiber enhancements I_{Nd_i} over each node. Indeed, at a formal level we can think of anomaly cancellation being satisfied by introducing $\mathfrak{su}(1)$ gauge algebras.

The $1, 2, ..., 2, 1$ Base

Compared with the case of affine ADE bases, there are comparatively more options available for fiber enhancements of the base $1, 2, ..., 2, 1$. In some sense, this is because these bases do not directly arise from the degeneration of a compact elliptic curve, but are better viewed as the degeneration of a cylinder.

[9] This same observation has already been made in the context of 4D superconformal $\mathcal{N} = 2$ quiver gauge theories [55]. Indeed, in this special case the condition of vanishing beta functions is identical to the condition that 6D anomalies cancel.

With this in mind, we now explain how fiber enhancements work for this choice of base. For a large number of -2 curves, the allowed enhancements take a rather simple form, whereas there are outlier LSTs for smaller numbers of -2 curves. In particular, when there are more than five -2 curves, the -2 curves necessarily hold $\mathfrak{su}(N_i)$ gauge algebras:

$$\begin{array}{ccccccccc} \mathfrak{g}_L & \mathfrak{su}(N_1) & \mathfrak{su}(N_2) & & \mathfrak{su}(N_{k-1}) & \mathfrak{su}(N_k) & \mathfrak{g}_R \\ 1 & 2 & 2 & \ldots & 2 & 2 & 1 \end{array}$$

The N_i are subject to the convexity conditions $2N_i \geq N_{i-1} + N_{i+1}$, with the understanding that $N_i = 1$ for a curve without a gauge algebra.

The -1 curves in this configuration may hold either $\mathfrak{sp}(M)$ or $\mathfrak{su}(M)$ gauge algebra. If the leftmost -1 curve holds $\mathfrak{sp}(M)$, anomaly cancellation imposes the additional conditions $2N_1 \geq 2M + N_2$, $2M + 8 \geq N_1$. If the leftmost -1 curve holds $\mathfrak{su}(M)$, anomaly cancellation imposes $2N_1 \geq M + N_2$, $M + 8 + \delta_{M,3} + \delta_{M,6} \geq N_1$. Finally, the -1 curve may be empty provided $N_1 \leq 9$. The story is mirrored for the rightmost -1 curve at the other end of the chain.

When there are exactly five -2 curves, we have two additional configurations:

$$\begin{array}{cccccccc} & & & \mathfrak{su}(2) & \mathfrak{so}(7) & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

and

$$\begin{array}{cccccccc} & & & \mathfrak{su}(2) & \mathfrak{g}_2 & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

When there are four -2 curves, we similarly have

$$\begin{array}{ccccccc} & & \mathfrak{su}(2) & \mathfrak{so}(7) & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

and

$$\begin{array}{ccccccc} & & \mathfrak{su}(2) & \mathfrak{g}_2 & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

When there are three -2 curves, we have several new configurations:

$$\begin{array}{ccccccc} & & \mathfrak{su}(2) & \mathfrak{so}(7) & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

$$\begin{array}{ccccccc} & & \mathfrak{su}(2) & \mathfrak{g}_2 & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

$$\begin{array}{ccccccc} \mathfrak{g}_L & \mathfrak{so}(7) & \mathfrak{su}(2) & & & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

and

$$\begin{array}{ccccccc} \mathfrak{g}_L & \mathfrak{g}_2 & \mathfrak{su}(2) & & & & \\ 1 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}$$

with $\mathfrak{g}_L = \mathfrak{sp}(M), M \leq 3$ in each of the last two cases.

When there are two -2 curves, we have:

$$\begin{array}{ccccc} \mathfrak{g}_L & \mathfrak{so}(7) & \mathfrak{su}(2) & & \\ 1 & 2 & 2 & 2 & 1 \end{array}$$

and

$$\overset{\mathfrak{g}_L}{1}\ \overset{\mathfrak{g}_2}{2}\ \overset{\mathfrak{su}(2)}{2}\ 1$$

with $\mathfrak{g}_L = \mathfrak{sp}(M), M \leq 3$ in each of the last two cases.

When there is only a single -2 curve, there are even more possibilities:

$$\overset{\mathfrak{g}_L}{1}\ \overset{\mathfrak{so}(8)}{2}\ \overset{\mathfrak{g}_R}{1}$$

with $\mathfrak{g}_L = \mathfrak{sp}(M_L), M_L \leq 2$, $\mathfrak{g}_R = \mathfrak{sp}(M_R), M_R \leq 2$.

$$\overset{\mathfrak{g}_L}{1}\ \overset{\mathfrak{so}(7)}{2}\ \overset{\mathfrak{g}_R}{1}$$

with $\mathfrak{g}_L = \mathfrak{sp}(M_L)$, $\mathfrak{g}_R = \mathfrak{sp}(M_R)$, $M_L + M_R \leq 4$ or $M_L = 4$, $M_R = 1$.

$$\overset{\mathfrak{g}_L}{1}\ \overset{\mathfrak{g}_2}{2}\ \overset{\mathfrak{g}_R}{1}$$

with $\mathfrak{g}_L = \mathfrak{sp}(M_L)$, $\mathfrak{g}_R = \mathfrak{sp}(M_R)$, $M_L + M_R \leq 4$.

$$\overset{\mathfrak{g}_L}{1}\ \overset{\mathfrak{su}(2)}{2}\ 1$$

with $\mathfrak{g}_L = \mathfrak{g}_2$ or $\mathfrak{so}(7)$.

2.7.3 Loop-like Bases

Finally, let us turn to the case of fiber enhancements for the loop-like bases. We have already discussed the case of an affine $\widehat{A}_k$ base of -2 curves in which we the allowed fiber enhancements are just a uniform I_N fiber. Otherwise, we induce some blowups. Now, if we allow for blowups, we can reach more general loop-like configurations. However, as we have already discussed near line (2.6.2), all of these cases consist of a single type of base node suspended between minimal links. This is a consequence of the fact that in a general 6D SCFT, there are nested containment relations on the minimal fiber types [10]:

$$\mathfrak{g}_1^{min} \subseteq ... \subseteq \mathfrak{g}_m^{min} \supseteq ... \supseteq \mathfrak{g}_k^{min}. \tag{2.7.9}$$

However, in a 6D LST, we must *also* demand periodicity of the full configuration. This forces a uniform fiber enhancement on each such node.

As a consequence, we can summarize all of these cases by keeping implicit the blowups associated with conformal matter. We have:

$$//\overset{\mathfrak{g}}{2}, ..., \overset{\mathfrak{g}}{2}// \tag{2.7.10}$$

where we allow for a general fiber enhancement to an ADE type simple Lie algebra $\mathfrak{g}$ over each of the -2 curves. For all cases other than the $\mathfrak{su}(N)$ gauge algebras, this in turn requires

further blowups in the base, i.e. we have a configuration with conformal matter in the sense of references [28, 29]. So in other words, all of these loop-like configurations are summarized by stating the number of -2 curves, and the choice of fiber type over any of the -2 curves.

2.8 Embeddings and Endpoints

In the previous sections we presented a general classification of LSTs in F-theory. One of the crucial ingredients we have used is that decompactifying any curve must return us to a collection of (possibly disconnected) SCFTs. Turning the question around, it is natural to ask whether *all* SCFTs embed in LSTs.

In this section we show that this is indeed the case. Moreover, there can often be more than one way to complete an SCFT to an LST. To demonstrate such an embedding, we will need to show that there exists a deformation of a given LST F-theory background which takes us to the requisite SCFT. This can involve both Kähler deformations, i.e. motion onto a partial tensor branch, and may also include complex structure deformations, i.e. a Higgsing operation.

With this in mind, we first demonstrate that all of the bases for 6D SCFTs embed in an LST base. A suitable tensor branch flow then takes us from the LST base back to the 6D SCFT base. Then, we proceed to show that the available fiber decorations for LSTs can be Higgsed down to the fiber decorations for an SCFT. The latter issue is somewhat non-trivial since the fiber decorations of an ADE-type base is comparatively less constrained when compared with their affine counterparts.

2.8.1 Embedding the Bases

We now show that all bases for 6D SCFTs embed in LST bases. To demonstrate that such an embedding is possible, it is convenient to use the terminology of "endpoints" for SCFTs introduced in reference [9], which we can also extend to the case of LSTs. Given a collection of curves for an SCFT base, we can consider blowdowns of all of the -1 curves of the configuration. Doing so, we shift the self-intersection of all curves touching this -1 curve according to the rule $x \to (x-1)$ for a curve of self-intersection $-x$. In the case where a -1 curve is interposed in between two curves, we have $x, 1, y \to (x-1), (y-1)$. After this first stage of blowdowns, we can then sometimes generate *new* -1 curves. Iteratively blowing down all such -1 curves, we eventually reach a configuration of curves which we shall refer to as an "endpoint." The set of all endpoints has been classified in reference [9], and they split up according to four general types:

$$\text{Trivial Endpoint: } 1 \to \mathbb{C}^2 \tag{2.8.1}$$

$$\text{A-type Endpoint: } x_1 ... x_k \tag{2.8.2}$$

$$\text{D-type Endpoint: } 2\overset{2}{x_1} ... x_k \tag{2.8.3}$$

$$\text{E-type Endpoint: } 2\overset{2}{2}222, \quad 22\overset{2}{2}222, \quad 222\overset{2}{2}222 \tag{2.8.4}$$

In fact, starting from such an endpoint we can generate all possible bases of 6D SCFTs by further blowups. Sometimes such blowups are required to define an elliptic Calabi-Yau, while some can be added even when an elliptic fibration already exists. By a similar token, we can also take a fixed base, and then decorate by appropriate fibers.

Now, a central feature of this procedure is that the resulting adjacency matrix retains the important property that it is positive definite. Similarly, if we instead have a positive semidefinite adjacency matrix, the resulting matrix will retain this property under further blowups (or blowdowns) of the base.

To demonstrate that we can always embed an SCFT in an LST, it will therefore suffice to show that there is *some* way to add additional curves to an SCFT endpoint such that the resulting LST defines a base. To illustrate the idea, suppose we have a 6D SCFT with a trivial endpoint. Then, before the very last stage of blowdowns, we have a single -1 curve, which we shall call Σ. If we return to the original SCFT, this curve will also be present, but its self-intersection will be different. Hence, to get an LST we can simply attach one more -1 curve to Σ. For example, the configuration $1, 1$ defines the base of an LST. Supplementing the fibers can always be done to realize the case of fiber decorations.

Consider next the case of A-type endpoints. Here, we can always attach a suitable number of "tails" of the form $1, 2..., 2$ to each curve such that blowing down these instanton links leads us to k curves of self-intersection -2, i.e. $2, ..., 2$. Attaching an additional -1 curve to the left and to the right, we get a base of the form $1, 2, ..., 2, 1$. We therefore conclude that adding such tails again allows an embedding in an LST.

By a similar token, we can append such tails to a D-type endpoint until we reach the Dynkin diagram with just -2 curves. Adding one more -1 curve to this configuration:

$$\text{D-type Endpoint} = 2\overset{2}{2}...2, 1 \tag{2.8.5}$$

leads to a blowdown eventually to the configuration $2, 1, 2$, which in turn blows down to $1, 1$. So again, we conclude that the adjacency matrix is positive semidefinite and we have arrived at an LST.

This leaves us with the E-type endpoints. In these cases, we just have a configuration of -2 curves, and the possible blowups are severely limited [9]. For example, there are no blowups of the E_8 Dynkin diagram. It therefore suffices to add one additional -2 curve to this configuration to reach its affine extension. Similar considerations also apply for the E_6 and E_7 configurations when no additional blowups are present, i.e. we simply proceed from the Dynkin diagram to its affine extension.

To round out the analysis, we need to demonstrate that if we perform any blowups of an E_6 or E_7 endpoint, we can again attach a -2 curve at the same location, without inducing any further blowups.[10] That this is indeed the case is conveniently summarized by simply writing down the possible blowups. For E_6, there are two other consistent bases, and for E_7 there is one. In both

[10]Note that attaching a -1 curve will not work since we would then blowdown to a configuration where the adjacency matrix is not positive semidefinite – a contradiction.

cases, we can indeed still add our -2 curve without inducing extraneous blowups:

$$\begin{array}{ccccc} & 2 & & & \\ 2 & 3 & 2 & & \\ 2 & 3 & 1 & 2 & \\ 1 & 1 & & 2 & \end{array}$$
$$232181232 \to 2315132 \to 22222 \tag{2.8.6}$$

$$\begin{array}{ccc} & 3 & \\ & 1 & \\ 2 & 2 & 2 \end{array}$$
$$223151322 \to 2222222. \tag{2.8.7}$$

Summarizing, we have just demonstrated that *all* bases for 6D SCFTs embed in an LST base.

LST Endpoints As a brief aside, one of the interesting features of this argument is that we have implicitly relied on the notion of an LST endpoint. Given that we have already classified all such bases, we can also ask about the possible endpoints for LSTs. Compared with the case of 6D SCFTs, the number of distinct endpoints are comparatively small. Roughly speaking, this is because of the positive semidefinite condition for our adjacency matrix, which in turn means that many configurations will blow down to a single curve of self-intersection zero (as in the $1, 2, ..., 2, 1$ configurations).

Combining the tensor-decoupling criterion with the demand that we have a positive semidefinite adjacency matrix means that the total number of endpoints are given by the Kodaira type intersections of -2 curves, an elliptic curve, as well as a single $\mathbb{P}^1$ of self-intersection zero, which we denote by P. In the latter two cases, we note that we only obtain a 6D LST by having a non-trivial elliptic fibration. Thus, we find the following list of LST endpoints:[11]

$$\text{LST Endpoints: } P, I_n, II, III, IV, I_n^*, II^*, III^*, IV^* \tag{2.8.8}$$

for $n \geq 0$.

2.8.2 Embedding the Fibers

Suppose next that we have supplemented a base by an additional curve. When we do this, additional non-trivial fibers are sometimes inevitable, and can in turn force additional structure on the elliptic fibers. To give a concrete example, consider the case of a 6D SCFT with base given by the E_6 Dynkin diagram of -2 curves. Fiber decorations for this model were studied in reference [10] where it was found that typically, additional flavors can be added so that an appropriate convexity condition on the ranks is obeyed. To extend this to an LST base, we cannot add a -1 curve (as the blowdown is inconsistent). Rather, we must add an additional -2 curve to reach the affine $\widehat{E}_6$ Dynkin diagram of -2 curves. When we do this, we must remember that the elliptic fibration of the resulting F-theory model also becomes rigid. So in other words, the available elliptic fibrations are further constrained. It is at this stage that we must include the effects of Higgsing as well as tensor branch deformations to reach the original 6D SCFT. That this is always possible follows from the fact that all representations of the Lie algebra $\mathfrak{e}_6$ embed

[11] Here we neglect the possibility of torsion in the normal direction to the compact curves of the base.

in representations of its affine extension $\widehat{\mathfrak{e}}_6$. Similar considerations apply for the fiber decorations of all of the E-type bases.

In the case of the A- and D-type bases, the analysis is comparatively simpler. The reason is that we can just add additional tails of the form $1, 2, ..., 2$ with trivial fibers and leave the fibers above the curves in the SCFT base as they were. One might worry here about the fact that curves of self-intersection -1 and -2 are not always allowed to have trivial fibers. For instance, the -1 curve in the sequence $2, 2, 3, 1, 5$ necessarily has a type II fiber. However, such subtleties do not arise in this case: one can always add small instanton links of the form $1, 2, ..., 2$ with trivial fibers to an A- or D-type base to get an LST.

Putting together our analysis of tensor branch flows and Higgs branch flows, we conclude that all 6D SCFTs embed in some LST. Indeed, it is also clear that there can sometimes be more than one such embedding.

2.9 T-duality

In the previous sections we used the geometry of F-theory compactifications to tightly constrain the structure of LSTs. In this section we turn the analysis around and show how the physics of little string theories suggests non-trivial geometric structures for elliptic Calabi-Yau threefolds in which the non-compact base has a negative semidefinite intersection form.

In physical terms, one of the key features of LSTs is that the description as a local quantum field theory must break down near the string scale. A sharp way to probe this structure is by compactifying on a circle. Recall that in T-duality, the theory compactified on a small S^1 of radius R is dual to another string theory compactified on an S^1 of radius $\widetilde{R} \sim \alpha'_{eff}/R$. Based on this, it is natural to expect that all LSTs have a similar T-duality.

This expectation suggests a non-trivial constraint on the geometry of an F-theory realization of an LST. Recall that F-theory compactified on an elliptically fibered Calabi-Yau threefold $X \to B$ leads, upon further compactification on an S^1, to M-theory compactified on the same Calabi-Yau threefold. In the M-theory description, the Kähler class becomes a dynamical modulus (which is taken to zero size to reach the F-theory limit). On the other hand, T-duality tells us that if we take this circle to be very small, we should expect to obtain another LST, this time compactified on a large radius circle. For this to be so, we must have available to us more than one way to reach an F-theory background from a given compactification of M-theory on X. In other words, physical considerations suggest the existence of *another* elliptic fibration for our Calabi-Yau threefold $X \to \widetilde{B}$, and the lift from M-theory to F-theory involves collapsing the Kähler class of this other elliptic fiber to zero size.

Our plan in this section will be to give further evidence that T-duality is realized in F-theory constructions of LSTs through the presence of a double elliptic fibration. We shall, however, mainly focus on particular examples of how T-duality is realized geometrically. After this, we give a sketch for how we expect this correspondence to work in general, leaving a complete proof to future work.

2.9.1 Examples

As a first example consider the T-duality between the LST of $k \geq 2$ NS5-branes in IIA (the $\mathcal{N} = (2,0)$ A-type LSTs), and that of k NS5-branes in IIB string theory (the $\mathcal{N} = (1,1)$ A-type LSTs). The IIA realization just follows from a base with -2 curves arranged in a loop, i.e. as the type I_{k-1} degeneration of an elliptic curve. The F-theory elliptic curve is then a smooth T^2, i.e. an I_0 fiber. Switching the roles of these two curves, we get the type IIB $\mathcal{N} = (1,1)$ LST, i.e. we have $(k-1)$ D7-branes wrapped over a T^2. There is a clear extension of this case to all of the ADE $\mathcal{N} = (2,0)$ LSTs in terms of the corresponding ADE 7-branes wrapped over a T^2.

Another class of examples we have already encountered several times involves the LSTs realized by M5-branes filling $\mathbb{R}^{5,1}$ and probing the geometry $S^1_\perp \times \mathbb{C}^2/\Gamma$ for Γ an ADE subgroup of $SU(2)$. Compactifying on a further circle, we can shrink the $S^1_\perp$ factor to reach IIA string theory. T-duality is then inherited from that of the physical superstring theory.

In the F-theory realization of these systems, we have a base quiver with conformal matter suspended in between the nodes:

$$//\underbrace{g \oplus g \oplus ... \oplus g \oplus g}_{k}//. \tag{2.9.1}$$

As we have already remarked in the classification of such structures, all fiber types and conformal matter are necessarily minimal; no deviations from this rigid structure are possible. This actually means that we can readily identify the other elliptic fiber of this model: It is given by a suitable multiple of an I_{kd_i} fiber, where the d_i denote the Dynkin labels of the affine extension of the g-type Lie algebra. Another consequence of this analysis is that sometimes, the absence of other fiber decorations for say the E-type affine bases means we cannot arbitrarily combine these two structures. Rather, since the only fiber enhancements over the E-type bases are I_k-type fibers (i.e. without inducing further blowups), there is again a clear exchange between the roles of the two (singular) elliptic curves.

Quite strikingly, this also entails the existence of several infinite classes of models which are self-T-dual upon toroidal reduction. These are models which have a double elliptic structure consisting of two copies of the *same* Kodaira fiber. For example, models which have an I_N Kodaira base, with gauge groups $\mathfrak{su}(N)$ on each -2 curve and bifundamentals at the intersections. As a more exotic example consider the blown up type IV degeneration: $\overset{3}{3}13$. On each -3 curve in this configuration there is an $\mathfrak{su}(3)$ gauge sector which arises precisely from a type IV fiber. Clearly, contracting the -1 curve, switching fiber and base, and then blowing up we obtain back the same model.[12]

As a final class of examples, consider the F-theory models with base:

$$[E_8]\underbrace{1,2,...,2,1}_{k}[E_8]. \tag{2.9.2}$$

This is realized in heterotic M-theory by a collection of k M5-branes in between the two heterotic walls. compactifying on a further circle and activating appropriate Wilson lines for the

[12] See [50] for a further analysis of this case.

background $E_8 \times E_8$ flavor symmetry, we reach –via T-duality of the physical superstring– the case of k NS5-branes of the $Spin(32)/\mathbb{Z}_2$ heterotic string theory. This system can be analyzed in perturbative string theory, and the T-dual LST is therefore realized by an $\mathfrak{sp}(k)$ gauge theory with 32 half hypermultiplets in the fundamental representation, and a single hypermultiplet in the two-index anti-symmetric representation.

Demonstrating the presence of the extra elliptic fibration for the F-theory model is somewhat more subtle in this case, but we can see it as descending from a $\mathbb{Z}_2$ quotient of the I_{2k}-type Kodaira configuration of -2 curves:

$$//\underbrace{2,2,...,2,2}_{2k}// \overset{\mathbb{Z}_2}{\to} \underbrace{1,2,...,2,1}_{k}. \tag{2.9.3}$$

In Appendix A.6 we present an explicit analysis of this case of the base $1,1$, and also explain its extension to the configuration $1,2,...,2,1$. The corresponding F-theory model is realized by a base given by P, a single $\mathbb{P}^1$ with self-intersection zero, with a fiber enhancement I_{2k}^{ns}, i.e. we get an $\mathfrak{sp}(k)$ 7-brane wrapped over P. We have already classified the matter enhancements for this case in section 2.7, and indeed, we find agreement with the purely heterotic analysis. Note that the $\mathfrak{sp}$-type algebra originates from the $\mathbb{Z}_2$ quotient of an A_{2k} algebra via its outer automorphism.

In fact, a similar observation allows us to extend this to some of the models in which we have a $\mathbb{P}^1$ of self-intersection zero. For example, under a further $\mathbb{Z}_2$ quotient, we can reach some of the gauge theories already encountered via more direct methods:

$$[\mathfrak{su}_8] \overset{\mathfrak{su}_N}{1}, \overset{\mathfrak{su}_N}{1} [\mathfrak{su}_8] \overset{\mathbb{Z}_2}{\to} [\mathfrak{su}_{16}] \overset{\mathfrak{su}_N}{0} [\mathfrak{su}_2], \tag{2.9.4}$$

where the $\mathfrak{su}_{16}$ flavor symmetry acts on the sixteen hypermultiplets in the fundamental representation, and the $\mathfrak{su}_2$ flavor symmetry acts on the hypermultiplets in the two index anti-symmetric representation. From this perspective, we can still recognize the quotient of an additional elliptic fiber. Schematically, we have:

$$//2,2// \overset{\mathbb{Z}_2}{\to} 1,1 \overset{\mathbb{Z}_2}{\to} 0. \tag{2.9.5}$$

2.9.2 Examples Involving Curves with Torsion Normal Bundle

Recall that in section 2.4 we found that the $\mathcal{N} = (1,1)$ LSTs admitted an ABCDEFG classification according to their corresponding affine Lie algebras. In the F-theory realization of these models, we also saw that the base contained a genus 1 curve with a torsion normal bundle. We encountered torsion normal bundles again in section 2.7.1 in our discussion of rank zero LSTs. We will now explain these examples in more detail.

By construction, examples in this class admit two genus one fibrations: $(T_F^2 \times S)/\mathbb{Z}_m \to (T_F^2 \times C)/\mathbb{Z}_m$, and $(T_F^2 \times S)/\mathbb{Z}_m \to S/\mathbb{Z}_m$. The first fibration has a section with monodromy over the central fiber of $\pi : (T_F^2 \times C)/\mathbb{Z}_m \to C$. The second fibration does not have a section (i.e., it is not an elliptic fibration, in the terminology of [56]), and we discuss its structure here.

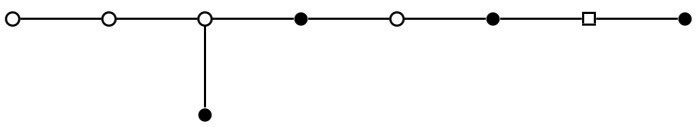
Recall that for a genus one fibration $X \to B$ without a section, there is an associated "Jacobian fibration" $J(X/B) \to B$ which has a section, and which has precisely the same τ function

describing the fibers.[13] As explained in [24, 56, 57], the set of X's which share a common Jacobian fibration (and are equipped with an action by the Jacobian fibration, as stressed in [58]) forms a group[14] which should be identified with the group of connected components of the gauge group in F-theory. That is, compactifying the F-theory model on a circle, there is a discrete choice of one of these X's to serve as the compactification space for M-theory, which is the hallmark of a discrete gauge choice.

In our examples, the action of $\mathbb{Z}_m$ on S has fixed points, leading to A_{m-1} singularities on the quotient $S/\mathbb{Z}_m$. (More generally, there can be $A_{\ell-1}$ singularities for any ℓ dividing m, due to fixed points of subgroups.) Over an $A_{\ell-1}$ point, we have taken the quotient of the fiber by a translation of order ℓ, so that fiber has multiplicity ℓ. Notice that even though the fiber is multiple, the total space is smooth.

This is precisely the situation analyzed by Mark Gross in [60], who showed that the Jacobian fibration is fibered over the same base, still with $A_{\ell-1}$ singularities. But once the Jacobian fibration has been taken, it is possible to resolve those singularities of the base (which corresponds physically to giving an expectation value to scalars in the corresponding tensor multiplets). We thus find that these theories are part of a larger family of LSTs, but at special values of the tensor moduli in the larger family, a finite gauge group appears, leading to additional 5D vacua (i.e., the fibrations without a section) corresponding to distinct sectors of Wilson line expectation values.

Let us illustrate this with two concrete examples drawn from Appendix A.7, where we work out the $\mathcal{N} = (1, 1)$ theories of BCFG type in detail. As a first example, consider a partially blown down graph of type $\widehat{E}_8$, illustrated here (and in Figure A.7.1):



where we blow down each solid circle to an A_1 singularity. As shown in Appendix A.7, there is a genus 1 fibration X over this base B with a fiber of multiplicity 2 at each A_1 singularity. The fiber never degenerates over this locus.

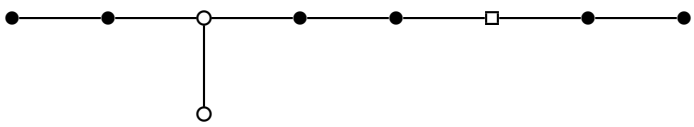
The Jacobian fibration $J(X/B)$ is simply a Weierstrass fibration over B whose fibers do not degenerate. There is no obstruction to resolving the singularities of B, moving out into the rest of the moduli space. In fact, this is part of the moduli space of the $\mathcal{N} = (2, 0)$ theory of type E_8, and we claim that when the tensors are tuned to blow down precisely the curves corresponding to solid nodes, a $\mathbb{Z}_2$ gauge symmetry appears in the theory. That gauge symmetry is necessary to

[13] A more common notation is $J(X)$, but since we are studying Calabi–Yau threefolds with more than one genus one fibration, we indicate the fibration for emphasis.

[14] This group has been incorrectly called the Tate–Shafarevich group in the literature [56, 57], for which the fourth author apologizes. The group contains the Tate–Shafarevich group [59], but it can also contain fibrations with isolated multiple fibers [60], a fact which arose in the proof of the finiteness theorem for elliptic Calabi–Yau threefolds [61,62]. The examples presented here have such isolated multiple fibers, and so do not belong to the Tate–Shafarevich group but rather to the larger group of Calabi–Yau threefolds sharing a common Jacobian fibration.

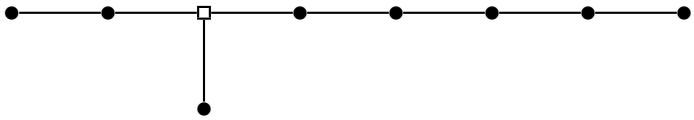
explain the additional 5D models which appear when the A_1's are blown down and a twist of the Jacobian fibration is possible. Presumably, moving away from this locus amounts to Higgsing the $\mathbb{Z}_2$ gauge symmetry. Incidentally, this example is the T-dual of the $\mathcal{N} = (1,1)$ model with group F_4.

As a second example (which is the T-dual of the $\mathcal{N} = (1,1)$ model with group G_2), consider another partially blown down graph of type $\widehat{E}_8$, illustrated here (and in Figure A.7.2):



where this time, we blow down each lined pair of solid circles to an A_2 singularity. There is a fibration with multiple fibers of multiplicity 3 over each of those points. The Jacobian fibration is again part of the $\mathcal{N} = (2,0)$ theory, with a different tuning of the moduli. With this tuning, we find a $\mathbb{Z}_3$ gauge symmetry.

A related example was given in section 2.7.1 using the action of $\mathbb{Z}_6$ on a nonsingular elliptic curve with $j = 0$. The quotient $S/\mathbb{Z}_6$ has central fiber which is a rational curve passing through surface singularities of types A_1, A_2, and A_5 with fibers of multiplicity 2, 3, and 6 over those points. Resolving the singularities leads to another affine $\widehat{E}_8$ diagram:



where again the solid circles represent the curves being blown down. Tuning the moduli of the $\mathcal{N} = (2,0)$ E_8 theory to blow down those curves leads to a locus in moduli with a $\mathbb{Z}_6$ gauge symmetry. Notice that this is the intersection of the previous two loci with $\mathbb{Z}_2$ and $\mathbb{Z}_3$ gauge symmetry.

2.9.3 Towards T-Duality in the More General Case

In the examples from the previous two subsections, we saw that the expected appearance of T-duality for an LST motivates the search for a double elliptic fibration structure in such F-theory models. When the base has a fibration by curves of genus 1, the origin of this second fibration is clear. When the fibration on the base is by curves of genus 0, however, the T-duality is not as readily manifest.

Example (2.9.2) does have manifest T-duality, as further analyzed in Appendix A.6. In this case, the Weierstrass model $y^2 - F(x, \psi, [s,t])$ can be regarded as a double cover of a $\mathbb{P}^1$-bundle

over the base B, where x denotes the coordinate on the $\mathbb{P}^1$. Since the double cover is branched at 4 points along each fiber $\mathbb{P}^1_F$ of this fibration, the total space gets a fibration by curves of genus 1.

Now the base B of that F-theory model has a fibration $\pi : B \to C$ of its own whose general fiber is another $\mathbb{P}^1$ which we shall call $\mathbb{P}^1_B$. In Appendix A.6, after making a birational modification of the base (i.e., blowing it up and down), we find that the double cover is branched along 4 points of each fiber $\mathbb{P}^1_B$ as well, and this implies that there is a *second* genus 1 fibration on the total space. Note that this echoes the discovery made in [15] that heterotic T-duality, when viewed from the perspective of F-theory, exchanges the roles of base and fiber in the heterotic weak coupling limit.

It is therefore natural to seek out a more general geometric exchange symmetry in LSTs with T-duality. We leave a more complete investigation of this possibility to future work.

2.10 Outliers and Non-Geometric Phases

Much as in the case of the classification of SCFTs achieved in reference [10], we view the F-theory realization of LSTs as providing a systematic approach to the construction of such models. In some cases, however, we indeed find a few small gaps between what is expected based on field theory considerations, and what can be obtained in geometric phases of F-theory.

Our plan in this short section will be to proceed mainly by effective field theory considerations to give a list of such outlying behavior, both for 6D SCFTs and LSTs. In this sense, we will present a complete classification of UV complete 6D phenomena decoupled from gravity, though we hasten to add that some of these putative theories may end up being inconsistent due to the lack of an embedding in an F-theory (or other string) construction.[15] In some cases, however, this also points to a few additional novel possibly non-geometric structures. Though we shall comment on possible ways to realize these models in F-theory, we leave a more complete analysis to future work.

2.10.1 Candidate LSTs and SCFTs

From a bottom up perspective, the primary constraints on the construction of consistent LSTs are the existence of a lattice of string charges with a negative semidefinite Dirac pairing, and possible gauge groups "decorating" the associated tensor multiplets.

In some cases, there is a clear indication from F-theory that certain bottom up considerations are too weak; For example, the phenomenon of a $-n$ curve for $n \geq 3$ always implies the existence of a non-trivial gauge group factor, a condition which is not obvious from any anomaly cancellation condition.

There are, however, two intermediate cases suggested by field theory which also have a potential realization in string theory. The first case deals with the gauge-gravitational anomaly

[15] For a recent example of a seemingly consistent 6D SCFT which is actually inconsistent, see e.g. Appendix A of reference [63].

cancellation condition (see (A.1.7)) imposed in all geometrically realized 6D SCFTs and LSTs. This condition has no field theoretic analog in flat space as there are no a priori restrictions on one-loop contribution to mixed gauge-gravitational anomaly. Non-vanishing of total mixed gauge-gravitational anomaly implies that the theory is inconsistent when put on a fixed curved spacetime background. It would clearly be troublesome if 6D SCFTs were anomalous in this way. Fortunately, it is always possible to cancel the one-loop contribution to this anomaly against a Green-Schwarz contribution, as was demonstrated in [64]. This argument does not apply to 6D LSTs, and indeed it can be checked explicitly in many examples of 6D LSTs that there is no way to cancel the one-loop mixed gauge-gravitational anomaly using the Green-Schwarz mechanism. This means that 6D LSTs cannot always be put on a fixed curved spacetime.

One class of such models would arise on:

$$12...21 \tag{2.10.1}$$

where we decorate all the tensors with $\mathfrak{su}$-type gauge groups along with anti-symmetric matter for the last $\mathfrak{su}$ gauge group and symmetric matter for the first $\mathfrak{su}$ gauge group. With the presence of this symmetric representation, (A.1.7) is violated. Nonetheless, this model can be constructed by putting type IIA on an interval $S^1/\mathbb{Z}_2$ with an $O8^-$ orientifold plane on each fixed point, D6s stretched along the interval, NS5s embedded in the D6s at various points along the interval, and two of the NS5s stuck respectively at the two fixed points.

More generally, to construct new examples of 6D SCFTs and LSTs which violate the 1-loop mixed gauge-gravitational anomaly condition (A.1.7) we can do the following: take any F-theory model having a -1 curve or P base curve associated with $\mathfrak{su}$-type gauge group and at least one anti-symmetric hyper and 16 fundamental hypers not transforming under any other gauge group. Then, replace this set of hypermultiplets with a single hypermultiplet in the symmetric representation. The resulting theory will satisfy gauge anomaly cancellation, but it will have a non-vanishing 1-loop gauge-gravitational anomaly.

The second case has to do with the condition of "normal crossing" which is present in all geometrically realized 6D SCFTs, and is only mildly violated in LSTs. For example, we have seen that an intersection with an order two tangency $4||1$ leads to a consistent LST base. In the bottom up perspective, we have a Dirac pairing which has 2 on the off-diagonal entries, and -4 and -1 on the diagonals. Generalizing, we can consider constructions such as:

$$4||2...2 \tag{2.10.2}$$

$$4||2...21 \tag{2.10.3}$$

$$4||2...2||4 \tag{2.10.4}$$

all of which have a negative definite Dirac pairing. Though we have not encountered these possibilities in our discussion of geometric phases of F-theory, the first two have been realized in IIA string theory via appropriate suspended brane configurations (see e.g. [18]), at least when there are non-trivial gauge group factors over the associated tensor multiplets. For example, we can have $\mathfrak{su}$-type on the -2 charge tensors, and $\mathfrak{so}$-type gauge groups on the -4 tensors, with $\mathfrak{sp}$-type gauge groups on the -1 tensors. The crucial ingredient appearing in these suspended brane models is an $O8^+$-orientifold, rather than an $O8^-$-orientifold. The T-dual description in

IIB string theory involves $O7^+$-orientifold planes, a case which leads to some non-geometric behavior, a point we shall return to later. The last example $4||2...2||4$ resists an embedding in IIA string theory, since it would appear to involve two $O8^+$-orientifold planes. Indeed, if we attempt to decorate these tensor multiplets by $\mathfrak{so}$-type algebras on the -4 tensors, and $\mathfrak{su}$-type algebras on the -2 tensors, we cannot cancel gauge theoretic anomalies.

Assuming that configurations such as $4||2$ can indeed occur in the construction of LSTs, it is natural to ask how many additional models can be obtained, at least from a bottom up perspective. Pairing each tensor multiplet with a simple gauge algebra, these are as follows:

$$\begin{array}{ccccccccc} \mathfrak{so}(M) & & \mathfrak{su}(N_1) & \mathfrak{su}(N_2) & & & \mathfrak{su}(N_k) & \mathfrak{sp}(N_R) \\ 4 & || & 2 & 2 & ... & & 2 & 1 \end{array} \tag{2.10.5}$$

$$\begin{array}{ccccccccc} & & & \mathfrak{sp}(N_T) & & & & & \\ & & & 1 & & & & & \\ \mathfrak{sp}(N_1) & & \mathfrak{so}(M_1) & & & \mathfrak{sp}(N_k) & \mathfrak{so}(M_k) & & \mathfrak{su}(N_R) \\ 1 & & 4 & & ... & 1 & 4 & || & 2 \end{array} \tag{2.10.6}$$

$$\begin{array}{ccccccccc} \mathfrak{su}(N') & \mathfrak{sp}(N_1) & \mathfrak{so}(M_1) & & & & \mathfrak{so}(M_k) & & \mathfrak{su}(N_R) \\ 2 & 1 & 4 & & ... & & 4 & || & 2 \end{array} \tag{2.10.7}$$

It should be noted that gauge and mixed anomalies strongly constrain the allowed gauge algebras in the above list. In particular, the configuration $\overset{\mathfrak{so}(M)}{4} || \overset{\mathfrak{su}(N)}{2}$ is constrained by mixed anomalies to have a bifundamental $(\mathbf{M}, \mathbf{N})$, requiring $N \leq M-8$, $2N \geq M$. These conditions lead to strong constraints on the rest of the gauge algebras in the aforementioned theories, as discussed in [11].

If one also relaxes the condition that there is a gauge group paired with each tensor, even more constructions are possible. In most of these cases, no known embedding in a string construction is available, so we suspect that at least some of these theories are actually inconsistent. Nevertheless, for the sake of completeness, we list them here:

$$2 \;||\; 4 \;||\; 2 \tag{2.10.8}$$

$$4 \; 1 \; 4 \;||\; 2 \; 2 \tag{2.10.9}$$

$$4 \;||\; 2 \; 2 \; ... \; 2 \;||\; 4 \tag{2.10.10}$$

$$4 \;||\; 2 \; 2 \; ... \; 2 \; 1 \tag{2.10.11}$$

$$4 \;||\; 2 \; 2 \; ... \; \overset{2}{2} \; 2 \tag{2.10.12}$$

$$1 \overset{1}{4} 1 \; ... \; 4 \; 1 \; 4 \;||\; 2 \tag{2.10.13}$$

$$1 \; 3 \; 1 \; ... \; 4 \; 1 \; 4 \;||\; 2 \tag{2.10.14}$$

$$1 \; 2 \; 3 \; 1 \; 4 \; ... \; 4 \; 1 \; 4 \;||\; 2 \tag{2.10.15}$$

$$1 \; 2 \; 2 \; 3 \; 1 \; 4 \; ... \; 4 \; 1 \; 4 \;||\; 2 \tag{2.10.16}$$

$$2 \; 1 \; 4 \; ... \; 4 \; 1 \; 4 \;||\; 2. \tag{2.10.17}$$

It is worth mentioning that a violation of normal crossing and a violation of gauge-gravitational anomaly cancellation do not appear simultaneously in any of these examples. The lists of LSTs and putative LSTs arising from these violations are small and tightly constrained.

This perspective on LSTs also points to the existence of some additional novel structures for

6D SCFTs. Indeed, starting from an LST, we can consider a combination of tensor-decoupling and Higgsing to reach some additional candidate SCFTs.

The theories of this type, which are consistent with anomaly cancellation and with no unpaired tensors, take the form:

$$\overset{\mathfrak{so}(M)}{4} \;||\; \overset{\mathfrak{su}(N_1)}{2} \; \overset{\mathfrak{su}(N_2)}{2} \; \ldots \; \overset{\mathfrak{su}(N_k)}{2} \tag{2.10.18}$$

$$\overset{\mathfrak{sp}(N')}{1} \; \overset{\mathfrak{so}(M)}{4} \;||\; \overset{\mathfrak{su}(N_1)}{2} \; \overset{\mathfrak{su}(N_2)}{2} \tag{2.10.19}$$

$$\overset{\mathfrak{so}(M_1)}{4} \; \overset{\mathfrak{sp}(N_1)}{1} \; \ldots \; \overset{\mathfrak{so}(M_k)}{4} \;||\; \overset{\mathfrak{su}(N_R)}{2} \tag{2.10.20}$$

$$\overset{\mathfrak{sp}(N_1)\mathfrak{so}(M_1)}{1} \; 4 \; \ldots \; \overset{\mathfrak{so}(M_k)}{4} \;||\; \overset{\mathfrak{su}(N_R)}{2} \tag{2.10.21}$$

Notice that (2.10.19) does not admit a known type IIA construction, whereas the other three do. Another curious thing to notice is that this model does not admit an embedding in a putative LST. The obvious examples of attaching an $\mathfrak{so}$ group to the left or attaching an $\mathfrak{su}$ group to the right are pathological because they necessarily have non-vanishing quartic part of gauge anomaly. Although we have only proved that every 6D SCFT can be embedded in a 6D LST for models arising geometrically within F-theory, we expect this statement to be true in general. Hence, we suspect that the putative SCFT (2.10.19) is inconsistent. Finally, concentrating only on positive-definiteness, we have the $4||2$ configurations:

$$1\ 4\ ||\ 2\ 2 \tag{2.10.22}$$

$$4\ ||\ 2\ 2\ \ldots\ 2 \tag{2.10.23}$$

$$1\ 4\ 1\ \ldots\ 4\ 1\ 4\ ||\ 2 \tag{2.10.24}$$

$$4\ 1\ 4\ \ldots\ 4\ 1\ 4\ ||\ 2 \tag{2.10.25}$$

$$3\ 1\ \ldots\ 4\ 1\ 4\ ||\ 2 \tag{2.10.26}$$

$$2\ 3\ 1\ 4\ \ldots\ 4\ 1\ 4\ ||\ 2 \tag{2.10.27}$$

$$2\ 2\ 3\ 1\ 4\ \ldots\ 4\ 1\ 4\ ||\ 2 \tag{2.10.28}$$

2.10.2 Towards an Embedding in F-theory

In the above, we encountered some additional candidate SCFTs and LSTs which appear to be consistent with effective field theory considerations. Additionally, some of these models admit an embedding in type IIA suspended brane constructions.

It is therefore important to ask to what extent we should expect F-theory to cover this and related examples. Though we leave a complete characterization to future work, there are some general ingredients we can already identify which point the way to incorporating these additional non-geometric structures.

As we have already mentioned, one crucial ingredient in the IIA realization of the $4||2$ configuration is the appearance of an $O8^+$-plane, which T-dualizes to a pair of $O7^+$-planes in type IIB string theory. Such orientifolds lead to the phenomena of "frozen singularities" in F-

theory [23–25]. These are models in which the monodromy of the axio-dilaton around the brane is consistent with that of an appropriate I_n^* singularity, but in which the corresponding gauge algebra is *not* $\mathfrak{so}(2n+8)$.

Another not entirely unrelated phenomenon we have encountered in the construction of the $\mathcal{N} = (1,1)$ LSTs, as well as in some of the low rank LSTs are models in which the normal bundle of a curve on the base is torsion of finite order. To produce a Weierstrass model, we have found it necessary to impose specific restrictions on the order of these torsion bundles, though the M-theory realization of these A-type $\mathcal{N} = (1,1)$ theories suggests a whole family of models parameterized by rational theta angles [47].

In fact, it is relatively straightforward to engineer all of the A-type $\mathcal{N} = (1,1)$ LSTs in type IIB, and to lift this back to F-theory. For example, in type IIB language, we have a stack of N D7-branes wrapped on a T^2. Switching on a background value for a flat RR and NS two-form potential, we get additional theories parameterized by the ratio of these two periods. In F-theory language, we see this by a choice of how we resolve the affine node of the $\widehat{A}_{N+1}$ Dynkin diagram. In physical terms, this resolution comes from compactifying an 8D model on an additional S^1. Going down on a T^2, we have Wilson lines for this affine node along the A- and B-cycles of the T^2.

The presence of such background B-fields also suggests that similar effects from discrete group actions may also make an appearance in the construction of the $4||2$ type configurations. For example, at the level of the effective field theory, we can consider a base:

$$\overset{2}{2}22...2 \tag{2.10.29}$$

with I_N-type fiber decorations on each -2 curve. Now, this field theory admits a $\mathbb{Z}_2$ automorphism in which we combine the outer automorphism of the D-type base with an outer automorphism on the $\mathfrak{su}_N$ factors on the leftmost -2 curve and the top -2 curve. At the level of gauge theory, the outer automorphisms of $\mathfrak{su}_N$ can take us to either an $\mathfrak{so}$ or $\mathfrak{sp}$-type algebra. Combining these two operations, we see that we get an effective field theory where there is one less tensor multiplet, and in which the BPS charge has doubled, and in which the Dirac pairing between this $\mathbb{Z}_2$ invariant combination and its neighbor is 2, leaving us with a configuration of tensor multiplets $4||2...2$. Taking into account the algebra assignment (i.e. the would-be fibers), we can in principle have either $\mathfrak{so}$ or $\mathfrak{sp}$ -type algebras, of which only the former is compatible with anomaly cancellation.[16] Similar considerations also apply for the LST tensor multiplet configurations such as $4||2...21$.

Let us stress that effective field theory considerations do not directly inform us of the actual non-geometric realization of these models in F-theory. Indeed, what is particularly remarkable is that *even if* we allow these additional structures, the total number of additional candidate SCFTs and LSTs is quite small, with the vast majority being covered by geometric phases of

[16]Let us recall that in F-theory, conjugation by an outer automorphism apparently leads to a geometrically ambiguous assignment for the quotient algebra. This ambiguity has been resolved by appealing to anomaly cancellation considerations, in which the opposite conclusion is reached, i.e. a non-split I_N fiber realizes an $\mathfrak{sp}$-type algebra [65]. However, effective field theory considerations suggest that when combined with a quotient on the tensor multiplets, there may be a generalization of this construction available in which we instead reach an $\mathfrak{so}$-type algebra. Incorporation of an $O7^+$ plane or of discrete B-fields might provide a route to such a generalization.

F-theory. This suggests that whatever the mild deformation of known F-theory backgrounds are that produce these models, the structures encountered in this chapter and in earlier work remain quite robust.

2.11 Conclusions

In this chapter we have given a systematic approach to realizing supersymmetric little string theories via compactifications of F-theory. Much as in the case of 6D SCFTs, these theories arise from working with F-theory on a non-compact base, in which some collection of curves are simultaneously collapsed to small size. The key difference with a 6D SCFT is that the intersection pairing for these curves defines a positive semidefinite quadratic form on the lattice of string charges. So, in contrast to the case of SCFTs the associated theories contain a dimensionful parameter which is naturally promoted to a non-dynamical tensor multiplet. After spelling out all necessary conditions to geometrically realize LSTs in F-theory, we have given a classification of all such theories. On the one hand, these theories can all be viewed as arising from extending 6D SCFTs by one or more additional curves. As such, they also admit an atomic classification, much as in reference [10]. We have also seen that the general expectation that all 6D SCFTs embed in an LST is indeed realized via the explicit embedding in an F-theory compactification. Finally, we have seen that T-duality of an LST is realized via a double elliptic fibration in the corresponding F-theory model. In the remainder of this section we discuss some avenues for future investigation.

Perhaps the most important issue left open by our analysis is the small gap between theories realized by geometric phases of F-theory, and the list of effective field theories which can potentially complete to an LST (or SCFT). It would be interesting to establish to what extent non-geometric deformations can enter in such F-theory models, and conversely, how many of these putatively consistent LSTs (and SCFTs) are actually excluded by further non-trivial consistency conditions.

One of the key simplifications in our analysis of LSTs is that decoupling any curve in the base takes us back to a collection of (possibly decoupled) 6D SCFTs and scale invariant theories (i.e. when we have free vector multiplets). This strongly suggests that the common notions of renormalization group flows for local quantum field theories extend to *non-local* LSTs. Developing the details of such a structure would provide a rather striking vantage point on what it means to "UV complete" a quantum field theory in the first place.

As a preliminary step in this direction, it is also natural to ask whether there is a notion of monotonic loss in the degrees of freedom in such conjectural flows from LSTs to SCFTs. For example, in many of the cases studied in this chapter, we can weakly gauge both diffeomorphisms as well as an $SU(2)$ field strength, which in the context of a 6D SCFT would be identified with the R-symmetry of the theory. It is tempting to conjecture that there is a formal extension of conformal anomalies to these cases as well. It would be interesting to study whether an extension of the methods presented in references [66, 67] (see also [68]) would provide insight into such generalizations of renormalization group flows.

Finally, one of the hallmarks of the systems we have encountered is the appearance of an effective T-duality upon compactification on a further circle. Given that there are now concrete methods for extracting the partition functions for some 6D SCFTs (see e.g. [69]), it would be quite natural to study this and related structures for LSTs.

Chapter 3

The Frozen Phase of F-theory

3.1 Introduction

F-theory [70–72] is a geometrical way to describe non-perturbative backgrounds of type IIB string theory, whose transition functions include S-duality in addition to the more usual symmetries. Supersymmetric backgrounds of F-theory describe a spacetime which includes the base of an elliptic Calabi–Yau variety, with a variable axio-dilaton field whose value is specified by the elliptic fibration. The degeneration loci of the fibration, called the irreducible components of the discriminant locus, are interpreted as seven-branes on which various gauge algebras are realized. Among these, one finds as particular examples the ordinary D7-branes and O7-planes of perturbative IIB theory.

The perturbative definition of O-planes, however, allows for several different variants. As was pointed out in the early days of F-theory [23], the one reproduced in conventional F-theory is the O7$_-$-plane, whose charge equals -4, in units where a (full) D7-brane has charge 1. Another type, called the O7$_+$-plane, has charge $+4$. At a fixed total charge, such an object allows for fewer deformations; for example, an O7$_-$ with 8 D7s on top, with total charge 4, can be deformed in various ways by pulling the D7s away, while a single O7$_+$ with the same charge does not allow for such a possibility. The F-theory description of the latter should hence involve a divisor which for some reason cannot be deformed, and was thus called *frozen* in [23], where this was also discussed in several dual frames. This phenomenon was then further investigated in [24].

Thus it was known for a long time that F-theory includes O7$_+$-planes but they were basically ignored in the vast existing literature on the compactifications of F-theory. One motivation for revisiting this issue at present rests in the classifications of six-dimensional superconformal theories (SCFTs). In a series of works initiated in [9], and in particular in [10], it was shown that *almost* all known 6d SCFTs at that time and a lot more were realizable using 6d compactifications of F-theory. However, if one compares this classification against the known examples constructed using massive IIA brane constructions [73, 19, 18, 74] and the purely-field theoretical analyses [75, 11], one recognizes that there are indeed cases not realized by conventional F-theory constructions.

A typical feature of these cases is that their massive IIA brane construction involves O8$_+$s.

By a T-duality, this is mapped to a IIB brane construction involving $O7_+$s. This motivated us to look at F-theory compactifications to six-dimensions in the presence of $O7_+$s.

At this point, it is natural to worry if there could be frozen singularities other than $O7_+$-planes which have not been studied in conventional F-theory. This question was settled, at least for supersymmetric seven-branes, in a recent re-analysis of 7-branes in F-theory [25] which concluded that the $O7_+$ is in fact the only type of frozen singularity in F-theory.[1] Therefore, the only ingredient missing in conventional F-theory compactifications to six-dimensions is the inclusion of $O7_+$-planes, and indeed including them we find F-theory realizations of 'missing' 6d SCFTs, as we will see later in the chapter.

Once we are convinced that $O7_+$-planes can be included in the F-theory construction, we realize that we need to revisit every part of the standard F-theory machinery, such as the assignment of the gauge algebras and of the matter content to the components of the discriminant and to their intersections, and the way the 6d anomalies cancel via the Green–Schwarz-West-Sagnotti effect [79, 80], derived geometrically for F-theory by Sadov in [52]. This chapter is the first attempt to provide such generalizations.

One unexpected consequence of the introduction of $O7_+$-planes is the following. To appreciate it, let us first recall the situation *without* $O7_+$-planes. In a conventional F-theory compactification without $O7_+$-plane, once one is given the geometry of the elliptically-fibered Calabi–Yau, there is a standard method to assign a unique set of gauge algebras and matter content to the geometry. In particular, under this standard assignment, each simple factor in the gauge algebra is associated to a single component of the discriminant divisor, and each component has at most one simple factor of gauge algebras associated to it. This choice corresponds to having zero holonomies of the gauge fields on these divisors themselves. We have the option of turning on the non-trivial gauge configurations, including the effects often called the T-branes [81], but we also have the standard option of not turning them on at all.

With $O7_+$-planes, however, we will often be forced to have at least some nontrivial gauge configurations on some of the components. More precisely, we even lose the concept of a unique, standard assignment of gauge algebras and matter content, since we do not even have a natural origin in the space of the all possible holonomies. Because of this, we often have multiple simple factors of gauge algebras on a single component of the discriminant locus, and also a single simple factor of gauge algebra shared across multiple components, as we will see later.

Unfortunately, at present, we do not have any algorithmic method to find consistent assignments given an elliptic Calabi–Yau and a specification of where the $O7_+$-planes are; we do not even have a method to tell if there are any consistent assignments at all. Therefore we are forced to rely on consistency checks via anomaly cancellation and dualities to backgrounds that are better understood.

The rest of the chapter is organized as follows. In Sec. 3.2, we study the properties of $O7_+$-planes in the context of F-theory, using string theory and M-theory dualities. This will let us figure out how to assign gauge algebras and matter content. In Sec. 3.3, we study the anomaly

[1] There are various other less-studied types of higher-codimension singularities one can incorporate in F-theory, such as the ones used by García-Etxebarria and Regalado [76] to construct 4d $\mathcal{N}=3$ SCFTs. Frozen versions of singularities also occur in M-theory [23,24], where they play an important role in M5-brane fractionation [24,77,78].

cancellation of F-theory models with O7$_+$-planes. We will see that the analysis of Sadov [52] can be naturally generalized by introducing a divisor which represents where O7$_+$-planes lie. Then in Sec. 3.4, we discuss some 6d SCFTs which can be realized only with O7$_+$-planes in F-theory construction, and in Sec. 3.5, we study the massless spectrum of a couple of compact six-dimensional models with O7$_+$-planes.

In Appendix B.1, we review the 8d compactifications with O7$_+$-planes, which is simpler than the 6d examples discussed in the main text. Finally in Appendix B.2, we give an alternative derivation, using intersecting brane models, of the spectrum of some compact models discussed in Sec. 3.5.

3.2 Frozen seven-branes and their properties

In this section, we use perturbative string techniques to obtain some properties of frozen singularities.

We start in section 3.2.1 with a lightning review of O-planes. We then discuss the basics of O7$_+$-planes in F-theory in Sec. 3.2.2, and in Sec. 3.2.3 we study the physics at individual intersection points of O7$_+$-planes and other seven-branes. To prepare ourselves for the analysis of an O7$_+$-plane which intersects with more than one seven-brane, we then need to have short digressions, on the T-duals of NS5- and D6-branes in Sec. 3.2.4 and on the phenomenon of shared gauge algebras in Sec. 3.2.5. We then come back to the case with O7$_+$-planes in Sec. 3.2.6. In the final subsection 3.2.7, we see that with O7$_+$-planes a shrunken divisor does not necessarily signify any singularity in the low energy physics.

3.2.1 Basics of orientifold planes

Let us start by a quick review of the basics of the orientifolds.[2]

Action on the closed strings: An orientifold is usually defined as a $\mathbb{Z}_2$ symmetry Π that includes world-sheet parity Ω. It can also include a spacetime involution σ. It is often necessary to also include an extra factor $(-)^{F_L}$ (where F_L is the left-moving spacetime fermion number) so that Π^2 acts as the identity. If locally σ is the reflection of $9-p$ coordinates, so that the *orientifold plane* Op (the fixed locus of σ)[3] has (spatial) dimension p, one needs to include $(-)^{F_L}$ if $p = 2, 3$

[2] A good review of the basics can also be found in [82]. More detailed and rigorous analysis of perturbative orientifolds were done e.g. in [83, 84], but we stick to the traditional, ad hoc approach in this chapter. The name *orientifold* itself was introduced in [85] by Dai, Leigh and Polchinski. The concept of the orientifold goes back further in history, see e.g. [86, 87] and references therein.

[3] We will also consider actions that include translations and thus have no fixed locus as in (3.2.5); the conclusions in (3.2.1) below also apply.

mod 4.[4] To summarize, locally the orientifold action is

$$\begin{array}{cccccc} \text{O9} & \text{O8} & \text{O7} & \text{O6} & \text{O5} & \cdots \\ \hline \Omega & \Omega R_9 & \Omega R_8 R_9 (-1)^{F_L} & \Omega R_7 R_8 R_9 (-1)^{F_L} & \Omega R_6 R_7 R_8 R_9 & \cdots \end{array}, \tag{3.2.1}$$

where R_p denotes a reflection of the p-th coordinate. This specifies the orientifold's action on closed strings. In this chapter, we will be interested in particular in O7s, with O6s and O8s making occasional appearances.

Action on the open strings: In presence of open strings, one also needs to decide its action on the Chan–Paton matrix λ, which appears in a superposition $\sum_{i,j} \lambda_{ij} |ij\rangle$ of the states $|ij\rangle$, that in turn can be interpreted as going from the i-th to the j-th brane in a stack (omitting other quantum numbers). Since the world-sheet parity Ω reverses orientation, it acts by transposing Λ, but it may also mix the states with a change of basis M: namely, $\lambda \to M \lambda^t M^{-1}$. Imposing that this action is an involution leads to the condition that

$$M^{-1} M^t = \mp 1 \,. \tag{3.2.2}$$

This sign choice leads to two different types of O-plane, which we call O$p_\pm$.[5]

The RR-charge: The RR charge can be computed through a one-loop computation, which contains $-\operatorname{tr} M^{-1} M^t$ in its Möbius strip contribution (see for example the reviews [82, 88]). In the end one concludes that the charge is $\pm 2^{p-5}$ that of a full Dp-brane: explicitly,

$$\begin{array}{c|cccccccc} p & 9 & 8 & 7 & 6 & 5 & 4 & 3 & \cdots \\ \hline \pm 2^{p-5} & \pm 16 & \pm 8 & \pm 4 & \pm 2 & \pm 1 & \pm \frac{1}{2} & \pm \frac{1}{4} & \cdots \end{array} \tag{3.2.3}$$

Thus, the Op_- has negative charge and the Op_+ has positive charge, as the name implies.

The gauge group: The gauge group is also influenced by the sign (3.2.2). If a stack of N Dp-branes is parallel to the Op-plane but not on top of it, the action will relate the strings ending on them to strings ending on an image stack in a different locus; the gauge group will be the usual $\mathrm{U}(N)$. On the other hand, if the stack is on top of the Op-plane, the action will relate the open string states to themselves, projecting out some of them. To read off the gauge group, we can consider the gauge field states $\lambda_{ij} \alpha^\mu_{-1/2} |0; ij\rangle$. Since $\Omega \alpha_{-1/2} \Omega = -\alpha_{-1/2}$, the surviving states will be those with Chan–Paton factors λ such that $\lambda = -M \lambda^t M^{-1}$. If the sign in (3.2.2) is -1, M is antisymmetric; by a change of basis ($\lambda \to C^{-1} \lambda C$, $M \to C M C^t$) it can be chosen to be

[4] To check this, one first uses the fact that a reflection R_I of the I-th spatial coordinate acts by Γ_I on the 10d Majorana spinor, which satisfies $(\Gamma_I)^2 = +1$. Therefore, $R^2_{I_1 \cdots I_p} = 1$ or $(-1)^{F_L + F_R}$ depending on whether $p = 0, 1$ or $2, 3$ mod 4, respectively. Then one compensates this $(-1)^{F_L+F_R}$ by the fact that $\Omega (-1)^{F_L} \Omega = (-1)^{F_R}$ and therefore $(\Omega (-1)^{F_L})^2 = (-1)^{F_L + F_R}$.

[5] In [23] and other older papers, O$p_\pm$-planes are called planes of type $\mathcal{O}^\mp$, with the opposite sign. We stick to the more modern conventions which are now standard.

of the form $J \equiv \left(\begin{smallmatrix} 0 & 1_N \\ -1_N & 0 \end{smallmatrix}\right)$, and thus λ will be in the $\mathfrak{sp}_N$ algebra.[6] If on the other hand the sign in (3.2.2) is $+1$, then M can be chosen to be 1_{2N}, and $\lambda \in \mathfrak{so}_{2N}$.

Summarizing, the choice (3.2.2) leads to two different orientifolds:

- Op_-, with $\mathfrak{so}_{2N}$ gauge algebra and charge -2^{p-5}, and
- Op_+, with $\mathfrak{sp}_N$ gauge algebra and charge $+2^{p-5}$.

Dq-branes intersecting Op-planes: More generally, if we also have a stack of Dq-branes which intersect our Op, there are subtle signs [8] coming from the fact that the strings from the Op- to the Dq-branes needed to be expanded to both integer and half-integer modes. In flat space (and vanishing B field), the number #ND of Neumann–Dirichlet directions (the number of directions transverse to the Dp and parallel to the Dq, or vice versa) has to be a multiple of 4, for unbroken supersymmetry. The result for the gauge algebra on the Dq-branes is then as follows:[7]

	Op_+	Op_-
#ND $= 0, 8$	symplectic	orthogonal
#ND $= 4$	orthogonal	symplectic

. (3.2.4)

T-duality: Let us next discuss the T-duality of orientifolds, since we often need to perform T-duality of the setup on $S^1/\mathbb{Z}_2$ where two fixed points support Op-planes, possibly of different types. Two most straightforward cases are when both fixed points have Op_- or both fixed points have Op_+. The T-dual is then simply O$(p+1)_-$ or O$(p+1)_+$ wrapped around S^1.

When one fixed point has Op_- and the other fixed point has Op_+, the T-dual is known to be a *shift-orientifold*, namely an orientifold whose spacetime action σ not only flips the coordinates transverse to the orientifold, but also translates a circle by half its radius

$$\sigma : (x_{p+1}, x_{p+2}, \cdots, x_9) \sim \left(x_{p+1} + \frac{R}{2}, -x_{p+2}, \cdots, -x_9\right) . \tag{3.2.5}$$

See Fig. 3.2.1 for a pictorial representation. Note that this action fixes no point.

The derivation of this fact can be found e.g. in [23, p. 41] or [89]. A rough argument goes as follows. We start from the shift-orientifold background (3.2.5), and T-dualize the x_{p+1} direction. Its T-dual should be a compactification on $S^1/\mathbb{Z}_2$. Therefore this should result in a combination of two Op-planes at two fixed points. The original shift-orientifold background did not have any D$(p+1)$-charge. Therefore, in the T-dual, we should have zero Dp-plane charge. This is only possible if one fixed point is Op_- and the other is Op_+.

[6] We follow the standard convention that $\mathfrak{sp}_1 = \mathfrak{su}_2$.

[7] The fields on the Dq stack get mapped to fields on another point of the stack, unless the Dq stack is completely embedded in the Op-plane. A priori this only restricts the behavior as a function of the coordinates of the gauge field, which would then locally remain of $\mathfrak{u}(2m)$ type. However, in situations where the divisor wrapped by the stack is compact, in most applications we want to keep only the zero-modes of the gauge field under its equation of motion, and this restricts the gauge group as in (3.2.4).

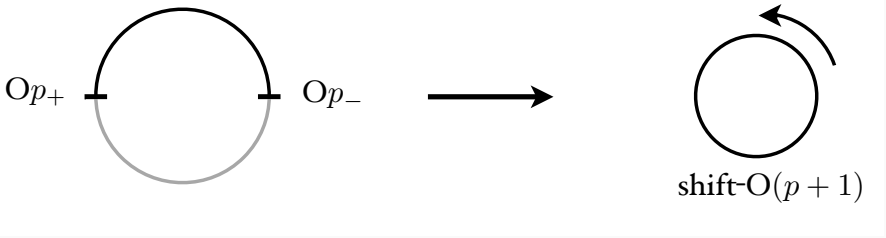


Figure 3.2.1: A model with two Op-planes with opposite sign is turned by T-duality.

Another intuitive argument is as follows. The shift operator $s : x_p \to x_{p+1}$ can be thought of as $e^{i\frac{R}{2}\widehat{p}}$, where $\widehat{p}$ is the momentum operator. Its T-dual is $\tilde{s} = e^{i\frac{1}{2R}\widehat{w}}$, where $\widehat{w}$ is the "winding operator", which measures the length of the string. $\tilde{s}$ gives 1 on strings of total length zero, such as those that begin and end on the same Op, but it gives -1 on the strings that begin and end on different Op's, signaling the fact that the two have different signs.

Other types of orientifolds: It is also known that there are $\tilde{\text{O}}p_\pm$-planes when $p \leq 6$, distinguished from the more ordinary O$p_\pm$-planes by the RR-torsion flux. As we will not use them heavily, we will not discuss them further.

3.2.2 Frozen divisors in F-theory

Our main interest lies in seven-branes in Type IIB theory and F-theory. An ordinary O7$_-$ without any D7-branes on top is known to lift to two I_1 divisors, due to quantum effects [90]. Similarly, with $n < 4$ D7-branes on top, the F-theory realization is given by $(n+2)$ I_1 divisors. With at least 4 D7-branes, it is interpreted in F-theory as an I^*_{n-4} divisor (where n is the number of D7-branes). Since string theory also has O7$_+$-planes, it is natural to ask how these are described in F-theory.

First of all, from (3.2.3) we see that O7$_\pm$ have charge equal to that of ± 4 full D7-branes. So an O7$_+$ has the same charge and tension as an O7$_-$ with 8 full D7-branes on top. In F-theory, they will give rise to the same monodromy [23, 91]; we expect both to be described by an I^*_4 divisor. However, the O7$_-$ with 8 D7 gives rise to an $\mathfrak{so}_{16}$ gauge algebra, while the O7$_+$ gives rise to none. A related difference is that the O7$_-$ with 8 D7 can be deformed by pulling the D7s away (which corresponds in F-theory to a complex structure deformation), while the O7$_+$ cannot. Thus an O7$_+$ is described by a I^*_4 singularity which for some reason cannot be deformed; we will call this a *frozen* singularity, and denote it by $\widehat{I}^*_4$.

More generally, an O7$_-$ with n D7s has the same charge and tension as an O7$_+$ with $(n-8)$ D7s; both are described by an I^*_{n-4} singularity, but in the latter case the gauge algebra is $\mathfrak{sp}_{n-4}$ rather than $\mathfrak{so}_{2n}$, and the deformations are correspondingly reduced. In this case too we say that the singularity is frozen, and we denote by $\widehat{I}^*_{n-4}$.

To be more expicit, an F-theory vacuum is typically described by the "Weierstrass coefficients" f and g which are sections of the line bundles $\mathcal{O}_B(-4K_B)$ and $\mathcal{O}_B(-6K_B)$ on the F-theory base B, and which lead to the equation

$$y^2 = x^3 + fx + g \tag{3.2.6}$$

for the total space of the elliptic fibration. Along a divisor D with a $\widehat{I}^*_{n-4}$ singularity, f vanishes to order 2, g vanishes to order 3, and the equation $4f^3 + 27g^2$ of the discriminant locus vanishes to order $(n-8)+10$, for a configuration with $n-8$ D7-branes on top of an O7$_+$. Although the "freezing" mechanism is not understood, it must prevent any deformation with lowers the order of vanishing of either f or g at all, or which lowers the order of vanishing of $4f^3 + 27g^2$ below 10.

Note that the Weierstrass coefficients are accompanied by periods of type IIB two-forms over appropriate two-cycles in B; for compactifications to 6d, the complex moduli provided by Weierstrass coefficients are paired with these periods of two-cycles to provide the two complex scalars in a hypermultiplet. In particular, by activating a vev represented by one of these two-form periods we may disturb the gauge group assigned to a divisor without changing the geometry of the divisor (which would have required a change of complex modulus). Such deformations are often described in the language of T-branes [81], for which a number of geometric tools have been developed [92–94].

As an exercise in using the rule (3.2.4), let us consider D3-branes embedded in the worldvolume of O7$_\pm$. Since #ND $= 4$, the gauge group on the embedded D3-branes is $\mathfrak{so}$ for O7$_+$ and $\mathfrak{sp}$ for O7$_-$. In particular, the smallest gauge algebra allowed is $\mathfrak{so}_1$ and $\mathfrak{sp}_1$, with one and two Chan-Paton indices, respectively. A bulk D3-brane has two Chan-Paton indices. Therefore, a bulk D3-brane can fractionate into two separate objects on O7$_+$ but not on O7$_-$. These D3-branes can be considered as point-like instantons of the gauge fields on O7$_\pm$, and therefore the D3-charges of the minimal-charge instanton on O7$_\pm$ differ by a factor of 2. This fact becomes important in the anomaly analysis in Sec. 3.3.1.

3.2.3 Intersections

As mentioned in the introduction, O7$_+$s are the only frozen F-theory singularities [25]. As our main interest lies in the compactification to 6d, we now want to understand their behavior when they intersect other singularities: namely, how they modify the gauge algebras of neighboring divisors, and the matter representations at intersections with them. We will do so by using perturbative string techniques, and dualities.

Some readers might want to study the simpler situation in 8d summarized in Appendix B.1, before considering the more interesting but complicated examples of 6d compactifications discussed here.

$\widehat{I}^*$–I intersection

Let us now start working out what happens when the frozen divisors intersect ordinary divisors. We will begin with the intersections of frozen $\widehat{I}_n^*$ with I_m divisors.

Let us first recall what this intersection gives in the unfrozen case, i.e. an I^*–I intersection. The intersection with the I^* induces on the I a so-called "Tate" monodromy, a nontrivial automorphism of the gauge algebra that reduces it [95].[8] This is expressed by saying that the divisor is *non-split*, and denoted by a superscript $^{\rm ns}$. Its effect on the gauge algebra is that it reduces from $\mathfrak{u}_{2m}$ to $\mathfrak{sp}_m$. We summarize this situation by writing

$$\begin{array}{cc} \mathfrak{so}_{2n+8} & \mathfrak{sp}_m \\ I_n^* & I_{2m}^{\rm ns}\,. \end{array} \tag{3.2.7}$$

As a warm-up, let us also see how it is reproduced by orientifolds. Consider an intersection of an O7$_-$ + $(n+4)$ D7 along directions 01256789 with m full D7s along directions 03456789. From (3.2.4) we see again that the gauge algebra on the m D7s is reduced to $\mathfrak{sp}_m$; see also footnote 7. We thus recover (3.2.7). Notice that the spacetime action of the orientifold projection can be interpreted as the Tate monodromy we mentioned above.

We can similarly work out what happens if the I^* divisor is replaced by its frozen $\widehat{I}^*$ counterpart: the configuration now involves an O7$_+$ + $(n-4)$ D7s, and $2m$ transverse D7s (see Fig. 3.2.2, where only directions 6789 are depicted). Looking again at (3.2.4), we see that the gauge algebra on the m D7s is reduced this time to $\mathfrak{so}_{2m}$. We conclude

$$\begin{array}{cc} \mathfrak{sp}_{n-4} & \mathfrak{so}_{2m} \\ \widehat{I}_n^* & I_{2m}^{\rm ns}\,. \end{array} \tag{3.2.8}$$

Thus, an $I^{\rm ns}$ divisor intersecting a frozen divisor has an $\mathfrak{so}$ gauge algebra, rather than a $\mathfrak{sp}$ gauge algebra. In both cases (3.2.7) and (3.2.8) there is a bifundamental at the intersection, due to the strings from one set of branes to the other.

I^*–I^*, I^*–$\widehat{I}^*$, $\widehat{I}^*$–$\widehat{I}^*$ intersections

We will now consider intersections between two I^* divisors, both frozen and unfrozen. We will see that using perturbative O7s we will have only partial success in understanding the full possibilities. This will lead us in later section to consider T-dualities with O6s and O8s.

I^*–I^* intersection: Let us again start by recalling what F-theory gives in the ordinary, unfrozen case. The intersection of two I^* divisors actually falls outside Kodaira's classification. To cure this, one can blow-up the base; this reveals a new divisor that touches both I^*'s, and that behaves

[8] This is not to be confused with the "Kodaira" monodromy, describing how the geometry changes when one goes around a singular divisor

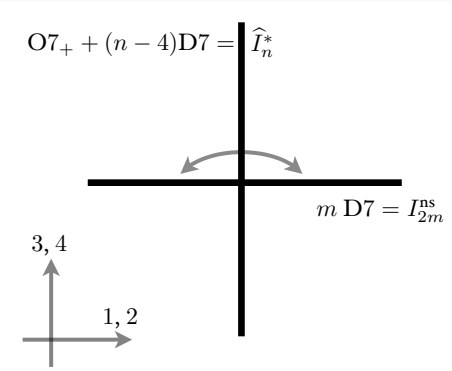


Figure 3.2.2: An O7–D7 intersection, interpreted in F-theory as an intersection between an $\widehat{I}^*_{n+4}$ and an $I^{\rm ns}_{2m}$.

like in (3.2.7):

$$\begin{matrix}\mathfrak{so}_{2k+8}\\ I^*_k\end{matrix} \quad \bullet \quad \begin{matrix}\mathfrak{so}_{2\ell+8}\\ I^*_\ell\end{matrix} \qquad \leftarrow \qquad \begin{matrix}\mathfrak{so}_{2k+8} & \mathfrak{sp}_{(k+\ell)/2} & \mathfrak{so}_{2\ell+8}\\ I^*_k & I^{\rm ns}_{k+\ell} & I^*_\ell\end{matrix} \tag{3.2.9}$$

where we assumed $k+\ell$ to be even, and the $\bullet$ denotes the bad singularity that we blew up. Physically, it signals a six-dimensional superconformal sector which is sometimes called $\mathbb{D}_{k+4}$–$\mathbb{D}_{\ell+4}$ *conformal matter*;[9] the blow-up represents moving along its tensor branch, namely the part of its moduli space where all six-dimensional tensor multiplets are abelian.

Let us now try to engineer an I^*–I^* intersection using O7s. The most natural generalization of Fig. 3.2.2 consists of two O7s that intersect transversally. This can be achieved by an orientifold projection in flat space that has more than one generator of the type we recalled in (3.2.1). For an intersection of two O7s, locally one takes the two generators

$$\Omega R_6 R_7 (-1)^{F_L}\,, \qquad \Omega R_8 R_9 (-1)^{F_L}\,. \tag{3.2.10}$$

We can see that in this situation there is an O7 on the locus $x^6 = x^7 = 0$, and one on the locus $x^8 = x^9 = 0$. (Notice that one is then also quotienting by their product $R_6 R_7 R_8 R_9$, so that at the intersection between the O7s there is in fact also a $\mathbb{Z}_2$ orbifold singularity.) Choosing the $\pm$ type of these two orientifold planes affects their charge and their action on Chan–Paton indices in the way we reviewed earlier; we will see shortly what their combined effect amounts to.

Another ingredient is that the projection on the closed $\mathbb{Z}_2$-twisted sector is reversed if two orientifolds of different type intersect [96]. This comes about by considering the exchange of closed strings between two crosscaps, one from one O7 and another from another O7. The

[9] In fact this superconformal theory depends only on $k+\ell$ and has $\mathfrak{so}(2k+2\ell+16)$ flavor symmetry. Thus we will simply call it $\mathbb{D}_{k+\ell+8}$ conformal matter in what follows. We use the blackboard letter $\mathbb{D}$ since the notation D_i denotes an i-th divisor in this chapter. One can also define $\mathbb{D}_{2n}$ as the 6d superconformal theory which has a one-dimensional tensor branch on which it becomes an $\mathfrak{sp}_{n-4}$ theory with $4n$ fundamentals with at least $\mathfrak{so}_{4n}$ flavor symmetry. For example, then, the $\mathbb{D}_8$ theory is the E-string theory.

sign of this diagram is reversed when two orientifolds are of different type, and the modular transformation of this diagram determines the orientifolding projection on the closed string $\mathbb{Z}_2$ twisted sector. In the end, one finds that an O7$_-$–O7$_+$ intersection has a six-dimensional tensor multiplet, while O7$_-$–O7$_-$ or O7$_+$–O7$_+$ intersection has a hypermultiplet:

	O7$_-$	O7$_+$
O7$_-$	hyper	tensor
O7$_+$	tensor	hyper

. (3.2.11)

As we mentioned, if D-branes are present, they will now feel the effect of both projections. Consider for example choosing both planes to be O7$_-$, with $k+4$ and $\ell+4$ D7s present on the $x^6 = x^7 = 0$ and $x^8 = x^9 = 0$ loci respectively. The first set of D7s, say, would be projected to $\mathfrak{so}_{2k+8}$ by the O7$_-$ parallel to it; but, recalling (3.2.4), it would also be projected to $\mathfrak{sp}_{k+4}$ by the O7$_-$ transverse to it. This means that it actually gets projected to the intersection of the two, $\mathfrak{u}_{k+4}$. In the language of F-theory branes, this gives

$$\begin{matrix} \mathfrak{u}_{k+4} & & \mathfrak{u}_{\ell+4} \\ I_k^* & \cdot & I_\ell^* \end{matrix} , \tag{3.2.12}$$

where the $\cdot$ now represents the hypermultiplet found in (3.2.11).[10] This hypermultiplet is uncharged under $\mathfrak{u}_{k+4} \oplus \mathfrak{u}_{\ell+4}$. The presence of this neutral hypermultiplet signals that the configuration (3.2.12) is obtained by moving along a particular direction in the *Higgs* branch of $\mathbb{D}_{k+4}$–$\mathbb{D}_{\ell+4}$ conformal matter whose tensor branch was depicted in (3.2.9). This particular direction in the Higgs branch is parametrized by vevs of the neutral hypermultiplet in (3.2.12). Another well-known direction in the Higgs branch, distinct from the one represented by (3.2.12), is provided by brane recombination, where the two I^* divisors merge.

$\widehat{I}^*$–$\widehat{I}^*$ intersection: For an O7$_+$–O7$_+$ projection, for the same reason we get

$$\begin{matrix} \mathfrak{u}_{k-4} & & \mathfrak{u}_{\ell-4} \\ \widehat{I}_k^* & \cdot & \widehat{I}_\ell^* \end{matrix} . \tag{3.2.13}$$

In analogy with our discussion below (3.2.12), it is natural to think that this is the Higgsing of a "frozen conformal matter"

$$\begin{matrix} \mathfrak{sp}_{k-4} & & \mathfrak{sp}_{\ell-4} \\ \widehat{I}_k^* & \bullet & \widehat{I}_\ell^* \end{matrix} , \tag{3.2.14}$$

and that upon blowing up (moving along the tensor branch) an $I^{\rm ns}_{k+\ell}$ with $\mathfrak{so}(k+\ell)$ gauge algebra would be created, which would behave as in (3.2.8). We will see later that this expectation is borne out.

[10] A warning is in order. The orientifold projection leaves the gauge algebra $\mathfrak{u}$ on I^*, but the $\mathfrak{u}_1$ part usually gets Higgsed and becomes massive by the Green-Schwarz mechanism, each $\mathfrak{u}_1$ eating a neutral hypermultiplet. This point was carefully analyzed in [97, Sec. 2]. When we compare this with explicit F-theory models, we can only expect to see $\mathfrak{su}_{n+4}$ and $\mathfrak{su}_{m+4}$ symmetry.

$\widehat{I}^*$–I^* intersection: For an O7$_+$–O7$_-$ intersection, on each set of D7s the two projections will be of the same type. For example, on the D7s on the O7$_-$, we have $\lambda = -M_1\lambda^t M_1^{-1} = -M_2\lambda M_2^{-1}$, with both M_i symmetric. We can make $M_1 = 1$ as in section 3.2.1; with the residual freedom in change of basis we can diagonalize M_2, but a priori it could have any number of positive and negative eigenvalues. If we also impose that the D7s can move off the O7$_-$, we obtain that $M_2 = \left(\begin{smallmatrix} 1_{\ell+4} & 0 \\ 0 & -1_{\ell+4} \end{smallmatrix}\right)$, and the gauge symmetry is $\mathfrak{so}_{\ell+4} \oplus \mathfrak{so}_{\ell+4}$. Similar considerations apply to the O7$_+$ + $(k-4)$D7s; hence we get

$$\begin{array}{ccc} \mathfrak{sp}_{k/2-2} \oplus \mathfrak{sp}_{k/2-2} & & \mathfrak{so}_{\ell+4} \oplus \mathfrak{so}_{\ell+4} \\ \widehat{I}^*_k & \circ & I^*_\ell \end{array} \tag{3.2.15}$$

where we assumed k to be even. Notice that in this case there is no neutral hypermultiplet at the origin, according to (3.2.11); we have included the symbol $\circ$ to mark this. So in this case we do not expect this configuration to be a Higgsing of a conformal one. This might look surprising, but it will become clearer in section 3.2.7 below, where we will see an alternative realization of the same setup (in the case $k = \ell$ is even).

3.2.4 NS5- and D6-branes

To go beyond the results in section 3.2.3, we will need to consider configurations which are dual to IIA in presence of NS5-branes. To set the stage, in this subsection we will discuss a situation without orientifolds.

We consider IIA on $\mathbb{R}^9 \times S^1$; let us say the S^1 corresponds to direction 4, and has periodicity R. Let us have a single NS5 whose worldvolume is in directions 056789, localized at $x^\alpha = x^4 = 0$, $\alpha = 1, 2, 3$. T-dualizing it along direction 4 turns it into an Euclidean Taub–NUT geometry. The space transverse to the NS5 is $\mathbb{R}^3 \times S^1$; T-duality turns the H flux of the NS5 into a Chern class that signals the S^1 is now Hopf-fibred over the S^2s at $x^\alpha x^\alpha = r^2$. The inverse images of these S^2s are thus copies of S^3. These shrink smoothly at $x^\alpha = 0$, so that locally around this point the fibration is $S^1 \hookrightarrow \mathbb{R}^3 \to \mathbb{R}^4$. One way to realize this fibration in coordinates is

$$\mathbb{H} \cong \mathbb{C}^2 \to \mathbb{R}^3 \tag{3.2.16}$$

$$q = \begin{pmatrix} z \\ w \end{pmatrix} \mapsto x^\alpha = q^\dagger \sigma^\alpha q \tag{3.2.17}$$

where σ^α are the Pauli matrices. So

$$x^1 + ix^2 = zw\,, \qquad x^3 = |z|^2 - |w|^2\,. \tag{3.2.18}$$

If we have several NS5s localized at several positions in the 3 direction ($x^3 = x^3_i$, $x^1 = x^2 = x^4 = 0$), T-duality turns the geometry into a multi-Taub–NUT geometry where the S^1 shrinks at the $x^3 = x^3_i$. The inverse image under the S^1 fibration of a path between two of these points is an S^2. We represent this in Fig. 3.2.3.

Let us now suppose some D6s are also stretched along the 0356789 directions. First let us imagine there are n D6 stretched along the entire 3 axis. Under T-duality along direction 4, they

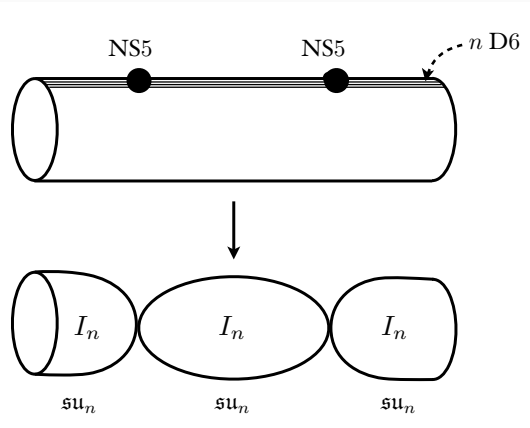


Figure 3.2.3: NS5-branes, D6-branes, and T-duality. The compact and noncompact directions of the cylinder are called respectively directions 4 and 3 in the text.

will turn into n D7s. More precisely, as Fig. 3.2.3 suggests, they will turn into a sequence of D7s wrapping the various S^2 on the Taub–NUT with multiplicity n. What the picture does not show is that these S^2s are holomorphic cycles. Locally around an NS5 at $x^\alpha = 0$, for example, we see from (3.2.18) that the locus $x^3 = 0$ is turned into $zw = 0$, which is the union of the curve $z = 0$ and of $w = 0$. In F-theory terms, this is a chain of intersecting I_n curves.

In the presence of a Romans mass, parameterized conventionally by an integer $2\pi F_0 \equiv n_0 \neq 0$, one can also have situations where D6s are ending on NS5s. Focusing on an NS5 on which a D6 ends from the right, we see again from (3.2.18) that T-duality turns it into the single curve $z = 0$. This would be one of the S^2s in Fig. 3.2.3. The number of D6s ending on an NS5 from the left minus the number of D6s from the right is proportional to F_0. We then have a chain of intersecting curves supporting I_n, I_{n+n_0}, I_{n+2n_0}, ...

Another possible generalization is to move the D6s in the x^4 direction, so that there is now a stack of n_j D6s at $x^4 = x^4_j$. On the IIB side, this corresponds to Wilson lines for the gauge field on the D7s.

3.2.5 Shared gauge algebras

From the setup of Fig. 3.2.3, we can also wonder what happens if we move only some of the D6s away from the NS5s in direction 4; say from an initial stack of n D6s we move m to the position $x^4 = x^4_0$. These D6s recombine: they no longer end on the NS5s. In field theory, this corresponds to a partial Higgsing

$$\mathfrak{su}_n \oplus \mathfrak{su}_n \;\to\; \mathfrak{su}_{n-m} \oplus \mathfrak{su}_m \oplus \mathfrak{su}_{n-m} \tag{3.2.19}$$

where the $\mathfrak{su}_m$ at the middle is the diagonal subalgebra of two copies of $\mathfrak{su}_m \subset \mathfrak{su}_n$.

Since the displacement has happened along the 4 direction, it is not immediately apparent on the IIB side: the T-dual still consists of two stacks of $m + n$ D7-branes meeting at a point, as in

section 3.2.4. The only consequence of the displacement is the presence of a Wilson line: there is a worldvolume gauge field with non-zero holonomy, $a = \frac{x_0^4}{l_s^2}\mathrm{diag}(0,\ldots,0,1,\ldots,1)d\tilde{x}^4$. Since direction $\tilde{4}$ shrinks at the intersection point, on both D7s there is a worldvolume $da = f$ field strength proportional to a δ-function supported on the intersection point.

By comparing with the IIA picture, we conclude that a Wilson line can partially break the gauge algebra on two intersecting D7s, as in (3.2.19): part of the gauge algebra can *recombine*. The $\mathfrak{su}_m$ algebra is now shared between the two intersecting divisors; this is summarized in Fig. 3.2.4. In what follows, we fill find other examples of such shared gauge algebras.

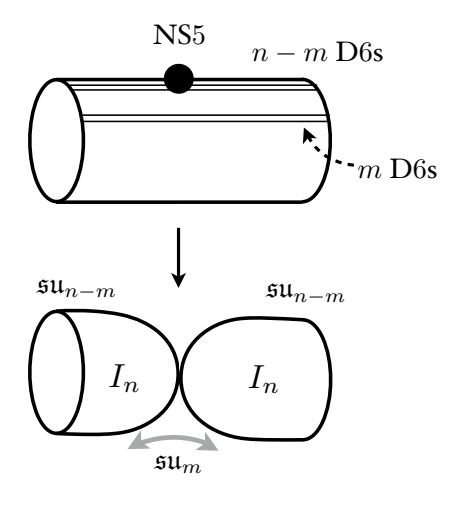


Figure 3.2.4: On the IIA side, we can move m of the D6s off the NS5s and make them recombine. On the IIB side, this corresponds to a gauge algebra $\mathfrak{su}_m$ that is shared between two curves meeting at a point. We denote this with a double-sided arrow.

If we move all the D6s off the NS5 (i.e. if $n = m$), only the shared gauge algebra is present. In this case, one might be puzzled by the fact that on the IIB side the Wilson line is now proportional to the identity. This would not seem to cause a Higgsing, while from the IIA picture it is clear that it does, since the D6s are away from the NS5.

To clarify this point, we need to identify the T-dual of the NS5 position in IIB. Since the NS5 position in IIA is shifted by a diffeomorphism in the x^4 position, its T-dual should be shifted in IIB by a gauge transformation for the NS=NS two-form field, namely $B \to B + d\Lambda$, for Λ a one-form. In fact this one-form was identified in [98, Sec. 2.2] explicitly. Thus more generally we conclude that, in the intersection between two curves $\mathcal{C}_1, \mathcal{C}_2$, there is a shared gauge algebra if on either curve there is an eigenvalue a_i of the Wilson line α on the curves that does *not* match with the pullback of Λ at large distance from the intersection:

$$a_i \neq \Lambda|_{\mathcal{C}_1} \,. \tag{3.2.20}$$

In F-theory language, we could consider a deformation of the Weierstrass coefficients which "recombined" two branes, i.e., smoothed the two divisors out into a single divisor. If instead

of this deformation, the corresponding periods of two-forms are activated, the gauge theory will recombine without any change in the geometry.

3.2.6 NS5-branes and O-planes

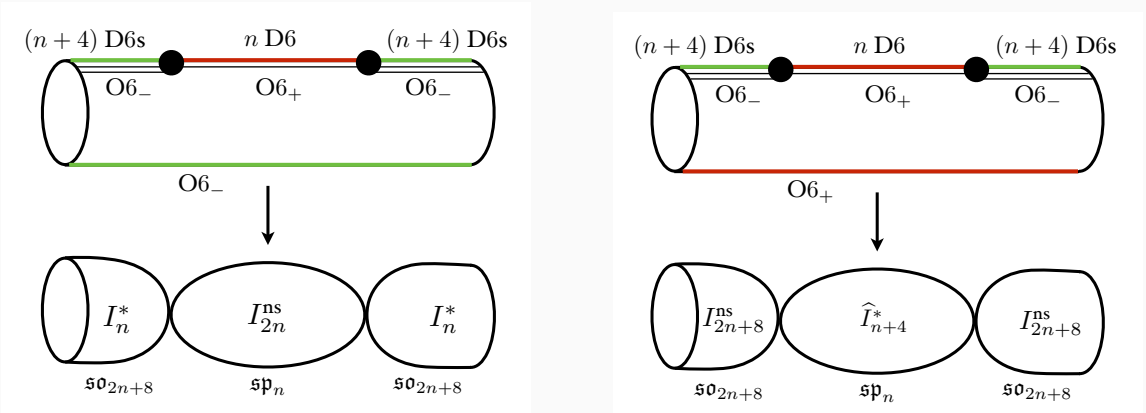
Having made a detour in the last two subsections, we now reintroduce O-planes in our story.

First we need to review the behavior of NS5s in presence of orientifolds. Like any other brane, any NS5 must come with a mirror image under the orientifold action. Each copy is usually called a *half-brane* to emphasize that it can become *full* if the two copies are brought to the O-plane. It turns out [99] that when this is done the two half-NS5s can be separated again: this time along the O-plane worldvolume, while staying on it. When this happens, the orientifold type changes between the two half-NS5s.

The situation relevant for our purposes consists in having an O6 defined by a reflection inverting directions 124, and for example two half-NS5s at two values of x^3. (Thus the O6-plane and the half-NS5s are stretched along the same directions as the D6 and NS5 in the previous subsection.) If the O6 is taken to be an O6$_-$ outside the two half-NS5s, its type will change to O6$_+$ inside. This leads to a sequence of gauge algebras

$$\mathfrak{so}_{2n+8}\ ,\ \mathfrak{sp}_n\ ,\ \mathfrak{so}_{2n+8}\ . \tag{3.2.21}$$

Actually, since direction 4 is compact, a reflection involving 124 will have a fixed point both at $x^4 = 0$ and at $x^4 = R/2$, the opposite locus on the circle. The O6-plane on that locus can be of both O6$_-$ and O6$_+$ type. We show both those cases in Fig. 3.2.5. In both cases the gauge algebras are still as in (3.2.21), since the difference with the case of Fig. 3.2.5(a) happens in a region where no D6s are present.



(a) (b)

Figure 3.2.5: Two configurations with O7$_\pm$-planes, and their T-duals. The dots now represent half-NS5s

Upon T-duality, we again find a chain of curves. To see what type of curves we have, we need to use the rules reviewed in section 3.2.1; see in particular Fig. 3.2.1. We learn from there that an orientifold with O6$_\pm$-planes on both sides of a circle gets T-dualized to an orientifold with an O7$_\pm$-plane, while a circle which has an O6$_+$ on one side and an O6$_-$ on the other gets T-dualized to a shift-orientifold. This is another realization of Tate monodromy, which we discussed at the beginning of section 3.2.3.

Thus, in the case of Fig. 3.2.5(a), after T-duality we end up with a curve $I_{2n}^{\rm ns}$ between two ordinary I_n^* curves. This is familiar from (3.2.9) with $m = n$, and is in agreement with the sequence of gauge algebras (3.2.21) we found in IIA.

In the case of Fig. 3.2.5(b), we have a frozen $\widehat{I}_{n+4}^*$ curve touching two $I_{2n+8}^{\rm ns}$ ones. The presence of the frozen singularity alters the usual F-theory rules: from the IIA picture, we see that as expected an $\widehat{I}_{n+4}^*$ curve supports an $\mathfrak{sp}_n$ gauge algebra; moreover, we also see that an $I_{2m}^{\rm ns}$ touching a frozen curve supports an $\mathfrak{so}_{2m}$. This can be generalized to

$$\begin{array}{ccc} \mathfrak{sp}_{k-4} & \mathfrak{so}_{k+\ell} & \mathfrak{sp}_{\ell-4} \\ \widehat{I}_k^* & I_{k+\ell}^{\rm ns} & \widehat{I}_\ell^* \end{array} \tag{3.2.22}$$

(with $k = n+4$). This is the theory on the tensor branch of (3.2.14), thus realizing the expectation discussed there.

If we put the half-NS5s back on top of each other, we recover a full NS5. We can now split it again by moving the two halves along the periodic 4 direction, together with some of the D6s, or by moving them in another direction, so that the degeneration induced by T-dualizing the NS5s no longer happens on the O6–D6 system. These two new configurations represent respectively the Higgsing in (3.2.12), and the one mentioned below it involving brane recombination. These two possibilities are depicted in Fig. 3.2.6.

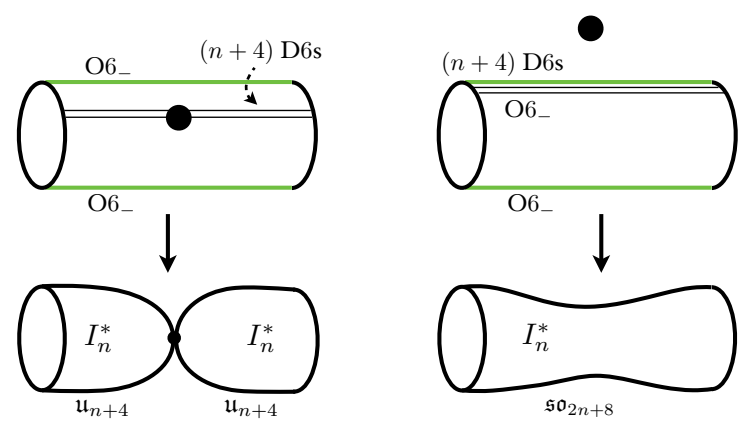


Figure 3.2.6: Two different ways of Higgsing $\mathbb{D}_{m+4}$–$\mathbb{D}_{n+4}$ conformal matter. The first reproduces (3.2.12); the second corresponds to brane recombination.

The setup of this section can also be decorated by adding m D6-branes at the bottom orientifold plane; this would add a gauge algebra $\mathfrak{so}_{2m}$ to Fig. 3.2.5(a), and $\mathfrak{sp}_m$ to Fig. 3.2.5(b).

On the F-theory side, this would correspond to the presence of a Wilson line, and to a gauge algebra that is shared among the three curves, in the language of section 3.2.5. Again, this can be realized through the T-brane-like phenomena of activating the two-form-period partner of a geometric deformation.

3.2.7 Smooth transitions

In the chains of curves considered so far, shrinking one or more of the curves leads to some strongly coupled physics. This is clear from the IIA picture, where it corresponds to making two or more NS5-branes coincide. In an effective field theory description, this often manifests itself in a gauge coupling becoming infinite. The positions of the NS5s parameterize a so-called tensor branch of a six-dimensional effective theory; these situations correspond to non-generic loci of the tensor branch.

For example, in the situations depicted in figures 3.2.3 and 3.2.5, there is a one-dimensional tensor branch, parameterized by a 6d tensor multiplet whose scalar ϕ corresponds to the distance between the two NS5s, and which in the 6d theory also plays the role of an inverse squared gauge coupling. At the origin $\{\phi = 0\}$, the gauge coupling diverges. At this strong coupling point it is expected that a CFT arises, describing two coincident NS5s on top of a D6 stack.

However, on the IIA side we can also consider placing the NS5s at different values of x^9 (the compact direction). In this case, bringing the NS5s at the same value of x^9 does not actually put them on top of each other; thus now we do not expect strong coupling physics at the origin of the tensor branch.

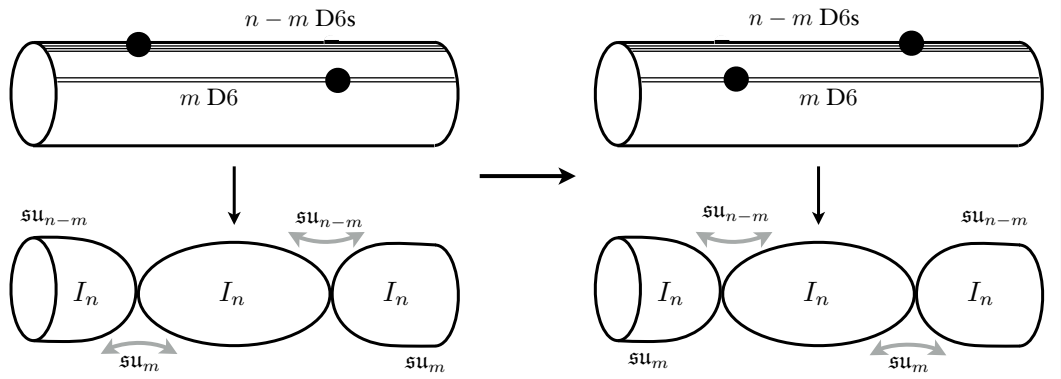


Figure 3.2.7: A smooth transition, in IIA and in F-theory.

A first example not involving orientifolds is shown in Fig. 3.2.7. In this case without frozen seven-branes as above, we can of course put all NS5-branes on the same stack of D6-branes so that this smooth transition does not happen.

When we start involving orientifolds, we can engineer more interesting situations. The example in Fig. 3.2.8 has a non-split I_{2n}^{ns} touching both a frozen and a non-frozen I^*. In this case we do not have any way to put all NS5-branes on the same side of the O6-planes. Note also that

in both sides of the figure the overall gauge algebra remains the same, but the roles of localized and shared simple subalgebras are exchanged.

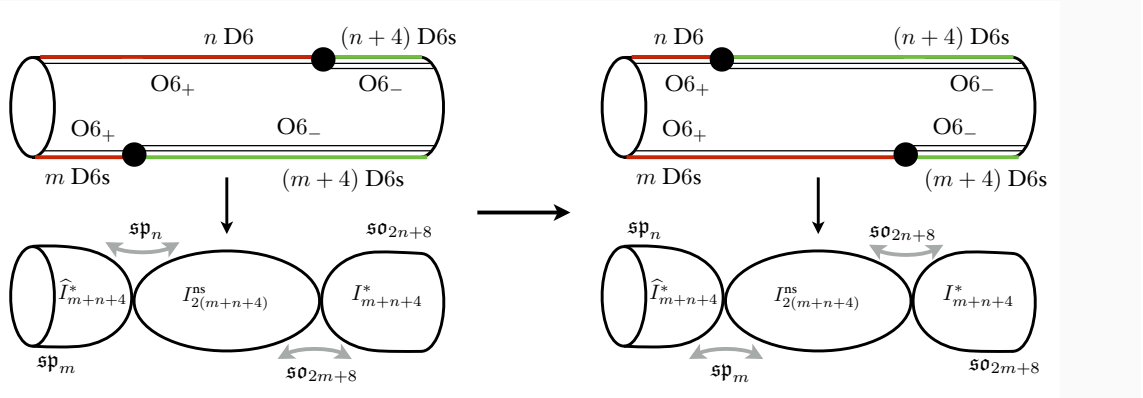


Figure 3.2.8: A configuration that produces a curve touching both an I^* and and an $\widehat{I}^*$. The gauge algebras $\mathfrak{sp}_n$ and $\mathfrak{so}_{2m+8}$ are shared between the first two and the second two curves respectively.

When the two NS5s are aligned, for $m = n$ we are in fact in the situation of (3.2.15), with $k = \ell = 2n+4$. This is in agreement with our observation made there (motivated by the absence of a hypermultiplet) that there is no conformal point at that intersection; in this case the transition is completely smooth, and there is no special point on the tensor branch.

In F-theory we are accustomed to getting conformal theories at the origin of the tensor branch; one reason for this is that one can engineer string states from D3-branes, and these strings become tensionless when we shrink a curve. In the situations of Fig. 3.2.7 and 3.2.8, in fact we cannot wrap a D3-brane on the middle curve: this is made clear by T-dualizing back to IIA, where it would become a D2-brane, which can terminate on either one or the other half-NS5, but not on both.

The situation in Fig. 3.2.8 is a simple illustration of the fact mentioned in the introduction that in the presence of O7$_+$ we lose the notion of a canonical assignment of gauge algebras and matter content. In this situation, this happens for two reasons. First, we can only take m D6-branes from bottom to top of the cylinder. After doing that, we are still left with 4 D6-branes ending on half-NS5-brane. This implies that there is no canonical 'zero' for the Wilson lines. Second, the half-NS5s are stuck at fixed values of x^4. This implies that there are fixed non-zero periods of NS-NS 2-form potential on the curves.

3.2.8 Tangential intersections and O8-planes

The discussion of I^*–I and $\widehat{I}^*$–I intersections in section 3.2.3 has an interesting exception, that occurs when the intersection is *tangential*. We discuss it now because T-duality helps in the analysis, as we will now see.

We start by considering O7s and D7s that again share the directions 056789, but which are extended in the remaining directions in a more complicated fashion than in section 3.2.3. Define

$z = x^1 + ix^2$, $w = x^3 + ix^4$, and let the orientifold act on the spacetime by $\sigma : z \leftrightarrow w$. The O7$_\mp$ will then be on the locus $z = w$; place again $n \pm 4$ D7s on top of it. Now also place m half-D7s on the locus $z = 0$; their m images will be on the locus $w = 0$. In this case, the gauge fields on the D7s on $z = 0$ will have a $\mathrm{U}(m)$ gauge field, which the O7 maps to a gauge field on the D7s on $w = 0$. To see why this is related to a tangential intersection, consider the invariant coordinates $v = z + w$, $u = zw$. The configuration we are considering is then mapped to an O7$_\mp$ + $(n \pm 4)$ D7s on the locus $v^2 = 4u$, and m D7s on $u = 0$. These two loci intersect tangentially. We can summarize this as follows:

$$\begin{array}{ccc} \mathfrak{so}_{2n+8} & & \mathfrak{su}_m \\ I_n^* & || & I_m\,; \end{array} \qquad\qquad \begin{array}{ccc} \mathfrak{sp}_{n-4} & & \mathfrak{su}_m \\ \widehat{I}_n^* & || & I_m\,, \end{array} \tag{3.2.23}$$

where we have used $||$ to denote tangency as in [1]. This coordinate change is illustrated in the top part of Fig. 3.2.9, in the O7$_+$ case.

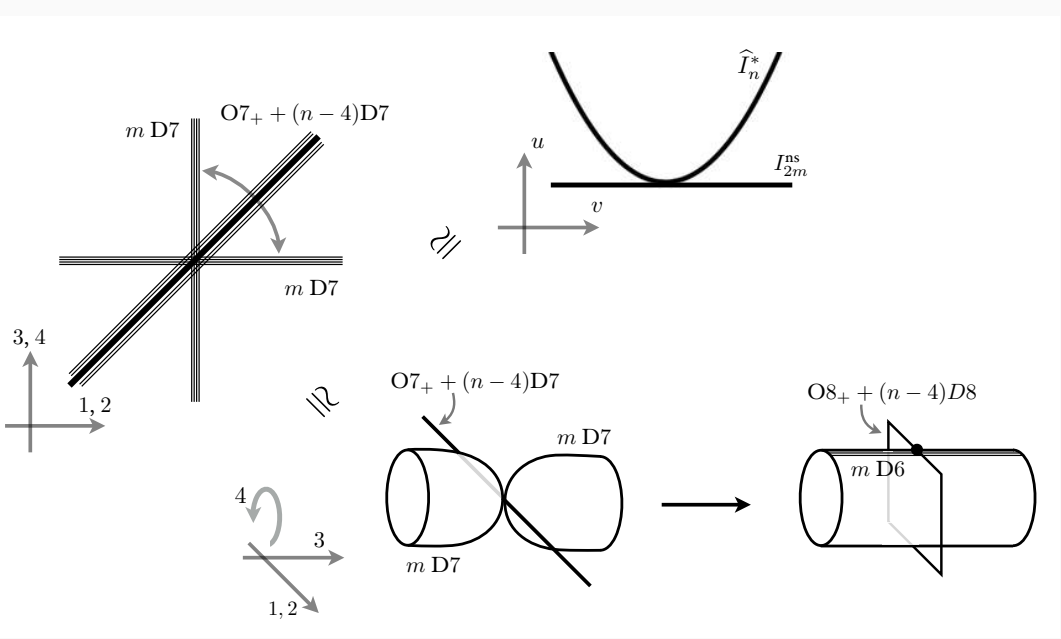


Figure 3.2.9: Various equivalent ways of seeing a tangential $\widehat{I}^*$–I intersection. As in recent figures, the dot on the bottom-right frame is a half-NS5.

An additional subtlety concerns the matter content in (3.2.23). One can in principle work this out directly in the original setup on the left of Fig. 3.2.9, but it is instructive to do it instead in a dual frame. First of all we change coordinates, using again (3.2.18); only this time we take $z = x^1 + ix^2$, $w = x^3 + ix^4$ introduced earlier, and define new coordinates $\tilde{x}^1 + i\tilde{x}^2 = zw$, $\tilde{x}^3 = |z|^2 - |w|^2$, with a fourth periodic coordinate $e^{i\tilde{x}^4} = \frac{z\overline{w}}{\overline{z}w}$. We are once again rewriting $\mathbb{R}^4$ as a fibration $S^1 \hookrightarrow \mathbb{R}^4 \to \mathbb{R}^3$. The orientifold is now defined by the involution $\sigma : \tilde{x}^3 \to -\tilde{x}^3, \tilde{x}^4 \to -\tilde{x}^4$; the O7-plane then sit at $\tilde{x}^3 = \tilde{x}^4 = 0$, while the D7s are on the locus $\tilde{x}^1 = \tilde{x}^2 = 0$. (Notice that the $\tilde{x}^4$ circle shrinks at $\tilde{x}^3 = 0$.) If we now T-dualize along direction 4, we end up with an O8 at $\tilde{x}^3 = 0$ with a half-NS5 stuck on it, and with D6s crossing it. All this is depicted on the

lower part of Fig. 3.2.9, again for the O8$_+$ case. At this point we can read off the matter content from a perturbative string computation similar to the one leading to (3.2.4), as already done in [19, 18]; the result is that in the tangential intersection (3.2.23) the $\mathfrak{u}_m$ has a hypermultiplet in the antisymmetric in the unfrozen case, and in the symmetric in the frozen case.

3.3 Anomaly analysis

In this section we discuss the cancellation of one-loop anomalies and the Green–Schwarz contributions in 6d compactifications with frozen seven-branes.

3.3.1 Anomaly cancellation with frozen singularities

A compactification of F-theory on an elliptically fibered Calabi–Yau threefold gives rise to an effective 6d gauge theory with $\mathcal{N} = (1, 0)$ supersymmetry at low energies. When there are no frozen singularities present, it is possible to turn off the holonomies of gauge fields on stacks of seven-branes, and the periods of 2-form NS-NS and R-R potentials. Then, each simple summand $\mathfrak{g}_i$ of the 6d gauge algebra is associated to a single irreducible component D_i of the discriminant locus of the elliptic fibration, and can be determined from the knowledge of the type of singular fiber over D_i along with the data of the monodromy of the elliptic fiber around D_i [72, 95, 65]. The matter content [100, 51] and the coupling of tensor multiplets [52] is encoded in the intersection numbers of various divisors in the base of the elliptic fibration. These data allow us to compute both the 1-loop contribution $I^8_{\text{1-loop}}$ to the anomaly polynomial, as well as the Green–Schwarz contribution I^8_{GS} to the anomaly polynomial. Combining these two, one finds that $I^8 = I^8_{\text{1-loop}} + I^8_{GS}$ vanishes for any smooth elliptically fibered Calabi–Yau threefold [53, 51].

Now let us include frozen singularities in the geometry. In this situation, it is not always possible to tune the above mentioned holonomies to zero. We do not have any canonical nonzero choice either. Because of the nonzero holonomies, one is forced to consider situations in which simple summands of the 6d gauge algebra are realized on divisors which are positive linear combinations of irreducible components of discriminant locus. We will call the divisors associated to simple summands of gauge algebra as *gauge divisors*.

We will not be able to list down all the possible 6d spectra that could result from a geometry, as that will require a systematic understanding of holonomies and fluxes in F-theory compactifications, which we do not have at present. Therefore, we suppose that an assignment of gauge algebras on the components of the discriminant is given, and study the Green–Schwarz contribution to the anomaly. We follow the work of Sadov [52] but we include the effects from the frozen singularities.

The 6d tensor multiplets descend from Kaluza-Klein reduction of the chiral 4-form $C^{(4)}$ of type IIB string theory. To determine the coupling of 6d tensor multiplets, we need to look at two couplings of $C^{(4)}$ in ten-dimensional type IIB string theory. One of them is a coupling to gravity and the other is a coupling to the gauge theory living on seven-branes.

Gauge Green–Schwarz terms: We start with the coupling of the gauge fields to the RR 4-form field $C^{(4)}$. When there are no O7$_+$-planes, the stack of seven-branes on D_a has a ten-dimensional coupling given by

$$\int C^{(4)}\nu(F_a)D_a \tag{3.3.1}$$

where F_a is the field strength valued in the "Kodaira" 8d gauge algebra $\mathfrak{k}_a$ on the D_a component of the discriminant, and $\nu(F_a)$ is the instanton number density, normalized so that it integrates to one on the standard BPST instanton embedded into $\mathfrak{k}_a$ with embedding index 1. This normalization reflects the familiar fact that an instanton in the worldvolume of a seven-brane has D3-charge 1.

When the component D_a carries a $\widehat{I}^*_{n+4}$ singularity, i.e. when it corresponds to an O7$_+$-plane with n D7-branes on top, the local 8d gauge algebra is $\mathfrak{k}_a = \mathfrak{sp}_n$, and the ten-dimensional coupling is

$$\int C^{(4)}\left(\frac{1}{2}\nu(F_a)\right)D_a. \tag{3.3.2}$$

Note a factor-of-two difference in the coefficient between (3.3.1) and (3.3.2). This is due to the fact that a bulk D3-brane can fractionate into two on O7$_+$, as reviewed in Sec. 3.2.2, and the gauge instanton in the standard normalization corresponds to the D3-brane of minimal possible charge.

Let us now write the 6d gauge algebra in the form $\oplus_i\mathfrak{g}_i$ where $\mathfrak{g}_i$ is simple. Each $\mathfrak{g}_i$ is shared on some of the D_a; we let $\mu_{i,a} = 1$ or 0 depending on whether $\mathfrak{g}_i$ is on D_a or not. An embedding $\rho_{i,a} : \mathfrak{g}_i \hookrightarrow \mathfrak{k}_a$ must exist whenever $\mu_{i,a} = 1$, and otherwise we let $\rho_{i,a}$ be the zero map. These embeddings have the properties

1. $\bigoplus_i \rho_{i,a}(\mathfrak{g}_i) \subset \mathfrak{k}_a$, and
2. $\mathfrak{g}_i$ is the diagonal in $\bigoplus_a \rho_{i,a}(\mathfrak{g}_i)$.

The Green–Schwarz coupling for the gauge fields is given in terms of the field strenghs F_i valued in $\mathfrak{g}_i$ by

$$\int C^{(4)}\sum_i \nu(F_i)\Sigma_i := \sum_{i,a}\int C^{(4)}\left(\sum_i \mu_{i,a}o_{i,a}\nu(F_i)\right)D_a \tag{3.3.3}$$

where we defined the i-th *gauge divisor* to be

$$\Sigma_i = \sum_a \mu_{i,a}o_{i,a}D_a, \tag{3.3.4}$$

and $o_{i,a}$ is the embedding index of $\mathfrak{g}_i \subset \mathfrak{k}_a$, multiplied by 1/2 when $\mathfrak{k}_a = \mathfrak{sp}_n$ is supported on a frozen singularity.

Note that even when there is no "sharing" (so the gauge divisors are $\Sigma_a = D_a$) and no O7$_+$-planes, $\mathfrak{g}_a$ could still be different from $\mathfrak{k}_a$, due to the "Tate monodromy" phenomenon [95].

Before proceeding, we point out here that the inverse square of the gauge coupling of $\mathfrak{g}_i$ is given by $\sum_a \mu_{i,a}o_{i,a}A_a$ where A_a is the area of D_a. This follows from the fact that the scalar

A_a and the 2-form $\int_{D_a} C^{(4)}$ are the bosonic components of a single supermultiplet, and therefore Green-Schwarz coupling $\int C^{(4)} \sum_a \mu_{i,a} o_{i,a} \nu(F_i) D_a$ comes with the coupling $\int \sum_a A_a \mu_{i,a} o_{i,a} \operatorname{tr} F_i \wedge *F_i$. This means in particular that when the gauge algebra $\mathfrak{g}_i$ is shared on multiple components, the gauge theory does *not* become singular when a single component D_a involved in the gauge divisor shrinks to zero size.

Gravitational Green–Schwarz terms: We turn our attention to the gravitational coupling. When there are no O7$_+$s, the stack of seven-branes on D_a has a ten-dimensional coupling to gravity given by

$$\int C^{(4)} \left(\frac{N_a}{12} \frac{p_1(T)}{4} \right) D_a \tag{3.3.5}$$

where N_a is the order of vanishing of discriminant Δ on D_a, $p_1(T)$ is the Pontryagin class of the tangent bundle of the worldvolume. We also slightly abuse notation and use D_a within the integral to represent the two-form determined by the divisor.[11] In particular, a D7-brane contributes $N_a = 1$ and an O7$_-$-plane contributes $N_a = 2$.

Now, the contribution of O7$_+$ to this gravitational coupling is opposite to that of O7$_-$; the "effective N_a" is -2. Since an $\widehat{I}_n^*$ singularity corresponds to O7$_+$ + $(n-4)$D7-branes, its "effective N_a" is $-2 + (n-4) = n-6$. In comparison, N_a of I_n^* is $n+6$. Hence, in the presence of O7$_+$ we need a correction term to the coupling, and it can be written as

$$\int C^{(4)} \left(\left(\frac{N_a}{12} - s_a \right) \frac{p_1(T)}{4} \right) D_a \tag{3.3.6}$$

where $s_a = 1$ when the curve D_a carries an O7$_+$ and $s_a = 0$ when it does not.

The cancellation: Combining (3.3.3) and (3.3.6), the full six-dimensional coupling relevant for the Green–Schwarz mechanism is

$$\int_B C^{(4)} \left(-(K+F) \frac{p_1(T)}{4} + \sum_i \Sigma_i \nu(F_i) \right) , \tag{3.3.7}$$

where the integral is performed only over the base B, $C^{(4)}$ has two legs on the base B, we have used the condition for unbroken supersymmetry (the Calabi–Yau condition) to substitute the canonical divisor K in place of $-\frac{1}{12} N_a D_a$, and

$$F = \sum_a s_a D_a \tag{3.3.8}$$

is the *frozen divisor*, signifying the divisor along which we find the frozen singularities.

[11] The couplings (3.3.1) and (3.3.5) follow in the case of N_a D7-branes by starting from the coupling $(\sum_p C^{(p)}) \widehat{A}(T)^{1/2} \operatorname{tr} e^F$ determined in [101] and extracting the necessary parts, using $\widehat{A}(T)|_4 = -p_1(T)/24$ and $\operatorname{tr} e^F|_4 = -\nu(F)$.

The contribution to anomaly polynomial is then a square of the coefficient of $C^{(4)}$, with a factor of $1/2$ in front, to take into account that the RR 4-form field is self-dual:

$$I^8_{GS} = -\frac{1}{2}\left(-(K+F)\frac{p_1(T)}{4} + \sum_i \Sigma_i \nu(F_i)\right)^2 . \tag{3.3.9}$$

It is a standard result (see e.g. [102–104, 97, 52]) that the one-loop anomaly of the 6d system is given by

$$I^8_{\text{1-loop}} = \frac{9-n_T}{32}p_1(T)^2 - \frac{N_i}{4}\nu(F_i)p_1(T) + \frac{M_{ij}}{2}\nu(F_i)\nu(F_j) \tag{3.3.10}$$

where n_T is the number of tensor multiplets, and N_i, M_{ij} are some numerical coefficients, assuming that the coefficient of $\operatorname{tr} R^4$ vanishes, i.e.

$$n_V - n_H - 29n_T + 273 = 0\,. \tag{3.3.11}$$

At the end of this subsection, we comment on how to obtain the numerical values N_i and M_{ij}.

We see that cancellation of gauge and gauge-gravity anomalies requires the following:

$$N_i = (K+F)\cdot\Sigma_i, \qquad\qquad M_{ij} = \Sigma_i\cdot\Sigma_j. \tag{3.3.12}$$

Here, $K\cdot\Sigma_i = K\cdot\sum\mu_{i,a}o_{i,a}D_a$ can be computed from the adjunction formula $2(g_a-1) = (K+D_a)\cdot D_a$, where g_a is the genus of the curve D_a.

If the 6d theory contains dynamical gravity and satisfies (3.3.11), then we obtain the following condition as well, from the vanishing of the coefficient of $(\operatorname{tr} R^2)^2$:

$$9 - n_T = (K+F)^2 \tag{3.3.13}$$

This condition (3.3.13) follows just from geometry, as we now demonstrate. If D_a carries a frozen singularity, then the singular fiber over D_a has Kodaira type $I^*_{n\geq 4}$. For these Kodaira fibers, it is known that $D_a\cdot(-2K-D_a) = 0$ [51]. Moreover, any two distinct components D_a and D_b of F must not intersect each other because two $I^*_{n\geq 4}$ singularities cannot intersect each other(in the absence of conformal matter). From these two facts, it follows that

$$(K+F)^2 = K^2 + \sum_a s_a D_a\cdot(2K+D_a) = K^2, \tag{3.3.14}$$

and the equality $K^2 = 9-n_T$.[12]

By now, the cancellation of the Green-Schwarz anomaly and of the one-loop anomaly in the conventional F-theory compactification without O7$_-$ is well-established. This allows us to read off N_i and M_{ij} for almost all the cases. First, for $i\neq j$, we have $M_{ij}=1$ for a bifundamental of $\mathfrak{su}$-$\mathfrak{su}$ or a half-bifundamental of $\mathfrak{so}$-$\mathfrak{sp}$. To read off N_i and M_{ii} (without suming over i),

[12] Compactification of $C^{(4)}$ on a base B produces $h^{1,1}(B)$ anti-symmetric 2-form potentials. One of them goes into the supergravity multiplet and the remaining $h^{1,1}(B)-1$ go into tensor multiplets; hence $n_T = h^{1,1}(B)-1$. Since B is the base of Calabi–Yau, $h^{1,0}(B) = h^{2,0}(B) = 0$ and it follows from Noether's formula that $K^2 = 10 - h^{1,1}(B) = 9-n_T$.

let us say that the given algebra is $\mathfrak{g}_i$ and the total set of hypermultiplets for $\mathfrak{g}_i$ is ρ. One then looks up the pair of $(\mathfrak{g}_i, \rho)$ e.g. in Eqs. (2.10)–(2.14) of [105], to find a conventional F-theory realization of the 6d gauge theory on a sphere of self-intersection $-n$. Then $M_{ii} = -n$ and $N_i = n - 2$. Essentially the only case not covered by this procedure is when $\mathfrak{g}_i = \mathfrak{su}(n)$, with one sym and $n - 8$ fundamentals. For this, one first Higgses the hypers in sym, to give $\mathfrak{so}(n)$ with $n - 8$ fundamentals. This has a well-known anomaly polynomial, which can be determined in the method just described above. Then one can convert it back the anomaly polynomial of the original $\mathfrak{su}(n)$ theory by using $\nu(\mathfrak{so}(n)) = 2\nu(\mathfrak{su}(n))$.

3.3.2 Matter content with frozen singularities

Transversal intersections: In the situation when there are no frozen singularities and each simple factor of gauge algebra $\mathfrak{g}_a$ is associated to a single irreducible component of discriminant locus D_a, Grassi and Morrison [51] wrote down the matter content charged under $\mathfrak{g}_a$ in terms of intersection numbers of combinations of D_a and K, assuming that every intersection among D_a and D_b are transversal. The geometry underlyng the derivation of those formulas, analyzed in the M-theory dual (and therefore on the Coulomb/tensor branch of the theories), consists of finding the curves in the total space upon which M2-branes can be wrapped, and finding the intersection numbers of those curves with the divisors which represent the Cartan subgroup of the original nonabelian gauge group, since those intersection numbers specify gauge charges. This was carried out in a number of works [95, 100, 106, 65, 54] which [51] relied upon.

Now we would like to understand the matter content in the presence of the frozen singularities. We do not have a geometric derivation for our proposed answer, since the M-theory geometry of frozen singularities is not well understood. However, as we have seen in detail, the effect of the frozen singularity in the anomaly contribution from the Green–Schwarz effect is summarized by the replacement of individual components D_a by the gauge divisor Σ_i, and of the canonical class K by $K' = K + F$. The one-loop contribution should then be able to exactly cancel this contribution. We thus propose that the corect answer for the matter content is to perform the same replacement in the results of [51].

We tabulate the results of this replacement, i.e., of our precise proposal for matter content, in Table 3.3.1. Two comments on the table are in order:

- The number associated to adjoint representation in the table is $n_H^{\mathsf{adj}} - 1$ where n_H^{adj} is the number of hypermultiplets charged in the adjoint representation. The -1 incorporates the contribution to the anomaly of a vector multiplet, which indeed comes with the opposite sign with respect to an adjoint hypermultiplet.

- The entry for $\mathfrak{so}_7$ in our table contains a refinement over [51], in which only the spinor representation was considered. But the coefficient of the spinor representation is negative whenever $(-2K' - \Sigma_i) \cdot \Sigma_i < 0$. In this case, a different representation with the same contribution to the anomaly needs to be used. One finds $\mathsf{sym}^2_{\mathrm{irr}}$ does the job.[13]

[13] We claim that similar modifications are unnecessary for $\mathfrak{so}_{n\geq 8}$. Even though the representation containing $\mathsf{sym}^2_{\mathrm{irr}}$ satisfies anomaly cancellation for any $\mathfrak{so}_{n\geq 7}$, the representation only makes sense for $\mathfrak{so}_7$, as we now demon-

$\mathfrak{g}_a$	ρ	Number of hypers in ρ
$\mathfrak{su}_2$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	fund	$(-8K'-2\Sigma_i)\cdot\Sigma_i$
$\mathfrak{su}_3$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	fund	$(-9K'-3\Sigma_i)\cdot\Sigma_i$
$\mathfrak{su}_n$, $n\geq 4$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	fund	$(-8K'-n\Sigma_i)\cdot\Sigma_i$
	asym^2	$-K'\cdot\Sigma_i$
$\mathfrak{sp}_n$, $n\geq 2$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	fund	$(-8K'-n\Sigma_i)\cdot\Sigma_i$
	$\mathsf{asym}^2_{\text{irr}}$	$\frac{1}{2}(-K'+\Sigma_i)\cdot\Sigma_i$
$\mathfrak{so}_7$, $(-2K'-\Sigma_i)\cdot\Sigma_i\geq 0$,	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	vect	$\frac{1}{2}(-3K'-\Sigma_i)\cdot\Sigma_i$
	spin	$(-4K'-2\Sigma_i)\cdot\Sigma_i$
$\mathfrak{so}_7$, $(-2K'-\Sigma_i)\cdot\Sigma_i\leq 0$,	adj	$\frac{1}{8}(-2K-2F+\Sigma_i)\cdot\Sigma_i$
	vect	$\frac{1}{4}(-16K'-7\Sigma_i)\cdot\Sigma_i$
	$\mathsf{sym}^2_{\text{irr}}$	$\frac{1}{8}(2K'+\Sigma_i)\cdot\Sigma_i$
$\mathfrak{so}_n$, $8\leq n\leq 14$,	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	vect	$\frac{1}{2}((4-n)K'+(6-n)\Sigma_i)\cdot\Sigma_i$
	spin_*	$\frac{1}{\dim(\mathsf{spin}_*)}(-32K'-16\Sigma_i)\cdot\Sigma_i$
$\mathfrak{so}_n$, $n\geq 15$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	vect	$(-4K'-\frac{n}{4}\Sigma_i)\cdot\Sigma_i$
$\mathfrak{e}_6$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	27	$(-3K'-2\Sigma_i)\cdot\Sigma_i$
$\mathfrak{e}_7$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	56	$\frac{1}{2}(-4K'-3\Sigma_i)\cdot\Sigma_i$
$\mathfrak{e}_8$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
$\mathfrak{f}_4$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	26	$\frac{1}{2}(-5K'-3\Sigma_i)\cdot\Sigma_i$
$\mathfrak{g}_2$	adj	$\frac{1}{2}(K'+\Sigma_i)\cdot\Sigma_i$
	7	$(-5K'-2\Sigma_i)\cdot\Sigma_i$

Table 3.3.1: Number of hypermultiplets for each relevant representation of each simple gauge algebra when frozen singularities are present. This includes the contribution of vector multiplet as a -1 hypermultiplet in adjoint. $K'=K+F$. The two different proposals for $\mathfrak{so}_7$ coincide when $(-2K'-\Sigma_i)\cdot\Sigma_i=0$. For $\mathfrak{so}_{n\geq 15}$ we have a further constraint that $\Sigma_i\cdot(-2K'-\Sigma_i)=0$, and for $\mathfrak{e}_8$ we have a further constraint that $(6K'+5\Sigma_i)\cdot\Sigma_i=0$.

For $\mathfrak{g}_a = \mathfrak{so}_{n\geq 15}$, we have the additional constraint $\Sigma_i \cdot (-2K' - \Sigma_i) = 0$. The physical meaning of this constraint is that the intersection points of Σ_i and $-2K' - \Sigma_i$ carry spinor representations, but it is impossible to satisfy anomaly cancellation for $\mathfrak{so}_{n\geq 15}$ if we have hypermultiplets transforming as spinors. There is a similar constraint for $\mathfrak{e}_8$ which states that $(6K' + 5\Sigma_i) \cdot \Sigma_i = 0$.

Tangential intersections: We know that this simple replacement cannot be the full story. We saw at the end of Sec. 3.2.8 that if a curve carrying frozen singularities intersects a curve carrying I_n singularity tangentially, then it traps a hypermultiplet in the two-index symmetric representation of $\mathfrak{su}_n$. In light of this, for $\mathfrak{g}_i = \mathfrak{su}_n$ we define t_a to be the number of tangential intersections of F with D_a. Let $t_i = \sum_a \mu_{i,a} t_a$, in terms of which we write our modified proposal for $\mathfrak{su}_n$ as

$$\rho = \left[\frac{1}{2}(K + F + \Sigma_i) \cdot \Sigma_i - t_i\right] \mathsf{adj} + (-8K - 8F - n\Sigma_i) \cdot \Sigma_i \, \mathsf{fund} + [(-K - F) \cdot \Sigma_i + t_i] \, \mathsf{asym}^2 + t_i \, \mathsf{sym}^2 \,. \quad (3.3.18)$$

This still satisfies anomaly cancellation because $\sigma = -\mathsf{adj} + \mathsf{asym}^2 + \mathsf{sym}^2$ has the property that $\mathrm{tr}_\sigma F^2$ and $\mathrm{tr}_\sigma F^4$ are both zero. This proposal gives correct predictions for models which have a perturbative dual in which case the spectrum can be determined by other methods.

3.4 Noncompact models

Now let us analyze how the anomaly cancellation works out in a few examples. We are particularly interested in 6d SCFTs which supplement the lists given in [9, 10]. As in [9, 10], we expect to model the tensor branch of a 6d SCFT by means of a contractible collection of curves in the F-theory base, with the difference that we will now allow frozen branes as well.

3.4.1 $\mathfrak{so}$-$\mathfrak{sp}$ chains

We first come back to the setup discussed in Sec. 3.2.6. In the type IIA frame, we consider the following chain:

$$\begin{array}{c|c|c|c} \mathrm{O6}_- & \mathrm{O6}_+ & \mathrm{O6}_- & \mathrm{O6}_+ \\ (n+4)\ \mathrm{D6s} & n\ \mathrm{D6s} & (n+4)\ \mathrm{D6s} & n\ \mathrm{D6s} \end{array} \quad (3.4.1)$$

strate. Suppose $(2K' + \Sigma_i) \cdot \Sigma_i \geq 8$, so that we have at least one $\mathsf{sym}^2_{\mathrm{irr}}$. Combining this inequality with the inequalities that the number of vectors are non-negative and the number of adjoints are ≥ -1, we obtain:

$$(-2K' + \Sigma_i) \cdot \Sigma_i \geq -8 \quad (3.3.15)$$

$$\left(-4K' - \frac{n}{4}\Sigma_i\right) \cdot \Sigma_i \geq 0 \quad (3.3.16)$$

$$(2K' + \Sigma_i) \cdot \Sigma_i \geq 8 \quad (3.3.17)$$

Combining the first and third inequalities, we find that $\Sigma_i^2 \geq 0$. Combining the second and third inequalities, we find that $(8 - n)\Sigma_i^2 \geq 32$. These two together imply that $n < 8$.

separated by half-NS5-branes. The leftmost and the rightmost stacks are semi-infinite. This realizes the 6d quiver theory with the structure

$$[\mathfrak{so}_{2n+8}] \quad \mathfrak{sp}_n \quad \mathfrak{so}_{2n+8} \quad [\mathfrak{sp}_n] \tag{3.4.2}$$

where the bracketed algebras are flavor symmetries.

We perform a T-duality to bring this setup into F-theory. The result depends on whether we have O6$_-$ or O6$_+$ on the other fixed locus, see Fig. 3.2.5. The first case is a familiar setup without frozen singularities:

$$\begin{array}{cccc} [\mathfrak{so}_{2n+8}] & \mathfrak{sp}_n & \mathfrak{so}_{2n+8} & [\mathfrak{sp}_n] \\ & 1 & 4 & \\ I_n^* & I_{2n}^{\text{ns}} & I_n^* & I_{2n}^{\text{ns}} \end{array} \tag{3.4.3}$$

where the first, the second, the third row shows the gauge algebra, the negative of the self-intersection number, and the singularity type, respectively. Denoting the two $\mathbb{CP}^1$'s in the middle by D_1 and D_2, the Green–Schwarz contribution to the anomaly is

$$-\frac{1}{2}\left(-\frac{p_1(T)}{4}\cdot K + (\nu(F_{\mathfrak{sp}})D_1 + \nu(F_{\mathfrak{so}})D_2)\right)^2. \tag{3.4.4}$$

In the second case we obtain a setup with frozen singularities:

$$\begin{array}{cccc} [\mathfrak{so}_{2n+8}] & \mathfrak{sp}_n & \mathfrak{so}_{2n+8} & [\mathfrak{sp}_n] \\ & 4 & 1 & \\ I_{2n+8}^{\text{ns}} & \widehat{I}_{n+4}^* & I_{2n+8}^{\text{ns}} & \widehat{I}_{n+4}^*. \end{array} \tag{3.4.5}$$

Note that the gauge group, matter content, and flavor symmetry group of (3.4.5) are identical to those of (3.4.3): only the F-theory realization is different.

Denoting the two middle $\mathbb{CP}^1$'s by $\tilde{D}_1$ and $\tilde{D}_2$ this time, and the canonical class by $\tilde{K}$ to distinguish it from the case above, the Green–Schwarz contribution is now

$$-\frac{1}{2}\left(-\frac{p_1(T)}{4}\cdot(\tilde{K}+\tilde{F}) + (\nu(F_{\mathfrak{sp}})\cdot\frac{1}{2}\tilde{D}_1 + \nu(F_{\mathfrak{so}})\cdot 2\tilde{D}_2)\right)^2. \tag{3.4.6}$$

where the factor $1/2$ in front of $\tilde{D}_1$ is due to the fractionation of D3-branes on O7$_+$, and the factor 2 in front of $\tilde{D}_2$ is due to the embedding index 2 of $\mathfrak{so}_{2n+8} \subset \mathfrak{su}_{2n+8}$. The frozen divisor $\tilde{F}$ is $\tilde{D}_1 + \tilde{D}_3$, where $\tilde{D}_3$ is the noncompact divisor on the far right.

The terms with $\operatorname{tr} F_{\mathfrak{sp}}^2$ and $\operatorname{tr} F_{\mathfrak{so}}^2$ in the two expressions (3.4.4) and (3.4.6) should agree, since they cancel the same 1-loop anomalies. Indeed, we can easily check that

$$\begin{pmatrix} K \\ D_1 \\ D_2 \end{pmatrix}\cdot(D_1, D_2) = \begin{pmatrix} \tilde{K}+\tilde{F} \\ \frac{1}{2}\tilde{D}_1 \\ 2\tilde{D}_2 \end{pmatrix}\cdot\left(\frac{1}{2}\tilde{D}_1, 2\tilde{D}_2\right) = \begin{pmatrix} -1 & 2 \\ -1 & 1 \\ 1 & -4 \end{pmatrix}. \tag{3.4.7}$$

In addition, as observed earlier, $K^2 = (\tilde{K}+\tilde{F})^2$.

3.4.2 $\mathfrak{su}$-$\mathfrak{su}$ chains

Let us next consider the IIA configurations shown in Fig. 3.4.1. The top row and the bottom row are distinguished by the type of the O8-plane; we add 16 D8-branes for the top row to have the same Romans mass for the both rows. The configurations on the left column contain an intersection of the type discussed in section 3.2.8. The configurations on the right column are obtained by moving the half-NS5-brane at the intersection of the 6-branes and the 8-branes away from the intersection. Gauge theoretically, this operation corresponds to giving a vev to hypermultiplets.

Using the discussion in section 3.2.8 and following [19, 18], we find that these configurations realize 6d quivers whose structures are summarized in Fig. 3.4.2. (We did not explicitly write in that figure the standard bifundamental matter hypermultiplets between two consecutive gauge factors.) The type of the O8-plane is correlated to the type of the two-index tensor representation of the $\mathfrak{su}_n$ gauge algebra. Higgsing is done by giving a vev to the hypermultiplet in the antisymmetric or symmetric two-index tensor representations of $\mathfrak{su}_n$, breaking it to $\mathfrak{sp}_{n/2}$ or $\mathfrak{so}_n$. Here for simplicity n is assumed to be even in the former case; if n is odd, the gauge algebra is $\mathfrak{sp}_{\lfloor n/2 \rfloor}$ and one needs to add a flavor to $\mathfrak{su}_{n-8}$.

We can discuss the F-theory duals by T-dualizing the original IIA configurations along the lines of section 3.2.8; the results are shown in Fig. 3.4.3. The top row and the bottom row are distinguished by whether we have the ordinary I_4^* singularity or the frozen $\widehat{I}_4^*$ singularity. For the left column, this noncompact divisor of I_4^* or $\widehat{I}_4^*$ is tangent to the divisor with I_n singularity. To go to the right column, we deform the divisors so that the tangent point is split to two transversal intersection points. This operation in turn changes the singularity type from I_n to I_n^{ns}. The two models on the bottom row realizes 6d quiver gauge theories (the tensor branches of 6d SCFTs) which were not previously possible in an ordinary F-theory compactification without frozen singularities.

Let us name the four divisors in each model as C_1, D_1, D_2, C_2 from the left to the right; $C_{1,2}$ are non-compact and $D_{1,2}$ are compact. The Green–Schwarz contribution to the anomaly can be written down as follows.

For the top row with the non-frozen I_4^* singularity, we have

$$-\frac{1}{2}\left(-K\frac{p_1(T)}{4}+D_1\nu(F_1)+D_2\nu(F_2)\right)^2 \tag{3.4.8}$$

both before and after the Higgsing. For the bottom row with the frozen I_4^* singularity, we have

$$-\frac{1}{2}\left(-(K+C_1)\frac{p_1(T)}{4}+D_1\nu(F_1)+D_2\nu(F_2)\right)^2 \tag{3.4.9}$$

where we used the fact that the frozen divisor is C_1. After the Higgsing, the Green–Schwarz contribution is

$$-\frac{1}{2}\left(-(K+C_1)\frac{p_1(T)}{4}+2D_1\nu(F_1)+D_2\nu(F_2)\right)^2 \tag{3.4.10}$$

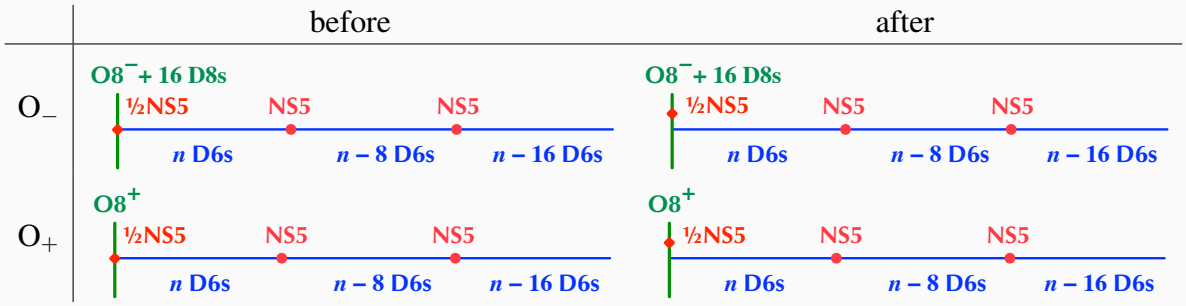


Figure 3.4.1: Four type IIA configurations.

	before	after
O$_-$	$[\mathfrak{su}_{16}], \mathfrak{su}_n + \text{asym}, \mathfrak{su}_{n-8}, [\mathfrak{su}_{n-16}]$	$[\mathfrak{su}_{16}], \mathfrak{sp}_{n/2}, \mathfrak{su}_{n-8}, [\mathfrak{su}_{n-16}]$
O$_+$	$\mathfrak{su}_n + \text{sym}, \mathfrak{su}_{n-8}, [\mathfrak{su}_{n-16}]$	$\mathfrak{so}_n, \mathfrak{su}_{n-8}, [\mathfrak{su}_{n-16}]$

Figure 3.4.2: Quivers. On the upper right corner, we assumed that n is even.

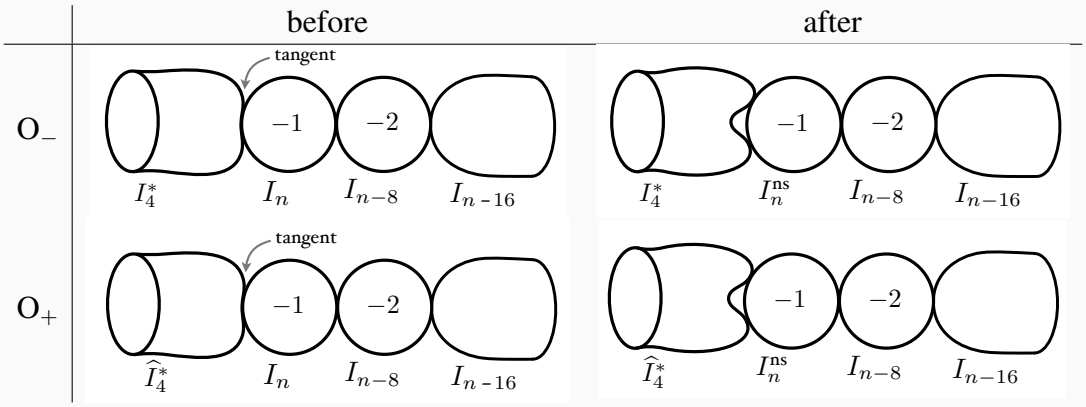


Figure 3.4.3: F-theory duals.

where the factor in front of D_1 is due to the embedding index of $\mathfrak{so}_n \subset \mathfrak{su}_n$.[14]. It is a straightforward exercise to show that these Green–Schwarz contributions correctly cancel the gauge squared and the gauge-gravity part of the one-loop anomalies.

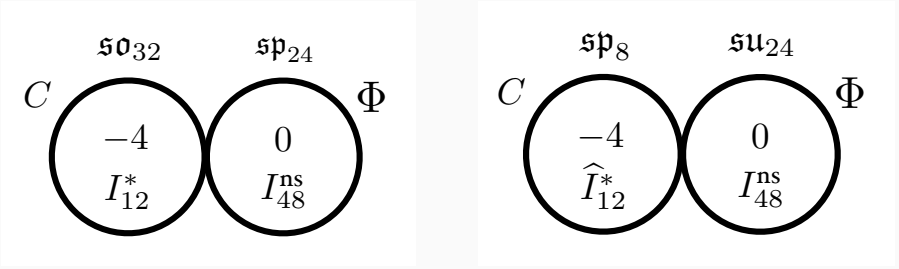
The construction discussed here gives a first indication of how the classification results of [9, 10] need to be modified in order to include frozen branes. We leave a thorough consideration of the effect of frozen branes on this classification to future work.

3.5 Compact models

In this section we discuss some compact models with O7$_+$-planes in F-theory language. The compact models we discuss are, if every O7 is of type O7$_-$, the very classic F-theory models which are dual to the $T^4/\mathbb{Z}_2$ compactifications of the $\mathfrak{so}_{32}$ heterotic string with various standard choices of gauge bundles, originally found in [108, 6, 15]. Their perturbative realizations were first considered by Bianchi and Sagnotti [7] before the second superstring revolution, during which these models were revisited by many others, including by Gimon and Polchinski [8].

3.5.1 The $\mathbb{F}_{-4}$ model and its flip

Without frozen 7-brane: Aspinwall and Gross considered the following model [6]: the F-theory base is the Hirzebruch surface $\mathbb{F}_{-4}$, which is a $\mathbb{CP}^1$ bundle over $\mathbb{CP}^1$ such that the base is a -4 curve. We have the I_{12}^* singularity along the -4 curve C and and a fiber Φ hosts an I_{48}^{ns} singularity; see Fig. 3.5.1(a).



(a) (b)

Figure 3.5.1: In 3.5.1(a), the compact model of [6] on the Hirzebruch surface $\mathbb{F}_{-4}$. In 3.5.1(b), a frozen version.

This model has the following massless matter content:

- $\mathfrak{so}_{32}$ on C and $\mathfrak{sp}_{24}$ on Φ,
- a half-hypermultiplet in $32 \otimes 48$,

[14] In particular it explains the superficially funny-looking η_{O8^+} in [107, (3.23)]

- a hypermultiplet in $\wedge^2 48$, together with
- one supergravity multiplet, one tensor multiplet and 20 neutral hypermultiplets.

Let us remind ourselves how this spectrum can be understood in a dual frame. We start from the heterotic or type I $\mathfrak{so}_{32}$ on a K3, realized as an elliptic fibration over $\mathbb{CP}^1$. The Green–Schwarz mechanism requires that the instanton number of the gauge bundle over K3 is 24. To keep the whole $\mathfrak{so}_{32}$ gauge algebra unbroken, we use 24 point-like instantons. We then collapse the whole 24 instantons to a point. This is known to generate $\mathfrak{sp}_{24}$ on the heterotic side [109]. The spectrum as written above can be found perturbatively on the type I side.

Assuming that the elliptic fiber has small area, we perform fibre-wise the duality between heterotic on T^2 and F-theory on an elliptically-fibered K3. This converts the whole to an elliptically-fibered K3 fibered over $\mathbb{CP}^1$. The $\mathfrak{so}_{32}$ gauge algebra is now realized on the base C as the I_{12}^* singularity, and the point-like instanton is on the fiber Φ as the I_{48}^{ns} singularity.

With a frozen 7-brane: Now, let us flip I_{12}^* to $\widehat{I}_{12}^*$. The anomaly cancellation suggests the following matter content:

- $\mathfrak{sp}_8$ on $\frac{1}{2}C$ and $\mathfrak{su}_{24}$ on 2Φ,
- a hypermultiplet in $16 \otimes 24$,
- two hypermultiplets in $\wedge^2 24$, together with
- one supergravity multiplet, one tensor multiplet and 20 neutral hypermultiplets.

This model is also shown in Fig. 3.5.1(b). It can be Higgsed to

- $\mathfrak{sp}_8$ on $\frac{1}{2}C$ and $\mathfrak{sp}_{12}$ on 2Φ,
- a hypermultiplet in $16 \otimes 24$,
- a hypermultiplet in $\wedge^2 24$, together with
- one supergravity multiplet, one tensor multiplet and 21 neutral hypermultiplets.

Here and below, we mean by the sentence "a gauge algebra $\mathfrak{g}$ on D" that the gauge divisor associated to $\mathfrak{g}$ is D, in the language of Sec. 3.3.

Let us give a derivation of these spectra, using the same duality as in the unfrozen case shown above. We again start from the heterotic or type I $\mathfrak{so}_{32}$ on a K3, realized as an elliptic fibration over $\mathbb{CP}^1$, but with the generalized Stiefel–Whitney class of $\mathrm{Spin}(32)/\mathbb{Z}_2$ being nonzero along the fiber, destroying the vector structure [23]. The maximal possible gauge algebra is now $\mathfrak{sp}_8$. We now need a gauge configuration of instanton number 12 on the K3 surface, since the embedding index of $\mathfrak{sp}_8 \subset \mathfrak{so}_{32}$ is two. We choose to put all 12 point-like instantons at the same place.

The spectrum can be determined perturbatively using the type I description.[15] We find that when the point-like instanton is on a generic point, the spectrum is as in the Higgsed case above, while when it is on a singularity of the form $\mathbb{C}^2/\mathbb{Z}_2$, the spectrum is the one before the Higgsing.

To go to the F-theory frame, we perform the fiber-wise duality as before. This time we use the frozen version reviewed in Appendix B.1, which relates heterotic or type I $\mathfrak{so}_{32}$ on T^2 without vector structure to F-theory on K3 with one frozen singularity. We now have the $\widehat{I}_{12}^*$ singularity on C and the $I_{48}^{\rm ns}$ singularity on Φ. The Higgsing distinguishing the two versions is related to how the residual part of the discriminant with the I_1 type singularity intersects with the fiber Φ.

3.5.2 The $\mathbb{CP}^1 \times \mathbb{CP}^1$ model and its flips

The next compact model has the following perturbative realizations: We first consider in type IIB theory the $T^2/\mathbb{Z}_2 \times T^2/\mathbb{Z}_2$ compactification with O7$_-$ at each $\mathbb{Z}_2$ singularity, together with 16 mobile D7-branes along the first $T^2/\mathbb{Z}_2$ and another 16 mobile D7-branes along the second $T^2/\mathbb{Z}_2$. We can then replace some of the O7$_-$-planes with the O7$_+$-planes. In Appendix B.2, we give the derivation of the spectrum of the models described below from the point of view of the intersecting brane models. The aim here is to understand the spectrum from the F-theory point of view.

Without frozen seven-branes: The perturbative type IIB setup can be promoted to an F-theory setup as follows. Since $T^2/\mathbb{Z}_2 \simeq \mathbb{CP}^1$, we take the F-theory base to be $\mathbb{CP}^1 \times \mathbb{CP}^1$. We pick divisors C and D wrapping each of the $\mathbb{CP}^1$ above. We let each divisor host an I_{12}^* singularity.

At the intersection we expect to have the conformal matter theory (see footnote 9) $\mathbb{D}_{32}$, where $\mathfrak{so}_{32} \times \mathfrak{so}_{32} \subset \mathfrak{so}_{64}$ is gauged.

This is not the end of the story, since the rest of the discriminant locus has singularities at three points on C and three points on D respectively. At each point, a blow-up reveals a -1 curve of $I_8^{\rm ns}$ singularities; see the discussion around (3.5.1) below. After the blow-down each gives rise to the conformal matter $\mathbb{D}_{16}$. For a summary see Fig. 3.5.2.

The matter content is then

- $\mathfrak{so}_{32}$ on C and $\mathfrak{so}'_{32}$ on D,
- the conformal matter $\mathbb{D}_{32}$ gauged by $\mathfrak{so}_{32} \times \mathfrak{so}'_{32} \subset \mathfrak{so}_{64}$
- three copies of $\mathbb{D}_{16}$ gauged by $\mathfrak{so}_{32}$,
- three copies of $\mathbb{D}_{16}$ gauged by $\mathfrak{so}'_{32}$,
- one supergravity multiplet, one tensor multiplet and 13 neutral hypermultiplets.

[15] An analysis after a T-dual along one direction in the T^2 without vector structure is given around equation (B.2.3) of Appendix B.2.2.

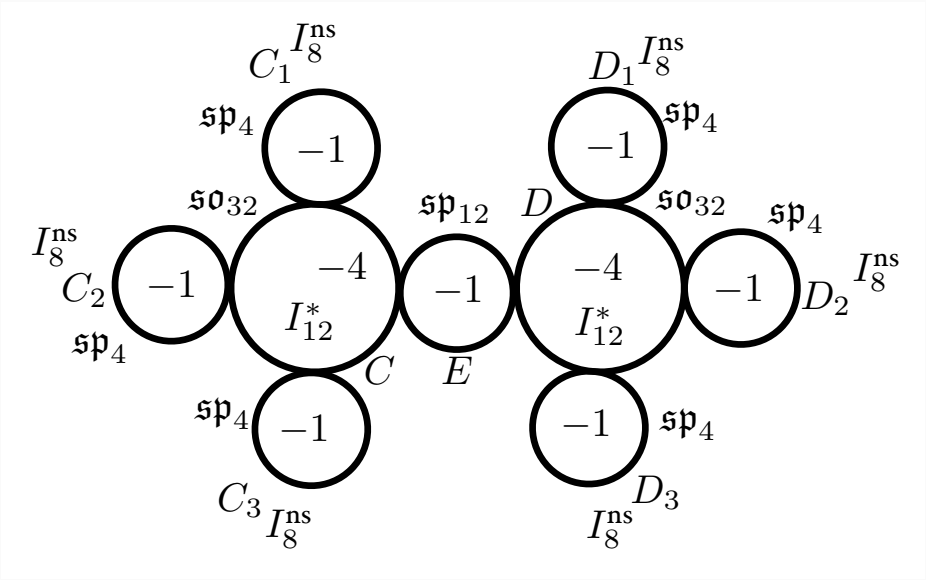


Figure 3.5.2: An F-theory description of the tensor branch of the perturbative model with two O7$_-$ [7, 8]. Upon shrinking E as well as C_1, C_2, C_3, D_1, D_2, D_3 and then Higgsing the resulting conformal matter theory, we recover the original perturbative model.

Upon Higgsing of the conformal matter theories, this results in the standard perturbative massless spectrum of the model:[16]

- $\mathfrak{u}_{16}$ on C and $\mathfrak{u}'_{16}$ on D,
- a hypermultiplet in $16 \otimes 16'$,
- two hypermultiplets in $\wedge^2 16$,
- two hypermultiplets in $\wedge^2 16'$,
- one supergravity multiplet, one tensor multiplet and 20 neutral hypermultiplets,

as can be found in the original papers [7, 8, 97], and reviewed in App. B.2.1. The F-theory interpretation of this Higgsed spectrum was given in [90, 110]; this study eventually led to a refined understanding of the relation between F-theory and O7$_-$ [111].

We note that the T-duality between Type IIB on $T^2/\mathbb{Z}_2 \times T^2/\mathbb{Z}_2$, which we used here, and Type I on $T^4/\mathbb{Z}_2$, as originally considered, was first discussed in [112]. We also mention that when each O7$_-$ has four D7s on top of it, then the perturbative orientifold construction can be subtly modified so that the system is slightly on the tensor branch side, rather than on the Higgs branch side, of the conformal point, as noticed early in the study of orientifolds [113, 114]

[16] The $\mathfrak{u}_1$ part of both $\mathfrak{u}_{16}$ are known to get Higgsed by the Green-Schwarz mechanism, eating one neutral hypermultiplet each, and becoming massive [97, Sec. 2]. Here we follow the older perturbative string terminology.

Detailed discussion of the geometry: Before discussing the case with frozen 7-branes, we pause here to record the details of the geometry. Let us choose coordinates $([s,t],[u,v])$ on $\mathbb{CP}^1 \times \mathbb{CP}^1$. We put I_{12}^* along $t=0$ and $v=0$. The equation defining the elliptic fibration was derived in [108] but we follow the notation of [15, Eq. (42)]:

$$y^2 = x^3 + tvp_{3,3}(s,t,u,v)x^2 + t^8v^8x, \tag{3.5.1}$$

where $p_{3,3}$ is bihomogeneous of degree $(3,3)$. This equation is not in Weierstrass form. By completing the cube, we find

$$f = t^2v^2\left(t^6v^6 - \frac{1}{3}p_{3,3}^2\right), \tag{3.5.2}$$

$$g = t^3v^3p_{3,3}\left(-\frac{1}{3}t^6v^6 + \frac{2}{27}p_{3,3}^2\right), \tag{3.5.3}$$

$$\Delta = t^{18}v^{18}(2t^3v^3 + p_{3,3})(2t^3v^3 - p_{3,3}). \tag{3.5.4}$$

We indeed see I_{12}^* along $t=0$ and $v=0$. Therefore we have the $\mathbb{D}_{32}$ conformal matter at $t=v=0$.

In addition, we see two components

$$2t^3v^3 = \pm p_{3,3} \tag{3.5.5}$$

of the residual discriminant. They intersect with $t=0$ at the three points $t=p_{3,3}=0$, and similarly with $v=0$ at the three points $v=p_{3,3}=0$.

Let us study the intersection points with $t=0$ in more detail. We locate one of the intersection points at $t=u=0$. In other words, assume we can write $p_{3,3} = u\widehat{p}_{3,2} + t\tilde{p}_{2,3}$. Multiplicities of f, g, and Δ at $t=u=0$ are easily seen to be 4, 6, and 20 so we have a conformal fixed point and we need to blow up. To perform the blowup, we work in the affine coordinate chart $v=s=1$. In one coordinate chart of the blowup, we have $t_1=t$, $u_1=u/t$, and the Weierstrass coefficients and discriminant become

$$f_1 = t_1^4 - \frac{1}{3}(u_1\widehat{p} + \tilde{p})^2, \tag{3.5.6}$$

$$g_1 = (u_1\widehat{p} + \tilde{p})\left(-\frac{1}{3}t_1^4 + \frac{2}{27}(u_1\widehat{p} + \tilde{p})^2\right), \tag{3.5.7}$$

$$\Delta_1 = t_1^8(2t_1^2 + u_1\widehat{p} + \tilde{p})(2t_1^2 - u_1\widehat{p} - \tilde{p}). \tag{3.5.8}$$

The exceptional divisor $t_1=0$ supports an I_8 fiber, since the orders of vanishing are $(0,0,8)$, and there is monodromy: the usual branch divisor $(g_1/f_1)|_{t_1}=0$ vanishes at $u_1=t_1=0$ in this chart and has a single order of vanishing. Thus, this is I_8^{ns} and the gauge algebra is $\mathfrak{sp}_4$. No matter is visible in this chart. Note that this branch point is the point at which the residual discriminant meets the exceptional divisor. The multiplicities at this point are 2, 3, 10 which is consistent with the enhancement from A_7 to D_8 which is expected at such a point. In the other coordinate chart

of the blowup, we have $t_2 = t/u, u_2 = u$. The Weierstrass coefficients and discriminant become

$$f_2 = t_2^2(u_2^4 - \frac{1}{3}(\widehat{p} + t_2\tilde{p})^2), \tag{3.5.9}$$

$$g_2 = t_2^3(\widehat{p} + t_2\tilde{p})\left(-\frac{1}{3}u_2^4 + \frac{2}{27}(\widehat{p} + t_2\tilde{p})^2\right), \tag{3.5.10}$$

$$\Delta_2 = t_2^{18}u_2^8(2u_2^2 + \widehat{p} + t_2\tilde{p})(2u_2^2 - \widehat{p} - t_2\tilde{p}) \tag{3.5.11}$$

and we indeed see the exceptional divisor $u_2 = 0$ meeting the original I_{12}^* at $t_2 = 0$. This intersection point also provides the second branch point defining the monodromy.

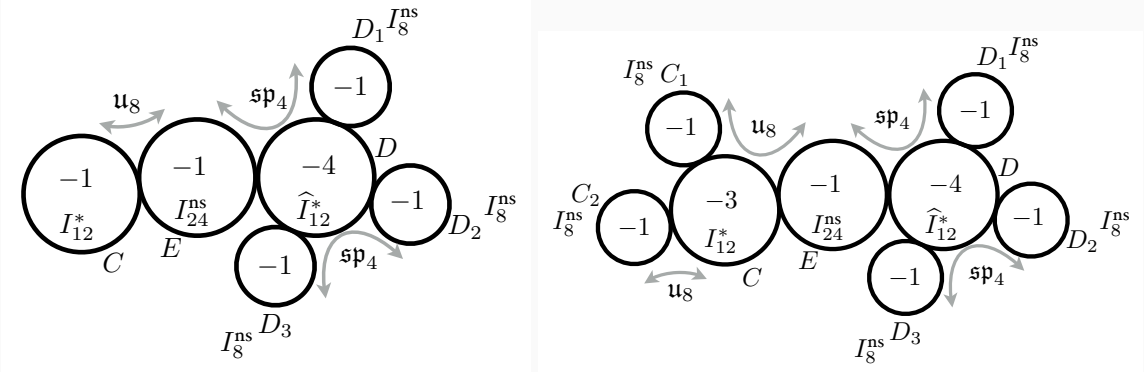


Figure 3.5.3: In 3.5.3(a), a compact model, obtained from the $\mathbb{CP}^1 \times \mathbb{CP}^1$ model of Fig. 3.5.2 by freezing one of the I_{12}^* curves and shrinking some I_8^{ns}. This model reproduces the spectrum of the perturbative model described in appendix B.2.2. In the sense of section 3.3, the gauge divisor for $\mathfrak{su}_8$ is $2(C + E)$, while the one for the $\mathfrak{sp}_4$s are $E + \frac{1}{2}D + D_1$ and $\frac{1}{2}D + D_2 + D_3$. In 3.5.3(b), an enhanced version of the same model, where a second $\mathfrak{u}_8$ gauge algebra appears. The gauge divisors are $E + \frac{1}{2}D + D_1$, $\frac{1}{2}D + D_2 + D_3$, $C + C_1 + 2E$, $C + C_2$.

With one frozen 7-brane: Now let us flip the I_{12}^* on C_2 to $\widehat{I}_{12}^*$. We then have the following setup, see Fig. 3.5.3(a):[17]

- $\mathfrak{u}_8$ on $2(C + E)$,
- $\mathfrak{sp}_4'$ on $E + \frac{1}{2}D + D_1$ and $\mathfrak{sp}_4''$ on $\frac{1}{2}D + D_2 + D_3$,
- hypermultiplets in $\mathsf{8} \otimes \mathsf{8}'$, and $\mathsf{8}' \otimes \mathsf{8}''$, $\mathsf{8}'' \otimes \mathsf{8}$,
- two hypermultiplets in $\wedge^2\mathsf{8}$,

[17] As in the previous footnote, we expect the $\mathfrak{u}_1$ part to become massive via the Green-Schwarz mechanism, eating a neutral hypermultiplet.

- one supergravity multiplet, 5 tensor multiplets and 16 neutral hypermultiplets.

This realizes the spectrum of the perturbative model described in appendix B.2.2. Notice that we have blown down some of the -1-curves of the unfrozen model in Fig. 3.5.2.

An enhanced variant of this model can be obtained by blowing up the -1-curves; see Fig. 3.5.3(b). It has[18]

- $\mathfrak{sp}_4$ on $E + \frac{1}{2}D + D_1$ and $\mathfrak{sp}'_4$ on $\frac{1}{2}D + D_2 + D_3$,
- $\mathfrak{u}''_8$ on $C + C_1 + 2E$ and $\mathfrak{u}'''_8$ on $C + C_2$,
- hypermultiplets in $\mathbf{8} \otimes \mathbf{8}'$, $\mathbf{8}' \otimes \mathbf{8}''$, $\mathbf{8}'' \otimes \mathbf{8}'''$, $\mathbf{8}''' \otimes \mathbf{8}$
- one supergravity multiplet, 7 tensor multiplets and 14 neutral hypermultiplets.

This enhanced model is related to the previous one by a combination of a Higgsing and of two 'small instanton' transitions, where two tensors are traded for two antisymmetric and two neutral hypermultiplets (for a total of $2 \times (28 + 1)$); indeed in such a transition a tensor is traded for 29 hypermultiplets. Other intermediate models are also possible, where only some of these transitions have taken place.

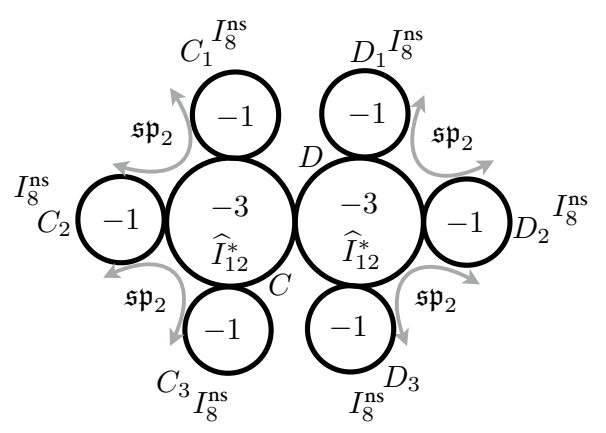


Figure 3.5.4: A compact model realizing the spectrum of the perturbative model with two O7$_+$, discussed in app. B.2.3. The gauge divisors are $C+C_1+C_2$, $C+C_2+C_3$, $D+D_1+D_2$, $D + D_2 + D_3$.

With two frozen seven-branes: We can further flip the I_{12}^* on C_1 to $\widehat{I}_{12}^*$. We then have the following setup, see Fig. 3.5.4:

[18] As in the previous two footnotes, we expect the $\mathfrak{u}_1$ parts to become massive via the Green-Schwarz mechanism, eating neutral hypermultiplets.

- $(\mathfrak{sp}_2)_{1,2,3,4}$, supported on $C+C_1+C_2$, $C+C_2+C_3$, $D+D_1+D_2$, $D+D_2+D_3$ respectively,
- hypermultiplets in $\mathsf{4}_i \otimes \mathsf{4}_j$ for $i < j$,
- one supergravity multiplet, 7 tensor multiplets, and 14 neutral hypermultiplets.

Chapter 4

State sum constructions of spin-TFTs and string net constructions of fermionic phases of matter

4.1 Introduction

The general purpose of this chapter is to explore the properties of spin-topological quantum field theories in $2+1$ dimensions [115] and their relation to fermionic gapped phases of matter [116, 117]. A concrete objective of this chapter is to leverage the relation between these two notions in order to produce explicit lattice Hamiltonians for new fermionic phases of matter.

Spin-topological quantum field theories are topological quantum field theories which are defined only on manifolds equipped with a spin structure. Fermionic phases of matter are phases defined by a microscopic local Hamiltonian which contains fermionic degrees of freedom. The relation between these two notions is most obvious for relativistic theories, thanks to the spin-statistics theorem. It is far from obvious for non-relativistic theories or discrete lattice systems [118].

Standard (unitary) TFTs in $2+1$ dimensions are rather well-understood in terms of properties of their line defects, which form a modular tensor category [119, 120]. A similar characterization for spin-TFTs is not as well developed. The expected relation to fermionic phases of matter suggests the existence of a formulation involving some kind of modular super-tensor category, involving vector spaces with non-trivial fermion number grading. We do not know how to give such a description or how to reconstruct a spin-TFT from this type of data.

Instead, we follow a different strategy: we encode a spin-TFT $\mathfrak{T}_s$ into the data of a "shadow" TFT $\mathfrak{T}_f$, a standard TFT equipped with an extra piece of data, a fermionic quasi-particle Π which fuses with itself to the identity. [1]

In the language of [118], the spin-TFT is obtained from its shadow by a procedure of "fermionic anyon condensation". Conversely, if we pick a spin manifold M and add up the $\mathfrak{T}_s$ partition function over all possible choices of spin structure η we recover the $\mathfrak{T}_f$ partition function.

[1] See appendix C.1 for a simple justification of this statement for TFTs which are associated to 2d RCFTs.

The relation between spin-TFTs and standard TFTs equipped with appropriate fermionic quasi-particles was also discovered in the mathematical literature [121]. This reference proposes a Reshetikhin-Turaev-like construction of a spin TFT partition function from the data of a modular tensor category equipped with invertible fermionic lines. The partition function of the spin TFT summed over the possible choices of spin structure reproduces the Reshetikhin-Turaev partition function of the underlying modular tensor category.

Similarly, the relation between fermionic phases of matter and bosonic phases equipped with a special fermionic quasi-particle was proposed in [122–124] as a form of "gauging fermionic parity".

It is natural to wonder if all spin TFTs should admit a shadow. We believe that should be the case. Given a spin TFT and a spin manifold, we can add up the partition function over all possible spin structures in order to define tentatively the partition function of its shadow. This procedure essentially corresponds to "gauging fermionic parity" and assigns to every spin manifold a partition function which does not depend on a choice of spin structure. The key question is if one can extend this definition to general manifolds which may not admit a spin structure. In $2+1$ dimensions TFTs can be reconstructed from the properties of their quasi-particles, which should be computable from the data of spin manifold partition functions.

The fermionic anyon condensation procedure computes the partition function of $\mathfrak{T}_s$ on a spin manifold M from partition functions of $\mathfrak{T}_f$ on M decorated by collections of fermionic quasi-particles. The calculation involves some crucial signs involving the Gu-Wen Grassmann integral [125] and a choice of spin structure on M.

A slightly more physical perspective on the construction can be given as follows. Consider some microscopic bosonic physical system S which engineers $\mathfrak{T}_f$ at low energy. Combine S with a system of free massive fermions. The Π quasi-particles in S can combine with the free fermions ψ to produce a a bosonic composite particle $\psi\Pi$. Condensation of $\psi\Pi$ produces a new, fermionic phase of matter which we identify as a physical realization of $\mathfrak{T}_s$.

We will focus in most of the chapter on theories which admit a state-sum construction. Concretely, that means that the shadow TFT $\mathfrak{T}_f$ is fully captured by the data of a spherical fusion category $\mathcal{C}_f$, which can be fed into the Turaev-Viro construction [12] of the partition function or the Levin-Wen construction [126] of a bosonic commuting projector Hamiltonian. We will learn how to modify these standard constructions to compute the partition function of $\mathfrak{T}_s$ on a spin manifold and a fermionic commuting projector Hamiltonian for $\mathfrak{T}_s$. This is an extension of the proposal of [127].

As an application of these ideas, we propose an explicit construction for all the fermionic SPT phases which are predicted by the spin-cobordism groups [128]. In particular, this includes phases which lie outside of the Gu-Wen super-cohomology construction. The classification of such phases has been previously proposed in [123], and our results agree with theirs. The novelty here is that we construct explicit state sums and Hamiltonians for all the phases and make explicit their dependence on spin structure. Furthermore, we give a cohomological description of the classification and determine explicitly the group structure of fermionic SPT phases under the stacking operation.

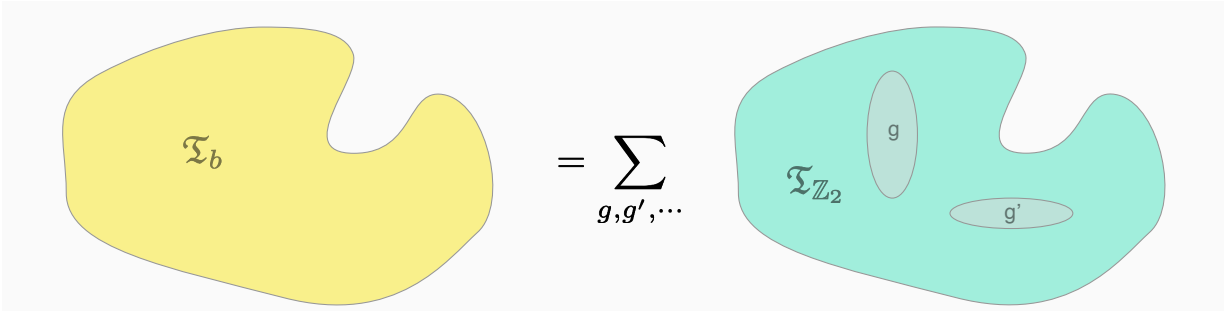


Figure 4.2.1: A graphical depiction of the map from $\mathfrak{T}_{\mathbb{Z}_2}$ to $\mathfrak{T}_b$. On the right we have the partition function of $\mathfrak{T}_{\mathbb{Z}_2}$ on a three-dimensional manifold, equipped with a $\mathbb{Z}_2$ flat connection. We represent the connection as a collection of domain walls implementing $\mathbb{Z}_2$ symmetry transformations g, g', etc. On the left we have the partition function of $\mathfrak{T}_b$, obtained by summing the $\mathfrak{T}_{\mathbb{Z}_2}$ partition function over all possible choices of $\mathbb{Z}_2$ flat connection.

4.2 Overview

4.2.1 One-form symmetries and their anomalies

In order to understand the relation between $\mathfrak{T}_s$ and $\mathfrak{T}_f$, it is useful to look at an analogous relation between standard "bosonic" TFTs. Consider TFTs $\mathfrak{T}_{\mathbb{Z}_2}$ endowed with a (non-anomalous) $\mathbb{Z}_2$ global symmetry, i.e. TFTs which are defined on manifolds equipped with a flat connection. The dimension of space-time is arbitrary at this stage. For a mathematical definition TFTs with symmetries in $2+1d$, see e.g. [129–131] and references therein.

Given such a TFT, we can build a new TFT $\mathfrak{T}_b$ by coupling the $\mathbb{Z}_2$ global symmetry to a dynamical gauge field. The partition function for $\mathfrak{T}_b$ on a manifold M is computed by summing up the $\mathfrak{T}_{\mathbb{Z}_2}$ partition functions over all possible inequivalent $\mathbb{Z}_2$ flat connections (with the same weight):

$$Z[M;\mathfrak{T}_b] = \frac{1}{|H^0(M,\mathbb{Z}_2)|} \sum_{[\alpha_1]\in H^1(M,\mathbb{Z}_2)} Z[M;\mathfrak{T}_{\mathbb{Z}_2};[\alpha_1]] \tag{4.2.1}$$

The theory $\mathfrak{T}_b$ is always equipped with a bosonic quasi-particle B, the Wilson line defect, which fuses with itself to the identity in a canonical way. We can recover $\mathfrak{T}_{\mathbb{Z}_2}$ from $\mathfrak{T}_b$ by condensing B. Intuitively, the insertion of $A = 1 \oplus B$ along a cycle in M forces the flat connection to be trivial along that cycle (i.e. the partition function vanishes unless the holonomy of the connection is trivial). Adding a sufficient number of A's to M will set the flat connection to zero.

In 2+1 dimensions, it is useful to think about this process as gauging a (non-anomalous) $\mathbb{Z}_2$ 1-form symmetry generated by B. By definition, a global $\mathbb{Z}_2$ 1-form symmetry is parameterized by an element of $H^1(M,\mathbb{Z}_2)$ [132]. Gauging this symmetry amounts to coupling the theory to a flat $\mathbb{Z}_2$-valued 2-form gauge field [2] $[\beta_2] \in H^2(M,\mathbb{Z}_2)$. Thus $\mathfrak{T}_b$ has more structure than an

[2] We will try be be careful and distinguish a 2-cocycle β_2 from its cohomology class $[\beta_2]$.

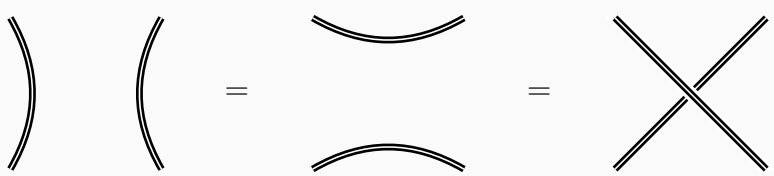


Figure 4.2.2: Wilson lines in $\mathbb{Z}_2$ gauge theory have trivial statistics and can be freely recombined. We use a double-line notation for quasi-particles and line defects to indicate a choice of framing, but the Wilson loops have no framing dependence, i.e. represent bosonic quasi-particles. In general, these abstract properties characterize the quasi-particle generators B of non-anomalous $\mathbb{Z}_2$ 1-form symmetries.

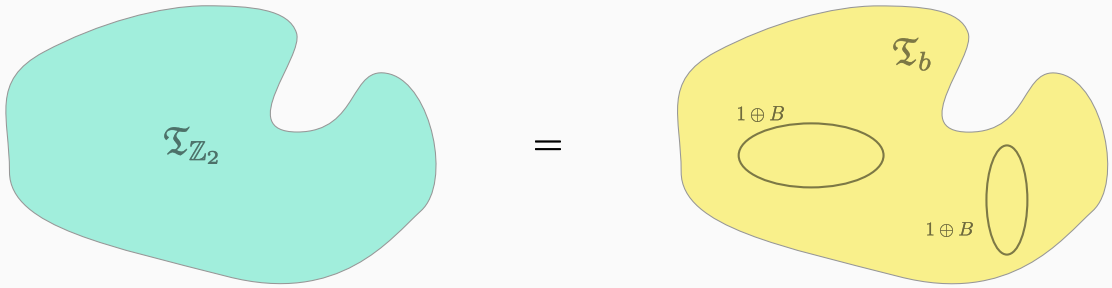


Figure 4.2.3: A graphical depiction of the map from $\mathfrak{T}_b$ to $\mathfrak{T}_{\mathbb{Z}_2}$. On the right we have the partition function of $\mathfrak{T}_b$ on a three-dimensional manifold, possibly decorated with Wilson line operators B along non-trivial cycles, dual to the domain walls of the previous picture. Abstractly, the choice of Wilson lines equips the manifold with a flat connection $[\beta_2]$ for a dual $\mathbb{Z}_2$ 1-form symmetry of $\mathfrak{T}_b$. Summing over all choices gives back the $\mathfrak{T}_{\mathbb{Z}_2}$ partition function.

ordinary TFT: it can associate a partition function to a manifold equipped with a 2-form gauge field $[\beta_2]$.

Concretely, we can triangulate the manifold M and represent β_2 as a 2-cocycle, an assignment of elements of $\mathbb{Z}_2$ to faces of the triangulation such that the sum over faces of each tetrahedron vanishes. [3] We can define the partition function $Z[M; \mathfrak{T}_b; \beta_2]$ of $\mathfrak{T}_b$ coupled to β_2 by decorating M with B lines which pass an (even) odd number of times through each face labelled by the

[3] An arbitrary $\mathbb{Z}_2$-valued function on faces is called a 2-cochain with values in $\mathbb{Z}_2$, and the condition that the sum over faces of each tetrahedron vanishes is written as $\delta\beta_2 = 0$, i.e. the 2-cochain is closed. A 1-form gauge transformation is parameterized by a 1-cochain λ_1, i.e. a $\mathbb{Z}_2$-valued function on the links, and transforms β_2 to $\beta_2 + \delta\lambda_1$.

(trivial) nontrivial element of $\mathbb{Z}_2$. [4]

$$Z[M;\mathfrak{T}_b;\beta_2] = \frac{1}{|H^0(M,\mathbb{Z}_2)|}\sum_{[\alpha_1]\in H^1(M,\mathbb{Z}_2)} (-1)^{\int_M \alpha_1\cup\beta_2} Z[M;\mathfrak{T}_{\mathbb{Z}_2};[\alpha_1]] \tag{4.2.2}$$

An even number of B lines enter each tetrahedron and can be connected to each other in any way we wish without changing the answer, thanks to the statistics and fusion properties of B. It is relatively straightforward to verify that the partition function does not change if we replace β_2 with a gauge-equivalent cocycle $\beta_2 + \delta\lambda_1$ or if we change the triangulation of M. In either case, the collection of B lines is deformed or re-organized. Thus

$$Z[M;\mathfrak{T}_b;\beta_2+\delta\lambda_1] = Z[M;\mathfrak{T}_b;\beta_2] \equiv Z[M;\mathfrak{T}_b;[\beta_2]] \tag{4.2.3}$$

Summing up this decorated partition function over all possible β_2 will insert enough A's to project us back to the partition function of $\mathfrak{T}_{\mathbb{Z}_2}$:

$$Z[M;\mathfrak{T}_{\mathbb{Z}_2}] = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|}\sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} Z[M;\mathfrak{T}_b;[\beta_2]] \tag{4.2.4}$$

We can introduce extra signs to select a specific $\mathbb{Z}_2$ flat connection α_1:

$$Z[M;\mathfrak{T}_{\mathbb{Z}_2};\alpha_1] = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|}\sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} (-1)^{\int_M \alpha_1\cup\beta_2} Z[M;\mathfrak{T}_b;[\beta_2]] \tag{4.2.5}$$

Vice versa, we can consider a theory $\mathfrak{T}_b$ equipped with a non-anomalous $\mathbb{Z}_2$ 1-form symmetry generated by some quasi-particle B. Gauging the $\mathbb{Z}_2$ 1-form symmetry with the same formulae 4.2.4 and 4.2.5 results into a new theory $\mathfrak{T}_{\mathbb{Z}_2}$ which is always equipped with a $\mathbb{Z}_2$ global symmetry generated by Wilson surfaces.

Now we can go back to $\mathfrak{T}_f$. By definition, this theory contains a particle which is a fermion. That is, a particle Π which is (1) an abelian anyon (2) generates a $\mathbb{Z}_2$ subgroup in the group of abelian anyons and (3) has topological spin -1. The first two conditions mean that $\mathfrak{T}_f$ has a 1-form $\mathbb{Z}_2$ symmetry, while the third one implies that this symmetry is anomalous, i.e. there is an obstruction to coupling the theory to a 2-form gauge field in $2+1$ dimensions in a gauge-invariant manner.

Concretely, in order to couple $\mathfrak{T}_f$ to the 2-cocycle β_2, we again pick a triangulation of M. Up to some choices of conventions for how to frame the quasi-particle worldlines, we can populate M with Π lines which pass an (even) odd number of times through each face labelled by the (trivial) non-trivial element of $\mathbb{Z}_2$, joined together inside each tetrahedron. This produces some tentative partition function $Z[M;\mathfrak{T}_f;\beta_2]$. The anomalous nature of the $\mathbb{Z}_2$ 1-form symmetry implies that the partition function changes by some signs when the 2-cocycle β_2 is replaced by a cohomologous one, i.e. when the Π lines are deformed and recombined. Signs may also arise

[4] We write concrete elements of $\mathbb{Z}_2$ additively. That is, the trivial element will be denoted 0, while the nontrivial one will be 1. In particular, when we discuss cochains with values in $\mathbb{Z}_2$, we will write the group operation additively.

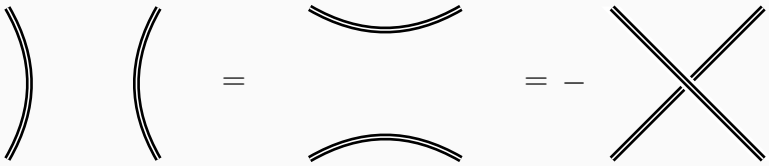


Figure 4.2.4: The Π lines have fermionic statistics and thus extra signs may occur as the worldlines are recombined.

when one re-triangulates M and, obviously, if we change our conventions of how to connect or frame the collection of Π lines representing β_2.

It is quite clear that the anomaly we encounter here does not depend on the specific choice of theory. If we are given two such TFTs $\mathfrak{T}_1$ and $\mathfrak{T}_2$, then their product $\mathfrak{T}_1 \times \mathfrak{T}_2$ has a standard, non-anomalous 1-form symmetry with generator $\Pi_1\Pi_2$. That means that we can define unambiguously the partition function for the product theory coupled to a background $\mathbb{Z}_2$ two-form connection, implemented by decorating M by a collection of $\Pi_1\Pi_2$ defects.

As we are considering a product theory and products of lines in the two factors, we can factor the partition function as

$$Z[M;\mathfrak{T}_1 \times \mathfrak{T}_2;\beta_2] = Z[M;\mathfrak{T}_1;\beta_2]Z[M;\mathfrak{T}_2;\beta_2] \tag{4.2.6}$$

Thus the individual partition functions can only change sign simultaneously under gauge transformations of changes of triangulation.

We would like to argue that we can pick our conventions of how to connect and frame Π lines in such a way that the gauge and re-triangulation anomalies coincide with the ones which emerged in the study of Gu-Wen fermionic SPT phases [125] and their relation to spin-TFTs [118]. The Gu-Wen Grassmann integral combined with a spin-structure-dependent sign gives a $\mathbb{Z}_2$-valued function $z_\Pi(M;\beta_2)$ of a triangulated manifold endowed with a cocycle β_2 and a spin structure. This function changes in a specific manner as one changes the cocycle by a gauge transformation $\beta_2 \to \beta_2 + \delta\lambda_1$ or the triangulation. We claim these are the same transformation rules as for $Z[M;\mathfrak{T};\beta_2]$.

In particular, the combination $z_\Pi(\beta_2)Z[M;\mathfrak{T}_f;\beta_2]$ is well-defined and gives us a spin-TFT with a bosonic $\mathbb{Z}_2$ one-form symmetry. Gauging that symmetry gives us the spin-TFT $\mathfrak{T}_s$, with a partition function

$$Z[M;\mathfrak{T}_s] = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|}\sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} z_\Pi(\beta_2)Z[M;\mathfrak{T}_f;\beta_2] \tag{4.2.7}$$

This is our basic prescription to recover $\mathfrak{T}_s$ from its shadow $\mathfrak{T}_f$.

An alternative way to express the expected anomalous transformation laws of $Z[M;\mathfrak{T}_f;\beta_2]$ is to say that the 1-form $\mathbb{Z}_2$ symmetry generated by the Π lines can only be gauged if we regard the (2+1)-dimensional theory $\mathfrak{T}_f$ as living on a boundary of a (3+1)-dimensional TFT containing

a 2-form gauge field β_2. Concretely, the action of this (3+1)-dimensional TFT is [118]

$$S_4 = i\pi \int_{M_4} \beta_2 \cup \beta_2. \tag{4.2.8}$$

This action is invariant under $\beta_2 \to \beta_2 + \delta\lambda_1$ if M_4 is closed, but on a general compact manifold it varies by a boundary term

$$S_4 \to S_4 + i\pi \int_{\partial M_4} \mathcal{A}(\beta_2, \lambda_1), \tag{4.2.9}$$

where the $\mathbb{Z}_2$-valued 3-cochain $\mathcal{A}$ is given by

$$\mathcal{A}(\beta_2, \lambda_1) = \lambda_1 \cup \beta_2 + \beta_2 \cup \lambda_1 + \lambda_1 \cup \delta\lambda_1. \tag{4.2.10}$$

Note that one cannot discard the first two terms in parentheses because the cup product is not supercommutative on the cochain level.

The anomalous nature of the 1-form $\mathbb{Z}_2$ symmetry means that when $\mathfrak{T}_f$ on $M = \partial M_4$ is coupled to a 2-form gauge field β_2, its partition function, with an appropriate choice of conventions for drawing and framing the Π lines encoding β_2, transforms under 1-form gauge symmetry precisely as in (4.2.9).

$$Z[M; \mathfrak{T}_f; \beta_2 + \delta\lambda_1] = (-1)^{\int_M \mathcal{A}(\beta_2, \lambda_1)} Z[M; \mathfrak{T}_f; \beta_2] \tag{4.2.11}$$

More generally, both gauge transformations and changes of triangulations can be interpreted as triangulated bordisms $M \times [0,1]$ with β_2 defined over the whole 4-manifold, interpolating between $\beta_2|_0$ and $\beta_2|_1$ at the two ends. Then $Z[M; \mathfrak{T}_f; \beta_2]$ changes under such manipulations as

$$Z[M; \mathfrak{T}_f; \beta_2|_1] = e^{i\pi \int_{M\times[0,1]} \beta_2 \cup \beta_2} Z[M; \mathfrak{T}_f; \beta_2|_0]. \tag{4.2.12}$$

4.2.2 Shadow of a product theory

The fermionic sign $z_\Pi(\beta_2)$ is almost multiplicative [118]:

$$z_\Pi(\beta_2) z_\Pi(\beta_2') = (-1)^{\int_M \beta_2 \cup_1 \beta_2'} z_\Pi(\beta_2 + \beta_2') \tag{4.2.13}$$

This observation allows us to re-write the product of two spin-TFT partition functions in a suggestive way

$$Z[M; \mathfrak{T}_s] Z[M; \mathfrak{T}_s'] = \frac{|H^0(M, \mathbb{Z}_2)|}{|H^1(M, \mathbb{Z}_2)|} \sum_{[\beta_2] \in H^2(M, \mathbb{Z}_2)} z_\Pi(\beta_2) Z[M; \mathfrak{T}_f \times_f \mathfrak{T}_f'; \beta_2] \tag{4.2.14}$$

with

$$Z[M;\mathfrak{T}_f\times_f\mathfrak{T}'_f;\beta_2]\equiv\frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|}\sum_{[\beta'_2]\in H^2(M,\mathbb{Z}_2)}(-1)^{\int_M(\beta_2+\beta'_2)\cup_1\beta'_2}Z[M;\mathfrak{T}_f;\beta_2+\beta'_2]Z[M;\mathfrak{T}'_f;\beta'_2] \tag{4.2.15}$$

This expression defines an operation $\times_f$ which maps the shadows of two spin-TFTs to the shadow of their product.

The physical interpretation of this formula is straightforward. The product of shadow theories $\mathfrak{T}_f$ and $\mathfrak{T}'_f$ is endowed with a bosonic $\mathbb{Z}_2$ 1-form symmetry generated by the product $\Pi\otimes\Pi'$ of the fermionic lines of the two theories. Gauging that symmetry leaves us with a new theory with fermionic 1-form symmetry generated by Π_1, which we interpret as the shadow of the product $\mathfrak{T}_s\times\mathfrak{T}'_s$ of the corresponding spin TFTs. This agrees with the stacking construction proposed in [124].

A simple check of this proposal is that the multiplication is associative: the product of three shadow theories has a $\mathbb{Z}_2\times\mathbb{Z}_2$ bosonic 1-form symmetry with non-trivial generators $\Pi\otimes\Pi'$, $\Pi'\otimes\Pi''$, $\Pi\otimes\Pi''$.

We will use this construction systematically in order to explore the group structure of fermionic SPT phases.

4.2.3 Gu-Wen and beyond

The starting point of the Gu-Wen construction of femionic SPT phases [125] is a group super-cohomology element (ν_3,n_2), i.e. a pair of cochains on BG with values in $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}_2$, respectively, satisfying

$$\delta n_2=0,\quad \delta\nu_3=\frac{1}{2}n_2\cup n_2. \tag{4.2.16}$$

Given a flat G-connection on M, one can pull back the cochains ν_3 and n_2 on BG to cochains on M which we can still denote as ν_3 and n_2. Then the Gu-Wen Grassmann integral $z_\Pi(n_2)$ can be combined with the product of ν_3 over all tetrahedra in M in order to give the partition function of an invertible spin-TFT with a symmetry G.

Our strategy to prove that $z_\Pi(\beta_2)$ captures the anomaly of fermionic 1-form symmetries will be to re-cast the Gu-Wen construction in this form, by defining an appropriate bosonic theory $\mathfrak{T}_f[\nu_3,n_2]$ such that the associated partition function reproduces the product of ν_3 over all tetrahedra in M.

The construction proceeds as follows. A 2-cocycle n_2 gives rise to a central extension

$$0\to\mathbb{Z}_2\to\widehat{G}\to G\to 0 \tag{4.2.17}$$

Consider a bosonic SPT phase for $\widehat{G}$, labelled by a $\widehat{G}$-cocycle $\widehat{\nu}_3$ with values in $\mathbb{R}/\mathbb{Z}$ [115, 133]. We can gauge the $\mathbb{Z}_2$ subgroup and get a bosonic TFT with symmetry G. The resulting theory is essentially an enriched version of the toric code, where the G symmetry acts on quasi-particles in a way determined by n_2 and $\widehat{\nu}_3$. If this theory has a bulk line defect Π which is a fermion and is acted upon trivially by the G symmetry, it is a candidate for a shadow of a Gu-Wen phase.

We will determine the condition for the bulk fermion Π to exist. The existence of Π will restrict $\widehat{\nu}_3$ to be a specific combination of n_2 and a group cochain ν_3 which satisfies (4.2.16). We will denote this bosonic TFT $\mathcal{G}_{\nu_3,n_2}$. The result is a one-to-one map between Gu-Wen fermionic SPT phases and bosonic SET phases of the form $\mathcal{G}_{\nu_3,n_2}$.

We will compute explicitly $Z[M;\mathcal{G}_{\nu_3,n_2};\beta_2]$ to find that it is only non-vanishing if β_2 equals the pull-back of n_2 to M, in which case the partition function is essentially equal to the product of ν_3 over all tetrahedra in M. This will verify that $Z[M;\mathfrak{T}_s]$ for these theories coincide with the Gu-Wen partition sum and $z_\Pi(\beta_2)$ is the correct kernel for fermionic anyon condensation.

Cobordism theory [128, 123] predicts the existence of a more general class of fermionic SPT phases protected by fermion number symmetry together with a global symmetry G, labelled by a triple (ν_3, n_2, π_1), where π_1 is a $\mathbb{Z}_2$-valued 1-cocycle on G, n_2 is a $\mathbb{Z}_2$-valued 2-cochain on G, and ν_3 is an $\mathbb{R}/\mathbb{Z}$-valued 3-cochain on G. We will show that n_2 is in fact a cocycle, and ν_3 and n_2 must again satisfy the Gu-Wen equations (4.2.16). Thus the set of fermionic SPT phases with symmetry G can be identified with the product of the set of Gu-Wen phases and the set $H^1(G,\mathbb{Z}_2)$ parameterized by π_1.

The meaning of π_1 is a group homomorphism from G to $\mathbb{Z}_2$, which is used to pull-back a certain "root" $\mathbb{Z}_2$ fermionic SPT phase along π_1. The "root" $\mathbb{Z}_2$ phase is expected to be the phase whose shadow is the toric code, enriched by the $\mathbb{Z}_2$ symmetry which exchanges the e and m quasi-particles. Such a $\mathbb{Z}_2$ symmetry is not manifest in the standard formulation of the toric code and only emerges at low energy. With a bit of effort, though, one can produce a microscopic description of the toric code with explicit $\mathbb{Z}_2$ symmetry [134], starting from an Ising fusion category.

We will verify that the $\mathbb{Z}_2$-equivariant toric code $\mathcal{I}$ is indeed the shadow of root fermionic SPT phase with $\mathbb{Z}_2$ global symmetry, by explicitly computing $\mathcal{I} \times_f \mathcal{I}$ and matching it with a Gu-Wen phase.

The Ising pull-back phases $\mathcal{I}_{\pi_1}$ can be combined with a standard Gu-Wen phase $\mathcal{G}_{\nu_3,n_2}$ to give a candidate $\mathcal{G}_{\nu_3,n_2} \times_f \mathcal{I}_\pi$ for the shadow of the most general fermionic SPT phase. We will verify this combination is indeed the most general symmetry-enriched version of the toric code which admits a suitable fermion Π.

Finally, we will compute the twisted products of general fermionic SPT phases with the help of a relation of the schematic form

$$\mathcal{I}_{\pi_1} \times_f \mathcal{I}_{\pi_1'} = \mathcal{G}_{\nu_3,n_2} \times_f \mathcal{I}_{\pi_1+\pi_1'} \tag{4.2.18}$$

where the Gu-Wen phase $\mathcal{G}_{\nu_3,n_2}$ is determined canonically from π_1 and π_1'.

This result explicitly realizes the group of fermionic SPT phases as an extension of $H^1(G,\mathbb{Z}_2)$ by the super-cohomology group of Gu-Wen phases (which itself is an extension of $H^2(G,\mathbb{Z}_2)$ by $H^3(G,\mathbb{Z}_2)$). This extension is nontrivial. That is, while the set of fermionic SPT phases is the product of the group $H^1(G,\mathbb{Z}_2)$ and the group of Gu-Wen phases with symmetry G, the abelian group structure on this set is not the product structure.

4.2.4 A Hamiltonian perspective

We would like to describe the relation between a gapped bosonic Hamiltonian which engineers the shadow bosonic TFT $\mathfrak{T}_f$ and a gapped fermionic Hamiltonian which can engineer the related spin TFT $\mathfrak{T}_s$. Again, it is useful to first look at a pair of bosonic Hamiltonians for $\mathfrak{T}_b$ and $\mathfrak{T}_{\mathbb{Z}_2}$, related by gauging standard or 1-form non-anomalous $\mathbb{Z}_2$ symmetries.

The procedure for gauging a standard on-site $\mathbb{Z}_2$ global symmetry of some lattice realization of $\mathfrak{T}_{\mathbb{Z}_2}$ is well understood. One extends the Hilbert space by adding $\mathbb{Z}_2$-valued edge variables playing the role of a flat $\mathbb{Z}_2$ connection α_1. Flatness is imposed locally by extra placquette terms in the Hamiltonian enforcing $\delta\alpha_1 = 0$. The Hamiltonian for $\mathfrak{T}_{\mathbb{Z}_2}$ deformed by the coupling to the flat connection can be denoted as $H_{\mathbb{Z}_2}[\alpha_1]$ and the Hamiltonian on the enlarged Hilbert space is schematically

$$H'_{\mathbb{Z}_2} = H_{\mathbb{Z}_2}\left[\frac{1+\widehat{\sigma}^z}{2}\right]. \tag{4.2.19}$$

Here $\widehat{\sigma}_e^{x,y,z}$ are Pauli matrices acting on the $\mathbb{Z}_2$ variables at the e-th edge. More explicitly, suppose $H_{\mathbb{Z}_2}[\alpha]$ is given as a sum of local terms:

$$H_{\mathbb{Z}_2}[\alpha] = \sum_v H^v_{\mathbb{Z}_2}[\alpha] \tag{4.2.20}$$

where $H^v_{\mathbb{Z}_2}$ acts nontrivially only on the degrees of freedom in a neighborhood of the vertex v. We can take $H^v_{\mathbb{Z}_2}[\alpha]$ to vanish if the connection is not flat in a neighbourhood of v.

Let P_f be a projector which enforces the flatness of $\mathbb{Z}_2$-valued edge variables at a face f. Concretely, denoting edges and faces as pairs and triples of vertices,

$$P_{012} = \frac{1}{2}(1 + \widehat{\sigma}^z_{01}\widehat{\sigma}^z_{12}\widehat{\sigma}^z_{20}). \tag{4.2.21}$$

Then the Hamiltonian in the enlarged Hilbert space is also a sum of local terms

$$H'_{\mathbb{Z}_2} = \sum_v H^v_{\mathbb{Z}_2} + \sum_f (1 - P_f) \tag{4.2.22}$$

The resulting enlarged Hilbert space is then projected to gauge-invariant states by a collection of projectors $U^{\mathbb{Z}_2}_v$ which act by a local $\mathbb{Z}_2$ transformation on the degrees of freedom at the lattice site v and shift the connection on the nearby edges. Concretely, we can write

$$U^{\mathbb{Z}_2}_v = \widehat{u}_v \prod_{v'} \widehat{\sigma}^x_{vv'} \tag{4.2.23}$$

Here $\widehat{u}_v$ acts on the local degrees of freedom of the original theory at v as a local $\mathbb{Z}_2$ symmetry transformations.

More generally, one can define operators $U^{\mathbb{Z}_2}_{\lambda_0}$ which implement gauge transformations $\alpha_1 \to \alpha_1 + \delta\lambda_0$ with a parameter λ_0 which is a $\mathbb{Z}_2$-valued 0-cochain. Absence of anomalies means that

$$U^{\mathbb{Z}_2}_{\lambda_0} U^{\mathbb{Z}_2}_{\lambda'_0} = U^{\mathbb{Z}_2}_{\lambda_0+\lambda'_0} \tag{4.2.24}$$

Hence the final Hilbert space $\mathcal{H}[\mathfrak{T}_b]$ is obtained by the projection

$$U^{\mathbb{Z}_2}_{\lambda_0}|\Psi_b\rangle = |\Psi_b\rangle \tag{4.2.25}$$

Thus we can define a Hamiltonian for $\mathfrak{T}_b$ as

$$H_{\mathbb{Z}_2\to b} = H'_{\mathbb{Z}_2} + \sum_v \frac{1}{2}(1 - U^{\mathbb{Z}_2}_v) \tag{4.2.26}$$

Wilson line quasi-particles can be added at the vertices of the lattice by flipping the sign of the Coulomb branch constraints there. For convenience, we will choose a branching structure on the lattice, taken to be triangular, and define the Hilbert space $\mathcal{H}[\mathfrak{T}_b;\beta_2]$ as

$$U^{\mathbb{Z}_2}_{\lambda_0}|\Psi_b;\beta_2\rangle = (-1)^{\int \lambda_0\cup\beta_2}|\Psi_b;\beta_2\rangle \tag{4.2.27}$$

i.e.

$$H_{\mathbb{Z}_2\to b}[\beta_2] = H'_{\mathbb{Z}_2} + \sum_v \frac{1}{2}\left(1 - (-1)^{\int \lambda^v_0\cup\beta_2}U^{\mathbb{Z}_2}_v\right) \tag{4.2.28}$$

where λ^v_0 is a delta function at the vertex v. Concretely, each face t with $\beta_2(t) = 1$ will contribute a Wilson loop at its first vertex. This makes the $\mathbb{Z}_2$ 1-form symmetry of $\mathfrak{T}_b$ manifest "on-site".

The construction can be readily generalized to non-anomalous symmetry realizations which do not act on-site. We can introduce a triangular lattice in the system, with a lattice scale which is much larger than the scale set by the gap in $\mathfrak{T}_{\mathbb{Z}_2}$, and add the $\mathbb{Z}_2$ connection to the edges of that lattice. Operators $U^{\mathbb{Z}_2}_{\lambda_0}$ with the correct properties will still be defined, up to exponentially suppressed effects.

Conversely, starting from a generic theory $\mathfrak{T}_b$ with non-anomalous 1-form symmetry, the Hilbert space of $\mathfrak{T}_{\mathbb{Z}_2}$ is obtained by first summing the Hilbert spaces of $\mathfrak{T}_b$ with one or none insertions of the B quasi-particle and then projecting to the sub-space which is fixed by the action of closed B string operators, i.e. closed B lines wrapping non-trivial cycles on the space manifold Σ.

We can obtain a more local description by enlarging further the original Hilbert space and the subsequent projector. If the theory $\mathfrak{T}_b$ is given in a form which allows a direct coupling to a 2-form connection on the lattice by a local Hamiltonian $H_b[\beta_2]$ we just make β_2 into a collection of dynamical $\mathbb{Z}_2$ variables attached to the faces of the lattice.

If not, we introduce a new triangular lattice in the system, with a lattice scale which is much larger than the scale set by the gap in $\mathfrak{T}_b$. We can attach a $\mathbb{Z}_2$ variable $\beta_2(t)$ to each face t of the lattice and denote as $\mathcal{H}[\beta_2]$ the space of ground states of $\mathfrak{T}_b$ with a B quasi-particle inserted in the middle of each face with $\beta_2(t) = 1$. In particular, $\mathcal{H}[0]$ is the usual space of ground states of $\mathfrak{T}_b$.

In either case, we define the enlarged Hilbert space as the direct sum $\mathcal{H}' = \oplus_{\beta_2}\mathcal{H}[\beta_2]$ over all 2-cocycles β_2. Concretely, the Hilbert space $\mathcal{H}[\beta_2]$ is realized as the space of zero energy states of a local Hamiltonian $H_b[\beta_2]$ acting on the microscopic Hilbert space. We can realize $\mathcal{H}'$ as the

space of zero energy states of a local Hamiltonian

$$H' = H_b \left[\frac{1 + \widehat{\sigma}^z}{2} \right]. \tag{4.2.29}$$

Here $\widehat{\sigma}_t^{x,y,z}$ are Pauli matrices acting on the $\mathbb{Z}_2$ variables at the t-th face.[5]

Due to the properties of the B quasi-particles, we must have unitary transformations

$$U_{\lambda_1} : \mathcal{H}[\beta_2] \to \mathcal{H}[\beta_2 + \delta\lambda_1] \tag{4.2.30}$$

which move B quasi-particles from one site to another or create or annihilate pairs of B quasi-particles. For example, if λ_1 is concentrated on one edge e, the corresponding unitary operator U_e will move, create or annihilate B particles in the two faces adjacent to that edge. In particular, it will anti-commute with the $\widehat{\sigma}^z$ variables for these two faces, commute with all others.

We expect the U_e operator to be an operator which only acts in the neighborhood of the edge e, i.e. local at the scale of our lattice. There is a certain degree of freedom in defining the U_{λ_1}. As the B quasi-particles are bosons, it should be possible to use that freedom to ensure that different ways to transport the B particles are all equivalent, i.e.

$$U_{\lambda_1} U_{\lambda_1'} = U_{\lambda_1 + \lambda_1'} \tag{4.2.31}$$

In other words, U_{λ_1} implement the 1-form gauge symmetry of the theory $\mathfrak{T}_b$, which should not be anomalous. In particular, for every edge e we have $U_e^2 = 1$, and $[U_e, U_e'] = 0$ for all e, e'. We must also ensure $U_{\delta\mu_0} = 1$ for all 0-cochains μ_0. This requirement means that 1-form symmetry transformations with parameters λ_1 and $\lambda_1 + \delta\mu_0$ are physically indistinguishable.

We want to define the Hilbert space for $\mathfrak{T}_{\mathbb{Z}_2}$ as the subspace of the enlarged Hilbert space fixed by the action of these unitary transformations. We can define a commuting projector Hamiltonian acting on the enlarged Hilbert space $\mathcal{H}'$ as

$$\sum_e \frac{1}{2}(1 - U_e). \tag{4.2.32}$$

It engineers the space of ground states of $\mathfrak{T}_{\mathbb{Z}_2}$.This construction makes the $\mathbb{Z}_2$ global symmetry of $\mathfrak{T}_{\mathbb{Z}_2}$ manifest: it acts on the face variables as $\prod_t \widehat{\sigma}_t^z$ and commutes with the Hamiltonian.

Note that the U_{λ_1} operators for closed 1-cochains, which satisfy $\delta\lambda_1 = 0$, can be identified with the closed B string operators we discussed originally, while the general U_{λ_1} operators are open B string operators. We can denote the closed string operators as $U_{\lambda_1}^{cl}$. They map each summand in the Hilbert space back to itself.

Thus we define a microscopic Hamiltonian for $\mathfrak{T}_{\mathbb{Z}_2}$ as

$$H_{b \to \mathbb{Z}_2} = H_b \left[\frac{1 + \widehat{\sigma}^z}{2} \right] + \sum_e \frac{1}{2}(1 - U_e) \tag{4.2.33}$$

[5] Since in two dimensions any 2-cochain is closed, there is no need for projectors in H'.

acting on the tensor product of the microscopic Hilbert space of $\mathfrak{T}_b$ and of the $\mathbb{Z}_2$ face degrees of freedom

Now consider the case of a fermionic $\mathbb{Z}_2$ 1-form symmetry, i.e. a $\mathbb{Z}_2$ 1-form symmetry with an anomaly (4.2.9). As a warm-up, we can focus on how to define consistently the action of closed Π string operators on the original Hilbert space of ground states for $\mathfrak{T}_f$. If we triangulate the space manifold and pick a 1-cocycle λ_1, i.e. a $\mathbb{Z}_2$-valued function on edges λ_1 satisfying $\delta\lambda_1 = 0$, we can draw a collection of non-intersecting Π lines which cross each edge e $\lambda_1(e)$ times modulo 2. We can relate different such collections for the same λ_1 without ever braiding the Π lines, and thus we should be able to define a corresponding composite string operator $V^{cl}_{\lambda_1}$ acting on the space of ground states of $\mathfrak{T}_f$.

If we compose two such closed string operators $V^{cl}_{\lambda_1}$ and $V^{cl}_{\lambda'_1}$, we get a collection of string which may have intersections. Resolving each intersection will cost us a -1 sign. The total number of intersections modulo 2 should coincide with $\int \lambda_1 \cup \lambda'_1$. Thus we expect to be able to consistently define the closed string operators in such a way that

$$V^{cl}_{\lambda_1} V^{cl}_{\lambda'_1} = (-1)^{\int_\Sigma \lambda_1 \cup \lambda'_1} V^{cl}_{\lambda_1+\lambda'_1} \tag{4.2.34}$$

In particular, there is no consistent way for a ground state to be fixed by all V^{cl}.

There is a natural way to correct the closed string operators in such a way that a consistent projection becomes possible: we can dress them by some extra sign $\sigma_2(\lambda_1)$ which also satisfies

$$\sigma_2(\lambda_1)\sigma_2(\lambda'_1) = (-1)^{\int_\Sigma \lambda_1 \cup \lambda'_1} \sigma_2(\lambda_1 + \lambda'_1) \tag{4.2.35}$$

If the space manifold is endowed with a spin structure, we can use the spin structure to define such a sign. Moreover, the Gu-Wen grassmann integral in two dimensions combined with a spin structure provides a local definition of precisely the same sign $\sigma_2(\lambda_1)$ provided we enlarge the Hilbert space with fermionic degrees of freedom living on faces [118]. In other words, $\sigma_2(\lambda_1)$ can be written as a product of local fermionic operators situated on the edges e for which $\lambda_1(e) \neq 0$.

In order to get a fully explicit and local definition of the space of $\mathfrak{T}_s$ ground states, we need to extend this logic to open Π string operators, or equivalently to V_{λ_1} for not necessarily closed 1-cochains λ_1.

We can proceed as before and consider the sum of Hilbert spaces $\mathcal{H}[\beta_2]$ over all 2-cocycles β_2, where $\mathcal{H}[\beta_2]$ is the space of ground states of $\mathfrak{T}_f$ with a Π quasi-particle inserted in the middle of each face with $\beta_2(t) = 1$. We can define as before unitary operators V_{λ_1} which re-arrange the location of the Π quasi-particles, but the fermionic nature of the quasi-particles, or the anomaly of the corresponding 1-form symmetry, indicates that the algebra of V_{λ_1} will only close up to signs:

$$V_{\lambda_1} V_{\lambda'_1} = V_{\lambda_1+\lambda'_1} (-1)^{\omega_\Sigma(\beta_2,\lambda_1,\lambda'_1)}. \tag{4.2.36}$$

Similar considerations as for the partition function show that the anomaly ω_Σ must be universal for all theories with a fermionic $\mathbb{Z}_2$ 1-form symmetry. We can get a concrete expression for as follows. Consider the 3+1d TFT (4.2.8) on $M_4 = \Sigma \times D^2$, coupled to the 2+1d TFT on $\partial M_4 = \Sigma \times S^1$. The operator V_{λ_1} in the 2+1d theory which implements the $\mathbb{Z}_2$ 1-form symmetry

transformations also shifts the 2-form gauge field β_2 by $\delta\lambda_1$. By continuing λ_1 into the bulk, we may regard V_{λ_1} as a boundary of a codimension-1 defect in the 3+1d TFT. By considering three such defects with parameters λ_1, λ_1' and $-(\lambda_1+\lambda_1')$ meeting at the origin of D^2, one can see that

$$V_{\lambda_1} V_{\lambda_1'} V_{-\lambda_1-\lambda_1'} = \int_\Sigma \omega(\beta_2, \lambda_1, \lambda_1'). \tag{4.2.37}$$

The 2-cochain ω is defined as a solution of the equation

$$\delta\omega(\beta_2, \lambda_1, \lambda_1') = \mathcal{A}(\beta_2, \lambda_1) + \mathcal{A}(\beta_2 + \delta\lambda_1, \lambda_1') - \mathcal{A}(\beta_2, \lambda_1 + \lambda_1'). \tag{4.2.38}$$

Using (4.2.9), we find

$$\omega(\beta_2, \lambda_1, \lambda_1') = \lambda_1 \cup \lambda_1' + \delta\lambda \cup_1 \lambda_1', \tag{4.2.39}$$

where $\cup_1$ is the Steenrod higher product [135] (see also Appendix B.1. of [136] for a brief summary). In particular, we see that $\omega_\Sigma = \int_\Sigma \omega$ does not depend on β_2 in this case.

Another manifestation of the anomaly is that the operators V_{λ_1} are not invariant under $\lambda_1 \mapsto \lambda_1 + \delta\mu_0$, where μ_0 is an arbitrary $\mathbb{Z}_2$-valued 0-cochain. Namely, by considering two defects implementing 1-form gauge transformations with parameters λ_1 and $\lambda_1 + \delta\mu_0$, we find

$$V_{\lambda_1+\delta\mu_0} V_{-\lambda_1} = (-1)^{\int_\Sigma (\mu_0 \cup \delta\lambda_1 + \beta_2 \cup \mu_0 + \mu_0 \cup \beta_2)}. \tag{4.2.40}$$

One way to deal with this anomaly would be to couple the system to the Hamiltonian version of Gu-Wen Grassmann integral. The Gu-Wen Grassmann integral on a bordism geometry $\Sigma \times [0,1]$ with β_2 and $\beta_2 + \delta\lambda_1$ at the two ends will provide dressing operators $U_{\lambda_1}^{z_\Pi}$ which should correct the V_{λ_1} to a set of commuting projectors. This is somewhat cumbersome, though, and we will propose a more direct alternative lattice construction.

We will promote the face variables $\beta_2(f)$ to occupation numbers n_f for fermionic degrees of freedom. Thus at each face we have a pair of fermionic creation and annihilation operators, or equivalently a pair of Majorana fermions γ_f, γ_f'. We combine the individual edge operators V_e with Majorana fermions on the two faces f_L and f_R sharing e and define new edge operators

$$U_e^f = \pm V_e \gamma_{f_L} \gamma_{f_R}' \tag{4.2.41}$$

in such a way as to make the following fermionic Hamiltonian well-defined

$$H_{f\to s} = H_f[n_f] + \sum_e \frac{1}{2}(1 - U_e^f) \tag{4.2.42}$$

The sign in the definition of U_e^f is determined by a certain 1-chain E with values in $\mathbb{Z}_2$. This chain encodes a choice of spin structure on Σ.

If $\mathfrak{T}_f$ admits a Levin-Wen construction, we will show how to incorporate directly the effect of the Π particles to get a string net construction for $\mathfrak{T}_s$.

4.2.5 Open questions and future directions

Classification of fermionic SPT phases can be generalized in several directions. Most obviously, one would like to classify SPT phases protected by $\widehat{G}$ which is a central extension of G by $\mathbb{Z}_2^f$. A natural guess is that the corresponding shadow theory must have both ordinary symmetry G and a fermionic one-form symmetry $\mathbb{Z}_2$, but with a mixed anomaly between the two.

The mixed anomaly is determined by the extension class of the short exact sequence $0 \to \mathbb{Z}_2 \to \widehat{G} \to G \to 0$. Concretely, this means that the shadow theory is described by a G-graded fusion category, but the crossing conditions for the fermion are modified by the 2-cocycle ψ_2 representing the extension class. Physically, intersections of domain walls implementing G symmetry transformations will carry non-trivial fermion number.

Following the approach of Appendix B, we get a generalization of the Gu-Wen equations:

$$\delta \nu_3 = \frac{1}{2} n_2 \cup n_2 + \frac{1}{2} \psi_2 \cup n_2, \quad \delta n_2 = 0. \tag{4.2.43}$$

It would be interesting to study the group structure on the space of such fermionic SPT phases.

Another possible generalization is to extend the discussion to unorientable theories. This is important for classifying fermionic SPT and SET phases with anti-unitary symmetries.

It would be very interesting to extend the shadow theory approach to fermionic phases in higher dimensions. For example, it has been proposed in [118] that 3+1d fermionic phases are related to bosonic phases with an anomalous 2-form $\mathbb{Z}_2$ symmetry, where the 5d anomaly action is

$$\int_{M_5} Sq^2 C_3, \tag{4.2.44}$$

with C being the background 3-form $\mathbb{Z}_2$ gauge field and Sq^2 denoting a Steenrod square. Gu-Wen equations in 3+1d can be interpreted as describing shadow theories of this sort, and it should be possible to use the methods of this chapter to produce more general SPT phases.

Optimistically, one might hope that every fermionic theory in every dimension has a bosonic shadow. Recent results of Brundan and Ellis [137] indicate that this is true in 2+1d. In particular, it would be very interesting to understand shadows of general spin-TFTs in 2+1d which have framing anomalies. This would require developing the theory of "super modular tensor categories."

Finally, we hope that the study of shadows of fermionic theories could shed light on the fermion doubling problem in lattice field theory.

4.3 Spherical fusion categories and fermions

The bosonic theory $\mathfrak{T}_f[\nu_3, n_2]$ we will associate to the Gu-Wen fermionic SPT phases belongs to the special class of TFTs which admit a Turaev-Viro state sum construction of the partition function [12] and a Levin-Wen string net construction of a microscopic lattice Hamiltonian [126].

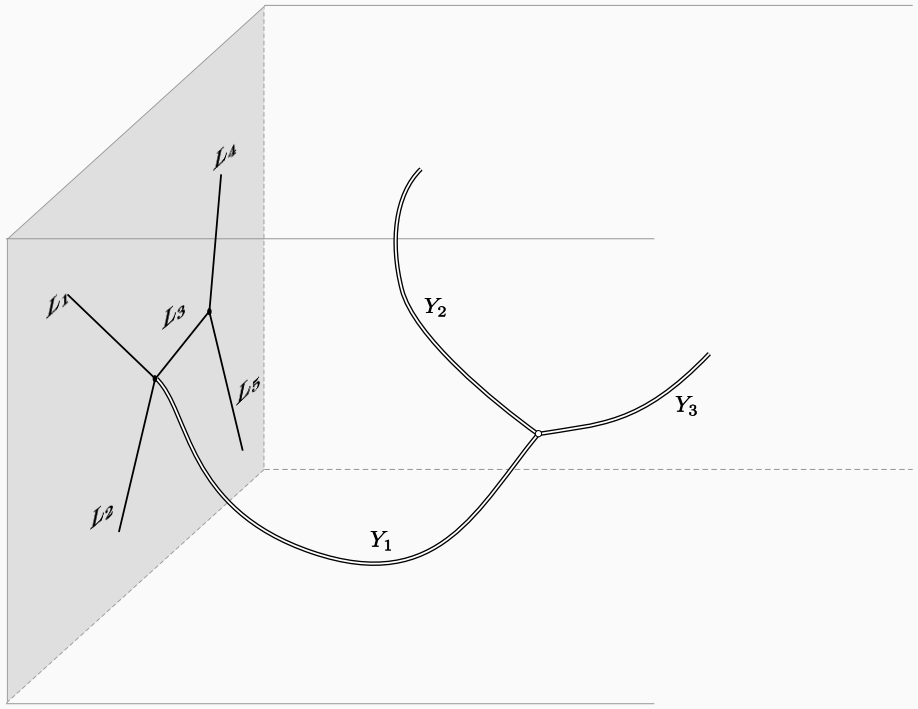


Figure 4.3.1: A topological field theory with a gapped boundary condition. Boundary lines are labelled by objects L_i in a spherical fusion category $\mathcal{C}$ which controls their topological fusion and junctions. Bulk lines are labelled by objects Y_a in a modular tensor category which can be recovered as the Drinfeld center $Z[\mathcal{C}]$ of the boundary lines. Junctions of lines are labelled by choices of local operators, i.e. elements in certain morphism spaces. We use a double-line notation to indicate the dependence of bulk lines on a choice of framing. The partition function can be computed by a Turaev-Viro state sum.

The Turaev-Viro construction allows one to define a large class of three-dimensional topological field theories. The mathematical input for the construction is a spherical fusion category $\mathcal{C}$. The output is the partition function of a topological field theory, whose quasi-particles are described by the modular tensor category $Z[\mathcal{C}]$, the Drinfeld center of $\mathcal{C}$.

The physical meaning of the mathematical input becomes manifest through the following observation: the Turaev-Viro construction produces topological field theories which admit a canonical topological boundary condition, which in turns supports topological line defects labelled by the objects in $\mathcal{C}$ [138].

This suggests the following physical statement: any (irreducible, unitary) three-dimensional topological field theory $\mathfrak{T}$ equipped with a topological boundary condition $\mathfrak{B}$ will admit a Turaev-Viro construction based on the category of topological line defects supported on $\mathfrak{B}$.

At first sight, it may appear surprising that the whole bulk topological field theory could be reconstructed from the properties of a single boundary condition. This is related to the cobordism

hypothesis [139]. There is a simple "swiss cheese" argument which demonstrates this fact in $2+1d$ and motivates the structure of the Turaev-Viro state sum model, which we review in a later section 4.4.

The same argument justifies the observation that several properties and enrichments of the bulk theory can be encoded in terms of the spherical fusion category $\mathcal{C}$. For example, if $\mathfrak{T}$ has a non-anomalous (0-form) symmetry group G then $\mathcal{C}$ will admit an extension to a G-graded category $\mathcal{C}_G$, which can be used to extend the Turaev-Viro construction to manifolds endowed with a G-valued flat connection [130].

With this motivation in mind, we can review some useful facts about spherical fusion categories and their physical interpretation.

4.3.1 Categories of boundary line defects

In the following we use the term topological field theory to denote the low energy/large distance effective field theory description of a gapped unitary quantum field theory. Similarly, a topological boundary condition is simply the low energy description of a gapped boundary condition.

The mathematical description of topological field theories involves a variety of operations which have an intuitive interpretation as a "fusion" of local operators or defects. The precise physical interpretation is that the local operators or defects to be fused are brought to relative distances which are still much larger than the gap, but smaller than the scale of the low energy effective field theory. This allows one to replace them by a single effective local operator or defect.

A gapped system may have multiple vacua, either due to spontaneous breaking of a symmetry or to accidental degeneracy. In the bulk theory, the presence of multiple vacua manifests itself in the existence of non-trivial local operators, whose expectation value labels different vacua. Mathematically, the local operators which survive at very low energy form a ring under the fusion operation described above (because of cluster decomposition). The identity operator can be decomposed into a sum of idempotents which project the system to a specific vacuum:

$$1 = \sum_v 1_v \qquad\qquad 1_v 1_{v'} = \delta_{v,v'} 1_v \tag{4.3.1}$$

The same idea applies to defects of lower co-dimension. As an example consider line defects, which could be the effective description of a quasi-particle or of a microscopic line defect. Line defects can be fused with each other and may support non-trivial local operators, including local operators which interpolate between two or more lines. Again, the existence of a local multiplicity of vacua for a line defect manifests itself in the existence of non-trivial idempotent local operators.

Mathematically, line defects can be organized into a fusion category. The objects in the category are the line defects themselves, and the morphisms are the local operators interpolating between two line defects. The physical fusion operation is encoded into a tensor product operation and accidental degeneracies into a sum operation. Line defects with no accidental degeneracy map to "simple" objects in the category.

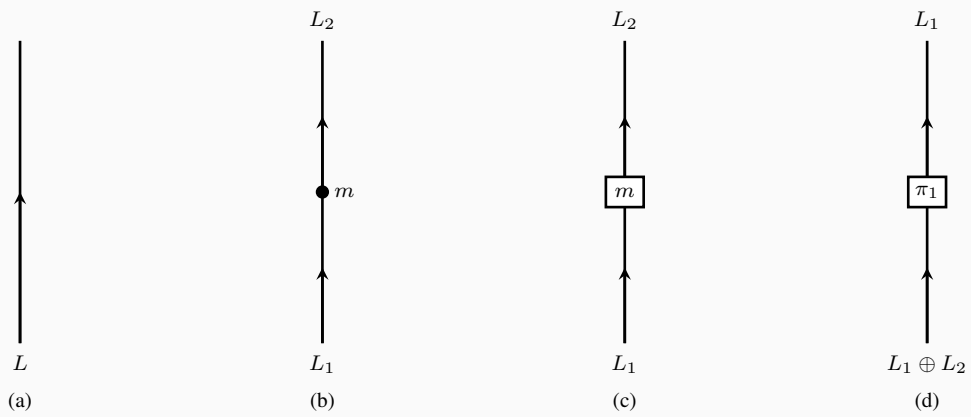


Figure 4.3.2: Data of a category $\mathcal{C}$: (a) A line defect (shown here with its orientation) is an object in $\mathcal{C}$. (b) A local operator between two lines is a morphism in $\mathcal{C}$. (c) Another representation of the previous figure (common in mathematics literature) in which morphisms are denoted by boxes. (d) The direct sum $L_1 \oplus L_2$ of two line defects can be projected to an individual summand by a local operator

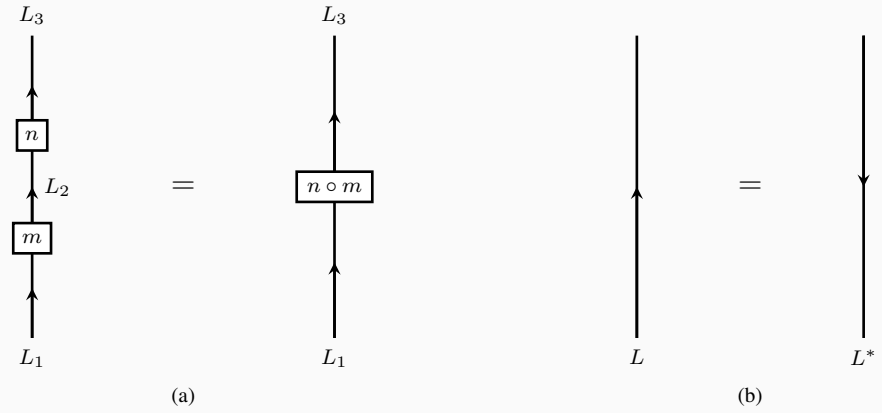


Figure 4.3.3: Various operations: (a) Fusion of local operators gives rise to composition of morphisms. (b) Changing the orientation of a line defect gives rise to the operation of taking *dual* of an object.

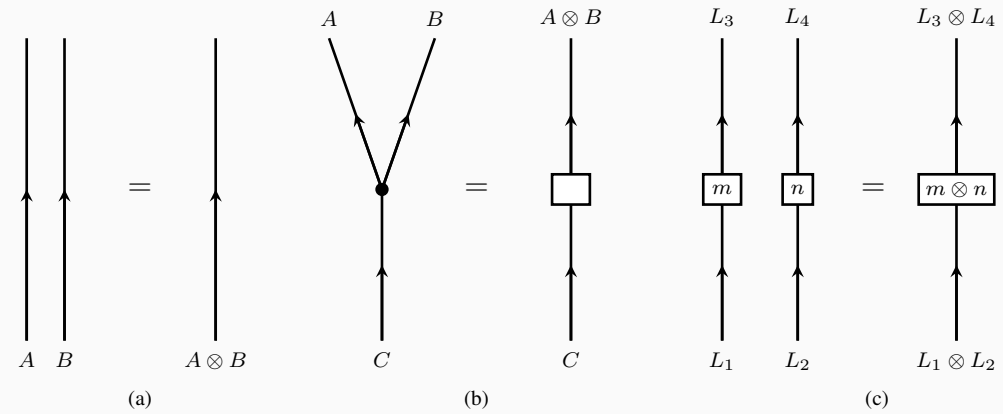


Figure 4.3.4: Fusion: (a) Fusion of line defects gives rise to the tensor product of objects. (b) Three line defects coming together with a local operator placed at the point of intersection can be interpreted as a morphism from one line defect to the tensor product of other two line defects. (c) Local operators between lines can also be fused to give rise to tensor product of morphisms.

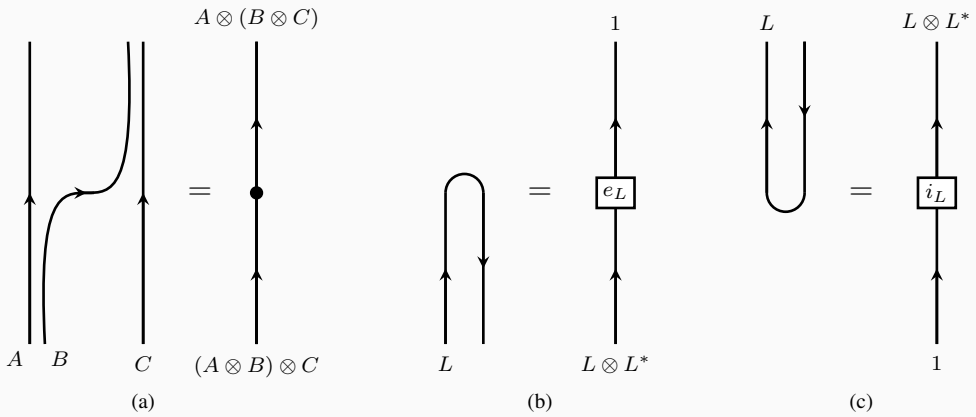


Figure 4.3.5: Canonical maps: (a) Placing the lines as shown and fusing them gives rise to a canonical *associator* map. (b) Folding a line as shown and fusing it with itself gives rise to a canonical *evaluation* map. (c) Folding a line as shown and fusing it with itself gives rise to a canonical *co-evaluation* map.

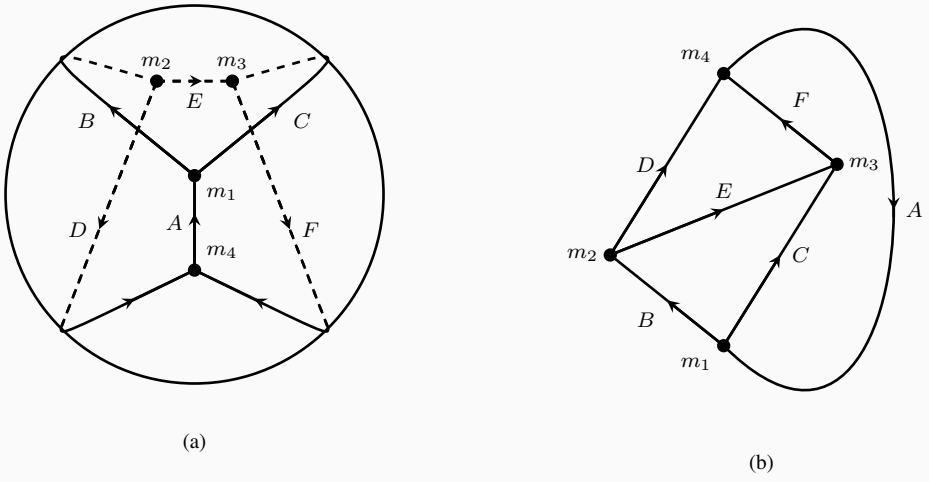


Figure 4.3.6: (a) A graph Γ of boundary line defects drawn on a sphere. (b) The same graph drawn on a plane obtained after removing a point from the sphere.

Depending on the dimension of space-time, the category of line defects will have further structures and constraints. Here we are interested in line defects which live on a gapped boundary condition. See Figures 4.3.2, 4.3.3, 4.3.4, 4.3.5 for examples.

If the boundary condition itself has a single vacuum, the boundary line defects are expected to form a spherical fusion category $\mathcal{C}$. The term spherical denotes a set of properties with a simple physical interpretation. Any graph Γ of line defects drawn on a two-sphere, with a specific choice of local operators at the vertices, will produce a state in the two-sphere Hilbert space of the bulk theory. As the latter space is one-dimensional, the graph will effectively evaluate to a number Z_Γ, which can be interpreted as the partition function of the theory for a three-ball decorated by Γ, normalized by the partition function of the bare three-ball. See Figure 4.3.6.

Mathematically, the graph is drawn on the plane as the evolution of a collection of lines, created, fused or annihilated at special points. The corresponding number is computed by Penrose calculus, as the composition of a sequence of maps associated to these individual processes, which form the data of the spherical fusion category. See Figure 4.3.7. The axioms of the spherical fusion category guarantee that the answer is independent of how we draw the graph. This evaluation map for graphs on the two-sphere is the basic ingredient in state sum constructions.

If we are given two topological field theories $\mathfrak{T}$ and $\mathfrak{T}'$, with gapped boundary conditions $\mathfrak{B}$ and $\mathfrak{B}'$ associated to spherical fusion categories $\mathcal{C}$ and $\mathcal{C}'$, the product of the two theories with the product boundary condition is associated to the product $\mathcal{C} \times \mathcal{C}'$ of the fusion categories.

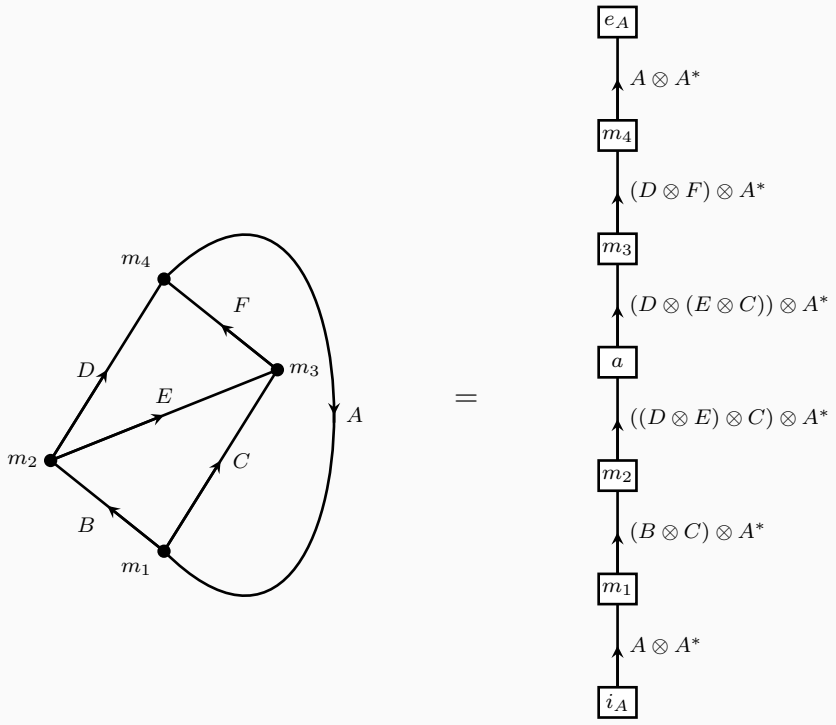


Figure 4.3.7: The computation of a graph on the plane involves the listed morphisms. Here a denotes the associator tensored with identity morphism for A^*. The final result is the partition function Z_Γ of the theory on a three-ball decorated by the graph Γ.

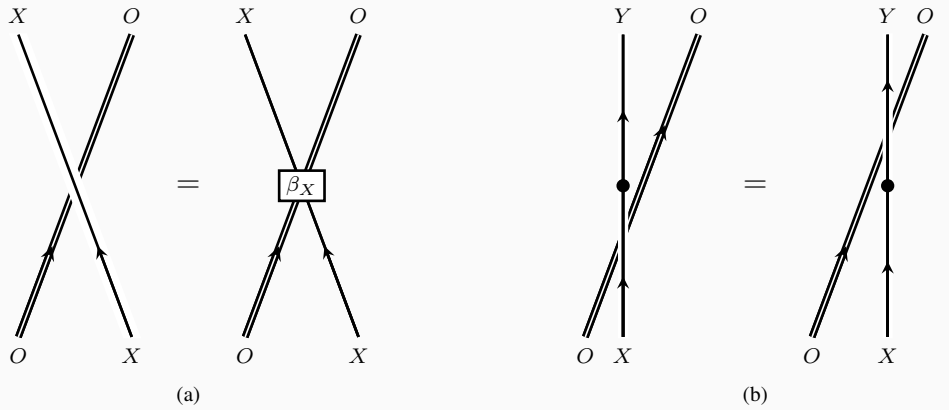


Figure 4.3.8: Bulk lines and Drinfeld Center: (a) Bringing a bulk line O to the boundary such that its image crosses a boundary line X gives rise to a canonical *half-braiding* given by morphism β_X. (b) Bringing O to the boundary in two different ways as shown in the figure is equivalent and hence β commutes with other morphisms.

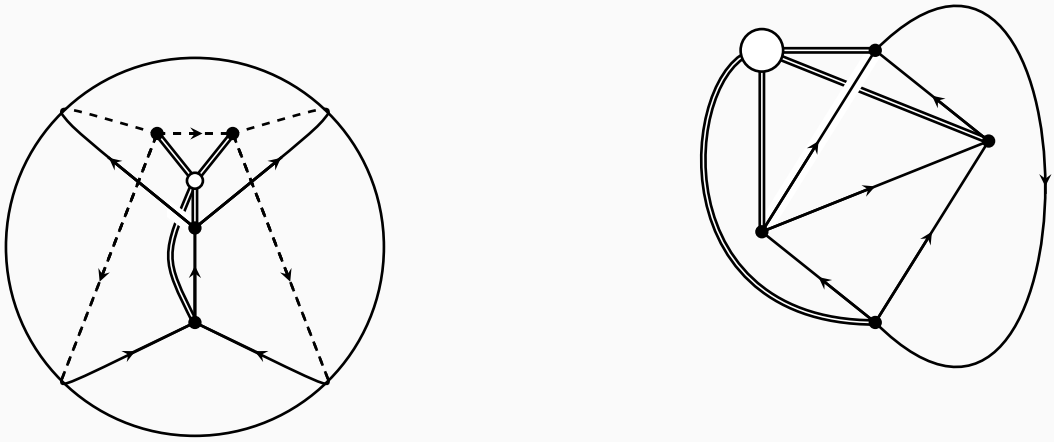
Bulk line defects can be fused with the boundary. If the boundary image crosses some pre-existing boundary line, the fusion produces some canonical local operator at the crossing. This physical process is encoded in the mathematical definition of Drinfeld center $Z[\mathcal{C}]$. An element of the center is a pair (O, β) of an object O in $\mathcal{C}$ together with a collection of crossing maps $\beta_X : O \otimes X \to X \otimes O$ for every other object X, satisfying certain axioms. See Figure 4.3.8.

These axioms have a simple interpretation. Consider a network of line defects in the three-ball, including boundary lines and bulk lines. If we project the network to a graph Γ on the boundary and evaluate Z_Γ, the answer will not depend on the choice of projection. See Figure 4.3.9. Every bulk line will thus map to an element of the center $Z[\mathcal{C}]$. Conversely, the Turaev-Viro construction gives an explicit definition of a bulk line defect for every element of the center $Z[\mathcal{C}]$. [6]

In particular, we can recognize the generators of bulk 1-form symmetries as special elements of the center. For example, a spherical fusion category $\mathcal{C}_b$ represents a bulk theory equipped with a bosonic $\mathbb{Z}_2$ one-form symmetry if we can find a generator $B = (b, \beta)$, an element of the center $Z[\mathcal{C}]$ such that $\beta_b = 1_{b\otimes b}$ and such that there is an isomorphism $\xi_b : b \otimes b \to I$ with $\xi \otimes 1 = 1 \otimes \xi$ in $\mathrm{Hom}(b \otimes b \otimes b, b)$. Essentially, this means that B lines fuse to the identity and can be freely re-connected in pairs. See Figure 4.3.10.

Similarly, a spherical fusion category $\mathcal{C}_f$ represents a bulk theory equipped with a fermionic $\mathbb{Z}_2$ one-form symmetry if we can find a generator $\Pi = (f, \beta)$, an element of the center $Z[\mathcal{C}]$ such that $\beta_f = -1_{f\otimes f}$ and such that there is an isomorphism $\xi_f : f \otimes f \to I$ with $\xi \otimes 1 = 1 \otimes \xi$ in $\mathrm{Hom}(f \otimes f \otimes f, f)$. Essentially, this means that f lines fuse to the identity and can be freely re-connected in pairs, at the price of a -1 sign for each crossing. See again Figure 4.3.10.

[6] From the point of view of the bulk theory, a gapped boundary condition can be characterized in terms the set of bulk lines which "condense" at the boundary, i.e. project to the trivial line on the boundary. They are a collection of mutually local bosons which is closed under fusion.



(a) (b)

Figure 4.3.9: (a) A three-ball partition function decorated by a graph Γ of bulk and boundary lines. (b) The graph is projected onto the sphere for evaluation. The different projections evaluate to the same result, thanks to the Drinfeld center axioms

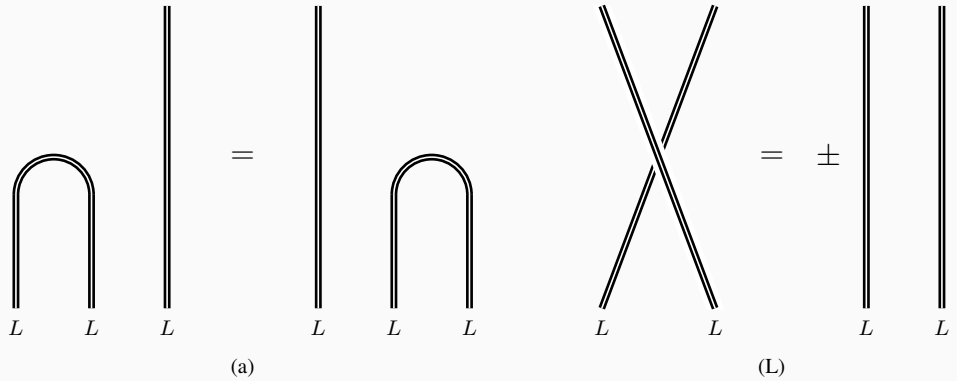


Figure 4.3.10: $\mathbb{Z}_2$ 1-form symmetries: (a) There exists a bulk line L with properties shown in the figure. (b) Half-braiding L lines across each other gives a factor of ± 1 when compared to L lines without braiding. The factor of $+1$ arises for a bosonic 1-form symmetry generator $L \equiv B$ and -1 arises for a fermionic 1-form symmetry generator $L \equiv \Pi$. This minus sign implies that the symmetry is anomalous.

More generally, a monoidal category equipped with such a Π is called a "monoidal Π-category" in [137].

A couple variants to this setup may be useful. If the boundary condition has some accidental degeneracy, we should consider a spherical multi-fusion category. Local operators on the boundary are morphisms from the trivial line defect to itself, which is thus not simple. The category $\mathcal{C}$ splits into multiple sub-categories $\mathcal{C}_{a,b}$ representing line defects which interpolate between vacua a and b. The objects in these categories fuse accordingly:

$$\mathcal{C}_{a,b} \otimes \mathcal{C}_{c,d} \in \delta_{b,c} \mathcal{C}_{a,d} \tag{4.3.2}$$

If the bulk theory and boundary condition have a non-anomalous discrete global symmetry G (possibly broken at the boundary), we will have a G-graded spherical fusion category (see e.g. [140]), with sub-categories $\mathcal{C}_g$ which fuse according to the group law:

$$\mathcal{C}_g \otimes \mathcal{C}_{g'} \in \mathcal{C}_{gg'} \tag{4.3.3}$$

The sector $\mathcal{C}_e$ labelled by the group identity e consists of standard boundary line defects while the other $\mathcal{C}_g$ contain the boundary version of g-twist line defects.

Note that we can define a G-graded product of G-graded spherical fusion categories by letting $(\mathcal{C} \times \mathcal{C}')_g \equiv \mathcal{C}_g \times \mathcal{C}'_g$. Physically, this corresponds to taking the direct product of two theories $\mathfrak{T}$ and $\mathfrak{T}'$ and of their corresponding boundary conditions $\mathfrak{B}$ and $\mathfrak{B}'$.

If we gauge the symmetry G (with Dirichlet boundary conditions for the gauge connection), all objects in $\mathcal{C}$ become true boundary line defects. Bulk line defects are now associated to the center of the whole $\mathcal{C}$. The center of $\mathcal{C}$ includes Wilson loops, of the form $(I^{\oplus n}, \beta_{X_g} = T_g)$, X_g is an arbitrary simple object of $\mathcal{C}_g$ and the matrices T_g define an n-dimensional representation of G.

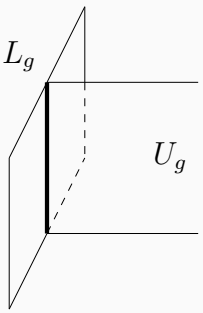


Figure 4.3.11: Lines L_g living at the intersection of a 0-form symmetry generator U_g and the boundary form the sub-category $\mathcal{C}_g$.

[7] If G is abelian, the Wilson loops are labelled by characters in the dual group G^* and generate a non-anomalous G^* 1-form symmetry.

We can also gauge a subgroup H of G. The resulting H gauge theory should have a residual global symmetry given by the quotient $G_H = N_G(H)/H$ of the normalizer of H by H. The corresponding G_H-graded category consists of

$$\mathcal{C}_{[g]} = \bigcup_{h\in H} \mathcal{C}_{hg} \tag{4.3.4}$$

Later in the chapter, we will find it useful to build some interesting G-graded categories starting from SPT phases for a central extension $\widehat{G}$ of G by an Abelian group and gauging the Abelian group as described above.

Although a 1-form symmetry generator B (or Π) for a G-graded theory is defined as a special element in $Z[\mathcal{C}_e]$, we will often be interested in 1-form symmetries which are compatible with turning on G-flat connections or even gauging G. We will see that this is the case if B (or Π) admits a lift to $Z[\mathcal{C}]$. The lift may not be unique and different lifts can be thought of as different ways to equip the theory with both G symmetry and 1-form $\mathbb{Z}_2$ symmetry.

4.3.2 Example: toric code

The simplest example of a category of boundary line defects occurs in the toric code, also known as topological $\mathbb{Z}_2$ gauge theory in 2+1d. Recall that the toric code has four quasi-particles, corresponding in the gauge theory to a trivial defect 1, a Wilson loop e, a flux line m and the fusion ϵ of the latter two. This topological field theory can be endowed with a $\mathbb{Z}_2$ global symmetry exchanging the e and m lines, which will be very important later on but which we ignore now.

The e and m lines are bosons, while ϵ is a fermion. Indeed, e generates a non-anomalous $\mathbb{Z}_2^e$ one-form symmetry and in the language of the introduction the toric code is the partner $\mathfrak{T}_b$ of a

[7] We are identifying here $\mathrm{Hom}_{\mathcal{C}}(I^{\oplus n}\otimes X_g, X_g\otimes I^{\oplus n})$ with $n\times n$ matrices.

trivial $\mathfrak{T}_{\mathbb{Z}_2}$. Symmetrically, m also generates a non-anomalous $\mathbb{Z}_2^m$ one-form symmetry (with a mixed anomaly with the $\mathbb{Z}_2^e$ symmetry).

On the other hand, ϵ generates precisely the sort of anomalous one-form $\mathbb{Z}_2^\epsilon$ symmetry we need for the shadow of a spin TFT. This will be an important example for us, especially after we make manifest the $\mathbb{Z}_2$ global symmetry exchanging e and m.

A $\mathbb{Z}_2$ gauge theory has two natural gapped boundary conditions: we can fix the flat connection at the boundary or let it free to fluctuate. The corresponding boundary conditions in the toric code, $\mathfrak{B}_e$ and $\mathfrak{B}_m$, condense either the e or the m particle. [8]

In either case, the category of boundary line defects consists of two simple objects, I and P, which fuse as

$$I \otimes I = I \qquad I \otimes P = P \qquad P \otimes I = P \qquad P \otimes P = I \tag{4.3.5}$$

All the associators and other data can be taken to be trivial.

The four elements in the center, say for $\mathfrak{B}_e$, can be described as

$$\begin{aligned} 1 &= (I; \beta_I = 1, \beta_P = 1) \\ e &= (I; \beta_I = 1, \beta_P = -1) \\ m &= (P; \beta_I = 1, \beta_P = 1) \\ \epsilon &= (P; \beta_I = 1, \beta_P = -1) \end{aligned} \tag{4.3.6}$$

We recognize the required properties for generators of bosonic or fermionic 1-form symmetries.

The toric code also offers a very simple example of gauging a $\mathbb{Z}_2$ symmetry at the level of spherical fusion categories: the trivial $\mathbb{Z}_2$ SPT phase is associated to a $\mathbb{Z}_2$-graded spherical fusion category, with $\mathcal{C}_0$ consisting of the identity object I and $\mathcal{C}_1$ consisting of P. Dropping the grading gives us the $\mathbb{Z}_2$ gauge theory/toric code.

4.3.3 Example: bosonic SPT phases and group cohomology

The group cohomology construction of bosonic SPT phases has precisely the form of a G-graded Turaev-Viro partition sum, based on a G-graded category $\mathcal{C}$ with a single (equivalence class of) simple object V_g in each $\mathcal{C}_g$ subcategory.

The associator is a map from $V_{gg'g''}$ to itself, which can be written as

$$e^{2\pi i \alpha_3(1,g,gg',gg'g'')} 1_{V_{gg'g''}}$$

where α_3 is a 3-cocycle on BG with values in $U(1)$. The cocycle condition is equivalent to the pentagon axiom for the associator. Re-definitions of the isomorphisms $V_g \otimes V_{g'} \simeq V_{gg'}$ used in the definition will shift α_3 by an exact cochain.

[8] In appendix C.3 we describe a fermionic boundary condition $\mathfrak{B}_\epsilon$ at which ϵ condenses.

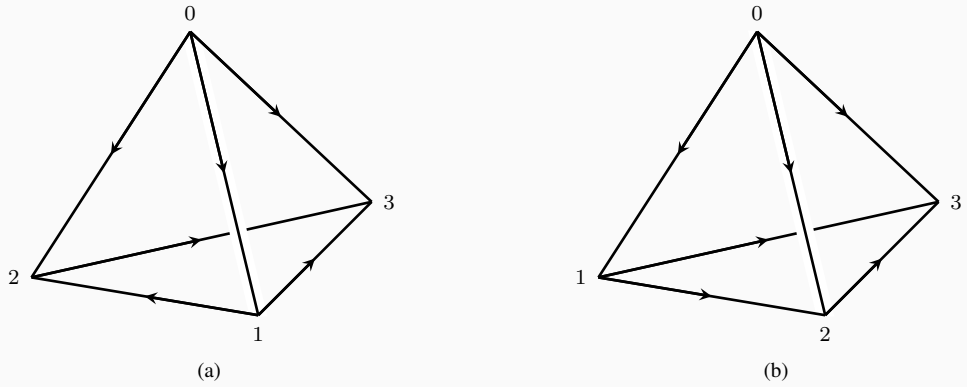


Figure 4.3.12: Given a tetrahedron with a labeling of vertices by $i \in \{0, 1, 2, 3\}$, we orient the edges such that vertex with label i has i incoming edges. This defines a *local order* on the tetrahedron. Orientation is defined by using right-hand rule going from 0 to 1 to 2. If the thumb points inward, we say that the tetrahedron is positively oriented as shown in (a). If the thumb points outward, we say that the tetrahedron is negatively oriented as shown in (b).

We refer the reader to Figure 4.3.13 for a graphical explanation of the relation between associators and cocycle elements. An illustrative example is the non-trivial group cocycle for $G = \mathbb{Z}_2$:

$$\alpha_3(0, \epsilon, \epsilon + \epsilon', \epsilon + \epsilon' + \epsilon'') = \frac{1}{2}\epsilon\epsilon'\epsilon'' \tag{4.3.7}$$

In terms of the cocycle ϵ_1 defined by the group element on edges of the tetrahedron, $\alpha_3^{\mathbb{Z}_2} = \frac{1}{2}\epsilon_1 \cup \epsilon_1 \cup \epsilon_1$. [9]

We can describe the corresponding G gauge theory simply by ignoring the G grading on $\mathcal{C}$. For future reference, it is useful to describe objects in the Drinfeld center of $\mathcal{C}$. The bulk defect lines (i.e. simple objects of the Drinfeld center) turn out to be labelled by a pair (g, χ), where g is an element of G and χ an irreducible projective representation of the stabilizer G_g of g in G [141].

The pair (g, χ) gives a center line of the form $(V_g^{\oplus n}, \beta_{g'} = \chi(g'))$. Notice that β only needs to be specified if g and g' commute, in which case it is a matrix multiple of the basis element of $\mathrm{Hom}_{\mathcal{C}}(V_g \otimes V_{g'}, V_{g'} \otimes V_g) \simeq \mathbb{C}$. The definition of the Drinfeld center requires

$$\chi(g'g'') = e^{2\pi i\alpha_3(1,g,gg',gg'g'') - 2\pi i\alpha_3(1,g',g'g,g'gg'') + 2\pi i\alpha_3(1,g',g'g'',g'g''g)}\chi(g')\chi(g'') \tag{4.3.8}$$

and fixes the group 2-cocycle associated to the projective representation in terms of α_3 and g. Physically, this is a g-twist line dressed by a Wilson line.

[9] An alternative expression for the cocycle can be given via the Bockstein homomorphism: $\epsilon_1 \cup \epsilon_1$ is equivalent modulo 2 to $\frac{1}{2}\delta\tilde{\epsilon}_1$, where $\tilde{\epsilon}_1$ is an integral lift of ϵ_1. Thus we can write $\alpha_3^{\mathbb{Z}_2} = \frac{1}{4}\tilde{\epsilon}_1 \cup \delta\tilde{\epsilon}_1$.

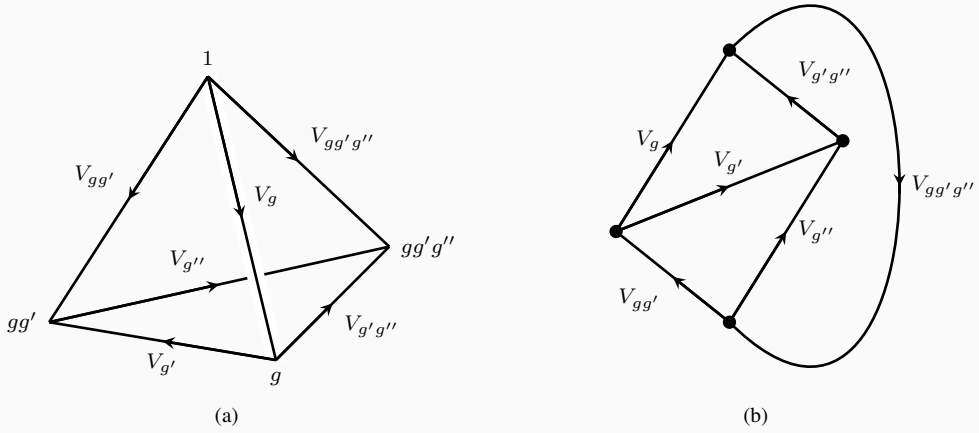


Figure 4.3.13: Bosonic SPT phases: (a) A positively oriented tetrahedron with generic simple elements on edges. We label the vertices such that an edge going from g to h is assigned the element $V_{g^{-1}h}$. (b) The planar graph dual to the tetrahedron with all the morphisms as canonical identity morphisms. The graph evaluates to the associator $\alpha_3(1, g, gg', gg'g'')$. Notice that the faces of the dual graph correspond to vertices of the tetrahedron and are ordered correspondingly. Edges are oriented so that the face to the left is comes before the face to the right. For example, the outer face in the dual graph is the first.

4.3.4 Example: G-equivariant $\mathbb{Z}_2$ gauge theory from a central extension

Consider a central extension

$$0 \to \mathbb{Z}_2 \to \widehat{G} \to G \to 0 \tag{4.3.9}$$

We can take a $\widehat{G}$ SPT phase and gauge the $\mathbb{Z}_2$ subgroup.

The result is a G-graded category with $\mathcal{C}_g$ consisting of two objects. If we denote the pre-images of g in $\widehat{G}$ as $(g,0)$ and $(g,1)$, then $\mathcal{C}_g$ consists of $V_{g,0}$ and $V_{g,1}$. The fusion rule is given by

$$V_{g,\epsilon} \otimes V_{g',\epsilon'} \simeq V_{gg',\epsilon+\epsilon'+n_2(g,g')}, \tag{4.3.10}$$

where n_2 is the $\mathbb{Z}_2$-valued group 2-cocycle corresponding to the central extension.

We can now ask if the $\mathbb{Z}_2$ gauge theory has $\mathbb{Z}_2$ 1-form symmetry generators which are compatible with the G global symmetry, i.e. map each $\mathcal{C}_g$ to itself. That means we should look for objects of the center $Z[\mathcal{C}]$ which project to either $V_{e,0}$ or $V_{e,1}$. The former case corresponds to the bare Wilson loop, which generates a bosonic $\mathbb{Z}_2$ 1-form symmetry.

The latter case is more interesting, as the 2-cocycle for $(e,1)$-twist lines may be non-trivial. A $(1,1)$-twist line will be a bosonic (fermionic) $\mathbb{Z}_2$ generator if we can find a 1-dimensional projective representation of $\widehat{G}$ with appropriate cocycle and $\chi((e,1)) = \pm 1$.

This is a somewhat intricate constraint on the $\widehat{G}$ 3-cocycle $\widehat{\alpha}_3$ defining the initial SPT phase. Up to a gauge transformation, this constraint has a neat solution: $\widehat{\alpha}_3$ must be given in terms of a group super-cohomology element (ν_3, n_2) as follows:

$$\widehat{\alpha}_3 = \nu_3 + \frac{1}{2} n_2 \cup \epsilon_1. \tag{4.3.11}$$

Here ν_3 is an $\mathbb{R}/\mathbb{Z}$-valued 3-cochain on BG satisfying the Gu-Wen equation (4.2.16), and where ϵ_1 is the $\mathbb{Z}_2$-valued 1-cochain which sends (g,ϵ) to ϵ. It is easy to see that $\delta\epsilon_1 = n_2$, and thus the cocycle condition $\delta\widehat{\alpha}_3 = 0$ follows from the Gu-Wen equations. The fermion Π corresponds to the projective representation $\chi((g,\epsilon)) = (-1)^\epsilon$.

Of course, the form given here for $\widehat{\alpha}_3$ can be modified by gauge transformations. For example, a transformation with parameter $\frac{1}{2}\epsilon_1 \cup_1 n_2$ would give another representative:

$$\widehat{\alpha}'_3 = \nu'_3 + \frac{1}{2} \epsilon_1 \cup n_2. \tag{4.3.12}$$

with $\nu'_3 = \nu_3 + \frac{1}{2} n_2 \cup_1 n_2$.

There are two complementary ways to arrive at this solution. In Appendix C.2 we give a derivation based on the analysis of anomalies in the $\mathbb{Z}_2$ gauge theory coupled to a G gauge field. In Figures 4.3.14 and 4.3.15 we give a graphical/physical proof of 4.3.12 using the spherical fusion category associated to $\widehat{G}$. Essentially, the existence of a Drinfeld center element of the form $(V_{e,1}; \beta)$ allows certain topological manipulations of planar graphs, relating two graph which encode the left and right side of equation 4.3.11.

In particular, we can define ν_3 in terms of the spherical fusion category data as a tetrahedron graph of $(g,0)$ lines, with n_2 extra fermion lines at each vertex, exiting from the earliest face

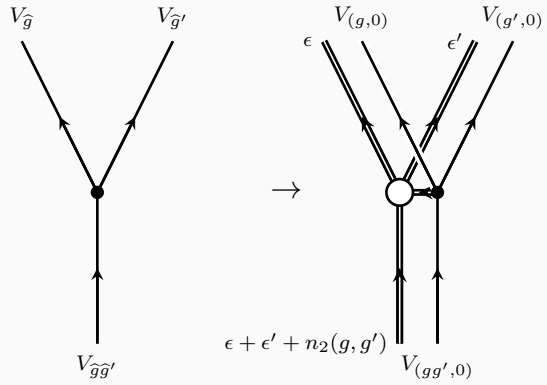


Figure 4.3.14: Gauge-fixing: A graphical representation of the partial gauge-fixing procedure used in computing the $\widehat{G}$ group cocycle. Left: A choice of gauge is the same as a choice of basis vector in the space of junctions between line defects in the full $\widehat{G}$ category. Right: we identify $V_{(g,\epsilon)} \simeq V_{(1,\epsilon)} \otimes V_{g,0}$ and identify $V_{(1,\epsilon)}$ with the corresponding elements I or Π of the center. We then express a general junction canonically in terms of a choice of junction between line defects labelled by G elements. The double lines denote the center elements. The empty circle represents any choice of how to connect the center lines in a planar way.

around the vertex and coming together to a common point where they are connected in a planar manner, as in Figure 4.3.15 (b).

In conclusion, we have a bijection between Gu-Wen fermionic SPT phases and potential shadows of G-symmetric spin-TFTs based on a $\mathbb{Z}_2$ theory.

Notice that the pair (ν_3, n_2) labels both the spherical fusion category and the choice of fermionic line, i.e. it labels the Π-category. The same spherical fusion category may admit multiple candidate fermionic lines. For example, if we are given a group homomorphism λ_1 from G to $\mathbb{Z}_2$, we can dress Π by a Wilson line for the corresponding representation, i.e. add a $(-1)^{\lambda_1}$ to β. Then the same choice of $\widehat{\alpha}_3$ will give a ν_3 which differs from the original by $\lambda_1 \cup n_2$.

As an example of the construction, consider $\widehat{G} = \mathbb{Z}_4$ as a $\mathbb{Z}_2$ central extension of $G = \mathbb{Z}_2$. Recall that $H^3(\mathbb{Z}_4, \mathbb{R}/\mathbb{Z}) = \mathbb{Z}_4$. We claim that the generator of this group corresponds to a shadow of a Gu-Wen fermionic SPT. Indeed, if $[\eta_1]$ is the generator of $H^1(G, \mathbb{Z}_2) = \mathbb{Z}_2$, then the extension class corresponding to $\widehat{G}$ can be written as $n_2 = [\frac{1}{2}\delta\tilde{\eta}_1]$, where $\tilde{\eta}_1$ is an integral lift of η_1. Concretely, η_1 is the $\mathbb{Z}_2$ cocycle defined by the G elements on the edges of the triangulation and $[\frac{1}{2}\delta\tilde{\eta}_1]$ measures the failure of the group law for a $\widehat{G}$ lift of the G elements.

Therefore a possible solution of the equation (4.2.16) is

$$\nu_3 = \frac{1}{8}\tilde{\eta}_1 \cup \delta\tilde{\eta}_1. \tag{4.3.13}$$

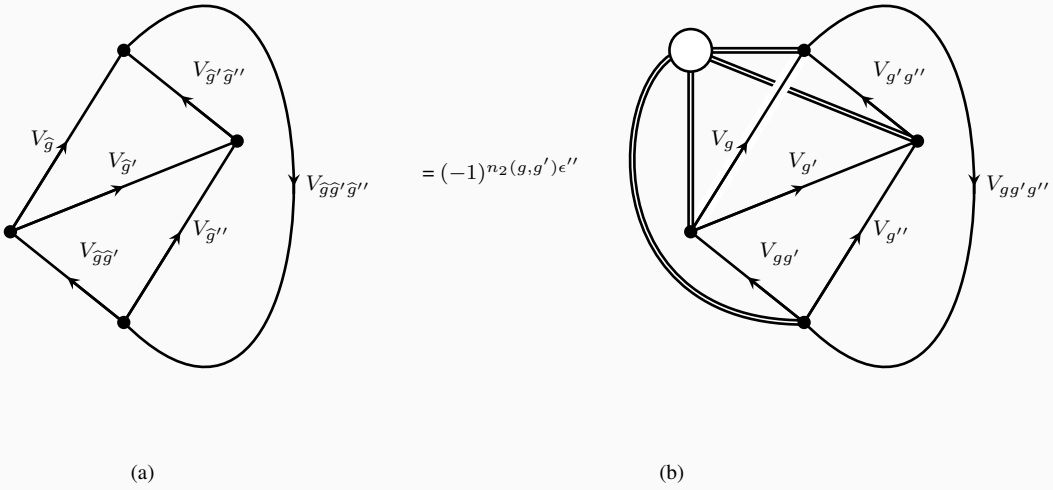


Figure 4.3.15: Gu-Wen fermionic SPT phases: (a) The planar graph dual to the tetrahedron which computes $\widehat{\alpha}$ in a gauge determined by the choice of morphism at the junctions. We partial gauge-fix as in the previous Figure. The resulting web of center lines can be simplified by bringing together all planar junctions and collapsing planar loops, up to resolving a single crossing (See next Figure 5.4.2). Up to the corresponding sign, we obtain: (b) a graph which depends on G elements only and defines ν_3.

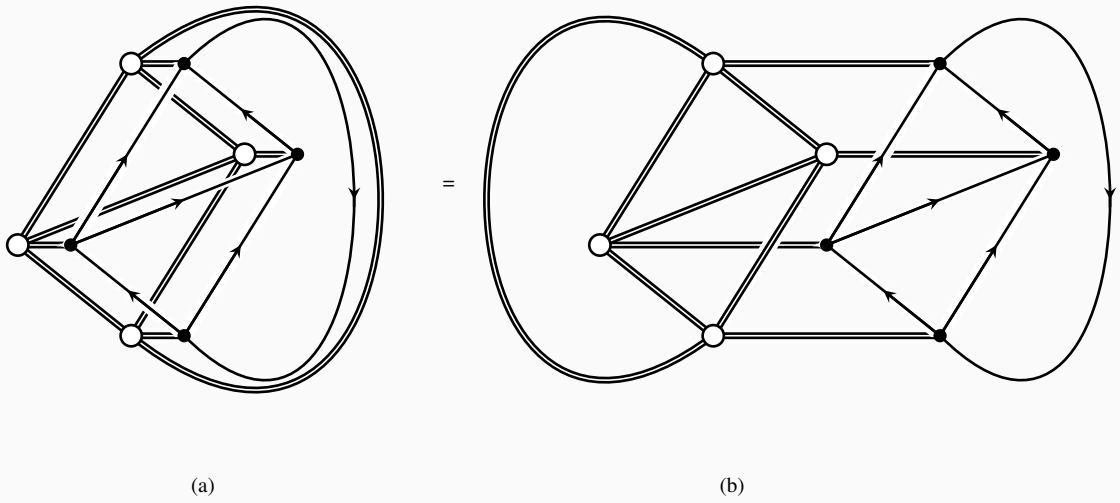


Figure 4.3.16: Intermediate computational steps relating $\widehat{\alpha}_3$ and ν_3. The planar intersections (white circles) of center lines can be collapsed together safely, but the non-planar intersection has to be resolved first, at the price of a sign $(-1)^{n_2(g,g')\epsilon''}$.

The corresponding 3-cocycle on $\widehat{G}$ is

$$\widehat{\alpha}_3 = \frac{1}{8}\tilde{\eta}_1 \cup \delta\tilde{\eta}_1 + \frac{1}{4}\delta\tilde{\eta}_1 \cup \epsilon_1. \tag{4.3.14}$$

Twice this cocycle is $\frac{1}{4}\tilde{\eta}_1 \cup \delta\tilde{\eta}_1 \sim \frac{1}{2}\epsilon_1^3$, which is a pull-back of a 3-cocycle on $G = \mathbb{Z}_2$ generating $H^3(\mathbb{Z}_2, \mathbb{Z}_2) \simeq \mathbb{Z}_2$. Therefore this cocycle represents the generator of $H^4(\widehat{G}, \mathbb{R}/\mathbb{Z})$. [10] This is the shadow of a Gu-Wen phase with symmetry $\mathbb{Z}_2$. It is an abelian phase, in the sense that the fusion rules of the shadow TFT are abelian (based on an abelian group $\mathbb{Z}_4$).

Another solution of the Gu-Wen equations with the same n_2 is

$$\nu_3 = -\frac{1}{8}\tilde{\eta} \cup \delta\tilde{\eta}. \tag{4.3.16}$$

It differs from (4.3.13) by a closed 3-cochain $\frac{1}{4}\tilde{\eta}\cup\delta\tilde{\eta}$ whose class is the generator of $H^3(G, \mathbb{R}/\mathbb{Z}) = \mathbb{Z}_2$. In physical terms, these two Gu-Wen phases (and their shadows) differ by tensoring with a bosonic SPT phase. Two more shadows of Gu-Wen phases are obtained by taking $\widehat{G} = \mathbb{Z}_2 \times \mathbb{Z}_2$. In this case $\widehat{\alpha}_3$ is a pull-back of a 3-cocycle on $G = \mathbb{Z}_2$, which is otherwise unconstrained. Overall, we get four Gu-Wen phases with symmetry $G = \mathbb{Z}_2$. They are all abelian phases and are naturally labeled by elements of $\mathbb{Z}_4$.

4.3.5 Example: $\mathbb{Z}_2$-equivariant toric code vs Ising

The toric code has a $\mathbb{Z}_2$ symmetry which exchanges e and m, which is not manifest as an on-site symmetry in the standard microscopic formulation of the theory.

The symmetry can be made manifest by extending the category of boundary line defects to a $\mathbb{Z}_2$-graded category which includes boundary twist lines for the $\mathbb{Z}_2$ symmetry and using the extended category as an input for a state sum or a string-net model.

As the $\mathbb{Z}_2$ symmetry exchanges the $\mathfrak{B}_e$ and $\mathfrak{B}_m$ boundary conditions, the boundary twist lines interpolate between $\mathfrak{B}_e$ and $\mathfrak{B}_m$.

Concretely the $\mathbb{Z}_2$-graded category can be identified with the Ising fusion category (see [142] or appendix B of [143] for a detailed discussion). There are three objects I, S, P fusing as $P \otimes S = S \otimes P = S$ and $S \otimes S = I \oplus P$. The object S belongs to $\mathcal{C}_1$, I and P to $\mathcal{C}_0$. The nontrivial associators are

$$a(P,S,P): \ (P \otimes S) \otimes P \to P \otimes (S \otimes P), \tag{4.3.17}$$
$$a(S,P,S): \ (S \otimes P) \otimes S \to S \otimes (P \otimes S), \tag{4.3.18}$$
$$a(S,S,S): \ (S \otimes S) \otimes S \to S \otimes (S \otimes S). \tag{4.3.19}$$

[10] Alternatively, we can re-write it directly in terms of the $\mathbb{Z}_4$ cocycle $\epsilon_1^{\mathbb{Z}_4} \equiv \tilde{\eta}_1 + 2\epsilon_1$. It is easy to verify that $\widehat{\alpha}_3$ is co-homologous to

$$\frac{1}{4}\epsilon_1^{\mathbb{Z}_4} \cup \epsilon_1^{\mathbb{Z}_4} \cup \epsilon_1^{\mathbb{Z}_4} = \frac{1}{4}\tilde{\eta}_1 \cup \tilde{\eta}_1 \cup \tilde{\eta}_1 + \frac{1}{2}\tilde{\eta}_1 \cup \tilde{\eta}_1 \cup \epsilon_1 + \frac{1}{2}(\epsilon_1 \cup \tilde{\eta}_1 + \tilde{\eta}_1 \cup \epsilon_1) \cup \tilde{\eta}_1, \tag{4.3.15}$$

modulo 1.

The first one, regarded as an endomorphism of S, is -1. The second one, regarded as an endomorphism of $I \oplus P$, is a vector $(1, -1)$. The last associator is determined by the pentagon equation only up to an overall sign: the associator morphism regarded as an endomorphism of $S \oplus S$ is a matrix

$$\lambda^{-1} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{4.3.20}$$

where $\lambda = \pm\sqrt{2}$.

The fusion rules can be explained as follows. The fusion rules for $\mathcal{C}_1$ are the usual fusion rules for the boundary lines on the $\mathfrak{B}_e$ boundary. Since S is the termination of a $\mathbb{Z}_2$ domain wall which implements the particle-vortex symmetry transformation, we must have $S \otimes S \supset I$.: this means that a domain wall shaped as a hemisphere ending on a $\mathfrak{B}_e$ boundary can be shrunk away. Finally, shrinking away the same hemispherical domain wall in the presence of a Wilson line P shows that $S \otimes S \supset P$. The associators are fixed by the pentagon equation, up to an ambiguity in the sign of λ [142].

This identification of the Ising category with the $\mathbb{Z}_2$ equivariant version of the toric code is consistent with the observation that gauging the $\mathbb{Z}_2$ symmetry of the toric code produces the quantum double of the 3d Ising TFT, i.e. a TFT whose category of bulk like defects is the product of the Ising modular tensor category and its conjugate.

The Ising modular tensor category has three simple objects $1, \sigma, \psi$ which fuse just as I, S, P above. The quantum double (i.e. the Drinfeld center of the Ising fusion category) has bulk quasi-particles which are the product of $1, \sigma, \psi$ and $1, \overline{\sigma}, \overline{\psi}$. The $\psi\overline{\psi}$ particle is a boson to be identified with the Wilson loop. The ψ and $\overline{\psi}$ fermions are two versions of the original ϵ particle. Thus, for a fixed λ, there is a two-fold ambiguity in the choice of the fermion Π for the Ising fusion category. More precisely, crossing either ψ or $\overline{\psi}$ with P gived -1, while crossing a fermion with S gives a phase ξ^2 satisfying [143]

$$\xi + \xi^{-1} = \lambda. \tag{4.3.21}$$

The two solutions of this equation correspond to taking $\Pi = \psi$ or $\Pi = \overline{\psi}$. It is easy to see that $\xi^4 = -1$, so taking into account both the freedom in choosing λ and the freedom in choosing Π we get four $\mathbb{Z}_2$-equivariant versions of the toric code with a fermionic $\mathbb{Z}_2$ 1-form symmetry. They can be labeled by ξ, which is a fourth root of -1. The four versions of the theory are on equal footing, since none of the four roots is preferred.

Recall that fermionic SPT phases with a unitary $\mathbb{Z}_2$ symmetry have a $\mathbb{Z}_8$ classification [**?**, **?**, 144]. Four of them correspond to Gu-Wen supercohomology phases. We will argue below that the shadows of the other four phases are given by the four versions of the Ising fusion category equipped with Π. The latter phases are non-abelian, in the sense that the fusion rules of the shadow TFT are not group-like.

4.3.6 Example: Ising pull-backs

If we are given a group G with a group homomorphism $\pi_1 : G \to \mathbb{Z}_2$, we can define a G-graded Ising-like category as follows.

If $\pi_1(g) = 0$, we take $\mathcal{C}_g$ to consist of two simple elements, I_g and P_g. If $\pi_1(g) = 1$, we take $\mathcal{C}_g$ to consist of a simple element S_g. We take the fusion rules to mimic the Ising category:

$$V_{g,\epsilon} \otimes V_{g,\epsilon'} = V_{gg',\epsilon+\epsilon'}, \tag{4.3.22}$$
$$V_{g,\epsilon} \otimes S_{g'} = S_{gg'} \tag{4.3.23}$$
$$S_g \otimes V_{g',\epsilon'} = S_{gg'}, \tag{4.3.24}$$
$$S_g S_{g'} = V_{gg',0} + V_{gg',1}, \tag{4.3.25}$$

where we denoted $I_g = V_{g,0}$ and $P_g = V_{g,1}$. The associators can be taken from the Ising category.

The center particle with boundary image P_1 and β taken from the fermion in the Ising category example equips this category with a fermionic 1-form symmetry. We will call the corresponding G-equivariant TFT an Ising pull-back and denote it $\mathcal{I}^{\xi}_{\pi_1}$. It depends on a parameter ξ satisfying $\xi^4 = -1$ as well as $\pi_1 : G \to \mathbb{Z}_2$. We will see below that it is a shadow of a fermionic SPT phase with symmetry $G \times \mathbb{Z}_2^f$.

A richer possibility is to consider a long exact sequence of groups

$$0 \to \mathbb{Z}_2 \to \widehat{G}_0 \to G \to \mathbb{Z}_2 \to 0, \tag{4.3.26}$$

where we denote the homomorphism from G to $\mathbb{Z}_2$ by π_1. The kernel of π_1 will be denoted G_0, then $\widehat{G}_0$ is a central extension of G_0 by $\mathbb{Z}_2$. Let n_2 be a 2-cocycle on G_0 corresponding to this central extension.

If $\pi_1(g) = 0$, we take $\mathcal{C}_g$ to have two simple objects, $V_{g,\epsilon}$, $\epsilon \in \mathbb{Z}_2$. If $\pi_1(g) = 1$, we take $\mathcal{C}_g$ to have a single simple object S_g. We again take the fusion rules to still mimic the Ising category: $S_g S_{g'} = V_{gg',0} + V_{gg',1}$, etc., but now require the $V_{\widehat{g}} \equiv V_{g,\epsilon}$ fusion to follow the $\widehat{G}_0$ multiplication rules.

It follows from the results of [145] that for any such long exact sequence there exists a fusion category with these fusion rules, provided a certain obstruction $[O_4] \in H^4(G, \mathbb{R}/\mathbb{Z})$ constructed from n_2 and π_1 vanishes. Possible associators depend are parameterized by a 3-cochain $\nu_3 \in C^3(G, \mathbb{R}/\mathbb{Z})$ such that $\delta\nu_3 = O_4$. As argued in appendix C.2, such a category has a fermion if and only if $[n_2]$ is a restriction of a class $[\beta_2]$ in $H^2(G, \mathbb{Z}_2)$, in which case ν_3 must satisfy the Gu-Wen equation (4.2.16). This TFT is a candidate for a shadow of a fermionic SPT phase with symmetry $G \times \mathbb{Z}_2^f$. One can view this theory as a G-equivariant version of the toric code, where some elements of G act by particle-vortex symmetry, and the fusion of G domain walls is associative only up to e and m lines. This failure of strict associativity is controlled by the extension class $n_2 \in H^2(G_0, \mathbb{Z}_2)$.

Thus we obtain categories labelled by a triple (ν_3, n_2, π_1) (and a choice of a fermion) which are shadows of fermionic TFTs with symmetry G. We will see below that all these TFTs are fermionic SPT phases, i.e. they are "invertible". On the other hand, one may argue that shadows of fermionic SPT phases with symmetry G must be G-equivariant versions of the toric code. Indeed, the component $\mathcal{C}_1$ of such a category must contain the identity object, the fermion Π, and no other simple objects, since condensing the fermion must give an invertible fermionic TFT. The fusion rules for $\mathcal{C}_1$ must have the same form as in the toric code, because Π generates a $\mathbb{Z}_2$ 1-form symmetry, and the associator for $\mathcal{C}_1$ must be trivial for Π to be a fermion. Thus $\mathcal{C}_1$

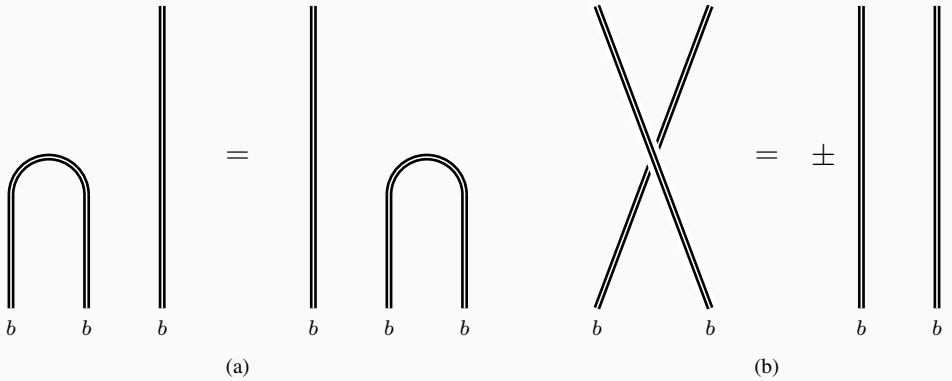


Figure 4.3.17: $\mathbb{Z}_2$ 1-form symmetries: (a) There exists a bulk line b with properties shown in the figure. (b) Half-braiding b lines across each other gives a factor of ± 1 when compared to b lines without braiding. The factor of $+1$ arises for a bosonic 1-form symmetry and -1 arises for a fermionic 1-form symmetry. This minus sign implies that the symmetry is anomalous.

describes the toric code, and $\mathcal{C} = \sum_g \mathcal{C}_g$ is a G-equivariant extension of the toric code.

4.3.7 Gauging one-form symmetries in the presence of gapped boundary conditions

Given a gapped boundary condition for $\mathfrak{T}_b$, we can derive in a simple manner a gapped boundary condition for $\mathfrak{T}_{\mathbb{Z}_2}$. Here we describe the process at the level of boundary line defects. In later sections we will test it at the level of partition sums and commuting projector Hamiltonians. [11]

We start from a spherical fusion category $\mathcal{C}_b$ equipped with a bosonic $\mathbb{Z}_2$ one-form symmetry generator $B = (b, \beta)$, an element of the center $Z[\mathcal{C}]$ such that $\beta_b = 1_{b\otimes b}$ and there is an isomorphism $\xi_b : b \otimes b \to I$ such that $\xi \otimes 1 = 1 \otimes \xi$ in $\mathrm{Hom}(b \otimes b \otimes b, b)$.

In the condensed matter language, our objective is to condense the anyon B. The general mathematical formalism for anyon condensation is described in [146]. It should be applied to the commutative separable algebra $A = I + B$. We will use a somewhat simplified procedure for concrete calculations.

We then define a new category $\mathcal{C}_0^{\mathbb{Z}_2}$ with the "same" objects and enlarged spaces of morphisms:

$$\mathrm{Hom}_{\mathcal{C}_0^{\mathbb{Z}_2}}(U_0, V_0) \equiv \mathrm{Hom}_{\mathcal{C}_b}(U, V) \oplus \mathrm{Hom}_{\mathcal{C}_b}(U, b \otimes V) \tag{4.3.27}$$

The new morphisms should be thought of as B-twisted sectors. The morphisms are composed with the help of the $b \otimes b \to 1$ map and the tensor product is defined with the help of β, as in Figure 5.3.7.

[11] Although we specialize here to a $\mathbb{Z}_2$ one-form symmetry, the same procedure works for a general Abelian group

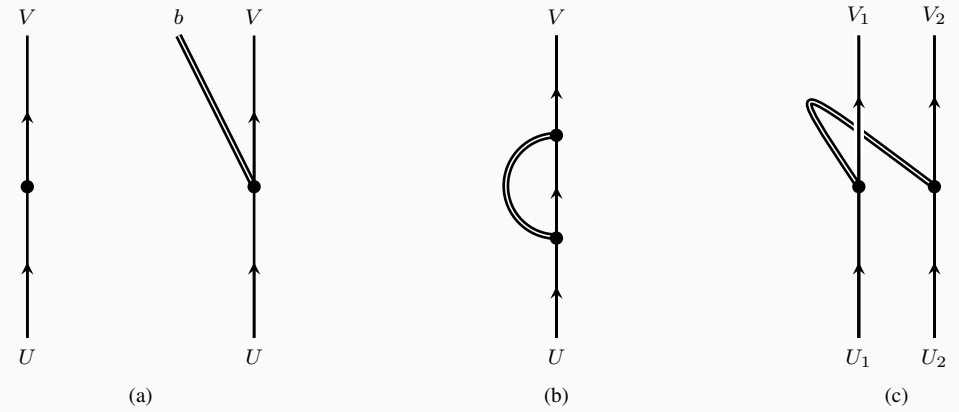


Figure 4.3.18: Construction of $\mathcal{C}_0^{Z_2}$: (a) A morphism can involve a b line or not. Notice that the direction of b line is irrelevant as it is equal to its dual. (b) Composition of two morphisms involving a b line is obtained by using the canonical map from $b \otimes b$ to identity to join the b lines. (c) Tensor product of two morphisms involving a b line is twisted by a half-braiding of b across V_1.

The image of simple objects under this map may not be simple: if X is simple, $b \otimes X$ is also simple and may or not coincide with X. In the former case, $\mathrm{Hom}_{\mathcal{C}_0^{\mathbb{Z}_2}}(X_0, X_0)$ is two-dimensional and X_0 will split into two simples $X_0^{\pm}$.

We then add to $\mathcal{C}_0^{\mathbb{Z}_2}$ the simple summands of the objects inherited from $\mathcal{C}_b$. Concretely, $X_0^{\pm}$ can be described as X_0 with the insertion of a projector $\pi_X^{\pm}$ along the line. The projectors will be linear combinations of the generator 1_X of $\mathrm{Hom}_{\mathcal{C}_b}(X, X)$ and the generator ξ_X of $\mathrm{Hom}_{\mathcal{C}_b}(X, b \otimes X)$. We can compute $\xi_X^2 = \eta_X 1_X$ and define projectors

$$\pi_X^{\pm} = \frac{1}{2}(1_X \pm (\eta_X)^{-1/2} \xi_X) \tag{4.3.28}$$

Notice that if b is the identity line in $\mathcal{C}_b$, the identity in $\mathcal{C}_0^{\mathbb{Z}_2}$ will itself split.

The final result will be a spherical (multi-)fusion category $\mathcal{C}_0^{\mathbb{Z}_2}$. We can extend $\mathcal{C}_0^{\mathbb{Z}_2}$ further to a $\mathbb{Z}_2$-graded category in the following manner. We extend the original category $\mathcal{C}_b$ to a new graded category $\mathcal{C}_b \times \mathbb{Z}_2$, a direct sum of 2 copies of $\mathcal{C}_b$. We extend the $\mathbb{Z}_2$ 1-form symmetry to $\mathcal{C}_b \times \mathbb{Z}_2$ by the center element $(b, \beta \times (-1)^{\epsilon})$, where B it taken to lie in $\mathcal{C}_b \times \{0\}$ and we twisted the original β by a sign when crossing a line in $\mathcal{C} \times \{1\}$.

Finally, we proceed as before using the extended center element. Objects in $\mathcal{C}_b \times \{\epsilon\}$ map to objects in $\mathcal{C}_{\epsilon}^{\mathbb{Z}_2}$. Concretely, the only difference between objects in $\mathcal{C}_0^{\mathbb{Z}_2}$ and $\mathcal{C}_1^{\mathbb{Z}_2}$ is an extra sign in the tensor product of morphisms which appear when the b line crosses a $\mathcal{C}_1^{\mathbb{Z}_2}$ object.

4.3.8 Example: 1-form symmetries in the toric code

Consider again the spherical fusion category $\mathcal{C}$ modelled on $\mathbb{Z}_2$, with two objects 1 and P fusing as $P \otimes P \simeq 1$ and trivial associators.

The Wilson loop in this $\mathbb{Z}_2$ gauge theory is the object $e = (1, \beta_P = -1)$ in the center of the category. It is a boson generating an "electric" $\mathbb{Z}_2^e$ 1-form symmetry. If we gauge this 1-form symmetry, we obtain a category $\mathcal{C}_e$ with elements I_0, P_0 with a two-dimensional space of morphisms. We can denote the generators of these morphisms as 1_1, ξ_1, 1_P, ξ_P. We have $\xi_1^2 = 1_1$ and $\xi_P^2 = 1_P$.

We can decompose $I_0 = I^{++} + I^{--}$ and $P_0 = P^{+-} + P^{-+}$. Working out the fusion rules, we find a multi-fusion category, with P^{+-} and P^{-+} being domain walls between the two vacua. Each vacuum has a trivial category of line defects.

Adding twisted sectors gives us two new objects, $I_1 = I^{+-} + I^{-+}$ and $P_1 = P^{++} + P^{--}$. Hence our final graded multi-fusion category has has four sectors, $\mathcal{C}_{\pm\pm}$, each consisting of an element of grading 0 and an element of grading 1. Physically, this is a boundary condition with two trivial vacua, each described by the trivial $\mathbb{Z}_2$-graded fusion category.

This makes sense. We obtained the toric code by gauging the $\mathbb{Z}_2$ global symmetry of a trivial theory. In the absence of boundary conditions, gauging the dual 1-form symmetry effectively ungauges the $\mathbb{Z}_2$ gauge theory. In the presence of boundary conditions, gauging the standard $\mathbb{Z}_2$ symmetry with Dirichlet b.c. leaves us with a bulk $\mathbb{Z}_2$ gauge theory with a residual $\mathbb{Z}_2$ global symmetry at the boundary. This can be thought of as a $\mathbb{Z}_2$ gauge theory coupled to a boundary $\mathbb{Z}_2$-valued sigma model. After we gauge the 1-form symmetry, the boundary sigma model remains and the extra $\mathbb{Z}_2$ global symmetry is spontaneously broken.

On the other hand, gauging the 1-form symmetry generated by m leads to a category $\mathcal{C}_m$ with two isomorphic simple elements I_0, P_0. This is again a trivial category of line defects. Adding twisted sectors, we find two more isomorphic objects, I_1 and P_1. We have obtained again the trivial $\mathbb{Z}_2$-graded fusion category.

4.3.9 The Π-product of shadows

The product of two theories $\mathfrak{T}_f$ and $\mathfrak{T}'_f$ is equipped with a bosonic line $\Pi\Pi'$ which generates a standard bosonic $\mathbb{Z}_2$ 1-form symmetry. If we gauge $\Pi\Pi'$, we obtain a new theory which we can denote as $\mathfrak{T}_f \times_f \mathfrak{T}'_f$. This new theory still has a fermionic 1-form symmetry, generated by Π, or equivalently Π' (the two coincide in the new theory). It is our candidate for the shadow of $\mathfrak{T}_s \times \mathfrak{T}'_s$.

The shadow product $\mathfrak{T}_f \times_f \mathfrak{T}'_f$ should be associative, as it corresponds to the product operation of the corresponding spin-TFTs. This is quite clear from the definition as well: the product of three shadows contain bosonic generators $\Pi\Pi'$, $\Pi'\Pi''$ and $\Pi\Pi''$ generating a $\mathbb{Z}_2 \times \mathbb{Z}_2$ 1-form symmetry. Gauging the two $\mathbb{Z}_2$ in any order should be equivalent to gauging both. In the language of anyon condensation, we are condensing the algebra $A = I + \Pi\Pi' + \Pi'\Pi'' + \Pi\Pi''$.

We would like to explore the group structure of the candidate fermionic SPT phases we have encountered until now. Recall that we have introduced two basic classes of fermionic SPT phases: Ising pull-backs $\mathcal{I}^{\xi}_{\pi_1}[G]$ and Gu-Wen phases $\mathcal{G}_{\nu_3,n_2}[G]$.

Gu-Wen SPT phases

As a simple example, consider two Gu-Wen phases $\mathcal{G}_{\nu_3,n_2}[G]$ and $\mathcal{G}_{\tilde{\nu}_3,\tilde{n}_2}[G]$. We can take the G-graded product of the corresponding categories. The result is a G-graded category with objects $V_{g,\epsilon,\tilde{\epsilon}}$ which fuse according to a $\mathbb{Z}_2 \times \mathbb{Z}_2$ extension G' of G, with cocycle $(n_2, \tilde{n}_2)$ and associators $\widehat{\alpha}_3\widehat{\tilde{\alpha}}_3$.

The bosonic symmetry generator is $V_{e,1,1}$, equipped with crossing $(-1)^{\epsilon+\tilde{\epsilon}}$. As we gauge the symmetry, we will extend the morphisms so that $V^0_{g,\epsilon,\tilde{\epsilon}}$ and $V^0_{g,\epsilon+1,\tilde{\epsilon}+1}$ become isomorphic. Keeping this identification into account, the resulting objects will fuse according to the $\mathbb{Z}_2$ extension $\widehat{G}$ of G associated to the cocycle $n_2 + \tilde{n}_2$.

Computing the associator of the new category takes a bit of effort. For concreteness, we can pick representative objects $V^0_{g,\epsilon,0}$. When we multiply them, we obtain, say, $V^0_{gg',\epsilon+\epsilon'+n_2(1,g,gg'),\tilde{n}_2(1,g,gg')}$ which has to be mapped back to $V^0_{gg',\epsilon+\epsilon'+n_2(1,g,gg')+\tilde{n}_2(1,g,gg'),0}$ by inserting $\tilde{n}_2(1,g,gg')$ extra intersections with $\Pi\tilde{\Pi}$ lines.

We can gauge fix and then compute the associator via the tetrahedron graph. We obtain $\nu_3\tilde{\nu}_3(-1)^{\epsilon_1\cup\tilde{n}_2+n_2\cup\epsilon_1}$ where ϵ_1 encodes the first $\mathbb{Z}_2$ grading of the elements placed on the edges. This differs from $\nu_3\tilde{\nu}_3(-1)^{(n_2+\tilde{n}_2)\cup\epsilon_1}$ by a sign

$$(-1)^{\epsilon_1\cup\tilde{n}_2+\tilde{n}_2\cup\epsilon_1} = (-1)^{\delta(\epsilon_1\cup_1\tilde{n}_2)+\epsilon_1\cup_1\delta\tilde{n}_2+\delta\epsilon_1\cup_1\tilde{n}_2} \tag{4.3.29}$$

The second term above is zero and the first term can be absorbed into a gauge redefinition of the associator. Hence, we obtain a new Gu-Wen super-cohomology phase $(\nu'_3, n_2 + \tilde{n}_2)$ with

$$\nu'_3 = \nu_3\tilde{\nu}_3(-1)^{n_2\cup_1\tilde{n}_2} \tag{4.3.30}$$

This is indeed the expected group law for Gu-Wen fermionic SPT phases.

The squared equivariant toric code

Another interesting example is the product of two equivariant toric codes. The resulting $\mathbb{Z}_2$-graded category has objects II, PI, IP, PP in $\mathcal{C}_0$ and SS in $\mathcal{C}_1$. The bosonic generator is then $\psi\psi$.

We choose the two fourth roots ξ_1 and ξ_2 of -1 which identify specific Ising Π-categories $\mathcal{I}^{\xi_1}$ and $\mathcal{I}^{\xi_2}$. Recall that only ξ^2 affects the crossing phases. A flip $\xi \to -\xi$ changes the associator SSS and thus effectively twists the category by a $\mathbb{Z}_2$ group cocycle, i.e. multiplies the theory by a bosonic $\mathbb{Z}_2$ SPT phase.

Gauging the 1-form symmetry leads one to identify the pairs $II^0 \simeq PP^0$ and $PI^0 \simeq IP^0$, while SS^0 will split into some S^0_+ and S^0_-.

Fusion of SS^0 with PI^0 from the left involves crossing $\psi\psi$ across PI and hence flips the sign of the non-trivial morphism of SS^0 to itself. On the other hand, fusion with PI^0 from the right flips the sign of the non-trivial morphism because of non-trivial PSP associators for Ising category. We thus learn that $PI^0 \otimes S^0_+ = S^0_+ \otimes PI^0 = S^0_-$ and $PI^0 \otimes S^0_- = S^0_- \otimes PI^0 = S^0_+$. These fusion rules do not depend on ξ_1 or ξ_2.

The fusion rules involving S^0_+ and S^0_-, on the other hand, are affected by the β_S crossing phases. We find that if $\xi_1 = \xi_2$, or more generally $\xi_1^2\xi_2^2 = -1$, we have $S^0_+ \otimes S^0_+ \simeq S^0_- \otimes S^0_- \simeq PI^0$ and $S^0_+ \otimes S^0_- \simeq S^0_- \otimes S^0_+ \simeq II^0$: the objects in the new category fuse according to a $\widehat{G} = \mathbb{Z}_4$ group law, generated, say, by S^0_+. We demonstrate an example of computation of fusion rules for this case in fig. 4.3.19.

The $\widehat{G} = \mathbb{Z}_4$ can be regarded as a $\mathbb{Z}_2$ central extension of $G = \mathbb{Z}_2$ with $V_{0,0} = II^0$, $V_{0,1} = PI^0$, $V_{1,0} = S^0_+$ and $V_{1,1} = S^0_-$. It can be easily checked that PI^0 equipped with crossing $(-1)^\epsilon$ is a fermionic bulk line. The result is the shadow of Gu-Wen phase for a $\mathbb{Z}_2$ global symmetry, with $\mathbb{Z}_4$ being the central extension.

On the other hand, if $\xi_1^2\xi_2^2 = 1$, we have $S^0_+ \otimes S^0_+ \simeq S^0_- \otimes S^0_- \simeq II^0$ and $S^0_+ \otimes S^0_- \simeq S^0_- \otimes S^0_+ \simeq PI^0$: the objects in the new category fuse according to a $\widehat{G} = \mathbb{Z}_2 \times \mathbb{Z}_2$ group law. We can set, say, $V_{0,0} = II^0$, $V_{0,1} = PI^0$, $V_{1,0} = S^0_+$ and $V_{1,1} = S^0_-$. The result is the shadow of a Gu-Wen phase for a $\mathbb{Z}_2$ global symmetry, with trivial central extension, i.e. a bosonic SPT phase.

We still need to compute the associator $\widehat{\alpha}_3(\xi_1, \xi_2)$. We can compute the associativity phases for $S^0_\pm$ from the associator for $SS \otimes SS \otimes SS$ or by evaluating some tetrahedron planar graphs. The general calculation is somewhat tedious and we will omit it. It should be obvious that if $\xi_1\xi_2 = 1$ all crossing or associator phases will cancel out among the two theories. Thus we expect to obtain a trivial associator as well as the trivial group extension. Thus we claim

$$\mathcal{I}^\xi \times_f \mathcal{I}^{\xi^{-1}} \simeq I \tag{4.3.31}$$

where I denotes the trivial $\mathbb{Z}_2$ SPT phase. In particular, this proves the claim that the Ising Π-category is the shadow of an SPT phase!

On the other hand, $\mathcal{I}^\xi \times_f \mathcal{I}^\xi$ will be a root Gu-Wen $\mathbb{Z}_2$ SPT phase, which one of the two being determined by the value of ξ^2, as the sign of ξ can be changed by adding a bosonic SPT phase. We can compute ν_3 for that phase by looking at graphs involving S^0_+ and identity lines, with Π lines emerging from junctions with two incoming S_+ lines. The only source of interesting phases is the crossing phase of the fermion and S_+. We find that if $\xi^2 = \pm i$, then $\nu_3 = \pm\frac{1}{4}\eta_1 \cup \eta_1 \cup \eta_1$.

Ising pull-back and Gu-Wen

We can combine the Ising pull-back category with homomorphism π_1 with a Gu-Wen phase. The G-graded product has objects $I_{g,\epsilon}$, $P_{g,\epsilon}$ or $S_{g,\epsilon}$ depending on the value of $\pi_1(g)$. Gauging the bosonic 1-form symmetry identifies $P^0_{g,\epsilon}$ with $I^0_{g,\epsilon+1}$ and $S_{g,\epsilon}$ with $S_{g,\epsilon+1}$.

We can restrict ourselves to objects $I^0_{g,\epsilon}$, or $S^0_{g,0}$. Effectively, the n_2 cocycle has been restricted to a cocycle n_2^0 on $G_0 = \ker \pi_1$. The fusion rules of this category mimic our example based on a long exact sequence

$$0 \to \mathbb{Z}_2 \to \widehat{G}_0 \to G \to \mathbb{Z}_2 \to 0 \tag{4.3.32}$$

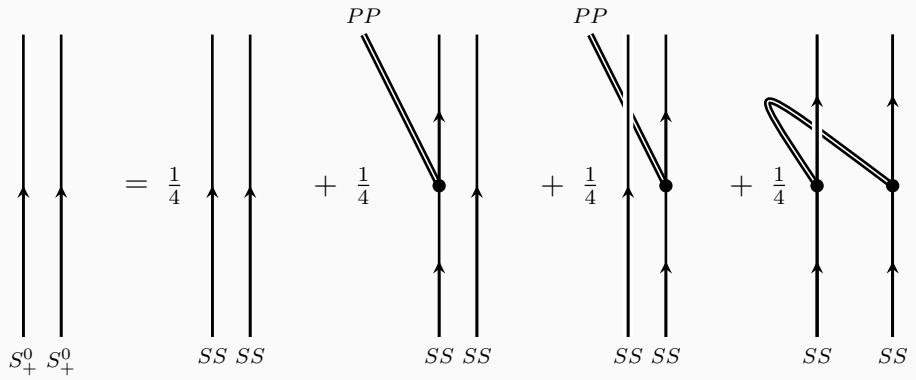


Figure 4.3.19: A sample computation of fusion rules in shadow product of two equivariant toric codes with $\xi_1 = \xi_2$: $S^0_+ \otimes S^0_+$ is by definition a sum of four terms which involve associators and crossings. II inside $SS \otimes SS$ is mapped to zero object as second term cancels against the third and the first term cancels against the fourth. PI is mapped to PI by the first and fourth terms and to IP by the second and third terms. Hence, $S^0_+ \otimes S^0_+ \simeq PI^0$.

One might wonder what happens to the rest of the data of the n_2 cocycle which is not captured by n_2^0. This data goes into the associators for the new category. In particular, it is possible to extract the values of $n_2(g, g')$ from (relative) signs of certain associators. This is of course true in the Gu-Wen case as well, where $n_2(g, g')$ is also encoded, for example, by the sign in the associator $(V_{g,0} \otimes V_{g',0}) \otimes V_{1,1} \to V_{g,0} \otimes (V_{g',0} \otimes V_{1,1})$.

The associators can be determined from tetrahedron graph by inserting the bosonic line $P_{1,1}$ at appropriate junctions. All of them can be written (modulo factors of square root of 2) as ν_3 times a sign which depends on n_2, the choice of morphism $S \otimes S \to (I, P)$ and ϵ grading of lines. We show two sample associators and their results in Figure 4.3.20.

Choosing $\epsilon = 1$, $\epsilon' = 0$ and g to be identity in Figure 4.3.20(a) tells us that the associator equals $(-1)^{n_2(g',g'')}$. This means that sign of this associator determines $n_2(g', g'')$ for such that $\pi_1(g') = 0$ and $\pi_1(g'') = 1$. Similarly, we could compute the associator of $I_{g,\epsilon}$, $S_{g',0}$ and $I_{g'',\epsilon''}$ and choosing g as identity, $\epsilon = 1$ and $\epsilon'' = 0$ would determine $n_2(g', g'')$ such that $\pi_1(g') = 1$ and $\pi_1(g'') = 0$. Determining $n_2(g', g'')$ such that $\pi_1(g') = 1$ and $\pi_1(g'') = 1$ is a bit more non-trivial. It is determined by the associator in Figure 4.3.20(b) when we choose $\epsilon = 1$, $m = n_2(g', g'')$ and the particular n for which the graph evaluates to a non-zero number. Notice that there is only one such n.

We will verify now that every long sequence example can be obtained in this manner.

Ising pull-back and long exact sequence with the same π_1

The product category has objects $V_{g_0,\epsilon}$, $PV_{g_0,\epsilon}$ and SS_{g_1}. The bosonic generator is associated to $P_{e,1}$. Condensation will identify $V_{g_0,\epsilon}$ and $PV_{g_0,\epsilon+1}$ and split SS_{g_1} to $S_{g_1,\epsilon}$.

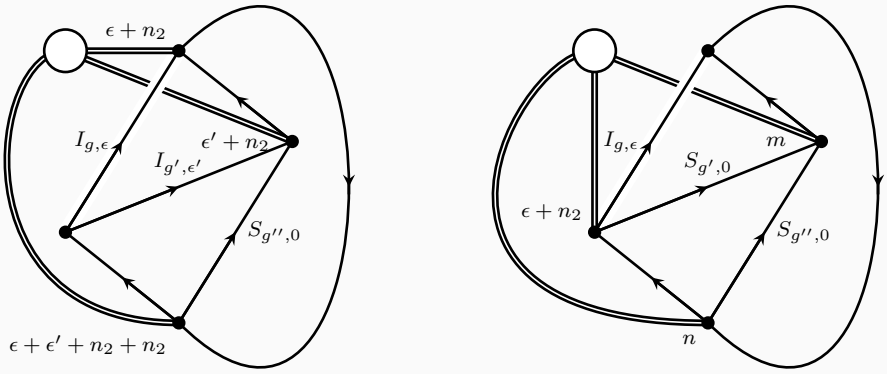


Figure 4.3.20: Two sample computations of associators for a phase corresponding to long exact sequence. The values at the starting of double lines encode the number of $P_{1,1}$ lines. We leave the argument of n_2 self-evident as it can be read from the diagram. m and n are numbers (defined modulo 2) associated to the choice of morphisms at the junctions where two S lines converge and diverge respectively. m is 1 if it corresponds to the morphism $S \otimes S \to P$ and 0 if it corresponds to $S \otimes S \to I$. n is defined similarly. The graph in (b) evaluates to a non-zero number only if $n_2(g', g'') + n_2(g, g'g'') + n_2(gg', g'') + \epsilon + m + n = 0$ which is the same as $n_2(g, g') + \epsilon + m + n = 0$. As a result of this, the double lines always come in pairs. The graphs in (a) and (b) imply that the associators respectively are $(-1)^{\epsilon(\epsilon' + n_2(g',g''))}\nu_3$ and $\lambda^{(m+1)(n+1)}(-1)^{\epsilon m}\nu_3$ where λ is a square root of 2.

It turns out that the consistency of fusion rules completely constrains them. First of all, we don't physically expect any of $S_{g,\epsilon} \otimes S_{g',\epsilon'}$ to be the zero object. This implies that they must fuse to a single object since the fusion of sums $(S_{g,0} \oplus S_{g,1}) \otimes (S_{g',0} \oplus S_{g',1})$ is equal to sum of four objects $V_{gg',0} \oplus V_{gg',0} \oplus V_{gg',1} \oplus V_{gg',1}$. Using similar arguments, we find that the fusion of two simple objects must be a single simple obejct. Second, $V_{g,0}$ and $V_{g,1}$ must map $S_{g',\epsilon}$ to different objects. If, on the contrary $V_{g,0} \otimes S_{g',\epsilon} \simeq V_{g,1} \otimes S_{g',\epsilon} \simeq S_{gg',\epsilon'}$, then we could fuse by $S_{g'',\epsilon''}$ from the right to find that the elements in subcategory associated to G_0 do not fuse according to a cocycle, leading to a contradiction. Third, the fusion of S elements with themselves must be captured by a cochain. This can be shown using a similar argument as above. This cochain can be combined with the cocycle for G_0 to give rise to a cochain for G governing the fusion rules for the full category. Associativity of fusion then implies that this cochain must be a cocycle n_2.

Thus, we see that this is a Gu-Wen extension example, with objects $V_{g_0,\epsilon}$ and $V_{g_1,\epsilon} = S_{g_1,\epsilon}$. As $\mathcal{I}^{\xi}_{\pi_1}$ and $\mathcal{I}^{\xi^{-1}}_{\pi_1}$ are inverse to each other, we can express any long exact sequence example as the Π-product of $\mathcal{I}^{\xi}_{\pi_1}$ and a Gu-Wen phase.

Product of long exact sequence examples

In a similar manner, we can verify that the $\times_f$ product of two long exact sequence examples is a new long exact sequence example. The product has a bosonic line $V_{e,1}\tilde{V}_{e,1}$. Let π_1 and $\tilde{\pi}_1$ be respectively the two homomorphisms.

- In the $\pi_1(g) = \tilde{\pi}_1(g) = 0$ sector, gauging the bosonic 1-form symmetry identifies $V_{g,\epsilon}\tilde{V}_{g,\epsilon'}$ with $V_{g,\epsilon+1}\tilde{V}_{g,\epsilon'+1}$ and we can choose representative objects as $V'_{g,\epsilon} = V_{g,\epsilon}\tilde{V}_{g,0}$.
- In the $\pi_1(g) = 0, \tilde{\pi}_1(g) = 1$ sector,$V_{g,\epsilon}\tilde{S}_g$ is identified with $V_{g,\epsilon+1}\tilde{S}_g$ and we choose representative object $S'_g = V_{g,0}\tilde{S}_g$.
- In the $\pi_1(g) = 1, \tilde{\pi}_1(g) = 0$ sector, $S_g\tilde{V}_{g,\epsilon'}$ is identified with $S_g\tilde{V}_{g,\epsilon'+1}$ and we choose representative object $S'_g = S_g\tilde{V}_{g,0}$.
- In the $\pi_1(g) = \tilde{\pi}_1(g) = 1$ sector, $S_g\tilde{S}_g$ splits into two objects (as in the product of two equivariant toric codes above) which we denote as $V'_{g,0}$ and $V'_{g,1}$.

The fusion rules of representative objects can be obtained analogously to the examples above.

This can be identified with a long exact sequence

$$0 \to \mathbb{Z}_2 \to G' \to G \to \mathbb{Z}_2 \to 0 \tag{4.3.33}$$

with $G \to \mathbb{Z}_2$ homomorphism $\pi'_1 = \pi_1 + \tilde{\pi}_1$. It is somewhat trickier to determine the G' central extension: while the restriction to $\pi_1(g_1) = \pi'_1(g_1) = \pi_1(g_2) = \pi'_1(g_2) = 0$ coincides with $n_2(g_1, g_2) + \tilde{n}_2(g_1, g_2)$, the rest of it depends on the details of the associators of the two initial categories.

We can attack the problem by specializing first to Ising pull-backs.

4.3.10 Triple products and quaternions

In consideration of our analysis, we expect some relation of the form

$$\mathcal{I}^{\xi}_{\pi_1}[G] \times_f \mathcal{I}^{\xi}_{\pi_1'}[G] = \mathcal{G}_{\nu_3(\pi_1,\pi_1',\xi),n_2(\pi_1,\pi_1',\xi)}[G] \times_f \mathcal{I}^{\xi^{-1}}_{\pi_1+\pi_1'}[G] \tag{4.3.34}$$

We switched the ξ phase for the Ising pull-back on the right hand side for future convenience.

In order to extract the Gu-Wen phase which appears in this expression, we consider the triple product

$$\mathcal{G}_{\nu_3(\pi_1,\pi_1'),n_2(\pi_1,\pi_1')}[G] = \mathcal{I}^{\xi}_{\pi_1}[G] \times_f \mathcal{I}^{\xi}_{\pi_1+\pi_1'}[G] \times_f \mathcal{I}^{\xi}_{\pi_1'}[G] \tag{4.3.35}$$

The details of the calculation only depend on the image of G group elements under π_1 and π_1'. Without loss of generality, we can do our computation for $G = \mathbb{Z}_2 \times \mathbb{Z}_2'$ with π_1 and π_1' being the projections into the first and second factor respectively. The general answer will be obtained by pulling back the $\mathbb{Z}_2 \times \mathbb{Z}_2'$ answer by $\pi_1 \times \pi_1'$.

This is a rather non-trivial calculation, but it is somewhat simplified by the permutation symmetry acting on the triple π_1, π_1', $\pi_1+\pi_1'$, although gauge-fixing choices may break the symmetry at intermediate stages of the calculation. The $[n_2]$ cocycle is actually independent from ξ^2: a shift of ξ^2 will be implemented by multiplying by the root Gu-Wen phase pulled back along π_1, π_1' and $\pi_1 + \pi_1'$, which shifts the cocycle by

$$\pi_1 \cup \pi_1 + \pi_1' \cup \pi_1' + (\pi_1 + \pi_1') \cup (\pi_1' + \pi_1') = \pi_1 \cup \pi_1' + \pi_1' \cup \pi_1 \tag{4.3.36}$$

which is exact.

It turns out to be possible to pick a gauge-fixing in which n_2 is at least cyclically symmetric. We take triple product of elements of various Ising categories in the order mentioned in (4.3.35). For instance, $\pi_1 = 0, \pi_1' = 1$ sector contains elements of the form ISS and PSS. The lines IPP, PIP and PPI give rise to a $\mathbb{Z}_2 \times \mathbb{Z}_2$ bosonic 1-form symmetry. There are two choices of junctions between these three lines. They correspond to canonical junctions between IPP, PIP and PPI lines taken in clockwise and counter-clockwise order respectively. Their product is clearly equal to 1 and their square is -1 as it involves a crossing. Hence, when we bring together these centre lines in calculations, we multiply the canonical junctions by i and $-i$ respectively. [12]

When we condense, the three lines generate three non-trivial morphisms such that the product of two of these gives rise to the third. We choose PPI to identify ISS with PSS, IPP to identify SIS with SPS, and PIP to identify SSI with SSP in a cyclic fashion. Similarly, we choose IPP to split ISS into ISS_+ and ISS_- etc. in a cyclic manner. We summarize our choice of objects in the condensed category:

- In $\pi_1 = \pi_1' = 0$ sector, the objects are III and IPI. IPI equipped with an appropriate crossing is the generator of fermionic 1-form symmetry. We rename III and IPI as V_1 and V_{-1} respectively.

[12] In the language of anyon condensation, this is the chocie of maps $A \otimes A \to A$ and $A \to A \otimes A$ with good properties.

- In $\pi_1 = 0, \pi_1' = 1$ sector, the objects are ISS_+ and ISS_-. We rename them as $V_{\pm i}$.
- In $\pi_1 = 1, \pi_1' = 0$ sector, the objects are SSI_+ and SSI_-. We rename them as $V_{\pm j}$.
- In $\pi_1 = \pi_1' = 1$ sector, the objects are SIS_+ and SPS_-. We rename them as $V_{\pm k}$.

Some of the computations of fusion rules are completely analogous to the case of squared equivariant toric code. These are $(\pm q)\otimes(-1) \simeq (-1)\otimes(\pm q) \simeq \mp q$, $q\otimes q \simeq -1$ and $q\otimes(-q) \simeq 1$ where q denotes either one of i, j and k.

The other computations are analogous but we have to be careful about choosing correct sign for the junctions of three bosonic lines. We show how these junctions arise in a sample computation in 4.3.21. The final result is captured by the quaternion group:

$$\begin{aligned} i^2 = j^2 = k^2 = -1 \\ ij = -ji = k \\ jk = -kj = i \\ ki = -ik = j \end{aligned} \tag{4.3.37}$$

This corresponds to the cocycle

$$n_2(\pi_1, \pi_1') = \pi_1 \cup \pi_1 + \pi_1' \cup \pi_1' + \pi_1 \cup \pi_1' \tag{4.3.38}$$

This describes the quaternion group as a $\mathbb{Z}_2$ central extension of $\mathbb{Z}_2 \times \mathbb{Z}_2'$!

The extension indeed enjoys S_3 permutation symmetry, up to a gauge transformation $i \to -i$, $j \to -j$, $k \to -k$ for odd permutations.

Computing $\nu_3(\pi, \pi', \xi)$ is of course rather more cumbersome. We leave it as an exercise for the enthusiastic reader.

4.3.11 The group structure of fermionic SPT phases

The dependence of $\mathcal{I}^\xi[\mathbb{Z}_2]$ on ξ is very mild: we can change ξ to another fourth root of -1 by multiplying it by one of the four $\mathbb{Z}_2$ Gu-Wen phases. Correspondingly, we can change ξ in $\mathcal{I}^\xi_{\pi_1}[G]$ by multiplying it by the pull-back along π_1 of one of the four $\mathbb{Z}_2$ Gu-Wen phases. Consequently, we can just stick to a specific choice of ξ in the following.

We expect all fermionic SPT phases to take the form $\mathcal{G}_{\nu_3,n_2}[G] \otimes \mathcal{I}^\xi_{\pi_1}[G]$. We could label such a phase by a triple (ν_3, n_2, π_1). It is natural to ask what is the group law for such phases.

We know that the product of two Gu-Wen phases is another Gu-Wen phase, with addition law

$$\mathcal{G}_{\nu_3,n_2} \times \mathcal{G}_{\nu_3',n_2'} = \mathcal{G}_{\nu_3+\nu_3'+\frac{1}{2}n_2\cup_1 n_2', n_2+n_2'} \tag{4.3.39}$$

This can be expressed as the statement that the group $\mathcal{G}[G]$ of Gu-Wen phases is a central extension

$$0 \to H^3[BG, U(1)] \to \mathcal{G}[G] \to H^2[BG, \mathbb{Z}_2] \to 0 \tag{4.3.40}$$

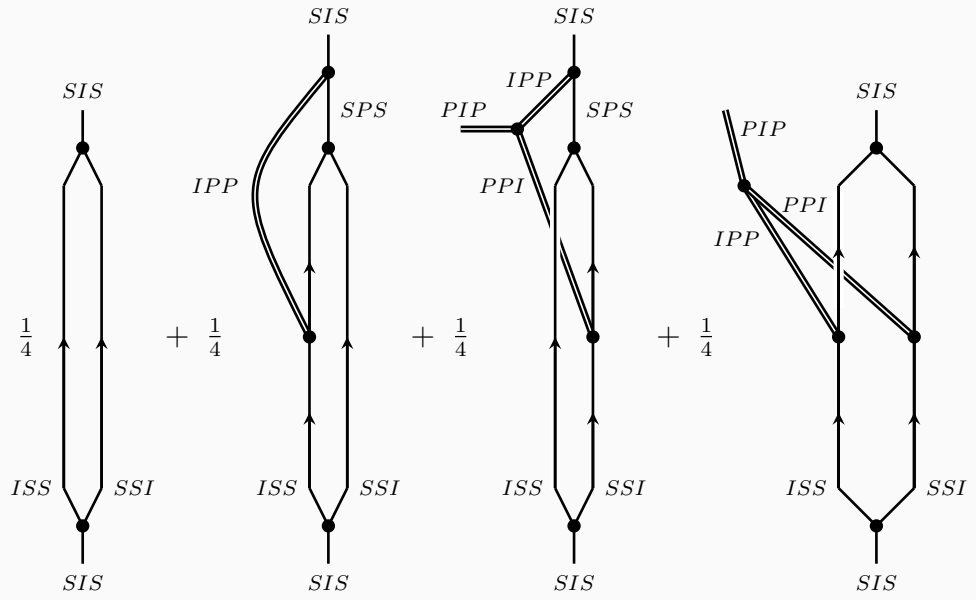


Figure 4.3.21: The figure depicts various terms in the computation of $ISS_+ \otimes SSI_+$. We need to convert SPS into SIS using the chosen isomorphisms. This results in junctions of the three bosonic lines with appropriate signs. After taking into account various factors from associators, crossings and junctions, we obtain $ISS_+ \otimes SSI_+ \simeq SIS_+$.

with cocycle $\frac{1}{2} n_2 \cup_1 n_2'$ valued in $H^3[BG, U(1)]$.

Similarly, when we add Ising pull-back phases the π_1 cocycles add up. Hence the group of fermionic SPT phases $\mathcal{F}[G]$ of the form (ν_3, n_2, π_1) will be a central extension

$$0 \to \mathcal{G}[G] \to \mathcal{F}[G] \to H^1[BG, \mathbb{Z}_2] \to 0 \tag{4.3.41}$$

The $\mathcal{G}[G]$-valued cocycle $\mathcal{G}_2$ for this extension can be computed by the relation

$$\mathcal{I}^{\xi}_{\pi_1}[G] \times_f \mathcal{I}^{\xi}_{\pi_1'}[G] = \mathcal{G}_2(\pi_1, \pi_1', \xi) \times_f \mathcal{I}^{\xi}_{\pi_1+\pi_1'}[G] \tag{4.3.42}$$

Comparing with our previous computation, the change $\xi^{-1} \to \xi$ on the right hand side shifts n_2 by $(\pi_1 + \pi_1') \cup (\pi_1 + \pi_1')$. Thus $\mathcal{G}_2(\pi_1, \pi_1', \xi)$ has cocycle

$$n_2(\pi_1, \pi_1') = \pi_1' \cup \pi_1 \tag{4.3.43}$$

This corresponds to the dihedral group extension of $\mathbb{Z}_2 \times \mathbb{Z}_2'$.

A standard presentation of the dihedral group is given by elements a and b such that $a^4 = b^2 = 1$ and $aba = b$. In our case, we can choose $V_a = SIS_+$ and $V_b = SSI_+$.

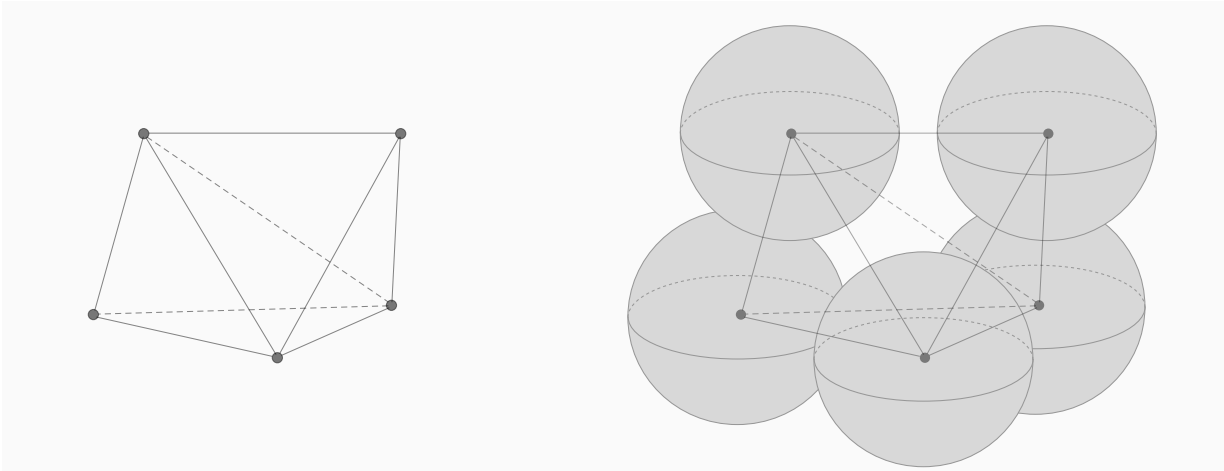


Figure 4.4.1: The first step of a swiss cheese construction: the manifold is triangulated, and spherical holes are opened up at the vertices of the triangulation.

4.3.12 Π-categories and Π-supercategories

There is a known relationship between Π-categories and super-categories which is analogous the the relation between $\mathcal{C}_b$ and $\mathcal{C}_{\mathbb{Z}_2}$ in the bosonic case [137].

Given a Π-category $\mathcal{C}_f$, we can build a super-category $\mathcal{C}_s$ whose even morphisms are $\mathrm{Hom}_{\mathcal{C}_f}(X, Y)$ and odd morphisms are $\mathrm{Hom}_{\mathcal{C}_f}(X, \Pi \otimes Y)$. This is a "$\Pi$-supercategory", i.e. a super-category equipped with an object Π with is odd-isomorphic to I. Vice versa, we can go from a Π-supercategory to a Π-category by dropping the odd morphisms.

In a previous work [118] , we sketched a state-sum construction based on spherical super-fusion categories. For simplicity, we assumed the spherical super-fusion category had no Majorana objects, i.e. irreducible objects with an even and an odd endomorphisms. If we take the "Π-envelope" of such a super-category we will get a Π-supercategory with an even and an odd copy of each irreducible object. Dropping odd morphisms we get a Π-category. The state-sum construction given in [118] builds up the spin TFT whose shadow is associated to this Π-category. In the next section, we will formulate the state-sum construction for general Π-categories. It should be possible to re-formulate it in terms of the associated super-categories, with or without Majorana objects.

4.4 Spherical fusion categories and state sum constructions

It is instructive to review the physical derivation of the Turaev-Viro construction for a 3d TFT with a single vacuum and a gapped boundary condition.

We begin with the observation that such a topological field theory $\mathfrak{T}$ associates a one-dimensional Hilbert space to a two-sphere. Thus a boundary $\mathfrak{B}$ with the topology of a two-sphere must create

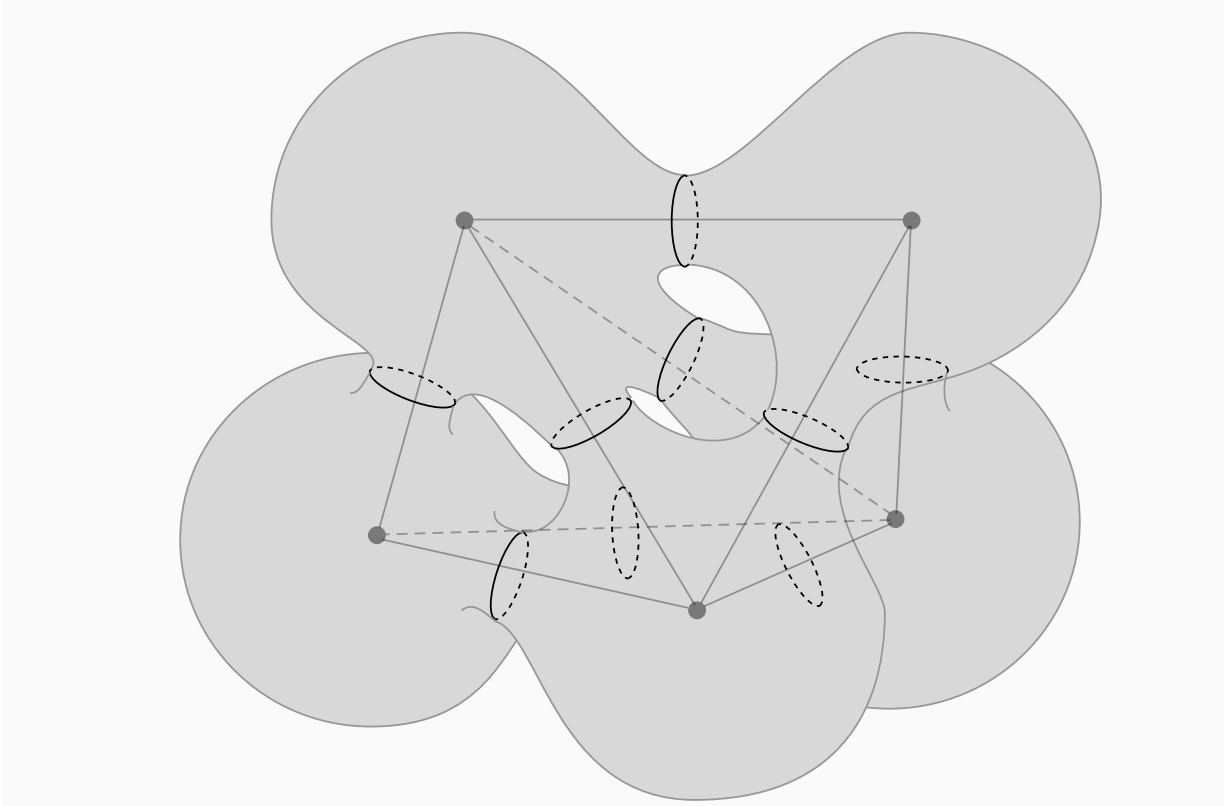


Figure 4.4.2: The second step of a swiss cheese construction: we connect the holes by tubes running along the edges of the triangulation. A special line is added around the tubes to make them trivial.

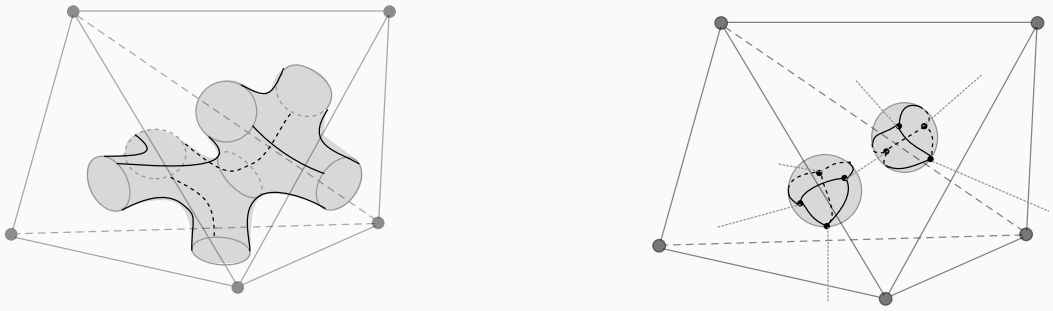


Figure 4.4.3: The third step of a swiss cheese construction. The complement of the holes is a collection of solid cylinders running through faces of the triangulation, fused at tetra-valent vertices inside the tetrahedra. The cross-section of the solid cylinders is a disk with three boundary punctures. We insert a complete basis of states along each cylinder.

a state in that Hilbert space which is proportional to the state created by a three-ball, with some specific proportionality constant $C_{\mathfrak{B}}$ which depends on the theory and on the boundary condition.

Consider a three-manifold M, say with no boundaries, for which we want to compute the partition function. Equip M with some triangulation. Up to a factor of $C_{\mathfrak{B}}$ for each vertex, the partition function of M will be the same as the partition function of a manifold M' obtained from M by removing a small ball around each vertex of the triangulation and replacing it with a spherical boundary of type $\mathfrak{B}$. See Figure 4.4.1.

We can enlarge the holes in M' until the spherical boundaries collide with each other, so that each hole almost fills the corresponding 3-cell in the cell-decomposition of M dual to the triangulation. The manifold M' looks like a foam of empty bubbles.

Next, we can "pop" the walls between bubbles. Concretely, this requires us to carve out parts of M' with the topology of a cylinder with $\mathfrak{B}$ boundaries at each end, i.e. $[0,1]_{\mathfrak{B},\mathfrak{B}} \times D^2$. The manifold is cut along the annulus $[0,1]_{\mathfrak{B},\mathfrak{B}} \times \partial D^2$. The path integral on the cylinder produces some state in the Hilbert space associated by the theory to the annulus with $\mathfrak{B}$ boundary conditions on the edge. We can replace the cylinder by some other geometry bounded by the same annulus, as long as they produce the same vector in the annulus Hilbert space.

An example of such geometry is half a solid torus, bounded by that annulus and by an annulus with $\mathfrak{B}$ boundary condition, decorated by some boundary line defect L_i running along the annulus. It is natural to expect such geometries to produce a basis in the Hilbert space as the choice of L_i is varied over all simple objects. [13] Thus the cylinder path integral should produce a state which can be decomposed as a linear combination of these elements, with some coefficients c_i. The correct choice of c_i is known to coincide with the quantum dimensions d_i.

We use the replacement of the cylinder geometry with the decorated half-solid torus to open holes in all the walls between bubbles, once for each 2-cell in the cell-decomposition of M dual to the triangulation. The result is a sum over manifolds M''_ℓ labelled by the choice ℓ of lines for each 2-cell. See Figure 4.4.2.

We can enlarge the holes in the walls until they almost fill the corresponding 2-cells in the cell-decomposition of M dual to the triangulation. The manifold M''_ℓ looks like the 1-skeleton of the cell decomposition. Each 1-cell between 2-cells associated to lines L_i, L_j, L_k corresponds to a component of the manifold with the cross-section of a disk with three punctures where the three lines L_i, L_i and L_k lie. See Figure 4.4.3.

Finally, we can cut the 1-cells by using the Hilbert space associated to the disk with three boundary punctures. We can identify this Hilbert space with the space V_{ijk} of local operators available at a junction between defects L_i, L_j, L_k by the state-operator map. Inserting a complete basis of states across each 1-cell we decompose the three-manifold to a collection of three-balls, with a tetrahedral graph of line defects drawn on the boundary. See again Figure 4.4.3. The partition function for each decorated three-ball can be evaluated using the data of $\mathcal{C}$.

These three steps express the original partition function as a state sum involving ingredients which can be computed fully in terms of the category $\mathcal{C}$ of boundary line defects. If we take the

[13] This should be analogous to the statement that solid tori with a bulk line defect give a basis of the Hilbert space associated to a torus.

basis of boundary line defects L_i to consist of the simple objects in $\mathcal{C}$ we obtain the Turaev-Viro state sum.

The physical construction suggests that more general choices of collections of objects in $\mathcal{C}$ should also reproduce the same partition sum, as long as one picks the correct c_i coefficients to reproduce the correct sum of simples $\sum_i d_i L_i$.

The construction can be extended to more general three manifolds, including a variety of extra topological defects. It is very simple to add boundaries with $\mathfrak{B}$ boundary conditions and arbitrary graphs of boundary line defects drawn on the boundary. This leads to the same state sum over a cell complex with boundary. [14]

Another important example are bulk topological line defects, which turn out to be labelled by elements of the Drinfeld center $Z[\mathcal{C}]$ of the spherical fusion category. Concretely, an element of the Drinfeld center is an object in $\mathcal{C}$ equipped with choice of canonical junction as it crosses any other element. The data encodes the image of the bulk line when brought to the boundary.

The bulk lines Y_α "run" along the 1-skeleton of the construction, resulting into a modification of the vector spaces which appear along the 1-cells to the spaces $V_{\alpha;ijk}$ of local operators available at a junction between defects Y_α,L_i, L_j, L_k.

4.4.1 Symmetries

The Turaev-Viro construction can be refined to deal with three manifolds equipped with a non-trivial flat connection for a discrete group G [130]. The starting point of such a construction is a G-graded spherical fusion category $\mathcal{C}$, which consists of a collection of sub-categories $\mathcal{C}_g$ labelled by elements of G. Essentially, the flat connection is represented on the triangulation by group elements on the edges of the triangulation and the state sum decorates edges labelled by g with objects in $\mathcal{C}_g$.

The output of the construction is a topological field theory with a non-anomalous G global symmetry. The theory is equipped with a topological boundary condition where the G symmetry may be broken.

As before, we expect the converse to be true as well. A topological theory endowed with a non-anomalous G global symmetry and a topological boundary condition admits topological domain walls U_g labelled by G elements, which fuse according to the group law and admit canonical topological junctions. The boundary condition will be support categories $\mathcal{C}_g$ of line defects at which a U_g domain wall ends. Together, the $\mathcal{C}_g$ form a G-graded spherical fusion category which can be used to reconstruct the topological theory.

In the absence of a flat connection, we can decorate all edges with the identity element e and the state sum reduces to the Turaev-Viro construction for $\mathcal{C}_e$. On the other hand, if we gauge the G symmetry (with Dirichlet boundary conditions) the resulting theory is given by the Turaev-Viro construction for the whole $\mathcal{C}$, forgetting about the grading.

[14] It is also possible to include other topological boundary conditions $\mathfrak{B}'$ (or interfaces) to the construction, but this requires extra data to be provided, in the form of the $\mathcal{C}$-module ($(\mathcal{C},\mathcal{C})$-bimodule) category of domain lines between $\mathfrak{B}$ and $\mathfrak{B}'$.

A topological field theory T may also have a non-anomalous 1-form global symmetry. Concretely, that means that there is a set of bulk line defects B_a which bosons, fuse according to a group law and braid trivially with each other. [15]

A non-anomalous 1-form global symmetry allows one to couple the theory to a background 2-form flat connection. We will show how to include this coupling in the Turaev-Viro construction, by modifying the vector spaces attached to faces of the triangulation according to the value of the background 2-form. [16]

We will demonstrate that the anomaly of a fermionic $\mathbb{Z}_2$ 1-form symmetry can be eliminated in a canonical way if the three-manifold is endowed with a spin structure.

4.4.2 Review of the Turaev-Viro construction

We refer to [138, 147, 148] for a very clear discussion of the Turaev-Viro construction and of its relation to 3d topological field theories $\mathfrak{T}$ equipped with a topological boundary condition $\mathfrak{B}$.

We denote the spherical fusion category as $\mathcal{C}$, with a finite set I of (equivalence classes of) simple objects V_i. Remember that the space of local operators at a junction with outgoing line defects $V_1, \cdots V_n$ is $\mathrm{Hom}_{\mathcal{C}}\left(1, V_1 \otimes \cdots \otimes V_n\right)$.

The building blocks of the Turaev-Viro construction are the spherical fusion category evaluation maps which assigns a complex number $Z(\Gamma)$ to a planar graph Γ on a two sphere with edges labelled by objects and vertices labelled by morphism in $\mathcal{C}$. More precisely, if we label the two ends of a segment e by dual objects V_e and V_e^*, a vertex v joining edges $e_1, \cdots e_n$ is labelled by a morphism

$$\varphi_v \in \mathrm{Hom}_{\mathcal{C}}\left(1, V_1 \otimes \cdots \otimes V_n\right) \tag{4.4.1}$$

where V_i are the objects associated to v and e_i.

The 3d partition function depends on a choice of manifold M, possibly decorated by bulk line defects T_α labelled by objects Y_α in the Drinfeld center $Z(\mathcal{C})$. The manifold may admit a boundary ∂M, possibly decorated by boundary line defects V_i.

The first step in the calculation is to give a combinatorial description $\mathcal{M}$ of M, which is essentially a decomposition of M into convex polytopes, say tetrahedra. The partition function is computed as a sum over different ways to decorate the edges of $\mathcal{M}$ by simple objects l in $\mathcal{C}$ (reversing the orientation of an edge conjugates the objects):

$$Z[\mathcal{M}, \{Y_\alpha\}] = \sum_l \frac{\prod_i d_i^{\mathrm{edges}(\mathcal{M},i)}}{\mathcal{D}^{2\mathrm{vertices}(\mathcal{M})}} Z[\mathcal{M}, \{Y_\alpha\}, l] \tag{4.4.2}$$

where we count bulk vertices with weight 1 and boundary vertices with weight $1/2$ in $\mathrm{vertices}(\mathcal{M})$,

[15] Gauging a 1-form symmetry in $2{+}1$ dimensions should be a special case of the operation of anyon condensation, which can be done to a theory which includes a topological line A with sufficiently nice properties, generalizing the properties of $A = \oplus_a B_a$.

[16] Standard and 1-form global symmetries can be combined into the notion of 2-group. It would be interesting to integrate this possibility in our story.

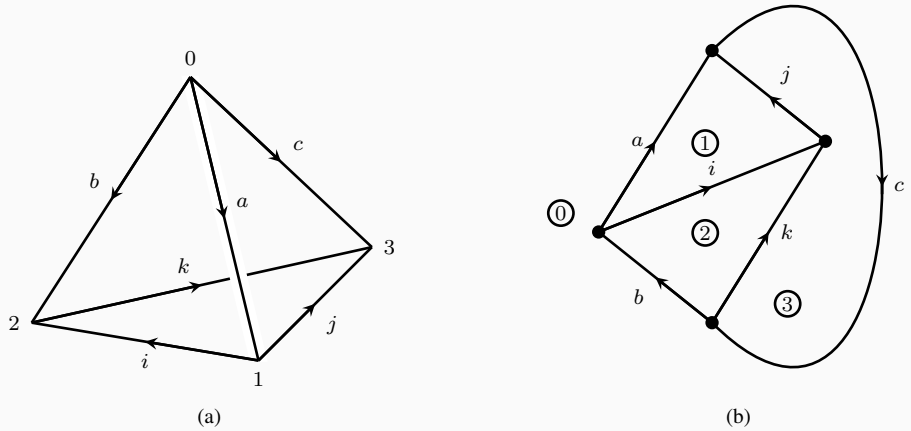


Figure 4.4.4: Left: The basic ingredient of the state sum is a tetrahedron decorated with lines. Right: The dual planar graph in the spherical fusion category. For clarity we denoted with circled numbers the tetrahedron vertices dual to each face.

and bulk edges with label i with weight 1 and boundary edges with label i with weight $1/2$ in $\mathrm{edges}(\mathcal{M}, i)$. The d_i and $\mathcal{D}$ are quantum dimensions and total dimension.

The partial partition functions $Z[\mathcal{M}, \{Y_\alpha\}, l]$ are computed by gluing together contributions of the individual polytopes of $\mathcal{M}$. Each face C of the triangulation with counterclockwise edges $e_1, \cdots, e_n$ is associated to a vector space

$$H(C, l) = \mathrm{Hom}_{\mathcal{C}}\left(1, l(e_1) \otimes \cdots \otimes l(e_n)\right) \tag{4.4.3}$$

and the partial partition function is valued in $H(\partial\mathcal{M}, l) = \prod_{C \in \partial\mathcal{M}} H(C, l)$. Pieces of a manifold are glued along faces C and $\overline{C}$ by contracting the elements of dual vector spaces $H(C, l)$ and $H(\overline{C}, l)$.

The contribution of an individual polytope is the output of the spherical fusion category evaluation map Z_Γ for a spherical graph Γ dual to the polytope. For example, a tetrahedron contribution is evaluated by the evaluation of a dual tetrahedral graph Γ, with vertices decorated by basis elements in $H(C, l)^*$. See Figure 4.4.4.

An important ingredient of the construction is a neat identity which holds for the evaluation maps. Consider a spherical graph Γ and cut it along the equator of the sphere. We can obtain two simpler graphs Γ_1 and Γ_2 by taking either half of Γ and bringing together the cut lines to a

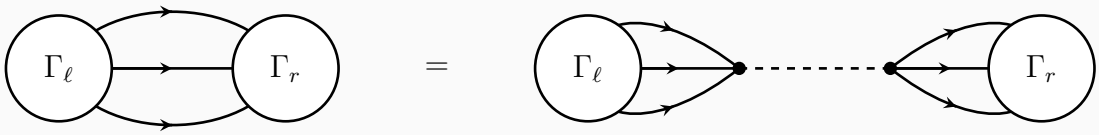


Figure 4.4.5: A crucial identity for spherical fusion category evaluation maps: a graph Γ (Left) can be split into two simpler graphs Γ_1 and Γ_2 (Right) with a sum over a complete set of dual local operators at the new junctions (dashed line).

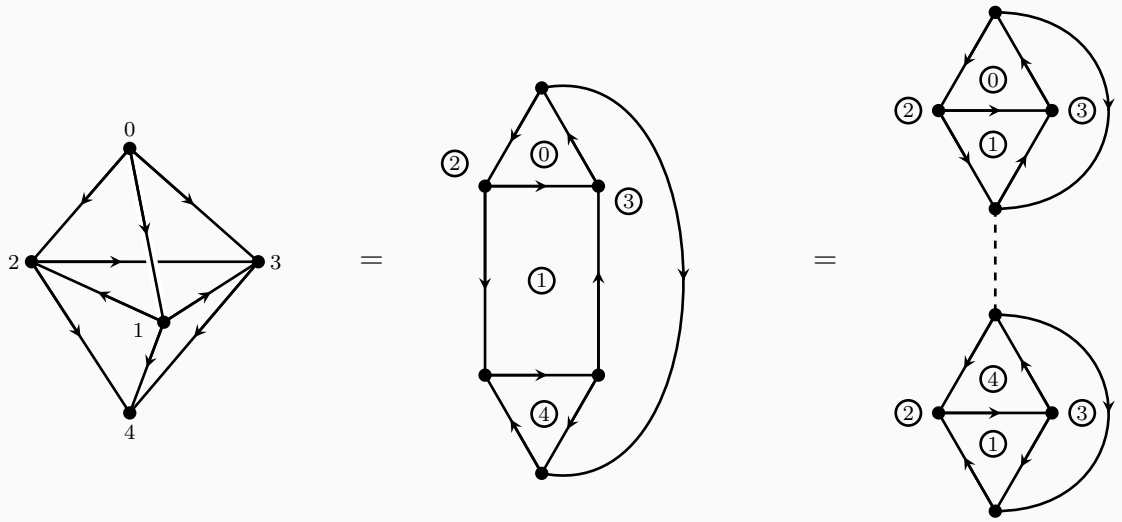


Figure 4.4.6: A triangular bi-pyramid (Right) can be obtained by gluing two tetrahedra. Correspondingly, the dual planar graph (Middle) can be obtained by fusing tetrahedral dual planar graphs along a pair of junctions (Right). For clarity, the faces dual to the original vertices are indicated by circled numbers.

common junction. The two new junctions support dual spaces of local operators V and V^*. Then

$$Z_\Gamma = Z_{\Gamma_1} \cdot Z_{\Gamma_2} \tag{4.4.4}$$

where the inner product denotes a sum over dual bases of local operators in V and V^*. See Figure 4.4.5.

This has a straightforward geometric interpretation: the polytope dual to Γ can be decomposed into the polytopes dual to Γ_1 and Γ_2, glued along the faces dual to the new junctions. The partition functions are glued by contracting the dual vector spaces associated to these faces. See Figure 4.4.6.

The bulk line defects affect the partition sum by modifying the vector spaces associated to the faces crossed by the lines. Essentially, they replace faces C with decorated faces D_α and

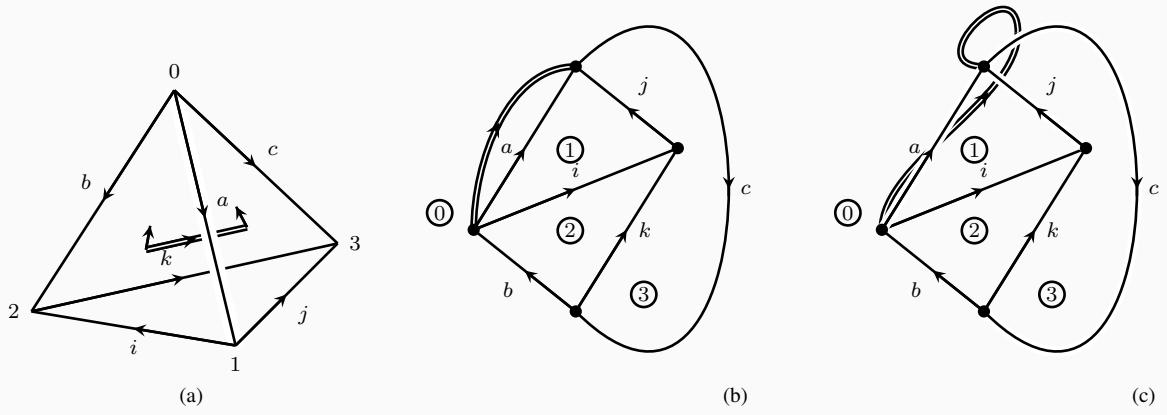


Figure 4.4.7: a) A tetrahedron with an extra quasi-particle transversing two faces. We indicated the choice of framing at each face. b) The dual planar graph in the spherical fusion category. We drew the center line along the simplest choice of path. Alternative paths which self-intersect or wind around the endpoints (c) would give answers which differ by framing phases

$H(C,l)$ with

$$H(D_\alpha, l) = \mathrm{Hom}_{\mathcal{C}}\left(1, Y_\alpha \otimes l(e_1) \otimes \cdots \otimes l(e_n)\right) \tag{4.4.5}$$

The computation of the contribution of a polyhedron with such modified faces involves adding an extra Y_α line attached to the appropriate vertices of Γ. If Y_α crosses some other line in Γ we can insert β_α there. The precise framed path followed by Y_α is immaterial because Y_α lies in the center. Changes of framing, though, change the answer appropriately. See Figure 4.4.7.

4.4.3 Adding a flat connection

Next, consider a G-graded spherical fusion category, a direct sum of sub-categories $\mathcal{C}_g$ with the property that $\mathcal{C}_g \otimes \mathcal{C}_{g'} \in \mathcal{C}_{gg'}$. We will denote the identity in G as 1. The Turaev-Viro construction applied to $\mathcal{C}_1$ gives a 3d TFT $\mathfrak{T}$ with boundary condition $\mathfrak{B}$, bulk lines in $Z[\mathcal{C}_1]$ and boundary lines in $\mathcal{C}_1$.

We can extend this theory to manifolds equipped with a G connection, simply representing the flat connection by edge elements g_e and by prescribing that an edge e of the triangulation is labelled by an object in $\mathcal{C}_{g_e}$ and building the partition sum as before. This endows $\mathfrak{T}$ with a global symmetry G.

Elements in $\mathcal{C}_g$ can be interpreted as lines lying at the intersection of the boundary $\mathfrak{B}$ with a U_g topological domain wall implementing the g symmetry. In general, there will not be any canonical choice of objects in $\mathcal{C}_g$ with trivial associators, meaning that the G global symmetry is broken at the boundary. [17]

Notice that bulk lines in the center $Z[\mathcal{C}_1]$ are not equipped with a canonical crossing through lines in $\mathcal{C}_g$. Physically, this corresponds to the fact that there is no canonical way for a bulk line to cross a U_g domain wall. [18]

The group cohomology construction of bosonic SPT phases is the simplest example of a G-graded Turaev-Viro partition sum, based on a G-graded category with a single (equivalence class of) simple object V_g in each $\mathcal{C}_g$ subcategory.

The evaluation of a tetrahedron of positive (negative) orientation produces directly the associator $\alpha_3^{\pm}$ and the partition sum immediately reproduces the SPT partition function.

4.4.4 Gauging standard global symmetries

Gauging the G global symmetry of $\mathfrak{T}$ should produce another topological field theory $\mathfrak{T}_G$. The partition function of $\mathfrak{T}_G$ should be obtained by summing the partition functions of $\mathfrak{T}$ over all possible choices of flat G connections. We expect this to coincide with Turaev-Viro applied to the whole $\mathcal{C}$, disregarding the G grading. This gives a sum over all flat connections rather than equivalence classes of flat connections, but the total quantum dimension should also change in such a way to compensate for that over-counting.

Notice that in the presence of a $\mathfrak{B}$ boundary the Turaev-Viro construction applied to the whole $\mathcal{C}$ will not sum over different choices of boundary lines. Correspondingly, the G gauge theory has Dirichlet b.c.: the flat connection is fixed at the boundary.

Gauging a theory with a standard Abelian global symmetry G should give a theory with a 1-form symmetry valued in the dual Abelian group G^*, generated by Wilson lines B_a. Using the definition of Wilson lines as center elements, we find that the insertion of a network of Wilson lines changes the sum over G flat connections α_1 by inserting a factor $e^{2\pi i \int \alpha_1 \cup \beta_2}$, where we are contracting the group elements in the flat connection α_1 with the characters in the background G^* 2-form flat connection β_2.

In order to see that in a fully explicit manner, it is useful to put a local order on the vertices of the triangulation and pick the first vertex in every face as a framing for the B_a lines. Then the decorated two-sphere graph associate to a tetrahedron has three B_a lines coming out of vertices into the first face, and one coming out of the 234 vertex towards the second face. In order to bring the B_a lines together and define a consistent graph, we need to have the B_a line from

[17] Depending on the $\mathcal{C}_g$ lines being dynamical or not in a UV completion of the theory, we can interpret the breaking as being spontaneous or explicit.

[18] It is possible to define a G action on the center $Z[\mathcal{C}_1]$, corresponding to surrounding a bulk line with an U_g domain wall, so that a canonical crossing morphism exists mapping $Y \otimes V_g$ to $V_g \otimes (g \circ Y)$. In the mathematical literature there is also the notion of G-center, corresponding to objects in $\mathcal{C}$ with a canonical crossing through objects in $\mathcal{C}_1$. These should correspond physically to bulk twist lines, at which U_g defects may end.

the 234 vertex cross the line between the first and second faces. The twisting factor contributes $e^{2\pi i(\widehat{\alpha}_1)_{12}(\widehat{\beta}_2)_{234}}$. This is precisely the contribution of a single tetrahedron to $e^{2\pi i \int \widehat{\alpha}_1 \cup \widehat{\beta}_2}$. [19]

4.4.5 Adding a 2-form flat connection

A theory endowed with a 1-form global symmetry Z can be coupled to a 2-form flat connection, say described by a 2-cocycle β_2 valued in Z. Concretely, β_2 should tell us which symmetry generators $B_a = (b_a, \beta_a)$ run through each face of the triangulation. Without loss of generality, we can take the lines entering each tetrahedron to join together at some interior point, thanks to $\delta\beta_2 = 0$. Gauge transformations on β_2 simply move around the lines or re-connect them.

Thus we have a Turaev-Viro construction of the partition function $Z[\beta_2]$: we simply replace the vector spaces $\mathrm{Hom}_{\mathcal{C}}\left(1, l(e_1) \otimes \cdots \otimes l(e_n)\right)$ with $\mathrm{Hom}_{\mathcal{C}}\left(1, b_a l(e_1) \otimes \cdots \otimes l(e_n)\right)$ and project the B_a lines for each tetrahedron to the surface, evaluating the corresponding graph as usual.

On general grounds, gauging a theory with a 1-form symmetry H should give a theory with a standard global symmetry valued in the dual Abelian group H^*. We have already described the process at the level of spherical fusion categories $\mathcal{C}_b$ and $\mathcal{C}^{\mathbb{Z}_2}$.

The evaluation of tetrahedra in the $\mathcal{C}_0^{\mathbb{Z}_2}$ category over objects inherited from $\mathcal{C}_b$ will precisely match the evaluation of tetrahedra in $\mathcal{C}_b$ coupled to a general β_2. Adding a $\mathbb{Z}_2$ flat connection α_1 simply adds the usual factor of $e^{\pi i \int \alpha_1 \cup \beta_2}$.

The only non-trivial step in identifying the Turaev-Viro partition sum of $\mathcal{C}^{\mathbb{Z}_2}$ as the result of gauging the 1-form symmetry of the Turaev-Viro partition sum of $\mathcal{C}_b$ is to observe that the sum over simple object of $\mathcal{C}_b$ of the images in $\mathcal{C}_0^{\mathbb{Z}_2}$

$$\sum_i d_i^{\mathcal{C}_b} (X_i)_0 \tag{4.4.6}$$

reproduces the correct sum of simple objects in $\mathcal{C}_0^{\mathbb{Z}_2}$.

4.4.6 Example: toric code

We can see now explicitly the equivalence between the toric code and $\mathbb{Z}_2$ gauge theory: the decoration I or P of the edges of the triangulation encodes a $\mathbb{Z}_2$ cochain and all tetrahedra contributions equal to 1. The total quantum dimension is 2, and the partition function is

$$\frac{1}{2^v} \sum_{\epsilon_1 | \delta\epsilon_1 = 0} 1 = \frac{1}{2} |H^1(M, \mathbb{Z}_2)|, \tag{4.4.7}$$

where v is the number of vertices. Remember that center of the category consists of objects $I = (I, \beta = 1)$, $e = (I, \beta_P = -1)$, $m = (P, \beta = 1)$, $\epsilon = (P, \beta_P = -1)$. The pre-factor 2^{-v} can be interpreted as the order of the $\mathbb{Z}_2$ gauge group .

[19] In general, one can interpret the Fourier transform kernel $e^{2\pi i \int \widehat{\alpha}_1 \cup \widehat{\beta}_2}$ as a very simple topological field theory with both a 1-form symmetry G^* and a standard symmetry G, generated by a spherical fusion category modelled on G and G^*-valued lines $(1, e^{2\pi i \chi \cdot h})$.

We see explicitly that adding an m line produces a vortex: the extra P line passing through a face breaks the flatness condition there. If we couple the system to the corresponding flat connection β_2^m, the partition sum becomes

$$\frac{1}{2^v}\sum_{\epsilon_1|\delta\epsilon_1=\beta_2^m} 1. \tag{4.4.8}$$

That is, it is equal to (4.4.7) if β_2^m is exact and equal to zero otherwise. This is somewhat boring, but consistent.

If we couple the system to a flat connection β_2^e, associated to the quasi-particle e, the partition sum becomes instead

$$\frac{1}{2^v}\sum_{\epsilon_1|\delta\epsilon_1=0} (-1)^{\int \epsilon_1\cup\beta_2^e} \tag{4.4.9}$$

The cup product emerges as before from the evaluation of the tetrahedron: with a canonical choice of framing, a single e line crosses a single edge as in Figure 4.4.8. If β_2^e is not exact, we can always find a dual 1-cocycle by which to shift ϵ_1 in order to switch the sign of the cup product and thus cancel all terms in pairs. (This is equivalent to the statement that the mod-2 intersection pairing on cohomology is non-degenerate.) Thus the sum is not-vanishing only if β_2^e is exact, in which case the integrand is a co-boundary and the sign drops out. This is consistent with the symmetry exchanging e and m.

Next, we can try to couple the system to a flat connection β_2^ϵ, associated to the fermion ϵ. The result should be anomalous, but yet instructive. The partition sum becomes

$$\frac{1}{2^v}\sum_{\epsilon_1|\delta\epsilon_1=\beta_2^\epsilon} (-1)^{\int \epsilon_1\cup\beta_2^\epsilon} \tag{4.4.10}$$

It is still true that if β_2^ϵ is not exact, we can always find a dual 1-cocycle by which to shift ϵ_1 in order to switch the sign of the cup product and cancel all terms in pairs. Thus the sum is not-vanishing only if β_2^ϵ is exact. If we write $\beta_2^\epsilon = \delta\lambda_1$, we can absorb λ into a shift of ϵ_1 to get

$$\frac{1}{2^v}\sum_{\epsilon_1|\delta\epsilon_1=0} (-1)^{\int \epsilon_1\cup\delta\lambda_1+\lambda_1\cup\delta\lambda_1} = (-1)^{\lambda_1\cup\delta\lambda_1}\frac{1}{2^v}\sum_{\epsilon_1|\delta\epsilon_1=0} 1 = \frac{1}{2}|H^1(M,\mathbb{Z}_2)||(-1)^{\lambda_1\cup\delta\lambda_1} \tag{4.4.11}$$

This is the toric code partition function (4.4.7) times $z_\Pi(\delta\lambda_1)$. In other words, the anomaly found in the toric code is precisely what we expected, at least for exact β_2^ϵ connections.

4.4.7 Gu-Wen Π-category

The $\mathbb{Z}_2$ gauge theory based on the $\widehat{G}$ SPT phase has a simple partition sum: lines are decorated by a fixed G flat connection which is lifted to a $\widehat{G}$ connection by some 1-cochain ϵ_1. The fusion

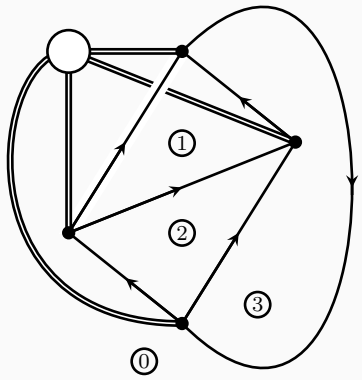


Figure 4.4.8: A canonical choice of framing for the dual tetrahedron graph dressed by 1-form symmetry generators.

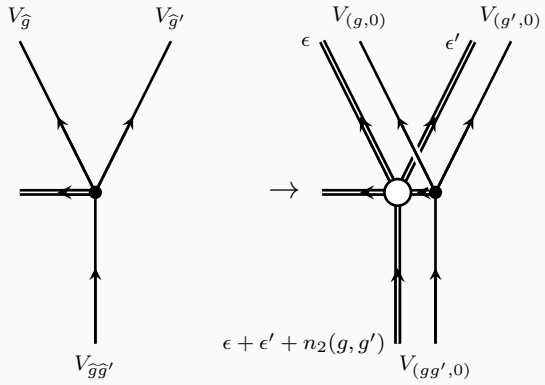


Figure 4.4.9: Gauge-fixing: A graphical representation of the partial gauge-fixing procedure used in computing the tetrahedron contribution of the $\widehat{G}$ Gu-Wen phase decorated by an extra 2-form flat connection β_2. Left: A choice of gauge is the same as a choice of basis vector in the space of junctions between line defects in the full $\widehat{G}$ category, possibly with an extra center line. Right: we identify $V_{(g,\epsilon)} \simeq V_{(1,\epsilon)} \otimes V_{g,0}$ and identify $V_{(1,\epsilon)}$ with the corresponding elements I or Π of the center. We then express a general junction canonically in terms of a choice of junction between line defects labelled by G elements. The double lines denote the center elements. The empty circle represents any choice of how to connect the center lines in a planar way.

rules imply that $\delta\epsilon_1$ equals the value of n_2 on faces. The partition sum is

$$\frac{1}{2^v} \sum_{\epsilon_1 | \delta\epsilon_1 = n_2} \prod \widehat{\alpha}_3. \tag{4.4.12}$$

We can pull out the ν_3 contribution and get

$$\frac{1}{2^v} \prod \nu_3 \sum_{\epsilon_1 | \delta\epsilon_1 = n_2} (-1)^{\int n_2 \cup \epsilon_1}. \tag{4.4.13}$$

From the fact that the mod-2 intersection pairing on cohomology is non-degenerate, one again deduces that the partition sum is non-zero only for G flat connections for which the pull-back of n_2 is exact, i.e. $n_2 = \delta\lambda_1$. Then we have

$$\frac{1}{2^v} (-1)^{\int \delta\lambda_1 \cup \lambda_1} \prod \nu_3 \sum_{\epsilon_1 | \delta\epsilon_1 = 0} 1 = z_\Pi(\delta\lambda_1) \tag{4.4.14}$$

Things become more interesting if we turn on the $\mathbb{Z}_2$ 2-form flat connection β_2 coupled to the fermionic 1-form symmetry generator. With appropriate gauge fixing, as in Figures 4.4.9 and

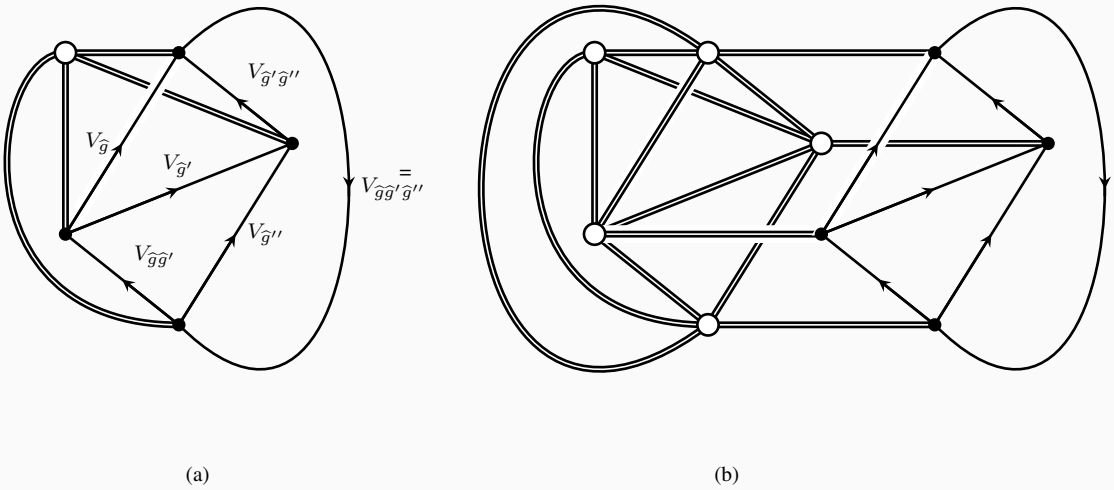


Figure 4.4.10: The computation of the tetrahedron contribution for the $\widehat{G}$ Gu-Wen phase decorated by an extra 2-form flat connection β_2.

4.4.10, the partition sum becomes

$$\frac{1}{2^v}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=n_2+\beta_2} (-1)^{\int n_2\cup\epsilon_1+\epsilon_1\cup\beta_2} \tag{4.4.15}$$

Now the partition sum is non-zero if the pull-back of n_2 is co-homologous to β_2. We can write $\beta_2 = n_2 + \delta\lambda_1$, shift ϵ_1 and obtain

$$\frac{1}{2^v}(-1)^{\int n_2\cup\lambda_1+\lambda_1\cup n_2+\lambda_1\cup\delta\lambda_1}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=0} (-1)^{\int n_2\cup\epsilon_1+\epsilon_1\cup n_2} \tag{4.4.16}$$

The sign in the sum is actually a boundary and drops out. We get

$$Z_f[\beta_2] = \frac{1}{2^v}(-1)^{\int n_2\cup\lambda_1+\lambda_1\cup n_2+\lambda_1\cup\delta\lambda_1}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=0} 1 \tag{4.4.17}$$

Crucially, the answer transforms under gauge transformations precisely as $z_\Pi(\beta_2)$. Furthermore, the product $Z_f[\beta_2]z_\Pi(\beta_2)$ simply coincides with the (spin-structure corrected) Gu-Wen SPT phase partition function: the $\prod \nu_3$ combines with the Gu-Wen grassmann integral in $z_\Pi(\beta_2)$. We accomplished our main objective.

4.4.8 State sums and spin-TFTs

We are ready to give our prescription for the Turaev-Viro partition sum of a spin TFT $\mathfrak{T}_s$ constructed from the spherical fusion category $\mathcal{C}_f$ for its shadow $\mathfrak{T}_f$.

Pick a spherical fusion Π-category $\mathcal{C}_f$. Define the decorated Turaev-Viro partition sum $Z_f[\beta_2]$ by adding fermionic Π lines through all faces where $\widehat{\beta}_2 = 1$. The lines will be framed as in the Gu-Wen Π-category calculation, going out of each dual vertex in the direction of the earliest face in the order. See Figure 4.4.8.

The amplitude for each tetrahedron is computed using the same projection on the two-sphere. The spin-TFT partition sum will be

$$Z[M;\mathfrak{T}_s] = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|}\sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} z_\Pi(\beta_2)Z[M;\mathfrak{T}_f;\beta_2] \tag{4.4.18}$$

The spin-structure dependence is hidden in $z_\Pi(\beta_2)$.

Notice that the calculation of $z_\Pi(\beta_2)$ can be integrated into the Turaev-Viro calculation. The Gu-Wen Grassmann integral can be given a super-vector space interpretation:

- We can assign fermion number 0 to the V_{ijk} spaces and fermion number 1 to the $V_{\Pi;ijk}$ spaces. This mimics the assignment of Grassmann variables θ_f, $\overline{\theta}_f$ to the faces of the triangulation.

- We can pick a specific order in the tensor product of face vector spaces which defined the Hilbert space associated to the boundary of a tetrahedron. The order mimics the choice of order for the Grassmann variables in the Gu-Wen integrand.
- When contracting pairs of dual vector spaces associated to each face, we keep track of the Koszul signs required to reorder the tensor product and bring the pair of spaces to be adjacent to each other. This mimics the choice of order for the Grassmann variables in the Gu-Wen integration measure

As the combinatorics of super-vector space tensor products reconstruct the Gu-Wen Grassmann integral, all which is left is the linear coupling of β_2 to the chain E of faces encoding the spin structure. We can write

$$Z[M;\mathfrak{T}_s] = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|} \sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} (-1)^{\int_E \beta_2} Z_{\text{super}}[M;\mathfrak{T}_f;\beta_2] \tag{4.4.19}$$

This has the form of a calculation in the spherical super-fusion category associated to the Π-category $\mathcal{C}_f$. It would be interesting to pursue this point further.

4.5 String net models

The same data which goes into the Turaev-Viro construction can also be used to give a local lattice Hamiltonian construction of the theory.

It is straightforward to give a physical motivation is analogous to the one we reviewed for the partition function. The basic step is to relate the Hilbert space $\mathcal{H}_\Sigma$ of the theory $\mathfrak{T}$ on some space manifold Σ and the Hilbert space $\mathcal{H}_{\Sigma'}$ on a manifold Σ' with an extra circular hole with boundary condition $\mathfrak{B}$.

In general, $\mathcal{H}_{\Sigma'}$ is larger than $\mathcal{H}_\Sigma$, but there will be maps i,π embedding $\mathcal{H}_\Sigma$ into $\mathcal{H}_{\Sigma'}$ and projecting $\mathcal{H}_{\Sigma'}$ to $\mathcal{H}_\Sigma$, which can be described in terms of three-manifolds with the topology of $\Sigma\times[0,1]$ minus a half-sphere. It is easy to see that $\pi\circ i$ is a multiple of the identity map, as it corresponds to a three-manifolds with the topology of $\Sigma\times[0,1]$ minus a contractible sphere with boundary condition $\mathfrak{B}$.

Thus we can describe $\mathcal{H}_\Sigma$ as the image in $\mathcal{H}_{\Sigma'}$ of the projector $i\circ\pi$, corresponding to a three-manifolds with the topology of $\Sigma\times[0,1]$ minus two half-spheres. Furthermore, the projector $i\circ\pi$ can be given a simple description in terms of the combination $\sum_i d_i L_i$ we encountered in explaining the Turaev-Viro construction, where the L_i are interpreted as closed line defects going around the circumference of the hole, acting on the Hilbert space.

More generally, we can triangulate Σ and carve out a circular hole at each vertex of the triangulation. Each hole will be associated to a separate projector $P_v = i_v\circ\pi_v$ and all projectors will commute. Thus $\mathcal{H}_\Sigma$ is obtained from $\mathcal{H}_{\Sigma'}$ as the ground state of a commuting projector Hamiltonian. See Figure 4.5.1.

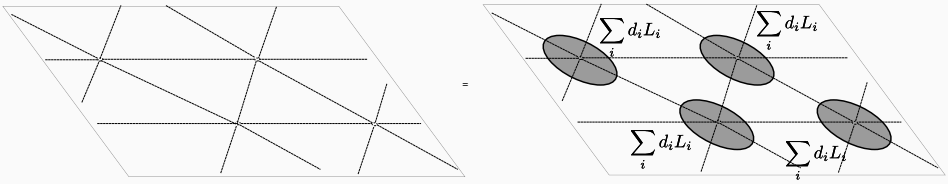


Figure 4.5.1: The Hilbert space associated to a surface Σ without holes (left) can be embedded in the Hilbert space associated to a surface Σ' with holes (right), as the image of projectors defined by the action of closed boundary lines $\sum_i d_i L_i$.

We can readily give a local description of $\mathcal{H}_{\Sigma'}$, by enlarging the holes until they almost fill the 2-cells dual to the vertices of the triangulation. We can continue out decomposition as we did for the partition function. At the next step we cut at 1-cells and replace Σ' with a collection Σ'' of disks associated to faces of the triangulation, with $\mathfrak{B}$ boundary conditions and three boundary lines for each disk. Each edge of the triangulation is associated with a pair of dual boundary lines in the disks corresponding to adjacent faces.

As long as the possible choices of lines run over all simples, or whatever other convenient basis of lines we employed in the state sum model, we can find maps i_e, π_e embedding $\mathcal{H}_{\Sigma'}$ in $\mathcal{H}_{\Sigma''}$. See Figure 4.5.2.

If the lines we selected are simple objects, the embeddings are actual isomorphisms, as both $i_e \circ \pi_e$ and $\pi_e \circ i_e$ turn out to be multiples of the identity as long as the pairs of simple lines corresponding to an edge of a triangulation are dual to each other. The Hilbert space $\mathcal{H}_{\Sigma''}$ is the tensor product of the corresponding morphism spaces V_{ijk} for each disk. [20] See again Figure 4.5.2.

The Hilbert space $\mathcal{H}_{\Sigma''}$ is the microscopic Hilbert space for the string net model. The decoration of edges by line defects L_i and the vector spaces V_{ijk} decorating the faces are the microscopic degrees of freedom. This is also the Hilbert space associated in the state sum construction to a boundary with topology Σ, triangulated and decorated by all possible simple simple lines.

The projectors P_v can be computed as the state sum partition function for a geometry M_v, consisting of a bi-pyramid made of tetrahedra to be glued on top of the triangles adjacent to v. Of course, the bi-pyramid contribution can be computed directly by the evaluation of the dual graph in the spherical fusion category. See Figure 4.5.3.

If we want to add a bulk quasi-particle Y as some point in Σ, say inside a face f of the triangulation, we simply replace V_{ijk} with the Hilbert space for a disk $V_{ijk;Y}$ with the extra bulk particle Y in the middle, as in the state-sum model. The projectors for the vertices around f are corrected by adding the quasi-particle to the state-sum calculation, going in and out the old and new f face. See Figure 4.5.4.

[20] If the lines L_i are not simple, the V_{ijk} are modules for the $\mathrm{Hom}(L_i, L_i)$ morphisms. Then $\pi_e \circ i_e$ is still a multiple of the identity but $i_e \circ \pi_e$ projects the naive tensor product of V_{ijk} spaces to the tensor product over $\mathrm{Hom}(L_i, L_i)$.

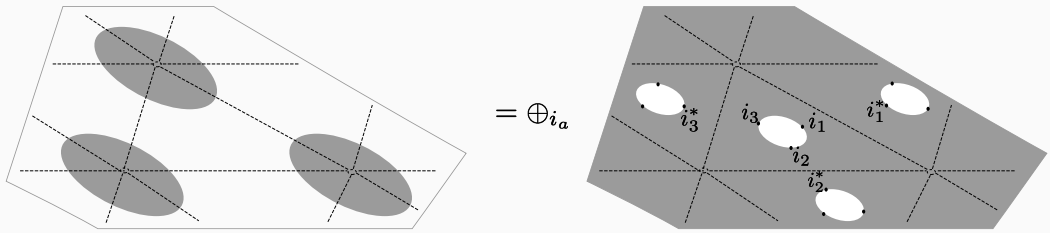


Figure 4.5.2: The Hilbert space associated to a surface Σ' with a regular arrangement of holes (left) can be identified with a direct sum of tensor product of Hilbert spaces associated to a collection of three-punctured disks. For clarity, we denote the boundary line defects L_{i_a} and $L_{i_a}^*$ simply as "i_a" and "i_a^*". Disks are in correspondence to faces of the triangulation. Pairs of dual defects are in correspondences to the edges of the triangulation.

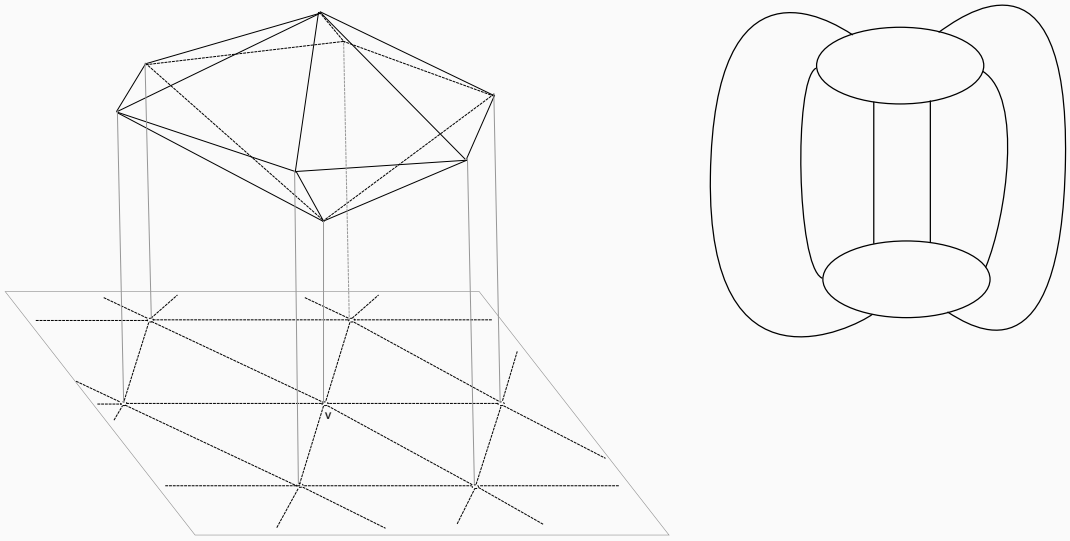


Figure 4.5.3: Left: The projector P_v can be computed as a state-sum model evaluation of a bi-pyramid. The bi-pyramid partition function is interpreted as a map from the dual of the vector space associated to the bottom faces to the vector space associated to the top faces. The action of the projector on the microscopic Hilbert space of the string-net model corresponds to gluing the bi-pyramid on top of the vertex v. Right: The bi-pyramid is computed in the spherical fusion category by an appropriate planar graph dual to the bipyramid surface. The oval faces are dual to the top and bottom vertices of the bi-pyramid.

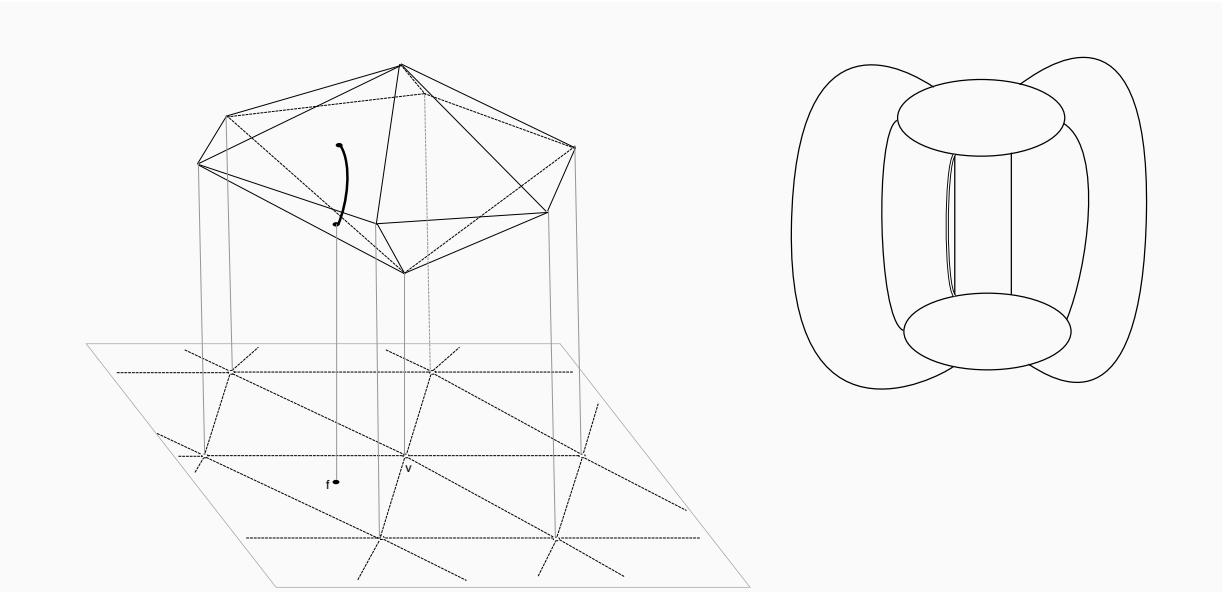


Figure 4.5.4: Left: The projector P_v in the presence of a quasi-particle at f can be computed as a state-sum model evaluation of a bi-pyramid with an extra bulk line. Right: The decorated bi-pyramid is computed in the spherical fusion category by an appropriate planar graph including the appropriate center line corresponding to the quasi-particle. The quasi-particle joins the junctions dual to the bi-pyramid faces above f. We selected a specific framing for the quasi-particle (which direction it exits and enters the junctions) it and kept it constant

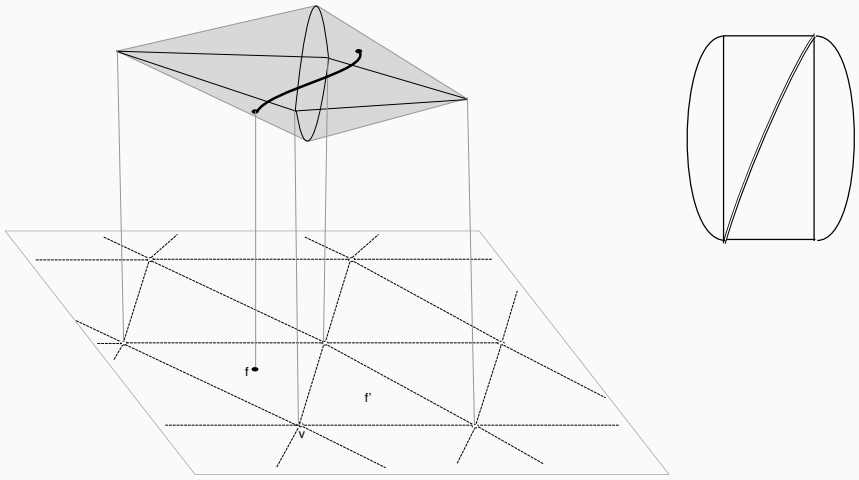


Figure 4.5.5: Left: A very economical description of the operator $U^Y_{f,f'}[\ell]$ for adjacent faces f and f'. The pillow-case geometry is the minimal way to interpolate between triangulations with quasi-particle insertions at f and f'. Right: The decorated pillowcase is computed in the spherical fusion category by an appropriate planar graph including the appropriate center line corresponding to the quasi-particle. The quasi-particle joins the junctions dual to the bottom bi-pyramid face above f and the top bi-pyramid face above f'. We selected a specific framing for the quasi-particle, pointing towards the vertex v.

We can also consider operators $U^Y_{f,f'}[\ell]$ corresponding to state-sum geometries which interpolate between the original triangulation and a triangulation where the quasi-particle Y has been moved to another face f' along some framed path ℓ in $\Sigma \times [0,1]$. See Figure 4.5.5 for a crucial example. Crucially, these operators will commute with the projectors in the Hamiltonian. Their algebra will mimic the topological properties of the corresponding quasi-particles.

4.5.1 Example: toric code

The string net model for the toric code, based on the category with objects I and P, is quite obviously a $\mathbb{Z}_2$ gauge theory: the configuration of edge decorations on the triangular lattice can be interpreted as a 1-cochain α_1 with values in $\mathbb{Z}_2$. The fusion constraint requires α_1 to satisfy $\delta\alpha_1 = 0$. Since all vector spaces V_{ijk} are one-dimensional or zero-dimensional, each allowed edge decoration corresponds to a basis vector in $\mathcal{H}_{\Sigma' s}$ which will be denoted $|\alpha_1\rangle$. The projector P_v simply acts as

$$|\alpha_1\rangle \to \frac{1}{2}|\alpha_1\rangle + \frac{1}{2}|\alpha_1 + \delta\lambda_0^v\rangle \tag{4.5.1}$$

where $\lambda_0^v(v') = \delta_{v,v'}$ is the 0-cochain supported on v. The product $\prod_v P_v$ projects to the subspace of gauge-invariant states.

It is interesting to decorate this picture with quasi-particles. The m quasi-particle at some face f simply deforms the fusion constraint at f. More generally, a configuration β_2^m of m quasi-particles imposes the constraint $\delta\alpha_1 = \beta_2^m$. The projectors are unchanged:

$$P_{\lambda_0}[\beta_2^m]|\alpha_1\rangle = \frac{1}{2}|\alpha_1\rangle + \frac{1}{2}|\alpha_1 + \delta\lambda_0\rangle \tag{4.5.2}$$

Similarly, the operators $U_{\lambda_1}^m$ which change the locations of m particles as $\beta_2^m \to \beta_2^m + \delta\lambda_1$ can be defined by combining individual U_e^m which act on the two faces adjacent to an edge e, built from a pillowcase geometry. The map only changes α_1 at the edge itself, and thus we have simply

$$U_{\lambda_1}^m|\alpha_1\rangle = |\alpha_1 + \lambda_1\rangle. \tag{4.5.3}$$

Clearly we have

$$U_{\lambda_1}^m U_{\lambda_1'}^m = U_{\lambda_1+\lambda_1'}^m. \tag{4.5.4}$$

Note that if λ_1 is exact, $\lambda_1 = \delta\mu_0$, we have

$$U_{\delta\mu_0}^m P_{\mu_0} = P_{\mu_0}. \tag{4.5.5}$$

Therefore on the image of $\prod_v P_v$ the operator $U_{\lambda_1}^m$ is invariant under $\lambda_1 \mapsto \lambda_1 + \delta\mu_0$. The ability to define operators $U_{\lambda_1}^m$ satisfying (4.5.4) and (4.5.5) indicates that the $\mathbb{Z}_2$ 1-form symmetry generated by the m particle is non-anomalous.

On the other hand, an e quasi-particle at a face f will not change the fusion constraint, but will change the form of the projectors for the vertices of f by adding some signs. Inspection of the dual bi-pyramid graph shows that the e center line only needs to cross other lines if it is framed towards v. In that case, we pick a sign -1 for each P line it crosses. See Figure 4.5.6.

For definiteness, we should pick a canonical framing for quasi-particles. For example, we can add a branching structure (local order of vertices) on the triangulation and frame quasi-particles towards the earliest vertex in each face. Then P_v has a sign only if v is the earliest vertex of f, appearing in front of the $|\alpha_1 + \delta\lambda_0^v\rangle$ term.

The projectors for a general configuration of e particles β_2^e become

$$P_{\lambda_0}[\beta_2^e]|\alpha_1\rangle = \frac{1}{2}|\alpha_1\rangle + \frac{1}{2}(-1)^{\int \lambda_0 \cup \beta_2^e}|\alpha_1 + \delta\lambda_0\rangle \tag{4.5.6}$$

Here we used the branching structure to define cup products. In a gauge theory language, the second term in the deformed projector inserts a Wilson line at the earliest vertex of the face f. The space of states in the presence of e particles is the image of the projector $\prod_v P_v[\beta_2^e]$. Here α_1 is closed because there are no m quasi-particles present.

The operators $U_{\lambda_1}^e$ which change the locations of e particles as $\beta_2^e \to \beta_2^e + \delta\lambda_1$ can be defined by combining individual U_e^e which act on the two faces adjacent to an edge e, built from a pillowcase geometry. See again Figure 4.5.6. The operator is diagonal in the $|\alpha_1\rangle$ basis:

$$U_{\lambda_1}^e|\alpha_1\rangle = (-1)^{\int \alpha_1 \cup \lambda_1}|\alpha_1\rangle \tag{4.5.7}$$

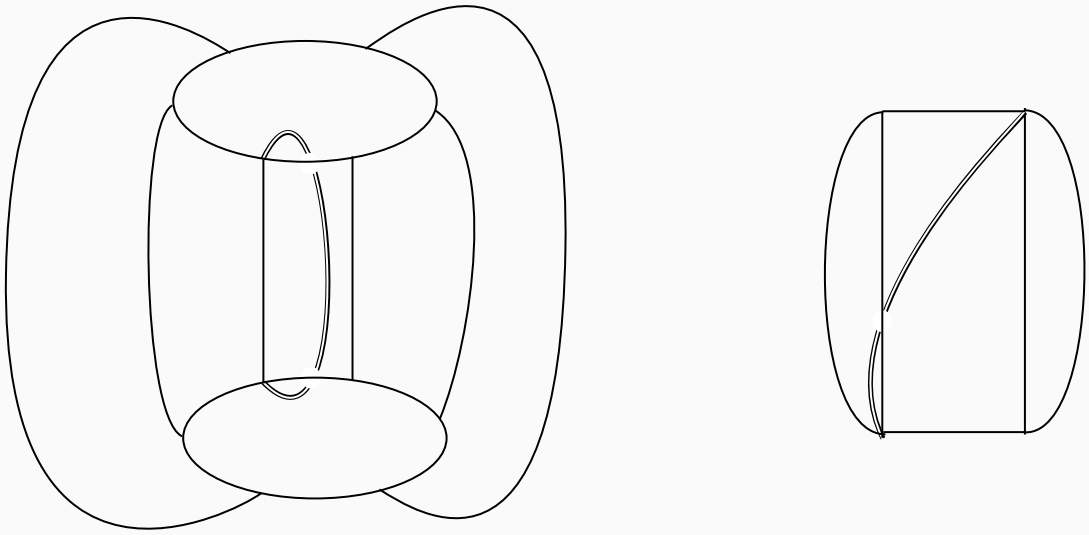


Figure 4.5.6: Left: The only bi-pyramid contributing non-trivial signs to P_v in the presence of an e particle. The quasi-particle is framed towards v and the decoration of the edges near v must flip from 1 to P or viceversa. Right: An example of a pillowcase contributing a non-trivial sign to U_e^e. There is a potential sign whenever the quasi-particle is framed towards an oval face, i.e. the earliest vertex of a face ("0") is opposite to the edge e. Then the sign measures the presence of P along the 01 edge of that face. This can be expressed as a cup product $\alpha_1 \cup \lambda_1^e$.

Again we have

$$U^e_{\lambda_1} U^e_{\lambda_1'} = U^e_{\lambda_1+\lambda_1'}, \quad U^e_{\delta\lambda_0} P^e_{\lambda_0}[\beta_2^e] = P^e_{\lambda_0}[\beta_2^e], \tag{4.5.8}$$

indicating that the $\mathbb{Z}_2$ 1-form symmetry generated by the e particle is non-anomalous.

Notice that the U^m and U^e operators do not commute, as expected from the braiding phase of e and m:

$$U^m_{\lambda_1} U^e_{\lambda_1'} = (-1)^{\int \lambda_1 \cup \lambda_1'} U^e_{\lambda_1'} U^m_{\lambda_1} \tag{4.5.9}$$

Neither is $U^e_{\lambda_1}$ invariant under $\lambda_1 \mapsto \lambda_1 + \delta\mu_0$ when $\delta\alpha_1 = \beta_2^m$ is non-vanishing. This indicates a mixed anomaly for the two $\mathbb{Z}_2$ 1-form symmetries.

Finally, we can insert ϵ particles. We both impose the constraint $\delta\alpha_1 = \beta_2$ and use the projectors

$$P_{\lambda_0}[\beta_2]|\alpha_1\rangle = \frac{1}{2}|\alpha_1\rangle + \frac{1}{2}(-1)^{\int \lambda_0 \cup \beta_2}|\alpha_1 + \delta\lambda_0\rangle \tag{4.5.10}$$

These correspond to a specific choice of framing of the ϵ line in the pillowcase geometry: it joins the two junctions along the most direct path compatible with the framing of the junctions, crossing a minimum number of other edges in the planar graph.

The operators which change the location of the ϵ particles around a single edge take the form

$$U^\epsilon_e|\alpha_1\rangle \equiv U^\epsilon_{\lambda_1^e}|\alpha_1\rangle = (-1)^{\int \alpha_1 \cup \lambda_1^e}|\alpha_1 + \lambda_1^e\rangle. \tag{4.5.11}$$

We can tentatively define a general operator rearranging ϵ particles:

$$U^{\epsilon}_{\lambda_1}|\alpha_1\rangle = (-1)^{\int \alpha_1 \cup \lambda_1}|\alpha_1 + \lambda_1\rangle \tag{4.5.12}$$

but the anomaly pops out as expected:

$$U^{\epsilon}_{\lambda_1} U^{\epsilon}_{\lambda'_1} = (-1)^{\int \lambda'_1 \cup \lambda_1} U^{\epsilon}_{\lambda_1 + \lambda'_1} \tag{4.5.13}$$

We can compute also

$$U^{\epsilon}_{\lambda_1} P_{\lambda_0}[\beta_2]|\alpha_1\rangle = \frac{1}{2}(-1)^{\int \alpha_1 \cup \lambda_1}|\alpha_1 + \lambda_1\rangle + \frac{1}{2}(-1)^{\int \lambda_0 \cup \beta_2}(-1)^{\int \delta\lambda_0 \cup \lambda_1}(-1)^{\int \alpha_1 \cup \lambda_1}|\alpha_1 + \delta\lambda_0 + \lambda_1\rangle \tag{4.5.14}$$

and check that it coincides with $P_{\lambda_0}[\beta_2 + \delta\lambda_1]U^{\epsilon}_{\lambda_1}|\alpha_1\rangle$, as expected.

Note that the pairing

$$\omega_\Sigma(\lambda_1, \lambda'_1) = \int_\Sigma \lambda'_1 \cup \lambda_1 \tag{4.5.15}$$

is symmetric modulo 2 for closed cochains λ_1, λ'_1 but not in general. Rather, one has

$$\omega_\Sigma(\lambda_1, \lambda'_1) - \omega_\Sigma(\lambda'_1, \lambda_1) = \int_\Sigma (\delta\lambda \cup_1 \lambda'_1 + \lambda_1 \cup_1 \delta\lambda'_1). \tag{4.5.16}$$

Thus some $U^{\epsilon}_{\lambda_1}$ and $U^{\epsilon}_{\lambda'_1}$ anti-commute rather than commute. Concretely, it is easy to check that U^{ϵ}_{e} and $U^{\epsilon}_{e'}$ anti-commute if the e and e' are adjacent to the same face and have the same orientation (induced by the branching structure) with respect to the face. They commute otherwise. Thus we cannot impose the constraint $U^{\epsilon}_{e} = 1$ on the states for all e.[21]

4.5.2 A fermionic dressing operator

To fix the sign problem in the algebra of the operators $U^{\epsilon}_{\lambda_1}$, let us place at each face f of the triangulation a pair of Majorana fermions γ_f and γ'_f. They are generators of a Clifford algebra $\mathrm{Cl}(2)$.

For any edge e, define

$$S_e = i\gamma_{f_L[e]}\gamma'_{f_R[e]} \tag{4.5.17}$$

where $f_{L,R}$ are the faces to the left and to the right of the edge (with respect to the branching structure orientation). We have the commutation relation

$$S_e S_{e'} = (-1)^{\int_\Sigma \delta_1(e) \cup \delta_1(e')} S_{e'} S_e, \tag{4.5.18}$$

where $\delta_1(e)$ is a 1-cochain supported on the edge e. In words: S_e operators commute unless the two edges share a face and have the same orientation with respect to the face, in which case they

[21] While it is true that the naive U^{ϵ}_{e} squares to 1 for all e, it is not true that the naive $U^{\epsilon}_{\lambda_1}$ squares to 1 for all λ_1. But this problem can be fixed by redefining $U^{\epsilon}_{\lambda_1}$ by suitable factors of i. The lack of commutativity for U^{ϵ}_{e} and $U^{\epsilon}_{e'}$ is a more serious problem.

anti-commute. We also have $S_e^2 = 1$ for all e. The crucial point is that the combined operators $S_e U_e^\epsilon$ commute with each other for all e.

Next we would like to define S_{λ_1} for a general 1-cochain λ_1, so that

$$S_{\lambda_1} S_{\lambda_1'} = (-1)^{\int_\Sigma \lambda_1' \cup \lambda_1} S_{\lambda_1 + \lambda_1'}. \tag{4.5.19}$$

We let

$$S_{\lambda_1} = (-1)^{\sum_{e<e' \in \Lambda} \int_\Sigma \delta_1(e) \cup \delta_1(e')} \prod_{e \in \Lambda} S_e. \tag{4.5.20}$$

Here Λ is the set of edges where $\lambda_1(e) = 1$, ordered in some way, and the product is ordered from right to left. The sign factor in (4.5.20) can also be described as follows: we include -1 for every pair of edges in Λ which share a face, have the same orientation (with respect to the branching structure), and whose order along the face agrees with the ordering of Λ. It is easy to check that S_{λ_1} does not depend on the choice or ordering of Λ and that (4.5.19) is satisfied. Thus, if we provisionally define $V_{\lambda_1} = S_{\lambda_1} U_{\lambda_1}^\epsilon$ on the tensor product of $\mathcal{H}_{\Sigma'}$ and the fermionic Fock space, we will have the relations

$$V_{\lambda_1} V_{\lambda_1'} = V_{\lambda_1 + \lambda_1'}. \tag{4.5.21}$$

A further issue which needs to be addressed is the behavior of V_{λ_1} under transformations $\lambda_1 \mapsto \lambda_1 + \delta\mu_0$, where μ_0 is a $\mathbb{Z}_2$-valued 0-cochain. A satisfactory generator of a $\mathbb{Z}_2$ 1-form symmetry must be invariant under such "symmetries of symmetries". Instead, in agreement with a general formula (4.2.40), we find

$$U_{\delta\mu_0}^\epsilon P_{\mu_0}[\beta_2] |\alpha_1\rangle = (-1)^{\int_\Sigma (\beta_2 \cup \mu_0 + \mu_0 \cup \beta_2)} P_{\mu_0}[\beta_2] |\alpha_1\rangle. \tag{4.5.22}$$

Thus the operator $U_{\delta\mu_0}^\epsilon$ is nontrivial even after projection to the physical Hilbert space $\mathcal{H}_\Sigma$.

Similarly, we can compute $S_{\delta\mu_0}$. To write down the answer, note that basis elements in the fermionic Fock space are naturally labeled by $\mathbb{Z}_2$-valued 2-cochains ν_2: for a given face f, the state $|\nu_2\rangle$ is an eigenstate of $i\gamma_f \gamma_f'$ with eigenvalue $(-1)^{\nu_2(f)}$. Then we get

$$S_{\delta\mu_0} |\nu_2\rangle = (-1)^{\int_\Sigma (\nu_2 \cup \mu_0 + \mu_0 \cup \nu_2 + C_2 \cup \mu_0) + \int_{\tilde{w}_2} \mu_0} |\nu_2\rangle. \tag{4.5.23}$$

Here $\tilde{w}_2$ is a particular $\mathbb{Z}_2$-valued 0-chain representing the 2nd Stiefel-Whitney class of Σ, and C_2 is a $\mathbb{Z}_2$-valued 2-cochain which takes value 1 on every face.

We can cancel all ν_2-dependent signs in $S_{\delta\mu_0}$ against β_2-dependent signs in $U_{\delta\mu_0}^\epsilon$ if we project to the subspace where $\nu_2 = \beta_2$. To eliminate state-independent signs, we choose a 1-chain E such that $\partial E = \tilde{w}_2$. As discussed in [118], such a E determines a spin structure on Σ. Then we define improved E-dependent Fock-space operators

$$S_{\lambda_1}^E = (-1)^{\int_\Sigma C_1 \cup \lambda_1 + \int_E \lambda_1} S_{\lambda_1}.$$

Here C_1 is a $\mathbb{Z}_2$-valued 1-cochain taking value 1 on every edge. It satisfies $\delta C_1 = C_2$. We also

define improved E-dependent dressed generators:

$$V^E_{\lambda_1} = U^\epsilon_{\lambda_1} S^E_{\lambda_1}.$$

On the projected Hilbert space, they satisfy

$$V^E_{\lambda_1} V^E_{\lambda_1'} = V^E_{\lambda_1+\lambda_1'}, \quad V^E_{\lambda_1+\delta\mu_0} = V^E_{\lambda_1}$$

for all 1-cochains λ_1 and all 0-cochains μ_0.

We can now define a commuting projector Hamiltonian for the phase $\mathfrak{T}_s$ on the projected Hilbert space as

$$H^E = \sum_e \frac{1}{2}(1 - V^E_e).$$

4.5.3 Fermionic dressing for general Π-categories

If the Drinfeld center of the fusion category $\mathcal{C}$ contains a fermion Π, we can define U^f_e operators in a manner completely analogous to the toric code example. The operator is evaluated by the pillowcase graphs framed as described above.

Because of the universality of factors associated to changes of framing and recombination of Π lines, we expect the same law as in the toric code.

$$U^f_e U^f_{e'} = (-1)^{\int \delta_1(e')_1 \cup \delta_1(e)} U^f_{e+e'} \tag{4.5.24}$$

More directly, we can compare the geometries associated associated to $U^f_e U^f_{e'}$ and $U^f_{e'} U^f_e$. The corresponding pairs of pillowcase geometries can be glued together to give the same geometry, but the framing of the center lines in the new geometry may or not agree. A careful analysis of all cases reproduces the expected multiplication law. See Figure 4.5.7. Therefore the same dressing by Majorana fermions will give commuting operators $U^s_e \equiv S_e U^f_e$.

From the toric code example, we also expect

$$U^f_{\delta\mu_0} = (-1)^{\int_\Sigma (\beta_2 \cup \mu_0 + \mu_0 \cup \beta_2)}. \tag{4.5.25}$$

This can be reproduced, with some effort, by counting the number of self-intersections of the Π lines obtained by merging the chain of pillowcase graphs for the sequence of edges around a single vertex.

Therefore the fermionic dressing ensures that $V^E_{\lambda_1} = U^f_{\lambda_1} S^E_{\lambda_1}$ is trivial when λ_1 is exact.

It follows that we can gauge the 1-form symmetry generated by Π by imposing the constraints $\beta_2(f) = \nu_2(f)$ for all faces f and $V^E_e = 1$ for all edges e in the tensor product of the Hilbert space of $\mathfrak{T}_f$ and the fermionic Fock space.

This is our final prescription for a microscopic Hamiltonian for $\mathfrak{T}_s$, built from the data of $\mathcal{C}_f$.

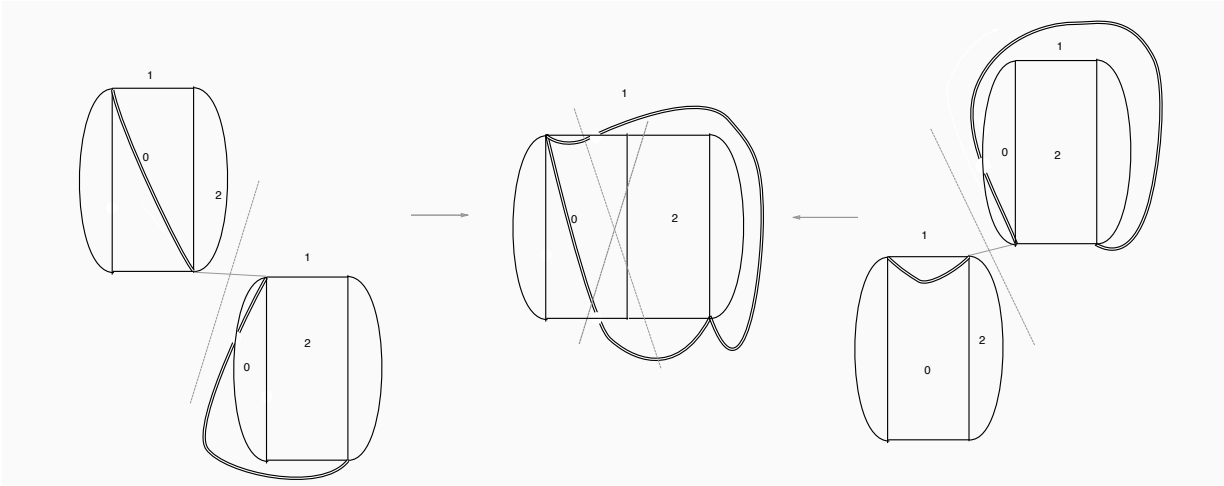


Figure 4.5.7: A comparison of $U_e^f U_{e'}^f$ and $U_{e'}^f U_e^f$ for edges e, e' adjacent to the same face f. Left and Right: The corresponding pairs of pillowcase geometries. Middle: The fused geometry. If the edges e and e' are the 01 and 12 edges of f, the center lines emanating from the fused junctions at f reconnect as shown in the middle. The two center lines can be identifies only up to a change of one unit of framing. Similar pictures with the outer face labelled by 0 or 2 match with no change of framing. Hence associativity fails only if e and e' are the 01 and 12 edges or vice-versa.

4.5.4 Including global symmetries

It is easy to extend the string-net construction to models with global non-anomalous symmetry G. Such a model is associated to a G-graded spherical fusion category $\mathcal{C} = \oplus_g \mathcal{C}_g$. In the Hamiltonian approach, G acts on-site and commutes with the Hamiltonian.

For example, we can model the lattice system on a discrete G-valued sigma model. That is, we put group variables g_v at vertices. Edges between vertices labeled by group elements g and g' are labeled by simple objects in $\mathcal{C}_{g'g^{-1}}$. We have commuting projectors $P_v^{g,g'}$ which change the group element at v from g to g', built from a state-sum bi-pyramid with central edge decorated by $g'g^{-1}$ [133].

Essentially by construction, adding G gauge fields on the edges and gauge-fixing the vertex group elements reproduces the string net model for the theory where G is gauged.

Bosonic SPT phases provide an obvious, well understood example of this construction. In this case $\mathcal{C}_g$ has a single simple object for all g, so the vertex variables g_v are the only variables. For an explicit expression for the projectors $P_v^{g,g'}$ see [133].

4.5.5 Example: the shadow of Gu-Wen phases

Our next example is the $\mathbb{Z}_2$ gauge theory associated to a $\mathbb{Z}_2$ central extension $\widehat{G}$ of a symmetry group G. In this case $\mathcal{C}_g$ has two simple objects which we denote $V_{g,\epsilon}$, $\epsilon \in \mathbb{Z}_2$, as before. They

fuse according to (4.3.10).

We decorate the vertices of the triangulation with group elements in G and the edges with $\mathbb{Z}_2$ variables ϵ_1 so that the edge objects are $V_{g^{-1}g',\epsilon_1(e)}$. The fusion rules imply that $\delta\epsilon_1 = n_2$, where n_2 is evaluated on the G group elements around each face.

The projectors $P_v^{g,g'}$ involve two terms, each computed as a product of $\widehat{\alpha}_3$. The two terms map a state with given $\mathbb{Z}_2$ decoration to two states with $\mathbb{Z}_2$ decorations which differ by a gauge transformation at v. This is expected, as we are defining an equivariant version of $\mathbb{Z}_2$ gauge theory.

With a bit of patience, one can disentangle the contribution of $\mathbb{Z}_2$ and G variables to the bi-pyramid graph of $P_v^{g,g'}$. For example, we can gauge-fix as in figure 4.4.9 every junction of the bi-pyramid graph. As we collapse the center line junctions to a single planar junction, we will get a factor of -1 from non-planar intersections. We can write these signs as $(-1)^{(n_2,\epsilon_1)}$, where the parenthesis indicates a certain bilinear pairing which is somewhat tedious to compute. This factor multiplies some expression $\tilde{P}_v^{g,g'}$ which depends on the G variables only.

We can populate the lattice with Π particles along some cocycle β_2. Now $\delta\epsilon_1 = n_2 + \beta_2$. We get deformed projectors $P_v^{g,g'}[\beta_2]$. Again, we can gauge-fix the junctions in a canonical way. The manipulation of Π lines will give some new signs $(-1)^{(n_2,\epsilon_1)+(\beta_2,\epsilon_1)'}$, multiplying the same $\tilde{P}_v^{g,g'}$ expression as before.

We can similarly compute the U_e^f operators which change β_2 in the two faces adjacent to e. The calculation involves the same pillowcase graphs as before. The G group elements do not change in the process, we only shift ϵ_1 by λ_1^e.

The contribution from Π lines crossing other lines is again $(-1)^{\int \epsilon_1 \cup \lambda_1^e}$, as in the toric code. Thus the fermionic dressing proceeds as before.

Chapter 5

Unoriented 3d TFTs

5.1 Introduction

It is an age-old problem to provide a complete definition of quantum field theories. A part of the problem is to understand on what kinds of manifolds can we put a quantum field theory. For instance, we can ask the following question: Given a theory that can be defined on orientable manifolds, what sort of extra data do we need in order to extend the definition of the theory to non-orientable manifolds? First of all, such an extension may not be possible. For instance, if the theory has a framing anomaly, then it will not be well-defined on non-orientable manifolds. This was recently explained in a footnote of [149]. Second, if such an extension is possible, then it need not be unique. That is, there can be different unoriented theories which reduce to the same oriented theory on orientable manifolds. We will see plenty of examples like this in this chapter.

On a non-orientable manifold, we can choose a consistent orientation everywhere if we remove a locus homologous to the Poincare dual of first Stiefel-Whitney class w_1. The induced local orientation flips as we cross this locus. In order to be able to define an unoriented theory in terms of the data of the oriented theory, we need the existence of orientation reversing codimension one defects which we place along this locus. These orientation reversing defects implement orientation reversing symmetries akin to the orientation preserving codimension one defects which implement a global symmetry transformation [150]. These defects can be placed on top of each other forming the structure of a group G with a homomorphism $\rho : G \to \mathbb{Z}_2$ whose kernel G_0 is the global symmetry group of the theory. The set $G_1 = G - G_0$ parametrizes the orientation reversing symmetries.

In this chapter, we explore the consequences of the existence of such orientation reversing defects in the context of 3d TFTs which admit a topological boundary condition. We restrict ourselves to the case in which the structure group of the TFT can be decomposed as $O(3) \times G$ for a finite global symmetry group G. In such cases, the properties of orientation reversing defects allow us to propose a generalization of Turaev-Viro state-sum construction of 3d TFTs [151, 138] to the unoriented case.[1] We check that this proposal indeed defines a 3d unoriented TFT. From now on, whenever we say "unoriented TFT", we mean this particular structure group.

[1] Historically, the original construction due to Turaev and Viro was based on a modular tensor category treated as a spherical fusion category. This construction produced theories which could be defined on an unoriented manifold.

For an oriented 3d TFT $\mathfrak{T}$ with global symmetry G, the input data for the construction is a G-graded spherical fusion category $\mathcal{C}$. We will find that an unoriented 3d TFT $\widetilde{\mathfrak{T}}$ extending $\mathfrak{T}$ is constructed in terms of a G-graded "twisted" spherical fusion category $\widetilde{\mathcal{C}}$ where $\mathcal{C}$ is embedded as a subcategory of $\widetilde{\mathcal{C}}$. In terms of the data of $\widetilde{\mathcal{C}}$, we give a prescription to construct the partition function of $\widetilde{\mathfrak{T}}$ on any (possibly non-orientable) 3-manifold.

We also apply these ideas to 3d Pin$^+$-TFTs (*i.e.* TFTs with structure group Pin$^+(3) \times G$) which are a generalization of 3d Spin-TFTs. Spin-TFTs are "fermionic" analogs of ordinary oriented TFTs as they are sensitive to the spin structure of the underlying orientable manifold. To define fermions on an unorientable manifold, we need to choose either a Pin$^+$-structure or a Pin$^-$-structure on the manifold. Correspondingly, the natural unoriented generalizations of Spin-TFTs are Pin$^+$-TFTs and Pin$^-$-TFTs.

In [3], a recipe was given to construct a 3d Spin-TFT from an ordinary 3d TFT with an anomalous $\mathbb{Z}_2$ 1-form symmetry. This ordinary TFT $\mathfrak{T}_f$ was called the shadow of the corresponding Spin-TFT $\mathfrak{T}_s$. The idea was to use a kernel TFT $\mathfrak{T}_k$ to connect the shadow theories with their spin counterparts. $\mathfrak{T}_k$ is a Spin-TFT with an anomalous $\mathbb{Z}_2$ 1-form symmetry. The diagonal $\mathbb{Z}_2$ 1-form symmetry in the product theory $\mathfrak{T}_f \times \mathfrak{T}_k$ is non-anomalous. This non-anomalous 1-form symmetry is then gauged to obtain the spin TFT $\mathfrak{T}_s$.

We extend their recipe by constructing shadows for Pin$^+$-TFTs. The Pin$^+$-shadows correspond to theories with anomalous $\mathbb{Z}_2$ 1-form symmetry and a certain time-reversal anomaly in the presence of a background 2-connection for the $\mathbb{Z}_2$ 1-form symmetry. The Pin$^+$-kernel TFT $\mathfrak{T}_k^+$ has a corresponding time-reversal anomaly which cancels the anomaly of the shadow. Hence, the resulting Pin$^+$-TFTs are time-reversal invariant and can be put on non-orientable manifolds without any ambiguity [152].

As an application, we construct a large class of Pin$^+$-SPT phases with global symmetry G. SPT phases are TFTs which are invertible under the product operation on TFTs. In the condensed matter literature, these are referred to as fermionic SPT phases protected by $G \times \mathbb{Z}_2^T$ with $T^2 = (-1)^F$. In the case when G is trivial, cobordism hypothesis predicts two Pin$^+$-SPT phases forming a $\mathbb{Z}_2$ group structure [128]. Our construction reproduces both of these SPT phases along with the $\mathbb{Z}_2$ structure.

This chapter is organized as follows. In section 6.5.31, we propose a Turaev-Viro construction for unoriented 3d TFTs. In section 6.5.32, we provide a construction of Pin$^+$-TFTs in terms of ordinary unoriented TFTs with a $\mathbb{Z}_2$ 1-form symmetry which is anomalous and has a mixed anomaly with time-reversal symmetry. In section 6.5.33, we construct a large class of Pin$^+$-SPT phases with global symmetry G and reproduce the $\mathbb{Z}_2$ group of Pin$^+$-SPT phases in the case of trivial G. In section 5.5, we present our conclusions and comment on future directions which include a strategy to classify all Pin$^+$-SPT phases with global symmetry G.

This construction was generalized to arbitrary spherical fusion categories but such theories could only be defined on oriented manifolds. It is this latter construction that we call "oriented Turaev-Viro construction" in this chapter. This chapter presents a further generalization of this setup which we call "unoriented Turaev-Viro construction". Our construction can be used to construct any unoriented 3d TFT with a topological boundary condition.

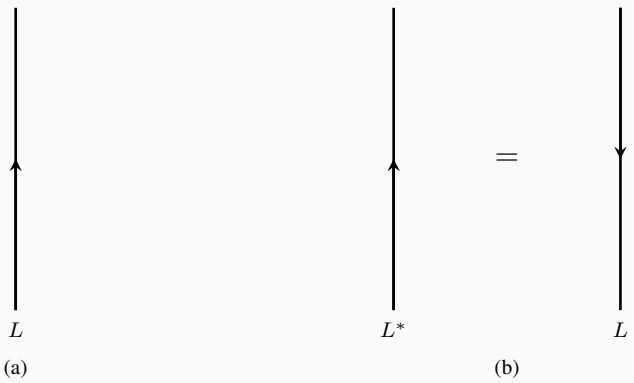


Figure 5.2.1: (a) A boundary line L. (b) The *dual* line L^* is defined by reversing the orientation of L.

5.2 Turaev-Viro construction

For an exhaustive review and physical understanding of the Turaev-Viro state-sum construction of oriented 3d TFTs, the reader is referred to [3]. In this section, we first review relevant aspects of this construction. Then, we propose a generalization of the construction to the unoriented case. We also provide a physical understanding of our proposal in terms of orientation reversing defects. We close the section by discussing invertible unoriented TFTs with global symmetry G, or in other words bosonic SPT phases protected by $G \times \mathbb{Z}_2^T$.

A reader only interested in the Turaev-Viro construction of unoriented 3d TFTs is referred to subsections 5.2.3 and 5.2.4.

5.2.1 Boundary line defects and spherical fusion category

In general, a boundary condition $\mathfrak{B}$ allows one to define a TFT $\mathfrak{T}$ on a manifold M with boundary B by placing $\mathfrak{B}$ on the boundary. The boundary condition is called topological if topological deformations of M (including topological deformations of B) leave the partition function of $\mathfrak{T}$ on M invariant.

Turaev-Viro procedure constructs an oriented 3d (unitary) TFT $\mathfrak{T}$ from the knowledge of a topological boundary condition $\mathfrak{B}$ of $\mathfrak{T}$ [153]. $\mathfrak{T}$ can be recovered from any one of its topological boundary conditions. For simplicity, we will assume that $\mathfrak{T}$ has a one-dimensional Hilbert space on S^2. The Turaev-Viro construction for such a TFT $\mathfrak{T}$ is phrased in terms of a (unitary) spherical fusion category $\mathcal{C}$.

The objects of $\mathcal{C}$ are line defects living on $\mathfrak{B}$. Such line defects are specified by a label L and an orientation along the line corresponding to L. If a line defect with a certain orientation is denoted as an object L in $\mathcal{C}$, the same line defect with opposite choice of orientation is denoted as the dual object L^*. See Figure 5.2.1.

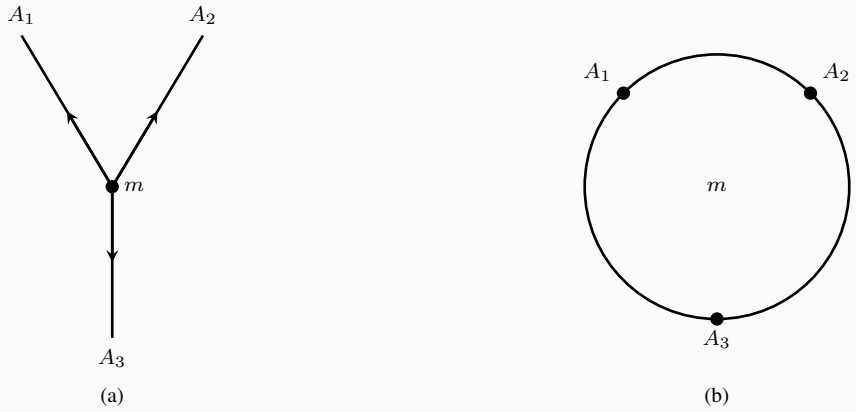


Figure 5.2.2: A morphism m between outgoing lines A_1, A_2 and A_3 corresponds to a state m in the Hilbert space on a disk with boundary punctures A_1, A_2 and A_3. Consider on a hemisphere geometry with a boundary on the spherical part and the disk shown in (b) being the cross-section. The state shown in (b) is produced on the cross-section if the boundary has the graph shown on (a) inserted on it such that A_i end on their respective punctures.

The morphisms m_{AB} from A to B in $\mathcal{C}$ are local operators living between two boundary lines. Thus, m_{AB} form a vector space. This vector space can also be identified with the Hilbert space of states on the disk with boundary punctures corresponding to A^* and B. Similarly, the local operators living at the junction of multiple outgoing lines A_i is the space of states on disk with boundary punctures corresponding to A_i. The space of states can be generated by placing a hemispherical cap on which the lines A_i emanate from a point on the boundary of the cap and go to their respective punctures. See Figure 5.2.2.

The composition of morphisms corresponds to fusion of local operators along the line. There is a tensor product corresponding to fusion of lines as they are brought together. There are also canonical associator, evaluation and coevaluation maps which physically correspond to placing the lines in a certain fashion and fusing them. See Figures 5.2.3 and 6.3.4. Using these canonical morphisms, we can assign a morphism m from $\otimes_i A_i$ to $\otimes_j B_j$ to any planar graph Γ of boundary line defects (with local operators at their junctions) such that Γ has incoming lines A_i and outgoing lines B_j. The canonical morphisms satisfy certain identities which guarantee that a topologically equivalent graph Γ' evaluates to the same morphism m.

Consider vacua i of $\mathfrak{B}$ which can be characterized by the expectation value of a line L_i. Such lines are called simple lines. Morphism space from L_i to L_j is empty for $i \neq j$ and is one-dimensional for $i = j$. The space of local operators living on L_i can be identified as $\mathbb{C}$ because there is a canonical identity operator living on L_i. Every line L can be written as a sum of simple lines $L = \oplus n_i L_i$ where n_i denotes the multiplicity of the simple line L_i in the sum. The identity line 1 can be treated as a special simple line which can be inserted anywhere without changing any answers. The duals of simple lines are simple as well.

Turaev-Viro construction uses $\mathcal{C}$ as an input and produces the partition of $\mathfrak{T}$ on any oriented manifold M as the output. We will describe the construction in a very hands-on fashion in the

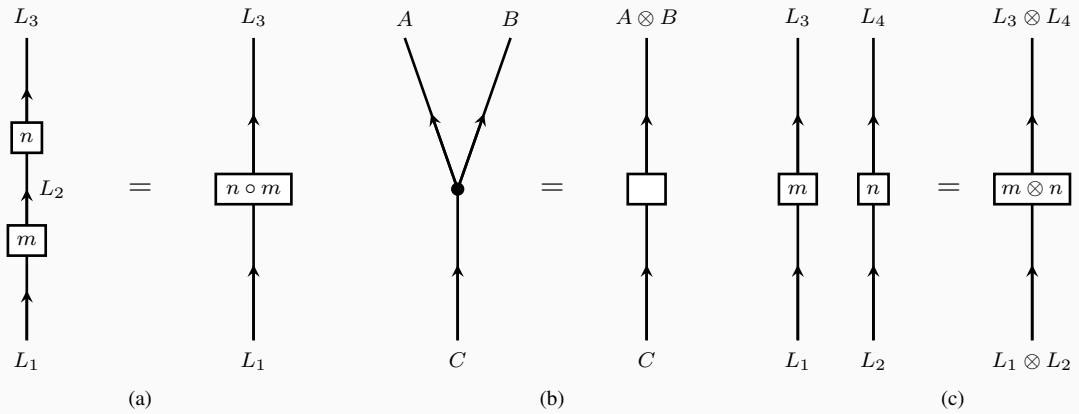


Figure 5.2.3: (a) Composition of morphisms. The box is our alternative notation for a morphism. (b,c) Tensor product of objects and morphisms.

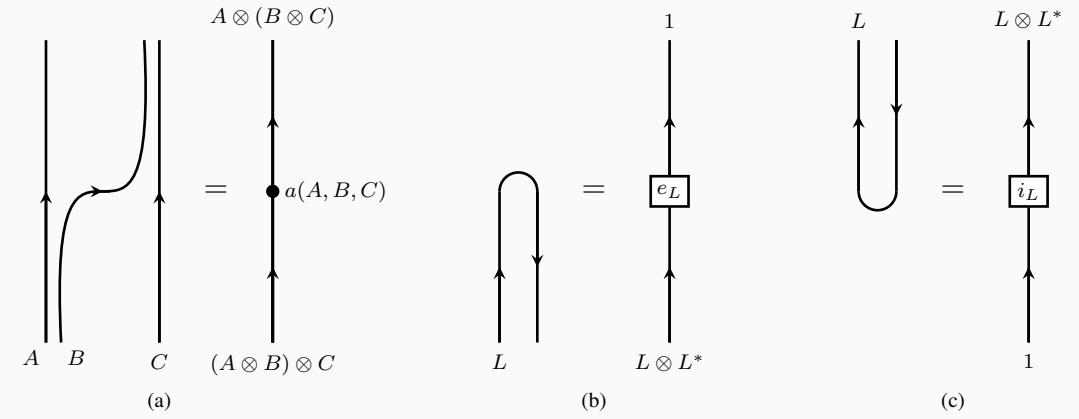


Figure 5.2.4: Canonical maps: (a) Associator $a(A, B, C)$, (b) Evaluation e_L, and (c) Coevaluation i_L.

Figure 5.2.5: (a) Graphical representation of the chosen basis. (b) Graphical representation of the dual basis.

next subsection. We will see that the basic object in the construction is a graph Γ in $\mathcal{C}$ drawn on the sphere. Γ can be projected down to a closed graph Γ_p drawn on the plane. Γ_p constructs a morphism from identity line to itself which evaluates to a definite number. This number is the partition function Z_Γ of $\mathfrak{T}$ on a 3-ball along with a network of boundary lines (and local operators at their junctions) Γ inserted at the boundary 2-sphere. The word "spherical" in spherical fusion category corresponds to certain axioms which guarantee that different projections to the plane evaluate to the same number.

This construction can be easily generalized to TFTs with a global symmetry group G. The symmetry manifests itself in the existence of codimension one topological defects U_g labeled by $g \in G$. Going across the locus of U_g implements a symmetry transformation on the system by g. These defects fuse according to the group law and can end on $\mathfrak{B}$ giving rise to new lines at the junction. Thus the category of boundary lines living on $\mathfrak{B}$ becomes graded by G, *i.e.* $\mathcal{C} = \oplus_g \mathcal{C}_g$.

5.2.2 Oriented Turaev-Viro

Let's look at the decomposition of the tensor product of two simple lines $L_i \otimes L_j = \oplus n_{ij}^k L_k$. This means that there is a n_{ij}^k dimensional space of morphisms from L_k to $L_i \otimes L_j$. We pick a basis of this space labeled by α. Similarly we pick a dual basis for the space of morphisms from $L_i \otimes L_j$ to L_k which we also label by α. See Figure 5.2.5. The completeness of the basis can be written graphically as in Figure 5.2.6(b). [2] We can transform to a basis labeled by α'. We denote the unitary matrix corresponding to the transformation as $(U_k^{ij})_{\alpha'\alpha}$. See Figure 5.2.7.

The associator induces an isomorphism between the morphism space from L_l to $(L_i \otimes L_j) \otimes L_k$ and the morphism space from L_l to $L_i \otimes (L_j \otimes L_k)$. In terms of our chosen basis, this isomorphism can be captured in terms of F-symbols $(F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)}$ which are defined in Figure

[2] We are assuming that the quantum dimensions of all L_i is 1 for simplicity. For generic quantum dimensions, we have to normalize these morphisms appropriately so that any graph of line defects and its topological deformations define the same morphism.

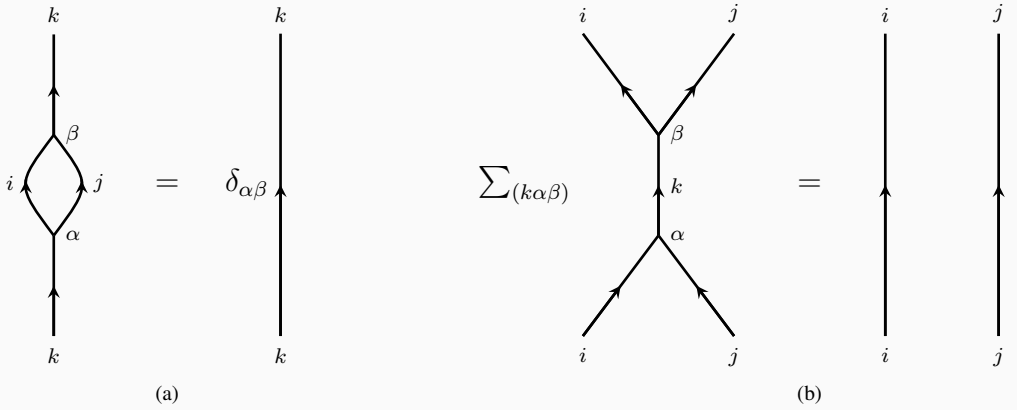


Figure 5.2.6: (a) Graphical representation of the fact that the two basis are dual to each other. (b) Completeness of the basis. Here, sum over $(k\alpha\beta)$ represents a sum over all such consistent triples.

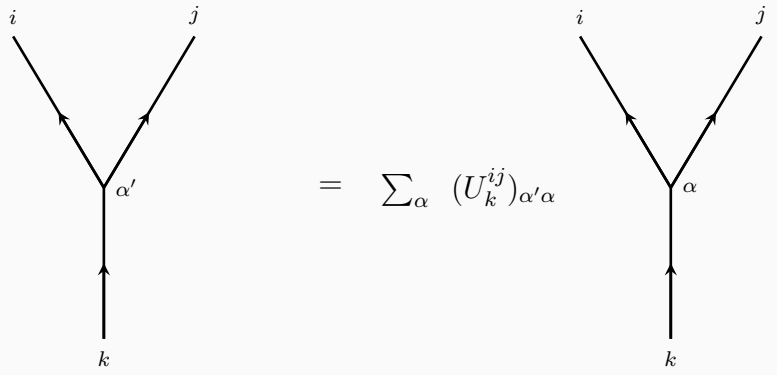


Figure 5.2.7: A change of basis via a unitary matrix.

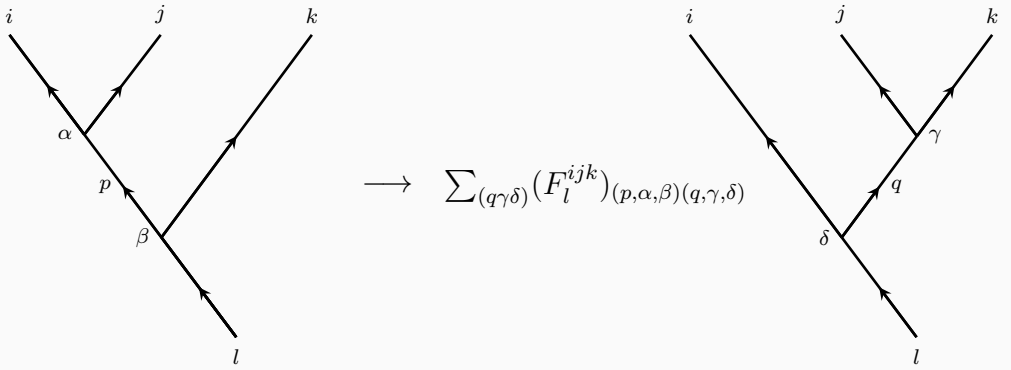
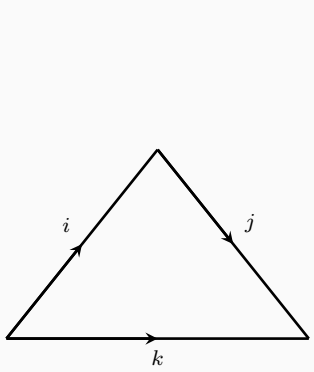
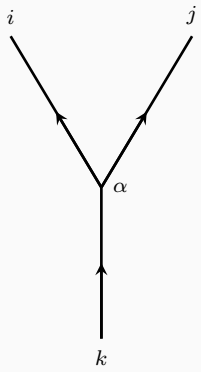


Figure 5.2.8: Definition of F-symbols.

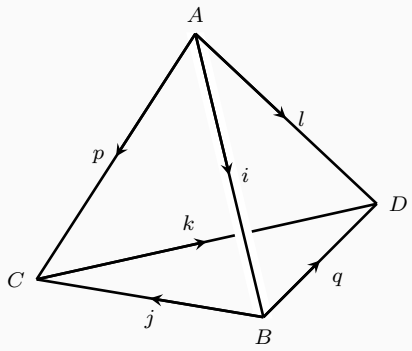


(a)

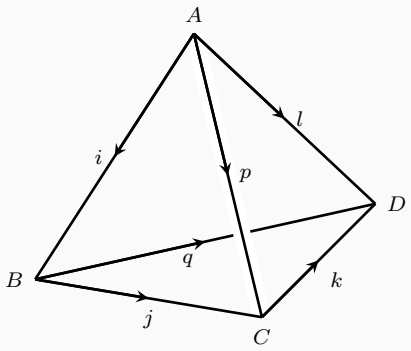


(b)

Figure 5.2.9: Labeling of faces.



(a)



(b)

Figure 5.2.10: A tetrahedron can have two chiralities - (a) positive and (b) negative. There is a label attached to every face but we don't show it in the figure for brevity.

6.3.8. Under a change of basis, F-symbols transform as

$$(F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} \to (F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} (U_q^{jk})^*_{\gamma'\gamma} (U_l^{iq})^*_{\delta'\delta} (U_p^{ij})_{\alpha'\alpha} (U_l^{pk})_{\beta'\beta} \tag{5.2.1}$$

We are now ready to describe Turaev-Viro prescription for the partition function of $\mathfrak{T}$ on a manifold M. Pick a branched triangulation T of M. A branched triangulation requires an ordering $>$ of the vertices of the triangulation. To an edge e between vertices a and b, a branched triangulation assigns a direction $a \to b$ if $a > b$. The G-connection α_1 on M assigns an element g_e of the group G to each directed edge e. We now label each directed edge e by a simple element living in $\mathcal{C}_{g_e}$. Pick a face f of T. Rotating it and flipping it, f looks like as shown in Figure 5.2.9(a). Then, we label f by some α corresponding to a morphism as shown in the Figure 5.2.9(b). Thus we have a labeling of edges and faces of a branched triangulation. Call one such labeling as $\tilde{l}$.

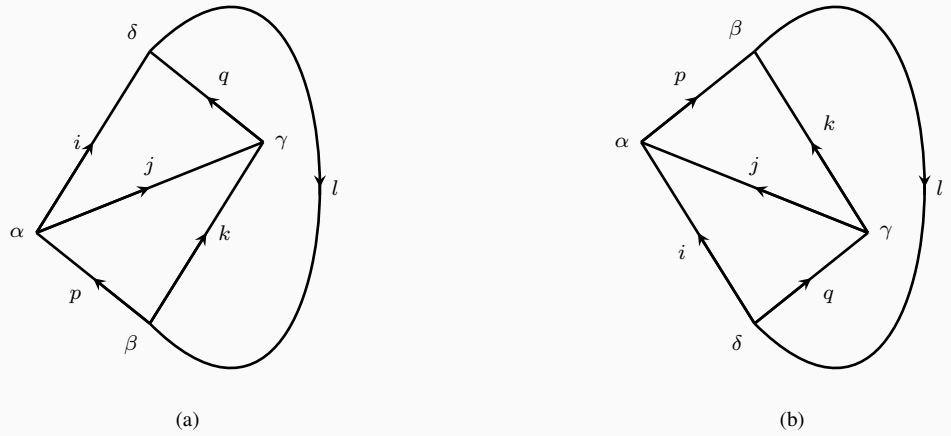


Figure 5.2.11: Graph attached to a tetrahedron: (a) positive chirality and (b) negative chirality.

Pick a tetrahedron t in $\tilde{l}$. To each t we assign a planar graph Γ_t in $\mathcal{C}$ and we call such a graph as a *tetrahedron graph*. Notice that t can have two chiralities - positive and negative as shown in Figure 5.2.10. Γ_t for a positive chirality t and a negative chirality t are shown in Figure 5.2.11. The first one evaluates to $(F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)}$ and the second one evaluates to $(F_l^{ijk})^*_{(p,\alpha,\beta)(q,\gamma,\delta)}$. Let's call this number as $n_t(\tilde{l})$ and define $N(l) = \prod_t n_t(\tilde{l})$. To each edge e of T, we can associate a number $d_e(\tilde{l})$ which is the quantum dimension of the simple line assigned to e in $\tilde{l}$. Define $d(\tilde{l}) = \prod_e d_e(\tilde{l})$. The partition function $Z(M)$ is then given by

$$Z(M) = D^{-2v} \sum_{\tilde{l}} N(\tilde{l}) d(\tilde{l}) \tag{5.2.2}$$

where $D = \sqrt{\sum_i d_i^2}$ is the total quantum dimension of $\mathcal{C}$ (where d_i is the quantum dimension of simple line L_i) and v is the number of vertices in T.

The invariance of $Z(M)$ under Pachner moves is guaranteed by the pentagon equation satisfied by the associators in $\mathcal{C}$. The pentagon equation says that the following morphism made by composing associators

$$((A\otimes B)\otimes C)\otimes D \to (A\otimes(B\otimes C))\otimes D \to A\otimes((B\otimes C)\otimes D) \to A\otimes(B\otimes(C\otimes D)) \tag{5.2.3}$$

and the following morphism made by composing associators

$$((A\otimes B)\otimes C)\otimes D \to (A\otimes B)\otimes(C\otimes D) \to A\otimes(B\otimes(C\otimes D)) \tag{5.2.4}$$

are equal. In terms of F-symbols, this means that

$$\sum_{r,\delta,\epsilon,\mu} (F_q^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} (F_m^{irl})_{(q,\epsilon,\gamma)(s,\mu,\nu)} (F_s^{jkl})_{(r,\delta,\mu)(t,\rho,\sigma)} = \sum_{\tau} (F_m^{pkl})_{(q,\beta,\gamma)(t,\rho,\tau)} (F_m^{ijt})_{(p,\alpha,\tau)(s,\sigma,\nu)} \tag{5.2.5}$$

One can check that (5.2.5) is invariant under an arbitrary gauge transformation (5.2.1).

5.2.3 Twisted spherical fusion category and orientation reversing defects

Consider an oriented theory defined by $\mathcal{C}$. We propose that an unoriented parent theory can be constructed in terms of a larger "twisted" spherical fusion category $\tilde{\mathcal{C}}$. This larger category is assembled from four pieces $\tilde{\mathcal{C}} = \tilde{\mathcal{C}}_{0,0} \oplus \tilde{\mathcal{C}}_{0,1} \oplus \tilde{\mathcal{C}}_{1,0} \oplus \tilde{\mathcal{C}}_{1,1}$ such that each of the subcategories $\tilde{\mathcal{C}}_{\epsilon,\epsilon'}$ is G-graded. $\tilde{\mathcal{C}}_{0,1}$ is a bimodule on which $\tilde{\mathcal{C}}_{0,0}$ acts from left and $\tilde{\mathcal{C}}_{1,1}$ acts from right. Similarly, $\tilde{\mathcal{C}}_{1,0}$ is a (non-empty) bimodule with $\tilde{\mathcal{C}}_{1,1}$ acting from the left and $\tilde{\mathcal{C}}_{0,0}$ acting from the right. An object in $\tilde{\mathcal{C}}_{0,1}$ fuses with an object in $\tilde{\mathcal{C}}_{1,0}$ to give an object in $\tilde{\mathcal{C}}_{0,0}$. Similarly, an object in $\tilde{\mathcal{C}}_{1,0}$ fuses with an object in $\tilde{\mathcal{C}}_{0,1}$ to give an object in $\tilde{\mathcal{C}}_{1,1}$. $\tilde{\mathcal{C}}$ can also be thought of as a 2-category made out of two objects '0' and '1'.

$\tilde{\mathcal{C}}_{0,0}$ is the same as the spherical fusion category $\mathcal{C}$. $\tilde{\mathcal{C}}_{1,1}$ has same objects as that of $\mathcal{C}$. Similarly, $\tilde{\mathcal{C}}_{1,0}$ is a copy of $\tilde{\mathcal{C}}_{0,1}$ at the level of objects. Graphically, we describe simple objects of $\tilde{\mathcal{C}}_{\epsilon,\epsilon'}$ as lines with the left plaquette labeled by ϵ and right plaquette labeled by ϵ'. In general, we draw graphs Γ in $\tilde{\mathcal{C}}$ in which we label the each plaquette by some ϵ. See Figure 5.2.12.

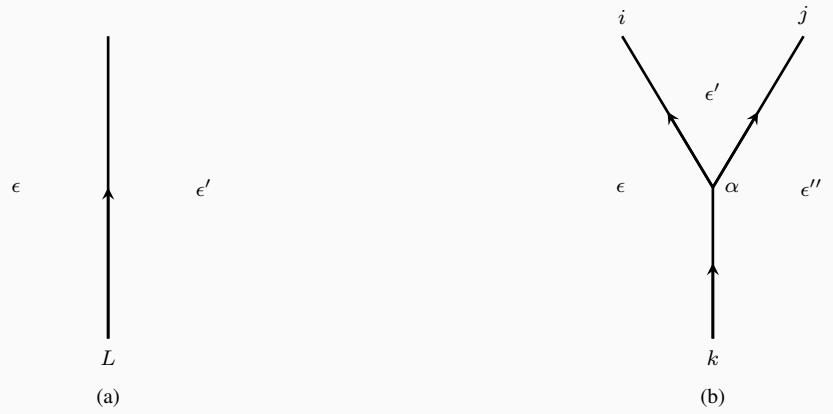


Figure 5.2.12: (a) A line in $\tilde{\mathcal{C}}_{\epsilon,\epsilon'}$. (b) A sample graph in $\tilde{\mathcal{C}}$ showing a morphism α in the morphism space $(V_k^{ij})_\epsilon$.

In our notation, the labels i, j, k etc. tell us that we have a line of $\tilde{\mathcal{C}}_{\epsilon,\epsilon'}$ with a specific value of $\epsilon+\epsilon'$ but do not determine the individual ϵ, ϵ'. The data of individual ϵ is captured in the labeling of plaquettes by ϵ, ϵ', ϵ'' etc. Thus the labeling of plaquettes is slightly redundant. We need only specify the label of a single plaquette and the labels for the other plaquettes can be determined from the labels i, j, k etc. In what follows, we will often just specify the label of the left-most plaquette.

$\tilde{\mathcal{C}}$ comes equipped with the data of an anti-linear isomorphism I between various morphism spaces. This map is easy to describe in terms of simple objects. It takes the morphism space $(V_k^{ij})_\epsilon$ from L_k to $L_i \otimes L_j$ with some labeling of plaquettes (such that the left-most plaquette is ϵ) to the morphism space $(V_k^{ij})_{\epsilon+1}$ from L_k to $L_i \otimes L_j$ but with the labeling on plaquettes flipped. Thus, we just have to pick a basis α of morphism spaces $(V_k^{ij})_0$. The basis $(V_k^{ij})_1$ is determined by the action of I and we label the resulting basis by the same labels α. See Figure 5.2.13. Thus, under a change of basis $(U_k^{ij})_{\alpha'\alpha}$ of $(V_k^{ij})_0$, the basis of $(V_k^{ij})_1$ transforms by $(U_k^{ij})^*_{\alpha'\alpha}$.

The associators are compatible with I. Let's denote F-symbol associated to a graph having the left-most plaquette 0 as $(F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)}$ as shown in Figure 5.2.14. Then, compatibility of associator and I implies that the F-symbol associated to the same graph but with the labels of all plaquettes flipped is $(F_l^{ijk})^*_{(p,\alpha,\beta)(q,\gamma,\delta)}$. Thus, the pentagon equation for the associator becomes

$$\sum_{r,\delta,\epsilon,\mu} (F_q^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} (F_m^{irl})_{(q,\epsilon,\gamma)(s,\mu,\nu)} (F_s^{jkl})^{s(i)}_{(r,\delta,\mu)(t,\rho,\sigma)} = \sum_{\tau} (F_m^{pkl})_{(q,\beta,\gamma)(t,\rho,\tau)} (F_m^{ijt})_{(p,\alpha,\tau)(s,\sigma,\nu)} \tag{5.2.6}$$

where $s(i) = *$ if i labels a simple object of $\tilde{\mathcal{C}}_{0,1}$ or $\tilde{\mathcal{C}}_{1,0}$ and $s(i) = 1$ otherwise. As the equation in terms of F-symbols looks different from (5.2.5), we refer to this equation as the twisted pentagon equation even though it still descends from the pentagon equation on the associators. This equation also appeared in [154].

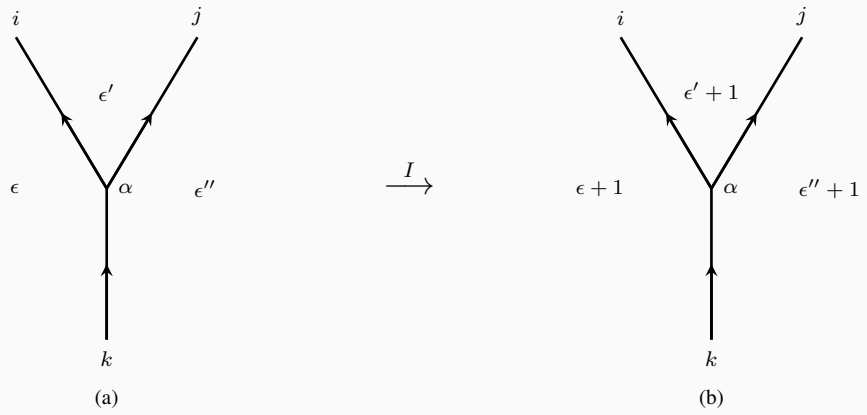


Figure 5.2.13: The action of anti-linear isomorphism I.

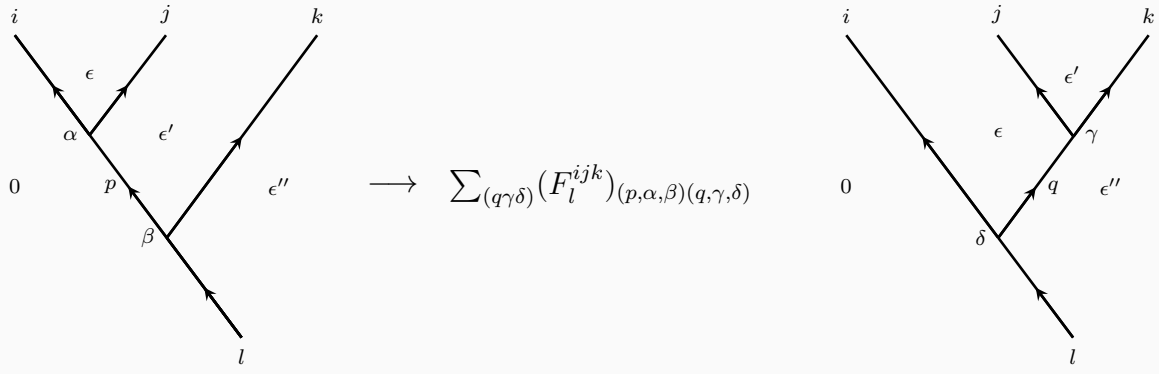


Figure 5.2.14: Definition of F-symbols for $\tilde{\mathcal{C}}$.

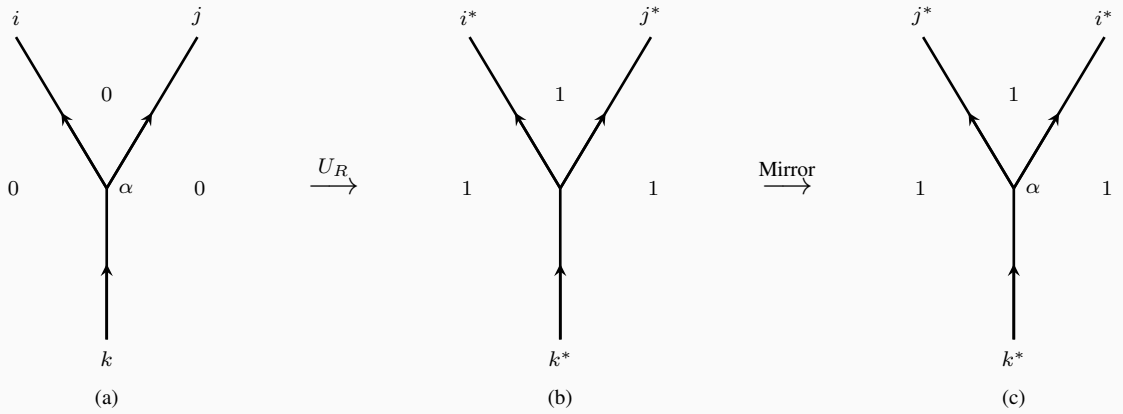


Figure 5.2.15: The action of U_R sends a graph (a) on $\mathfrak{B}$ to a graph (b) on $\mathfrak{B}'$ but in the "wrong" orientation. This means that the tensor product in graph (b) is taken from right to left. In order to get back to our convention of tensor product from left to right, we take a mirror of graph (b) and obtain graph (c). The resulting graph (c) is read in $\mathcal{C}'$.

Notice that the gauge transformations on F-symbols now take the following form

$$(F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} \to (F_l^{ijk})_{(p,\alpha,\beta)(q,\gamma,\delta)} (U_q^{jk})^{*s(i)}_{\gamma'\gamma} (U_l^{iq})^*_{\delta'\delta} (U_p^{ij})_{\alpha'\alpha} (U_l^{pk})_{\beta'\beta} \tag{5.2.7}$$

and one can check that (5.2.6) is invariant under this gauge transformation.

The identity line of $\mathcal{C} = \tilde{\mathcal{C}}_{0,0}$ can be inserted anywhere in plaquettes labeled by 0 without changing any answers. Similarly, the identity line of $\tilde{\mathcal{C}}_{1,1}$ can be inserted anywhere in plaquettes labeled by 1 without changing any answers. For each object in $\tilde{\mathcal{C}}_{0,1}$, we define a dual object in $\tilde{\mathcal{C}}_{1,0}$ and vice-versa. These duals are slightly different from the usual duals in a spherical fusion category $\mathcal{C}$. That is, given a line L in $\tilde{\mathcal{C}}_{0,1}$, the evaluation maps take $L \otimes L^*$ to identity in $\tilde{\mathcal{C}}_{0,0}$ and take $L^* \otimes L$ to identity in $\tilde{\mathcal{C}}_{1,1}$. Similar statements hold true if we flip 0 and 1 or replace evaluation with co-evaluation.

Given a graph Γ in $\tilde{\mathcal{C}}$ drawn on the sphere, different projections to planar graphs Γ_p must be equivalent. In other words, we demand that there are conditions on $\tilde{\mathcal{C}}$ similar to that of a spherical structure on a spherical fusion category $\mathcal{C}$.

We now turn our attention to the physical interpretation of the structure we have just described. An unoriented 3d TFT $\tilde{\mathfrak{T}}$ has an orientation reversing defect U_R implementing a reflection transformation. This defect can fuse with other orientation preserving defects U_g to form more orientation reversing defects U_{Rg}. The fusion of these defects froms a group $\tilde{G} = G \times \mathbb{Z}_2$ and there is a canonical homomorphism ρ_1 from $\tilde{G}$ to $\mathbb{Z}_2$ whose kernel is G.

U_R can be fused with the boundary $\mathfrak{B}$ to give a new boundary $\mathfrak{B}'$. Under such a fusion, the orientation of the boundary flips as well. There is a spherical fusion category $\mathcal{C}'$ associated to $\mathfrak{B}'$ which is identified as $\tilde{\mathcal{C}}_{1,1}$. If there is a line L on $\mathfrak{B}$, then fusion of $\mathfrak{B}$ with U_R flips its orientation and we obtain the line L^* on $\mathfrak{B}'$. Consider a morphism from L_k to $L_i \otimes L_j$ on $\mathfrak{B}$. Slapping U_R on top of it, we send each line to its dual and $\mathfrak{B}$ to $\mathfrak{B}'$. However, since this process flips the

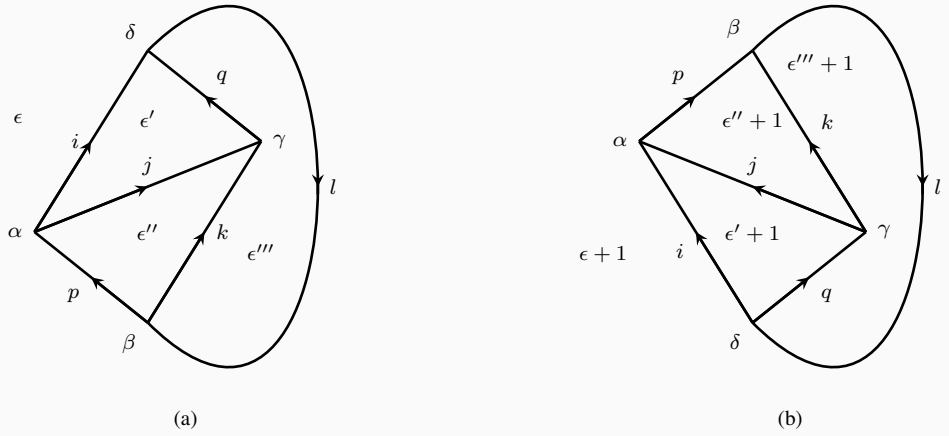


Figure 5.2.16: The symmetry of the theory under a reflection guarantees that the above two graphs evaluate to the same number.

orientation of the boundary, we have to take a mirror of the resulting configuration of lines to read it in terms of $\mathcal{C}'$. See Figure 5.2.15. Thus, fusion with U_R provides a linear isomorphism from V_k^{ij} in $\mathcal{C}$ to $V_{k^*}^{j^*i^*}$ in $\mathcal{C}'$. This is the origin of the anti-linear isomorphism I in $\tilde{\mathcal{C}}$ described above.

U_{Rg} can end on the boundary giving an interface between $\mathfrak{B}$ and $\mathfrak{B}'$ and an interface between $\mathfrak{B}'$ and $\mathfrak{B}$. The lines living on these interfaces give rise to $\tilde{\mathcal{C}}_{0,1}$ and $\tilde{\mathcal{C}}_{1,0}$ respectively. Together they form a "twisted" spherical fusion category $\tilde{\mathcal{C}}$ described above.

The label 0 and 1 of plaquettes in a graph in $\tilde{\mathcal{C}}$ corresponds respectively to the boundary $\mathfrak{B}$ and $\mathfrak{B}'$ in the physical setting. A graph Γ in $\tilde{\mathcal{C}}$ drawn on a sphere computes the partition function of $\tilde{\mathfrak{T}}$ on a 3-ball with a network of boundary lines specified by Γ. The bulk of the 3-ball contains orientation reversing defects which end on the boundary at the location of lines living in $\tilde{\mathcal{C}}_{0,1}$ or $\tilde{\mathcal{C}}_{1,0}$.

Given such a 3-ball with Γ on the boundary, we can bubble a U_R in the bulk of the 3-ball and bring it to the boundary. This sends Γ to Γ' (after taking the mirror) and both of these graphs must evaluate to the same number. This is the origin of the compatibility between associators and I. See Figure 5.2.16.

5.2.4 Unoriented Turaev-Viro

In this subsection, we generalize the Turaev-Viro prescription to compute the partition function of an unoriented theory $\tilde{\mathfrak{T}}$ on an unoriented 3-manifold M. We will assume that the reader has read subsection 5.2.2 before reading this subsection and so we will sometimes cut corners in what follows.

Fix an orientation $\mathcal{O}$ on $\mathbb{R}^3$. An unoriented 3-manifold M can be constructed as follows. We pick open sets of $\mathbb{R}^3$ and glue them along codimension one loci using piecewise-linear maps. This gives us a locus $\mathcal{L}$ in M which is defined by the property that the transition functions are orientation reversing. The Poincare dual of this locus is a representative of first Stiefel-Whitney class and we call it w_1. We assign a local orientation $\mathcal{O}_t$ to any small tetrahedron t in $M - \mathcal{L}$ by first using the local chart to map it to a tetrahedron in $\mathbb{R}^3$ where we have already picked an orientation $\mathcal{O}$. $\mathcal{O}_t$ remains invariant under deformations of t inside $M - \mathcal{L}$.

Pick a branched triangulation T of M. w_1 assigns a number p_e valued in $\{0, 1\}$ to every edge e. And the G-connection α_1 on M assigns an element g_e of the group G to each directed edge e.

Let's extract a set of labels $S_{0,0}$ such that each label in the set corresponds to a simple object of $\tilde{\mathcal{C}}_{0,0}$. Similarly, extract a set of labels $S_{0,1}$ such that each label in the set corresponds to a simple object of $\tilde{\mathcal{C}}_{0,1}$. Define an involution $*$ on $S_{0,0}$ induced by taking the dual of simple objects. Similarly, define an involution $*$ on $S_{0,1}$ under which i is sent to j if the $*$ operation of $\tilde{\mathcal{C}}$ sends the object L_i in $\tilde{\mathcal{C}}_{0,1}$ to the object L_j in $\tilde{\mathcal{C}}_{1,0}$. There is also a G-grading on both of these sets descending from the G-grading of simple objects.

We now label each directed edge e by a label in $(S_{0,p_e})_{g_e}$. Pick a face f of T. We can label f by some label α just as in the oriented case. Thus we obtain a labeling of edges and faces of a branched triangulation. Call one such labeling as $\tilde{l}$.

Now pick a tetrahedron t in $M - \mathcal{L}$ in the labeling $\tilde{l}$. To each such t we will assign a planar graph Γ_t in $\tilde{\mathcal{C}}$. If the chirality of t matches the local orientation $\mathcal{O}_t$, we assign the graph shown in Figure 5.2.17(a) with $\epsilon = \epsilon' = \epsilon'' = \epsilon''' = 0$ and if the chirality doesn't match the local orientation we assign the graph shown in Figure 5.2.17(b) with $\epsilon = \epsilon' = \epsilon'' = \epsilon''' = 0$. To define Γ_t for a t intersecting $\mathcal{L}$, we choose a small neighborhood U_t of t such that $\mathcal{L}$ looks like a wall cutting U_t into two parts. On one side of the wall, we assign 0 to every vertex and on the other side we assign 1. We assign a global orientation $\mathcal{O}_{U_t}$ to U_t given by local orientation $\mathcal{O}_{t'}$ of any tetrahedron t' lying completely on one side of the wall where vertices are labeled by 0. We now assign the graph shown in Figure 5.2.17(a) with $\epsilon = 0$ and arbitrary ϵ', ϵ'', ϵ''' if chirality of t matches $\mathcal{O}_{U_t}$ and the graph shown in Figure 5.2.17(b) with $\epsilon = 0$ and arbitrary ϵ', ϵ'', ϵ''' if it does not.

Notice that if we flip the choice of 0 and 1 that we assigned to the sides of the wall and apply the above presciption, then Γ_t is flipped to the "reflected" graph Γ'_t which is the graph obtained by acting U_R on Γ_t. See Figure 5.2.16. Γ_t and Γ'_t evaluate to the same number.

Also notice that if we take a tetrahedron t'' in U_t whose vertices are all labeled by 1 and assign to it a new graph $\Gamma'_{t''}$ by matching its chirality with $\mathcal{O}_{U_t}$ instead, then $\Gamma'_{t''}$ will be the "reflected" version of the old graph $\Gamma_{t''}$ that we assigned in the starting of last paragraph by matching its chirality with the local orientation, and hence $\Gamma'_{t''}$ will evaluate to the same number as $\Gamma_{t''}$. Thus,

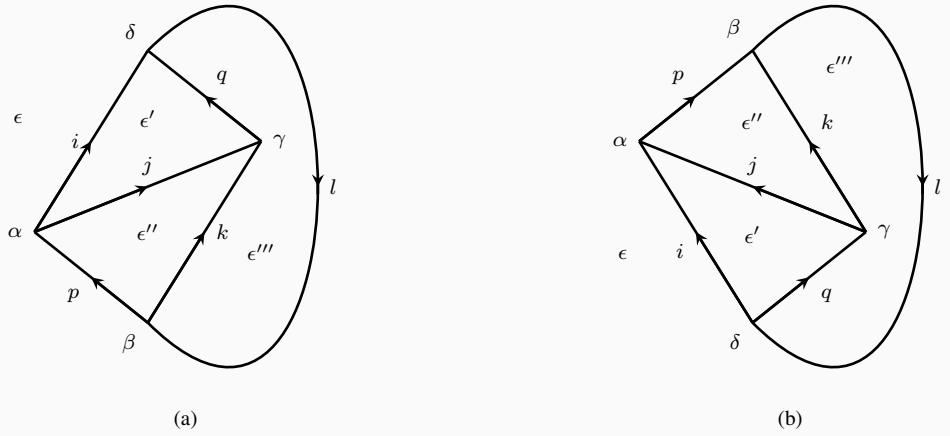


Figure 5.2.17: The possible graphs that we attach to a tetrahedron.

we see that we could have also given the prescription to compute the partition function in various patches U_i by using the local orientations and assigning $\{0, 1\}$ to the two sides produced by an intersection of U_i with the w_1 wall. The tetrahedra lying the intersection of U_i and U_j would give the same contribution in each patch. Thus, we would just have to make sure that we "glue" the tetrahedra in various intersections properly.

Returning back to our original prescription, we just repeat what we already said for the oriented case. Let's call the evaluation of Γ_t as $n_t(\tilde{l})$ and define $N(l) = \prod_t n_t(\tilde{l})$. To each edge e of T, we can associate a number $d_e(\tilde{l})$ which is the quantum dimension of the simple line assigned to e in $\tilde{l}$. Define $d(\tilde{l}) = \prod_e d_e(\tilde{l})$. The partition function $Z(M)$ is then given by

$$Z(M) = D^{-2v} \sum_{\tilde{l}} N(\tilde{l}) d(\tilde{l}) \tag{5.2.8}$$

where $D = \sqrt{\sum_i d_i^2}$ is what we dub as the total quantum dimension of $\tilde{\mathcal{C}}$ (where d_i is the quantum dimension of simple line L_i) and v is the number of vertices in T. We would like to emphasize that we are picking labels i only in "half" of $\tilde{\mathcal{C}}$ *i.e.* $\tilde{\mathcal{C}}_{0,0}$ and $\tilde{\mathcal{C}}_{0,1}$. Hence, the total quantum dimension only involves square of quantum dimensions of half of the simple lines.

The invariance of $Z(M)$ under Pachner moves and under change of representative of w_1 is guaranteed by the twisted pentagon equation (5.2.6) satisfied by the F-symbols in $\tilde{\mathcal{C}}$. In the rest of the chapter, by "twisted spherical fusion category" we will mean the data of $\tilde{\mathcal{C}}_{0,0} \oplus \tilde{\mathcal{C}}_{0,1}$ and

we will often repackage this data as $\mathcal{C}' = \oplus_{\tilde{g}} \mathcal{C}'_{\tilde{g}} = \oplus_g \mathcal{C}'_g \oplus_g \mathcal{C}'_{Rg}$ where $\tilde{g} \in \tilde{G} = G \times \mathbb{Z}_2$ and R is the generator of $\mathbb{Z}_2$ in $\tilde{G}$. $\mathcal{C}'_g = (\tilde{\mathcal{C}}_{0,0})_g$ and $\mathcal{C}'_{Rg} = (\tilde{\mathcal{C}}_{0,1})_g$. We also define a homomorphism ρ (also called ρ_1) from $\tilde{G}$ to $\mathbb{Z}_2$ which sends $G \times \{e\}$ to 0 and $G \times \{R\}$ to 1.

5.2.5 Example: Bosonic SPT phases

Bosonic SPT phases protected by $G = G_0 \times \mathbb{Z}_2^T$ are invertible unoriented TFTs with global symmetry G_0. Such a phase is constructed by a twisted spherical fusion category $\mathcal{C}$ having a single simple object L_g in each subcategory $\mathcal{C}_g$. The fusion rules are $L_g \otimes L_{g'} \simeq L_{gg'}$.

F-matrices define a $U(1)$ valued function of three group elements $\alpha_3(g, g', g'') = F_{g,g',g'';gg'g''}$. The twisted pentagon equation (5.2.6) translates to

$$\frac{\alpha_3^{s(g)}(g', g'', g''')\alpha_3(g, g'g'', g''')\alpha_3(g, g', g'')}{\alpha_3(gg', g'', g''')\alpha_3(g, g', g''g''')} = 1 \tag{5.2.9}$$

where $s(g) = (-1)^{\rho(g)}$. This means that α_3 is a ρ-twisted group cocycle. On the other hand, gauge transformations (5.2.7) become

$$\alpha_3(g, g', g'') \to \alpha_3(g, g', g'')\frac{\beta_2^{s(g)}(g', g'')\beta_2(g, g'g'')}{\beta_2(g, g')\beta_2(gg', g'')} \tag{5.2.10}$$

which corresponds to adding an exact ρ-twisted cocycle $\delta\beta_2$ to α_3.

This means that the bosonic SPT phases are classified by the ρ-twisted group cohomology $H^3(BG, U(1)_\rho)$ [155]. A background connection α_1 for G_0 on M combines with w_1 to give a background connection for G which is represented as a map from M to BG. An element of $H^3(BG, U(1)_\rho)$ is then pulled back to a density on M which can be integerated on M to produce the partition function $Z(M, \alpha_1)$.

5.3 Pin$^+$-TFTs

We start this section by reviewing the construction of Spin-TFTs from their shadows [3]. We will argue that the Pin$^+$-shadows must have an additional kind of anomaly which was not present in the case of Spin-shadows. Incorporating this addtional anomaly will allow us to generalize the shadow construction to Pin$^+$-TFTs. We finish the section by showing how to take a product of Pin$^+$-TFTs at the level of shadows.

5.3.1 Review of Spin case

[3] provided a recipe to construct a 3d Spin-TFT $\mathfrak{T}_s$ from its shadow $\mathfrak{T}_f$. The shadow is an ordinary TFT with an anomalous $\mathbb{Z}_2$ 1-form symmetry. This manifests itself in the existence of a bulk line Π which fuses with itself to the identity and has certain properties. See Figure 5.3.1.

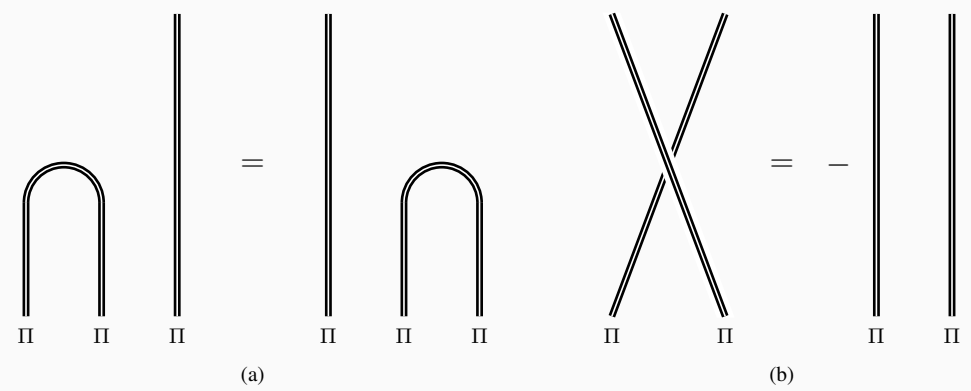


Figure 5.3.1: (a) A property of bulk line Π generating an anomalous $\mathbb{Z}_2$ 1-form symmetry. (b) Half-braiding Π lines across each other gives a factor of -1 when compared to Π lines without braiding.

We want to couple $\mathfrak{T}_f$ to a background 2-connection β_2 for the 1-form symmetry. We can do so by inserting Π lines inside a triangulated manifold such that an even number of Π lines cross a face having $\beta_2 = 0$ and an odd number of Π lines cross a face having $\beta_2 = 1$. Since Π has a non-trivial crossing with itself, topologically different ways of gluing Π lines inside the tetrahedron will differ by signs. Hence, we need to pick a convention of how we will glue the Π lines crossing these faces inside each tetrahedron when we say that $\mathfrak{T}_f$ is coupled to a background 2-connection β_2. Once we have picked this convention, the partition function will not be invariant under gauge transformations of β_2.

After fixing the convention, the change in the partition function under gauge transformations is independent of the theory, however. To see this, consider the product $\mathfrak{T} = \mathfrak{T}_1 \times \mathfrak{T}_2$ of two shadow theories $\mathfrak{T}_1$ and $\mathfrak{T}_2$, and couple it to a background 2-connection for the diagonal $\mathbb{Z}_2$ 1-form symmetry. The partition function would then be the product

$$Z(M, \beta_2) = Z_1(M, \beta_2) Z_2(M, \beta_2) \tag{5.3.1}$$

and a gauge transformation would leave Z invariant. This is because resolving a crossing of the product line $\Pi_1\Pi_2$ gives no minus sign as the signs from crossing of Π_1 and crossing of Π_2 cancel each other.

The strategy of [3] was to compute this anomaly for a simple shadow theory, namely the shadow of Gu-Wen fermionic SPT phases. The anomaly under $\beta_2 \to \beta_2 + \delta\lambda_1$ turns out to be

$$Z_f(M, \beta_2) \to (-1)^{\int_M \lambda_1 \cup \beta_2 + \beta_2 \cup \lambda_1 + \lambda_1 \cup \delta\lambda_1} Z_f(M, \beta_2) \tag{5.3.2}$$

This transformation is the same as the transformation of a spin-structure η_1 dependent sign $z(M, \eta_1, \beta_2)$. This sign can be written as [156]

$$z(M, \eta_1, \beta_2) = (-1)^{\int_M \eta_1 \cup \beta_2 + \int_N \beta_2 \cup \beta_2 + w_2 \cup \beta_2} \tag{5.3.3}$$

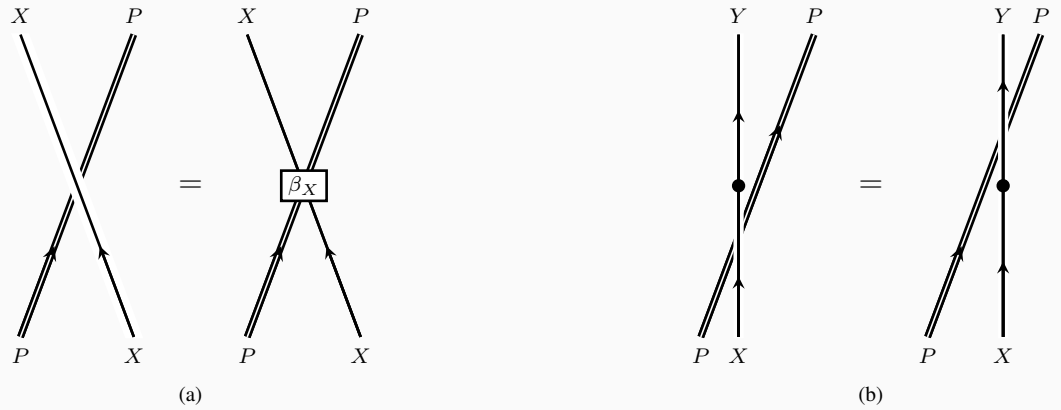


Figure 5.3.2: Properties of boundary image of bulk lines. These properties imply that the bulk lines are elements of the Drinfeld center of the spherical fusion category formed by boundary lines.

where N is a 4-manifold whose boundary is M and w_2 is a representative of second Stiefel-Whitney class. For oriented manifolds, this sign is independent of N because $\beta_2 \cup \beta_2 + w_2 \cup \beta_2$ is exact if β_2 is a cocycle. It is easy to see from this expression that

$$z(M, \eta_1, \beta_2 + \delta\lambda_1) = (-1)^{\int_M \lambda_1 \cup \beta_2 + \beta_2 \cup \lambda_1 + \lambda_1 \cup \delta\lambda_1} z(M, \eta_1, \beta_2) \tag{5.3.4}$$

Here we have used a representation of spin structure as an equivalence class of 1-cochains η_1 satisfying $\delta\eta_1 = w_2$ under the equivalence relation given by addition of exact 1-cochains to η_1 [156].

Thus combining the shadow theory with this sign gives a theory with a non-anomalous $\mathbb{Z}_2$ 1-form symmetry. The spin theory $\mathfrak{T}_s$ is obtained from this by gauging this 1-form symmetry

$$Z_s(M, \eta_1) = \frac{|H^0(M, \mathbb{Z}_2)|}{|H^1(M, \mathbb{Z}_2)|} \sum_{[\beta_2] \in H^2(M, \mathbb{Z}_2)} z(M, \eta_1, \beta_2) Z_f(M, \beta_2) \tag{5.3.5}$$

So, we have a one-to-one correspondence $[\mathfrak{T}_s] \leftrightarrow (\mathfrak{T}_f, \Pi)$ where $[\mathfrak{T}_s]$ denotes the equivalence class of Spin-TFTs under permutations of spin structures $\eta_1 \to \eta_1 + \alpha_1$ where $[\alpha_1] \in H^1(M, \mathbb{Z}_2)$.

We would like to have a Turaev-Viro construction for $Z_f(M, \beta_2)$. To this end, we should understand how to encode the Π line in terms of the spherical fusion category $\mathcal{C}$. Notice that Π is mapped to a boundary line P by bringing it to the boundary. If we bring Π to the boundary such that it crosses a boundary line X, we obtain a canonical isomorphism $\beta_X : X \otimes P \to P \otimes X$. Bringing Π to the boundary in topologically equivalent ways should lead to same answers. Hence, (P, β) can be moved across other morphisms. See Figure 5.3.2.

Mathematically, this means that Π is an element (P, β) of Drinfeld center of $\mathcal{C}$. This element fuses with itself to identity and $\beta_P = -1$. The Turaev-Viro construction for $Z_f(M, \beta_2)$ is achieved by inserting a Π line emanating from every vertex whose dual face has $\beta_2 = 1$. See

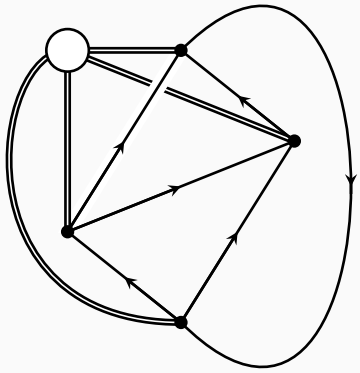


Figure 5.3.3: The basic tetrahedron graph in the Turaev-Viro construction of the partition function $Z_f(M, \beta_2)$ of the shadow theory in the presence of a background 2-connection β_2. A Π line (shown as double line in the figure) leaves the vertex if the dual face has $\beta_2 = 1$. We let such lines meet without crossing each other in the region denoted by a disk in the graph. Different ways of joining the lines in the disk are equivalent because of the property of Π lines shown in Figure 5.3.1(a).

Figure 5.3.3.

5.3.2 Fermion in Pin$^+$-theories

Pin$^+$-TFTs are a generalization of Spin-TFTs to the unoriented case. Spinors can be defined on an n-dimensional non-orientable manifold by using transition functions valued in Pin$^+$(n) group or Pin$^-$(n) group, both of which are double covers of $O(n)$. They are distinguished by the value of R^2 acting on spinors where R is a spatial reflection. $R^2 = +1$ for Pin$^+$-group and $R^2 = -1$ for Pin$^-$-group. In terms of time reversal symmetry T, the action on spinors is $T^2 = -1$ for the Pin$^+$ case and $T^2 = +1$ for the Pin$^-$ case. A Pin$^+$-structure exists only if the second Stiefel-Whitney class $[w_2]$ vanishes. On the other hand, a Pin$^-$-structure exists only if $[w_2 + w_1^2]$ vanishes where $[w_1]$ is the first Stiefel-Whitney class. Two Pin$^+$ or Pin$^-$-structures differ by an element of $H^1(M, \mathbb{Z}_2)$.

In the Pin$^+$ case, there is a choice in defining the action of reflection in i-th spatial direction on spinors. We can either multiply the spinor by the gamma matrix γ_i or by $-\gamma_i$. This suggests that in a Pin$^+$-shadow there are two canonical choices m_R and $n_R = -m_R$ of local operators at the junction of a Π line and R-defect. These operators square to 1. The orientation preserving defects always have a single canonical local operator at the junction.

Now we will argue that, in general, there must be a locus $\mathcal{L}$ embedded inside the locus $\mathcal{M}$ of orientation reversing defects which implements the transformation $m \leftrightarrow n$. Moreover, the homology class of $\mathcal{L}$ must be the Poincare dual of $[w_1^2]$. Choose a locus $\mathcal{L}'$ embedded inside $\mathcal{M}$ whose homology class is the dual of $[w_1^2]$. Now bubble a fermion line near $\mathcal{M}$ and move it such that it intersects $\mathcal{M}$ in two junctions. See Figure 5.3.4(a). The local opeartors at the two junctions must be inverses of each other. Take one of these junctions around a cycle C in $\mathcal{M}$. If the cycle intersects $\mathcal{L}'$, then fusing the fermion line with itself at the end of this process gives a crossing of fermion line which provides a factor of -1. See Figure 5.3.4(b). Topological invariance demands that C must intersect $\mathcal{L}$ as well so that the fusion of the local operators at the end of the process provides a factor of -1 which cancels the sign from the crossing. Similarly, if C doesn't intersect $\mathcal{L}$, it doesn't intersect $\mathcal{L}'$ either. Hence, $\mathcal{L}$ and $\mathcal{L}'$ are in the same homology class. Thus, we can choose to identify $\mathcal{L}$ with the representative w_1^2.

We will see in the next section that this flip $m \leftrightarrow n$ as Π crosses $\mathcal{L}'$ is responsible for the presence of mixed anomaly between time reversal symmetry and $\mathbb{Z}_2$ 1-form symmetry in Pin$^+$-shadows.

5.3.3 Shadows of Pin$^+$-TFTs

Just as in the spin case, to define what we mean by a Pin$^+$-shadow $\mathfrak{T}_f$ coupled to a background β_2, we need to pick a convention for configuring Π lines. In addition to this, we also need to choose whether we will put m or n on the junctions when Π crosses R-defect. The Pin$^+$-TFT $\mathfrak{T}_+$ is obtained as

$$Z_+(M, \eta_1) = \frac{|H^0(M, \mathbb{Z}_2)|}{|H^1(M, \mathbb{Z}_2)|} \sum_{[\beta_2] \in H^2(M, \mathbb{Z}_2)} z_+(M, \eta_1, \beta_2) Z_f(M, \beta_2) \tag{5.3.6}$$

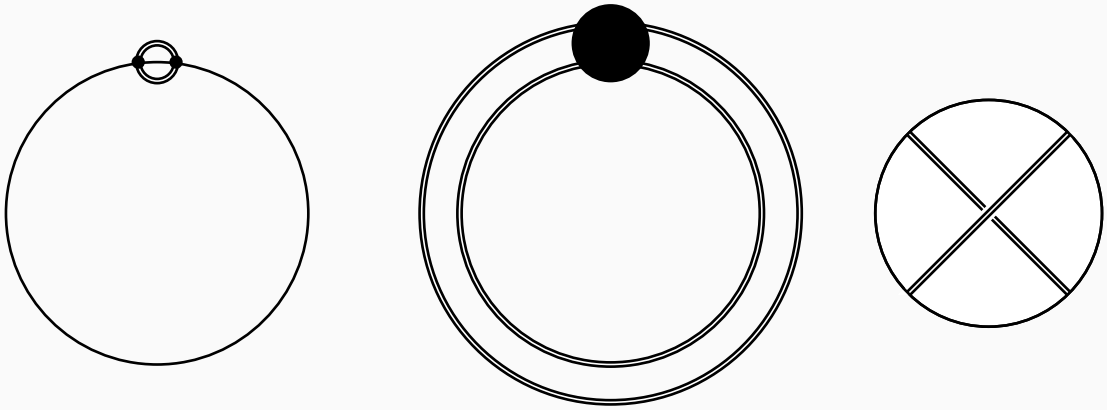


Figure 5.3.4: (a) The big circle is a cartoon representing the locus $\mathcal{M}$ dual to w_1. A Π line is bubbled nearby and dipped into this locus. The operators at the two marked junctions are inverses of each other. (b) Taking the left half of Π line around a loop C in $\mathcal{M}$ which intersects once the locus $\mathcal{L}'$ dual to w_1^2, brings us to the configuration shown in the figure. We omit $\mathcal{M}$ in this figure for brevity. The region denoted by a black disk is shown in (c). That is, the Π lines are glued inside this black region in such a way that there is a non-trivial half-braiding (*i.e.* crossing) of the Π lines. The reason for the appearance of this crossing is that the normal direction to $\mathcal{M}$ is reflected across $\mathcal{L}'$ and hence the top and bottom parts of the (left half of) Π loop are exchanged as C crosses $\mathcal{L}'$.

where the sign which cancels the anomaly for $\mathbb{Z}_2$ 1-form symmetry of $\mathfrak{T}_f$ can be defined as

$$z_+(M,\eta_1,\beta_2) = (-1)^{\int_M \eta_1\cup\beta_2+\int_N \beta_2\cup\beta_2+(w_1^2+w_2)\cup\beta_2} \tag{5.3.7}$$

where $\partial N = M$ and η_1 parametrizes Pin$^+$-structures. The expression is independent of N as $\beta_2\cup\beta_2+(w_1^2+w_2)\cup\beta_2$ is exact if β_2 is a cocycle. Fliping the choice of local operator changes the partition function as $Z_f(M,\beta_2)\to(-1)^{\int_M w_1\cup\beta_2}Z_f(M,\beta_2)$. This can be absorbed into a permutation of Pin$^+$-structures $\eta_1\to\eta_1+w_1$. Thus, as in the spin case, we have a one-to-one correspondence $[\mathfrak{T}_+]\leftrightarrow(\mathfrak{T}_f,\Pi)$ where $[\mathfrak{T}_+]$ denotes the equivalence class of Pin$^+$-TFTs under permutations of Pin$^+$-structures.

Now, notice that Pin$^+$-shadows have a time reversal anomaly in the presence of a background 2-connection β_2. As we add δv_0 to w_1, we add δu_1 to w_1^2 where $u_1 = w_1\cup v_0+v_0\cup w_1+v_0\cup\delta v_0$. This corresponds to moving $\mathcal{M}$ and $\mathcal{L}'$. But during such movements, $\mathcal{L}'$ will cross some Π lines encoding β_2 and the partition function will change as

$$Z_f(M,\beta_2)\to(-1)^{\int_M u_1\cup\beta_2}Z_f(M,\beta_2) \tag{5.3.8}$$

Under this transformation, the sign also transforms in the same way

$$z_+(M,\eta_1,\beta_2)\to(-1)^{\int_M u_1\cup\beta_2}z_+(M,\eta_1,\beta_2) \tag{5.3.9}$$

and the corresponding Pin$^+$-TFTs have no time-reversal anomaly.

The signs z_+ written above implies the following anomaly under $\beta_2\to\beta_2+\delta\lambda_1$

$$Z_f(M,\beta_2)\to(-1)^{\int_M \lambda_1\cup\beta_2+\beta_2\cup\lambda_1+\lambda_1\cup\delta\lambda_1+w_1^2\cup\lambda_1}Z_f(M,\beta_2) \tag{5.3.10}$$

where w_1 is a representative of first Stiefel-Whitney class. As the anomaly is universal, we will verify that this is the correct anomaly by computing the anomaly directly for shadows of Pin$^+$ generalization of Gu-Wen fermionic SPT phases in the next section.

To obtain the Turaev-Viro construction for $Z_f(M,\beta_2)$, we need to know how to encode the Π line in terms of the data of $\mathcal{C}$. As in the spin case, Π is mapped to some boundary line P with canonical isomorphisms $\beta_X : X\otimes P\to P\otimes X$. However, unlike the spin case, Π is not an element of Drinfeld center of $\mathcal{C}$. Rather, we need to insert extra signs whenever we move Π across $\mathcal{L}'$. This descends to the statement that (P,β) is an element of a twisted Drinfeld center which is defined in Figure 5.3.5.

5.3.4 Product of Pin$^+$-TFTs

In this subsection, we want to figure out the shadow of the product of two Pin$^+$-TFTs. This will lead to the definition of a product on the shadow theories which we will call the *shadow product*.

First, notice that [3]

$$z_+(M,\eta_1,\beta_2)z_+(M,\eta_1,\beta_2') = z_+(M,\eta_1,\beta_2+\beta_2')(-1)^{\int_M \beta_2\cup_1\beta_2'} \tag{5.3.11}$$

[3] See Appendix B of [156] for an introduction to higher cup products like $\cup_1$ used in the following equation.

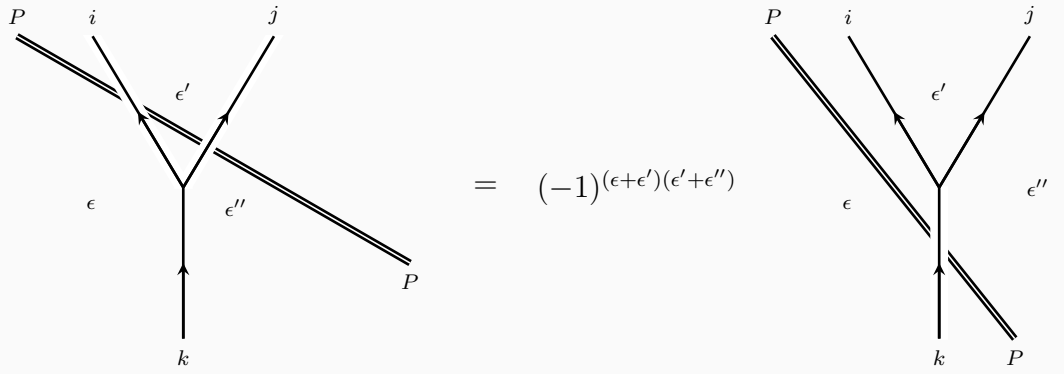


Figure 5.3.5: The equations defining twisted Drinfeld center.

Now consider two Pin$^+$-TFTs $\mathfrak{T}_+$ and $\mathfrak{T}'_+$ with their corresponding shadows $\mathfrak{T}_f$ and $\mathfrak{T}'_f$. Using the above, we can write the partition function of the product theory as

$$Z_+(M)Z'_+(M) = \frac{|H^0(M,\mathbb{Z}_2)|^2}{|H^1(M,\mathbb{Z}_2)|^2} \sum_{[\beta_2],[\beta'_2]} z_+(M,\beta_2+\beta'_2) Z_f(M,\beta_2) Z'_f(M,\beta'_2) (-1)^{\int_M \beta_2 \cup_1 \beta'_2} \tag{5.3.12}$$

which can be massaged as

$$Z_+(M,\eta_1)Z'_+(M,\eta_1) = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|} \sum_{[\beta_2]\in H^2(M,\mathbb{Z}_2)} z_+(M,\eta_1,\beta_2) \tilde{Z}_f(M,\beta_2) \tag{5.3.13}$$

with

$$\tilde{Z}_f(M,\beta_2) = \frac{|H^0(M,\mathbb{Z}_2)|}{|H^1(M,\mathbb{Z}_2)|} \sum_{[\beta'_2]\in H^2(M,\mathbb{Z}_2)} (-1)^{\int_M (\beta_2+\beta'_2)\cup_1 \beta'_2} Z_f(M,\beta_2+\beta'_2) Z'_f(M,\beta'_2) \tag{5.3.14}$$

being the partition function of the shadow corresponding to the product theory. We denote this shadow theory as the shadow product $\mathfrak{T}_f \times_f \mathfrak{T}'_f$.

Physically, we are constructing the shadow of the product by gauging the diagonal $\mathbb{Z}_2$ 1-form symmetry in the product of the shadow theories. Notice that this 1-form symmetry is non-anomalous and hence gauging it makes sense.

To implement the shadow product in the Turaev-Viro description, we first take a graded product of $\mathcal{C} \times_G \mathcal{C}'$ of $\mathcal{C}$ and $\mathcal{C}'$. This means that $(\mathcal{C} \times_G \mathcal{C}')_g = \mathcal{C}_g \times \mathcal{C}'_g$. Now we need a notion of gauging the line $b = \Pi\Pi'$ in the Drinfeld center of $\mathcal{C} \times_G \mathcal{C}'$. In general, we can consider the following problem. Take a theory $\mathfrak{T}_b$ specified by a twisted spherical fusion category $\mathcal{C}_b$ having a non-anomalous $\mathbb{Z}_2$ 1-form symmetry. This means that there exists a line b in the Drinfeld center of $\mathcal{C}_b$ which fuses with itself to identity and has the properties shown in Figure 5.3.6. We want to construct the twisted spherical fusion category for the theory $\mathfrak{T}_{\mathbb{Z}_2}$ obtained after gauging the 1-form symmetry generated by b.

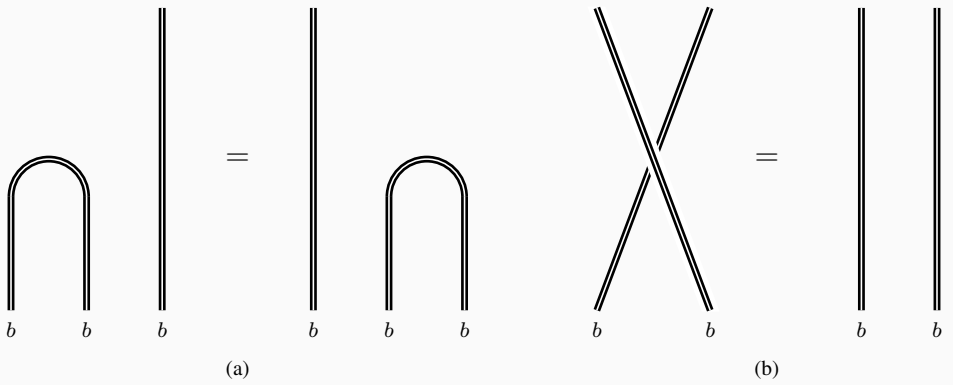


Figure 5.3.6: The properties of a bulk line b generating a non-anomalous $\mathbb{Z}_2$ 1-form symmetry are very similar to that of Π. The only difference is that crossing b lines doesn't lead to a sign.

b is invisble in the gauge theory. This means that a morphism from A to $b \otimes B$ in $\mathcal{C}_b$ has to be regarded as a morphism from A to B in $\mathcal{C}_{\mathbb{Z}_2}$. And the morphisms from A to B in $\mathcal{C}_b$ are also morphisms from A to B in $\mathcal{C}_{\mathbb{Z}_2}$. The composition and tensor product of new morphisms are defined as shown in the Figure 5.3.7.

Let's try to understand what happens to the simple objects under this operation. If L is a simple object in $\mathcal{C}_b$, $M = b \otimes L$ is simple as well. If M is not isomorphic to L, then the morphism from L to $b \otimes M$ in $\mathcal{C}_b$ provides an isomorphism from L to M in $\mathcal{C}_{\mathbb{Z}_2}$ combining them into a single simple object in $\mathcal{C}_{\mathbb{Z}_2}$. If M is isomorphic to L, then the morphism from L to $b \otimes M$ in $\mathcal{C}_b$ provides an additional endomorphism ξ_L of L in $\mathcal{C}_{\mathbb{Z}_2}$. Since there are two independent morphisms from L to itself in $\mathcal{C}_{\mathbb{Z}_2}$, it must split into two simple objects L^+ and L^- which can be constructed by inserting a projector

$$\pi_L^{\pm} = \frac{1}{2}(1 \pm \xi_L) \tag{5.3.15}$$

on L.

5.4 Fermionic SPT phases

In this section, we discuss Pin$^+$-SPT phases. We also explicitly compute the partition function on an arbitrary manifold M of a certain Pin$^+$-shadow which gives rise to the Pin$^+$ Gu-Wen phases. We can read the anomaly of Pin$^+$-shadows from the expression for the partition function. The anomaly matches the expectation of the previous section. We finish the section by reproducing $\mathbb{Z}_2$ group of Pin$^+$-SPT phases without any global symmetry.

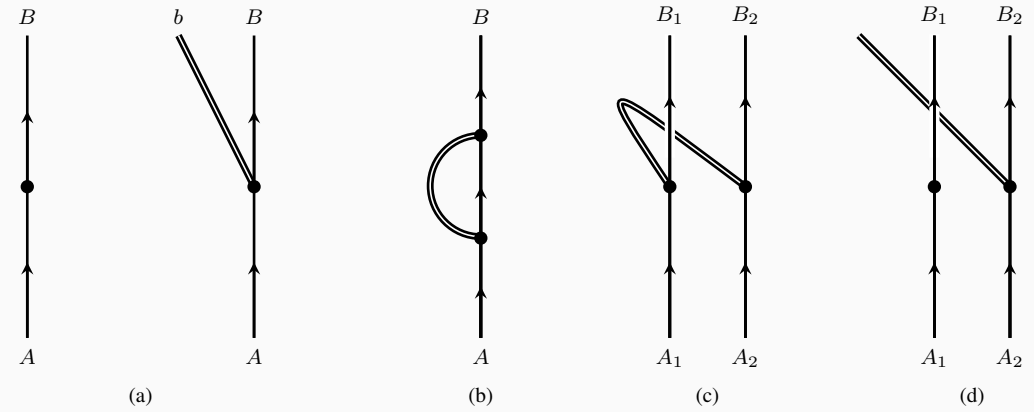


Figure 5.3.7: Construction of $\mathcal{C}_{\mathbb{Z}_2}$: (a) Two types of morphisms from A to B. (b) Composition of two morphisms of the second type. (c) Tensor product of two morphisms of the second type. (d) Tensor product of a morphism of first type on the left and of second type on the right will similarly involve a crossing of b line. And the tensor product of second type on left and first type on right doesn't involve any crossing.

5.4.1 Gu-Wen phases

In this subsection, we will discuss Pin^+ Gu-Wen SPT (f-SPT) phases with global symmetry G. Gu-Wen fermionic SPT phases were first described in [157] and explored further in [156].

The twisted spherical fusion category for these phases is such that $\mathcal{C}_g$ has two simple objects $L_{g,0}$ and $L_{g,1}$ for any g in $G \times \mathbb{Z}_2^R$. The fusion rule is

$$L_{g,\epsilon} \otimes L_{g',\epsilon'} \simeq L_{gg',\epsilon+\epsilon'+n_2(g,g')} \tag{5.4.1}$$

where n_2 is a $\mathbb{Z}_2$-valued group cocycle, *i.e.* it is an element of $H^2(B(G \times \mathbb{Z}_2^R), \mathbb{Z}_2)$. $H^2(B(G \times \mathbb{Z}_2^R), \mathbb{Z}_2)$ is also the group of central extensions of the form

$$0 \to \mathbb{Z}_2 \to \widehat{G} \to G \times \mathbb{Z}_2^R \to 0 \tag{5.4.2}$$

Thus, we can view $\mathcal{C}$ as descending from $\widehat{\mathcal{C}}$ which is a $\widehat{G}$-graded category with a single simple object in each grade. One obtains $\mathcal{C}$ by forgetting the sub-grading corresponding to the $\mathbb{Z}_2$ subgroup appearing in the above central extension. More physically, $\widehat{\mathcal{C}}$ can be viewed as generalizing the notion of unoriented bosonic SPT phases to bosonic SPT phases with more complicated structure group. Forgetting the $\mathbb{Z}_2$ grading corresponds to gauging the $\mathbb{Z}_2$ symmetry. The associator of elements in $\mathcal{C}$ can be read from the associator in $\widehat{\mathcal{C}}$ which we denote as $\widehat{\alpha}_3$. It is an element of $H^3(B\widehat{G}, U(1)_\rho)$. As a note, we will denote an arbitrary element of $G \times \mathbb{Z}_2^R$ by g in what follows.

We demand the existence a fermionic line Π in the twisted Drinfeld center of $\mathcal{C}$ which fuses with itself to the identity. For the 1-form symmetry generated by this line to be compatible with G, the line must be of the form $(L_{e,0}, \beta)$ or $(L_{e,1}, \beta)$. The former case cannot lead to a fermionic line. Hence, Π must be of the form $(L_{e,1}, \beta)$. The existence of such a line will

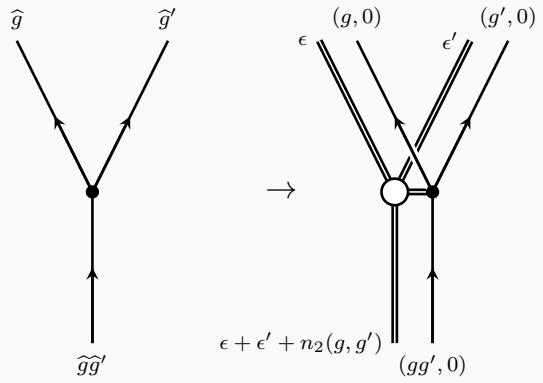
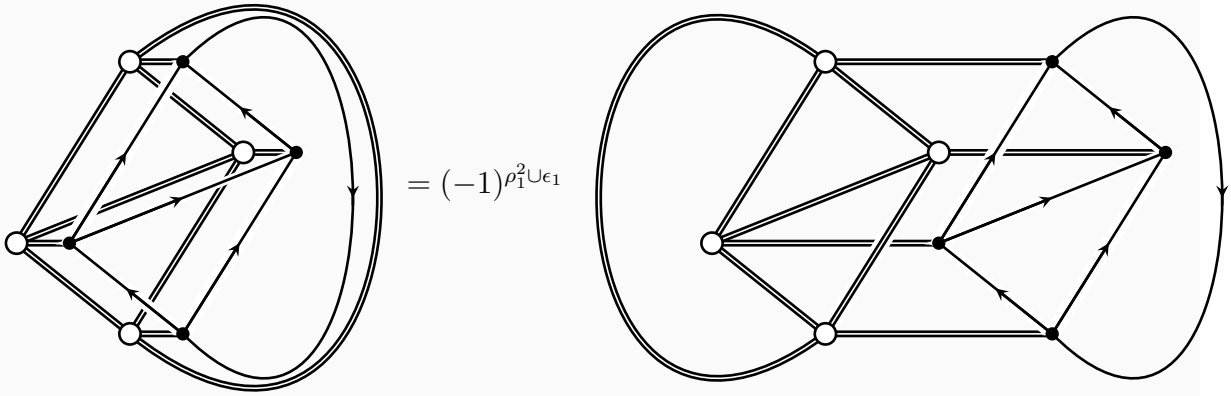


Figure 5.4.1: We choose our basis for morphism space $L_{gg,\epsilon+\epsilon'+n_2(g,g')} \to L_{g,\epsilon} \otimes L_{g',\epsilon'}$ such that the basis for different (ϵ, ϵ') are related by crossing of a Π line as shown in the figure. Here a label ϵ adjacent to double line denotes that the double line is Π if $\epsilon = 1$ and the double line is the identity line if $\epsilon = 0$.



(a) (b)

Figure 5.4.2: Intermediate computational steps relating $\widehat{\alpha}_3$ and ν_3. The sign arises from dragging a Π line over a vertex. We can further resolve the crossing on the right hand side to make contact with ν_3 which is defined in Figure 5.4.3.

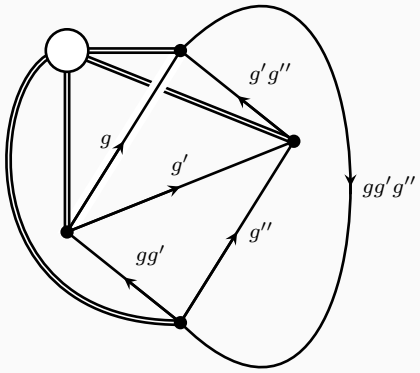


Figure 5.4.3: The definition of $\nu_3(g, g', g'')$.

put some constraints on the form of $\mathcal{C}$ which we now explore. First, we choose our basis of morphisms as shown in the Figure 5.4.1. Consider the basic graph dual to the tetrahedron. Using our basis, it can be written as in Figure 5.4.2(a). This, in turn, can be manipulated to the final graph shown in Figure 5.4.3 which we define to be $\nu_3(g, g', g'')$. During this manipulation we obtain a sign from resolving a crossing and another sign from moving a Π line across a vertex. See Figure 5.4.2(b). Thus, we see that

$$\widehat{\alpha}_3 = \nu_3(-1)^{(n_2+\rho_1^2)\cup\epsilon_1} \tag{5.4.3}$$

where ϵ_1 is a $\mathbb{Z}_2$-valued co-chain which sends (g, ϵ) to ϵ.

We find that a Pin$^+$ Gu-Wen phase is specified by a double (ν_3, n_2) where ν_3 satisfies

$$\delta\nu_3 = (-1)^{n_2\cup n_2+\rho_1^2\cup n_2} \tag{5.4.4}$$

However, there is a redundancy in such a description. We will see in subsection 5.4.3 that the phase defined by $\nu_3 = 1$ and $n_2 = \rho_1^2$ is the same as the trivial phase specified by $\nu_3 = 1$ and $n_2 = 0$.

To completely specify the Pin$^+$ Gu-Wen phase, we also need to pick a specific Π line. The twisted Drinfeld center equations (see Figure 5.3.5 with $i = (g, \epsilon)$ and $j = (g', \epsilon')$) for such a

line tell us that

$$\begin{aligned}\widehat{\alpha}_3(g,\epsilon;g',\epsilon';e,1)\beta(g',\epsilon')\widehat{\alpha}_3^{-1}(g,\epsilon;e,1;g',\epsilon')\beta(g,\epsilon)\widehat{\alpha}_3(e,1;g,\epsilon;g',\epsilon') \\ = (-1)^{\rho_1(g)\rho_1(g')}\beta(gg',\epsilon+\epsilon'+n_2(g,g'))\end{aligned} \tag{5.4.5}$$

Using (5.4.3), we see that it reduces to

$$\beta(gg',\epsilon+\epsilon'+n_2(g,g')) = (-1)^{n_2(g,g')}\beta(g,\epsilon)\beta(g',\epsilon') \tag{5.4.6}$$

Using the fact that Π is fermion tells us that

$$\beta(g,\epsilon+1) = -\beta(g,\epsilon) \tag{5.4.7}$$

Feeding it back, we obtain that

$$\beta(gg',0) = \beta(g,0)\beta(g',0) \tag{5.4.8}$$

The only solution to this equation that works uniformly for any group G is

$$\beta(g,\epsilon) = (\pm 1)^{\rho_1(g)}(-1)^{\epsilon} \tag{5.4.9}$$

Since flipping the sign of all the β in the orientation reversing sector doesn't change the resulting Pin$^+$-TFT, we can choose $L_{e,1}$ equipped with

$$\beta(g,\epsilon) = (-1)^{\epsilon} \tag{5.4.10}$$

as the fermion.

For the rest of this subsection, we note that we can write $H^2(B(G\times\mathbb{Z}_2^R),\mathbb{Z}_2)$ in terms of group cohomology of G. Let's denote an arbitrary element of $G\times\mathbb{Z}_2^R$ as g_1, g_2 etc. We also denote an arbitrary element of G as g and R as the generator of $\mathbb{Z}_2^R$. We have the gauge transofrmations

$$n_2(g_1,g_2) \to n_2(g_1,g_2) + n_1(g_1) + n_1(g_1g_2) + n_1(g_2) \tag{5.4.11}$$

Pick n_1 such that $n_1(g) = 0$ for all g and $n_1(R) + n_1(gR) = n_2(g,R)$. Thus we have fixed a gauge such that $n_2(g,R) = 0$ for all g. Then using the cocycle condition

$$n_2(g_2,g_3) + n_2(g_1g_2,g_3) + n_2(g_1,g_2g_3) + n_2(g_1,g_2) = 0 \tag{5.4.12}$$

we find that we can express n_2 as

$$n_2 = \tilde{m}_2 + \rho_1\cup\tilde{m}_1 + \rho_1\cup\rho_1 \tag{5.4.13}$$

where m_2 parametrizes an element of $H^2(BG,\mathbb{Z}_2)$, m_1 parametrizes an element of $H_1(BG,\mathbb{Z}_2)$ and $\tilde{m}_{1,2}$ denotes the pullback of $m_{1,2}$ from G to $G\times\mathbb{Z}_2^R$. This analysis establishes that

$$H^2(B(G\times\mathbb{Z}_2^R),\mathbb{Z}_2) = H^2(BG,\mathbb{Z}_2)\times H^1(BG,\mathbb{Z}_2)\times\mathbb{Z}_2 \tag{5.4.14}$$

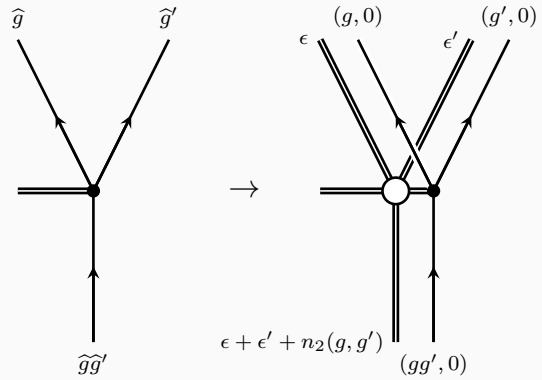


Figure 5.4.4: Gauge fixing: We choose basis for various morphisms related in the way shown in the figure.

5.4.2 Anomaly for Pin$^+$-shadows

In this subsection we will compute the partition function $Z_f(M,\beta_2)$ for a Pin$^+$ Gu-Wen phase. The explicit expression will allow us to compute the anomaly under a gauge transformation $\beta_2 \to \beta_2 + \delta\lambda_1$. As the anomaly is universal, this will justify our prescription (5.3.6) for constructing Pin$^+$-TFTs in terms of their shadows.

In the presence of a background β_2, the basic tetrahedron graph is as shown in Figure 5.4.5(a). This can be gauge fixed as shown in Figure 5.4.4. After the gauge fixing, we can move the Π lines to the position shown in Figure 5.4.5(b). This implies that the partition function can be written as

$$\frac{1}{2^v}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=n_2+\beta_2} (-1)^{\int_M n_2\cup\epsilon_1+w_1^2\cup\epsilon_1+\epsilon_1\cup\beta_2} \tag{5.4.15}$$

This expression is non-zero only when the G-connection is such that $n_2 = \beta_2 + \delta\alpha_1$. By shifting $\epsilon_1 \to \epsilon_1 + \alpha_1$, the above expression can be re-written as

$$\frac{1}{2^v}(-1)^{\int_M n_2\cup\alpha_1+\alpha_1\cup n_2+\alpha_1\cup\delta\alpha_1+w_1^2\cup\alpha_1}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=0} (-1)^{\int_M n_2\cup\epsilon_1+\epsilon_1\cup n_2+w_1^2\cup\epsilon_1} \tag{5.4.16}$$

The sign inside the sum is exact and hence we obtain

$$Z_f(M,\beta_2) = \frac{1}{2^v}(-1)^{\int_M n_2\cup\alpha_1+\alpha_1\cup n_2+\alpha_1\cup\delta\alpha_1+w_1^2\cup\alpha_1}\prod \nu_3 \sum_{\epsilon_1|\delta\epsilon_1=0} 1 \tag{5.4.17}$$

Shifting $\beta_2 \to \beta_2 + \delta\lambda_1$ is the same as shifting $\alpha_1 \to \alpha_1 + \lambda_1$ under which the partition function changes as

$$Z_f(M,\beta_2+\delta\lambda_1) = (-1)^{\int_M \lambda_1\cup\beta_2+\beta_2\cup\lambda_1+\lambda_1\cup\delta\lambda_1+w_1^2\cup\lambda_1} Z_f(M,\beta_2) \tag{5.4.18}$$

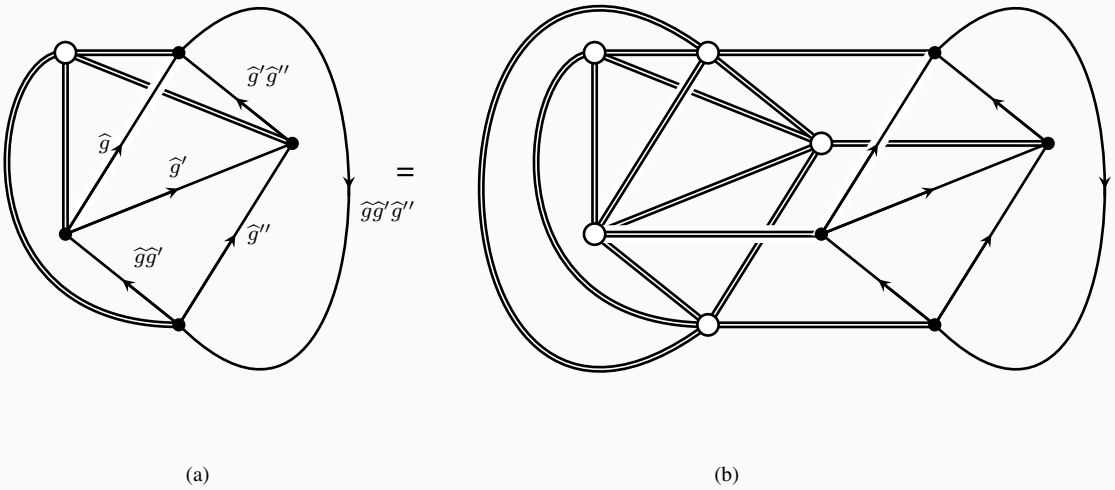


Figure 5.4.5: The graph on the right can be obtained from graph on the left after gauge fixing and deforming the Π lines.

which matches the expectation in (5.3.10) exactly.

5.4.3 Group structure of Gu-Wen phases

Now we would like to compute the product of two Gu-Wen phases labeled by (ν_3, n_2) and (ν_3', n_2'). The G-graded product of corresponding categories has 4 simple objects in each grade $L_{g,\epsilon,\epsilon'}$ which fuse according to the cocycle (n_2, n_2') and have associators $\widehat{\alpha}_3\widehat{\alpha}_3'$. The non-anomalous $\mathbb{Z}_2$ 1-form symmetry is generated by $L_{e,1,1}$ which has crossing $(-1)^{\epsilon+\epsilon'}$.

Gauging the symmetry identifies $L_{g,\epsilon,\epsilon'}$ with $L_{g,\epsilon+1,\epsilon'+1}$. We pick representative objects $L_{g,\epsilon,0}$ in each grade and compute the associator of $L_{g,\epsilon,0}$, $L_{g',\epsilon',0}$ and $L_{g'',\epsilon'',0}$ via the tetrahedron graph. Multiplying two representative objects $L_{g,\epsilon,0}$ and $L_{g',\epsilon',0}$, we obtain $L_{gg',\epsilon+\epsilon'+n_2(g,g'),n_2'(g,g')}$ which can be mapped back to the representative object $L_{gg',\epsilon+\epsilon'+n_2(g,g')+n_2'(g,g'),0}$ by inserting $n_2'(g,g')$ number of $\Pi\Pi'$ lines emanating from the corresponding vertex. The representative objects thus fuse according to the cocycle $n_2 + n_2'$. Now, we gauge fix as in the previous subsection. Then, doing same manipulations as in the previous subsection, we find that the tetrahedron graph evaluates to

$$\nu_3\nu_3'(-1)^{n_2\cup\epsilon_1+w_1^2\cup\epsilon_1+\epsilon_1\cup n_2'} \tag{5.4.19}$$

Upto a gauge redefintion, it can be written as

$$(\nu_3\nu_3'(-1)^{n_2\cup_1 n_2'})(-1)^{(n_2+n_2'+w_1^2)\cup\epsilon_1} \tag{5.4.20}$$

Thus the product is a Gu-Wen phase with $\tilde{\nu}_3 = \nu_3\nu_3'(-1)^{n_2\cup_1 n_2'}$ and $\tilde{n}_2 = n_2 + n_2'$.

However, notice that substituting $\nu_3 = 1, n_2 = 0$ in (5.4.17) and writing $w_1^2 = \delta\sigma_1$ gives

$$Z_f(M,\beta_2) = \frac{1}{2^v}(-1)^{\int_M \alpha_1\cup\delta\alpha_1+\delta\sigma_1\cup\alpha_1}\sum_{\epsilon_1|\delta\epsilon_1=0} 1 \tag{5.4.21}$$

and substituting $\nu_3 = 1, n_2 = \rho_1^2$ gives

$$Z_f(M,\beta_2) = \frac{1}{2^v}(-1)^{\int_M(\alpha_1+\sigma_1)\cup\delta\alpha_1}\sum_{\epsilon_1|\delta\epsilon_1=0} 1 \tag{5.4.22}$$

which are the same expressions! Thus, the Gu-Wen phase labeled by $(\nu_3 = 1, n_2 = \rho_1^2)$ is the trivial phase. The reader might complain that (5.4.21) does not seem to describe a *trivial* phase. We would like to stress that this is the partition function of the shadow theory describing the trivial Pin$^+$-TFT. The trivial Pin$^+$-TFT is obtained by combining a non-trivial shadow with a non-trivial sign.

Thus, the group $\mathcal{GW}(G)$ of Pin$^+$ Gu-Wen phases with global symmetry G can be described as follows. Consider the set parametrized by (ν_3, n_2) with ν_3 parametrizing elements of $H^3(B(G\times\mathbb{Z}_2^R), U(1)_\rho)$ and n_2 parametrizing elements of $H^2(B(G\times\mathbb{Z}_2^R),\mathbb{Z}_2)$. Provide it a group structure given by

$$(\nu_3, n_2)(\nu_3', n_2') = (\nu_3\nu_3'(-1)^{n_2\cup_1 n_2'}, n_2 + n_2') \tag{5.4.23}$$

Finally, quotient it by the $\mathbb{Z}_2$ subgroup generated by $(\nu_3, n_2) = (1, \rho_1^2)$.

An alternative description can be given by first defining a group $H(G) = H^3(B(G \times \mathbb{Z}_2^R), U(1)_\rho)/\mathbb{Z}_2$ where the $\mathbb{Z}_2$ is generated by the cocycle $(-1)^{\rho_1^2 \cup_1 \rho_1^2}$. Then, $\mathcal{GW}(G)$ is a central extension

$$0 \to H(G) \to \mathcal{GW}(G) \to H^2(BG, \mathbb{Z}_2) \times H^1(BG, \mathbb{Z}_2) \to 0 \tag{5.4.24}$$

with cocycle valued in $H(G)$ being $(-1)^{n_2 \cup_1 n_2'} \in H^3(B(G \times \mathbb{Z}_2^R), U(1)_\rho)$ where n_2 and n_2' are valued in $H^2(BG, \mathbb{Z}_2) \times H^1(BG, \mathbb{Z}_2)$ as in (5.4.13) but without the ρ_1^2 summand.

5.4.4 $\mathbb{Z}_2^R$ version of Ising

As an application of our formalism, we would like to construct all Pin$^+$-SPT phases wth global symmetry group G being the trivial group $\{id\}$. There is only one Gu-Wen phase in this class, which is the trivial phase. There is a non-trivial phase in this class which is given by the $\mathbb{Z}_2^R$ analogue of a $\mathbb{Z}_2$ graded spherical fusion category $\mathcal{I}$ which is known as the Ising fusion category. Below we recall the construction of $\mathcal{I}$ and its $\mathbb{Z}_2^R$ cousin. It turns out that the analysis for both the cases is similar and we treat both of them together.

We are looking for a $\mathbb{Z}_2$ graded (twisted) spherical fusion category such that $\mathcal{C}_0 = \{I, P\}$ and $\mathcal{C}_1 = \{S\}$ are the simple objects. The fusion rules are

$$P \otimes P \simeq I \tag{5.4.25}$$
$$S \otimes P \simeq S \tag{5.4.26}$$
$$P \otimes S \simeq S \tag{5.4.27}$$
$$S \otimes S \simeq I \oplus P \tag{5.4.28}$$

The F-symbols can be bootstrapped from these fusion rules by using (twisted) pentagon equation and taking advantage of the gauge freedom.

When the $\mathbb{Z}_2$ grading corresponds to a $\mathbb{Z}_2$ global symmetry, the non-trivial F-symbols are determined to be

$$(F_S^{PSP})_{(S)(S)} = -1 \tag{5.4.29}$$
$$(F_P^{SPS})_{(S)(S)} = -1 \tag{5.4.30}$$
$$(F_S^{SSS})_{(I)(I)} = (F_S^{SSS})_{(P)(I)} = (F_S^{SSS})_{(I)(P)} = \pm\frac{1}{\sqrt{2}} \tag{5.4.31}$$
$$(F_S^{SSS})_{(P)(P)} = \mp\frac{1}{\sqrt{2}} \tag{5.4.32}$$

When the $\mathbb{Z}_2$ grading corresponds to $\mathbb{Z}_2^R$ orientation reversing symmetry, the non-trivial F-

symbols are determined to be

$$(F_S^{PSP})_{(S)(S)} = -1 \tag{5.4.33}$$

$$(F_P^{SPS})_{(S)(S)} = -1 \tag{5.4.34}$$

$$(F_S^{SSS})_{(I)(I)} = (F_S^{SSS})_{(P)(I)} = (F_S^{SSS})_{(I)(P)} = \frac{1}{\sqrt{2}} \tag{5.4.35}$$

$$(F_S^{SSS})_{(P)(P)} = -\frac{1}{\sqrt{2}} \tag{5.4.36}$$

That is, the choice of sign becomes a gauge freedom in the $\mathbb{Z}_2^R$ case.

The fermion line is given by an element of (twisted) Drinfeld center of the form (P, β). Solving the Drinfeld center equations for the $\mathbb{Z}_2$ case, we obtain

$$\beta_P = -1 \tag{5.4.37}$$

$$\beta_S = \pm i \tag{5.4.38}$$

Thus, there are two choices for the fermion line Π. Given the choice in picking the associator and the choice in Π, we can construct four Spin-TFTs with global symmetry $\mathbb{Z}_2$.

On the other hand, solving the twisted Drinfeld center equations for the $\mathbb{Z}_2^R$ case, we obtain

$$\beta_P = -1 \tag{5.4.39}$$

$$\beta_S = \pm 1 \tag{5.4.40}$$

However, as we know from before, flipping the sign of all the β in the orientation reversing sector doesn't change the resulting Pin$^+$-TFT and we can fix $\beta_S = +1$. Hence, in the $\mathbb{Z}_2^R$ case, there are no choices and we obtain only one Pin$^+$-TFT which we call $\mathcal{I}_+$.

5.4.5 Pin$^+$-SPT phases with no global symmetry

Cobordism hypothesis predicts a $\mathbb{Z}_2$ group of Pin$^+$-SPT phases [128]. We have already found the trivial phase as a Gu-Wen phase. We claim that the non-trivial phase corresponds to the Pin$^+$-TFT $\mathcal{I}_+$ that we encountered in last subsection. To justify this, we will show that the square of $\mathcal{I}_+$ is the trivial Gu-Wen phase. This will prove that $\mathcal{I}_+$ is indeed an SPT phase and provide an explicit construction of Pin$^+$-SPT phases without global symmetry. The existence of this non-trivial phase was also discussed in [158].

The graded product of $\mathcal{I}_+$ with itself has simple objects II, PI, IP, PP in the trivial grade and a simple object SS in the non-trivial grade. Gauging the 1-form symmetry generated by $\Pi\Pi$, we obtain a category $\mathcal{C}$ with $\mathcal{C}_0$ having simple objects II, PI and $\mathcal{C}_1$ having simple objects SS^+, SS^-. SS^+ and SS^- are constructed by using projectors obtained by using the non-trivial endomorphism of SS. See Figure 5.4.6.

$SS^+ \otimes PI$ involves the F-symbol F^{PSP} which flips the sign of ξ_S and hence $SS^+ \otimes PI \simeq SS^-$. On the other hand, $PI \otimes SS^+$ involves β_P and hence $PI \otimes SS^+ \simeq SS^-$. The computation

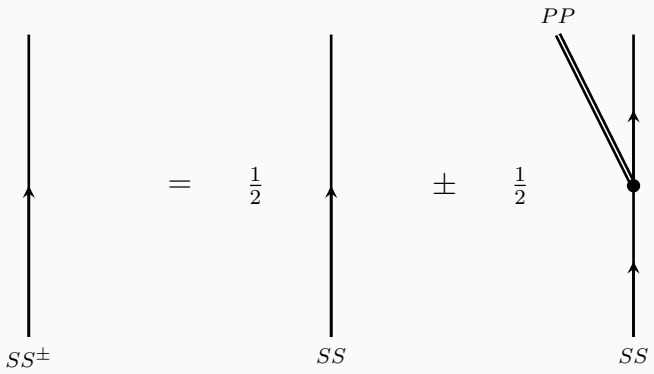


Figure 5.4.6: Definition of $SS^{\pm}$.

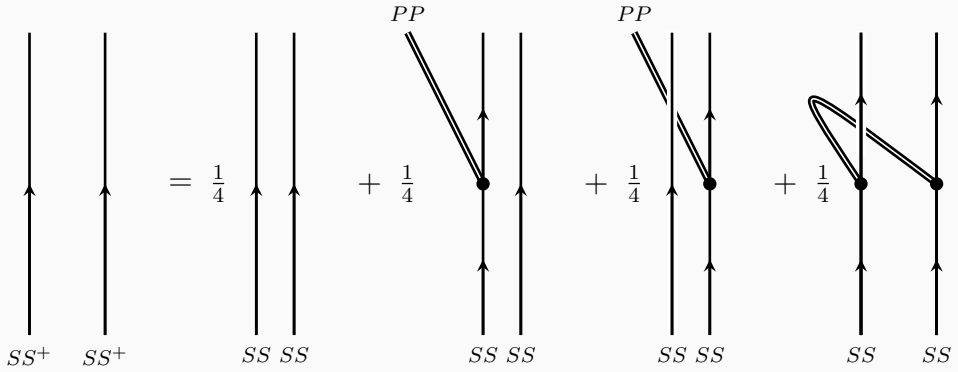


Figure 5.4.7: Computation of $SS^+ \otimes SS^+$ is by definition a sum of four terms which involve associators and crossings. II inside $SS \otimes SS$ is mapped to II by first and fourth terms and to PP (which is isomorphic to II in the new category) by the second and third terms. Similarly, PI is mapped to the zero object as the four terms cancel in pairs. Hence, $SS^+ \otimes SS^+ \simeq II$.

of $SS^+ \otimes SS^+$ can be done in a similar but more involved manner which we explain in Figure 5.4.7. We find that $SS^+ \otimes SS^+ \simeq II$. All the statements above hold true if we replace SS^+ with SS^-. Thus, $\mathcal{C}$ has the fusion rules of the Gu-Wen phase which is trivial.

For a general G, we can consider the pullback of $\mathcal{I}_+$ along ρ_1 which we denote as $\mathcal{I}_+(G)$. $\mathcal{I}_+(G)_g$ has two simple elements I_g, P_g if $\rho_1(g) = 0$ and has a single simple object S_g if $\rho_1(g) = 1$. The fusion rules and associators are just pulled back from $\mathcal{I}_+$. Clearly, $\mathcal{I}_+(G)$ will also square to 0 as our argument above is independent of G-grading.

This allows us to construct $\mathcal{GW}(G) \times \mathbb{Z}_2$ worth of Pin$^+$-SPT phases with global symmetry G. We suspect that this is not the full classification and comment on how to complete the classification in the next section.

5.5 Conclusion and future directions

In this chapter we discussed the generalization of Turaev-Viro construction of oriented 3d TFTs to unoriented 3d TFTs. We proposed that the input data of this construction in the unoriented case should be a "twisted" spherical fusion category in which the pentagon equation for the F-symbols is modified.

As a generalization of the construction of [3], we also proposed a construction for Pin$^+$-TFTs in terms of their shadows. The shadows are ordinary unoriented TFTs with a $\mathbb{Z}_2$ 1-form symmetry which is anomalous and has a mixed anomaly with time-reversal symmetry.

Combining the above two ingredients, we were able to give explicit constructions of a large class of invertible Pin$^+$-TFTs with global symmetry G. Such theories are known as Pin$^+$-SPT phases. We also reproduced the $\mathbb{Z}_2$ group of Pin$^+$-SPT phases without any global symmetry.

There are plenty of interesting directions in which this work can be extended in the future and we make some very speculative comments about them in what follows. Perhaps the most immediate future direction is to use the machinery developed in this chapter to provide a classification of Pin$^+$-SPT phases for an arbitrary group G which admit a topological boundary condition. The author suggests to look at a spherical fusion category graded by $\mathbb{Z}_2 \times \mathbb{Z}_2^R$ with simple elements I, P in the $(0,0)$ grade, I_1, P_1 in the $(0,1)$ grade, S in the $(1,0)$ grade and S_1 in the $(1,1)$ grade. The fusion rules mimic the Ising category. Is it possible to find a consistent set of F-symbols? If yes, then the class of Pin$^+$-SPT phases we presented in this chapter is not the full answer. It should then be possible to finish the classification, in a spirit similar to the one in [3], by pulling back this $\mathbb{Z}_2 \times \mathbb{Z}_2^R$ phase and combining it with the class of phases presented in this chapter.

It would be very interesting to provide a construction (Turaev-Viro-like or some other construction) for TFTs with more general struture groups. For instance, one could mix $O(n)$ and G or mix Pin$^+(n)$ and G in the fermionic case. It seems that a proper treatment of these generalizations should involve a rich interplay of symmetry defects along with higher codimension defects living in the worldvolume of symmetry defects.

Let us comment about the Pin$^-(n) \times G$ case. It seems natural that the kernel for Pin$^-$-TFTs would be the sign

$$z_-(M, \eta_1, \beta_2) = (-1)^{\int_M \eta_1 \cup \beta_2 + \int_N \beta_2 \cup \beta_2 + (w_1^2 + w_2) \cup \beta_2} \tag{5.5.1}$$

which seems to be the same expression as (5.3.7) but this time we take η_1 to parametrize Pin$^-$-structures. This would suggest that the corresponding shadow theory has no mixed anomaly between time reversal and $\mathbb{Z}_2$ 1-form symmetry. Also, the anomaly for the 1-form symmetry should now be

$$Z_f(M, \beta_2) \to (-1)^{\int_M \lambda_1 \cup \beta_2 + \beta_2 \cup \lambda_1 + \lambda_1 \cup \delta \lambda_1} Z_f(M, \beta_2) \tag{5.5.2}$$

However, for Pin$^+$ case, we saw in Figure 5.3.4 that moving the fermion Π across the locus dual to w_1^2 should change the operator at the junction of Π line and the orientation reversing defects. The argument given there was that this sign was needed to cancel the sign coming from the crossing of Π lines. This lead to different anomalies than the ones we want for the Pin$^-$ case. So, in the Pin$^-$ case, we do not want such a change in the sign of the corresponding local operator. The author suspects that in this case the sign coming from crossing of Π lines will be canceled by factors coming from *patching* of Π with $R\Pi$ where $R\Pi$ is Π line with a reflected framing. This

would make sure that Π is an element of Drinfeld center rather than a twisted Drinfeld center, which would in turn imply the anomalies given above. It would be interesting to work out the details and provide a Turaev-Viro construction for Pin^- shadows.

Of course, this means that one will have to first understand how to compute (in terms of the twisted spherical fusion category) the extra data attached to a bulk line which corresponds to patching the line with itself but with reflected framing. In other words, this corresponds to a generalization of Moore-Seiberg data [159, 160] to the unoriented case. A step towards this was recently taken in [161].

A puzzle here is that there should be no non-trivial Pin^--SPT phase according to [128]. So, somehow the Pin^--TFTs produced by the potential Pin^--shadows having $\mathbb{Z}_2^R$ version of Ising as their twisted spherical fusion category should be trivial.

Another interesting direction to pursue would be to see if it is possible to find a generalization of Turaev-Viro construction which could construct anomalous 3d TFTs. Such TFTs live at the boundary of a 4d SPT phase. Hence, such TFTs should not admit topological boundaries of their own but they can admit interfaces to other 3d TFTs with the same anomaly. Perhaps it is possible to choose a simple TFT in each anomaly class and build a Turaev-Viro construction using a topological interface between the TFT we want to construct and the simple TFT. See [152, 149, 162, 163] for recent interesting work on anomalous unoriented 3d TFTs.

Finally, it would be interesting to concretely construct a time-reversal invariant commuting projector Hamiltonian using the data of twisted spherical fusion category. This Hamiltonian goes into the string-net construction of fermionic phases of matter. See [126], [3] for more details.

Chapter 6

On finite symmetries and their gauging in two dimensions

6.1 Introduction

Let us start by considering a two-dimensional theory T with $\mathbb{Z}_n$ symmetry. We can gauge it to get the gauged theory $T/\mathbb{Z}_n$. This gauged theory is known to have a new $\mathbb{Z}_n$ symmetry, and re-gauging it gives us back the original theory: $T/\mathbb{Z}_n/\mathbb{Z}_n = T$ [13]. It is also well-known that this phenomenon generalizes to any arbitrary finite Abelian group G. That is, gauging a theory T with G symmetry results in a theory T/G with a new finite Abelian group symmetry $\widehat{G}$ such that gauging the new theory by the new symmetry takes us back to the original theory: $T/G/\widehat{G} = T$. One natural question arises: can it be generalized to higher dimensions? Yes, according to [150], where the generalized concept of p-form symmetries was introduced. Our investigation starts from a related but different question: can it be generalized to non-Abelian finite groups?

The answer is again yes [164], and it again requires a generalization of the concept of symmetries, but in a direction different from that of [150]. To explain this, let us pose what we have said above in a different way. Traditionally, we say that T has symmetry G if we can find unitary operators U_g labeled by elements $g \in G$ whose action on the Hilbert space commutes with the Hamiltonian. When G is Abelian, the information about G is captured in T/G via unitary operators $U_{\widehat{g}}$ labeled by elements $\widehat{g} \in \widehat{G}$ which commute with the Hamiltonian of T/G. When G is non-Abelian, we will argue that the information of G in T/G is captured by operators U_i which still commute with the Hamiltonian but these operators are now in general non-unitary. These operators can be constructed by wrapping a Wilson line for G along the spatial circle. Hence, Wilson lines should be thought of as generalized symmetries for the theory T/G. In fact, we will also argue that there is a natural notion of gauging this symmetry formed by Wilson lines such that gauging T/G results back in the original theory T.

This raises the following question: How do we specify a generalized symmetry that a theory can admit? In this chapter, we give an answer to this question: a general finite symmetry of a two-dimensional theory is specified by a structure which is known to mathematicians in the

name of *unitary fusion categories*. We prefer to call it *symmetry categories*.[1] For the gauged theory T/G for possibly non-Abelian group G, the Wilson line operators form $\mathrm{Rep}(G)$, which is a symmetry category formed by the representations of G. Similarly, a general symmetry category $\mathcal{C}$ physically corresponds to more general line operators of T.

We also discuss how a symmetry category $\mathcal{C}$ can be gauged. It turns out that there is no canonical way of gauging a generic symmetry category. Pick one way $\mathcal{M}$ of gauging the symmetry $\mathcal{C}$ of a theory T. Denote the gauged theory by $T/\mathcal{M}$ and its symmetry category by $\mathcal{C}'$. It then turns out that there exists a dual way $\mathcal{M}'$ of gauging $\mathcal{C}'$ such that $T/\mathcal{M}/\mathcal{M}'$ is equivalent to T. This generalizes the fact that regauging the gauged theory with symmetry $\mathrm{Rep}(G)$ gives back the original theory with symmetry G.

This generalization of the concept of symmetry allows us not only to perform the re-gauging of non-Abelian gauge theories, but also to answer various other questions. First of all, we will see that symmetry categories $\mathcal{C}$ capture symmetries together with their anomalies. Then, the machinery we spell out allows us to compute what is the symmetry of the gauge theory T/H when we gauge a subgroup H of a possibly anomalous flavor symmetry G. For example, if we gauge a non-anomalous $\mathbb{Z}_2$ subgroup of $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ with a suitable choice of the anomaly, we can get non-anomalous non-Abelian symmetry D_8 and Q_8, the dihedral group and the quaternion group of order 8.[2] In general, the symmetry $\mathcal{C}$ of the gauged theory is neither a group nor $\mathrm{Rep}(G)$ for a finite group, and we need the concept of symmetry categories to describe it.

There are also vast number of symmetry categories not related to finite groups, formed by topological line operators of two-dimensional rational conformal field theories (RCFTs). In particular, any unitary modular tensor category, or equivalently any Moore-Seiberg data, can be thought of as a symmetry category, by forgetting the braiding.

We should emphasize here that this generalization of the concept of symmetry from that defined by groups to that defined by categories was already done long ago by other authors, belonging to three somewhat independent lines of studies, namely in the study of the subfactors and operator algebraic quantum field theories, in the study of representation theory, and in the study of RCFTs. Each of the communities produced vast number of papers, and not all of them can be cited here. We recommend to the readers textbooks by Bischoff, Longo, Kawahigashi and

[1] We do not claim that this is the ultimate concept for the 0-form finite symmetry in two dimensions; there still might be a generalization in the future. For example, in other spacetime dimensions, a proposed generalization was to use the concept of p-groups, see e.g. [165], and one might want to unify the two approaches. We also note that this extension of the concept of the symmetry in two dimensions from groups to categories was already proposed by many other authors in the past, and we are merely shedding a light to it from a slightly different direction. We will come back to this point later in the Introduction.

[2] Recently in [166], Gaiotto, Kapustin, Komargodski and Seiberg performed an impressive study of the phase structure of thermal 4d $\mathfrak{su}(2)$ Yang-Mills theory. One important step in the analysis is the symmetry structure of the thermal system, which is essentially three-dimensional. As a dimensional reduction from 4d, the system has a $\mathbb{Z}_2 \times \mathbb{Z}_2$ 0-form symmetry and a $\mathbb{Z}_2$ 1-form symmetry, with a mixed anomaly. Then the authors gauged the $\mathbb{Z}_2$ 1-form symmetry, and found that the total 0-form symmetry is now D_8. This D_8 was then used very effectively to study the phase diagram, but that part of their paper does not directly concern us here. Their analysis of turning an anomalous Abelian symmetry by gauging a non-anomalous subgroup into a non-Abelian symmetry is a 3d analogue of what we explain in 2d. See their Sec. 4.2, Appendix B and Appendix C. Clearly an important direction to pursue is to generalize their and our constructions to arbitrary combinations of possibly-higher-form symmetries in arbitrary spacetime dimensions, but that is outside of the scope of this chapter.

Rehren [167] and Etingof, Gelaki, Nikshych and Ostrik [168] from the first two communities and the articles by Carqueville and Runkel [169] and by Brunner, Carqueville and Plencner [170] from the third community as the starting points.

Our first aim in this chapter is then to summarize the content of these past works in a way hopefully more accessible to other researchers of quantum field theory, emphasizing the point of view related to the modern study of symmetry protected topological phases. What we explain in this first part of the chapter is not new, except possibly the way of the presentation, and can all be found in the literature in a scattered form.

Our second aim is to axiomatize two-dimensional topological quantum field theories (TFTs) whose symmetry is given by a symmetry category $\mathcal{C}$. This is a generalization of the work by Moore and Segal [171], where two-dimensional TFTs with finite group symmetry were axiomatized. We write down basic sets of linear maps between (tensor products of) Hilbert spaces on S^1 and basic consistency relations among them which guarantee that a unique linear map is associated to any surface with m incoming circles and n outgoing circles together with arbitrary network of line operators from $\mathcal{C}$.

The rest of the chapter is organized as follows. First in Sec. 6.2, as a preliminary, we recall how gauging of a finite Abelian symmetry G can be undone by gauging the new finite Abelian symmetry $\widehat{G}$, and then briefly discuss how this can be generalized to non-Abelian symmetries G, by regarding $\mathrm{Rep}(G)$ as a symmetry. This effort of generalizing the story to a non-Abelian group makes the possibility and the necessity of a further generalization to symmetry categories manifest. We exploit this possibility and describe the generalization in detail in subsequent sections.

Second, we have two sections that form the core of the chapter. In Sec. 6.3, we introduce the notion of symmetry categories, and discuss how we can regard as symmetry categories both a finite group G with an anomaly α and the collection $\mathrm{Rep}(G)$ of representations of G. We then explain in Sec. 6.4 that physically distinct gaugings of a given symmetry category $\mathcal{C}$ correspond to indecomposable module categories $\mathcal{M}$ of $\mathcal{C}$, and we describe how to obtain the new symmetry $\mathcal{C}'$ of the theory $T/\mathcal{M}$ for a given theory T with a symmetry $\mathcal{C}$.

Third, in Sec. 6.5, we give various examples illustrating the notions introduced up to this point. Examples include the form of new symmetry categories $\mathcal{C}'$ when we gauge a non-anomalous subgroup H of an anomalous finite group G, and the symmetry categories of RCFTs.

Fourth, in Sec. 6.6, we move on to the discussion of the axioms of two-dimensional TFTs whose symmetry is given by a symmetry category $\mathcal{C}$. We also construct the gauged TFTs $T/\mathcal{M}$ given an original TFT T with a symmetry category $\mathcal{C}$ and a gauging specified by its module category $\mathcal{M}$. Sections 6.5 and 6.6 can be read independently.

Finally, we conclude with a brief discussion of what remains to be done in Sec. 6.7. We have an appendix D.1 where we review basic notions of group cohomology used in the chapter.

Before proceeding, we note that we assume that the space-time is oriented throughout the chapter. We also emphasize that all the arguments we give, except in Sec. 6.6, apply to non-topological non-conformal 2d theories.

6.2 Re-gauging of finite group gauge theories

6.2.1 Abelian case

Let us start by reminding ourselves the following well-known fact [13]:

Let T be a 2d theory with flavor symmetry given by an Abelian group G. Let us assume that G is non-anomalous and can be gauged, and denote the resulting theory by T/G. Then this theory has the flavor symmetry $\widehat{G}$, which is the Pontrjagin dual of G, such that $T/G/\widehat{G} = T$.

Recall the definition of the Pontrjagin dual $\widehat{G}$ of an Abelian group G. As a set, it is given by

$$\widehat{G} = \{\chi : G \to U(1) \mid \chi \text{ is an irreducible representation } \}. \tag{6.2.1}$$

Note that χ is automatically one-dimensional. Therefore the product of two irreducible representations is again an irreducible representation, which makes $\widehat{G}$ into a group. G and $\widehat{G}$ are isomorphic as a group but it is useful to keep the distinction because there is no canonical isomorphism between them.

In the literature on 2d theories, gauging of a finite group G theory is more commonly called as orbifolding by G, and the fact above is often stated as follows: a G-orbifold has a dual $\widehat{G}$ symmetry assigning charges to twisted sectors, and orbifolding again by this dual $\widehat{G}$ symmetry we get the original theory back. This dual $\widehat{G}$ symmetry is also known as the quantum symmetry in the literature.

This fact can be easily shown as follows. Let $Z_T[M, A]$ denote the partition function of T on M with the external background G gauge field A. Here A can be thought of as taking values in $H^1(M, G)$. Then the partition function of the gauged theory T/G on M is given by $Z_{T/G}[M] \propto \sum_A Z_T[M, A]$. Here and in the following we would be cavalier on the overall normalization of the partition functions. More generally, with the background gauge field B for the dual $\widehat{G}$ symmetry, the partition function is given by

$$Z_{T/G}[M, B] \propto \sum_A e^{i(B,A)} Z_T[M, A] \tag{6.2.2}$$

where $B \in H^1(M, \widehat{G})$ and $e^{i(B,A)}$ is obtained by the intersection pairing

$$e^{i(-,-)} : H^1(M, \widehat{G}) \times H^1(M, G) \to H^2(M, U(1)) \simeq U(1). \tag{6.2.3}$$

The equation (6.2.2) says that the partition function of T/G is essentially the discrete Fourier transform of that of T, and therefore we dually have $T = T/G/\widehat{G}$:

$$Z_T[M, A] \propto \sum_B e^{i(A,B)} Z_{T/G}[M, B]. \tag{6.2.4}$$

This statement was generalized to higher dimensions in e.g. [150]:

Let T be a d-dimensional theory with p-form flavor symmetry given by an Abelian group G. Let us assume that G is non-anomalous and can be gauged, and denote the resulting theory by T/G. Then this theory has the dual $(d-2-p)$-form flavor symmetry $\widehat{G}$, such that $T/G/\widehat{G} = T$.

The derivation is entirely analogous to the 2d case, except that now $A \in H^{p+1}(M, G)$ and $B \in H^{d-1-p}(M, \widehat{G})$.

6.2.2 Non-Abelian case

The facts reviewed above means that the finite Abelian gauge theory T/G still has the full information of the original theory T, which can be extracted by gauging the dual symmetry $\widehat{G}$. It is natural to ask if this is also possible when we have a non-Abelian symmetry G, which we assume to be an ordinary 0-form symmetry.

This is indeed possible[3] by suitably restating the derivation above, but we will see that we need to extend the concept of what we mean by *symmetry*. To show this, we first massage (6.2.4) in a suitable form which admits a straightforward generalization. Let us consider the case of $Z_T[M, A = 0]$ for illustration. By Poincare duality, B can also be represented as an element of $H_1(M, \widehat{G})$. Then, (6.2.4) can be rewritten as

$$Z_T[M] \propto \sum_{\widehat{g}_1, \cdots, \widehat{g}_n} Z_{T/G}[M, \widehat{g}_1, \cdots, \widehat{g}_n] \tag{6.2.5}$$

where $i \in \{1, \cdots, n\}$ labels generators of $H_1(M)$ and $\widehat{g}_i$ is an element of $\widehat{G}$ associated to the cycle labeled by i. Each summand on the right hand side, $Z_{T/G}[M, \widehat{g}_1, \cdots, \widehat{g}_n]$, is then the expectation value of Wilson loops in representations labeled by $\widehat{g}_i$ placed along the cycle i.

Now, we can sum the $\widehat{G}$ elements for each i separately to obtain

$$Z_T[M] \propto Z_{T/G}[M, W_1^{\text{reg}}, \cdots, W_n^{\text{reg}}] \tag{6.2.6}$$

where W_i^{reg} denotes the insertion of a Wilson line in the regular representation along the cycle i. This is because, for an abelian G, the regular representation is just the sum of representations corresponding to elements $\widehat{g}$ of $\widehat{G}$.

The relation (6.2.6) says that by inserting all possible Wilson lines on all possible cycles, we are putting the delta function for the original gauge field A. We now note that the relation (6.2.6) holds for a non-Abelian G as well, if we insert W_i^{reg} not only for the generators of $H_1(M)$ but for the generators of $\pi_1(M)$. This can be seen by the fact that $\mathrm{tr}\, g$ in the regular representation is nonzero if and only if g is the identity.

[3] That this is possible was already shown for two-dimensional theories in [164], as an example of a much more general story, which we will also review in the forthcoming sections. Here we describe the construction in an elementary language.

The identity (6.2.6) means that the ungauged theory T can be recovered from the gauged theory T/G by inserting line operators W_i^{reg} in an appropriate manner. This is analogous to the construction of the gauged theory T from the ungauged theory T by inserting line operators representing the G symmetry in an appropriate manner. Given the importance of Wilson lines in recovering the information of the ungauged theory, we assign the status of dual *symmetry* to Wilson lines.

Let us phrase it another way. When G is Abelian, the dual $(d-2)$-form $\widehat{G}$ symmetries can be represented by 1-cycles labeled by elements of $\widehat{G}$, forming a group. When G is non-Abelian, the dual *symmetry* can still be represented by 1-cycles labeled by representations $\text{Rep}(G)$ of G. We can still multiply lines, corresponding to the tensor product of the representations, and this operation reduces to group multiplication of $\widehat{G}$ in the abelian case. But $\text{Rep}(G)$ is not a group if G is non-Abelian. Therefore this is not a flavor symmetry *group*, it is rather a flavor symmetry *something*. We summarize this observation as follows:

Let T be a d-dimensional theory with 0-form flavor symmetry given by a possibly non-Abelian group G. Let us assume that G is non-anomalous and can be gauged, and denote the resulting theory by T/G. Then this theory has $\text{Rep}(G)$ as the dual $(d-2)$-form flavor symmetry 'something', such that $T/G/\text{Rep}(G) = T$.

6.3 Symmetries as categories in two dimensions

Any finite group, possibly non-Abelian, can be the 0-form symmetry *group* of a theory. In addition to this, we saw in the last section that $\text{Rep}(G)$, the representations of G, can also be the $(d-2)$-form symmetry *something* of a d-dimensional theory. We do not yet have a general understanding of what should be this *something* in general d dimensions for general combinations of various p-form symmetries. However, at least for d=2 and p=0, we already have a clear concept for this *something* in the literature, which includes both groups G and representations of groups $\text{Rep}(G)$ and much more. In this section we explain what it is.

6.3.1 Basic notions of symmetry categories

In two dimensions, a 0-form finite symmetry element can be represented by a line operator with a label a. Inserting this line operator on a space-like slice S corresponds to acting on the Hilbert space associated to S by a possibly non-unitary operator U_a corresponding to the symmetry element. Moreover, U_a commutes with the Hamiltonian H associated to any foliation of the two-dimensional manifold. In addition, U_a cannot change under a continuous deformation of its path. Therefore the line operator under consideration is automatically topological.

Topological line operators, in general, form a structure which mathematicians call a *tensor category*. We want to restrict our attention to topological line operators describing a *finite* symmetry. Such topological line operators form a structure which mathematicians call a *unitary fusion category*. We call it instead as a *symmetry category*, to emphasize its role as the finite symmetry of unitary two-dimensional quantum field theory. We start by stating the slogan, and then fill in the details:

Finite flavor symmetries of 2d theories are characterized by symmetry categories.

Objects

The objects in a symmetry category $\mathcal{C}$ correspond to topological line operators generating the symmetry. More precisely, a theory T admitting the symmetry $\mathcal{C}$ will admit topological line operators labeled by the objects of $\mathcal{C}$. Henceforth, we will drop the adjective topological in front of line operators.

For any line operator labeled by an object a, we have a partition function of T with the line inserted along an *oriented* path C, which we can denote as

$$\langle \cdots a(C) \cdots \rangle \tag{6.3.1}$$

where the dots stand for additional operators inserted away from C.

Morphisms

The morphisms in a symmetry category $\mathcal{C}$ correspond to topological local point operators which can be inserted between two lines. More precisely, consider two labels a and b, and a path C such that up to a point $p \in C$ we have the label a and from the point p we have the label b. Then, we can insert a possibly-line-changing topological operator labeled by m at p. We call such m a morphism from a to b, and denote this statement interchangeably as either

$$m : a \to b \qquad \text{or} \qquad m \in \mathrm{Hom}(a, b). \tag{6.3.2}$$

The set $\mathrm{Hom}(a, b)$ is taken to be a complex vector space and it labels (a subspace of) topological local operators between line operators corresponding to a and b in T. From now on, we would drop the adjective topological in front of local operators.

Existence of a trivial line

$\mathcal{C}$ contains an object 1, which labels the trivial line of T. We have

$$\langle \cdots 1(C) \cdots \rangle = \langle \cdots\cdots \rangle. \tag{6.3.3}$$

Additive structure

Given two objects a and b, there is a new object $a \oplus b$ in $\mathcal{C}$. In terms of lines in T, we have

$$\langle \cdots (a \oplus b)(C) \cdots \rangle = \langle \cdots a(C) \cdots \rangle + \langle \cdots b(C) \cdots \rangle. \tag{6.3.4}$$

We abbreviate $a \oplus a$ by $2a$, $a \oplus a \oplus a$ by $3a$, etc. The linear operator $U_{a \oplus b}$ acting on the Hilbert space obtained by wrapping $a \oplus b$ on a circle is then given by $U_a + U_b$.

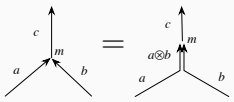


Figure 6.3.1: Two lines with the labels a and b can be fused to form a line with the label $a \otimes b$.

Tensor structure

Given two objects a and b, we have an object $a \otimes b$ in $\mathcal{C}$. This corresponds to considering two parallel-running line operators a and b as one line operator. The linear operator $U_{a\otimes b}$ acting on the Hilbert space obtained by wrapping $a \otimes b$ on a circle is then given by $U_a U_b$.

The trivial object 1 acts as an identity for this tensor operation. That is, there exist canonical isomorphisms $a \otimes 1 \simeq a$ and $1 \otimes a \simeq a$ for each object a. We can always find an equivalent category in which these isomorphisms are trivial, that is $a \otimes 1 = 1 \otimes a = a$. Hence, we can assume that these isomorphisms have been made trivial in $\mathcal{C}$. Henceforth, the unit object will also be referred to as the identity object.

Consider three lines $C_{1,2,3}$ meeting at a point p, with $C_{1,2}$ incoming and C_3 outgoing. We can put the label a, b, c on $C_{1,2,3}$, respectively. We demand that the operators we can put at the junction point p is given by $m \in \mathrm{Hom}(a \otimes b, c)$. This label $a \otimes b$ corresponds to a composite line as can be seen by the following topological deformation shown in Fig. 6.3.1.

The definition of $a \otimes b$ here includes a choice of the implicit junction operator where the lines labeled by a, b and $a \otimes b$ meet. In this chapter, whenever we draw figures with such implicit junction operators, we always choose the operator to be the one labeled by the identity morphism $\mathrm{id} : a \otimes b \to a \otimes b$.

Simplicity of the identity, semisimplicity, and finiteness

The simple objects $a \in \mathcal{C}$ are objects for which $\mathrm{Hom}(a, a)$ is one-dimensional. In general, for any object x, there is always a canonical *identity morphism* from x to x which labels the identity operator on the line labeled by x. For a simple object a, the existence of the identity morphism implies that there is a natural isomorphism $\mathrm{Hom}(a, a) \simeq \mathbb{C}$ as an algebra. We assume for simplicity that the identity object 1 is simple.

We also assume that every object x has a decomposition as a finite sum

$$x = \bigoplus_a N_a a \tag{6.3.5}$$

where N_a is a nonnegative integer and a is *simple*. In other words, every object x is *semisimple*.

Finally, we assume *finiteness*, that is the number of isomorphism classes of simple objects is finite. Below, we will be somewhat cavalier on the distinction between simple objects and isomorphism classes of simple objects.

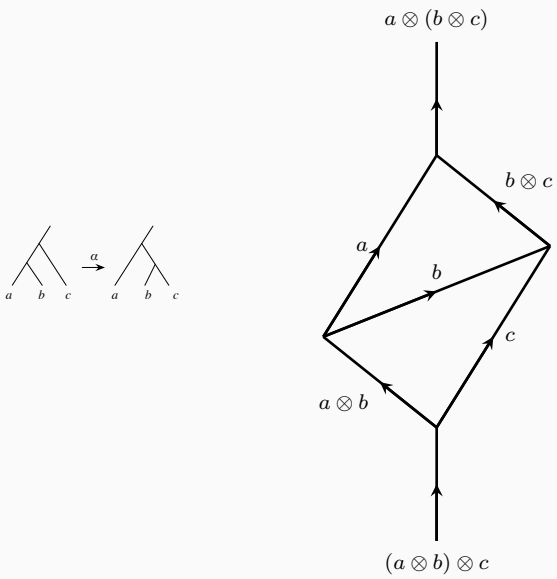


Figure 6.3.2: Left: The associator relates two different orders to tensor three lines. Right: Equivalently, the local operator labeled by the *associator* morphism is obtained by squeezing the region between $(a \otimes b) \otimes c$ and $a \otimes (b \otimes c)$ shown above to a point.

Associativity structure

The data in a symmetry category $\mathcal{C}$ includes certain isomorphisms implementing associativity of objects

$$\alpha_{a,b,c} \in \mathrm{Hom}((a \otimes b) \otimes c, a \otimes (b \otimes c)) \tag{6.3.6}$$

which we call *associators*.[4] Fusion matrices F for the Moore-Seiberg data, and the (quantum) 6j symbols for the (quantum) groups are used in the literature to capture the data of associators.

The associator $\alpha_{a,b,c}$ corresponds to a local operator which implements the process of exchanging line b from the vicinity of a to the vicinity of c. See Figure 6.3.2. They satisfy the pentagon identity which states the equality of following two morphisms

$$\begin{aligned}((a \otimes b) \otimes c) \otimes d \to (a \otimes (b \otimes c)) \otimes d \to a \otimes ((b \otimes c) \otimes d) \to a \otimes (b \otimes (c \otimes d)) \\ = ((a \otimes b) \otimes c) \otimes d \to (a \otimes b) \otimes (c \otimes d) \to a \otimes (b \otimes (c \otimes d))\end{aligned} \tag{6.3.7}$$

where each side of the equation stands for the composition of the corresponding associators. The pentagon identity ensures that exchanging two middle lines b, c between two outer lines a, d in two different ways is the same, see Fig. 6.3.3.

[4] Since we can and do choose the identity morphisms $1 \otimes a \to a$ and $a \otimes 1 \to a$ to be trivial, the associator $\alpha_{a,b,c}$ is also trivial when any of a, b, c is trivial.

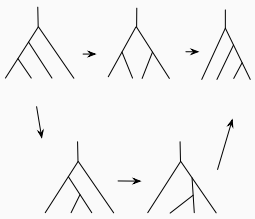


Figure 6.3.3: The pentagon identity guarantees that two distinct ways to rearrange the order of the tensoring of four lines lead to the same result.

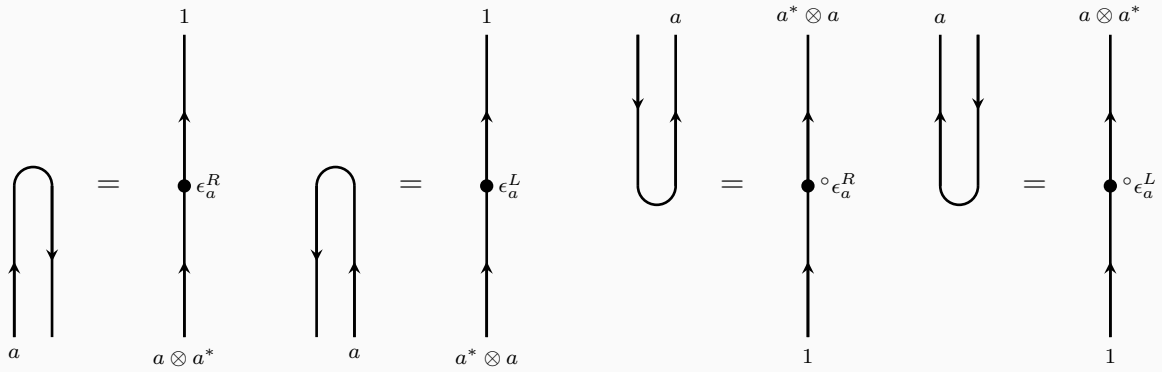


Figure 6.3.4: Folding a line and squeezing it gives rise to local operators labeled by evaluation and co-evaluation morphisms.

Dual structure

For every object a, $\mathcal{C}$ contains a dual object a^*. The line labeled by the dual object has the property that

$$\langle \cdots a(C) \cdots \rangle = \langle \cdots a^*(\tilde{C}) \cdots \rangle. \tag{6.3.8}$$

Here, $\tilde{C}$ denotes the same path C but with a reverse orientation, and the morphisms attached at the junctions on $\tilde{C}$ need to be changed appropriately as we explain below at the end of this subsubsection.

We require that the dual of the dual is naturally isomorphic to the original object: $(a^*)^* \simeq a$. The dual operation also changes the order of the tensoring:

$$(a \otimes b)^* = b^* \otimes a^*. \tag{6.3.9}$$

We demand that there are *evaluation* morphisms

$$\epsilon_a^R : a \otimes a^* \to 1, \quad \epsilon_a^L : a^* \otimes a \to 1 \tag{6.3.10}$$

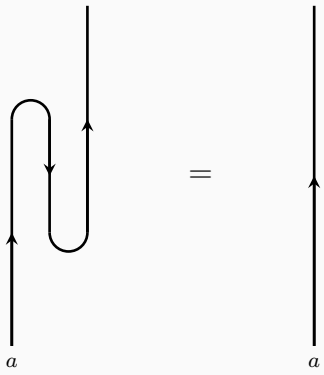


Figure 6.3.5: Consistency condition on evaluation and co-evaluation morphisms resulting from a topological deformation.

and *co-evaluation* morphisms[5]

$$ {}^{\circ}\epsilon^R_a : 1 \to a^* \otimes a, \quad {}^{\circ}\epsilon^L_a : 1 \to a \otimes a^*. \tag{6.3.11} $$

These label local operators corresponding to the process of folding a line operator a. See Figure 6.3.4.

We note that ϵ^R_a and $\epsilon^L_{a^*}$ are not necessarily equal. However, we require as part of definition of dual structure that they are related as follows

$$ \epsilon^R_a = \epsilon^L_{a^*} \circ (p_a \otimes 1) = \epsilon^L_{a^*} \circ (1 \otimes p^{-1}_{a^*}) \tag{6.3.12} $$

where p_a is an isomorphism from a to a and p_{a^*} is an isomorphism from a^* to a^*. Similarly, we require

$$ {}^{\circ}\epsilon^R_a = (1 \otimes p^{-1}_a) \circ {}^{\circ}\epsilon^L_{a^*} = (p_{a^*} \otimes 1) \circ {}^{\circ}\epsilon^L_{a^*} \tag{6.3.13} $$

The data of p_a and p_{a^*} is referred to in the literature as a *pivotal structure* on the fusion category $\mathcal{C}$.

The evaluation and co-evaluation morphisms have to satisfy the following consistency condition with the associator

$$ (\epsilon^R_a \otimes 1) \circ {\alpha_{a,a^*,a}}^{-1} \circ (1 \otimes {}^{\circ}\epsilon^R_a) = 1 \tag{6.3.14} $$

as morphisms from a to a. This ensures that a line with two opposite folds in the right direction can be unfolded as shown in Figure 6.3.5. A similar identity is satisfied by ϵ^L_a and ${}^{\circ}\epsilon^L_a$ which ensures that two opposite folds in the left direction can be unfolded.

Using evaluation and co-evaluation morphisms of different parity we can construct loops of

[5]We use the convention that when something is denoted by x, co-something is denoted by ${}^{\circ}x$. This usage is unconventional, in particular for the case of coproduct for which Δ is definitely the standard notation, but it reduces the amount of notations that one has to remember.

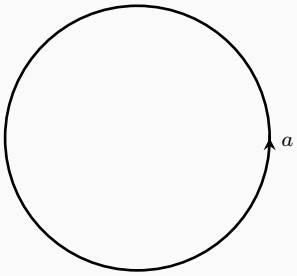


Figure 6.3.6: A loop of line a constructed by composing evaluation and co-evalutaion morphisms. The loop, if it contains no other operators in it, can then be shrunk and the partition function with the loop is equal to $\dim_{\mathrm{CC}} a = \dim a$ times the partition function without the loop, with all other insertions unchanged.

lines

$$(\dim_{\mathrm{CC}} a)\mathrm{id} : 1 \xrightarrow{\circ\epsilon_a^R} a^* \otimes a \xrightarrow{\epsilon_a^L} 1, \tag{6.3.15}$$

$$(\dim_{\mathrm{C}} a)\mathrm{id} : 1 \xrightarrow{\circ\epsilon_a^L} a \otimes a^* \xrightarrow{\epsilon_a^R} 1. \tag{6.3.16}$$

These are morphisms from 1 to 1 and hence they are proportional to the identity morphism. The proportionality factors define two numbers: the counter-clockwise dimension $\dim_{\mathrm{CC}} a$ of a and the clockwise dimension $\dim_{\mathrm{C}} a$ of a. See Figure 6.3.6.

Since we can replace the label a by a^* at the cost of flipping the orientation of line, we must have

$$\dim_{\mathrm{CC}} a = \dim_{\mathrm{C}} a^*, \tag{6.3.17}$$

$$\dim_{\mathrm{CC}} a^* = \dim_{\mathrm{C}} a. \tag{6.3.18}$$

Indeed, this follows from (6.3.12) and (6.3.13).

In fact, it turns out that we can further argue that

$$\dim_{\mathrm{CC}} a = \dim_{\mathrm{C}} a \equiv \dim a. \tag{6.3.19}$$

To see this, place a small counter-clockwise loop of line a around the "north pole" on the sphere. Let there be no other insertions anywhere on the sphere. This evaluates to $\dim_{\mathrm{CC}} a \times Z_{S^2}$ where Z_{S^2} is the partition function on sphere. Now we can move the line such that it looks like a small clockwise loop around the "south pole" on the sphere. This evaluates to $\dim_{\mathrm{C}} a \times Z_{S^2}$. Equating the two expressions we find (6.3.19) [6]. This is a further constraint on $\mathcal{C}$. If the fusion category $\mathcal{C}$ satisfies (6.3.19), then $\mathcal{C}$ is called a *spherical fusion category* in the literature.

Since ϵ_a^R and $\epsilon_{a^*}^L$ are not necessarily equal, we have to specify whether a folding of line a to the right should be read as the morphism ϵ_a^R or the morphism $\epsilon_{a^*}^L$. Similarly there is a specification

[6] The authors thank Shu-Heng Shao for discussion related to this point.

of ${}^{\circ}\epsilon^R_a$ vs. ${}^{\circ}\epsilon^L_{a^*}$. This issue can be dealt with in two ways, which are technically equivalent but have a rather different flavor.

One method: One perspective is to regard that a line is always labeled by the pair (the local orientation, an object in $\mathcal{C}$). Then, a pair $(\uparrow, a)$ and $(\downarrow, a^*)$ are isomorphic but not actually the same. We note that this distinction needs to be made even when $a \simeq a^*$. Then we make the rule that when a vertical line is labeled by $(\uparrow, a)$ up to some point and then labeled by $(\downarrow, a^*)$ from that point, we insert the pivotal structure $p_a \in \mathrm{Hom}(a, a)$ at that point. This approach would be preferred by those who have no trouble with adding local orientations as a new datum to a topological line operator.

Another method: Another perspective is to think that the change between ϵ^R_a and $\epsilon^L_{a^*}$ and between ${}^{\circ}\epsilon^R_a$ vs. ${}^{\circ}\epsilon^L_{a^*}$ is canceled by changing the nearby morphisms. This method might be preferred by those who do not want to add local orientation as a new datum to a topological line operator.

We emphasize that, in this approach, the operation of exchanging a by a^* with a reversed orientation does not change the local operators at the junctions. Instead it changes the way the local operators at the junctions are read as morphisms in the associated symmetry category $\mathcal{C}$.

The following moves are sufficient to specify what happens in any situation:

1. Consider a morphism $\epsilon^R_a \circ \alpha^{-1}_{a,a^*,c} \circ (1 \otimes m) : a \otimes b \to c$ where $m : b \to a^* \otimes c$. This is equal to $\epsilon^L_{a^*} \circ \alpha^{-1}_{a,a^*,c} \circ (1 \otimes n) : a \otimes b \to c$ where $n = (p^{-1}_{a^*} \otimes 1) \circ m : b \to a^* \otimes c$.

2. Consider a morphism $\epsilon^R_{a^*} \circ \alpha_{c,a^*,a} \circ (m \otimes 1) : b \otimes a \to c$ where $m : b \to c \otimes a^*$. This is equal to $\epsilon^R_a \circ \alpha_{c,a^*,a} \circ (n \otimes 1) : b \otimes a \to c$ where $n = (p_{a^*} \otimes 1) \circ m : b \to c \otimes a^*$.

3. Consider a morphism $(1 \otimes m) \circ \alpha_{a^*,a,b} \circ {}^{\circ}\epsilon^R_a : b \to a^* \otimes c$ where $m : a \otimes b \to c$. This is equal to $(1 \otimes n) \circ \alpha_{a^*,a,b} \circ {}^{\circ}\epsilon^L_{a^*} : b \to a^* \otimes c$ where $n = m \circ (p^{-1}_a \otimes 1) : a \otimes b \to c$.

4. Consider a morphism $(m \otimes 1) \circ \alpha^{-1}_{b,a,a^*} \circ {}^{\circ}\epsilon^R_{a^*} : b \to c \otimes a^*$ where $m : b \otimes a \to c$. This is equal to $(n \otimes 1) \circ \alpha^{-1}_{b,a,a^*} \circ {}^{\circ}\epsilon^L_a : b \to c \otimes a^*$ where $n = m \circ (1 \otimes p_a) : b \otimes a \to c$.

These moves follow from (6.3.12) and (6.3.13). We draw a picture of the first move in Figure 6.3.7. The other three moves are also described by similar pictures.

Unitary structure

The unitary structure requires an existence of a conjugate-linear involution sending $m \in \mathrm{Hom}(a, b)$ to $m^\dagger \in \mathrm{Hom}(b, a)$, generalizing the Hermitian conjugate in the standard linear algebra.

We require that the evaluation and the coevaluation morphisms are related by this conjugate operation:

$$ {}^{\circ}\epsilon^R_{a^*} = (\epsilon^L_a)^\dagger, \qquad {}^{\circ}\epsilon^L_{a^*} = (\epsilon^R_a)^\dagger. \tag{6.3.20}$$

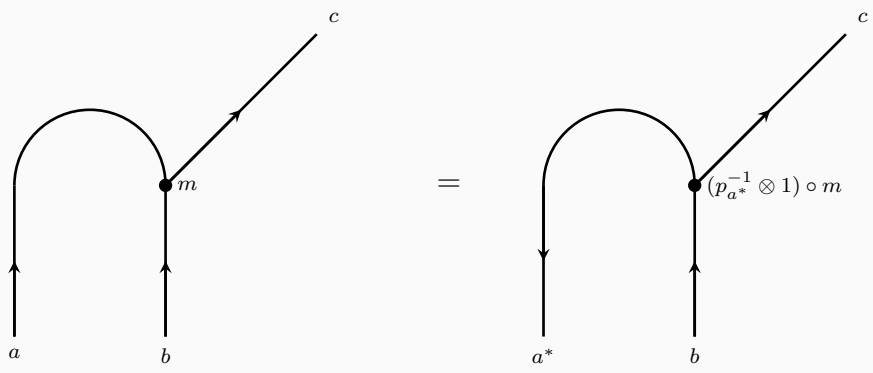


Figure 6.3.7: The fold in the diagram on left hand side is specified as ϵ_a^R and the fold on the right is specified as $\epsilon_{a^*}^L$. This changes the morphism from m on the left side to $(p_{a^*}^{-1} \otimes 1) \circ m$ on the right side. These two diagrams provide two different categorical representations of the same physical configuration.

We further require $m^\dagger \circ m \in \mathrm{Hom}(a, a)$ to be positive semi-definite in the following sense: Since we assumed the semisimplicity and the finiteness of the number of simple objects, $\mathrm{Hom}(a, a)$ can naturally be identified with a direct sum of a matrix algebra. Then we require $m^\dagger \circ m$ to have non-negative eigenvalues.

The above positivity condition requires $\dim a > 0$ for all a. As we will see in Sec. 6.3.3 below, the unitarity implies sphericity.

6.3.2 Comments

We have several comments:

- Note that category theorists do not like unitary structures, since it is specific to the base field $\mathbb{C}$ while they would like to keep everything usable for arbitrary base field. For this reason they often distinguish various concepts of $*$ operations and various structures satisfied by them, such as rigid structure, pivotal structure, spherical structure and pseudo-unitary structure.[7] If we consider unitary 2d quantum field theories (or more precisely its Wick-rotated versions which are reflection-positive), the unitary structure is the most natural one.

  Operator algebraic quantum field theorists in fact work in this setting, since for them the existence of the positive-definite inner product on the Hilbert space is paramount. Unfortunately their papers often phrase purely categorical results in the operator algebra theoretic

[7] The rigid structure posits the existence of the left dual a^* and the right dual ${}^*a$, satisfying various conditions. It can be shown that ${}^{**}a \simeq a^{**}$, and $a^{****} \simeq a$. The pivotal structure is a collection of isomorphisms $a^{**} \simeq a$. In our description, the pivotal structure relates ϵ_a^L and $\epsilon_{a^*}^R$. A pivotal structure is called spherical if $\dim a = \dim a^*$ for all a. A spherical structure is called pseudo-unitary if $|\dim a|$ is the largest eigenvalue of $(N_a)_b^c$ for all simple a. For the definition of N_{ab}^c, see Sec. 6.3.3.

language, which makes them somewhat harder for outsiders to digest. From this point of view their review article [167] is very helpful, where a concise translation between terminologies of two different schools is given.

- Every property given above, except the simplicity of identity, semisimplicity and finiteness, is a straightforward expression of how topological lines and the junction operators associated to symmetries should behave. We impose the simplicity of identity, semisimplicity and finiteness to make the situation tractable. When the semisimplicity is dropped, the category is called a finite tensor category; when the simplicity of identity is dropped, the category is called a finite multi-fusion category; when both are dropped, it is called a finite multi-tensor category. When finiteness is dropped, we simply drop the adjective "finite".

  Indeed, if we consider all topological lines in a given 2d theory and all topological operators on topological lines, they might not in general form a unitary fusion category. Rather, our point of view is that we take a subset of topological lines and subspaces of topological operators on the lines so that they form a unitary fusion category, and then it can be thought of as a symmetry of the 2d theory.

- A very similar categorical structure was introduced by Moore and Seiberg [172] in the analysis of 2d RCFTs and 3d TFTs. In the category theory they are now called unitary modular tensor categories. In fact, the unitary modular tensor categories *are* also unitary fusion categories, where the latter description is obtained by forgetting the braiding.

6.3.3 More notions of symmetry categories

Before discussing examples, it is useful to set up a few more notions:

'Homomorphisms' between symmetry categories

In the case of two groups G_1 and G_2, we have the concept of homomorphisms $\varphi : G_1 \to G_2$, preserving the group multiplication. Similarly, we can talk about *symmetry functors* $\varphi : \mathcal{C}_1 \to \mathcal{C}_2$ between two symmetry categories, together with the data specifying how the structures listed above are mapped. Among them are isomorphisms

$$\epsilon_{a,b} \in \mathrm{Hom}(\varphi(a) \otimes \varphi(b), \varphi(a \otimes b)) \tag{6.3.21}$$

which tell us how the tensor structure of $\mathcal{C}_1$ is mapped into the tensor structure of $\mathcal{C}_2$. For example, the morphisms $\epsilon_{a,b}$ map the associator of $\mathcal{C}_1$ to the associator of $\mathcal{C}_2$.

Two symmetry functors $\varphi, \varphi' : \mathcal{C}_1 \to \mathcal{C}_2$ are considered equivalent when there is a set of isomorphisms

$$\eta_a \in \mathrm{Hom}(\varphi(a), \varphi'(a)) \tag{6.3.22}$$

such that

$$\eta_{a\otimes b}\epsilon_{a,b} = \epsilon'_{a,b}(\eta_a \otimes \eta_b). \tag{6.3.23}$$

When a symmetry functor has an inverse, it is called an equivalence between symmetry categories.

Products of symmetry categories

In the case of two groups G_1 and G_2, their product $G_1 \times G_2$ is also a group. Similarly, given two symmetry categories $\mathcal{C}_1$ and $\mathcal{C}_2$, we denote their product as $\mathcal{C}_1 \boxtimes \mathcal{C}_2$, whose simple objects are given by $a_1 \boxtimes a_2$ where $a_{1,2}$ are simple objects of $\mathcal{C}_{1,2}$, respectively. This product is called Deligne's tensor product of categories.

Fusion rule of unitary fusion categories

A symmetry category comes with a lot of structures. Sometimes it is useful to forget about most of them as follows. For each isomorphism class of simple objects a, introduce a symbol $[a]$, and define their multiplication by $[a][b] := \sum_c N_{ab}^c [c]$ when $a \otimes b = \bigoplus_c N_{ab}^c c$. This makes non-negative integral linear combinations of $[a]$'s into an algebra over $\mathbb{Z}_+$ with a specific given basis. We call this algebra $R(\mathcal{C})$ the fusion ring of the symmetry category $\mathcal{C}$. In the case of modular tensor categories, this algebra is also called the Verlinde algebra. We would often call this algebra as just the fusion rule of $\mathcal{C}$.

Let n be the number of isomorphism classes of simple objects. Then we can regard $(N_a)_b^c$ as $n \times n$ matrices and $[a] \mapsto N_a$ is the adjoint representation of the fusion ring.

Determination of dimensions of objects

The dimensions are fixed by N_{ab}^c. To see this, consider the n-dimensional vector $v := (\dim a)_a$ where a runs over the isomorphism classes of simple objects. Its entries are positive real numbers thanks to the unitarity. Furthermore, v is the simultaneous eigenvector of all N_a's with eigenvalues $\dim a$. Then by the Perron-Frobenius theorem, $\dim a$ is the largest eigenvalue of the matrix N_a, which is guaranteed to be positive. The argument above applies both to $\dim_{\mathrm{C}}$ and $\dim_{\mathrm{CC}}$, and therefore the sphericity is implied by the unitarity. Unitarity also guarantees $\dim a = \dim a^*$.

We define the total dimension of the symmetry category $\mathcal{C}$ by the following formula:

$$\dim \mathcal{C} = \sum_a (\dim a)^2. \tag{6.3.24}$$

Here the sum runs over the isomorphism classes of simple objects.

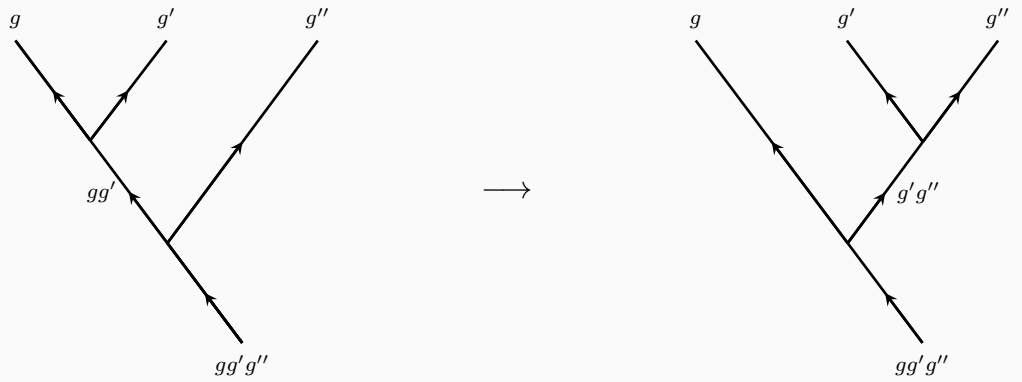


Figure 6.3.8: The same background connection represented in two different ways as networks of topological line operators. If the symmetry is anomalous, they lead to different partition functions.

6.3.4 Groups and representations of groups as symmetry categories

Symmetry categories $\mathcal{C}(G,\alpha)$

As an example, let us recast an ordinary group G as a symmetry category. We first regard each element $g \in G$ as a simple object denoted by the same letter in the category. We define

$$g \otimes g' := gg', \qquad g^* := g^{-1}. \tag{6.3.25}$$

Taking α_{g_1,g_2,g_3} to be the identity maps, they clearly form a unitary fusion category, which we denote by $\mathcal{C}(G)$.

More generally, the pentagon identity among α_{g_1,g_2,g_3} says that α is a 3-cocycle on G valued in $\mathrm{U}(1)$.[8] Denote the resulting fusion category by $\mathcal{C}(G,\alpha)$. In the literature it is often denoted Vec_G^α. We clearly have $\dim g = 1$. Thus the total dimension of this symmetry category is the order of the group.

When α_1 and α_2 differ by a coboundary of a 2-cochain ϵ, we can construct an equivalence of categories between $\mathcal{C}(G,\alpha_1)$ and $\mathcal{C}(G,\alpha_2)$ using the functor specified using the same ϵ in (6.3.21). This means that in the definition of $\mathcal{C}(G,\alpha)$, one can regard $\alpha \in H^3(G,\mathrm{U}(1))$.

It is also clear that any unitary fusion category whose simple objects are all invertible can be made to be of this form. Summarizing,

A symmetry category $\mathcal{C}$ whose simple lines are all invertible is equivalent to $\mathcal{C}(G,\alpha)$ where G is a finite group and α is an element in $H^3(G,\mathrm{U}(1))$.

[8] As already noted in footnote 4, in our convention α_{g_1,g_2,g_3} is trivial whenever any of $g_{1,2,3}$ is the identity. Such a cocycle is called normalized. It is a well-known fact in group theory that group cohomology can be computed by restricting every cochains involved to be normalized.

$\mathcal{C}(G,\alpha)$ and the anomaly

As is by now familiar, this cohomology class $\alpha \in H^3(G, \mathrm{U}(1))$ specifies the anomaly of G flavor symmetry in two dimensions [173, 155]. One way to see it is as follows [150]: Insert a network of lines with trivalent junctions between them on the spacetime manifold Σ. Let the lines be labeled by simple objects of $\mathcal{C}(G,\alpha)$, that is by group elements. And let every junction of the form $g \otimes g' \to gg'$ be labeled by the identity morphism $gg' \to gg'$. Such a configuration can also be thought of as reproducing the effects of a background connection on Σ which has holonomies given by g on crossing transversely a line labeled by g. Now, consider a local region looking like the left hand side of Figure 6.3.8. Move the lines such that now it looks like the right hand side of Figure 6.3.8. This changes the partition function by $\alpha(g, g', g'')$. The new background connection is just a gauge transform of the original background connection. Hence, we see that α precisely captures the anomaly in the flavor symmetry. Morally, this means the following:

A symmetry category $\mathcal{C}$ includes the specification of its anomaly.

Fixing a group G, the set of its anomalies forms an Abelian group. Notice that, in our language, $\mathcal{C}(G,\alpha)$ for different α have the same fusion ring $R = \mathbb{Z}_+ G$. Thus, we can ask the following more general question: does the set of symmetry categories $\mathcal{C}$ with the same fusion ring R form an Abelian group? The answer is that we need a coproduct on R. To see this, let us recall why the anomaly of a flavor symmetry forms an Abelian group from the perspective of quantum field theory.

In general, given two theories $T_{1,2}$, we can consider the product theory $T_1 \times T_2$ which is just two decoupled theories considered as one. When T_i has flavor symmetry group G_i, the product $T_1 \times T_2$ has flavor symmetry group $G_1 \times G_2$. When $G_1 = G_2 = G$, we can take the diagonal G subgroup of $G \times G$ and regard $T_1 \times T_2$ to have flavor symmetry G. Now, when T_i has the anomaly α_i, we define the anomaly of $T_1 \times T_2$ to be the sum $\alpha_1 + \alpha_2$. This abstractly defines the addition operation on the anomaly.

The crucial step that does not directly generalize to symmetry categories is the existence of the diagonal subgroup $G \subset G \times G$. In order to define the addition operation on the set of fusion categories sharing the same fusion rule R, similarly we need a coproduct $R \to R \otimes R$.

$\mathcal{C}(G,\alpha)$ and the G-SPT phases

Next, fixing a 3-cocycle α, let us ask what is the autoequivalence of $\mathcal{C}(G,\alpha)$, that is, the self equivalence that preserves the structure as a symmetry category.

Pick an autoequivalence $\varphi : g \to \varphi(g)$ with the associated $\epsilon_{g,h} \in \mathrm{Hom}(\varphi(g)\varphi(h), \varphi(gh))$. Clearly φ is an automorphism of G. Fixing φ to be the identity, ϵ needs to be a 2-cocycle so that it does not change α. Furthermore, two such ϵ's are considered equivalent when they differ by a 2-coboundary, due to (6.3.23). Therefore ϵ can be thought of as taking values in $H^2(G, \mathrm{U}(1))$. Summarizing,

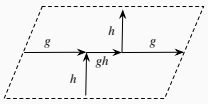


Figure 6.3.9: The lines representing a commuting holonomy (g, h) on a torus.

Autoequivalence of $\mathcal{C}(G, 0)$ is the semidirect product $\mathrm{Aut}(G) \ltimes H^2(G, \mathrm{U}(1))$.

The $\mathrm{Aut}(G)$ part is clear: it just amounts to renaming the topological lines associated to the group operation. How should we think of $H^2(G, \mathrm{U}(1))$? It is telling us that instead of choosing the identity operator as the implicit junction operator for $g \otimes g' \to gg'$ as done in Fig. 6.3.1, we can choose $\epsilon_{g,g'}$ times the identity. This will not change the associator but will change the partition function associated to a background connection on Σ. This corresponds to coupling a two-dimensional theory with $\mathcal{C}(G, \alpha)$ symmetry with a two-dimensional bosonic symmetry protected topological (SPT) phase, which is specified by the 2-cocycle ϵ protected by the flavor symmetry G. The 2-cocycle ϵ is also known as a discrete torsion of G.

As an example, consider a torus with holonomies g, h around two 1-cycles. They can be represented using the topological lines as in Fig. 6.3.9. There, we resolved the intersection to two trivalent junctions.

We now change the operators at the two junctions to $\epsilon_{g,h}$ and $\epsilon_{g^{-1},h^{-1}}$ given by the values of the 2-cocycle. In total the phase of the partition function changes by

$$c_{g,h} = \epsilon_{g,h}/\epsilon_{h,g} \tag{6.3.26}$$

which is the standard relation between the discrete torsion phase c on the torus and the 2-cocycle ϵ [174]. We can thus generalize as follows:

Autoequivalences of a symmetry category $\mathcal{C}$ generalize the notion of renaming and multiplying by SPT-phases for a group symmetry.

We need to keep in mind however that the phases introduced by ϵ in the general case do not have an interpretation of multiplying a SPT phase protected by $\mathcal{C}$, since the product of two theories with symmetry $\mathcal{C}$ has symmetry $\mathcal{C} \boxtimes \mathcal{C}$ but is not guaranteed to have symmetry $\mathcal{C}$, as already discussed above.

$\mathrm{Rep}(G)$ as symmetry category

Next, let us discuss $\mathrm{Rep}(G)$ for a finite group G. Its structure as a symmetry category is straightforward: the objects are representations of G, the morphism space $\mathrm{Hom}(R, S)$ between two representations consists of intertwiners, as the tensor product $\otimes$ and the associator $\alpha_{R,S,T}$ we use the ordinary operations, the dual is the complex conjugate, etc. The simple objects are the

irreducible representations. The dimension of a representation R as defined above equals the ordinary definition. Then, the total dimension of $\mathrm{Rep}(G)$ is the sum of the square of the dimensions of the irreducible representations, therefore

$$\dim \mathrm{Rep}(G) = |G|. \tag{6.3.27}$$

Clearly, when G is an Abelian group, we have a natural equivalence $\mathcal{C}(\widehat{G}) = \mathrm{Rep}(G)$. This means in particular that the symmetry categories $\mathrm{Rep}(G)$ for distinct Abelian G are distinct. In fact the fusion rings $R(\mathrm{Rep}(G))$ are already distinct.

For non-Abelian G, the situation is more delicate. For example, take the dihedral group D_8 with eight elements (i.e. the symmetry of a square in three-dimensional space) and the quaternion group Q_8 (i.e. the group formed by eight quaternions $\pm 1, \pm i, \pm j, \pm k$). Their character tables are the same. In particular, they both have four one-dimensional irreducible representations 1, a, b and ab and a unique two-dimensional irreducible representation m, with $m \otimes m \simeq 1 \oplus a \oplus b \oplus ab$. Therefore, their fusion rings are the same. In this case, however, the symmetry categories are different, since they are known to have distinct associator $\alpha_{m,m,m}$, see e.g. [175]. In this sense, the relation between $\mathrm{Rep}(D_8)$ and $\mathrm{Rep}(Q_8)$ is analogous to the relation between the symmetry categories $\mathcal{C}(G, \alpha)$ associated to the same group but with a different anomaly.

It is also known that there are cases where two distinct groups lead to the same symmetry category [176, 177]. Let $\mathbb{F}_2$ be the finite field with two elements. The symplectic group $\mathrm{Sp}(2d, \mathbb{F}_2)$ acts on $(\mathbb{F}_2)^{2d}$. The extension of $(\mathbb{F}_2)^{2d}$ by $\mathrm{Sp}(2d, \mathbb{F}_2)$ with this natural action is classified by the cohomology $H^2(\mathrm{Sp}(2d, \mathbb{F}_2), (\mathbb{F}_2)^{2d})$, see Appendix A, in particular (D.1.11). It is known that when $d \geq 3$, there is a suitable nonzero $\nu \in H^2(\mathrm{Sp}(2d, \mathbb{F}_2), (\mathbb{F}_2)^{2d})$ such that $G_d = \mathrm{Sp}(2d, \mathbb{F}_2) \ltimes (\mathbb{F}_2)^{2d}$. and $G'_d = \mathrm{Sp}(2d, \mathbb{F}_2) \ltimes_\nu (\mathbb{F}_2)^{2d}$ we have $\mathrm{Rep}(G_d) = \mathrm{Rep}(G'_d)$ as a symmetry category. The smallest example is when $d = 3$, where $|G_3| = |G'_3| = (2^6 - 1)(2^4 - 1)(2^2 - 1)2^{3^2} \cdot 2^{3\cdot 2} = 92897280$.

As already roughly discussed in Sec. 6.2.2, a theory T/G where T has flavor symmetry G has a symmetry $\mathrm{Rep}(G)$, and re-gauging $\mathrm{Rep}(G)$ we get back the original theory T. Equivalently, if a theory T' has a symmetry $\mathrm{Rep}(G)$, gauging $\mathrm{Rep}(G)$ of T' has the symmetry G. In the example above, we said $\mathrm{Rep}(G_d) = \mathrm{Rep}(G'_d)$ for different G_d and G'_d. This means that there are (at least) two distinct ways of gauging $\mathcal{C} = \mathrm{Rep}(G_d) = \mathrm{Rep}(G'_d)$, producing two different theories with different symmetries G_d and G'_d. Let us next study what is going on in detail.

6.4 Gaugings and symmetry categories

Given a theory T with non-anomalous G symmetry, we can form the gauged theory $T' = T/G$, and T' has $\mathrm{Rep}(G)$ as the symmetry category. We would like to precisely state the process of gauging $\mathrm{Rep}(G)$ of T' and getting back T.

To do this, it is helpful first to state two procedures, one of obtaining $\mathrm{Rep}(G)$ from G and another of obtaining G from $\mathrm{Rep}(G)$, in a manner that treats G and $\mathrm{Rep}(G)$ on an equal footing. This is done using the concept of module categories and bimodule categories.

Once we understand this process, it is straightforward to study how to gauge a more general symmetry category $\mathcal{C}$. We will also be able to answer how the symmetry category $\mathcal{C}$ changes under gauging in the general setting.

6.4.1 Module and bimodule categories

Readers would be familiar with the action of a group or an algebra on a vector space. Then the vector space is called a representation or equivalently a module. When there are two commuting actions, one from the left and another from the right, it is often called a bimodule. Module categories and bimodule categories are categorified versions of this construction. Our expositions will be brief; for more details, please consult Chapter 7 of the textbook [168].

The categorified version of a linear space is an additive category. An additive category $\mathcal{M}$ is a category with the additive structure whose morphism spaces are vector spaces. Here and in the following, we assume objects of $\mathcal{M}$ to be of the form $\bigoplus_m N_m m$ where m runs over simple objects and N_m are non-negative integers. We also assume that the number of isomorphism classes of simple objects is finite.

The simplest additive category is Vec, the category of finite dimensional vector spaces. There is only one simple object $\mathbb{C}$, and every other object is isomorphic to $n\mathbb{C} = \mathbb{C}^n$.

Take $\mathcal{C}$ to be a symmetry category. A left-module category over $\mathcal{C}$ is an additive category $\mathcal{M}$ such that for $a \in \mathcal{C}$ and $m \in \mathcal{M}$ the product $a \otimes m \in \mathcal{M}$ is defined, together with the associativity morphisms

$$\alpha_{a,b,m} \in \mathrm{Hom}((a \otimes b) \otimes m, a \otimes (b \otimes m)) \tag{6.4.1}$$

satisfying the pentagon identity. Direct sums of module categories can be easily defined. When $\mathcal{C}$ is a symmetry category it is known that every module category can be decomposed into a direct sum of indecomposable module categories.

Physically, a module category specifies a topological boundary condition. Each object m in the module category specifies the boundary condition and the "type of flux" it carries. A line a can end on m at the cost of transforming m to $a \otimes m$ which changes the flux at the boundary by a. If $a \otimes m \simeq m$ for all $a \in \mathcal{C}$ then m, along with its direct sums with itself, describes an indecomposable module category by itself which corresponds to a boundary condition that absorbs all the flux and hence we can refer to it as the "Dirichlet boundary condition". In general, an indecomposable module category is a generalization of mixed Dirichlet-Neumann boundary conditions familiar from gauge theory. In the literature, the structure on Vec of a module category for $\mathcal{C}$ is often called a fiber functor of $\mathcal{C}$, but we try not to use this terminology.

We can similarly define a right-module category over $\mathcal{C}$. A $(\mathcal{C}_1, \mathcal{C}_2)$ bimodule category is an additive category $\mathcal{M}$ which has a left action of $\mathcal{C}_1$ and a right action of $\mathcal{C}_2$, together with further compatibility morphisms

$$\alpha_{c_1,m,c_2} \in \mathrm{Hom}((c_1 \otimes m) \otimes c_2, c_1 \otimes (m \otimes c_2)). \tag{6.4.2}$$

A bimodule category has the physical interpretation that it describes a topological interface where lines in the symmetry categories $\mathcal{C}_{1,2}$ can end from the left and from the right, respectively.

Recall that a representation of an algebra A on a vector space V can be thought of as a homomorphism from A to the algebra of matrices $\mathrm{Hom}(V,V)$. Similarly, given an additive category $\mathcal{M}$ we can form the symmetry category $\mathrm{Func}(\mathcal{M},\mathcal{M})$ of additive self functors, and then giving the structure of a module category on $\mathcal{M}$ is the same as giving a homomorphism of symmetry categories from $\mathcal{C}$ to $\mathrm{Func}(\mathcal{M},\mathcal{M})$.

The commutant of $\mathcal{C}$ within $\mathrm{Func}(\mathcal{M},\mathcal{M})$ is denoted by $\mathcal{C}^*_{\mathcal{M}}$ and called the dual of $\mathcal{C}$ with respect to $\mathcal{M}$. When $\mathcal{M}$ is indecomposable, it is known that $(\mathcal{C}^*_{\mathcal{M}})^*_{\mathcal{M}} \simeq \mathcal{C}$. $\mathcal{M}$ is naturally a $(\mathcal{C},\mathcal{C}')$ bimodule category where $\mathcal{C}'$ is $\mathcal{C}^*_{\mathcal{M}}$ with the order of the tensor product reversed. Since we will only use $\mathcal{C}'$ and not $\mathcal{C}^*_{\mathcal{M}}$ in the following, we will call $\mathcal{C}'$ the dual of $\mathcal{C}$ with respect to $\mathcal{M}$. Two symmetry categories $\mathcal{C}$ and $\mathcal{C}'$ such that $\mathcal{C}'$ is the dual of $\mathcal{C}$ acting on $\mathcal{M}$ are called *categorically Morita equivalent*. We will come back to this notion in Sec. 6.4.7.

6.4.2 Duality of $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$

Let us discuss examples of module categories. We can give Vec a structure of a module category over $\mathcal{C}(G)$: for $g \in \mathcal{C}(G)$ and $\mathbb{C} \in \mathrm{Vec}$, we just define $g \otimes \mathbb{C} = \mathbb{C}$, and let the associator be trivial. Note that this construction does not work for $\mathcal{C}(G,\alpha)$ with nonzero anomaly $\alpha \in H^3(G,\mathrm{U}(1))$, since the pentagon will not be satisfied.

We can also give Vec a structure of a module category over $\mathrm{Rep}(G)$: for $R \in \mathrm{Rep}(G)$ and $\mathbb{C} \in \mathrm{Vec}$, we just define $R \otimes \mathbb{C} = R \in \mathrm{Vec}$, and let the associator be the natural one induced from the tensor product structure of G. We note that for $\mathcal{C} = \mathrm{Rep}(G_d) = \mathrm{Rep}(G'_d)$ at the end of Sec. 6.3.4, the module category structures on Vec as $\mathrm{Rep}(G_d)$ and $\mathrm{Rep}(G'_d)$ are different, since the associators (6.4.1) as morphisms in Vec turn out to be distinct.

Furthermore, we note that Vec is a $(\mathcal{C}(G),\mathrm{Rep}(G))$ bimodule category. The only additional ingredient is the associator

$$\alpha_{g,\mathbb{C},R} \in \mathrm{Hom}((g \otimes \mathbb{C}) \otimes R, g \otimes (\mathbb{C} \otimes R)) = \mathrm{Hom}(R,R). \tag{6.4.3}$$

The compatibility condition just says that $\alpha_{g,\mathbb{C},R}$ are representation matrices of g on R, chosen compatible with the tensor product. Physically, we can regard Vec as the boundary between a system with G symmetry on the left and a gauged system with $\mathrm{Rep}(G)$ symmetry on the right.

In fact, $\mathrm{Rep}(G)$ is characterized as the maximal set of right actions compatible with the left action of $\mathcal{C}(G)$, and $\mathcal{C}(G)$ as the maximal set of left actions compatible with the right action of $\mathrm{Rep}(G)$. As we will explicitly confirm in Sec. 6.4.5, $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$ are dual to each other with respect to Vec:

The symmetry category $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$ are dual with respect to the module category Vec.

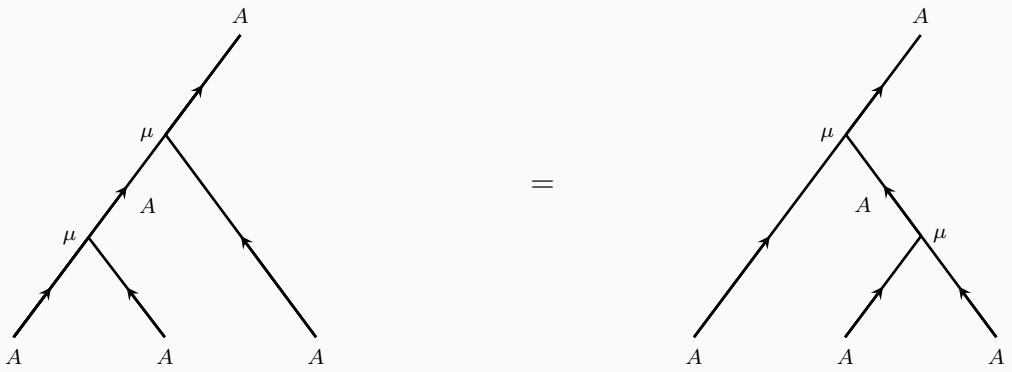


Figure 6.4.1: Associativity axiom for an algebra object $A \in \mathcal{C}$

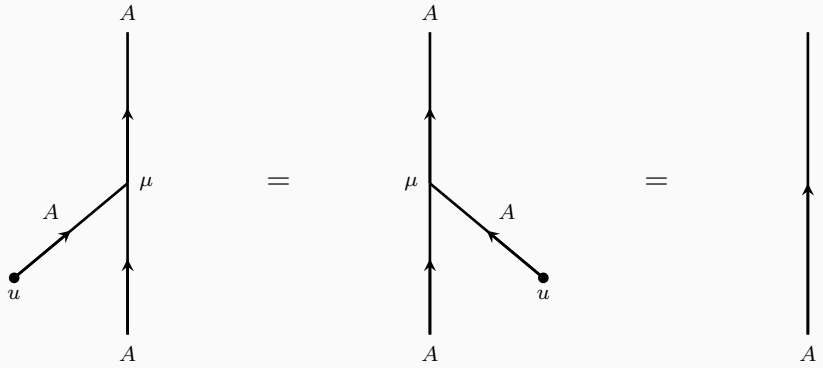


Figure 6.4.2: Unit axiom for an algebra object $A \in \mathcal{C}$.

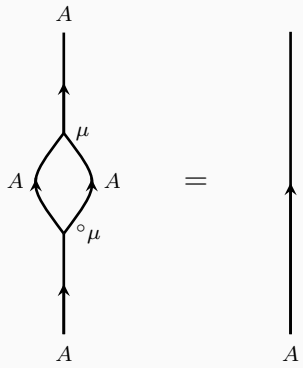


Figure 6.4.3: Separability axiom for an algebra object $A \in \mathcal{C}$.

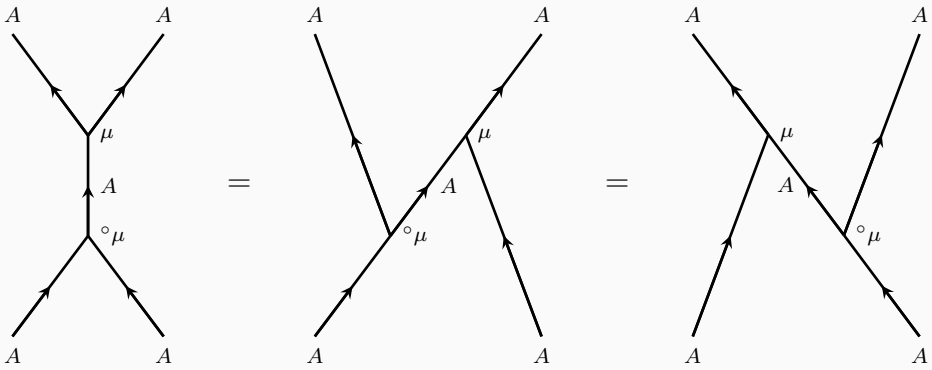


Figure 6.4.4: Frobenius axiom for an algebra object $A \in \mathcal{C}$.

6.4.3 Gauging by an algebra object

How is this duality of $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$ related to gauging? To understand this, we need a digression. Take $A = \bigoplus_g g$ in $\mathcal{C}(G)$. The group multiplication gives a multiplication morphism $\mu : A \otimes A \to A$, and and the map $1 \mapsto \bigoplus_g \delta_{1g} g$ defines a unit morphism $u : 1 \to A$. These two operations satisfy the properties in Figures 6.4.1 and 6.4.2. In general, if we have an object A in a general symmetry category $\mathcal{C}$ along with morphisms μ and u which satisfy these properties, then A is called an *algebra* in $\mathcal{C}$.

We also have corresponding co-morphisms. The map $g \mapsto \frac{1}{|G|} \bigoplus_h (h \otimes h^{-1}g)$ defines a co-multiplication $^\circ\mu : A \to A \otimes A$ and the projection $g \mapsto \delta_{1g} 1$ defines a co-unit morphism $^\circ u : A \to 1$. $^\circ\mu$ satisfies a co-associativity axiom which is given by turning the graphs in Figure 6.4.1 upside down and changing the directions of all the arrows. Similarly, $^\circ\mu$ and $^\circ u$ together satisfy a co-unit axiom which is given by turning the graphs in Figure 6.4.2 upside down and changing the directions of all the arrows. In general, if we have an object $A \in \mathcal{C}$ with morphisms $^\circ\mu$ and $^\circ u$ satisfying these axioms, we say that A is a *co-algebra* in $\mathcal{C}$.

Moreover, in the group case we can explicitly check that A satisfies additional properties

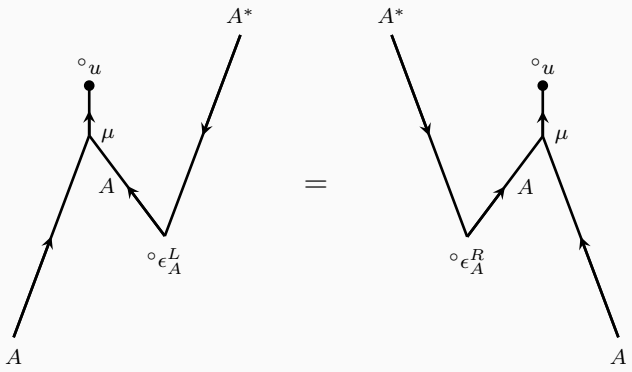


Figure 6.4.5: Symmetricity axiom for an algebra object $A \in \mathcal{C}$. Here $^\circ\epsilon_a$ denotes the co-evaluation map for line a.

which relate the algebra and co-algebra structure on A. These are shown in Figures 6.4.3, 6.4.4 and 6.4.5. An object $A \in \mathcal{C}$ satisfying these properties is called as a *symmetric Frobenius algebra object* in $\mathcal{C}$.

Now, given a two-dimensional theory T with flavor symmetry G, the partition function of the gauged theory T/G can be described using this object A. Namely, $Z_{T/G}(M)$ on a manifold M is defined by the partition function of T on M with a *fine-enough* trivalent mesh of topological lines all labeled by A. By fine-enough, we mean a mesh which can be obtained as the dual graph of a triangulation of the 2d manifold M. The relations above guarantee that the result does not depend on the choice of the mesh as long as it is fine enough. Dually, these relations guarantee invariance under changes of triangulations by Pachner moves. Furthermore, by decomposing $A = \bigoplus_g g$, one can see that this trivalent mesh is a superposition of various G bundles over M, giving us back the standard definition of the gauged theory.

The crucial idea of [178, 169] is to take this as the definition of gauging in the generalized sense. Namely, given a theory T with a symmetry category $\mathcal{C}$, pick a symmetric Frobenius algebra object $A \in \mathcal{C}$. Then, the gauged theory T/A is defined exactly as in the previous paragraph.

Note that there can be multiple possible choices of A for a given $\mathcal{C}$. For example, when $\mathcal{C} = \mathcal{C}(G)$, we can pick any subgroup $H \subset G$ and take $A_H := \bigoplus_{h\in H} h$. When $\mathcal{C} = \mathcal{C}(G, \alpha)$, we can check that A_H is an algebra only when the anomaly $\alpha \in H^3(G, \mathrm{U}(1))$ when restricted to H is trivial. We can also twist the multiplication morphism $m : A \otimes A \to A$ by using

$$\epsilon_{g,h} \in \mathrm{Hom}(g \otimes h, gh) \tag{6.4.4}$$

where $\epsilon_{g,h} \in H^2(G, \mathrm{U}(1))$. The choice of A then can be thought of as the choice of a gauge-able subsymmetry together with the choice of the discrete torsion [164]. We summarize:

Gauging a gauge-able subpart of $\mathcal{C}$ is done by inserting a fine-enough mesh of an algebra object A in $\mathcal{C}$. The choice of A in $\mathcal{C}$ generalizes the notion of choosing a non-anomalous subgroup and the discrete torsion in the case of group symmetry.

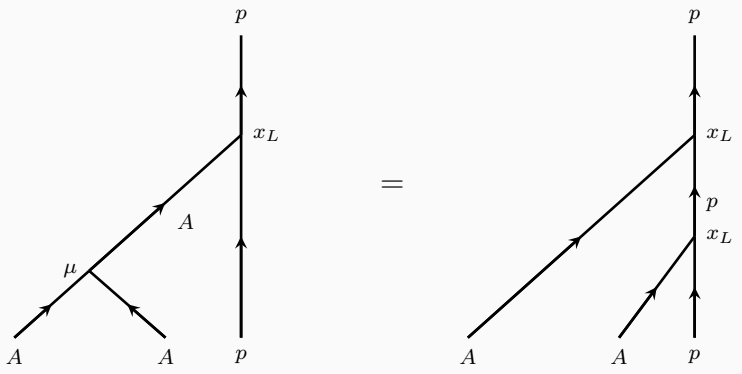


Figure 6.4.6: Associativity axiom for a left A-module (p, x_L).

We have some comments:

- There is in general no canonical maximal A in $\mathcal{C}$. For example, given $\mathcal{C}(G, \alpha)$ for a non-trivial anomaly, there is not necessarily a unique maximal non-anomalous subgroup.
- In the case of $\mathcal{C}(G, \alpha)$, finding possible algebras A is equivalent to finding a symmetry subcategory of the form $\mathcal{C}(H, 0)$. This is however not the general situation. Algebras A specifying possible gaugings do not necessarily correspond to symmetry subcategories.
- Note also that two distinct A and A' might give rise to the same gauged theory. We will come back to this question in Sec. 6.4.6.

6.4.4 Symmetries of the gauged theory from bimodules for the algebra object

What is the symmetry category of the gauged theory T/A? Since T/A is just the original theory T with a fine mesh of A, we can still consider topological lines $p \in \mathcal{C}$ of the original theory. But for p to be topological in the presence of an arbitrary mesh of A, the lines of A need to be able to end on p both on the left and the right consistently.

First of all, there must be a morphism $x_L : A \otimes p \to p$ such that conditions in Figures 6.4.6 and 6.4.7 are satisfied. These properties make p into a left A-module in $\mathcal{C}$. Similarly, there must be a morphism $x_R : p \otimes A \to p$ which satisfies similar conditions and makes p into a right A-module as well. Moreover, there must be morphisms ${}^{\circ}x_L : p \to A \otimes p$ and ${}^{\circ}x_R : p \to p \otimes A$ which satisfy co-conditions obtained by reflecting the graphs in Figures 6.4.6 and 6.4.7 upside down and reversing all the arrows. This would make p into a left and right A-comodule. It turns out that for a symmetry category, we can always get the co-module structure by combining the module structure with the co-evaluation map for A or A^*. Hence, from now on we restrict our attention only to the module structure. On top of all these conditions, p must also satisfy conditions which allow us to commute the left and right actions of A on p. This is the condition shown in Figure 6.4.8 and it makes p into an (A, A)-bimodule in $\mathcal{C}$.

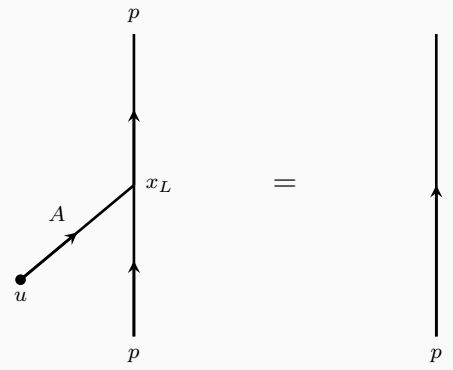


Figure 6.4.7: Unit axiom for a left A-module (p, x_L).

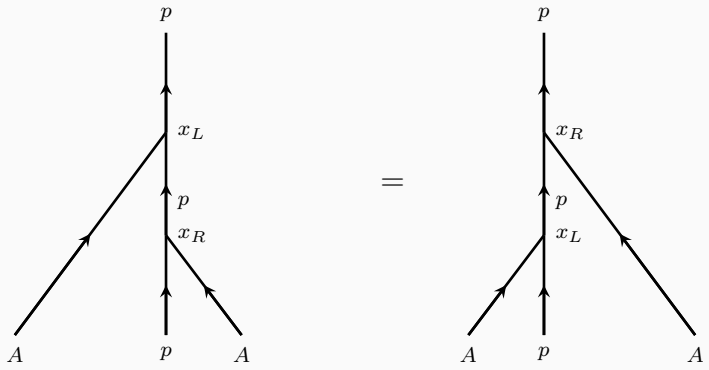


Figure 6.4.8: Definition of (A, A)-bimodule (p, x_L, x_R).

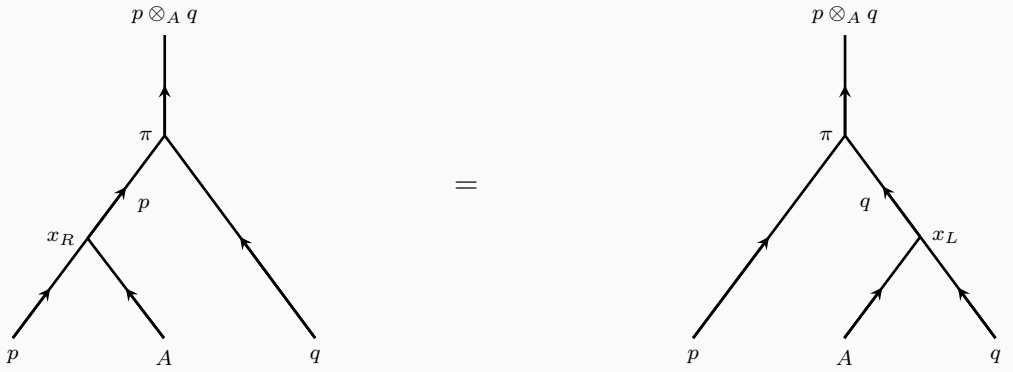


Figure 6.4.9: The compatibility condition defining balanced tensor product of two A-bimodules p and q via π.

The topological operators between two lines p and q, both of which are bimodules, need to be compatible with the action of A from both sides. That is, the insertion of the bimodule changing operator and the action of A must commute. They give rise to the category $\mathrm{Bimod}_{\mathcal{C}}(A)$ of (A, A) bimodules of A within $\mathcal{C}$. Note that the concept of the category of (A, A) bimodules is different from the concept of bimodule categories over $\mathcal{C}$ that we encountered above.

The tensor product in the category $\mathrm{Bimod}_{\mathcal{C}}(A)$ is written as $\otimes_A$ and it ensures that any insertions of A between p and q in $p \otimes_A q$ can be removed. The product $p \otimes_A q$ is given as a subobject of $p \otimes q$ defined by the most general projection π

$$\pi : p \otimes q \to p \otimes_A q \tag{6.4.5}$$

which satisfies the equation

$$(p \otimes A) \otimes q \overset{x_R}{\to} p \otimes q \overset{\pi}{\to} p \otimes_A q = (p \otimes A) \otimes q \overset{\alpha}{\to} p \otimes (A \otimes q) \overset{x_L}{\to} p \otimes q \overset{\pi}{\to} p \otimes_A q \tag{6.4.6}$$

where each side of the equation stands for the composition of the morphisms shown, which are the associators α, the projection π and the morphisms x_L and x_R defining the action of A on p and q. See Figure 6.4.9. The equation tells us that the right action of A on p is *balanced* against the left action of A on q.

The left action of A on $p \otimes_A q$ is defined by the compatibility condition

$$A \otimes (p \otimes q) \overset{\pi}{\to} A \otimes (p \otimes_A q) \overset{x_L}{\to} p \otimes_A q = A \otimes (p \otimes q) \overset{\alpha}{\to} (A \otimes p) \otimes q \overset{x_L}{\to} p \otimes q \overset{\pi}{\to} p \otimes_A q \tag{6.4.7}$$

where each side of the equation means the composition of the appropriate morphisms, namely the associators, the projection, and the left action of A on p. The reader can draw a figure for (6.4.7) in a similar fashion as to Figure 6.4.9 for (6.4.6). Similarly, we define the right action of A on $p \otimes_A q$. These actions manifestly commute and convert $p \otimes_A q$ into an A-bimodule.

The dual of π is ${}^{\circ}\pi : p \otimes_A q \to p \otimes q$ and we have $\pi \circ {}^{\circ}\pi = 1$. The associator in $\mathrm{Bimod}_{\mathcal{C}}(A)$

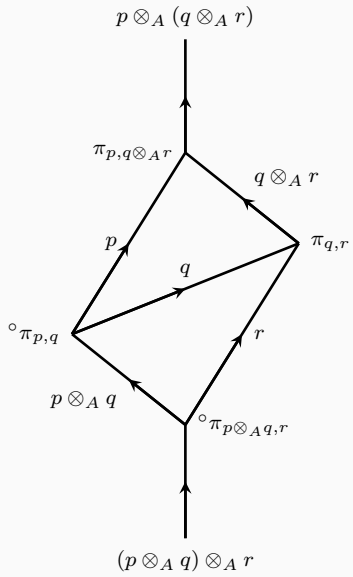


Figure 6.4.10: The associator in $\mathrm{Bimod}_{\mathcal{C}}(A)$.

is then defined as

$$\tilde{\alpha}_{p,q,r} = \pi_{p,q\otimes_A r} \circ (1 \otimes \pi_{q,r}) \circ \alpha_{p,q,r} \circ ({}^\circ\pi_{p,q} \otimes 1) \circ ({}^\circ\pi_{p\otimes_A q,r}), \tag{6.4.8}$$

see Figure 6.4.10. This ensures that any insertions of A can be removed in the diagram defining the associator.

Similarly, the evaluation and co-evaluation maps in $\mathrm{Bimod}_{\mathcal{C}}(A)$ are defined as $\tilde{\epsilon}_p^{L,R} = u \circ \epsilon_p^{L,R} \circ {}^\circ\pi$ and $\widetilde{{}^\circ\epsilon}_p^{L,R} = \pi \circ {}^\circ\epsilon_p^{L,R} \circ {}^\circ u$ where u and ${}^\circ u$ are unit and co-unit morphisms defining A. The argument above suggests that we should take $\mathrm{Bimod}_{\mathcal{C}}(A)$ to be the symmetry category of the gauged theory T/A.

Now, consider two half-spaces, and put the mesh only on the right half. We now have a domain wall between the original theory T on the left and the gauged theory T/A on the right. On this domain wall, the lines of the theory T, in particular those from $\mathcal{C}$, should be able to end on the left and the lines of the theory T/A, namely those from $\mathrm{Bimod}_{\mathcal{C}}(A)$, should be able to end on the right. This suggests that the domain wall is described by the category $\mathrm{Mod}_{\mathcal{C}}(A)$, the category of right A-modules. Note that $\mathrm{Mod}_{\mathcal{C}}(A)$ has a natural left action of $\mathcal{C}$ and a natural right action of $\mathrm{Bimod}_{\mathcal{C}}(A)$, making it naturally a $(\mathcal{C}, \mathrm{Bimod}_{\mathcal{C}}(A))$ bimodule category. Almost by definition, $\mathrm{Bimod}_{\mathcal{C}}(A)$ is the dual of $\mathcal{C}$ with respect to the module category $\mathrm{Mod}_{\mathcal{C}}(A)$.

6.4.5 Gauging of $\mathcal{C}(G)$ to get $\mathrm{Rep}(G)$ and vice versa

As an exercise, let us describe the process explicitly when $\mathcal{C} = \mathcal{C}(G)$ and $\mathrm{Rep}(G)$.

From $\mathcal{C}(G)$ to $\mathrm{Rep}(G)$: We have already defined an algebra A in $\mathcal{C}(G)$ in Sec. 6.4.3. Let us determine $\mathrm{Mod}_{\mathcal{C}(G)}(A)$, the category of right modules of A.

Consider an object $M = \bigoplus N_g g$ where N_g are non-negative integers. We constrain M by demanding that it forms a right-module for A. First of all, we need a morphism $x : M \otimes A \to M$, which is characterized by morphisms $x_{g,g'} : N_g g \otimes g' \to N_{gg'} gg'$, each of which can be though of as an $N_{gg'} \times N_g$ matrix. These matrices have to satisfy two equations. The first one involves the product axiom of A and it tells us that

$$x_{gg',g''} x_{g,g'} = x_{g,g'g''}. \tag{6.4.9}$$

The second one involves the unit axiom of A and it tells us that

$$x_{g,e} = 1. \tag{6.4.10}$$

Combining these two, we find that

$$x_{g,g'} x_{gg',g'^{-1}} = 1 \tag{6.4.11}$$

which implies that $N_g = N$ for all g and

$$x_{g,g'} = x_{e,gg'} (x_{e,g})^{-1}. \tag{6.4.12}$$

So, we find that the right modules are $M = N \bigoplus_g g = NA$ with arbitrary invertible $N \times N$ matrices $x_{e,g}$.

Now, we find the morphisms in $\mathrm{Mod}_{\mathcal{C}(G)}(A)$. The first observation is that all $(NA, x_{e,g})$ are isomorphic to $(NA, 1)$. A morphism from right to left is provided by sending Ng inside $(NA, 1)$ to Ng inside $(NA, x_{e,g})$ by the matrix $x_{e,g}$. The condition that this morphism commutes with the action of A turns out to be (6.4.12) and hence this morphism is indeed a module morphism. The morphism from left to right via the inverse matrix is the inverse of this module morphism. So, we can restrict our attention to objects $NA \equiv (NA, 1)$.

Let us then find module morphisms from NA to $N'A$. Such a morphism φ is specified by an $N \times N'$ matrix φ_g sending Ng inside NA to $N'g$ inside $N'A$. The condition that φ be a module morphism implies that φ is a constant $N \times n'$ matrix independent of g. Thus, we identify $\mathrm{Mod}_{\mathcal{C}(G)}(A)$ as the category Vec.

Next, let us determine $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$. By a similar analysis as above, we can write such a bimodule as $(NA, (x_L)_{g,e}, (x_R)_{e,g})$ where $(x_R)_{e,g}$ and $(x_L)_{g,e}$ encode the morphisms $M \otimes A \to M$ and $A \otimes M \to M$. As before we can restrict our attention to $(NA, (x_L)_{g,e}, 1)$. Demanding that the left and right actions of A are compatible, we find that

$$(x_L)_{g,e} (x_L)_{g',e} = (x_L)_{gg',e} \tag{6.4.13}$$

which means that the objects of $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ are identified as representations of G. One can check that the morphisms of $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ are precisely the intertwiners between the representations. Thus $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ is equivalent to $\mathrm{Rep}(G)$. We can also work out $\otimes_A$ and it agrees with the tensor product on $\mathrm{Rep}(G)$.

From $\mathrm{Rep}(G)$ **to** $\mathcal{C}(G)$**:** Now the algebra object $A \in \mathrm{Rep}(G)$ is the regular representation. The regular representation is spanned by basis vectors $\widehat{g}$ in one-to-one correspondence with the group elements g. If we denote the action of g by $\rho(g)$ then $\rho(g)\widehat{h} = \widehat{gh}$. The algebra multiplication takes $\widehat{g} \otimes \widehat{h}$ to $\delta_{gh}\widehat{g}$ which is an intertwiner. The unit morphism $1 \to A$ corresponds to choosing $\sum_g \widehat{g}$ in A.

First, we look for right modules of A. Choose some representation R of G. We denote the action of g on $\vec{q} \in R$ as $g\vec{q}$. Note that $g\vec{q} \neq 0$ for any g and any non-zero $\vec{q}$ because otherwise this would imply that $\vec{q} = g^{|G|}\vec{q} = 0$. The right-action of A on R must satisfy $g(\vec{p}\,\widehat{h}) = (g\vec{p})\widehat{h}$. We must also have that $(\vec{p}\,\widehat{h})\,\widehat{k} = \delta_{hk}\vec{p}\,\widehat{h}$ and $\vec{p}(\sum_g \widehat{g}) = \vec{p}$. Now, start with some arbitrary vector $\vec{q} \neq 0$ such that $\vec{q}\,\widehat{g} \neq 0$ for some g. Then, it must be true that $\vec{p} := (g^{-1}\vec{q})\,\widehat{1} \neq 0$. Then, $\vec{p}\,\widehat{1} = \vec{p}$ and $\vec{p}\,\widehat{g} = 0$ for all $g \neq 1$. This implies that for every g we obtain a vector $\vec{p}_g = g\vec{p}$ which has the following properties: $\vec{p}_g\,\widehat{h} = \delta_{gh}\vec{p}_g$. Also, all $\vec{p}_g$ must be linearly independent because $\sum \alpha_g \vec{p}_g = 0$ can be hit by $\widehat{h}$ from the right to yield $\alpha_h = 0$. Thus the set formed by $\vec{p}_g$ for all g is the regular representation A of G and R breaks up as a sum of regular representations. Thus, the objects of $\mathrm{Mod}_{\mathrm{Rep}(G)}(A)$ are all isomorphic to NA for some non-negative integer N. One can easily work out the module morphisms as well and then one finds that $\mathrm{Mod}_{\mathrm{Rep}(G)}(A)$ is equivalent to Vec.

Next, we look for bi-modules of A. Pick A as a particular right module for A. It can be converted into a bimodule V_g by simply stating that $\widehat{h}\,\vec{p}_1 = \delta_{gh}\vec{p}_1$. The G-equivariance then determines that $\widehat{h}\,\vec{p}_k = \delta_{kg,h}\vec{p}_k$. Thus objects of $\mathrm{Bimod}_{\mathrm{Rep}(G)}(A)$ are labeled by non-negative integers n_g for each g. One can similarly figure out bimodule morphisms leading to the result that $\mathrm{Bimod}_{\mathrm{Rep}(G)}(A)$ is equivalent to $\mathcal{C}(G)$. Again, the tensor product $\otimes_A$ can be worked out and it equals the group multiplication.

6.4.6 Gaugings and module categories

Let us come back to the general question. We said that given a theory T with a symmetry $\mathcal{C}$ and a symmetric Frobenius algebra object $A \in \mathcal{C}$, we can put a fine mesh of A on the two-dimensional manifold to define the gauged theory T/A. We now ask the question: when do two algebra objects A and A' give rise to the same gauged theory $T/A \simeq T/A'$? For this, the symmetries of the gauged theories should be the same, $\mathrm{Bimod}_{\mathcal{C}}(A) \simeq \mathrm{Bimod}_{\mathcal{C}}(A')$ and the properties of the domain walls should also be the same, $\mathrm{Mod}_{\mathcal{C}}(A) \simeq \mathrm{Mod}_{\mathcal{C}}(A')$.

Let us see that conversely, if $\mathrm{Mod}_{\mathcal{C}}(A) \simeq \mathrm{Mod}_{\mathcal{C}}(A')$, the gauged theories are the same $T/A \simeq T/A'$. Indeed, the equivalence is captured by an (A, A') bimodule m in $\mathcal{C}$ which sends an object M' of $\mathrm{Mod}_{\mathcal{C}}(A')$ to an object M of $\mathrm{Mod}_{\mathcal{C}}(A)$ via $M = m \otimes_{A'} M'$ where $\otimes_{A'}$ denotes a tensor product in the category $\mathrm{Bimod}_{\mathcal{C}}(A')$. There is an inverse (A', A) bimodule n providing the inverse map. Physically, m corresponds to a topological interface between T/A and T/A' which can be fused with boundary conditions of T/A' to yield boundary conditions of T/A. n is the inverse interface. Algebras A and A' such that $\mathrm{Mod}_{\mathcal{C}}(A) \simeq \mathrm{Mod}_{\mathcal{C}}(A')$ are called *Morita equivalent*.

We claim that the existence of such an invertible interface guarantees that T/A and T/A' are isomorphic theories. For instance, consider the Hilbert space V of T/A on S^1. It can be mapped

to the Hilbert space V' of T/A' on S^1 by considering a cylinder geometry of infinitesimal time with the insertion of a wrapped domain wall n in between. Similarly, we have an inverse map from V' to V constructed similarly from m. Now, we can compose these maps by putting one cylinder on top of the other. As the interfaces are topological, they can be moved towards each other and ultimately fused away as they are inverses of each other. Thus, we are left with a unitary evolution on an infinitesimal length cylinder geometry. We can now take the infinitesimal to zero in all the steps above to obtain an isomorphism between V and V'. Similar arguments can be used to show that any kind of data of T/A is isomorphic to the same kind of data of T/A'.

We would like to stress that this equivalence of T/A and T/A' is different from our claim that T and T/A contain the same amount of information. For example, in the latter case, the Hilbert spaces would be different, as we will detail in Sec. 6.4.8.

All of this means that what really characterizes different gaugings are not the algebra objects used in the gaugings themselves but the associated module categories $\mathrm{Mod}_{\mathcal{C}}(A)$ over $\mathcal{C}$. We can then ask the question: does every module category $\mathcal{M}$ over $\mathcal{C}$ come from a symmetric Frobenius algebra object A in $\mathcal{C}$, so that $\mathcal{M}$ arises as the domain wall between the original theory T and the gauged theory T/A? The answer is yes [168, 179].

The construction goes roughy as follows. Given a module category $\mathcal{M}$ over $\mathcal{C}$, consider the morphism spaces $\mathrm{Hom}(c \otimes m, n)$ for $c \in \mathcal{C}$ and $m, n \in \mathcal{M}$. There is an object called the internal hom and denoted by $\underline{\mathrm{Hom}}(m, n)$ in $\mathcal{C}$ which satisfies

$$\mathrm{Hom}_{\mathcal{M}}(c \otimes m, n) \simeq \mathrm{Hom}_{\mathcal{C}}(c, \underline{\mathrm{Hom}}(m, n)). \tag{6.4.14}$$

At the level of objects, this is given by

$$\underline{\mathrm{Hom}}(m, n) = \bigoplus_c (\dim \mathrm{Hom}_{\mathcal{M}}(c \otimes m, n))c \tag{6.4.15}$$

where c runs over the isomorphism classes of simple objects in $\mathcal{C}$ as can be simply verified by computing $\mathrm{Hom}_{\mathcal{C}}(c, \underline{\mathrm{Hom}}(m, n))$ from this equation.

Now, the internal homs can be concatenated naturally, in the sense that there is a natural morphism

$$\mu : \underline{\mathrm{Hom}}(m, n) \otimes \underline{\mathrm{Hom}}(n, o) \to \underline{\mathrm{Hom}}(m, o) \tag{6.4.16}$$

and in particular $\underline{\mathrm{Hom}}(m, m)$ has the multiplication. It can be shown [168] that $A = \underline{\mathrm{Hom}}(m, m)$ for a simple object $m \in \mathcal{M}$ is an algebra object such that $\mathrm{Mod}_A(C) \simeq \mathcal{M}$. Furthermore, A is automatically symmetric Frobenius if $\mathcal{C}$ is a symmetry category [179]. Summarizing, we have

Distinct choices of gaugings of a theory with symmetry category $\mathcal{C}$ are in one to one correspondence with choices of indecomposable module categories $\mathcal{M}$ over $\mathcal{C}$ characterizing the domain wall between the original theory and the gauged theory. The gauged theory has the symmetry $\mathcal{C}'$ which is the dual of $\mathcal{C}$ with respect to $\mathcal{M}$.

We denote the gauged theory corresponding to a module category $\mathcal{M}$ by $T/\mathcal{M}$. We would like to note that a decomposable module category corresponds to a direct sum of algebras $A = A_1 \oplus A_2$ where the algebra structure of A comes from algebra structure of A_1 and the algebra structure

of A_2. Gauging a theory T using A produces $T/A = T/A_1 \oplus T/A_2$ because a mesh of A is the same as a mesh of A_1 plus a mesh of A_2.

6.4.7 (Re-)gauging and its effect on the symmetry category

Now we can give a unified description of the (re-)gauging process. Consider a theory T with a symmetry $\mathcal{C}$. Pick an indecomposable module category $\mathcal{M}$ of $\mathcal{C}$. There is an algebra object $A \in \mathcal{C}$ such that $\mathcal{M} \simeq \mathrm{Mod}_{\mathcal{C}}(A)$, and then the gauged theory $T/\mathcal{M} := T/A$ has the symmetry $\mathcal{C}' = \mathrm{Bimod}_{\mathcal{C}}(A)$ such that $\mathcal{M}$ is a natural $(\mathcal{C}, \mathcal{C}')$ bimodule category.

Now, pick an indecomposable module category $\mathcal{M}'$ of $\mathcal{C}'$ and repeat the same process. We get the gauged theory $T/\mathcal{M}/\mathcal{M}'$ which has the symmetry $\mathcal{C}''$ such that $\mathcal{M}'$ is a $(\mathcal{C}', \mathcal{C}'')$ bimodule category. We now see that the 'set' of all symmetry categories can be subdivided into 'subsets' consisting of symmetry categories that can be converted into each other by a sequence of gauging.

Note that this double gauging can be done in one step, since from two bimodule categories $\mathcal{M}$ and $\mathcal{M}'$ we can form the tensor product $\mathcal{M} \boxtimes_{\mathcal{C}'} \mathcal{M}'$ which is a $(\mathcal{C}, \mathcal{C}'')$ bimodule category. Notationally it is convenient to write a $(\mathcal{C}, \mathcal{C}')$ bimodule category $\mathcal{M}$ as a morphism $\mathcal{M} : \mathcal{C} \to \mathcal{C}'$. Then we can regard $\mathcal{M} \boxtimes_{\mathcal{C}'} \mathcal{M}'$ as $\mathcal{M}' \circ \mathcal{M}$. Then $T/\mathcal{M}/\mathcal{M}' \simeq T/(\mathcal{M}' \circ \mathcal{M})$. Any multiple gauging can be done in one step.

In particular, $\mathcal{M}^{\mathrm{op}}$ given by reversing the order of the tensor product is naturally a $(\mathcal{C}', \mathcal{C})$ bimodule category, and $\mathcal{M}^{\mathrm{op}} \circ \mathcal{M}$ is the identity. Therefore we have $T/\mathcal{M}/\mathcal{M}^{\mathrm{op}} = T$. The example we started this chapter with, $T/G/\widehat{G} = T$ for an Abelian group G, is a special instance of this construction.

It is instructive to phrase the re-gauging process in terms of the algebra A as well, whose detail can be found in [169] with plenty of helpful drawings. The category of bimodules $\mathrm{Bimod}_{\mathcal{C}}(A)$ have an algebra object $B = A^* \otimes A$. B is also trivially an algebra in $\mathcal{C}$ and its algebra multiplication is given by the evaluation map. Moreover, we can embed A into B as $A \to A^* \otimes A \otimes A \to A^* \otimes A = B$ where the first map is the co-evaluation map and the second map is the multiplication for A. We are looking for B-bimodules in $\mathrm{Bimod}_{\mathcal{C}}(A)$ which can be identified with B-bimodules in $\mathcal{C}$ because A embeds inside B and hence a B-bimodule structure carries an A-bimodule structure. Now, B-bimodules in $\mathcal{C}$ are trivially obtained from any object $a \in \mathcal{C}$ by the mapping $A^* \otimes a \otimes A$. This mapping gives us an equivalence of $\mathcal{C}$ and the category of B-bimodules in $\mathrm{Bimod}_{\mathcal{C}}(A)$.

Physically, this means that the lines of $T/\mathcal{M}/\mathcal{M}^{\mathrm{op}}$ come as lines of T dressed by A^* on one side and by A on the other. The local operators in $T/\mathcal{M}/\mathcal{M}^{\mathrm{op}}$ also correspond to a local operator of T which appears at the junction of "middle" lines and the A-lines on the side are just joined smoothly. Let's insert a network of lines of $T/\mathcal{M}/\mathcal{M}^{\mathrm{op}}$ on a surface Σ and compute the partition function. The network will look like a network of lines of T with additional A loops, one for each "plaquette". These A loops can be shrunk away leaving the partition function of a network of lines of T. Thus we see explicitly how we obtain the same theory on regauging. Later on, we will discuss 2d TFTs with $\mathcal{C}$ symmetry and we will give a prescription for constructing the gauged TFT which uses heavily the algebra A rather than $\mathcal{M}$. From this argument, it is clear that regauging the gauged TFT will give back the original TFT.

Summarizing, we have the following:

Denote by $\mathcal{M} : \mathcal{C} \to \mathcal{C}'$ when the symmetry category $\mathcal{C}$ has a gauging $\mathcal{M}$ under which the dual symmetry is $\mathcal{C}'$. Two consecutive gaugings $\mathcal{M} : \mathcal{C} \to \mathcal{C}'$ and $\mathcal{M}' : \mathcal{C}' \to \mathcal{C}''$ can be composed to $\mathcal{M}' \circ \mathcal{M} : \mathcal{C} \to \mathcal{C}''$, and any gauging $\mathcal{M} : \mathcal{C} \to \mathcal{C}'$ has an inverse $\mathcal{M}^{-1} : \mathcal{C}' \to \mathcal{C}$ such that $\mathcal{M}^{-1} \circ \mathcal{M}$ is the identity.

Now, let us ask when are two symmetry categories $\mathcal{C}$ and $\mathcal{C}'$ dual with respect to some indecomposable module category $\mathcal{M}$. It is known [145] that this happens if and only if $Z(\mathcal{C}) \simeq Z(\mathcal{C}')$, where $Z(\mathcal{C})$ is the Drinfeld center of $\mathcal{C}$. This has the following physical interpretation.

Recall first that any symmetry category where every simple object is invertible is of the form $\mathcal{C} = \mathcal{C}(G, \alpha)$. From the same data of G and $\alpha \in H^3(G, \mathrm{U}(1))$, we can construct a 3d TFT called the Dijkgraaf-Witten theory [173]. On the boundary of the Dijkgraaf-Witten theory, we can put a 2d theory with symmetry G with an anomaly α.

There is a generalization of this construction that gives a 3d TFT starting from any symmetry category $\mathcal{C}$, sometimes called the generalized Turaev-Viro construction. We will just call them the Dijkgraaf-Witten theory associated to $\mathcal{C}$. The Drinfeld center $Z(\mathcal{C})$ is the braided tensor category which captures the properties of the line operators of this 3d TFT, and this 3d TFT can have a boundary where a 2d theory with symmetry $\mathcal{C}$ lives. For a recent exposition, see e.g. [3].

Since T/A is obtained by just adding a mesh of A on T, T/A and T can be put on boundaries of the same 3d TFT. Therefore their Drinfeld centers should be the same. Conversely, specifying an isomorphism $Z(\mathcal{C}) \simeq Z(\mathcal{C}')$ corresponds to specifying a 1d domain wall on a 2d boundary of a single 3d TFT, such that we have lines from $\mathcal{C}$ on the left and lines from $\mathcal{C}'$ on the right of the domain wall. This corresponds to specifying a $(\mathcal{C}, \mathcal{C}')$ bimodule implementing the gauging. Summarizing,

Two symmetry categories $\mathcal{C}$ and $\mathcal{C}'$ can be transformed to each other by an appropriate gauging if and only if the 3d TFTs associated to them are equivalent, i.e. $Z(\mathcal{C}) \simeq Z(\mathcal{C}')$.

The fact $Z(\mathcal{C}) \simeq Z(\mathcal{C}')$ also implies that the total dimension of $\mathcal{C}$ and $\mathcal{C}'$ are the same,

$$\dim \mathcal{C} = \dim \mathcal{C}'. \tag{6.4.17}$$

This is because it is known that $\dim Z(\mathcal{C}) = |\dim \mathcal{C}|^2$.

Another related point is the following. This mathematical process of choosing a module category or an algebra object corresponding to it for a symmetry category and taking the dual category can be performed for a modular tensor category. A modular tensor category describes topological line operators in a three-dimensional theory and therefore this operation should be thought of as creating a new three-dimensional theory from an old one. This operation is in fact known under the name of *anyon condensation* [180].

6.4.8 The effect of the gauging on Hilbert space on S^1

Let us now discuss the Hilbert space on S^1 of a theory with symmetry $\mathcal{C}$, and how it is modified by the gauging.

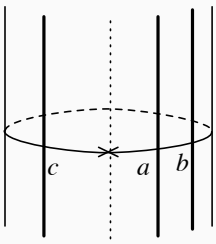


Figure 6.4.11: An example of a circle with transverse line operators. The point marked with an X labels the choice of the base point. In the diagrams here and below, the time flows upwards, and the lines also carry arrows in the upward direction unless otherwise mentioned.

Backgrounds for $\mathcal{C}$ symmetry

Consider a cylinder of the form $S^1 \times \mathbb{R}_t$, with lines a_i in $\mathcal{C}$ transverse to the constant time slice $t = 0$. Let us assume that other data about the circle like its size, spin-structure etc. have been fixed so that we can associate a Hilbert space of states to this circle. See Figure 6.4.11 for an example. In the figures here and below, the time flows from the bottom to the top.

We would like to identify this configuration of lines with an object a of $\mathcal{C}$ and say that the circle carries the *background* a for the symmetry $\mathcal{C}$. When $\mathcal{C} = \mathcal{C}(G)$, the backgrounds are labeled by group elements and correspond to holonomy of the G-connection around the circle. As the lines are topological, we can try to define such an a by fusing them together. Clearly, there are choices in the precise way one performs this procedure: there is a choice of the *base-point* from which we start taking the tensor product, and there is also a choice in the relative order of fusion of the lines, which is related to the associators in $\mathcal{C}$. To spell this out more fully, we need some notations and operations.

Let us denote the Hilbert space for a cylinder with just one line operator $a \in \mathcal{C}$ by V_a. We require $V_{a\oplus b} = V_a \oplus V_b$. A morphism $m : a \to b$ defines an operator $Z(m) : V_a \to V_b$. This is given by the geometry shown in Fig. 6.4.12. Note that in a non-topological theory, the vertical height of the cylinder containing the topological line and/or the topological operator needs to be taken to zero. The same comment applies to all the figures discussed below in this section. Now, two choices of fusing a, b and c in Fig. 6.4.11 can be taken care of by using the associator α and Z:

$$Z(\alpha_{a,b,c}) : V_{(a\otimes b)\otimes c} \to V_{a\otimes(b\otimes c)}. \tag{6.4.18}$$

To track the change of the base point, we need to introduce the map $X_{a,b} : V_{a\otimes b} \to V_{b\otimes a}$ defined by Fig. 6.4.13, together with its inverse. They are associated to the intersection of the topological line and the *auxiliary line*, which is the trajectory of the base point.

The operations $X_{a,b}$ and $Z(\alpha)$ satisfy many consistency relations, which we spell out in full in Sec. 6.6.2. With them, we can keep track of the change of the base point and the change of the order of fusing the lines in a consistent manner. This allows us to associate a Hilbert space to a circle with multiple transverse line operators $a_1, \ldots, a_k$.

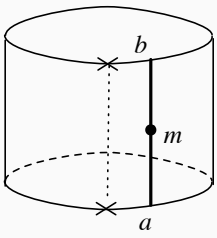


Figure 6.4.12: A morphism $m : a \to b$ defines an operator $Z(m) : V_a \to V_b$.

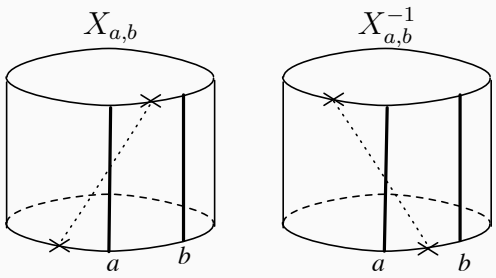


Figure 6.4.13: $X_{a,b}$ and its inverse tracks the change of the base points.

The operators X and Z can also be used to express how a line operator wrapped around S^1 acts on the Hilbert space. Indeed, given two morphisms $m : a \to c \otimes d$ and $n : d \otimes c \to b$, we can define the operator $U_{c,d}(m, n)$ using the network drawn in Fig. 6.4.14. Explicitly, this is given by

$$U_{c,d}(m, n) = Z(n)X_{c,d}Z(m). \tag{6.4.19}$$

The action U_a of a line operator labeled by a wrapped around on S^1 on the Hilbert space V_1 on S^1 without any transverse line, referred to in Sec. 6.3.1, is a special example of this construction: $U_a = U_{a^*,a}({}^\circ\epsilon_a, \epsilon_a)$ where ${}^\circ\epsilon_a$, ϵ_a are the co-evaluation map and the evaluation map introduced in (6.3.10), (6.3.11).

The Hilbert space of the gauged theory

So far we discussed the Hilbert spaces V_a of the original theory T with the symmetry $\mathcal{C}$, where $a \in \mathcal{C}$ is the label of the line operator transverse to the constant time slice S^1. Let us now discuss how Hilbert spaces W_p of the gauged theory T/A can be found, where $p \in \mathcal{C}'$.

A line in the gauged theory is given by an (A, A) bimodule in $\mathcal{C}$. Each such bimodule p can be viewed as an object in $\mathcal{C}$ and hence has a Hilbert space V_p associated to it by the ungauged theory T. We construct the Hilbert space W_p associated to p by the gauged theory T/A as a subspace of V_p in the following manner.

In the case of $\mathcal{C}(G)$, one is traditionally instructed to project this space to the subspace left

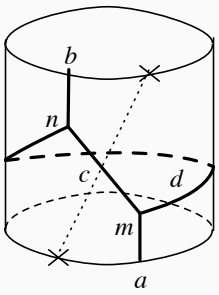


Figure 6.4.14: The action $U_{c,d}(m,n)$ of a line operator wrapped around S^1.

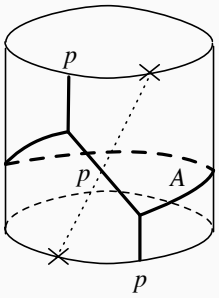


Figure 6.4.15: The projector defining the Hilbert space of the gauged theory is a specific instance of an action of a wrapped line operator. The morphisms used at the trivalent vertices are those specifying the bimodule structure of p.

invariant by the action of the group G. For a general symmetry category $\mathcal{C}$ and its gauging A, we need a projector $P : V_p \to W_p$. Such a projector is naturally given by $P := U_{p,A}({}^\circ x_R, x_L)$ where ${}^\circ x_R : p \to p \otimes A$ and $x_L : A \otimes p \to p$ are the morphisms defining the module and comodule structures on p, see Fig. 6.4.15. This projector was written down in [181].

Let us now show that this projector to W_p agrees with the traditional definition in the case of $\mathcal{C}(G)$, when p is the identity object in the category of bimodules. This means that $p = A = \bigoplus_g g$ as an object in $\mathcal{C}(G)$. Therefore we are going to project $V_A = \bigoplus_g V_g$. The operator P restricted to V_0 is given by $|G|^{-1} \sum_g g$, and this is indeed a projector to the G-invariant subspace. The action of P on the whole V_A can be found similarly, and we find that

$$W_1 = PV_A \simeq \bigoplus_C (G_c\text{-invariant subspace of } V_c) \tag{6.4.20}$$

where C runs over conjugacy classes of G, $c \in C$ is a representative element, and G_c is the commutant of c in G. This is as it should be.

Summarizing, we have the following statement:

The Hilbert space of the gauged theory T/A is obtained by taking the invariant part of the ungauged theory T under the projector naturally defined by wrapping A around S^1.

6.5 More examples of symmetry categories and their gauging

In the previous sections we reviewed the general theories of symmetry categories and their gauging. So far, however, we only saw basic examples where symmetry categories are either of the form $\mathcal{C}(G,\alpha)$ where G is a finite group and α its anomaly, or of the form $\mathrm{Rep}(G)$ where G is a group. In this section we discuss many other examples and their gauging.

6.5.1 Symmetry category with two simple lines

The simplest symmetry category consists of just one simple object 1 and its multiples. This can be thought of as a symmetry of any 2d theory, but it is not very interesting.

Let us then consider the simplest nontrivial symmetry category, consisting of simple objects 1 and x. The dual of x can only be x, and then $x\otimes x$ can only contain one copy of 1. Therefore the tensor product can only be of the form $x\otimes x = nx\oplus 1$, where n is a non-negative integer.

When $n=0$, then the simple objects 1 and x form the group $\mathbb{Z}_2$. As we already discussed, there are two possible symmetry categories $\mathcal{C}(\mathbb{Z}_2,\alpha)$ where $\alpha\in H^3(\mathbb{Z}_2,\mathrm{U}(1))=\mathbb{Z}_2$ determines the anomaly.

When $n=1$, the dimension of x is easily determined to be $(1+\sqrt{5})/2$. The only nontrivial condition is the pentagon identity of the associator $a_{x,x,x}$, and can be solved uniquely [182]. This symmetry category in fact has a braiding which turns it into a modular tensor category describing an anyon system. Recently this is known under the name of Fibonacci anyons because fusing n copies of x generates the $F_{n-1}1\oplus F_n x$ where F_n is the nth Fibonacci number. For this reason we denote this category Fib.

It is known that we cannot have $n>1$, as shown by Ostrik [183]. Therefore, possible symmetry categories with two simple lines are just three, $\mathcal{C}(\mathbb{Z}_2,0)$, $\mathcal{C}(\mathbb{Z}_2,1)$ and Fib. Ostrik also classified all possible symmetry categories with three simple objects [184].

6.5.2 Symmetry category of $\mathrm{SU}(2)$ WZW models and other RCFTs

Next, we review the symmetry category of RCFTs, following the construction of [185]. Let us take a rational chiral algebra $\mathcal{A}$ in two dimensions and consider a conformal field theory T which corresponds to the diagonal modular invariant of this algebra $\mathcal{A}$. As is well-known, from a chiral vertex operator a corresponding to an irreducible representation of $\mathcal{A}$, we can construct a topological line operator a of this theory T. Because the theory T corresponds to the diagonal modular invariant, chiral and antichiral vertex operators generate the same line operators. Therefore the theory has topological lines generated by irreducible representations of a single copy of $\mathcal{A}$. They of course are specified by the Moore-Seiberg data, or equivalently they

form a unitary modular tensor category $\mathcal{C}$. We can forget the braiding of $\mathcal{C}$ and regard it as the symmetry category of this theory T.

The essential observation of [185] is that the choice of gauge-able subpart of $\mathcal{C}$, or equivalently the choice of the module category $\mathcal{M}$ over $\mathcal{C}$, is in one to one correspondence with the choice of modular invariants of the chiral algebra $\mathcal{A}$. In particular, all modular invariants, including the exceptional ones, arise as the result of a generalized gauging.

Let us describe them in more detail in the case of $\mathrm{SU}(2)$ WZW models, following [186]. The chiral algebra is $\widehat{\mathrm{SU}}(2)_k$, which has $k+1$ irreducible representations

$$V_j, \qquad j = 0, 1/2, \cdots, k/2 \tag{6.5.1}$$

with the fusion rule

$$V_j \otimes V_{j'} = V_{|j-j'|} \oplus V_{|j-j'|+1} \oplus \cdots \oplus V_m \tag{6.5.2}$$

where $m = \min(j+j', k-(j+j'))$. We have

$$\dim V_j = \frac{q^{j+1/2} - q^{-j-1/2}}{q^{1/2} - q^{-1/2}} \qquad \text{where} \quad q = e^{2\pi i/(k+2)}. \tag{6.5.3}$$

They form a symmetry category we denote by $\mathrm{Rep}(\widehat{\mathrm{SU}}(2)_k)$. The object V_0 is the identity. It is clear that the objects V_j with integral j form a symmetry subcategory, and can be denoted by $\mathrm{Rep}(\widehat{\mathrm{SO}}(3)_k)$. In particular, when $k=3$, this is equivalent to the symmetry category Fib discussed above.

Since $V_{k/2} \otimes V_{k/2} = V_0$, the simple lines V_0 and $V_{k/2}$ form a sub-symmetry category. From our general discussion above, this is equivalent to $\mathcal{C}(\mathbb{Z}_2, \alpha)$ where $\alpha \in H^3(\mathbb{Z}_2, \mathrm{U}(1)) = \mathbb{Z}_2$. This α is determined in terms of the associator, or equivalently the fusion matrix or the quantum 6j symbol involving four $V_{k/2}$, and is known to be $\alpha = k \in \mathbb{Z}_2$. In particular $\mathrm{Rep}(\widehat{\mathrm{SU}}(2)_1) = \mathcal{C}(\mathbb{Z}_2, 1)$. This also means that the subsymmetry formed by V_0 and $V_{k/2}$ is gauge-able when k is even. Gauging it we obtain the modular invariant of type $D_{k/2+2}$.

The $E_{6,7,8}$ modular invariants correspond to algebra objects

$$A = \begin{cases} V_0 \oplus V_3, & k = 10 \quad (E_6), \\ V_0 \oplus V_4 \oplus V_8, & k = 16 \quad (E_7), \\ V_0 \oplus V_5 \oplus V_9 \oplus V_{14}, & k = 28 \quad (E_8). \end{cases} \tag{6.5.4}$$

The type X_n of the modular invariants, or equivalently of the possible gauging, specifies the corresponding module category structure as follows: the isomorphism classes of simple objects in the module category are labeled by the nodes of the Dynkin diagram of X_n, and the edges describe how $V_{1/2}$ acts on the simple objects.

6.5.3 Gauging a subgroup of a possibly-anomalous group

Generalities

We now turn our attention to a more traditional setup of gauging a subgroup H of a bigger group G. We will soon see that already in this traditional-looking setup we encounter various surprises.

We start by specifying the anomaly of the bigger group; we start from a symmetry category $\mathcal{C}(G,\alpha)$ where $\alpha \in H^3(G,\mathrm{U}(1))$. Possible gaugings are classified by their module categories, as already discussed. They turn out to be in one-to-one correspondence with a pair (H,ψ) where H is a subgroup such that the restriction of α to H is trivial, and ψ is an element of $H^2(H,\mathrm{U}(1))$.[9] This result agrees with the more traditional viewpoint: we choose a non-anomalous subgroup H and then choose the discrete torsion ψ.

From the general machinery described above, the gauged theory has a symmetry category which is the dual of $\mathcal{C}(G,\alpha)$ with respect to (H,ψ). Let us denote the resulting symmetry category by $\mathcal{C}' := \mathcal{C}(G,\alpha;H,\psi)$. When $\alpha = 0$, $H = G$, $\psi = 0$, we already know that $\mathcal{C}' = \mathrm{Rep}(G)$. Furthermore, when G is abelian, $\mathcal{C}' = \mathcal{C}(\widehat{G},0)$. The explicit structure of $\mathcal{C}'$ in the general case can be determined by realizing it as a category of bimodules $\mathrm{Bimod}_{\mathcal{C}}(A)$ for the algebra object A corresponding to (H,ψ).

Let us see what we can say generally. Firstly, there are two general facts:

- the dimensions of lines in $\mathcal{C}'$ are all integral; such symmetry categories are called as integral symmetry categories.
- the total dimension $\dim \mathcal{C}' := \sum_a (\dim a)^2$ is the same as the original one: $\dim \mathcal{C}' = |G|$.

Secondly, there are cases where the dual symmetry $\mathcal{C}(G,\alpha;H,\psi)$ itself is of the form $\mathcal{C}(G',\alpha')$ for some group G' and the anomaly α'. There is a theorem by Naidu, Nikshych [187, 188] and Uribe [189] determining exactly when this happens, and if so, explicitly the form of G' and α'. The general formula is too complicated to reproduce in full here. A necessary condition is that H is an Abelian normal subgroup. When α is trivial this in fact suffices. In the next subsubsection we describe its explicit structure.

Gauging a normal Abelian subgroup of a non-anomalous group

Statement: Let us choose a group G and its normal Abelian subgroup H. We then gauge H. The gauged theory then has a symmetry group G' with an anomaly $\alpha' \in H^3(G',\mathrm{U}(1))$, given as follows.

The fact that H is a normal Abelian subgroup means that G is an extension

$$0 \to H \to G \to K = G/H \to 0 \tag{6.5.5}$$

[9] More precisely, when α is nontrivial, ψ is an element of a torsor over $H^2(H,\mathrm{U}(1))$.

and as such it is determined by an an action of K on H by inner automorphisms in G, and an element $\kappa \in H^2(K,H)$ defined using the group action. The group G is a crossed product, $G = H \rtimes_\kappa K$. Let us identity $G = H \times K$ *as a set*. Then the group structure is given as follows:

$$(h,k)(h',k') = (h(k \triangleright h')\kappa(k,k'), kk') \tag{6.5.6}$$

where $k \triangleright h'$ is the action of k on h' and $\kappa(k,k')$ is an H-valued 2-cocycle of K.

Denote by $\widehat{H}$ the dual group of H. There is a natural action of K on $\widehat{H}$ given by $k \triangleright \rho(h) = \rho(k^{-1} \triangleright h)$ for arbitrary elements $k \in K$, $h \in H$ and $\rho \in \widehat{H}$. Under this action,

$$G' = \widehat{H} \rtimes K \tag{6.5.7}$$

with the trivial two-cocycle in $H^2(K, \widehat{H})$, and α' is given by

$$\alpha' = \alpha_\kappa \qquad \text{where} \quad \alpha_\kappa((\rho,k),(\rho',k'),(\rho'',k'')) = \rho''(kk' \triangleright \kappa(k,k')) \tag{6.5.8}$$

Note that the nontriviality κ of the crossed product in the original G side is traded for the nontriviality of the anomaly α_κ on the G' side. Summarizing, we have

Let us gauge a normal Abelian subgroup H of a symmetry group G. G is then necessarily of the form $G = H \rtimes_\kappa K$, where $\kappa \in H^2(K,H)$. When G has no anomaly, the gauged theory has the symmetry group $G' = \widehat{H} \rtimes K$, and the resulting anomaly α' is given in terms of κ as in (6.5.8).

Derivation: Let us now derive the description of G' given in the last paragraph. Our starting category is $\mathcal{C}(G)$ and we want to gauge it by the algebra object $A = \bigoplus h$ where $h \in H$. The simple objects of gauged category $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ are bimodules which can be seen to form the set $\widehat{H} \times K$ using arguments very similar to those in Sec. 6.4.5. An object (ρ,k) in $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ is built from the object $\bigoplus_h (h,k)$ in $\mathcal{C}(G)$. Our choice of the bimodule structure on (ρ,k) is that the right action by A is trivial and the left action by A is given in terms of morphisms $(x_L)_{h,h'} : (h,e) \otimes (h',k) \to (hh',k)$ satisfying the familiar condition

$$(x_L)_{h,h'} = (x_L)_{hh',e}((x_L)_{h',e})^{-1} \tag{6.5.9}$$

with

$$(x_L)_{h,e} = \rho(h). \tag{6.5.10}$$

The balanced tensor product of (ρ,k) and (ρ',k') in $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ is given in terms of projectors $\pi_{h,h'} : (h,k) \otimes (h',k') \to (h\kappa(k,k')(k \triangleright h'), kk')$. The equation (6.4.6) tells us that

$$\pi_{h(k\cdot h'),e} = \pi_{h,h'}\rho'(h') \tag{6.5.11}$$

Demanding the right action on $(\rho,k) \otimes_A (\rho',k')$ to be trivial leads us to the condition that

$$\pi_{h,h'} = \pi_{h,e} \tag{6.5.12}$$

which can be substituted into (6.5.11) to simplify it to

$$\pi_{h(k\triangleright h'),e} = \pi_{h,e}\rho'(h'). \tag{6.5.13}$$

Via (6.4.7), the left action ρ'' on $(\rho,k)\otimes_A(\rho',k')$ satisfies

$$\rho''(h)\pi_{e,e} = \pi_{h,e}\rho(h) \tag{6.5.14}$$

which can be combined with (6.5.13) to yield

$$\rho''(h) = \rho(h)\rho'(k^{-1}\triangleright h). \tag{6.5.15}$$

In particular we have

$$\pi_{h,e} = \rho'(k^{-1}\triangleright h). \tag{6.5.16}$$

The equation (6.5.15) means that $\mathrm{Bimod}_{\mathcal{C}(G)}(A)$ is equivalent to $\mathcal{C}(G',\alpha')$ where $G' = \widehat{H}\rtimes K$ for some yet to be determined α'.

The associator can be computed from the graph in Figure 6.4.10. Let the objects p,q,r be $(k,\rho),(k',\rho'),(k'',\rho'')$. It suffices to restrict each object to the sub-object (k,e) in $\mathcal{C}(G)$. Without loss of generality, we can assume $\pi_{e,e}=1$ because factors of $\pi_{e,e}$ are canceled by factors of ${}^\circ\pi_{e,e}$. Then the only contribution comes from what is denoted as ${}^\circ\pi_{p\otimes_A q,r}$ in Figure 6.4.10 and we find the anomaly α' given in (6.5.8).

Examples: As an example, consider $G=\mathbb{Z}_{2n}$ generated by x with $x^{2n}=1$, and gauge the $\mathbb{Z}_2$ subgroup generated by x^n. When n is odd, $G=\mathbb{Z}_2\times\mathbb{Z}_n$, and the dual symmetry is clearly just $G'=\widehat{\mathbb{Z}}_2\times\mathbb{Z}_n$ without any anomaly, since the part $\mathbb{Z}_n$ does not matter. When n is even, G is a nontrivial extension $0\to\mathbb{Z}_2\to G=\mathbb{Z}_{2n}\to\mathbb{Z}_n\to 0$, corresponding to a nonzero $\kappa\in H^2(\mathbb{Z}_n,\mathbb{Z}_2)$. This means that the dual is $G'=\widehat{\mathbb{Z}}_2\times\mathbb{Z}_n$, with a nontrivial anomaly α_κ as given above.

As another example, consider $G=D_{2n}$, the dihedral group of $2n$ elements, generated by two elements r,s such that $r^n=s^2=1$, $srs^{-1}=r^{-1}$. In particular, let $n=2m$. Then $x:=r^m$ generates the center $\mathbb{Z}_2=\langle x\rangle$ of D_{2n}. Let us then gauge the center. Since the extension $0\to\mathbb{Z}_2\to D_{2n}\to D_n\to 0$ is nontrivial, the dual group $G'=\widehat{\mathbb{Z}_2}\times D_n$ has a nontrivial anomaly α_{κ_D}, given in terms of a nonzero $\kappa_D\in H^2(D_n,\mathbb{Z}_2)$ describing the extension. In particular, for $D_{2n}=D_8$, the dual group $G'=\widehat{\mathbb{Z}_2}\times\mathbb{Z}_2\times\mathbb{Z}_2$ is Abelian. Dually, this means that by gauging $\widehat{\mathbb{Z}_2}$ of the Abelian group G' with an anomaly α_{κ_D} turns the symmetry into a non-Abelian group D_8.

As a final example in this subsection, consider $G=Q_8$, the quaternion group of eight elements, formed by eight quaternions ±1, $\pm i$, $\pm j$, $\pm k$. This is naturally a subgroup of $\mathrm{SU}(2)$ since quaternions of absolute value 1 form the group $\mathrm{SU}(2)$, and as such the lift to $\mathrm{SO}(3)$ of a finite subgroup of $\mathrm{SO}(3)$, this case $D_4=\mathbb{Z}_2\times\mathbb{Z}_2$. This means that we have a nontrivial extension

$$0\to\mathbb{Z}_2\to Q_8\to D_4\to 0. \tag{6.5.17}$$

This extension is again nontrivial, whose class $\kappa_Q\in H^2(\mathbb{Z}_2\times\mathbb{Z}_2,\mathbb{Z}_2)$ is different from κ_D in the

case of D_8. The dual group is then $G' = \widehat{\mathbb{Z}_2} \times \mathbb{Z}_2 \times \mathbb{Z}_2$ but with a different anomaly α_{κ_Q}.

Dually, we can say as follows. The same Abelian group, $\widehat{\mathbb{Z}_2} \times \mathbb{Z}_2 \times \mathbb{Z}_2$ with two different anomalies κ_D and κ_Q dualizes, under gauging of $\widehat{\mathbb{Z}_2}$, into two different non-Abelian groups D_8 and Q_8.

6.5.4 Integral symmetry categories of total dimension 6

Let us study the symmetry categories of total dimension 6 in detail. We already know a few such symmetry categories, $\mathcal{C}(\mathbb{Z}_2 \times \mathbb{Z}_3, \alpha)$, $\mathcal{C}(S_3, \alpha)$ and $\mathrm{Rep}(S_3)$, where S_3 is the symmetric group acing on three objects. Let us study what the gauging of their subgroups leads to. We will see that there are in fact two more integral symmetry categories of total dimension 6.

From $\mathcal{C}(\mathbb{Z}_2 \times \mathbb{Z}_3, \alpha)$: Here the anomaly is determined by $\alpha \in H^3(\mathbb{Z}_2 \times \mathbb{Z}_3, \mathrm{U}(1)) = \mathbb{Z}_2 \times \mathbb{Z}_3$.

- $\mathbb{Z}_1$ is always gaugeable,
- $\mathbb{Z}_2$ is gaugeable only when α is from $\mathbb{Z}_3$ and then the dual is itself,
- $\mathbb{Z}_3$ is gaugeable only when α is from $\mathbb{Z}_2$ and then the dual is itself,
- $\mathbb{Z}_6$ is gaugeable only when α is trivial.

So there is nothing particularly interesting going on here.

From $\mathrm{Rep}(S_3)$: Any possible gauging of $\mathrm{Rep}(S_3)$ can always be done by first gauging $\mathrm{Rep}(S_3)$ back to $\mathcal{C}(S_3)$ and then gauge one of its subgroup. Therefore we do not have to study it separately.

From $\mathcal{C}(S_3, \alpha)$: Here the anomaly is determined by $\alpha \in H^3(S_3, \mathrm{U}(1)) = \mathbb{Z}_2 \times \mathbb{Z}_3$. Let us denote by a and b the generators of $\mathbb{Z}_2$ and of $\mathbb{Z}_3$, respectively.

- $\mathbb{Z}_1$ is always gaugeable and the dual is itself.
- The subgroup $\mathbb{Z}_2$ is gaugeable only when $\alpha = b^i$ with $i = 0, 1, 2$. The dual is *not* of the form $\mathcal{C}(G', \alpha')$ because this subgroup is not normal. When $\alpha = 0$ the dual turns out to be $\mathrm{Rep}(S_3)$. When $i = 1, 2$, the duals *cannot* be $\mathrm{Rep}(S_3)$, since if so, a further gauging will produce $\mathcal{C}(S_3, b^0)$ from $\mathcal{C}(S_3, b^{1,2})$. But this is impossible, since these two symmetry categories have different number of possible gaugings.
- The normal subgroup $\mathbb{Z}_3$ is gaugeable only when $\alpha = a^{0,1}$. Gauging it leads back to itself, with the same anomaly.
- S_3 is gaugeable only when α is trivial. The dual is $\mathrm{Rep}(S_3)$.

From the analysis above, we find that symmetry categories $\mathcal{C}(S_3, b^{1,2}; \mathbb{Z}_2, 0)$ obtained by gauging the $\mathbb{Z}_2$ subgroup of S_3 with a nontrivial anomaly $\alpha = b^{1,2}$ is neither of the form $\mathcal{C}(G, \alpha)$ nor of the form $\mathrm{Rep}(S_3)$. It turns out that they have the same fusion rule as $\mathrm{Rep}(S_3)$, namely there are simple objects $1, x$ of dimension 1 and a of dimension 2, such that $x^2 = 1$, $ax = a$, $a^2 = 1 + x + a$. These are the smallest integral symmetry categories which is neither $\mathcal{C}(G, \alpha)$ nor $\mathrm{Rep}(G)$.

It is known that the symmetry categories we listed so far exhaust all possible integral symmetry categories of dimension 6. This was shown in [190].

6.5.5 Integral symmetry categories of total dimension 8

Let us next have a look at symmetry categories of dimension 8. There are five finite groups G of order 8, namely the three Abelian ones $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_4$, $(\mathbb{Z}_2)^3$ and two non-Abelian ones D_8 and Q_8. Correspondingly, we already see that there are symmetry categories $\mathcal{C}(G, \alpha)$ constructed from these group, where the possible anomalies are given as follows:

G	$\mathbb{Z}_8$	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\mathbb{Z}_2^3$	D_8	Q_8
$H^3(G, \mathrm{U}(1))$	$\mathbb{Z}_8$	$\mathbb{Z}_2^2 \times \mathbb{Z}_4$	$\mathbb{Z}_2^7$	$\mathbb{Z}_2^2 \times \mathbb{Z}_4$	$\mathbb{Z}_8$

(6.5.18)

We also know two other symmetry categories of total dimension 8, namely the representation categories $\mathrm{Rep}(D_8)$ and $\mathrm{Rep}(Q_8)$.

These two representation categories have the same fusion rules: there are four dimension-1 simple objects 1, a, b, ab forming an Abelian group $\mathcal{A} = \mathbb{Z}_2 \times \mathbb{Z}_2$, and one dimension-2 simple object m, such that the fusion rule is commutative, $a \otimes m = b \otimes m = m$, and

$$m \otimes m = 1 \oplus a \oplus b \oplus ab. \tag{6.5.19}$$

There are in fact two more symmetry categories, known as KP and TY with this fusion rule [175]. For all these four cases, it is known that we can gauge the subsymmetry $\mathbb{Z}_2 = \{1, a\}$ and obtain $\mathcal{C}(D_8, \alpha)$ where $\alpha \in H^3(D_8, \mathrm{U}(1))$ is chosen depending on the four cases.

The four symmetry categories with the above fusion rule have a nice uniform description due to Tambara and Yamagami, which is applicable to a more general case based on any Abelian group $\mathcal{A}$. The gauging of $\mathbb{Z}_2 = \{1, a\}$ leading to $\mathcal{C}(D_8, \alpha)$ also has an explanation in the larger context of Tambara-Yamagami categories. We will study them in more detail in the next subection.

The only remaining choice of the fusion rule of an integral symmetry category of total dimension 8 has the following form [191]: there are four dimension-1 simple objects 1, c, c^2, c^3 forming an Abelian group $\mathcal{A} = \mathbb{Z}_4$, and one dimension-2 simple object m, such that $c \otimes m = m \otimes c = m$ and

$$m \otimes m = 1 \oplus c \oplus c^2 \oplus c^3. \tag{6.5.20}$$

The result of Tambara and Yamagami [175] implies that there are four symmetry categories with this fusion rule, distinguished by two sign choices. These categories do not have a common

name; let us temporarily call it $\mathrm{S}_{\pm\pm}$. This completes the list of the integral symmetry category of total dimension 8.

Before moving on, we have two comments. First, every integral symmetry categories we saw so far, i.e. those symmetry categories for which dimensions of objects are integers, can be obtained by gauging a non-anomalous subgroup of a possibly anomalous group. This property fails when the total dimension is larger. Indeed, some of the Tambara-Yamagami categories we discuss next are integral but cannot be obtained by gauging a non-anomalous subgroup of a possibly anomalous group.

Second, the symmetry category KP is of a historical interest, since it is the category of representations of the first non-commutative non-cocommutative Hopf algebra that appeared in the literature, constructed by Kac and Paljutkin in 1966 [192]. One way to construct a symmetry category is to pick a Hopf algebra H and take the category of its representations. When H is commutative, the symmetry category is of the form $\mathcal{C}(G)$, and when H is cocommutative, the symmetry category is of the form $\mathrm{Rep}(G)$. When we take the dual of a Hopf algebra, this naturally interchanges $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$. Therefore considering Hopf algebras is a unified framework in which $\mathcal{C}(G)$ and $\mathrm{Rep}(G)$ can be treated symmetrically. That said, to treat the symmetries of two dimensional theories and their gauging, we need to deal with symmetry categories in general and we cannot stop at the level of the Hopf algebras. The symmetry categories which are categories of representations of Hopf algebras can be characterized as symmetry categories which has Vec as a module category. But even the familiar $\mathcal{C}(G,\alpha)$ with a nontrivial α does not have Vec as a module category!

6.5.6 Tambara-Yamagami categories

To construct a Tambara-Yamagami category, we start from an Abelian group $\mathcal{A}$. The simple objects of the category are elements $a \in \mathcal{A}$ of dimension 1 together with an object m of dimension $|\mathcal{A}|$, with the commutative fusion rule

$$a \otimes m = m, \qquad m \otimes m = \bigoplus_{a\in\mathcal{A}} a. \tag{6.5.21}$$

Tambara and Yamagami showed in [175] that any symmetry category with this fusion ring is given by the choice of a symmetric nondegenerate bicharacter $\chi : \mathcal{A}\times\mathcal{A} \to \mathrm{U}(1)$ and the choice of the sign of $\tau = \pm 1/\sqrt{|\mathcal{A}|}$. The nontrivial associators are given in terms of χ and τ:

$$a_{a,m,b} = \chi(a,b), \tag{6.5.22}$$

$$a_{m,a,m} = \bigoplus_{b} \chi(a,b)\mathrm{id}_b, \tag{6.5.23}$$

$$a_{m,m,m} = \tau(\chi(a,b)^{-1})_{a,b} \in \mathrm{Hom}(\bigoplus_{a} m, \bigoplus_{b} m). \tag{6.5.24}$$

Let us denote the resulting symmetry category by $\mathrm{TY}(\mathcal{A},\chi,\tau)$.

In our case where $\mathcal{A} = \mathbb{Z}_2\times\mathbb{Z}_2$ generated by a and b, we just have two possible symmetric nondegenerate bicharacters χ up to the action of $\mathrm{SL}(2,\mathbb{Z}_2)$. Explicitly, two such choices are

specified by

$$\chi(a,a)=1, \quad \chi(b,b)=-\chi(a,b)=\pm 1. \tag{6.5.25}$$

We denote the choices by $\chi_\pm$. The choice of τ is $\tau = \pm 1/2$. Then we have the following correspondence:

	χ	τ
$\mathrm{Rep}(D_8)$	χ_+	$+1/2$
$\mathrm{Rep}(Q_8)$	χ_+	$-1/2$
KP	χ_-	$+1/2$
TY	χ_-	$-1/2$

. (6.5.26)

Here we are slightly abusing the notation such that TY alone stands for a specific symmetry category with total dimension 8, while $\mathrm{TY}(\mathcal{A},\chi,\tau)$ refers to a general construction.

Another Tambara-Yamagami category of total dimension 8 is based on $\mathcal{A}=\mathbb{Z}_4=\{1,c,c^2,c^3\}$. Any non-degenerate symmetric bicharacter is of the form

$$\chi_\pm(c^k,c^l)=(\pm i)^{kl}. \tag{6.5.27}$$

Together with the choice of the sign of τ, we have four Tambara-Yamagami categories $\mathrm{S}_{\pm\pm}$ based on $\mathbb{Z}_4$. We note that the bicharacter χ_+ is trivial on the subgroup $\mathbb{Z}_2=\{1,a=c^2\}$.

It is known that for all eight Tambara-Yamagami categories described above, we can gauge the subsymmetry $\mathbb{Z}_2=\{1,a\}$ and obtain $\mathcal{C}(D_8,\alpha)$ where $\alpha\in H^3(D_8,\mathrm{U}(1))$ is chosen depending on the four cases. This fact is a specific instance of a general theorem determining when a Tambara-Yamagami category is obtained by gauging a subgroup of a possibly non-anomalous group [193]. They showed that this occurs if and only if $\mathcal{A}$ has a Lagrangian subgroup H for χ, i.e. there is a subgroup $H\subset\mathcal{A}$ such that i) the restriction of χ on H is trivial and ii) $\mathcal{A}/H\simeq\widehat{H}$ via the pairing induced by χ. In the four cases above, H is given by $\mathbb{Z}_2=\{1,a\}$.

Let us explicitly describe below that gauging the subgroup H of the Tambara-Yamagami symmetry category $\mathcal{C}=\mathrm{TY}(\mathcal{A},\chi,\tau)$ produces a symmetry category of the form $\mathcal{C}(G,\alpha)$. The algebra object we use to gauge the system is $A=\bigoplus_{h\in H}h$.

We note that $\mathcal{A}$ fits in the extension $0\to H\to\mathcal{A}\to\widehat{H}\to 0$. We fix a specific section $s:\widehat{H}\to\mathcal{A}$ and denote by κ the two-cocycle in $C^2(\widehat{H},H)$ characterizing this extension. Then the symmetric nondegenerate bicharacter χ on $\mathcal{A}$ defines a symmetric map

$$\chi:\widehat{H}\times\widehat{H}\to\mathrm{U}(1) \tag{6.5.28}$$

satisfying the condition

$$\chi(\rho+\rho',\sigma)=\chi(\rho,\sigma)+\chi(\rho',\sigma)+\sigma(\kappa(\rho,\rho')). \tag{6.5.29}$$

Simple (A,A) bimodules turn out to be isomorphic to either of the following two types:

- $X_{\rho,\sigma}$ for $\rho,\sigma\in\widehat{H}$. As an object in $\mathcal{C}$, it is $\bigoplus_{h\in H}h\sigma$. The right action of A is trivial, and the left action of A is given by $\rho(a)\mathrm{id}: a\otimes h\sigma\to ah\sigma$.

- $Y_{\rho,\sigma}$ for $\rho, \sigma \in \widehat{H}$. As an object in $\mathcal{C}$, it is just m. The left action and the right action of A is given by $\rho(a)\mathrm{id} : a \otimes m \to m$ and $\sigma(a)\mathrm{id} : m \otimes a \to m$.

The tensor product $\otimes_A$ in the category of bimodules, together with the projections π (6.4.5) in the definition of $\otimes_A$ is then given as follows:

$$X_{\rho,\sigma} \otimes_A X_{\rho',\sigma'} = X_{\rho\rho',\sigma\sigma'} \tag{6.5.30}$$

where the projections π are trivial,

$$X_{\rho,\sigma} \otimes_A Y_{\rho',\sigma'} = Y_{\rho\rho',\sigma\sigma'} \tag{6.5.31}$$

where the projections π are given by $\rho'(a)\mathrm{id} : a\sigma \otimes m \to m$,

$$Y_{\rho,\sigma} \otimes_A X_{\rho',\sigma'} = Y_{\rho(\sigma')^{-1},\sigma(\rho')^{-1}} \tag{6.5.32}$$

where the projections π are given by $\sigma(a)\rho'(a)^{-1}\mathrm{id} : m \otimes a\sigma' \to m$, and

$$Y_{\rho,\sigma} \otimes_A Y_{\rho',\sigma'} = X_{\rho(\sigma')^{-1},\sigma(\rho')^{-1}} \tag{6.5.33}$$

where the projections π are given by

$$\pi = \bigoplus_{a\in H} \sigma'(a)^{-1}\mathrm{id}_{a\sigma(\rho')^{-1}} : m \otimes m \to \bigoplus_{x\in\mathcal{A}} x. \tag{6.5.34}$$

The equations (6.5.30),(6.5.31),(6.5.32),(6.5.33) show that the simple objects $X_{\rho,\sigma}$ and $Y_{\rho,\sigma}$ form the group

$$G = (\widehat{H} \times \widehat{H}) \rtimes \mathbb{Z}_2 \tag{6.5.35}$$

where the $\mathbb{Z}_2$ acts by

$$(\rho, \sigma) \mapsto (-\sigma, -\rho). \tag{6.5.36}$$

The anomaly α can be computed using the projections given above and the associators (6.5.22), (6.5.23) and (6.5.24) of the original category. We find

$$\alpha(X_{\rho,\sigma}, X_{\rho',\sigma'}, X_{\rho'',\sigma''}) = \rho''(\kappa(\sigma, \sigma')), \tag{6.5.37}$$
$$\alpha(X_{\rho,\sigma}, X_{\rho',\sigma'}, Y_{\rho'',\sigma''}) = \rho''(\kappa(\sigma, \sigma')), \tag{6.5.38}$$
$$\alpha(X_{\rho,\sigma}, Y_{\rho',\sigma'}, X_{\rho'',\sigma''}) = \chi(\sigma, \sigma''), \tag{6.5.39}$$
$$\alpha(Y_{\rho,\sigma}, X_{\rho',\sigma'}, X_{\rho'',\sigma''}) = (\rho'\rho''\sigma^{-1})(\kappa(\sigma', \sigma'')), \tag{6.5.40}$$
$$\alpha(X_{\rho,\sigma}, Y_{\rho',\sigma'}, Y_{\rho'',\sigma''}) = \sigma''(\kappa(\sigma, (\sigma')^{-1}\rho'')), \tag{6.5.41}$$
$$\alpha(Y_{\rho,\sigma}, X_{\rho',\sigma'}, Y_{\rho'',\sigma''}) = \chi(\sigma', \sigma(\rho')^{-1}(\rho'')^{-1}), \tag{6.5.42}$$
$$\alpha(Y_{\rho,\sigma}, Y_{\rho',\sigma'}, X_{\rho'',\sigma''}) = (\sigma'(\rho'')^{-1})(\kappa(\sigma(\rho')^{-1}, \sigma'')), \tag{6.5.43}$$
$$\alpha(Y_{\rho,\sigma}, Y_{\rho',\sigma'}, Y_{\rho'',\sigma''}) = \mathrm{sgn}(\tau)\chi(\sigma(\rho')^{-1}, \sigma'(\rho'')^{-1}). \tag{6.5.44}$$

As a check of the computation, we can directly confirm that these define a 3-cocycle on G.

In the eight cases $\mathrm{Rep}(D_8)$, $\mathrm{Rep}(Q_8)$, KP, TY and $\mathrm{S}_{\pm\pm}$ we discussed above, we always have

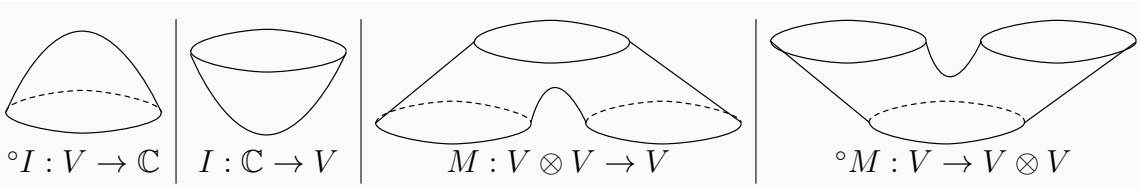


Figure 6.6.1: The basic building blocks of a 2d TFT.

$H = \mathbb{Z}_2$ and the resulting group $G = (\mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes \mathbb{Z}_2$ is D_8. To see this, regard $\mathbb{Z}_2 \times \mathbb{Z}_2$ as the group of flipping the coordinates x and y of $\mathbb{R}^2$ generated by

$$(x, y) \mapsto (-x, y), \qquad (x, y) \mapsto (x, -y) \tag{6.5.45}$$

respectively, and $\mathbb{Z}_2$ acting on $\mathbb{Z}_2 \times \mathbb{Z}_2$ to be the exchange of x and y given by

$$(x, y) \mapsto (y, x). \tag{6.5.46}$$

Dually, with a suitably chosen α on D_8 and gauging the $\mathbb{Z}_2$ subgroup flipping the x coordinate, we get the four symmetry categories given above.

6.6 2d TFT with $\mathcal{C}$ symmetry and their gauging

6.6.1 2d TFTs without symmetry

As a warm-up, let us recall the structure of 2d TFTs without any symmetry. We follow the exposition in [171] closely, see in particular their Appendix A.

We start with a vector space V of states on S^1 and one wants to define a consistent transition amplitude

$$Z_\Sigma : V^{\otimes m} \to V^{\otimes n} \tag{6.6.1}$$

corresponding to a given topological surface Σ with m incoming circles and n outgoing circles. We need four basic maps $^\circ I$, I, M, $^\circ M$ corresponding to four basic geometries given in Fig. 6.6.1.

First, we construct maps $^\circ IM : V \otimes V \to \mathbb{C}$ and $^\circ MI : \mathbb{C} \to V \otimes V$ as in Fig. 6.6.2. This inner product must be non-degenerate because it just corresponds to a cylinder geometry which pairs a state on one circle with a dual state on the other circle. Using it, we can identify V and V^*. Then, $^\circ I$ is an adjoint of I and $^\circ M$ is an adjoint of M. Therefore, to every property involving M, we can write down a corresponding property involving $^\circ M$, and similarly for statements about I and $^\circ I$. This allows us to reduce the number of independent statements we need to write down roughly by half; we do not repeat these adjoint statements below.

We consider M as giving a product on V. There is no order on the two incoming circles of a pair of pants and hence the product is commutative, see Fig. 6.6.3. We can also see that M is associative from Fig. 6.6.4 and that I is a unit of the multiplication M from Fig. 6.6.5. Also,

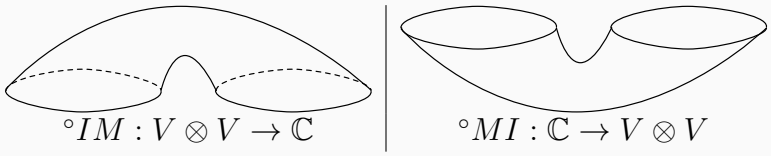


Figure 6.6.2: The pairing of V with itself.

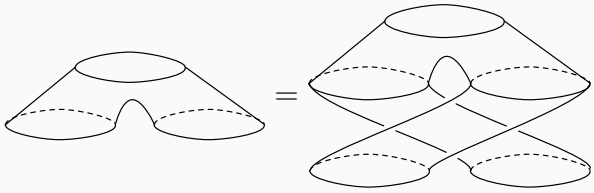


Figure 6.6.3: The product M is commutative.

by composing these inner products with the product, we see that the product is invariant under permuting three legs, see Fig. 6.6.6

After these preparations, let us associate *a* map $Z_\Sigma : V^{\otimes m} \to V^{\otimes n}$ to a surface Σ with m incoming circles and n outgoing circles. We pick a time coordinate $t : \Sigma \to [0, 1]$ such that at $t = 0$ we start with m initial circles and at $t = 1$ we finish with n final circles. As time goes from 0 to 1, the number of circles generically stay constant but can either increase or decrease by one unit at specific times $0 = t_0 < t_1 < t_2 < \cdots < t_p = 1$. Cut Σ once in each interval (t_i, t_{i+1}). This divides Σ into p pieces. The geometry of each piece contains some cylinders, which correspond to trivial transition amplitude, and exactly one non-trivial geometry out of the four non-trivial cases shown in Fig. 6.6.1. This gives us *an* expression for Z_Σ in terms of the four maps $\circ I, I, M, \circ M$.

However, one could choose a different time t' which starts with same m initial circles and ends with same n final circles. In general, this would lead to a different cutting of Σ and a different compositions of four maps $\circ I, I, M, \circ M$. We need to make sure that they agree.

We can continuously deform the time function t to obtain the time function t'. The critical

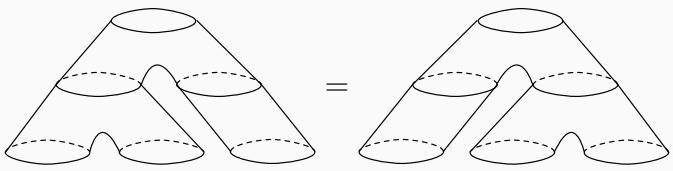


Figure 6.6.4: The product M is associative.

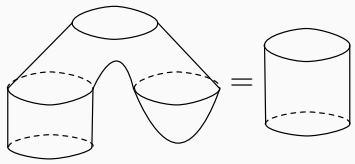


Figure 6.6.5: I is a unit of the product M.

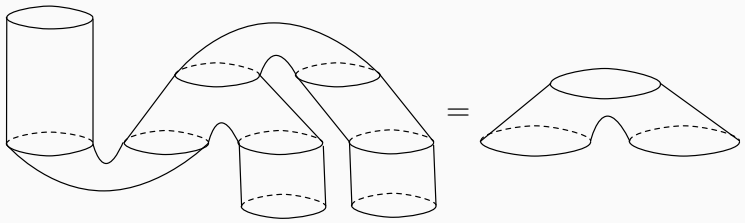


Figure 6.6.6: The product is invariant under exchanging an incoming circle and an outgoing circle.

points t_i will move under this deformation and will cross each other. It is also possible for two critical points to meet and annihilate each other or for two critical points to pop out of nowhere. We therefore need to ensure that Z_Σ remains invariant when t_i and t_{i+1} cross each other, and when two critical points are created or annihilated. For this, we just need to ensure that the two-step composition from the cut between t_{i-1} and t_i to the cut between t_{i+1} and t_{i+2} remains invariant under these processes.

All possible types of the topology changes were enumerated carefully in Appendix A of [171]. The cases are the following and their adjoints:

1. The creation or the annihilation of two critical points as shown in Fig. 6.6.5, or

2a. The exchange of two critical points as shown in Fig. 6.6.4, which we already encountered, or

2b. the situation Fig. 6.6.7 where the number of intermediate circles changes from one to three, or

2c. the situation Fig. 6.6.8 where the A-cycle and the B-cycle of a torus is exchanged. In more detail, on one side, a circle consisting of segments a, b, c, d in this order splits to two circles consisting of a, b and c, d, which are now along the A-cycle. They then merge into a circle consisting of four segments with the order b, a, d, c. On the other side, the two circles in the intermediate stage consists of segments b, c and d, a, and are along the B-cycle.

The invariance of Z_Σ under the change 1 is the unit property itself, and the invariance under the change 2a is the associativity itself. The invariance under the change 2b can be reduced to

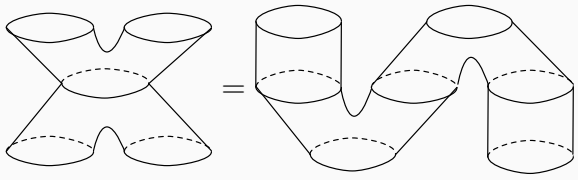


Figure 6.6.7: One possible topology change.

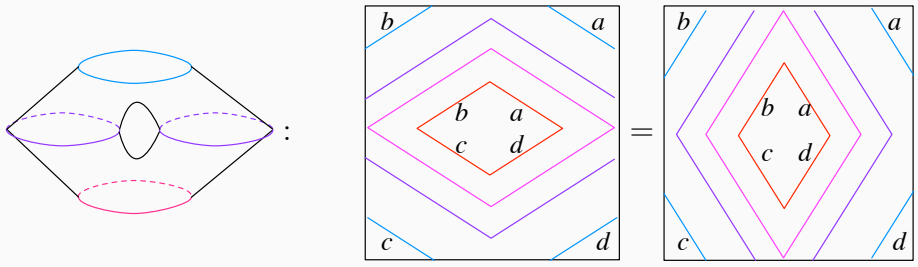


Figure 6.6.8: Another possible topology change concerns a torus with two holes. On the two figures on the right, the time flows from inside to the outside, and the parallel edges of the boundary need to be identified to form a torus. On one side, the intermediate two circles are along the A-cycle, and on the other side, they are along the B-cycle.

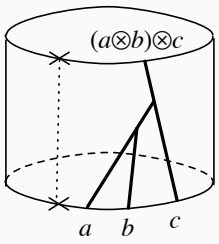


Figure 6.6.9: The Hilbert space of a circle with multiple line operators are identified with the Hilbert space of a circle with the fused line operator.

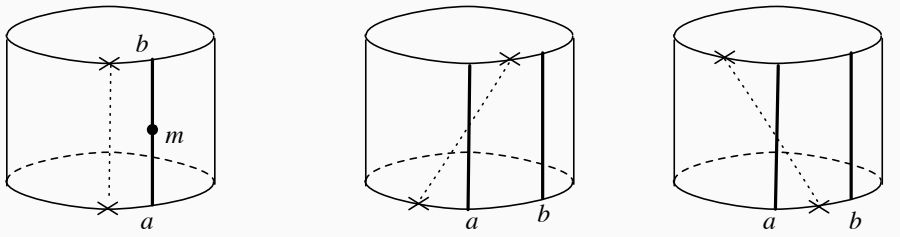


Figure 6.6.10: Basic operations on the cylinder.

associativity by using the cyclic invariance of the product, shown in Fig. 6.6.6. Finally, under the topology change 3b, the map Z_Σ is trivially invariant.

In total, we have shown that a 2d TFT with no symmetry is completely defined by a vector space V with the four maps $^\circ I, I, M, ^\circ M$ with the conditions described above. Such a vector space is known as a *commutative Frobenius algebra* V.

6.6.2 TFT with $\mathcal{C}$ symmetry on a cylinder

Let us now move on to the discussion of TFTs with symmetry given by a symmetry category $\mathcal{C}$. In this subsection we start with the simplest geometry, namely cylinders. We already discussed basics in Sec. 6.4.8. As mentioned there, we choose a *base point* along each constant-time cicle, and call its trajectory the *auxiliary line*.

Basic ingredients: We first associate the Hibert space V_a for a circle with a single insertion of a line labeled by $a \in \mathcal{C}$. We require $V_{a\oplus b} = V_a \oplus V_b$. We now associate a Hilbert space $V_{a,b,c,\dots}$ for a circle with insertions of transverse lines a, b, c, $\dots$ by fusing them in a fixed particular order, starting from the closest line on the right of the base point and then toward the right :

$$V_{a,b,c,\dots} := V_{(\cdots((a\otimes b)\otimes c)\cdots)}. \tag{6.6.2}$$

The case with three lines is shown in Fig. 6.6.9.

We have two basic operations we can perform on the cylinder, see Fig. 6.6.10. One is to insert a morphism $m : a \to b$, which defines an operator $Z(m) : V_a \to V_b$. Another is to move the base point to the right and to the left, which defines morphisms $X_{a,b} : V_{a\otimes b} \to V_{b\otimes a}$ and $Y_{b,a} : V_{b\otimes a} \to V_{a\otimes b}$.

Assignment of a map to a given network: With these basic operations, we can assign *a* map $V_{a,b,\dots} \to V_{c,d,\dots}$ for a cylinder equipped with an arbitrary network of lines and morphisms from the symmetry category $\mathcal{C}$, where an incoming circle have insertions a, b, $\dots$ and an outgoing circle have insertions c, d, $\dots$.

We choose a time function t on it, and we call any time t_i a *critical point* when either of the following happens: i) there is an insertion of a morphism on a line, ii) there is a fusion of two lines a, b into one line $a \otimes b$ or vice versa, and iii) a line crosses an auxiliary line. Note that we do not allow the auxiliary line to bend backward in time, as part of the definition.

We order $0 = t_0 < t_1 < \cdots < t_{p-1} < t_p = 1$ so that the incoming circle is at $t = 0$ and the outgoing circle is at $t = 1$. Each critical point of type i) gives a factor of $Z(m)$, that of type ii) gives a factor of $Z(\alpha)$ where α is an appropriate associator, and that of type iii) gives a factor of X or Y. Then we define the map $V_{a,b,\dots} \to V_{c,d,\dots}$ associated with this time function t to be the composition of factors corresponding to these critical points.

Consistency of the assignment: We now need to show that this assignment is consistent. There are three types of changes under which the assignment needs to be constant, namely

- the change of the time function t,
- the change of the positions of the auxiliary line, and
- the change of the network in a disk region that does not change the morphism within it.

The third point might need some clarification. In the symmetry category $\mathcal{C}$, a topologically different network can correspond to the same morphism. Then we need to ensure that if we replace a subnetwork on a cylinder accordingly, the resulting map on the Hilbert space should also be the same, see Fig. 6.6.11 This is not just a change in the time function, therefore we need to guarantee the invariance separately.

The auxiliary line might cut though the subdiagram, as also shown in Fig. 6.6.11, but this does not have to be treated separately, since we can first move the auxiliary line outside of the disk region, assuming that it is shown that the auxiliary lines can be moved.

Then this third type of change can be just taken care of by assuming that we can fuse two local operators, leading to the following constraint, see Fig. 6.6.12:

$$Z(n)Z(m) = Z(n \circ m). \tag{6.6.3}$$

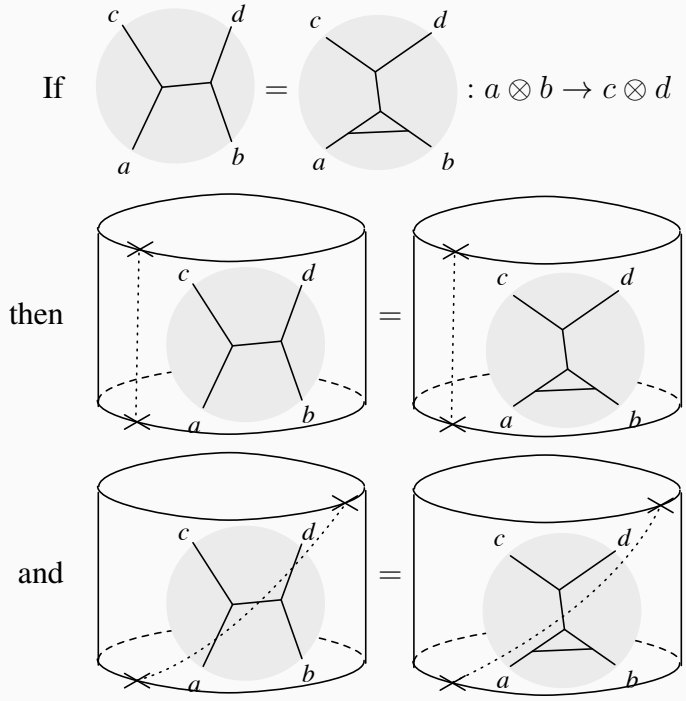


Figure 6.6.11: A local change in the network should not affect the map on the Hilbert space if the two subnetworks give the same morphism.

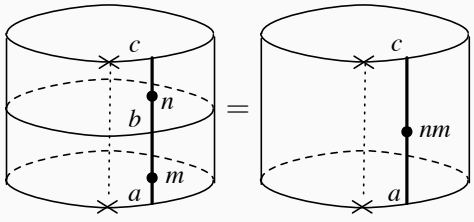


Figure 6.6.12: Two morphisms can be combined.

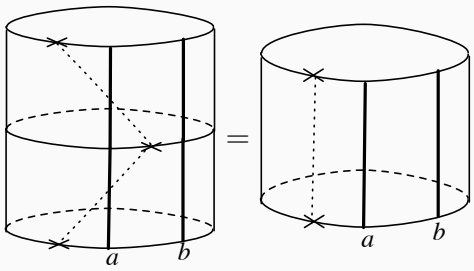


Figure 6.6.13: Moving the auxiliary line back and forth should do nothing.

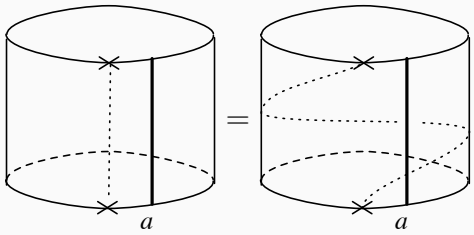


Figure 6.6.14: Winding the auxiliary line all the way around, represented by $X_{a,1}$, should not do anything.

Next, let us take care of the second type of change, where we move the auxiliary lines keeping the network and the time function fixed. First, moving the auxiliary line back and forth in succession should not do anything, so we have

$$X_{a,b} = Y_{b,a}^{-1}, \tag{6.6.4}$$

see Fig. 6.6.13. Rotating the base point all the way around should not do anything either, therefore we have

$$X_{a,1} = \mathrm{id}, \tag{6.6.5}$$

see Fig. 6.6.14.

Then we should be able to move the morphisms across the auxiliary line, leading to two relations, as illustrated in Fig. 6.6.15:

$$X_{a',b} Z(m \otimes 1) = Z(1 \otimes m) X_{a,b}, \tag{6.6.6}$$
$$X_{a,b'} Z(1 \otimes n) = Z(n \otimes 1) X_{a,b} \tag{6.6.7}$$

for $m : a \to a'$ and $n : b \to b'$. We can also fuse two lines before crossing the auxiliary line, see Fig. 6.6.16. This leads to the constraint

$$X_{b,c\otimes a} Z(\alpha_{b,c,a}) X_{a,b\otimes c} Z(\alpha_{a,b,c}) = Z({\alpha_{c,a,b}}^{-1}) X_{a\otimes b,c}. \tag{6.6.8}$$

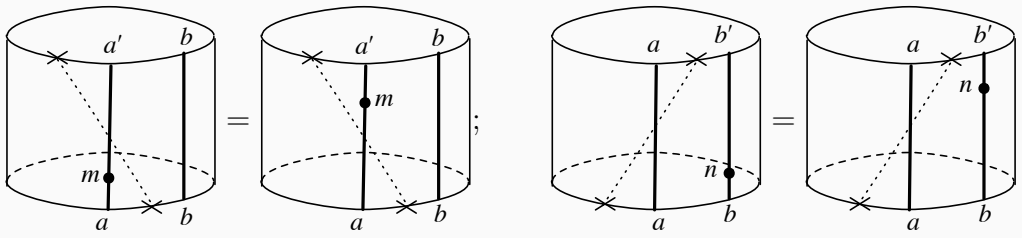


Figure 6.6.15: Morphisms can be crossed across the auxiliary line.

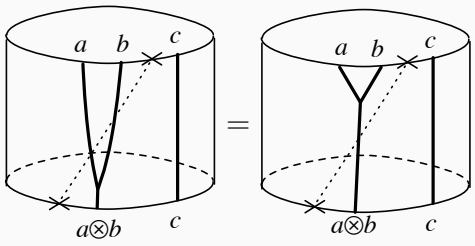


Figure 6.6.16: Crossing the fused line over the auxiliary line should be the same with crossing two lines separately.

Finally, on the cylinder, the change in the time function itself does not do much, and possible changes are already all covered. Thus, we see that to define a consistent TFT with $\mathcal{C}$ symmetry on a cylinder, we need the data of an additive functor $Z : \mathcal{C} \to \mathrm{Vec}$ with morphisms $X_{a,b} : V_{a\otimes b} \simeq V_{b\otimes a}$ satisfying (6.6.5), (6.6.6), (6.6.7) and (6.6.8).

Generalized associators on the cylinder: The relations so far guarantees that we can always move the base point and change the order of the tensoring of lines in a consistent manner. . For example, the relation (6.6.8) means that there is a single well-defined isomorphism between $V_{(a\otimes b)\otimes c}$ and $V_{(c\otimes a)\otimes b}$. We introduce a notation

$$\mathcal{A}_{(a\otimes b)\otimes c\to(c\otimes a)\otimes b} : V_{(a\otimes b)\otimes c} \to V_{(c\otimes a)\otimes b} \tag{6.6.9}$$

for it, and call it a generalized associator on the cylinder. We similarly introduce generalized associators for an arbitrary motion of the base point and an arbitrary rearrangement of parentheses. Each such generalized associator have multiple distinct-looking expressions in terms of sequences of $Z(\alpha)$, X and X^{-1}, but they give rise to the same isomorphism.

6.6.3 TFT with $\mathcal{C}$ symmetry on a general geometry

Basic data: Let us discuss now the TFT with $\mathcal{C}$ symmetry on a general geometry. The four basic geometries are given in Fig. 6.6.17. For a pair of pants, we need to join the two auxiliary

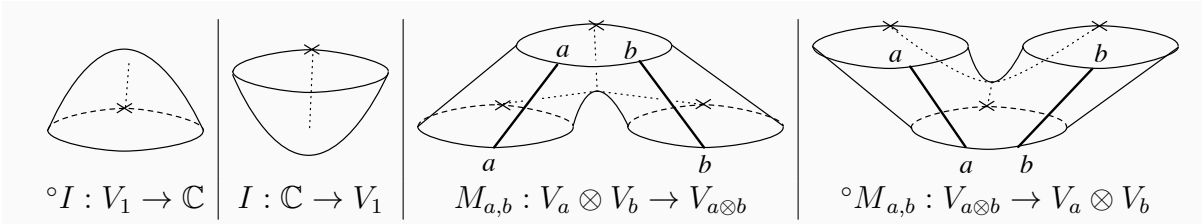


Figure 6.6.17: The basic building blocks of a 2d TFT with $\mathcal{C}$ symmetry.

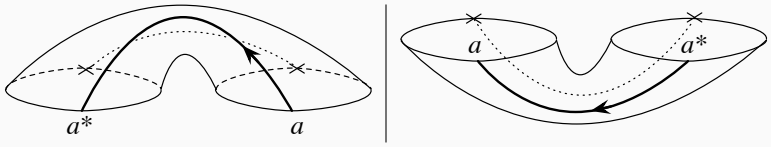


Figure 6.6.18: The pairing of V_a and V_{a^*}.

lines coming from each leg into a single auxiliary line. We take the point where this happens to coincide with the critical point where two circles join to form a single circle. In what follows, we will refer to the initial two legs of a pair of pants as the initial legs and the final leg as the product leg.

We can now associate to any geometry Σ with m initial legs and n final legs with an arbitrarily complicated network of lines and morphisms from $\mathcal{C}$ a linear map as follows. We first choose a time function $t : \Sigma \to [0, 1]$. We call a time value t_i critical when any of the following happens: i) the topology of the constant time slice change, ii) there is a morphism, or iii) a line crosses a auxiliary line. We order the critical times so that $0 = t_0 < t_1 < t_2 < \cdots < t_p = 1$. We cut Σ once in each interval (t_i, t_{i+1}), and associate to each critical time t_i one of the basic linear maps. We then compose them. We now need to guarantee that this assignment is consistent.

Basic consistency conditions: Let us first enumerate basic consistency conditions. First, we define the pairing of V_a and V_{a^*} as in Fig. 6.6.18:

$$^\circ I Z(\epsilon_a^L) M_{a^*,a} : V_{a^*} \otimes V_a \to \mathbb{C}, \tag{6.6.10}$$

$$^\circ M_{a,a^*} Z(^\circ\epsilon_a^L) I : \mathbb{C} \to V_a \otimes V_{a^*}. \tag{6.6.11}$$

Then we require that

$$\textit{this pairing is non-degenerate and can be used to identify } (V_a)^* \simeq V_{a^*}. \tag{6.6.12}$$

Under this pairing, the product $M_{a,b}$ and the coproduct $^\circ M_{b^*,a^*}$ are adjoint, etc. This again allows us to reduce the number of cases need to be mentioned below roughly by half.

Before proceeding, we note that we used ϵ_a^L, $^\circ\epsilon_a^R$ to define the pairing. We can also use $\epsilon_{a^*}^R$,

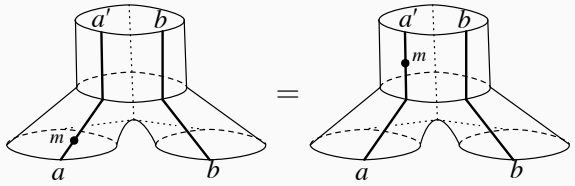


Figure 6.6.19: A morphism can be moved across the product.

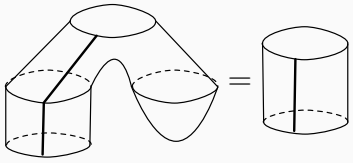


Figure 6.6.20: I is a unit of the product $M_{a,1}$.

${}^{\circ}\epsilon^{R}_{a^*}$ to define a slightly different pairing. Exactly which pairing to be used in each situation can be determined by fully assigning orientations to every line involved in the diagram. Below, we assume that every line carries an upward orientation, unless otherwise marked in the figure.

Second, a morphism can be moved across the product, see Fig. 6.6.19:

$$M_{a,b}(Z(m)\otimes \mathrm{id}_{V_b}) = Z(m\otimes \mathrm{id}_b)M_{a,b}. \tag{6.6.13}$$

Third, the map I defined by the bowl geometry gives the unit, see Fig. 6.6.20:

$$M_{a,1}(v\otimes I) = v, \qquad v\in V_a. \tag{6.6.14}$$

Fourth, it is twisted commutative:

$$X_{a,b}M_{a,b}(v\otimes w) = M_{b,a}(w\otimes v), \qquad v\in V_a, \quad w\in V_b \tag{6.6.15}$$

as illustrated in Fig. 6.6.21. Fifth, it is associative up to the associator:

$$Z(\alpha_{a,b,c})M_{a\otimes b,c}(M_{a,b}\otimes \mathrm{id}_c) = M_{a,b\otimes c}(\mathrm{id}_a\otimes M_{b,c}), \tag{6.6.16}$$

as shown in Fig. 6.6.22.

Sixth, we want to formulate that the product is symmetric under the cyclic permutation of three circles. To do this we first introduce a slightly generalized form of the product shown in Fig. 6.6.23:

$$M_{(a\otimes c^*,c\otimes b^*)\to a\otimes b^*} : V_{a\otimes c^*}\otimes V_{c\otimes b^*}\to V_{a\otimes b^*} \tag{6.6.17}$$

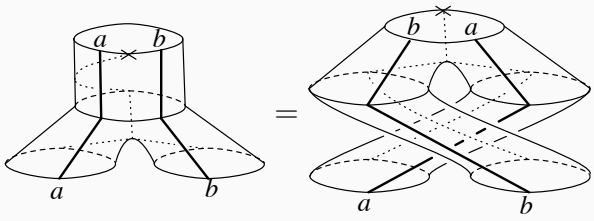


Figure 6.6.21: The product M is twisted commutative.

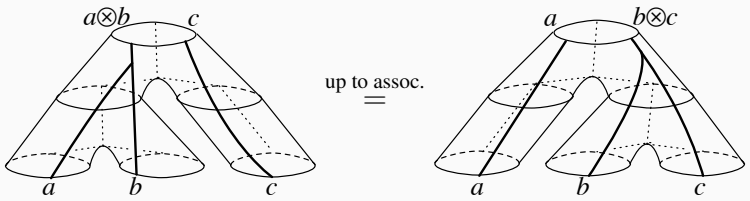


Figure 6.6.22: The product M is associative up to the associator.

given by

$$M_{(a\otimes c^*,c\otimes b^*)\to a\otimes b^*} = Z(\mathrm{id}_a \otimes \epsilon_c^L \otimes \mathrm{id}_{b^*})\mathcal{A}_{(a\otimes c^*)\otimes(c\otimes b^*)\to a\otimes(c^*\otimes c)\otimes b^*} M_{a\otimes c^*,c\otimes b^*} \tag{6.6.18}$$

where $\epsilon_c^L : c^* \otimes c \to 1$ is the evaluation morphism and $\mathcal{A}$ is the generalized associator introduced in the last subsection.

The generalized product has an alternative definition as given in Fig. 6.6.24, where the line c

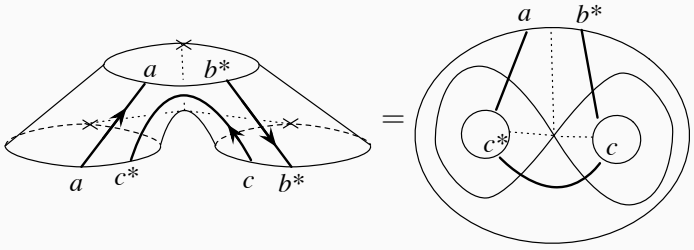


Figure 6.6.23: A slightly generalized version of the product operation. In the figure on the right, the time flows from inside to the outside.

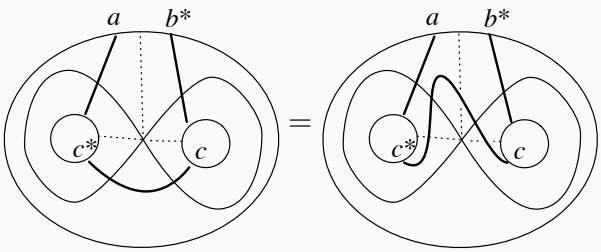


Figure 6.6.24: An alternative definition of the generalized product. The time flows from the inside to the outside.

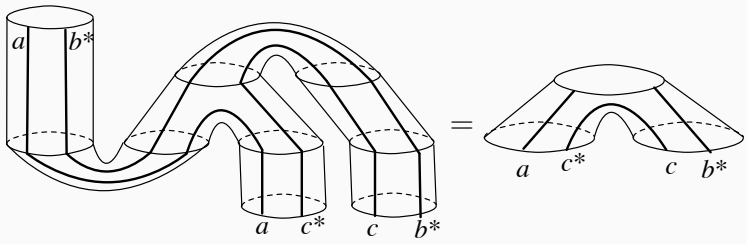


Figure 6.6.25: The product is invariant under exchanging an incoming circle and an outgoing circle.

crosses three auxiliary lines. This gives an alternative expression

$$M_{(a\otimes c^*,c\otimes b^*)\to a\otimes b^*} = Z(\mathrm{id}_{a\otimes b^*}\otimes \epsilon_c^R)\mathcal{A}_{(c^*\otimes a)\otimes(b^*\otimes c)\to(a\otimes b^*)\otimes(c\otimes c^*)}M_{c^*\otimes a,b^*\otimes c}(X_{a,c^*}\otimes X_{c,b^*}) \quad (6.6.19)$$

and we demand

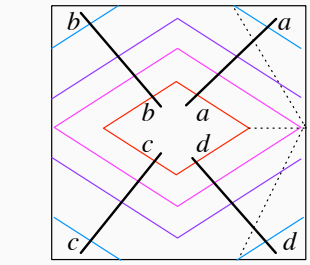
the right hand sides of the equations (6.6.18) *and* (6.6.19) *are the same.* (6.6.20)

We can now formulate the cyclic symmetry of the product:

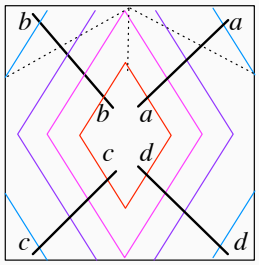
$M_{(a\otimes c^*,c\otimes b^*)\to a\otimes b^*}$ *and* $M_{(c\otimes b^*,b\otimes a^*)\to c\otimes a^*}$ *are related by the inner products,* (6.6.21)

see Fig. 6.6.25. We can in fact derive this relation for general a, b, c just from the subcase when $c = 1$ and the relations already mentioned. We keep the general case for cosmetic reasons, since it looks more symmetric.

Seventh, we need a consistency on the torus. An incoming circle consisting of four segments with lines a, b, c, d can first split into two circles with two segments a, b and c, d each and then rejoins to form a circle with four segments in the order b, a, d, c; another way this happens is that the two intermediate circles have segments b, c and d, a, see Fig. 6.6.26. They each determine



$V_{(a\otimes b)\otimes(c\otimes d)} \to V_{(b\otimes a)\otimes(d\otimes c)}$



$V_{(b\otimes c)\otimes(d\otimes a)} \to V_{(c\otimes b)\otimes(a\otimes d)}$

Figure 6.6.26: Two ways a circle splits into two and then rejoins. They should be equal up to the action of X and the associators.

maps $(M\Delta)_{a,b;c,d} : V_{(a\otimes b)\otimes(c\otimes d)} \to V_{(b\otimes a)\otimes(d\otimes c)}$ and $(M\Delta)_{b,c;d,a} : V_{(b\otimes c)\otimes(d\otimes a)} \to V_{(c\otimes b)\otimes(a\otimes d)}$ given by

$$(M\Delta)_{a,b;c,d} := M_{b\otimes a,d\otimes c}(X_{a,b}\otimes X_{c,d})\Delta_{a\otimes b,c\otimes d}, \tag{6.6.22}$$

$$(M\Delta)_{b,c;d,a} := M_{c\otimes b,a\otimes d}(X_{b,c}\otimes X_{d,a})\Delta_{b\otimes c,d\otimes a}. \tag{6.6.23}$$

We then demand that they are equal up to the generalized associators:

$$\mathcal{A}_{(b\otimes a)\otimes(d\otimes c)\to(c\otimes b)\otimes(a\otimes d)}(M\Delta)_{a,b;c,d} = (M\Delta)_{b,c;d,a}\mathcal{A}_{(a\otimes b)\otimes(c\otimes d)\to(b\otimes c)\otimes(d\otimes a)}. \tag{6.6.24}$$

Consistency in the general case: We finally finished writing down basic moves. Now we can analyze the general moves. We again have three cases:

- the change of the time function t on the surface Σ,
- the change of the positions of the auxiliary line, and
- the change of the network in a disk region that does not change the morphism.

Let us start by discussing the third case. This is in fact automatic once the first two cases are taken care of, since any disk region can be put into a cylinder under a topological change, and then the auxiliary line can be moved off away from it. Then all we have to assume is that $Z(m)$ fuses appropriately, (6.6.3).

The change in the position of the auxiliary line can happen in the following three ways:

- The auxiliary line can move within a single cylinder. This was already discussed in the last subsection.
- When a circle with line a and a circle with line b join to form a circle, the order of a, b and the base point x in the product leg can either be x, a, b or a, b, x. The invariance under this is the twisted commutativity (6.6.15).

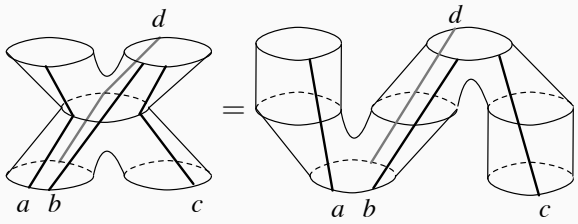


Figure 6.6.27: One possible topology change. The line d is on the back side of the figures.

- A line $c \in \mathcal{C}$ can cross the trivalent vertex of the auxiliary line. This move changes the number of the intersection of the line c with the three auxiliary lines in one of the two ways, $0 \leftrightarrow 3$ or $1 \leftrightarrow 2$. One example of the move $0 \leftrightarrow 3$ is the equality (6.6.20) of the two definitions (6.6.18) and (6.6.19) of the generalized product. The move $1 \leftrightarrow 2$ can be deduced by combining the twisted commutativity.

Finally we need to take care of the changes in the time function t. One possible change is that a morphism and a product can happen in two different orders. The invariance under this move is (6.6.13). Then there are topological changes in the cutting of the surface, which again comes in the following varieties:

1. The creation or the annihilation of two critical points does Fig. 6.6.20. The consistency under this change is the unit property (6.6.14).

2a. The exchange of two critical points does Fig. 6.6.22. The consistency under this change is the associativity (6.6.16).

2b. The number of intermediate circles changes from one to three. One example is drawn in Fig. 6.6.27. The consistency under this change can be reduced to the cyclic symmetry of the generalized product (6.6.21).

2c. How the torus is decomposed is changed as in Fig. 6.6.26, for which we assigned a basic relation (6.6.24).

Summarizing, a TFT with $\mathcal{C}$ symmetry is captured by the data $(V; Z, X; {}^\circ I, I, M, {}^\circ M)$ satisfying the various relations listed above. Namely, on the cylinder, we have (6.6.3), (6.6.5), (6.6.6), (6.6.7), (6.6.8), and on the general geometry, we have in addition (6.6.12), (6.6.13), (6.6.14), (6.6.15), (6.6.16), (6.6.20), (6.6.21), and (6.6.24), and finally, diagrams turned upside down correspond to adjoint linear maps.

6.6.4 Gauged TFT with the dual symmetry

Now we would like to discuss the definition of the TFT T/A gauged by an algebra object A in terms of the ungauged TFT T. We start from the data $(V; Z, X; {}^\circ I, I, M, {}^\circ M)$ for the original theory T.

The Hilbert space of the gauged theory was introduced in Sec. 6.4.8. See that section for some necessary background. We use the action by A on V_p, depicted in Fig. 6.4.15. This is given by

$$P := U_{p,A}({}^\circ x_R, x_L) := Z(x_L) X_{p,A} Z({}^\circ x_R) \tag{6.6.25}$$

where $x_L : A \otimes p \to p$ and ${}^\circ x_R : p \to p \otimes A$ are the morphisms defining the A-bimodule structure on p. As we already discussed, P turns out to be a projector, and we define W_p to be the projection PV_p. Now, we define the data $(W; \tilde{Z}, \tilde{X}, {}^\circ\tilde{I}, \tilde{I}, \tilde{M}, {}^\circ\tilde{M})$ for T/A in terms of corresponding data for T.

The morphism map $\tilde{Z}$: We define the new $\tilde{Z}$ to be the restriction of the old Z. We need to check that if the initial state lay in $W_p \subset V_p$ then the final state also necessarily lies in W_q, for a bimodule morphism $p \to q$. This can be checked by gluing a cylinder on top of the final state which corresponds to the action of A. The wrapped A line can then be taken across any bimodule morphism until it wraps the p line at the start of the cobordism. But then the wrapped A line has no effect and can be removed because the initial state we started with is invariant under the action of A.

The base point-change map $\tilde{X}$: We now want to define the new $\tilde{X}_{p,q} : W_{p\otimes_A q} \to W_{q\otimes_A p}$. We use π and ${}^\circ\pi$ (see Sec. 6.4.4) to define $\tilde{X}_{p,q} = Z(\pi) X_{p,q} Z({}^\circ\pi)$. This is well defined because a wrapped A line at the end of the cobordism can be moved to an A line propagating between a and b at the start of the cobordism which can be removed because of the definition of $\otimes_A$.

The unit and counit maps $\tilde{I}$ and ${}^\circ\tilde{I}$: Now the new unit morphism $\tilde{I}$ would be a map from $\mathbb{C}$ to $W_A \subset V_A$. We define it as $\tilde{I} = Z(u)I$ where $u : 1 \to A$ is the unit morphism in the definition of A. Similarly, ${}^\circ\tilde{I} = {}^\circ I Z(v)$ where $v : A \to 1$ is the co-unit morphism in the definition of A. These are well-defined as can be shown by manipulations similar to those we are now going to perform for the definition of $\tilde{M}$.

The product and coproduct maps $\tilde{M}$ and ${}^\circ\tilde{M}$: The new $\tilde{M}$ is defined analogously as $\tilde{M}_{p,q} = Z(\pi) M_{p,q}$. This $\tilde{M}$ can be shown to be well defined as a map $W_p \otimes W_q \to W_{p\otimes_A q}$ by a series of manipulations using a lot of properties of A. See Figure 6.6.28. We represent a pair of pants as a 3-punctured plane for ease of illustration. The lower punctures correspond to input legs and the upper puncture corresponds to the product leg of the pair of pants. Unlabeled lines correspond to A. To explain various manipulations, let us refer to manipulations involving $=_i$ as "step i".

- Step 1 just tells us that the left action of A on $p \otimes_A q$ is defined by the left action on p, and the right action on $p \otimes_A q$ is defined by the right action on q.
- In step 2, we introduce an A line wrapping the leg carrying q. We can do that because the input state in that leg is invariant under the action of A.
- In step 3, we first use the fact that q is an A-bimodule and then use the fact that q is a right A-module.

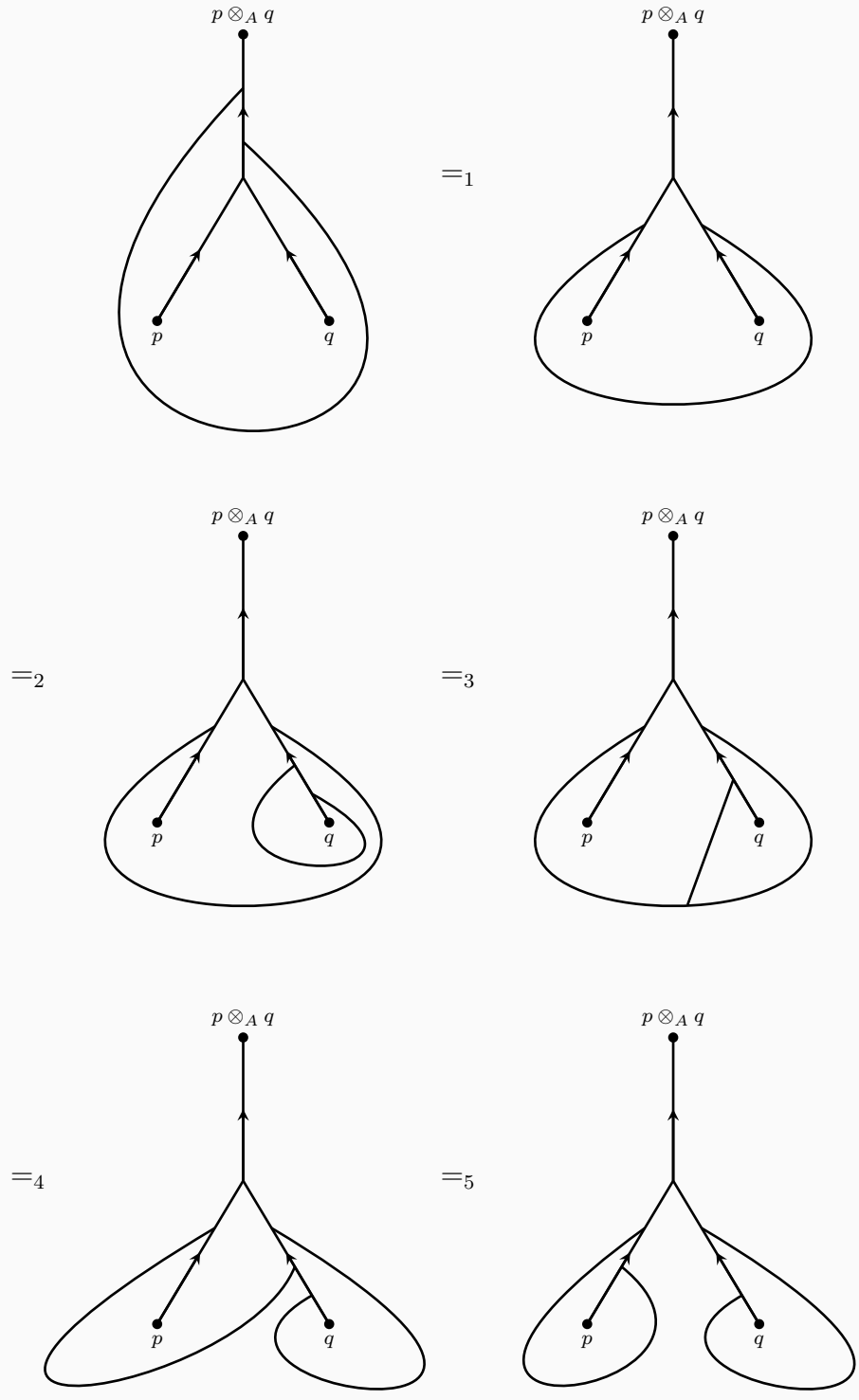


Figure 6.6.28: The new product $\tilde{M}$ and its well-definedness. For details of the manipulation, see the main text.

- In step 4, we use the fact that q is a left module for A.
- Finally, in step 5, we first use the fact that q is a bimodule, then the fact that the tensor product is $\otimes_A$, and then the fact that p is a bimodule.
- Ultimately, we can simply remove both the A lines because the input states are invariant under the action of A.

We define ${}^\circ\tilde{M}$ as the adjoint of $\tilde{M}$.

To complete the definition of T/A, we have to check that the operations defined above satisfy the various conditions that we described in the last section. Most of them just concern some trivial topological manipulations of lines and are manifestly satisfied. Some others, such as (6.6.14), can be checked using manipulations similar to the ones we have been doing in this sub-section. Yet some others, like the complicated relation (6.6.20) and (6.6.24)), require us to simplify a lot of π and ${}^\circ\pi$, but this can be done. This completes the definition of T/A.

6.7 Conclusions

In this chapter we reviewed the notion of unitary fusion categories, or symmetry categories as we prefer to call them, and how they formalize the generalized notion of finite symmetries of a two-dimensional system. We studied various explicit examples of such symmetry categories, some of which are related rather directly to finite groups and some of which are not. We then studied how a symmetry category can be gauged and be re-gauged back. We also defined 2d topological quantum field theories admitting a symmetry given by a symmetry category. Many questions remain. Here we mention just two.

The first is how to generalize the constructions discussed in this chapter to higher dimensions. In a sense, this is a merger of the generalized symmetry in two dimensions in the sense of this chapter, and of the generalized symmetry in the sense of [150]. That there should be something that combines both is clear: even in general spacetime dimension d, the 0-form symmetries can be any non-Abelian group G, possibly with an anomaly specified by $H^d(G, \mathrm{U}(1))$ in the bosonic case and by subtler objects in the fermionic case. Then the $(d-2)$-form symmetry needs to be extended at least to allow $\mathrm{Rep}(G)$. When $d = 3$, it seems that the notion of 1-form symmetries needs to be extended at least to include modular tensor categories, with an action of the 0-form symmetry group G with an anomaly. What should be the notion in $d = 4$ and higher?

The second is to actually construct two-dimensional systems T for a given symmetry category $\mathcal{C}$. For any group G without an anomaly, there is the trivial theory where the Hilbert space is always one-dimensional. How about the other cases? We can roughly classify symmetry categories $\mathcal{C}$ as follows, depending on the simplest possible theories T that have $\mathcal{C}$ as a symmetry:

1. The simplest $\mathcal{C}$-symmetric theories have one-dimensional Hilbert space. These would be $\mathcal{C}$-symmetry protected topological (SPT) phases.
2. The simplest $\mathcal{C}$-symmetric theories have finite-dimensional Hilbert space. These would be $\mathcal{C}$-symmetry enriched topological (SET) phases.

3. The simplest $\mathcal{C}$-symmetric theories have infinite-dimensional Hilbert space. Taking the low energy limit, these would be $\mathcal{C}$-symmetric conformal field theories (CFTs).

4. There is no $\mathcal{C}$-symmetric theory.

Clearly this classification forms a hierarchy, and it would be nice if we have a uniform construction that tells easily which stage of the above classification a given symmetry category $\mathcal{C}$ belongs to. There is a recent paper in this general direction [194], where a construction of 2d theory starting from any given symmetry category $\mathcal{C}$ was discussed. We hope to see more developments in the future.

Actually, there are various symmetry categories $\mathcal{C}$ constructed in the subfactor theory, e.g. what is called the Haagerup fusion category, for which no $\mathcal{C}$-symmetric theory is known. If a theory symmetric under the Haagerup fusion category could be constructed, it would be considered as a huge breakthrough.

References

[1] L. Bhardwaj, M. Del Zotto, J. J. Heckman, D. R. Morrison, T. Rudelius, and C. Vafa, "F-theory and the Classification of Little Strings," *Phys. Rev.* **D93** no. 8, (2016) 086002, arXiv:1511.05565 [hep-th].

[2] L. Bhardwaj, "Classification of 6d $\mathcal{N} = (1, 0)$ gauge theories," arXiv:1502.06594 [hep-th].

[3] L. Bhardwaj, D. Gaiotto, and A. Kapustin, "State sum constructions of spin-TFTs and string net constructions of fermionic phases of matter," arXiv:1605.01640 [cond-mat.str-el].

[4] L. Bhardwaj, "Unoriented 3d TFTs," *JHEP* **05** (2017) 048, arXiv:1611.02728 [hep-th].

[5] L. Bhardwaj and Y. Tachikawa, "On finite symmetries and their gauging in two dimensions," arXiv:1704.02330 [hep-th].

[6] P. S. Aspinwall, "Point - like Instantons and the $Spin(32)/\mathbb{Z}_2$ Heteroti String," *Nucl. Phys.* **B496** (1997) 149–176, arXiv:hep-th/9612108.

[7] M. Bianchi and A. Sagnotti, "Twist Symmetry and Open String Wilson Lines," *Nucl. Phys.* **B361** (1991) 519–538.

[8] E. G. Gimon and J. Polchinski, "Consistency Conditions for Orientifolds and D Manifolds," *Phys. Rev.* **D54** (1996) 1667–1676, arXiv:hep-th/9601038 [hep-th].

[9] J. J. Heckman, D. R. Morrison, and C. Vafa, "On the Classification of 6D SCFTs and Generalized ADE Orbifolds," *JHEP* **05** (2014) 028, arXiv:1312.5746 [hep-th]. [Erratum: *JHEP* **06** (2015) 017].

[10] J. J. Heckman, D. R. Morrison, T. Rudelius, and C. Vafa, "Atomic Classification of 6D SCFTs," *Fortsch. Phys.* **63** (2015) 468–530, arXiv:1502.05405 [hep-th].

[11] L. Bhardwaj, "Classification of 6d $\mathcal{N} = (1, 0)$ gauge theories," *JHEP* **11** (2015) 002, arXiv:1502.06594 [hep-th].

[12] V. G. Turaev and O. Y. Viro, "State sum invariants of 3 manifolds and quantum 6j symbols," *Topology* **31** (1992) 865–902.

[13] C. Vafa, "Quantum Symmetries of String Vacua," *Mod. Phys. Lett.* **A4** (1989) 1615.

[14] E. Witten, "Some comments on string dynamics," in *Future perspectives in string theory. Proceedings, Conference, Strings'95, Los Angeles, USA, March 13-18, 1995*. 1995. arXiv:hep-th/9507121.

[15] P. S. Aspinwall and D. R. Morrison, "Point - like instantons on K3 orbifolds," *Nucl. Phys.* **B503** (1997) 533–564, arXiv:hep-th/9705104.

[16] N. Seiberg, "New theories in six-dimensions and matrix description of M theory on T^5 and $T^5/\mathbb{Z}_2$," *Phys. Lett.* **B408** (1997) 98–104, arXiv:hep-th/9705221.

[17] K. A. Intriligator, "New string theories in six-dimensions via branes at orbifold singularities," *Adv. Theor. Math. Phys.* **1** (1998) 271–282, arXiv:hep-th/9708117.

[18] A. Hanany and A. Zaffaroni, "Branes and six-dimensional supersymmetric theories," *Nucl. Phys.* **B529** (1998) 180–206, arXiv:hep-th/9712145.

[19] I. Brunner and A. Karch, "Branes at orbifolds versus Hanany Witten in six-dimensions," *JHEP* **03** (1998) 003, arXiv:hep-th/9712143.

[20] A. Kapustin, "On the universality class of little string theories," *Phys. Rev.* **D63** (2001) 086005, arXiv:hep-th/9912044.

[21] K. A. Intriligator, "Compactified little string theories and compact moduli spaces of vacua," *Phys. Rev.* **D61** (2000) 106005, arXiv:hep-th/9909219.

[22] O. Aharony, M. Berkooz, D. Kutasov, and N. Seiberg, "Linear dilatons, NS five-branes and holography," *JHEP* **10** (1998) 004, arXiv:hep-th/9808149.

[23] E. Witten, "Toroidal Compactification Without Vector Structure," *JHEP* **02** (1998) 006, arXiv:hep-th/9712028.

[24] J. de Boer, R. Dijkgraaf, K. Hori, A. Keurentjes, J. Morgan, D. R. Morrison, and S. Sethi, "Triples, Fluxes, and Strings," *Adv. Theor. Math. Phys.* **4** (2002) 995–1186, arXiv:hep-th/0103170.

[25] Y. Tachikawa, "Frozen," arXiv:1508.06679 [hep-th].

[26] N. Seiberg, "Nontrivial fixed points of the renormalization group in six-dimensions," *Phys. Lett.* **B390** (1997) 169–171, arXiv:hep-th/9609161.

[27] D. Gaiotto and A. Tomasiello, "Holography for (1,0) theories in six dimensions," *JHEP* **12** (2014) 003, arXiv:1404.0711 [hep-th].

[28] M. Del Zotto, J. J. Heckman, A. Tomasiello, and C. Vafa, "6d Conformal Matter," *JHEP* **02** (2015) 054, arXiv:1407.6359 [hep-th].

[29] J. J. Heckman, "More on the Matter of 6D SCFTs," *Phys. Lett.* **B747** (2015) 73–75, arXiv:1408.0006 [hep-th].

[30] M. Del Zotto, J. J. Heckman, D. R. Morrison, and D. S. Park, "6D SCFTs and Gravity," *JHEP* **06** (2015) 158, arXiv:1412.6526 [hep-th].

[31] S. Hohenegger, A. Iqbal, and S.-J. Rey, "Instanton-Monopole Correspondence from M-Branes on $\mathbb{S}^1$ and Little String Theory," arXiv:1511.02787 [hep-th].

[32] N. Seiberg and W. Taylor, "Charge Lattices and Consistency of 6D Supergravity," *JHEP* **06** (2011) 001, arXiv:1103.0019 [hep-th].

[33] È. B. Vinberg, "Discrete linear groups that are generated by reflections," *Izv. Akad. Nauk SSSR Ser. Mat.* **35** (1971) 1072–1112.

[34] M. Del Zotto, J. J. Heckman, D. S. Park, and T. Rudelius, "On the Defect Group of a 6D SCFT," arXiv:1503.04806 [hep-th].

[35] D. R. Morrison and W. Taylor, "Classifying bases for 6D F-theory models," *Central Eur. J. Phys.* **10** (2012) 1072–1088, arXiv:1201.1943 [hep-th].

[36] D. R. Morrison and W. Taylor, "Toric bases for 6D F-theory models," *Fortsch. Phys.* **60** (2012) 1187–1216, arXiv:1204.0283 [hep-th].

[37] C. Beasley, J. J. Heckman, and C. Vafa, "GUTs and exceptional branes in F-theory - II: Experimental predictions," *JHEP* **01** (2009) 059, arXiv:0806.0102 [hep-th].

[38] C. Córdova, "Decoupling gravity in F-theory," *Adv. Theor. Math. Phys.* **15** (2011) 689–740, arXiv:0910.2955 [hep-th].

[39] J. J. Heckman and H. Verlinde, "Evidence for F(uzz) Theory," *JHEP* **01** (2011) 044, arXiv:1005.3033 [hep-th].

[40] M. Bertolini, P. R. Merkx, and D. R. Morrison, "On the global symmetries of 6D superconformal field theories," arXiv:1510.08056 [hep-th].

[41] S. Mori, "Threefolds whose canonical bundles are not numerically effective," *Ann. of Math. (2)* **116** no. 1, (1982) 133–176.

[42] Y. Kawamata, "The cone of curves of algebraic varieties," *Ann. of Math. (2)* **119** (1984) 603–633.

[43] S. Donaldson and S. Sun, "Gromov-Hausdorff limits of Kähler manifolds and algebraic geometry," *Acta Math.* **213** no. 1, (2014) 63–106, arXiv:1206.2609 [math.DG].

[44] S. Donaldson and S. Sun, "Gromov-Hausdorff limits of Kähler manifolds and algebraic geometry, II," arXiv:1507.05082 [math.DG].

[45] M. Gross and P. M. H. Wilson, "Large complex structure limits of $K3$ surfaces," *J. Differential Geom.* **55** no. 3, (2000) 475–546, arXiv:math.DG/0008018.

[46] C. Schoen, "On fiber products of rational elliptic surfaces with section," *Math. Z.* **197** (1988) 177–199.

[47] E. Witten, "New "gauge" theories in six dimensions," *JHEP* **01** (1998) 001, arXiv:hep-th/9710065.

[48] N. Nakayama, "On Weierstrass models," in *Algebraic geometry and commutative algebra, Vol. II*, pp. 405–431. Kinokuniya, 1988.

[49] M. R. Douglas and G. W. Moore, "D-branes, quivers, and ALE instantons," arXiv:hep-th/9603167.

[50] D. R. Morrison, D. S. Park, and W. Taylor, "to appear,".

[51] A. Grassi and D. R. Morrison, "Anomalies and the Euler characteristic of elliptic Calabi-Yau threefolds," *Commun. Num. Theor. Phys.* **6** (2012) 51–127, arXiv:1109.0042 [hep-th].

[52] V. Sadov, "Generalized Green-Schwarz mechanism in F theory," *Phys. Lett.* **B388** (1996) 45–50, arXiv:hep-th/9606008 [hep-th].

[53] A. Grassi and D. R. Morrison, "Group representations and the Euler characteristic of elliptically fibered Calabi–Yau threefolds," *J. Algebraic Geom.* **12** (2003) 321–356, arXiv:math.AG/0005196.

[54] D. R. Morrison and W. Taylor, "Matter and singularities," *JHEP* **01** (2012) 022, arXiv:1106.3563 [hep-th].

[55] N. Nekrasov and V. Pestun, "Seiberg-Witten geometry of four dimensional N=2 quiver gauge theories," arXiv:1211.2240 [hep-th].

[56] V. Braun and D. R. Morrison, "F-theory on Genus-One Fibrations," *JHEP* **08** (2014) 132, arXiv:1401.7844 [hep-th].

[57] D. R. Morrison and W. Taylor, "Sections, multisections, and $U(1)$ fields in F-theory," *J. Singularities, to appear* , arXiv:1404.1527 [hep-th].

[58] M. Cvetič, R. Donagi, D. Klevers, H. Piragua, and M. Poretschkin, "F-theory vacua with $\mathbb{Z}_3$ gauge symmetry," *Nucl. Phys.* **B898** (2015) 736–750, arXiv:1502.06953 [hep-th].

[59] I. Dolgachev and M. Gross, "Elliptic threefolds. I. Ogg-Shafarevich theory," *J. Algebraic Geom.* **3** no. 1, (1994) 39–80, arXiv:alg-geom/9210009.

[60] M. Gross, "Elliptic three-folds. II. Multiple fibres," *Trans. Amer. Math. Soc.* **349** no. 9, (1997) 3409–3468.

[61] A. Grassi, "Log contractions and equidimensional models of elliptic threefolds," *J. Algebraic Geom.* **4** no. 2, (1995) 255–276, arXiv:alg-geom/9305003.

[62] M. Gross, "A finiteness theorem for elliptic Calabi-Yau threefolds," *Duke Math. J.* **74** no. 2, (1994) 271–299, arXiv:alg-geom/9305002.

[63] K. Ohmori, H. Shimizu, Y. Tachikawa, and K. Yonekura, "6d $\mathcal{N}=(1,0)$ theories on S^1/T^2 and class S theories: part II," arXiv:1508.00915 [hep-th].

[64] K. Ohmori, H. Shimizu, Y. Tachikawa, and K. Yonekura, "Anomaly polynomial of general 6d SCFTs," *PTEP* **2014** no. 10, (2014) 103B07, arXiv:1408.5572 [hep-th].

[65] P. S. Aspinwall, S. H. Katz, and D. R. Morrison, "Lie groups, Calabi-Yau threefolds, and F theory," *Adv. Theor. Math. Phys.* **4** (2000) 95–126, arXiv:hep-th/0002012.

[66] J. J. Heckman, D. R. Morrison, T. Rudelius, and C. Vafa, "Geometry of 6D RG Flows," *JHEP* **09** (2015) 052, arXiv:1505.00009 [hep-th].

[67] J. J. Heckman and T. Rudelius, "Evidence for C-theorems in 6D SCFTs," *JHEP* **09** (2015) 218, arXiv:1506.06753 [hep-th].

[68] C. Cordova, T. T. Dumitrescu, and K. Intriligator, "Anomalies, Renormalization Group Flows, and the a-Theorem in Six-Dimensional (1,0) Theories," arXiv:1506.03807 [hep-th].

[69] B. Haghighat, G. Lockhart, and C. Vafa, "Fusing E-strings to heterotic strings: $E + E \to H$," *Phys. Rev.* **D90** no. 12, (2014) 126012, arXiv:1406.0850 [hep-th].

[70] C. Vafa, "Evidence for F theory," *Nucl. Phys.* **B469** (1996) 403–418, arXiv:hep-th/9602022.

[71] D. R. Morrison and C. Vafa, "Compactifications of F theory on Calabi-Yau threefolds. 1," *Nucl. Phys.* **B473** (1996) 74–92, arXiv:hep-th/9602114.

[72] D. R. Morrison and C. Vafa, "Compactifications of F theory on Calabi-Yau threefolds. 2.," *Nucl. Phys.* **B476** (1996) 437–469, arXiv:hep-th/9603161.

[73] I. Brunner and A. Karch, "Branes and Six-Dimensional Fixed Points," *Phys. Lett.* **B409** (1997) 109–116, arXiv:hep-th/9705022 [hep-th].

[74] A. Hanany and A. Zaffaroni, "Issues on Orientifolds: on the Brane Construction of Gauge Theories with SO(2n) Global Symmetry," *JHEP* **07** (1999) 009, arXiv:hep-th/9903242.

[75] U. H. Danielsson, G. Ferretti, J. Kalkkinen, and P. Stjernberg, "Notes on Supersymmetric Gauge Theories in Five-Dimensions and Six-Dimensions," *Phys. Lett.* **B405** (1997) 265–270, arXiv:hep-th/9703098 [hep-th].

[76] I. García-Etxebarria and D. Regalado, "Exceptional $\mathcal{N} = 3$ Theories," *JHEP* **12** (2017) 042, arXiv:1611.05769 [hep-th].

[77] K. Ohmori, H. Shimizu, Y. Tachikawa, and K. Yonekura, "6D $\mathcal{N} = (1,0)$ Theories on T^2 and Class S Theories: Part I," *JHEP* **07** (2015) 014, arXiv:1503.06217 [hep-th].

[78] N. Mekareeya, K. Ohmori, H. Shimizu, and A. Tomasiello, "Small Instanton Transitions for M5 Fractions," *JHEP* **10** (2017) 055, arXiv:1707.05785 [hep-th].

[79] M. B. Green, J. H. Schwarz, and P. C. West, "Anomaly Free Chiral Theories in Six-Dimensions," *Nucl. Phys.* **B254** (1985) 327–348.

[80] A. Sagnotti, "A Note on the Green-Schwarz Mechanism in Open String Theories," *Phys. Lett.* **B294** (1992) 196–203, arXiv:hep-th/9210127 [hep-th].

[81] S. Cecotti, C. Cordova, J. J. Heckman, and C. Vafa, "T-Branes and Monodromy," *JHEP* **07** (2011) 030, arXiv:1010.5780 [hep-th].

[82] A. Dabholkar, "Lectures on Orientifolds and Duality," in *High-Energy Physics and Cosmology. Proceedings, Summer School, Trieste, Italy, June 2-July 4, 1997*. 1997. arXiv:hep-th/9804208 [hep-th]. http://alice.cern.ch/format/showfull?sysnb=0278186.

[83] J. Distler, D. S. Freed, and G. W. Moore, "Orientifold précis," in *Mathematical foundations of quantum field theory and perturbative string theory*, vol. 83 of *Proc. Sympos. Pure Math.*, pp. 159–172. Amer. Math. Soc., Providence, RI, 2011. arXiv:0906.0795 [hep-th].

[84] D. Gao and K. Hori, "On the Structure of the Chan-Paton Factors for D-Branes in Type II Orientifolds," arXiv:1004.3972 [hep-th].

[85] J. Dai, R. G. Leigh, and J. Polchinski, "New Connections Between String Theories," *Mod. Phys. Lett.* **A4** (1989) 2073–2083.

[86] A. Sagnotti, "Open Strings and Their Symmetry Groups," in *Nato Advanced Summer Institute on Nonperturbative Quantum Field Theory (Cargese Summer Institute) Cargese, France, July 16-30, 1987*, pp. 0521–528. 1987. arXiv:hep-th/0208020 [hep-th].

[87] P. Horava, "Strings on World Sheet Orbifolds," *Nucl. Phys.* **B327** (1989) 461–484.

[88] C. Angelantonj and A. Sagnotti, "Open Strings," *Phys. Rept.* **371** (2002) 1–150, arXiv:hep-th/0204089 [hep-th]. [Erratum: Phys. Rept.376,no.6,407(2003)].

[89] A. Hanany and J. Troost, "Orientifold Planes, Affine Algebras and Magnetic Monopoles," *JHEP* **08** (2001) 021, arXiv:hep-th/0107153.

[90] A. Sen, "F Theory and Orientifolds," *Nucl. Phys.* **B475** (1996) 562–578, arXiv:hep-th/9605150 [hep-th].

[91] K. Landsteiner and E. Lopez, "New curves from branes," *Nucl. Phys.* **B516** (1998) 273–296, hep-th/9708118.

[92] L. B. Anderson, J. J. Heckman, and S. Katz, "T-Branes and Geometry," *JHEP* **05** (2014) 080, arXiv:1310.1931 [hep-th].

[93] A. Collinucci and R. Savelli, "T-branes as branes within branes," *JHEP* **09** (2015) 161, arXiv:1410.4178 [hep-th].

[94] L. B. Anderson, J. J. Heckman, S. Katz, and L. Schaposnik, "T-Branes at the Limits of Geometry," *JHEP* **10** (2017) 058, arXiv:1702.06137 [hep-th].

[95] M. Bershadsky, K. A. Intriligator, S. Kachru, D. R. Morrison, V. Sadov, and C. Vafa, "Geometric singularities and enhanced gauge symmetries," *Nucl. Phys.* **B481** (1996) 215–252, arXiv:hep-th/9605200 [hep-th].

[96] J. Polchinski, "Tensors from K3 orientifolds," *Phys. Rev.* **D55** (1997) 6423–6428, arXiv:hep-th/9606165 [hep-th].

[97] M. Berkooz, R. G. Leigh, J. Polchinski, J. H. Schwarz, N. Seiberg, and E. Witten, "Anomalies, dualities, and topology of $D = 6$ $\mathcal{N} = 1$ superstring vacua," *Nucl. Phys.* **B475** (1996) 115–148, arXiv:hep-th/9605184 [hep-th].

[98] E. Witten, "Branes, Instantons, And Taub-NUT Spaces," *JHEP* **06** (2009) 067, arXiv:0902.0948 [hep-th].

[99] N. J. Evans, C. V. Johnson, and A. D. Shapere, "Orientifolds, branes, and duality of 4D gauge theories," *Nucl. Phys.* **B505** (1997) 251–271, arXiv:hep-th/9703210.

[100] S. H. Katz and C. Vafa, "Matter from Geometry," *Nucl. Phys.* **B497** (1997) 146–154, arXiv:hep-th/9606086.

[101] M. B. Green, J. A. Harvey, and G. W. Moore, "I-Brane Inflow and Anomalous Couplings on D-Branes," *Class. Quant. Grav.* **14** (1997) 47–52, arXiv:hep-th/9605033.

[102] J. Erler, "Anomaly Cancellation in Six-Dimensions," *J. Math. Phys.* **35** (1994) 1819–1833, arXiv:hep-th/9304104 [hep-th].

[103] J. H. Schwarz, "Anomaly - Free Supersymmetric Models in Six-Dimensions," *Phys.Lett.* **B371** (1996) 223–230, arXiv:hep-th/9512053 [hep-th].

[104] N. Seiberg and E. Witten, "Comments on string dynamics in six-dimensions," *Nucl. Phys.* **B471** (1996) 121–134, arXiv:hep-th/9603003.

[105] H.-C. Kim, S. Kim, and J. Park, "6D Strings from New Chiral Gauge Theories," arXiv:1608.03919 [hep-th].

[106] K. A. Intriligator, D. R. Morrison, and N. Seiberg, "Five-dimensional supersymmetric gauge theories and degenerations of Calabi–Yau spaces," *Nucl.Phys.* **B497** (1997) 56–100, arXiv:hep-th/9702198 [hep-th].

[107] F. Apruzzi and M. Fazzi, "AdS$_7$/CFT$_6$ with orientifolds," *JHEP* **01** (2018) 124, arXiv:1712.03235 [hep-th].

[108] P. S. Aspinwall and M. Gross, "The SO(32) Heterotic String on a K3 Surface," *Phys. Lett.* **B387** (1996) 735–742, arXiv:hep-th/9605131.

[109] E. Witten, "Small instantons in string theory," *Nucl. Phys.* **B460** (1996) 541–559, arXiv:hep-th/9511030.

[110] A. Sen, "F Theory and the Gimon-Polchinski Orientifold," *Nucl. Phys.* **B498** (1997) 135–155, arXiv:hep-th/9702061 [hep-th].

[111] A. Sen, "Orientifold Limit of F Theory Vacua," *Phys. Rev.* **D55** (1997) R7345–R7349, arXiv:hep-th/9702165 [hep-th].

[112] A. Sen and S. Sethi, "The Mirror Transform of Type I Vacua in Six-Dimensions," *Nucl. Phys.* **B499** (1997) 45–54, arXiv:hep-th/9703157 [hep-th].

[113] J. D. Blum and A. Zaffaroni, "An Orientifold from F Theory," *Phys. Lett.* **B387** (1996) 71–74, arXiv:hep-th/9607019 [hep-th].

[114] A. Dabholkar and J. Park, "A Note on Orientifolds and F - Theory," *Phys. Lett.* **B394** (1997) 302–306, arXiv:hep-th/9607041 [hep-th].

[115] R. Dijkgraaf and E. Witten, "Topological gauge theories and group cohomology," *Comm. Math. Phys.* **129** no. 2, (1990) 393–429.

[116] L. Fidkowski and A. Kitaev, "Topological phases of fermions in one dimension," 1008.4138. http://arxiv.org/abs/1008.4138.

[117] X. Chen, Z.-C. Gu, and X.-G. Wen, "Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order," *Physical review b* **82** no. 15, (2010) 155138.

[118] D. Gaiotto and A. Kapustin, "Spin tqfts and fermionic phases of matter," 1505.05856. http://arxiv.org/abs/1505.05856.

[119] A. Kitaev, "Anyons in an exactly solved model and beyond," cond-mat/0506438. http://arxiv.org/abs/cond-mat/0506438.

[120] G. Moore and N. Seiberg, "Classical and quantum conformal field theory," *Comm. Math. Phys.* **123** no. 2, (1989) 177–254.

[121] A. Beliakova, C. Blanchet, and E. Contreras, "Spin modular categories," 1411.4232. http://arxiv.org/abs/1411.4232.

[122] Z.-C. Gu, Z. Wang, and X.-G. Wen, "Lattice model for fermionic toric code," 1309.7032. https://arxiv.org/abs/1309.7032.

[123] M. Cheng, Z. Bi, Y.-Z. You, and Z.-C. Gu, "Towards a complete classification of symmetry-protected phases for interacting fermions in two dimensions," 1501.01313. http://arxiv.org/abs/1501.01313.

[124] T. Lan, L. Kong, and X.-G. Wen, "A theory of 2+1d fermionic topological orders and fermionic/bosonic topological orders with symmetries," 1507.04673. https://arxiv.org/abs/1507.04673.

[125] 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," 1201.2648. http://arxiv.org/abs/1201.2648.

[126] M. A. Levin and X.-G. Wen, "String net condensation: A Physical mechanism for topological phases," *Phys. Rev.* **B71** (2005) 045110, arXiv:cond-mat/0404617 [cond-mat].

[127] Z.-C. Gu, Z. Wang, and X.-G. Wen, "A classification of 2d fermionic and bosonic topological orders," 1010.1517. http://arxiv.org/abs/1010.1517.

[128] A. Kapustin, R. Thorngren, A. Turzillo, and Z. Wang, "Fermionic Symmetry Protected Topological Phases and Cobordisms," *JHEP* **12** (2015) 052, arXiv:1406.7329 [cond-mat.str-el]. [JHEP12,052(2015)].

[129] V. Turaev, "Homotopy field theory in dimension 3 and crossed group-categories," math/0005291. http://arxiv.org/abs/math/0005291.

[130] V. Turaev and A. Virelizier, "On 3-dimensional homotopy quantum field theory, i," 1202.6292. http://arxiv.org/abs/1202.6292.

[131] V. Turaev and A. Virelizier, "Surgery hqft," 1303.1331. http://arxiv.org/abs/1303.1331.

[132] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, "Generalized global symmetries," 1412.5148. http://arxiv.org/abs/1412.5148.

[133] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, "Symmetry protected topological orders and the group cohomology of their symmetry group," 1106.4772. http://arxiv.org/abs/1106.4772.

[134] L. Chang, M. Cheng, S. X. Cui, Y. Hu, W. Jin, R. Movassagh, P. Naaijkens, Z. Wang, and A. Young, "On enriching the levin-wen model with symmetry," 1412.6589. http://arxiv.org/abs/1412.6589.

[135] N. E. Steenrod, "Products of cocycles and extensions of mappings," *Annals of Mathematics* (1947) 290–320.

[136] A. Kapustin and N. Seiberg, "Coupling a qft to a tqft and duality," 1401.0740. http://arxiv.org/abs/1401.0740.

[137] J. Brundan and A. P. Ellis, "Monoidal supercategories," 1603.05928. http://arxiv.org/abs/1603.05928.

[138] A. Kirillov, Jr. and B. Balsam, "Turaev-Viro invariants as an extended TQFT," *ArXiv e-prints* (Apr., 2010) , arXiv:1004.1533 [math.GT].

[139] J. Lurie, "On the classification of topological field theories," 0905.0465. http://arxiv.org/abs/0905.0465.

[140] V. Turaev and A. Virelizier, "On the graded center of graded categories," *ArXiv e-prints* (Aug., 2012) , arXiv:1208.5696 [math.QA].

[141] A. Coste, T. Gannon, and P. Ruelle, "Finite group modular data," hep-th/0001158. http://arxiv.org/abs/hep-th/0001158.

[142] D. Tambara and S. Yamagami, "Tensor categories with fusion rules of self-duality for finite abelian groups," *J. Algebra* **209** no. 2, (1998) 692–707.

[143] V. Drinfeld, S. Gelaki, D. Nikshych, and V. Ostrik, "On braided fusion categories i," *Selecta Mathematica* **16** no. 1, (2010) 1–119.

[144] Z.-C. Gu and M. Levin, "The effect of interactions on 2d fermionic symmetry-protected topological phases with z2 symmetry," 1304.4569. http://arxiv.org/abs/1304.4569.

[145] P. Etingof, D. Nikshych, and V. Ostrik, "Fusion categories and homotopy theory," *Quantum Topol.* **1** no. 3, (2010) 209–273.

[146] L. Kong, "Anyon condensation and tensor categories," 1307.8244. http://arxiv.org/abs/1307.8244.

[147] B. Balsam, "Turaev-Viro invariants as an extended TQFT II," *ArXiv e-prints* (Oct., 2010) , arXiv:1010.1222 [math.QA].

[148] B. Balsam, "Turaev-Viro invariants as an extended TQFT III," *ArXiv e-prints* (Dec., 2010) , arXiv:1012.0560 [math.QA].

[149] Y. Tachikawa and K. Yonekura, "On time-reversal anomaly of 2+1d topological phases," arXiv:1610.07010 [hep-th].

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

[151] V. G. Turaev and O. Y. Viro, "State sum invariants of 3-manifolds and quantum 6j-symbols," *Topology* **31** no. 4, (1992) 865–902.

[152] E. Witten, "The "Parity" Anomaly On An Unorientable Manifold," arXiv:1605.02391 [hep-th].

[153] J. Fuchs, C. Schweigert, and A. Valentino, "Bicategories for boundary conditions and for surface defects in 3-d TFT," *Commun. Math. Phys.* **321** (2013) 543–575, arXiv:1203.4568 [hep-th].

[154] M. Cheng, Z.-C. Gu, S. Jiang, and Y. Qi, "Exactly solvable models for symmetry-enriched topological phases," *arXiv preprint arXiv:1606.08482* (2016) .

[155] 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.* **B87** no. 15, (2013) 155114, arXiv:1106.4772 [cond-mat.str-el].

[156] D. Gaiotto and A. Kapustin, "Spin TQFTs and fermionic phases of matter," *Int. J. Mod. Phys.* **A31** no. 28n29, (2016) 1645044, arXiv:1505.05856 [cond-mat.str-el].

[157] 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.* **B90** no. 11, (2014) 115141, arXiv:1201.2648 [cond-mat.str-el].

[158] M. Cheng, Z. Bi, Y.-Z. You, and Z.-C. Gu, "Towards a Complete Classification of Symmetry-Protected Phases for Interacting Fermions in Two Dimensions," *ArXiv e-prints* (Jan., 2015) , arXiv:1501.01313 [cond-mat.str-el].

[159] G. Moore and N. Seiberg, "Classical and quantum conformal field theory," *Communications in Mathematical Physics* **123** no. 2, (1989) 177–254.

[160] G. Moore and N. Seiberg, "Lectures on rcft," in *Physics, Geometry and Topology*, pp. 263–361. Springer, 1990.

[161] M. Barkeshli, P. Bonderson, C.-M. Jian, M. Cheng, and K. Walker, "Reflection and time reversal symmetry enriched topological phases of matter: path integrals, non-orientable manifolds, and anomalies," arXiv:1612.07792 [cond-mat.str-el].

[162] C. Wang and M. Levin, "Anomaly indicators for time-reversal symmetric topological orders," arXiv:1610.04624 [cond-mat.str-el].

[163] Y. Tachikawa and K. Yonekura, "More on time-reversal anomaly of 2+1d topological phases," arXiv:1611.01601 [hep-th].

[164] I. Brunner, N. Carqueville, and D. Plencner, "Discrete Torsion Defects," *Commun. Math. Phys.* **337** (2015) 429–453, arXiv:1404.7497 [hep-th].

[165] E. Sharpe, "Notes on Generalized Global Symmetries in QFT," *Fortsch. Phys.* **63** (2015) 659–682, arXiv:1508.04770 [hep-th].

[166] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, "Theta, Time Reversal, and Temperature," arXiv:1703.00501 [hep-th].

[167] M. Bischoff, R. Longo, Y. Kawahigashi, and K.-H. Rehren, "Tensor Categories of Endomorphisms and Inclusions of Von Neumann Algebras," arXiv:1407.4793 [math.OA].

[168] P. Etingof, S. Gelaki, D. Nikshych, and V. Ostrik, *Tensor categories*, vol. 205 of *Mathematical Surveys and Monographs*. AMS, Providence, RI, 2015.

[169] N. Carqueville and I. Runkel, "Orbifold completion of defect bicategories," *Quantum Topol.* **7** (2016) 203, arXiv:1210.6363 [math.QA].

[170] I. Brunner, N. Carqueville, and D. Plencner, "A Quick Guide to Defect Orbifolds," *Proc. Symp. Pure Math.* **88** (2014) 231–242, arXiv:1310.0062 [hep-th].

[171] G. W. Moore and G. Segal, "D-Branes and K-Theory in 2D Topological Field Theory," arXiv:hep-th/0609042.

[172] G. W. Moore and N. Seiberg, "Lectures on RCFT," in *Strings '89, Proceedings of the Trieste Spring School on Superstrings*. World Scientific, 1990. http://www.physics.rutgers.edu/~gmoore/LecturesRCFT.pdf.

[173] R. Dijkgraaf and E. Witten, "Topological Gauge Theories and Group Cohomology," *Commun. Math. Phys.* **129** (1990) 393.

[174] M. R. Douglas, "D-Branes and Discrete Torsion," arXiv:hep-th/9807235 [hep-th].

[175] D. Tambara and S. Yamagami, "Tensor categories with fusion rules of self-duality for finite abelian groups," *J. Algebra* **209** (1998) 692–707.

[176] P. Etingof and S. Gelaki, "Isocategorical groups," *Internat. Math. Res. Notices* (2001) 59–76, arXiv:math.QA/0007196.

[177] M. Izumi and H. Kosaki, "On a subfactor analogue of the second cohomology," *Rev. Math. Phys.* **14** (2002) 733–757.

[178] J. Fuchs, I. Runkel, and C. Schweigert, "TFT Construction of RCFT Correlators 1. Partition Functions," *Nucl. Phys.* **B646** (2002) 353–497, arXiv:hep-th/0204148 [hep-th].

[179] G. Schaumann, "Traces on module categories over fusion categories," *J. Algebra* **379** (2013) 382–425, arXiv:1206.5716 [math.QA].

[180] L. Kong, "Anyon condensation and tensor categories," *Nucl. Phys.* **B886** (2014) 436–482, arXiv:1307.8244 [cond-mat.str-el].

[181] I. Brunner, N. Carqueville, and D. Plencner, "Orbifolds and Topological Defects," *Commun. Math. Phys.* **332** (2014) 669–712, arXiv:1307.3141 [hep-th].

[182] G. W. Moore and N. Seiberg, "Classical and Quantum Conformal Field Theory," *Commun. Math. Phys.* **123** (1989) 177.

[183] V. Ostrik, "Fusion categories of rank 2," *Math. Res. Lett.* **10** (2003) 177–183, arXiv:math.QA/0203255.

[184] V. Ostrik, "Pivotal fusion categories of rank 3," *Mosc. Math. J.* **15** (2015) 373–396, 405, arXiv:1309.4822.

[185] J. Frohlich, J. Fuchs, I. Runkel, and C. Schweigert, "Defect Lines, Dualities, and Generalised Orbifolds," arXiv:0909.5013 [math-ph].

[186] V. Ostrik, "Module categories, weak Hopf algebras and modular invariants," *Transform. Groups* **8** (2003) 177–206, arXiv:math.QA/0111139.

[187] D. Naidu, "Categorical Morita equivalence for group-theoretical categories," *Comm. Algebra* **35** (2007) 3544–3565, arXiv:math.QA/0605530.

[188] D. Naidu and D. Nikshych, "Lagrangian subcategories and braided tensor equivalences of twisted quantum doubles of finite groups," *Comm. Math. Phys.* **279** (2008) 845–872, arXiv:0705.0665 [math.QA].

[189] B. Uribe, "On the classification of pointed fusion categories up to weak morita equivalence," arXiv:1511.05522.

[190] P. Etingof, S. Gelaki, and V. Ostrik, "Classification of fusion categories of dimension pq," *Int. Math. Res. Not.* **57** (2004) 3041–3056, arXiv:math.QA/0304194.

[191] H. I. Blau, "Fusion rings with few degrees," *J. Algebra* **396** (2013) 220–271.

[192] G. I. Kac and V. G. Paljutkin, "Finite ring groups," *Trans. Moscow Math. Soc.* (1966) 251–294. originally in Russian, *Trudy Moskov. Mat. Obšč.* **15** (1966) 224–261.

[193] E. Meir and E. Musicantov, "Module categories over graded fusion categories," *J. Pure Appl. Algebra* **216** (2012) 2449–2466, arXiv:1010.4333 [math.QA].

[194] M. Buican and A. Gromov, "Anyonic Chains, Topological Defects, and Conformal Field Theory," arXiv:1701.02800 [hep-th].

[195] M. Bershadsky and C. Vafa, "Global anomalies and geometric engineering of critical theories in six-dimensions," arXiv:hep-th/9703167 [hep-th].

[196] J. McOrist, D. R. Morrison, and S. Sethi, "Geometries, Non-Geometries, and Fluxes," *Adv. Theor. Math. Phys.* **14** (2010) 1515–1583, arXiv:1004.5447 [hep-th].

[197] D. R. Morrison, "On K3 surfaces with large Picard number," *Invent. Math.* **75** (1984) 105–121.

[198] K. Oguiso, "On Jacobian fibrations on the Kummer surfaces of the product of nonisogenous elliptic curves," *J. Math. Soc. Japan* **41** no. 4, (1989) 651–680.

[199] A. Clingher and C. F. Doran, "Modular invariants for lattice polarized $K3$ surfaces," *Michigan Math. J.* **55** no. 2, (2007) 355–393, arXiv:math.AG/0602146.

[200] A. Kumar, "$K3$ surfaces associated with curves of genus two," *Int. Math. Res. Not.* **2008** no. 6, (2008) Art. ID rnm165, arXiv:math.AG/0701669.

[201] A. Clingher and C. F. Doran, "Lattice polarized K3 surfaces and Siegel modular forms," *Adv. Math.* **231** no. 1, (2012) 172–212, arXiv:1004.3503 [math.AG].

[202] A. Malmendier and D. R. Morrison, "K3 surfaces, modular forms, and non-geometric heterotic compactifications," *Lett. Math. Phys.* **105** (2015) 1085–1118, arXiv:1406.4873 [hep-th].

[203] M. Reid, "Young person's guide to canonical singularities," in *Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985)*, vol. 46 of *Proc. Sympos. Pure Math.*, pp. 345–414. Amer. Math. Soc., 1987.

[204] S. Chaudhuri, G. Hockney, and J. D. Lykken, "Maximally Supersymmetric String Theories in $D < 10$," *Phys. Rev. Lett.* **75** (1995) 2264–2267, arXiv:hep-th/9505054 [hep-th].

[205] A. Mikhailov, "Momentum Lattice for CHL String," *Nucl. Phys.* **B534** (1998) 612–652, arXiv:hep-th/9806030 [hep-th].

[206] M. Bershadsky, T. Pantev, and V. Sadov, "F Theory with Quantized Fluxes," *Adv. Theor. Math. Phys.* **3** (1999) 727–773, arXiv:hep-th/9805056 [hep-th].

[207] P. Berglund, A. Klemm, P. Mayr, and S. Theisen, "On Type IIB Vacua with Varying Coupling Constant," *Nucl. Phys.* **B558** (1999) 178–204, arXiv:hep-th/9805189.

[208] S. Katz, D. R. Morrison, S. Schafer-Nameki, and J. Sully, "Tate's algorithm and F-theory," *JHEP* **08** (2011) 094, arXiv:1106.3854 [hep-th].

[209] P. S. Aspinwall and D. R. Morrison, "Non-Simply-Connected Gauge Groups and Rational Points on Elliptic Curves," *JHEP* **07** (1998) 012, arXiv:hep-th/9805206.

[210] G. Pradisi, "Magnetic Fluxes, NS NS B Field and Shifts in Four-Dimensional Orientifolds," in *String Phenomenology. Proceedings, 2nd International Conference, Durham, UK, July 29-August 4, 2003*, pp. 304–314. 2003. arXiv:hep-th/0310154 [hep-th].

[211] A. Kapustin and R. Thorngren, "Higher symmetry and gapped phases of gauge theories," arXiv:1309.4721 [hep-th].

[212] L. Evens, *The cohomology of groups*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1991. Oxford Science Publications.

[213] D. Handel, "On products in the cohomology of the dihedral groups," *Tohoku Math. J. (2)* **45** (1993) 13–42.

[214] M. D. F. de Wild Propitius, "Topological Interactions in Broken Gauge Theories," arXiv:hep-th/9511195 [hep-th]. PhD thesis at Amsterdam U., 1995.

[215] T. Hayami and K. Sanada, "Cohomology ring of the generalized quaternion group with coefficients in an order," *Comm. Algebra* **30** (2002) 3611–3628.

APPENDICES

Appendix A

Appendix to Chapter 2

A.1 Brief Review of Anomaly Cancellation in F-theory

The allowed gauge algebras and matter content for a given tensor branch structure is heavily constrained by anomaly cancellation. In 6D, anomalies are related to four-point amplitudes of external currents J_a associated with continuous symmetry group G_a. In such a four-point amplitude, insertions of external currents J_a for a given gauge group G_a must come in pairs, so we need only consider the anomalies related to the four point functions $\langle J_a J_a J_a J_a \rangle$ (gauge anomaly cancellation) and $\langle J_a J_a J_b J_b \rangle$ (mixed gauge anomaly cancellation).

For a representation R of some gauge group G, we introduce constants Ind_R, x_R, and y_R relating the quadratic and quartic Casimirs of G as follows:

$$\mathrm{Tr}_R F^2 = Ind_R \mathrm{tr} F^2\,, \quad \mathrm{Tr}_R F^4 = x_R \mathrm{tr} F^4 + y_R (\mathrm{tr} F^2)^2. \tag{A.1.1}$$

Here, tr indicates the trace in a defining representation of the group[1].

The constraints on gauge and mixed anomalies take the form [52, 195, 53, 51]

$$Ind_{Adj_a} - \sum_R Ind_{R_a} n_{R_a} = 6(10-n)\frac{\Omega_{IJ} a^I b_a^J}{\Omega_{IJ} a^I a^J} \tag{A.1.2}$$

$$y_{Adj_a} - \sum_R y_{R_a} n_{R_a} = -3(10-n)\frac{\Omega_{IJ} b_a^I b_a^J}{\Omega_{IJ} a^I a^J} \tag{A.1.3}$$

$$x_{Adj_a} - \sum_R x_{R_a} n_{R_a} = 0 \tag{A.1.4}$$

$$\sum_{R,R'} Ind_{R_a} Ind_{R'_b} n_{R_a R'_b} = (10-n)\frac{\Omega_{IJ} b_a^I b_b^J}{\Omega_{IJ} a^I a^J}. \tag{A.1.5}$$

[1] For $SU(N)$ and $Sp(N)$, the defining representation is simply the fundamental representation. For $SO(5)$ and $SO(6)$, it is the spinor representation. For $SO(N)$, $N \geq 7$, it is the fundamental representation, though normalized with an additional factor of 2 so that $\mathrm{Tr}_f F^2 = 2\mathrm{tr} F^2$, $\mathrm{Tr}_f F^4 = 2\mathrm{tr} F^4$. A complete list including exceptional gauge groups can be found in Table 2 of [51].

Here, Ω_{IJ} is the natural metric on the space of antisymmetric tensors, and a^I, b_a^J are related to the anomaly 8-form I_8 via

$$I_8 = \frac{1}{2}\Omega_{IJ}X^I X^J\,, \quad X^I = \frac{1}{2}a^I \mathrm{tr}R^2 + 2b_a^I \mathrm{tr}F_a^2. \tag{A.1.6}$$

These field theoretic conditions can be translated into F-theory language as restrictions on the allowed gauge algebras and matter for a given base. Any gauge algebra summand $\mathfrak{g}_a$ in the theory is paired with a tensor multiplet, which is in turn associated with a curve Σ_a in the base. In these terms, the anomaly cancellation conditions become

$$Ind_{Adj_a} - \sum_R Ind_{R_a} n_{R_a} = 6(K\cdot\Sigma_a) = 6(2g_a - 2 - \Sigma_a\cdot\Sigma_a) \tag{A.1.7}$$

$$y_{Adj_a} - \sum_R y_{R_a} n_{R_a} = -3(\Sigma_a\cdot\Sigma_a) \tag{A.1.8}$$

$$x_{Adj_a} - \sum_R x_{R_a} n_{R_a} = 0 \tag{A.1.9}$$

$$\sum_{R,R'} Ind_{R_a} Ind_{R'_b} n_{R_a R'_b} = \Sigma_a\cdot\Sigma_b. \tag{A.1.10}$$

Here, K is the canonical divisor of the base and g_a is the genus of Σ_a. The adjunction formula $K\cdot\Sigma_a = 2(g_a-1) - \Sigma_a\cdot\Sigma_a$ has been used in the second equality of (A.1.7).

Gauge anomaly cancellation may be used to constrain the gauge groups paired with a given tensor node, and mixed anomaly cancellation constrains the gauge groups allowed on neighboring tensor nodes. However, tensors need not be paired with gauge groups. In particular, curves of self-intersection 0, -1, or -2 can be devoid of a gauge group entirely. Curves of self-intersection 0 cannot touch any other curves, so this case is relatively uninteresting. Curves of self-intersection -1 and -2, on the other hand, can touch other curves.

A.2 Matter for Singular Bases and Tangential Intersections

In this Appendix, we review the relationship between matter, singularities and tangential intersections. Further information can be found in [54].

Consider some curve Σ. When this curve becomes singular, there are two distinct notions of genus. Geometric genus, denoted p_g, is the topological genus of the curve after all singularities have been resolved. Arithmetic genus, denoted g, is the quantity related to the intersection theory of the curve by

$$2g - 2 = K\cdot\Sigma + \Sigma^2. \tag{A.2.1}$$

The arithmetic genus is the one that shows up in the adjunction formula and hence enters the anomaly cancellation equation (A.1.7). These two notions of the genus are related via

$$g = p_g + \sum_P \frac{m_P(m_P-1)}{2}. \tag{A.2.2}$$

Here, the sum runs over all singular points P of the curve and m_P is the multiplicity of the singularity at P. For a curve with a nodal singularity (of Kodaira type I_1) or a cusp singularity (of Kodaira type II), there is a single singular point of multiplicity $m_P = 2$. Hence, in each of these cases we have $g = p_g + 1$. The only values of g and p_g that can actually arise in our classification are $g = 1$ and $p_g = 0$.

Although p_g does not show up directly in the anomaly cancellation conditions, it still determines the F-theory matter content. Namely, p_g is precisely the number of adjoint hypermultiplets charged under the gauge algebra paired with this curve. For *smooth* LST bases, we see that there is one adjoint whenever the genus $g = p_g$ of the curve is 1. For singular curves, on the other hand, the arithmetic genus g can be 1 even without any adjoint hypermultiplets. For curves with gauge algebra $\mathfrak{su}(N)$, one instead finds symmetric and antisymmetric representations, as noted in [52].

We now consider the Kodaira type III base configuration, which consists of two -2 curves intersecting tangentially. First, we show that two curves meeting tangentially can indeed carry fiber types I_0^* and I_4, which correspond to gauge algebras $\mathfrak{g}_2/\mathfrak{so}(7)/\mathfrak{so}(8)$ and $\mathfrak{su}(4)$, respectively. For this, we consider the Weierstrass model:

$$y^2 = x^3 + (zu + \sigma^2)x^2 + (zu + \sigma^2)^2 z^2 v x + (zu + \sigma^2)^3 z^4 w \tag{A.2.3}$$

Here z and σ are local coordinates on the base and u, v, and w are functions that do not vanish at $(z, \sigma) = (0, 0)$. From this, we compute f, g, and Δ to be

$$\begin{aligned} f &= (zu + \sigma^2)^2(-1/3 + z^2 v) \\ g &= (zu + \sigma^2)^3(2/27 - (1/3)z^2 v + z^4 w) \\ \Delta &= z^4(zu + \sigma^2)^6(-v^2 - 18z^2 wv + 4z^2 v^3 + 27z^4 w^2 + 4w). \end{aligned} \tag{A.2.4}$$

From this, we see that the curve $z = 0$ has Kodaira type I_4 whereas the curve $zu + \sigma^2$ has Kodaira type I_0^*. These curves meet at $(z, \sigma) = (0, 0)$, where they are tangent to each other. The multiplicities of f, g, and Δ at $(0, 0)$ are 2, 3, and 10, respectively. On the other hand, had we attempted to intersect transversely curves with fibers of singularity types I_0^* and I_4, we would have found that the multiplicities of f, g, and Δ would have been 4, 6, and 12, respectively. Hence, curves with these fiber types cannot meet transversely, only tangentially.

We can enhance the gauge algebras on these tangent curves to $\mathfrak{so}(12)$ and $\mathfrak{su}(6)$, corresponding to Kodaira fiber types I_2^* and I_6, respectively. Here, the Weierstrass model is

$$y^2 = x^3 + (zu + \sigma^2)x^2 + (zu + \sigma^2)^3 z^3 v x \tag{A.2.5}$$

This yields

$$\begin{aligned} f &= -\frac{1}{3}(zu + \sigma^2)^2 + v z^3 (zu + \sigma^2)^3 \\ g &= \frac{2}{27}(zu + \sigma^2)^3 - \frac{1}{3} v z^3 (zu + \sigma^2)^4 \\ \Delta &= -v^2 z^6 (zu + \sigma^2)^8 + 4v^3 z^9 (zu + \sigma^2)^9. \end{aligned} \tag{A.2.6}$$

Thus, from the degrees of vanishing of f, g, and Δ, we see the curve $z = 0$ has Kodaira type I_6 whereas the curve $zu + \sigma^2$ has Kodaira type I_2^*.

We also get $\mathfrak{so}(13)$ and $\mathfrak{su}(7)$ from Kodaira fiber types I_3^* and I_7, respectively, using the Weierstrass model

$$y^2 = x^3 + (zu + \sigma^2)x^2 + w(zu + \sigma^2)^6 z^7. \tag{A.2.7}$$

This model has

$$\begin{aligned} f &= -\frac{1}{3}(zu + \sigma^2)^2 \\ g &= \frac{2}{27}(zu + \sigma^2)^3 + wz^7(zu + \sigma^2)^6 \\ \Delta &= 4wz^7(zu + \sigma^2)^9 + 27w^2z^{12}(zu + \sigma^2)^{16}. \end{aligned} \tag{A.2.8}$$

The fact that we got $\mathfrak{so}(13)$ in this case and $\mathfrak{so}(12)$ in the previous case follows from Table 7 of [51].

Using these Weierstrass models, we are able to realize the gauge algebra enhancements on two tangent -2 curves discussed in section 2.7.

A.3 Novel DE Type Bases

Appendices B and C [10] provided a complete list of DE type bases that can be used to construct 6D SCFTs. All of these bases can show up in LSTs as well when certain non-DE type links are suitably attached. However, there are also some novel DE type bases, which blow down to 0.

We use an abbreviated notation to describe these bases. Namely, we specify curves of self-intersection -4, -6, -8, and -12 according to

$$D \simeq 4 \tag{A.3.1}$$

$$E_6 \simeq 6 \tag{A.3.2}$$

$$E_7 \simeq 8 \tag{A.3.3}$$

$$E_8 \simeq 12 \tag{A.3.4}$$

The allowed types of conformal matter between DE type nodes are represented as follows:

$$D1D \simeq D \overset{1,1}{\oplus} D \tag{A.3.5}$$

$$E_6131E_6 \simeq E_6 \overset{2,2}{\oplus} E_6 \tag{A.3.6}$$

$$E_712321E_6 \simeq E_7 \overset{3,3}{\oplus} E_6 \tag{A.3.7}$$

$$E_712321E_7 \simeq E_7 \overset{3,3}{\oplus} E_7 \tag{A.3.8}$$

$$E_71231D \simeq E_7 \overset{3,2}{\oplus} D \tag{A.3.9}$$

$$E_812231D \simeq E_8 \overset{4,2}{\oplus} D \tag{A.3.10}$$

$$E_61315131E_6 \simeq E_6 \overset{3,3}{\bigcirc} E_6 \tag{A.3.11}$$

$$E_613151321E_7 \simeq E_6 \overset{3,4}{\oplus} E_7 \tag{A.3.12}$$

$$E_6131513221E_8 \simeq E_6 \overset{3,5}{\oplus} E_8 \tag{A.3.13}$$

$$E_7123151321E_7 \simeq E_7 \overset{4,4}{\oplus} E_7 \tag{A.3.14}$$

$$E_71231513221E_8 \simeq E_7 \overset{4,5}{\oplus} E_8 \tag{A.3.15}$$

$$E_812231513221E_8 \simeq E_8 \overset{5,5}{\oplus} E_8. \tag{A.3.16}$$

We omit the subscripts above the $\oplus$ when dealing with the minimal type of conformal matter:

$$D \oplus D \simeq D \overset{1,1}{\oplus} D \tag{A.3.17}$$

$$D \oplus E_6 \simeq D \overset{2,2}{\oplus} E_6 \tag{A.3.18}$$

$$D \oplus E_7 \simeq D \overset{2,3}{\oplus} E_7 \tag{A.3.19}$$

$$D \oplus E_8 \simeq D \overset{2,4}{\oplus} E_8 \tag{A.3.20}$$

$$E_6 \oplus E_6 \simeq E_6 \overset{2,2}{\oplus} E_6 \tag{A.3.21}$$

$$E_6 \oplus E_7 \simeq E_6 \overset{3,3}{\oplus} E_7 \tag{A.3.22}$$

$$E_6 \oplus E_8 \simeq E_6 \overset{3,5}{\oplus} E_8 \tag{A.3.23}$$

$$E_7 \oplus E_7 \simeq E_7 \overset{3,3}{\oplus} E_7 \tag{A.3.24}$$

$$E_7 \oplus E_8 \simeq E_7 \overset{4,5}{\oplus} E_8 \tag{A.3.25}$$

$$E_8 \oplus E_8 \simeq E_8 \overset{5,5}{\oplus} E_8 \tag{A.3.26}$$

Using this notation, we now list the novel DE type bases for LSTs. All of these bases are

positive semi-definite with a single zero eigenvalue. We begin with configurations with only D nodes:

$$D \overset{3,3}{\oplus} D \tag{A.3.27}$$

$$D \overset{3,3}{\bigcirc} D \tag{A.3.28}$$

The configurations with only E_6 nodes:

$$E_6 \overset{5,5}{\oplus} E_6 \tag{A.3.29}$$

The configurations with D and E_6:

$$D \overset{3,5}{\oplus} E_6 \tag{A.3.30}$$

$$D \oplus E_6^{\oplus 2} \overset{3,3}{\oplus} E_6 \tag{A.3.31}$$

$$D \oplus E_6^{\oplus 2} \overset{3,3}{\bigcirc} E_6 \tag{A.3.32}$$

$$D \oplus E_6 \overset{3,4}{\oplus} E_6 \tag{A.3.33}$$

$$D \oplus E_6 \overset{3,2}{\oplus} D \tag{A.3.34}$$

$$D^{\oplus 2} \oplus E_6 \oplus D \tag{A.3.35}$$

The configurations with D and E_7:

$$D \overset{3,3}{\oplus} E_7^{\oplus n} \overset{3,3}{\oplus} D\,, \quad n = 1, 2, ... \tag{A.3.36}$$

$$D \oplus E_7^{\oplus 3} \overset{4,4}{\oplus} E_7 \tag{A.3.37}$$

$$D \oplus E_7^{\oplus 2} \overset{4,5}{\oplus} E_7 \tag{A.3.38}$$

$$D^{\oplus 2} \oplus E_7 \oplus D \tag{A.3.39}$$

$$D \overset{2,4}{\oplus} E_7 \oplus D \tag{A.3.40}$$

$$D \oplus 1\overset{2}{3}21 \oplus E_7 \tag{A.3.41}$$

The configurations with D and E_8:

$$D \overset{3,5}{\oplus} E_8^{\oplus n} \overset{5,3}{\oplus} D\,, \quad n = 1, 2, ... \tag{A.3.42}$$

$$D^{\oplus 3} \oplus E_8 \oplus D^{\oplus 2} \tag{A.3.43}$$

$$D^{\oplus 4} \oplus E_8 \oplus D \tag{A.3.44}$$

The configurations with E_6 and E_8:

$$E_6 \overset{5,5}{\oplus} E_8^{\oplus n} \overset{5,5}{\oplus} E_6\,, \quad n = 1, 2, ... \tag{A.3.45}$$

$$E_6 \overset{4,5}{\oplus} E_8^{\oplus 2} \oplus E_6^{\oplus 2} \tag{A.3.46}$$

The configurations with D, E_6 and E_7:

$$D \oplus E_7 \overset{4,4}{\oplus} E_6 \tag{A.3.47}$$

$$D \oplus E_7^{\oplus 2} \overset{4,3}{\oplus} E_6 \tag{A.3.48}$$

$$D \oplus E_6^{\oplus 2} \overset{3,5}{\oplus} E_7 \tag{A.3.49}$$

$$D \oplus E_6^{\oplus 3} \overset{3,4}{\oplus} E_7 \tag{A.3.50}$$

$$D \oplus E_6^{\oplus 4} \oplus E_7 \tag{A.3.51}$$

$$D \oplus E_7^{\oplus 2} \oplus E_6^{\oplus 2} \tag{A.3.52}$$

$$D \oplus E_6 \oplus E_7 \oplus E_6^{\oplus 2} \tag{A.3.53}$$

The configurations with D, E_6 and E_8:

$$D \overset{4,5}{\oplus} E_8^{\oplus n} \overset{5,5}{\oplus} E_6\,, \quad n = 1, 2, ... \tag{A.3.54}$$

$$D \oplus E_6^{\oplus 6} \oplus E_8 \tag{A.3.55}$$

$$D^{\oplus 2} \oplus E_8^{\oplus 2} \oplus E_6^{\oplus 2} \tag{A.3.56}$$

$$D^{\oplus 3} \oplus E_8 \overset{5,4}{\oplus} E_6 \tag{A.3.57}$$

The configurations with D, E_7 and E_8:

$$D \oplus E_7^{\oplus 6} \overset{4,5}{\oplus} E_8 \tag{A.3.58}$$

$$D^{\oplus 2} \oplus E_8^{\oplus 3} \oplus E_7^{\oplus 2} \tag{A.3.59}$$

The configurations with E_6, E_7 and E_8:

$$E_6 \overset{4,5}{\oplus} E_8 \oplus E_7 \oplus E_6 \tag{A.3.60}$$

$$E_6 \overset{4,5}{\oplus} E_8^{\oplus 3} \oplus E_7^{\oplus 2} \tag{A.3.61}$$

The configurations with D, E_6, E_7 and E_8:

$$D \oplus E_8 \oplus E_7^{\oplus 2} \oplus E_6 \tag{A.3.62}$$

$$D^{\oplus 2} \oplus E_8 \oplus E_7 \oplus E_6 \tag{A.3.63}$$

A.4 Novel Non-DE Type Bases

The following is the list of novel bases constructed solely from curves of self-intersection -1, -2, -3, and -5:

$$12....21 \tag{A.4.1}$$

$$5 \oplus 1^{\oplus 5} \tag{A.4.2}$$

$$\overset{1}{1}5\underset{1}{1}2 \tag{A.4.3}$$

$$\overset{1}{2}1\overset{}{5}12 \tag{A.4.4}$$

$$1312 \tag{A.4.5}$$

$$\overset{1}{1}31 \tag{A.4.6}$$

$$\overset{1}{1}31\overset{}{5}12 \tag{A.4.7}$$

$$\overset{1}{1}31\overset{}{5}\underset{1}{1}2 \tag{A.4.8}$$

$$1\overset{1}{2}31\overset{}{5}12 \tag{A.4.9}$$

$$1\overset{1}{2}31\overset{}{5}\underset{1}{1}2 \tag{A.4.10}$$

$$12\overset{1}{2}31\overset{}{5}12 \tag{A.4.11}$$

$$12\overset{1}{2}31\overset{}{5}\underset{1}{1}2 \tag{A.4.12}$$

$$\overset{1}{2}31\overset{}{5}12 \tag{A.4.13}$$

$$\overset{1}{2}3\overset{1}{1}51 \tag{A.4.14}$$

$$3131512 \tag{A.4.15}$$

$$313\overset{1}{1}51 \tag{A.4.16}$$

$$215131512 \tag{A.4.17}$$

$$2151315\overset{1}{1} \tag{A.4.18}$$

$$\overset{1}{1}513\overset{1}{1}51 \tag{A.4.19}$$

$$\overset{1}{1}5131512 \tag{A.4.20}$$

$$\overset{1}{1}513\overset{1}{1}51 \tag{A.4.21}$$

$$1315131512 \tag{A.4.22}$$

$$13151315\overset{1}{1} \tag{A.4.23}$$

$$12315131512 \tag{A.4.24}$$

$$123151315\overset{1}{1} \tag{A.4.25}$$

$$122315131512 \tag{A.4.26}$$

$$1223151315\overset{1}{1} \tag{A.4.27}$$

$$151231512 \tag{A.4.28}$$

$$1512315\overset{1}{1} \tag{A.4.29}$$

$$512231512 \tag{A.4.30}$$

$$5122315\overset{1}{1} \tag{A.4.31}$$

$$12321512 \tag{A.4.32}$$

$$123215\overset{1}{1} \tag{A.4.33}$$

$$1321512 \tag{A.4.34}$$

$$13215\overset{1}{1} \tag{A.4.35}$$

$$\overset{1}{1}322 \tag{A.4.36}$$

$$\overset{2}{2}2...21 \tag{A.4.37}$$

$$151\overset{1}{3}22 \tag{A.4.38}$$

$$15123\overset{1}{2} \tag{A.4.39}$$

$$512222 \tag{A.4.40}$$

$$151222 \tag{A.4.41}$$

$$3122 \tag{A.4.42}$$

$$1315122 \tag{A.4.43}$$

$$12315122 \tag{A.4.44}$$

$$122315122 \tag{A.4.45}$$

$$215122 \tag{A.4.46}$$

$$\overset{1}{1}5122 \tag{A.4.47}$$

$$1315132 \tag{A.4.48}$$

$$12315132 \tag{A.4.49}$$

$$122315132 \tag{A.4.50}$$

$$215132 \tag{A.4.51}$$

$$\overset{1}{1}5\overset{1}{1}32 \tag{A.4.52}$$

$$32132 \tag{A.4.53}$$

$$12231321 \tag{A.4.54}$$

$$23123 \tag{A.4.55}$$

$$12313221 \tag{A.4.56}$$

$$1313221 \tag{A.4.57}$$

$$1512231 \tag{A.4.58}$$

$$13213 \tag{A.4.59}$$

$$123131 \tag{A.4.60}$$

$$13131 \tag{A.4.61}$$

$$312321 \tag{A.4.62}$$

$$3132151 \tag{A.4.63}$$

$$2313151 \tag{A.4.64}$$

$$151232151 \tag{A.4.65}$$

$$1232151321 \tag{A.4.66}$$

$$5123132 \tag{A.4.67}$$

$$3215123 \tag{A.4.68}$$

$$123151231 \tag{A.4.69}$$

$$13151231 \tag{A.4.70}$$

$$1321513221 \tag{A.4.71}$$

$$5131322 \tag{A.4.72}$$

$$13215131 \tag{A.4.73}$$

$$3221513 \tag{A.4.74}$$

1512313	(A.4.75)
5122313	(A.4.76)
1231512321	(A.4.77)
12231512321	(A.4.78)
131512321	(A.4.79)
12315132151	(A.4.80)
1315132151	(A.4.81)
315123151	(A.4.82)
122315132215	(A.4.83)
321513151	(A.4.84)
315123215	(A.4.85)
513215132	(A.4.86)
51223151321	(A.4.87)
512315123	(A.4.88)
151231513221	(A.4.89)
3131513221	(A.4.90)
1512315131	(A.4.91)
5122315131	(A.4.92)
512321513	(A.4.93)
123151313	(A.4.94)
13151313	(A.4.95)
513151232	(A.4.96)
15131513215	(A.4.97)
15131513151	(A.4.98)
12231513151321	(A.4.99)
12315131513221	(A.4.100)
1315131513221	(A.4.101)
123151315131	(A.4.102)
13151315131	(A.4.103)
51315131513	(A.4.104)

A.5 Novel Gluings

At times, non-DE type side links or noble atoms can attach to DE type nodes in ways they could not for 6D SCFTs. For instance, the side link 2151321 could never attach to a D node in a 6D SCFT, since it induces four blow-downs on the -4 curve. However, this is allowed for LSTs. The full list of these novel gluings of one side link to a single node is as follows:

For gluing to a D-type node (i.e. a -4 curve), we have:

$$2221 \oplus D \tag{A.5.1}$$
$$2151321 \oplus D \tag{A.5.2}$$
$$3215131 \oplus D \tag{A.5.3}$$
$$1512321 \oplus D \tag{A.5.4}$$
$$3151231 \oplus D \tag{A.5.5}$$
$$3\overset{1}{1}5131 \oplus D \tag{A.5.6}$$
$$22\overset{1}{3}1 \oplus D \tag{A.5.7}$$
$$\overset{1}{1}51321 \oplus D \tag{A.5.8}$$
$$151315131 \oplus D \tag{A.5.9}$$
$$512315131 \oplus D \tag{A.5.10}$$

For gluing to an E_6-type node (i.e. a -6 curve), we have:

$$222221 \oplus E_6 \tag{A.5.11}$$
$$223\overset{1}{1}5131 \oplus E_6 \tag{A.5.12}$$
$$231512321 \oplus E_6 \tag{A.5.13}$$
$$232151321 \oplus E_6 \tag{A.5.14}$$
$$321513221 \oplus E_6 \tag{A.5.15}$$
$$3\overset{1}{1}513221 \oplus E_6 \tag{A.5.16}$$
$$23\overset{1}{1}51321 \oplus E_6 \tag{A.5.17}$$
$$15131513221 \oplus E_6 \tag{A.5.18}$$
$$51231513221 \oplus E_6 \tag{A.5.19}$$
$$3151315121 \oplus E_6 \tag{A.5.20}$$
$$23151315131 \oplus E_6 \tag{A.5.21}$$

For gluing to an E_7-type node (i.e. a -8 curve), we have:

$$\underbrace{22222221}_{8} \oplus E_7 \tag{A.5.22}$$
$$2231513151321 \oplus E_7 \tag{A.5.23}$$
$$223\overset{1}{1}513221 \oplus E_7 \tag{A.5.24}$$
$$2231\overset{1}{5}13221 \oplus E_7 \tag{A.5.25}$$

while for an E_8-type node (i.e. a -12 curve), we have:

$$\underbrace{222222222221}_{12} \oplus E_8. \tag{A.5.26}$$

We remark that in the case of the E_7 and E_8 nodes, we can also delete some of the aforementioned curves, using instead a "primed" node, (i.e. by adding a small instanton link elsewhere).

A.6 T-Duality in the $1, 2, \ldots, 2, 1$ Model

Let us consider the LST whose F-theory base B has a chain of rational curves consisting of two curves of self-intersection -1 at the ends of the chain, and $k \geq 0$ curves of self-intersection -2 in between meet each other. The union of those $k+2$ curves on the base deforms to a nonsingular rational curve of self-intersection 0, and provides a fibration $\pi : B \to C$ on the base whose general fiber is $\mathbb{P}^1$. The total space of the elliptic fibration looks like a one parameter family of elliptic K3 surfaces (over general fibers of π) degenerating to a pair of dP_9's with k intermediate elliptic ruled surfaces (over $\pi^{-1}(0)$), which (when $k = 0$) is precisely the degeneration which appears in the analysis of heterotic / F-theory duality in reference [72].

Remarkably, the total space of this elliptic fibration admits a *second* elliptic fibration, at least birationally. The computation which shows this was written out in reference [196], although it has some earlier antecedents in the math literature [197–200].[2]

Consider a base of the form $\mathbb{P}^1 \times \mathbb{C}$ where the homogeneous coordinates on $\mathbb{P}^1$ are $[\sigma, \tau]$ on the coordinate on $\mathbb{C}$ is ψ. We write a Weierstrass equation of the form

$$Y^2 = X^3 + a\sigma^4\tau^4 X + (\psi^{k+1}\sigma^5\tau^7 + c\sigma^6\tau^6 + \sigma^7\tau^5), \tag{A.6.1}$$

where a and c are constants. (In [196] only the $\tau = 1$ affine chart appears.) Notice that we get a small instanton with instanton number $k+1$ when $\psi = \sigma = 0$, so to get the $1, 2, \ldots, 2, 1$ model we should blowup the point $\psi = \sigma = 0$ $k+1$ times. We keep this implicit in what follows.

Note that we have Kodaira type II^* at both $\sigma = 0$ and $\tau = 0$, and since those represent non-compact curves, we see $E_8 \times E_8$ global symmetry.

Now there is a remarkable coordinate change:

$$X = st^{-5}x^2 \tag{A.6.2}$$
$$Y = t^{-8}x^2y \tag{A.6.3}$$
$$\sigma = t^{-3}x \tag{A.6.4}$$
$$\tau = t. \tag{A.6.5}$$

(This is only a "rational map" because of the division by powers of t.) The result of the substitu-

[2] There was a subsequent extension of this computation to a more general case [200–202], which will undoubtedly be useful for understanding additional T-dualities of LSTs.

tion is

$$t^{-16}x^4y^2 = s^3t^{-15}x^6 + ast^{-13}x^6 + (\psi^{k+1}t^{-8}x^5 + ct^{-12}x^6 + t^{-16}x^7). \tag{A.6.6}$$

If we multiply the resulting equation by t^{16}/x^4, we obtain

$$y^2 = s^3tx^2 + ast^3x^2 + (\psi^{k+1}t^8x + ct^4x^2 + x^3) \tag{A.6.7}$$

$$= x^3 + (s^3t + ast^3 + ct^4)x^2 + \psi^{k+1}t^8x. \tag{A.6.8}$$

We interpret this as a family elliptic curves over $\mathbb{P}^1_{[s,t]} \times \mathbb{C}$. (Note that this is a very different base, with coordinates $[s,t]$ which mix the former base and fiber coordinates.) Remarkably, this is the equation for the F-theory dual of the $Spin(32)/\mathbb{Z}_2$ heterotic string with a small instanton at $t = \psi = 0$ (as derived in [72, 15])! We have a global symmetry algebra $\mathfrak{so}(32)$ along $t = 0$. Note that we should also blow up $t = \psi = 0$.

Thus, in a quite subtle way there is T-duality for this pair of models, in which the two different elliptic fibrations on the semi-local total space are exchanged, after a birational change. Note that the coordinate change given above is (rationally) invertible. The inverse is:

$$x = \sigma\tau^3 \tag{A.6.9}$$

$$y = \sigma^{-2}\tau Y \tag{A.6.10}$$

$$s = \sigma^{-2}\tau^{-1}X \tag{A.6.11}$$

$$t = \tau. \tag{A.6.12}$$

A.7 F-Theory Construction of $\mathcal{N} = (1,1)$ LSTs

The M-theory construction for little string theories with $\mathcal{N} = (1,1)$ supersymmetry is described in [47]. Such a theory can be seen as M-theory compactified on a spacetime of the form $(\mathbb{C}^2 \times S^1)/\Gamma$ for $\Gamma \subset SU(2)$ a finite subgroup. The action on S^1 is by rotations, and there is a subgroup Γ' of Γ which acts trivially on S^1, leading to a short exact sequence of groups

$$0 \to \Gamma' \to \Gamma \to \mathbb{Z}_r \to 0. \tag{A.7.1}$$

Geometrically, there is an action of $\mathbb{Z}_r$ on the ALE space $\mathbb{C}^2/\Gamma'$ corresponding to an automorphism of the corresponding Lie group; the gauge group of the little string theory is the subgroup commuting with the outer automorphism.

The mathematical description of these groups has been known for a long time, and is nicely summarized in a table on p. 376 of [203] which we reproduce as Table A.7.1. We have made a minor correction to the table (already noted in footnote 15 of [47]), and we have added the information about the gauge group and theta angle as discussed in [47]. In the table, the cyclic group action on $\mathbb{C}^3$ with coordinates (x,y,z) is described in terms of exponents $(\frac{a}{r}, \frac{b}{r}, \frac{c}{r})$ of the generators; the notation also indicates how the equation transforms under the action.

To construct these theories using F-theory, we use a base B which is a neighborhood of an elliptic curve Σ whose normal bundle is a torsion line bundle of order r. The base has a finite

unramified cover of degree r which is a product $\Sigma \times \mathbb{C}$, and the cyclic group $\mathbb{Z}_r$ will act on the elliptic fibration over $\Sigma \times \mathbb{C}$ (which we take to be in Weierstrass form). Thus, the classification is analogous – we must find cyclic group actions on Weierstrass elliptic fibrations which induce the corresponding actions on the ADE singularities.

Note that, as observed in section 2.7.1, this construction requires $r \in \{2, 3, 4, 6\}$, so that most of the instances of case (1) are ruled out. In fact, we have been unable to find a conventional F-theory construction of any instance of case (1) (which would correspond, in the interpretation of [47], to an $SU(n)$ theory with a rational theta angle). Instead, as mentioned in section 2.10.2, we anticipate an F-theory construction for these models involving B-field expectation values.

In Table A.7.2, we present explicit forms of these group actions in cases (3), (4), and (6), using a Weierstrass equation φ with variables (x, y, t). The quotient can be described in terms of a Weierstrass equation Φ whose variables (X, Y, T) are expressed in terms of (x, y, t) in the table.

We now explain the geometry of the group actions by means of figures illustrating cases (2), (3), (4), and (6). We will explain the cases in the reverse order from the one given in Table A.7.1.

We begin with case (6) leading to gauge symmetry F_4, illustrated in Figure A.7.1. The M-theory construction only involved the Dynkin diagram of the singularities, but here we must consider the entire Kodaira fiber as represented by the affine Dynkin diagram. Thus, in the bottom half of Figure A.7.1 we see $\widehat{E}_6$, with the affine node represented by a square rather than a circle. This diagram has been obtained as a double cover of the affine E_8 diagram shown in the top half of Figure A.7.1, in which the solid circles represent curves along which we branch as we construct a double cover. The two curves on the left are duplicated in the bottom half because they don't meet the trivalent vertex at a branch point of the double cover. After taking the double cover, the solid curves have become -1 curves and are to be blown down. (Alternatively, the solid curves can be contracted to A_1 singularities prior to taking the double cover.)

Note that the quotient involves an affine E_8 diagram rather than an affine E_7 diagram. This is because the $\mathbb{Z}_2$ action on the "extra" curve in the Kodaira fiber (corresponding to the image of the affine node) has two fixed points, one giving an E_7 singularity and the other giving an A_1 singularity. The two together fit into an affine E_8 diagram.

We next treat case (4) leading to gauge symmetry G_2, depicted in Figure A.7.2. The bottom

	r	Type	φ	Description	Gauge Group	θ
(1)	any	$\frac{1}{r}(1,-1,0;0)$	$xy + z^n$	$A_{n-1} \xrightarrow{r\text{-to-}1} A_{rn-1}$	$SU(n)$	$\in \pi\mathbb{Q}$
(2)	4	$\frac{1}{4}(1,3,2;2)$	$x^2 + y^2 + z^{2n-1}$	$A_{2n} \xrightarrow{4\text{-to-}1} D_{2n+3}$	$Sp(n)$	π
(3)	2	$\frac{1}{2}(0,1,1;0)$	$x^2 + y^2 + z^{2n}$	$A_{2n-1} \xrightarrow{2\text{-to-}1} D_{n+2}$	$Sp(n)$	0
(4)	3	$\frac{1}{3}(0,1,2;0)$	$x^2 + y^3 + z^3$	$D_4 \xrightarrow{3\text{-to-}1} E_6$	G_2	0
(5)	2	$\frac{1}{2}(1,1,0;0)$	$x^2 + y^2 z + z^n$	$D_{n+1} \xrightarrow{2\text{-to-}1} D_{2n}$	$SO(2n+1)$	0
(6)	2	$\frac{1}{2}(1,0,1;0)$	$x^2 + y^3 + z^4$	$E_6 \xrightarrow{2\text{-to-}1} E_7$	F_4	0

Table A.7.1: Cyclic actions on ALE spaces

	Type	φ	(X, Y, T)	Φ
(3)	$\frac{1}{2}(0,1,1)$	$-y^2 + x^3 + ux^2 + wt^{2n}$	(t^2x, t^3y, t^2)	$-Y^2 + X^3 + uTX^2 + wT^{n+3}$
(4)	$\frac{1}{3}(1,0,2)$	$-y^2 + x^3 + wt^3$	(t^4x, t^6y, t^3)	$-Y^2 + X^3 + wT^5$
(6)	$\frac{1}{2}(0,1,1)$	$-y^2 + x^3 + wt^4$	(t^2x, t^3y, t^2)	$-Y^2 + X^3 + wT^5$

Table A.7.2: Group actions on Weierstrass models. Here u and w represent invariant functions of t which do not vanish at $t = 0$.

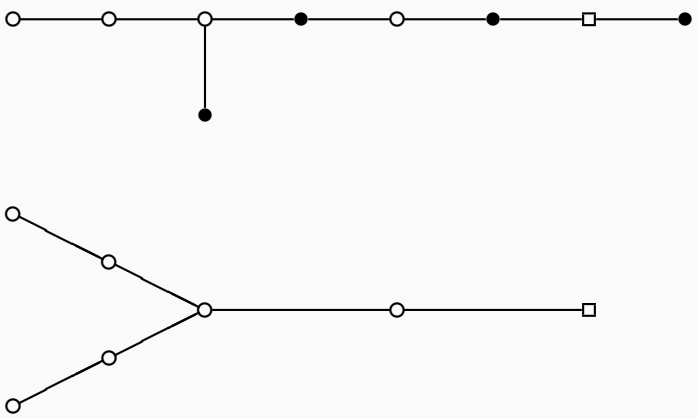


Figure A.7.1: F_4

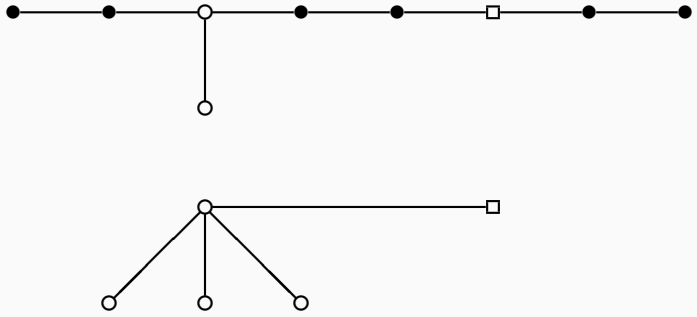


Figure A.7.2: G_2

half of the figure this time is $\widehat{D}_4$ with the affine node again represented by a square. This diagram has been obtained as a *triple* cover of the affine E_8 diagram shown in the top half of Figure A.7.2, in which the solid circles represent curves along which we branch as we construct a triple cover.

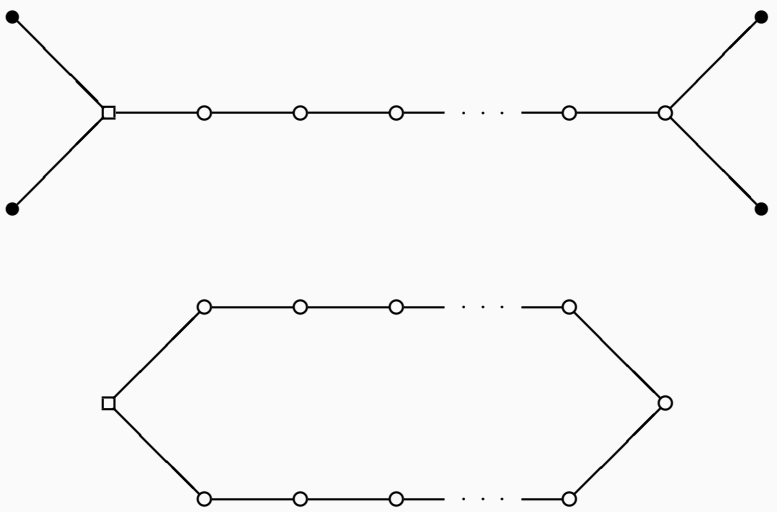


Figure A.7.3: $Sp(n)$, $\theta = 0$

(For an adjacent pair of solid curves, the branching is via $1/3$ on one curve and $2/3$ on the other.) The lowest curve in the top half of the diagram is triplicated in the bottom half because it doesn't meet the trivalent vertex at a branch point of the triple cover. After taking the triple cover, the solid curves have become -1, -2 pairs and are to be blown down. (Alternatively, the solid curves can be contracted to A_2 singularities prior to taking the triple cover.) Note that the $\mathbb{Z}_3$ action on the image of the affine node has two fixed points, one giving an E_6 singularity and the other giving an A_2 singularity; the two together fit into an affine E_8 diagram.

Now we consider case (3), depicted in Figure A.7.3. The bottom half is $\widehat{A}_{2n-1}$, obtained as a branched double cover of the top half, which is $\widehat{D}_{n+4}$. The square denotes the affine node (or its image), and the solid circles in the top half denote curves along which the double cover is branched. This time, the $\mathbb{Z}_2$ action on the the affine node has two fixed points unrelated to the singularity we are studying, creating the two extra solid curves at the left of the diagram.

Finally, we treat case (2), depicted in Figure A.7.4, which is the most complicated case. Here, we have a $\mathbb{Z}_4$ action on $\widehat{A}_{2n}$ at the bottom of the figure, and we describe the quotient in two stages: the quotient by the $\mathbb{Z}_2$ subgroup (shown in the middle of the figure as an $\widehat{A}_{4n+1}$ diagram) and the quotient by the full $\mathbb{Z}_4$ (shown at the top of the figure as a $\widehat{D}_{2n+5}$ diagram). The diagram at the top includes a resolved A_3 singularity (three connected solid dots), two resolved A_1 singularities (at the far left) which are to be branched along during the first double cover, and points marked by $\times$ which are A_1 singularities not branched during the first double cover (from top to middle) but branched during the second double cover (from middle to bottom).

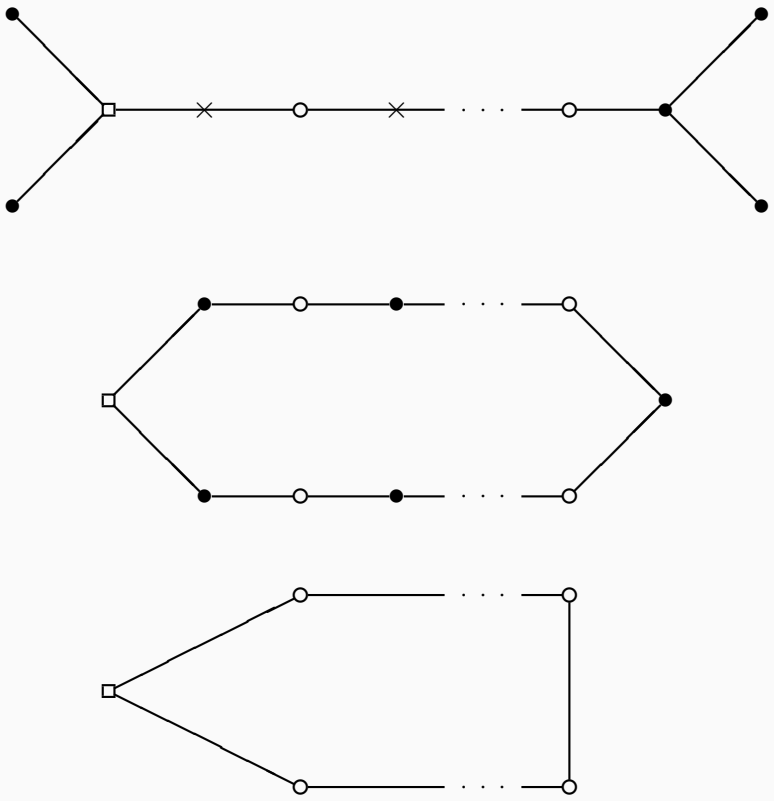


Figure A.7.4: $Sp(n)$, $\theta = \pi$

Appendix B

Appendix to Chapter 3

B.1 Dimension eight

There are three families [23] of vacua with 16 supercharges in dimension eight. The standard one has gauge algebra of rank 20, the next one has gauge algebra of rank 12, and the final one has gauge algebra of rank 4.

The rank-12 case was originally found in the context of heterotic string by Chaudhuri, Hockney and Lykken in [204] and is called the CHL string. An easy generalization leads to the rank-4 case. The moduli space of these systems and the possible enhancements of gauge algebras are studied in detail in [205].

In this appendix, we give an F-theory description of three cases: they are models on elliptically-fibered K3 with 0, 1, or 2 frozen seven-branes.

B.1.1 IIB with seven-branes

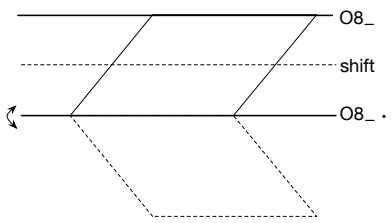
Let us start by the perturbative IIB setup on the orientifold $T^2/\mathbb{Z}_2$. We can either have zero $\mathrm{O7}_+$, one $\mathrm{O7}_+$ or two $\mathrm{O7}_+$:

(B.1.1)

with 16, 8 or 0 D7-branes, respectively. The first one, under T-duality, maps to 2 $\mathrm{O8}_-$ in type IIA, and then 1 $\mathrm{O9}_-$ in type IIB. The last one, under T-duality, maps to $\mathrm{O8}_-$ and $\mathrm{O8}_+$, or to a shift-orientifold of type IIA, and then a shift-orientifold of type IIB, without any D9-brane.

The second one is more peculiar. One T-duality should combine a pair of two $\mathrm{O7}_-$s to $\mathrm{O8}_-$, while the other pair of $\mathrm{O7}_-$ and $\mathrm{O7}_+$ to a shift orientifold. The resulting geometry is shown

below:



(B.1.2)

Namely, we consider a T^2 whose complex structure modulus is of the form $\tau \in \frac{1}{2}+i\mathbb{R}$, and take the $\mathbb{Z}_2$ flip along the horizontal axis. Then we have just one O8$_-$ locus and a shift-orientifold locus. Another T-duality leads to the $\mathrm{Spin}(32)/\mathbb{Z}_2$ bundle without vector structure [23].

B.1.2 F-theory interpretation

The F-theory representation of the rank-20 case is the standard F-theory compactification on the elliptically-fibered K3 surface.

The F-theory representation of the rank-12 case is given by an elliptic K3 compactification with a single frozen seven-brane.[1] We use projective coordinates $[z, w]$ on $\mathbb{CP}^1$ and locate the frozen brane at $z = 0$:

$$y^2 = x^3 + u_3(z, w)zx^2 + v_4(z, w)z^4x + w_5(z, w)z^7. \tag{B.1.3}$$

Here we have used the "Tate form" [95, 208] to present the equation, which involves arbitrary homogeneous polynomials u_3, v_4, and w_5 of the labeled degrees. By a change of variables, the equation can be put into Weierstrass form:

$$y^2 = \widehat{x}^3 + z^2(-\frac{1}{3}u_3^3 + z^2v_4)\widehat{x} + z^3(\frac{2}{27}u_3^3 - \frac{1}{3}z^2u_3v_4 + z^4w_5), \tag{B.1.4}$$

from which we can read off the equation of the discriminant locus

$$\Delta = z^{10}\left(4u_3^3w_5 - u_3^2v_4^2 - 18z^2u_3v_4w_5 + 4z^2v_4^3 + 27z^4w_5^2.\right) \tag{B.1.5}$$

Generically, in addition to the frozen seven-brane of type $\widehat{I_4^*}$ at $z = 0$, which makes no contribution to enhanced gauge symmetry, there are 14 additional zeros of the discriminant, which correspond to 14 seven-branes of type I_1 (i.e., 14 individual D7-branes) also contributing no enhanced gauge symmetry. Tuning the coefficients can lead to enhanced gauge symmetry.

The brane counting becomes clear if we explicitly include a Kodaira fiber of type I_0^* supporting an $\mathfrak{so}_8$ gauge algebra: this "uses up" 6 of the 14 D7-branes, but can be interpreted as an O7$_-$-plane on top of a stack of 4 D7-branes, which is the quantum splitting of the O7$_-$-plane [90].

The F-theory representation of the rank-4 8D vacuum with 16 supercharges involves two frozen seven-branes, which we can locate at $z = 0$ and $w = 0$, respectively. The equation for

[1] Note that this is a substantially different description than the ones proposed in [206] and [207], where a torsion flux on the base $\mathbb{CP}^1$ was proposed. It is possible that they are all dual descriptions.

these models (in Tate form) is

$$y^2 = x^3 + u_2(z,w)zwx^2 + v_0(z,w)z^4w^4x \tag{B.1.6}$$

with frozen brane-locus $\delta = zw$. Note that $v_0(z,w)$ is constant, and the x^0 term in the equation vanishes due to degree considerations. This implies that $(x,y) = (0,0)$ is a section which has order 2 in the Mordell–Weil group, and suggests a subtle modification of the F-theory gauge group.[2]

B.2 $\mathbb{CP}^1 \times \mathbb{CP}^1$ model and its flips via branes

B.2.1 Unflipped case

The original model considered by Bianchi–Sagnotti and Gimon–Polchinski was given in terms of Type I on $T^4/\mathbb{Z}_2$. It has $\mathrm{O9}_-$ with 16 D9s in the bulk, with 16 $\mathrm{O5}_-$ at the $\mathbb{Z}_2$ fixed points, and 16 D5s.[3]

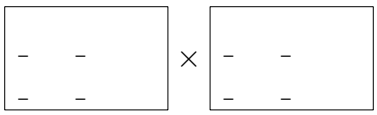
Let us determine its massless spectrum. From the bulk closed string modes, we have one supergravity multiplet, one tensor, and four neutral hypermultiplets. From the $\mathbb{Z}_2$ twisted closed strings, one neutral hypermultiplet arises from each $\mathbb{Z}_2$ singularity.

As for the open strings, $\mathrm{O9}_-$ wants to make the gauge algebra on D9 orthogonal. Therefore the bulk of the 9-brane has $\mathrm{Spin}(32)/\mathbb{Z}_2$ as the gauge group. But $\mathrm{O5}_-$ wants to make the gauge algebra on D9 symplectic. This gives a localized $\mathrm{Spin}(32)/\mathbb{Z}_2$ holonomy around the intersection point, and the massless gauge algebra on D9 that can remain is $\mathfrak{u}_{16}$, the intersection of $\mathfrak{sp}_{16}$ and $\mathfrak{so}_{32}$. This will keep charged hypermultiplets in $2 \cdot \wedge^2 16$. One can do the same analysis on the D5-branes, and get the same answer, when all the D5s are on a single $\mathrm{O5}_-$. Finally, the 5-9 strings give hypermultiplets in 16×16. The spectrum is then

- gauge algebras $\mathfrak{u}_{16} \times \mathfrak{u}'_{16}$,
- charged hypermultiplets in $2 \cdot \wedge^2 16 \oplus 16 \times 16' \oplus 2 \cdot \wedge^2 16'$,
- one supergravity multiplet, one tensor multiplet, and 20 neutral hypermultiplets.

Anomalies correctly cancel [97].

We can take T-duality along two directions and bring this model to the type IIB $T^4/\mathbb{Z}_2$ orientifolds with seven-branes, with the structure below:



(B.2.1)

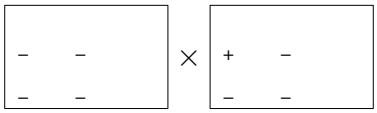
[2] We are assuming here that the torsion in the Mordell–Weil group is calculated for frozen F-theory models in the same way it is calculated for conventional F-theory models [209]. We leave detailed investigations of this for the future.

[3] Here the number of D-branes is counted in terms of Type IIB or Type IIA RR-charge, in a way invariant under T-duality. In simple orientifold models, this number equals the number of mobile D-branes or the rank of the gauge groups, but in more complicated models such as those discussed in this chapter, they can be different.

where the first T^2 has the coordinate u, the second has the coordinate v, with the orientifolding action sending $u \to -u$ and $v \to -v$ individually. The spectrum above are when all 16 D7s along v are on $u = 0$ and when all 16 D7s along u are on $v = 0$.

B.2.2 Singly-flipped case

For this, we consider the setup

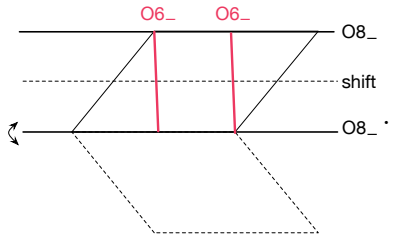


(B.2.2)

with 16 D7s perpendicular to the first T^2 and 8 D7s perpendicular to the second T^2.

To deduce the open string spectrum on the 8 D7s perpendicular to the second T^2, we just T-dualize one direction of the first T^2 and apply the rules of [74]. When the 8D7s are on a generic point, one gets $\mathfrak{sp}_4$ with a full antisymmetric tensor (both the traceless part and a singlet), and with 16 fundamentals. If they are on $\mathrm{O7}_-$, it gets enhanced to $\mathfrak{u}_8$ with $2 \cdot \wedge^2 \mathbf{8}$, and if they are on $\mathrm{O7}_+$, it gets enhanced to $\mathfrak{sp}_4 \times \mathfrak{sp}_4$ with a bifundamental.

For the 16 D7s on the first T^2, we can take the T-dual of the second T^2:



(B.2.3)

This T-duality was derived from the worldsheet point of view in [210].

When 16 D7s are on a single generic point on T^2, the T-dual is just 8 D6s suspended between two D8s that are in fact *the same* due to the funny geometry. This is $\mathfrak{sp}_4$ with one asym and 16 flavors. When they are all on an $\mathrm{O7}_-$, this gets enhanced to $\mathfrak{u}_8$ with $2 \cdot \wedge^2 \mathbf{8}$. Although we started from 32 Chan-Paton indices but we got just $\mathfrak{u}_8$. We give two other explanations to this somewhat unexpected fact:

- If we T-dualize the second torus twice, this describes instantons (or 5-branes) in the $\mathrm{Spin}(32)/\mathbb{Z}_2$ gauge fields on T^2 without vector structure. As discussed in [23], a minimal flat $\mathrm{Spin}(32)/\mathbb{Z}_2$ configuration without vector structure is in $\mathrm{SU}(2)$ embedded in $\mathfrak{so}_{32}$ as $\mathfrak{sp}_1 \times \mathfrak{sp}_8$. Then the instanton needs to be embedded into this $\mathfrak{sp}_8$; a single such instanton counts as two instantons in the original $\mathrm{Spin}(32)/\mathbb{Z}_2$. In other words, two small instantons of $\mathrm{Spin}(32)/\mathbb{Z}_2$ needs to move together.

- In the original 7-brane description, there are four intersections with transverse O7s; one is with $\mathrm{O7}_+$ and three are with $\mathrm{O7}_-$. The former has a monodromy that squares to -1 and the latter has a monodromy that squares to 1. But one cannot embed them into $\mathrm{O}(1)$: they

are not consistent, since the four monodromies need to multiply to one. To compensate this, one needs an additional flat $\mathrm{SO}(3)$ background on the 7-brane.

Summarizing, when 16 D7s perpendicular to the first T^2 are on a single $\mathrm{O7}_-$ and 8 D7s perpendicular to the second T^2 are on a single $\mathrm{O7}_+$, the spectrum is

- gauge algebras $\mathfrak{u}_8 \times \prod_{i=1,2}(\mathfrak{sp}_4)_i$,
- charged hypermultiplets in $2 \cdot \wedge^2 \mathsf{8} \oplus (\bigoplus_{i=1,2} \mathsf{8} \otimes \mathsf{8}_i) \oplus \mathsf{8}_1 \otimes \mathsf{8}_2$,
- one supergravity multiplet, 5 tensor multiplets, and 16 neutral hypermultiplets.

B.2.3 Doubly-flipped case

Let us finally consider

$$\begin{array}{|cc|}\hline + & - \\ - & - \\ \hline \end{array} \times \begin{array}{|cc|}\hline + & - \\ - & - \\ \hline \end{array} \tag{B.2.4}$$

with 8 D7-branes on each T^2. Using the analysis as in case II, we see that when 8 D7s are on a single $\mathrm{O7}_+$, the gauge algebra is $\mathfrak{sp}_2 \times \mathfrak{sp}_2$. Considering D7s on both T^2, we have $(\mathfrak{sp}_2)^4$ in total. The matter spectrum can be worked out as before:

- gauge algebras $\prod_{i=1}^4 (\mathfrak{sp}_2)_i$,
- charged hypermultiplets in $\bigoplus_{i<j} \mathsf{4}_i \otimes \mathsf{4}_j$,
- one supergravity multiplet, 7 tensor multiplets, and 14 neutral hypermultiplets.

The anomaly cancels; although there are 8 additional tensors, they do not participate in the gauge anomaly cancellation. This is as it should be, since they are localized on the intersections of $\mathrm{O7}_-$ and $\mathrm{O7}_+$, and bifundamentals are supported away from them.

Appendix C

Appendix to Chapter 4

C.1 Spin-TFTs from Rational spin-CFTs

Unitary rational conformal field theory is a rich source of examples of topological field theories in $2+1$ dimensions. The Hilbert space of the topological field theory is identified with the space of conformal blocks for the chiral algebra the conformal field theory.

If we endow the Riemann surface with a spin structure, we can consider conformal blocks for a chiral super-algebra $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{A}_1$, which includes both bosonic currents of integral spin and fermionic currents of half-integral spin. The conformal blocks for $\mathcal{A}$ can be naturally identified with the Hilbert space of a spin-TFT $\mathfrak{T}_s$.

We expect the shadow $\mathfrak{T}_f$ to be the TFT associated to the bosonic sub-algebra $\mathcal{A}_0$. Notice that $\mathcal{A}_1$ is a module for $\mathcal{A}_0$ and thus gives a quasi-particle in $\mathfrak{T}_f$, a fermion. As $\mathcal{A}_1$ currents fuse to $\mathcal{A}_0$ currents, the fermionic quasi-particle fuses to the identity. We identify it with Π. Thus fermionic anyon condensation is related to fermionic current algebra extensions, in the same way as the standard anyon condensation is related to standard current algebra extensions.

It should be possible to pursue this analogy further and derive from spin-RCFTs appropriate axioms for "super modular tensor categories".

We can give a few well-known examples of this construction.

C.1.1 Ising model and a chiral fermion

The Ising modular tensor category is naturally associated to the current algebra of a $c = \frac{1}{2}$ Virasoro minimal model.

The current algebra is generated by the stress tensor and can also be described as the coset $\frac{SU(2)_1 \times SU(2)_1}{SU(2)_2}$. It has three modules, which we can denote as $M_1 \equiv \mathcal{A}_0$, M_σ and M_ψ, of conformal dimension 0, $\frac{1}{16}$ and $\frac{1}{2}$. The latter is associated to the fermionc quasi-particle of the Ising modular tensor category.

We can consider the super-algebra $\mathcal{A}$ consisting of $\mathcal{A}_0$ and $\mathcal{A}_1 = M_\psi$. This is simply the algebra generated by a free chiral fermion $\psi(z)$. This algebra has a single Neveu-Schwarz (NS) module, $\mathcal{A}$ itself. It corresponds to the identity quasi-particle in the bulk spin-TFT. On the other hand, M_σ is a Ramond module. It must lie at the end of a bulk defect with a non-bounding spin structure.

The Ising 3d TFT is the shadow of the spin-TFT associated to a free chiral fermion. We could denote it as $\mathfrak{T}_s^\psi$.

C.1.2 Π-product of Ising models and multiple chiral fermions

In order to find the shadow of the product of two free chiral fermions, we are supposed to gauge the $\mathbb{Z}_2$ 1-form symmetry generated by $\Pi\Pi'$. This is the same as a simple current extension.

Consider two copies of the $c = \frac{1}{2}$ Virasoro minimal model. The module corresponding to $\Pi\Pi'$ is $M_\psi \otimes M_{\psi'}$. We thus consider the chiral algebra $\mathcal{A}_0 = M_1 \otimes M_{1'} \oplus M_\psi \otimes M_{\psi'}$. In other words, the chiral algebra generated by $\psi\partial\psi$, $\psi\psi'$ and $\psi'\partial\psi'$.

By bosonization, we identify that with the algebra $\mathcal{A}_0$ defined by a free boson current $\partial\phi$ and vertex operators $e^{2ni\phi}$, of dimension $2n^2$. In other words, this is the $U(1)_4$ current algebra. It has 4 modules generated respectively by 1, $e^{\frac{i\phi}{2}}$, $\psi = e^{i\phi}$ and $e^{-\frac{i\phi}{2}}$.

The chiral super-algebra generated by $\mathcal{A}_0$ and by ψ can be identified with $U(1)_1$ and is associated to the simplest spin Chern-Simons TFT. Again, it has a single quasi-particle, the identity, and an extra Ramond line defect associated to the module generated by $e^{\frac{i\phi}{2}}$. We can also identify it as the square of $\mathfrak{T}_s^\psi$.

More generally, a set of N free chiral fermions has bosonic 1-form symmetry generators $\psi_i\psi_j$. Adjoining them to the identity module gives us the $SO(N)_1$ WZW current algebra, with a module M_ψ generated by the ψ_i and one or two modules generated by twist fields, associated to the spinor representation(s). This is the shadow of the N-th power of $\mathfrak{T}_s^\psi$.

C.1.3 $U(1)_{4k}$ Chern-Simons theories

Consider the current algebra $\mathcal{A}_0 = U(1)_{4k}$ for odd k, generated by the bosonic current $\partial\phi$ and vertex operators $e^{2\sqrt{k}ni\phi}$. This is associated to an $U(1)_{4k}$ Chern-Simons theory.

The algebra $\mathcal{A}_0$ has modules M_m generated by $e^{\frac{m}{2\sqrt{k}}i\phi}$, for $-2k < m \leq 2k$. In particular, $M_\Pi \equiv M_{2k}$ is generated by $e^{\sqrt{k}i\phi}$, which has half-integral dimension $k/2$.

The chiral super-algebra $\mathcal{A}$ generated by $\partial\phi$ and vertex operators $e^{\sqrt{k}ni\phi}$ is associated to an $U(1)_k$ spin Chern-simons theory. It has NS modules generated by $e^{\frac{m}{\sqrt{k}}i\phi}$ for $-k/2 < m < k/2$ and Ramond modules generated by $e^{\frac{2m+1}{2\sqrt{k}}i\phi}$.

In particular, $U(1)_{4k}$ is the shadow of $U(1)_k$.

C.2 G-equivariant toric code

C.2.1 Symmetries of the toric code and their anomalies

The toric code (also known as $\mathbb{Z}_2$ topological gauge theory in 2+1d with a trivial Dijkgraaf-Witten class) can be described by a Euclidean action

$$S_{toric} = \pi i \int_M b \cup \delta a, \tag{C.2.1}$$

where a and b are $\mathbb{Z}_2$-valued 1-cochains on a triangulation of an oriented 3-manifold M. One may call a the gauge field, then b is a Lagrange multiplier field imposing the constraint $\delta a = 0$. The model has $\mathbb{Z}_2 \times \mathbb{Z}_2$ 0-form gauge symmetry:

$$a \mapsto a + \delta\lambda_a, \quad b \mapsto b + \delta\lambda_b, \quad \lambda_a, \lambda_b \in C^0(M, \mathbb{Z}_2). \tag{C.2.2}$$

As for global symmetries, the toric code has a $\mathbb{Z}_2$ 0-form global symmetry F_0 exchanging a and b. We will call it particle-vortex symmetry, since the Wilson line for a represents an electrically charged particle, while the Wilson line for b represents a vortex excitation. This symmetry is not manifest in the action since the cup product is not supercommutative on the cochain level. There is also $\mathbb{Z}_2 \times \mathbb{Z}_2 = F_1$ 1-form global symmetry

$$a \mapsto a + \alpha, \quad b \mapsto b + \beta, \quad \alpha, \beta \in Z^1(M, \mathbb{Z}_2). \tag{C.2.3}$$

Crucially, F_0 acts on F_1 by a nontrivial automorphism exchanging α and β. The combined symmetry is described by a 2-group (or equivalently a crossed module) [211]. In general, the equivalence class of a 2-group $\mathbb{F}$ is described by a pair of groups F_0, F_1, where F_1 is abelian, an action π of F_0 on F_1, and a Postnikov class $\gamma \in H^3_\pi(F_0, F_1)$. The Postnikov class describes the failure of the fusion of F_0 domain walls to be associative "on the nose". In the case of the toric code we can use Shapiro's lemma [212] to compute $H^n_\pi(F_0, F_1) = H^n(\ker \pi, \mathbb{Z}_2) = 0$ for $n > 0$. Hence the Postnikov class is necessarily trivial.

Global symmetries may have 't Hooft anomalies, i.e. obstructions to gauging. Such anomalies can always be canceled by coupling the theory to a topological gauge theory in one dimension higher. Thus anomalies are classified by topological actions for the gauge fields in one dimension higher. In the case of a 2-group symmetry, such actions have been classified in [211]. The gauge fields are a 1-form F_0 gauge field A and a 2-form F_1 gauge field $\mathsf{B} = (B_a, B_b)$. More precisely, A is a 1-cocycle with values in F_0, while B is a twisted 2-cocycle with values in a local system (i.e. flat bundle) with fiber F_1. The twist arises from the fact that F_0 acts nontrivially on F_1. The most general action in 4d representing the anomaly for a 2-group symmetry is

$$S_{anomaly} = 2\pi i \int_{M_4} \left(\mathfrak{P}_q(\mathsf{B}) + \langle \mathsf{B}, \cup c_2(\mathsf{A})\rangle + \omega_4(\mathsf{A})\right). \tag{C.2.4}$$

The notation is as follows. We regard the pair (A, B) as map from M_4 to the classifying space $B\mathbb{F}$ of the 2-group, which is a bundle over BF_0 with fiber BF_1. The action depends on a quadratic function $q : F_1 \to \mathbb{R}/\mathbb{Z}$ invariant under the F_0 action, a class $c_2 \in H^2_\pi(F_0, F_1^*)$, and a class

$\omega_4 \in H^4(F_0, \mathbb{R}/\mathbb{Z})$. $\mathfrak{P}$ denotes the Pontryagin square (a cohomological operation associated to the quadratic function q which maps a twisted 2-cocycle B to an $\mathbb{R}/\mathbb{Z}$-valued 4-cocycle $\mathfrak{P}(\mathsf{B})$). In the case of interest to us, both $H^2_\pi(F_0, F_1^*)$ and $H^4(F_0, \mathbb{R}/\mathbb{Z})$ vanish (the former by Shapiro's lemma again), so there are no anomalies for F_0 or a mixed anomaly between F_0 and F_1. On the other hand, there exist F_0-invariant quadratic functions on F_1, so the 1-form symmetry F_1 could be anomalous. In fact, it is easy to see that the anomaly is nontrivial and corresponds to the quadratic function

$$q : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{R}/\mathbb{Z}, \quad q : (x, y) \mapsto \frac{1}{2}xy. \tag{C.2.5}$$

Indeed, let the F_0 gauge field A be trivial, so that B is an ordinary 2-cocycle (B_a, B_b) with values in $F_1 = \mathbb{Z}_2 \times \mathbb{Z}_2$. If we perform the shifts (C.2.3) with not-necessarily-closed 1-cochains α and β, the action (C.2.1) transforms as follows:

$$S_{toric} \mapsto S_{toric} + \pi i \int_M (b \cup \delta\alpha + \beta \cup \delta a + \beta \cup \delta\alpha). \tag{C.2.6}$$

To cancel the terms which depend on b and a we couple the action to the 2-form gauge fields B_a and B_b which transform as $B_a \mapsto B_a + \delta\alpha$, $B_b \mapsto B_b + \delta\beta$ and define

$$S'_{toric} = \pi i \int_M (b \cup \delta a + b \cup B_a + B_b \cup a). \tag{C.2.7}$$

The new action transforms as

$$S'_{toric} \mapsto S'_{toric} + \pi i \int_M (\beta \cup B_a + B_b \cup \beta + \beta \cup \alpha), \tag{C.2.8}$$

which is precisely the boundary term in the variation of

$$\pi i \int_{M_4} B_b \cup B_a. \tag{C.2.9}$$

This is nothing but the Pontryagin square of $\mathsf{B} \in Z^2(M_4, F_1)$ for the quadratic function (C.2.5).[1]

To summarize, the anomaly action for the toric code is

$$S_{anomaly} = 2\pi i \int_{M_4} \mathfrak{P}_q(\mathsf{B}). \tag{C.2.10}$$

In particular, the anomaly for the diagonal subgroup of F_1 is obtained by letting $B_a = B_b = B$. Note that this subgroup is F_0-invariant. The corresponding anomaly action is

$$S_{anomaly} = \pi i \int_{M_4} B \cup B, \tag{C.2.11}$$

which means that the toric code is a shadow of a fermionic theory. The worldline ot the corre-

[1] In general, if A is nontrivial, B is a twisted 2-cocycle, and the Pontryagin square for twisted cocycles is more difficult to write down.

sponding fermion Π is represented by the Wilson line $\exp(i\pi \int (a+b))$.

C.2.2 G-equivariant toric code

We are now ready to promote the toric code to a G-equivariant toric code, i.e. couple it to a G gauge field A. Mathematically, this means embedding G into the symmetry of the toric code. Since the "target" symmetry is a 2-group $\mathbb{F}$ rather than a group, this means specifying a homotopy class of maps $BG \to B\mathbb{F}$. Physically, we make the fields A and B functions of A so that under G gauge transformations A and B transform by F_0 and F_1 gauge transformations, respectively. Such am embedding is characterized by a homomorphism $\pi : G \to F_0$ and a cohomology class $[\Lambda_2] \in H^2_\pi(G, F_1)$. That is, we set

$$\mathsf{A} = \pi(A), \quad \mathsf{B} = \Lambda_2(A). \tag{C.2.12}$$

The corresponding anomaly action is obtained by substituting into (C.2.10):

$$S_{anomaly} = \int_{M_5} \mathfrak{P}_q(\Lambda_2(A)). \tag{C.2.13}$$

This means that the symmetry G is free of 't Hooft anomalies if and only if $\mathfrak{P}_q(\Lambda_2(A))$ is cohomologically trivial, i.e. if and only if there exists a 3-cochain $\nu_3 \in Z^3(G, \mathbb{R}/\mathbb{Z})$ such that

$$\delta\nu_3 = \mathfrak{P}_q(\Lambda_2). \tag{C.2.14}$$

Then G gauge-invariance can be restored by modifying the 3d action by a term $2\pi i \int_M \nu_3(A)$.

Let us specialize this to the case of trivial π. Then A is trivial, and $\mathsf{B} = \Lambda_2(A)$ is an ordinary (not twisted) 2-cocycle on M with values in $F_1 = \mathbb{Z}_2 \times \mathbb{Z}_2$. We can write $\Lambda_2 = (\beta_2^a, \beta_2^b)$, where $\beta_2^a, \beta_2^b \in Z^2(G, \mathbb{Z}_2)$. The condition (C.2.14) simplifies to

$$\delta\nu_3 = \frac{1}{2}\beta_2^b \cup \beta_2^a. \tag{C.2.15}$$

The action of the equivariant toric code in this case is

$$2\pi i \int_M \left(\frac{1}{2} b \cup \delta a + \frac{1}{2} b \cup \beta_2^a(A) + \frac{1}{2}\beta_2^b \cup a + \nu_3(A) \right). \tag{C.2.16}$$

Regarding b as a Lagrange multiplier, we see that it imposes a constraint $\delta a = \beta_2^a(A)$. This means that the pair (A, a) is a 1-cocycle with values in $\widehat{G}$, where $\widehat{G}$ is a central extension of G by $\mathbb{Z}_2$ whose extension class is β_2^a. The part of the action independent of b can then be interpreted as an integral of a pull-back of a 3-cocycle $\widehat{\nu}_3$, where $\widehat{\nu}_3$ is given by

$$\widehat{\nu}_3 = \nu_3 + \frac{1}{2}\beta_2^b \cup \epsilon, \tag{C.2.17}$$

where ϵ is a $\mathbb{Z}_2$-valued 1-cochain on $\widehat{G}$ trivializing the pull-back of β_2^a and restricting to the identity on the central $\mathbb{Z}_2$ subgroup of $\widehat{G}$. The corresponding fusion category is a twisted Drinfeld

double of $\widehat{G}$, with the twist given by $\widehat{\nu}_3$. Essentially, we have shown that this is the most general G-equivariant extension of the toric code where G acts trivially on the toric code quasi-particles (none of the elements of G exchange e and m). Note that the model considered in section (4.3.4) has this general form, but in addition has $\beta_2^a = \beta_2^b = \beta_2$. We will see shortly that this constraint arises if we require the model to contain a fermion.

Now suppose π is nontrivial. Let $G_0 = \ker \pi$. It is easy to check that the action of G on $F_1 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ is induced from the trivial action of G_0 on $\mathbb{Z}_2$. .[2] Therefore by Shapiro's lemma $H^2_\pi(G, F_1) \simeq H^2(G_0, \mathbb{Z}_2)$. Thus for nontrivial π G-equivariant extensions of the toric code are labeled by a central extension of G_0 together with a trivialization ν_3 of the corresponding Pontryagin square $\mathfrak{P}(\Lambda_2)$. The model considered in section (4.3.6) is of this form. Below we will determine the condition on Λ_2 imposed by the existence of a fermion.

C.2.3 One-form symmetries of the equivariant toric code

Consider enlarging the symmetry group G to a 2-group $\mathbb{G}$, such that the group of 1-form symmetries is $\mathbb{Z}_2$. Since $\mathbb{Z}_2$ has no nontrivial automorphisms, the 2-group structure of $\mathbb{G}$ is controlled by a Postnikov class $\Gamma \in H^3(G, \mathbb{Z}_2)$. Enhancing the symmetry of the toric code from G to $\mathbb{G}$ involves extending the map $BG \to B\mathbb{F}$ to a map $B\mathbb{G} \to B\mathbb{F}$. Since the Postnikov class γ of $\mathbb{F}$ is trivial, this is only possible if Γ is trivial. Physically, since the fusion of F_0 domain walls in the toric code is associative "on the nose", this remains true even after we reinterpret them as G domain walls via a homomorphism $\pi : G \to F_0$.

Specifying the homotopy class of a map $B\mathbb{G} \to B\mathbb{F}$ is equivalent to specifying B and A as functions of $B \in Z^2(M, \mathbb{Z}_2)$ and $A \in Z^1(M, G)$ in a way compatible with gauge transformations of B and A. This means

$$\mathsf{A} = \pi(A), \quad \mathsf{B} = \Lambda_2(A) + \rho(B), \tag{C.2.18}$$

where π and Λ_2 are as before, and ρ is a nonzero homomorphism from $\mathbb{Z}_2$ to F_1 which is invariant with respect to the G action on F_1 induced by $\pi : G \to F_0$.

The anomaly for the 2-group $\mathbb{G}$ is obtained by substituting the expressions for B and A into (C.2.10). Using the properties of the Pontryagin square, we get

$$S_{anomaly} = 2\pi i \int_{M_4} \left(\mathfrak{P}_q(\Lambda_2(A)) + b_q(\Lambda(A), \cup\rho(B) + \frac{1}{2}\rho(B) \cup \rho(B) \right), \tag{C.2.19}$$

where b_q is an F_0-invariant bilinear form on F_1 associated to the quadratic function q. Explicitly:

$$b_q(x_1, y_1; x_2, y_2) = \frac{1}{2}(x_1 y_2 + y_1 x_2). \tag{C.2.20}$$

We will assume as before that anomalies for G are absent, i.e. $\mathfrak{P}(\Lambda_2)$ is cohomologically trivial. Then the first term in (C.2.19) is exact. The second term describes the mixed anomaly between $\mathbb{Z}_2$ 1-form symmetry and G, so it must be exact for $\mathbb{Z}_2$ to be a global 1-form symmetry of the G-equivariant toric code. This means that the cohomology class of $\rho(1)$ must be orthogonal to

[2] We are grateful to V. Ostrik for pointing this out.

the cohomology class of Λ_2 with respect to b_q. Finally, the last term describes the anomaly of the 1-form $\mathbb{Z}_2$ symmetry.

Let us focus on fermionic $\mathbb{Z}_2$ 1-form symmetries. Such symmetries must have a nontrivial anomaly, so $q(\rho(1)) \neq 0$. This uniquely fixes ρ : it must send 1 to the generator of $F_1^{F_0}$. That is, the fermion must be represented by a Wilson line $\exp(\pi i \int (a+b))$. The orthogonal complement of $F_1^{F_0}$ is $F_1^{F_0}$ itself, therefore the mixed anomaly is absent if and only if $[\Lambda_2] \in H^2_\pi(G, F_1)$ is in the image of the map $H^2(G, F_1^{F_0}) \to H^2_\pi(G, F_1)$. That is, we must have

$$\Lambda_2 = \tilde{\Lambda}_2 + \delta_\pi \tilde{\psi}_1, \tag{C.2.21}$$

where $\tilde{\Lambda} \in Z^1(G, F_1^{F_0})$ and $\tilde{\psi}_1 \in C^1_\pi(G, F_1)$. The anomaly action for symmetry G is now exact, and if we let $M_4 = \partial M_3$, it becomes

$$S_{anomaly} = 2\pi i \int_{M_3} \Big(\nu_3(A) + b_q(\tilde{\psi}_1(A), \rho(B))\Big) + \pi i \int_{M_4} B \cup B. \tag{C.2.22}$$

The first term is not invariant under 0-form and 1-form gauge symmetries, but it does not represent a true anomaly: it can be removed by modifying the action of the equivariant toric code by a local counterterm

$$S_3^{ct}(A, B) = -2\pi i \int_{M_3} \Big(\nu_3(A) + b_q(\tilde{\psi}_1(A), \rho(B))\Big). \tag{C.2.23}$$

The last term in (C.2.22) is the correct anomaly for the 1-form $\mathbb{Z}_2$ symmetry to be fermionic.

The 1-cochain $\tilde{\psi}_1 \in C^1(G, F_1)$ does not affect the cohomology class of the B-field and can be removed by a 1-form gauge transformation with a parameter $-\tilde{\psi}_1$. (This transformation also shifts ν_3). Then $\Lambda_2 = \tilde{\Lambda}_2 \in Z^2(G, \mathbb{Z}_2)$, and the constraint on ν_3 simplifies:

$$\delta\nu_3 = \mathfrak{P}_q(\tilde{\Lambda}_2) = \frac{1}{2}\tilde{\Lambda}_2 \cup \tilde{\Lambda}_2. \tag{C.2.24}$$

This is nothing but the Gu-Wen equation. The counterterm action also takes a simple form:

$$S_3^{ct}(A, B) = -2\pi i \int_{M_3} \nu_3(A). \tag{C.2.25}$$

We still retain the ability to perform 1-form symmetry transformations valued in $F_1^{F_0} \simeq \mathbb{Z}_2$. Indeed, while such transformations shift $\tilde{\psi}_1$, they do not affect S_3^{ct} (C.2.23), since $F_1^{F_0}$ is an isotropic subgroup of F_1. Under a transformation with a parameter $\lambda_1 \in C^1(G, \mathbb{Z}_2)$ the data $(\nu_3, \tilde{\Lambda}_2)$ transform as follows:

$$\nu_3 \mapsto \frac{1}{2}\lambda_1 \cup \delta\lambda_1, \quad \tilde{\Lambda}_2 \mapsto \tilde{\Lambda}_2 + \delta\lambda_1. \tag{C.2.26}$$

Changing ν_3 by exact cocycles also does not affect the action.

We conclude that G-equivariant versions of the toric code with a fermionic $\mathbb{Z}_2$ 1-form symmetry are labeled by triples $(\pi, \tilde{\Lambda}_2, \nu_3)$, where π is a homomorphism $G \to \mathbb{Z}_2$ and $(\nu_3, \tilde{\Lambda}_2) \in$

$C^3(G, U(1)) \times Z^2(G, \mathbb{Z}_2)$ satisfy the Gu-Wen equations. We also described the identifications on this set which do not affect the model.

C.3 Wen plaquette model

The Wen placquette model for the toric code is defined on a square lattice with $\mathbb{Z}_2$ variables at each site. The commuting projectors are $P_{i,j} = \sigma^x_{i,j}\sigma^y_{i+1,j}\sigma^y_{i,j+1}\sigma^x_{i+1,j+1}$, associated to the plaquettes of the lattice. This is a realization of the toric code.

The model realizes the $\mathbb{Z}_2$ symmetry of the equivariat toric code, but not in an on-site manner: the $\mathbb{Z}_2$ symmetry maps to translations of the lattice by one unit.

The e and m quasi-particles are described by switching the sign of a plaquette at even or odd locations on the lattice. The corresponding string operators can be taken to be, say, products of the form $\sigma^x_{i,j}\sigma^y_{i,j+1}\sigma^x_{i,j+2}\cdots$ which create a particle in the plaquette to the left and below the beginning of the string. A similar effect is archived by $\sigma^y_{i,j}\sigma^x_{i,j+1}\sigma^y_{i,j+2}\cdots$.

The combination $\sigma^z_{i,j}\sigma^z_{i,j+1}\sigma^z_{i,j+2}\cdots$ creates a pair of e and m particles at neighbouring plaquettes, i.e. an ϵ particle.

C.3.1 Fermionic "boundary condition"

Consider the model restricted to the upper half-plane. The bulk plaquettes do not gap the system completely. At the boundary, degrees of freedom survive which roughly correspond to a $\mathbb{Z}_2$ spin chain of twice the lattice spacing: the boundary operators $S_i = \sigma^y_{i,0}\sigma^x_{i+1,0}$ commute with all bulk plaquette operators. They anti-commute with nearest neighbours and commute with all others.

The $\mathfrak{B}_e$ gapped boundary condition where e condenses is easily described: we can add $\sum_i S_{2i}$ to the Hamiltonian. This commute with e string operators ending on the boundary. Similarly, we obtain $\mathfrak{B}_m$ by adding $\sum_i S_{2i+1}$ to the Hamiltonian. This commute with m string operators ending on the boundary. [3]

Both choices break explicitly the translation symmetry along the boundary by one unit, which maps one into the other. This is compatible with the action of the $\mathbb{Z}_2$ symmetry of the toric code. A boundary condition defined by a bosonic boundary Hamiltonian which preserves the translation symmetry will flow to a $\mathbb{Z}_2$-symmetric boundary condition for the equivariant toric code. If gapped, it must coincide with a direct sum of $\mathfrak{B}_e$ and $\mathfrak{B}_m$, i.e. it must spontaneously break the $\mathbb{Z}_2$ symmetry.

Simple choices of bosonic boundary conditions, such as adding $\sum_i S_i$ to the Hamiltonian, leave the $\mathbb{Z}_2$ symmetry unbroken and give a gapless critical Ising model at the boundary. The Ising model is coupled to the bulk toric code in a straightforward manner: the bulk $\mathbb{Z}_2$ gauge theory couples to the Ising symmetry of the gapless theory. In particular, e lines can end on $\sigma(z, \overline{z})$

[3] In a gauge theory description, these boundary conditions are either Dirichlet, i.e. fix the connection at the boundary, or Neumann, i.e. leave the connection free to fluctuate at the boundary.

operators, m lines on the dual $\mu(z,\overline{z})$ operators and ϵ lines on the fermionic local operators $\psi(z)$ and $\overline{\psi}(\overline{z})$.

As the result of fermionic anyon condensation of ϵ is the root $\mathbb{Z}_2$ fermionic SPT phase, we should be able to produce a boundary condition $\mathfrak{B}_\epsilon$ at which ϵ ends and the $\mathbb{Z}_2$ symmetry is broken. The boundary condition should be related to a $\mathbb{Z}_2$-invariant interface between the toric code and the root $\mathbb{Z}_2$ fermionic SPT phase.

We can define $\mathfrak{B}_\epsilon$ as follows. First, we add a Majorana mode γ_i at each boundary lattice site. Next, we add the following commuting projectors to the Hamiltonian: $\sum_i \gamma_i\gamma_{i+1}S_i$. This Hamiltonian gaps the system. Indeed, if we fermionize locally the boundary Ising degrees of freedom in terms of new Majoranas c_i, the commuting projectors become $\sum_i \gamma_i\gamma_{i+1}c_ic_{i+1}$ and we have massive ground states where all the $c_i\gamma_i$ have the same sign.

The boundary Hamiltonian commutes with the ϵ line defects $c_i\sigma^z_{i,j}\sigma^z_{i,j+1}\sigma^z_{i,j+2}\cdots$ ending at the boundary. This motivates the identification with $\mathfrak{B}^\epsilon$. [4]

Although the new Hamiltonian is naively invariant under translations along the boundary, in oder to define the actual Hilbert space we need to pair up the Majorana modes in some manner. This breaks the $\mathbb{Z}_2$ symmetry, as expected.

There is a neat way to restore it: we can place on the lower half plane some gapped system which has Majorana boundary oscillators. An example is an infinite collection of Kitaev chains extended along the vertical direction. Concretely, we can put a pair of Majorana modes $\gamma_{i,j}$ and γ'_{ij} at each site in the lower half plane and build the Hamiltonian with projectors $\gamma'_{i,j}\gamma_{i,j-1}$. These boundary plaquettes commute with the $\gamma_i \equiv \gamma_{i,-1}$ oscillators used in the boundary Hamiltonian.

This is a microscopic description of the expected gapped interface between the root $\mathbb{Z}_2$ SPT phase and the equivariant toric code.

C.3.2 Anyon condensation

It is straightforward to implement the bosonic anyon condensation in the plaquette model.

We can populate the lattice with an arbitrary number of e particles by removing from the Hamiltonian the odd plaquettes, using the Hamiltonian $-\sum_{i,j|i+j\text{even}} P_{ij}$.

The "edge operators" U^b can be taken to be σ^x_{ij} for even $i+j$ and σ^y_{ij} for odd ij: they commute with the Hamiltonian and move or annihilate e particles along diagonals in the lattice. They clearly all commute.

Adding the U^b edge operators to the Hamiltonian eliminates all the spin degrees of freedom and returns the trivial theory, as expected.

Fermionic anyon condensation is a bit more subtle. We can populate the lattice with an

[4] In a gauge theory language, we expect $\mathfrak{B}^\epsilon$ to correspond to a deformed Neumann boundary condition, with extra boundary action given by the quadratic refinement of the intersection pairing associated to a spin structure on the boundary.

arbitrary number of ϵ particles if we use the Hamiltonian

$$H^\epsilon = \sum_{i,j} P_{2i,j} P_{2i+1,j} = \sum_{ij} \sigma^x_{2i,j} \sigma^z_{2i+1,j} \sigma^y_{2i+2,j} \sigma^y_{2i,j+1} \sigma^z_{2i+1,j+1} \sigma^x_{2i+2,j+1} \tag{C.3.1}$$

We can visualize the ϵ particles as living in the middle of vertical edges at odd horizontal locations.

The operator $\sigma^z_{2i+1,j}$ commute with the new Hamiltonian but anti-commute with the four P operators around the vertex. It moves an ϵ particle vertically by one unit.

The operator $\sigma^x_{2i-1,j} \sigma^z_{2i} \sigma^y_{2i+1}$ anti-commutes with the four P operators above it. It moves an ϵ particle horizontally by one unit.

As expected these V^ϵ operators do not commute: vertical operators commute with each other, but horizontal operators anti-commute with horizontal neighbours and vertical operators immediately below.

We can add Majorana pairs. It is convenient to denote them as $\gamma_{2i+1,j}$ and $\gamma'_{2i+1,j}$.

The dressed hopping operators U^f take the form: $U^f_{2i+1,j} \equiv i\sigma^z_{2i+1,j} \gamma_{2i+1,j} \gamma'_{2i+1,j}$ and $U^f_{2i,j} \equiv i\sigma^x_{2i-1,j} \sigma^z_{2i,j} \sigma^y_{2i+1,j} \gamma_{2i-1,j} \gamma_{2i+1,j}$.

The U^f operators square to 1. The product of the operators around a closed path is

$$P_{2i-1,j-1} P_{2i,j-1} \gamma'_{2i-1,j} \gamma_{2i-1,j-1} \gamma'_{2i+1,j} \gamma_{2i+1,j-1} \tag{C.3.2}$$

This will become 1 as soon as we impose the Coulomb branch constraints

$$C_{2i-1,j-1} \equiv P_{2i-1,j-1} \gamma'_{2i-1,j} \gamma_{2i-1,j-1} = i(-1)^i \tag{C.3.3}$$

Overall, we only need to impose the $C_{2i+1,j+1}$, $U^f_{2i+1,j}$ and $U^f_{2i,j}$ projectors, as they imply the original $P_{2i,j} P_{2i+1,j} = 1$ constraints.

It is straightforward, if tedious, to show that we can use the $U^f_{2i+1,j} = 1$ and $U^f_{2i,j} = 1$ constraints to gauge-fix the spin variables. In terms of dressed fermionic operators

$$\Gamma_{2j+1,j} = \gamma_{2j+1,j} \sigma^x_{2i+1,j} \sigma^y_{2i+2,j} \qquad \Gamma'_{2j+1,j} = \gamma'_{2j+1,j} \sigma^y_{2i+1,j} \sigma^x_{2i+2,j} \tag{C.3.4}$$

commuting with the U^f projectors, the Gauss law constraints involve $\Gamma'_{2j+1,j} \Gamma_{2j+1,j-1}$ and make the system into a collection of vertical Kitaev chains.

Appendix D

Appendix to Chapter 6

D.1 Group cohomology

In this appendix, we collect various standard facts about group cohomology.

Definition: Given a finite group G and its module A, we define n-cochains $C^n(G, A)$ as functions $G^n \to A$. The differential is given by

$$df(g_1, \ldots, g_{n+1}) = g_1 f(g_2, \ldots, g_{n+1}) + \sum_{i=1}^{n} (-1)^i f(g_1, \ldots, g_i g_{i+1}, \ldots, g_{n+1}) + (-1)^{n+1} f(g_1, \ldots, g_n). \quad \text{(D.1.1)}$$

The differential squares to zero: $d^2 = 0$. Then we define the group cohomology $H^i(G, A)$ as the cohomology of this differential. Explicitly, the first few differentials are given by

$$df(g, h) = gf(h) - f(gh) + f(g), \quad \text{(D.1.2)}$$
$$df(a, b, c) = af(b, c) - f(ab, c) + f(a, bc) - f(a, b), \quad \text{(D.1.3)}$$
$$df(x, y, z, w) = xf(y, z, w) - f(xy, z, w) + f(x, yz, w) - f(x, y, zw) + f(x, y, z). \quad \text{(D.1.4)}$$

Some points on notation: It does not lead to any loss of generality if we assume that every cochain/cocycle/coboundary is *normalized*, i.e. it is zero whenever at least one of $g_i = 1$. See e.g. [155]. We have assumed throughout the chapter that every cochain is normalized. We are often interested in $H^i(G, \mathrm{U}(1))$ for $i = 2, 3$ where the action of G on $\mathrm{U}(1)$ is taken to be trivial. Henceforth, we will assume the trivial action whenever we write $\mathrm{U}(1)$. It is also convenient to treat $\mathrm{U}(1)$ elements as phases and in this case the $+$ sign in above definitions should be replaced by the usual multiplication of phases. For instance we have,

$$df(g, h) = \frac{f(h)f(g)}{f(gh)}. \quad \text{(D.1.5)}$$

We have used the product notation throughout the chapter in the context of group cohomology valued in $\mathrm{U}(1)$.

Pull-back: Recall that given a map $M_1 \to M_2$ between two manifolds, one can *pull-back* n-forms on M_2 to n-forms on M_1. The analogous statement in group cohomology is that given a group homomorphism $G \to G'$, we obtain a module homomorphism $H^i(G', A) \to H^i(G, A)$. Explicitly, let $h : G \to G'$, then $\tilde{h} : H^i(G', A) \to H^i(G, A)$ is given by

$$\tilde{h}(\alpha)(g_1, \cdots, g_i) = \alpha\left(h(g_1), \cdots, h(g_i)\right). \tag{D.1.6}$$

Cup product: There is an operation called *cup product* $C^i(G, A) \times C^j(G, A) \to C^{i+j}(G, A)$ when A is a ring. If $\alpha \in H^i(G, A)$ and $\beta \in H^j(G, A)$, then the cup product is defined as

$$(\alpha \cup \beta)(g_1, \cdots, g_{i+j}) = \alpha(g_1, \cdots, g_i)\beta(g_{i+1}, \cdots, g_{i+j}). \tag{D.1.7}$$

It can be easily checked that this product descends to a product on cohomologies.

One-dimensional representations of G: Let us ask what is the meaning of $H^1(G, \mathrm{U}(1))$. The 1-cochains are maps from G to $\mathrm{U}(1)$ and imposing the cocycle condition turns them into group homomorphisms. Hence $H^1(G, \mathrm{U}(1))$ is the group formed by one-dimensional representations of G. In particular, when G is a finite Abelian group, then $H^1(G, \mathrm{U}(1)) \simeq \widehat{G}$ is the dual group.

Projective representations of G: Now, let us ask what is the meaning of $\epsilon \in H^2(G, \mathrm{U}(1))$. We want to interpret the $\epsilon(g_1, g_2)$ as the phases defining a projective representation of G. The cocyle condition reads

$$\epsilon(g_1, g_2)\epsilon(g_1 g_2, g_3) = \epsilon(g_2, g_3)\epsilon(g_1, g_2 g_3) \tag{D.1.8}$$

which is the associativity condition such phases are required to satisfy. Such a cocycle can be shifted by a coboundary of the form

$$d\beta(g_1, g_2) = \frac{\beta(g_1 g_2)}{\beta(g_1)\beta(g_2)} \tag{D.1.9}$$

which corresponds to rephasing of the group action on the projective representation. Thus, we see that $H^2(G, \mathrm{U}(1))$ classifies the phases encountered in projective representations upto rephasing. The usual representations correspond to the trivial element of $H^2(G, \mathrm{U}(1))$.

Crossed products and extensions of groups: Consider an Abelian group H and a (possibly non-Abelian) group K. Consider an action of K on H and use it to define $H^2(K, H)$. An element $\kappa \in H^2(K, H)$ can be used to define a group extension G of K by H, that is there is a short exact sequence

$$0 \to H \to G \to K \to 0 \tag{D.1.10}$$

and G is called the κ-cross product of K and H and it is written as $G = H \rtimes_\kappa K$. Explicitly, the group G as a set is the direct product $H \times K$ with the group multiplication given as follows

$$(h_1, k_1)(h_2, k_2) = (h_1 + (k_1 \triangleright h_2) + \kappa(k_1, k_2), k_1 k_2). \tag{D.1.11}$$

Here, $\triangleright$ denotes the action of K on H via inner automorphism in G. The reader can verify that the associativity of the group multiplication is ensured by the cocycle condition on κ. Shifting κ by a coboundary changes G upto isomorphism. Hence, group extensions of K by an Abelian group H are classified by a group action of K on H along with an element in $H^2(K, H)$ defined using the group action.

Bicharacters on G: Consider an Abelian group G and a trivial module A of G. Given a cohomology element in $H^2(G, A)$ represented by a cocycle $\epsilon(g, h)$, one can form $\alpha(g, h) = \epsilon(g, h) - \epsilon(h, g)$ which is an antisymmetric function on G. This is indeed well defined because adding a coboundary to ϵ doesn't change α. When A is $\mathrm{U}(1)$ this α is a bicharacter, and there's a bijection between an antisymmetric bicharacter on G and $H^2(G, \mathrm{U}(1))$.

A useful isomorphism: Recall the familiar statement that any closed n-form is exact *locally*. In group cohomology, the analogous statement is that

$$H^i(G, \mathbb{R}) = 1, \tag{D.1.12}$$

that is $H^i(G, \mathbb{R})$ is trivial. We can use this to obtain $H^{i+1}(G, \mathbb{Z}) \simeq H^i(G, \mathrm{U}(1))$. This follows from the long exact sequence associated to the sequence $0 \to \mathbb{Z} \to \mathbb{R} \to \mathrm{U}(1) \to 0$.

Explicit group cohomologies:

- $H^*(\mathbb{Z}_n, \mathbb{Z}) = \mathbb{Z}[x_2]/(n x_2)$.
- $H^2((\mathbb{Z}_n)^k, \mathrm{U}(1)) = (\mathbb{Z}_n)^{k(k-1)/2}$.
- $H^3((\mathbb{Z}_n)^k, \mathrm{U}(1)) = (\mathbb{Z}_n)^{k+k(k-1)/2+k(k-1)(k-2)/6}$, with generators given by

$$\alpha^{(i)}(a, b, c) = e^{2\pi i a_i (b_i + c_i - \langle b_i + c_i \rangle)/n^2}, \tag{D.1.13}$$
$$\alpha^{(i,j)}(a, b, c) = e^{2\pi i a_i (b_j + c_j - \langle b_j + c_j \rangle)/n^2}, \tag{D.1.14}$$
$$\alpha^{(i,j,k)}(a, b, c) = e^{2\pi i a_i b_j c_k / n} \tag{D.1.15}$$

  where $a, b, c = \{0, 1, \ldots, n-1\}$ and $\langle a \rangle$ is the mod n function to $\{0, \ldots, n-1\}$. In particular, $H^3(\mathbb{Z}_n, \mathrm{U}(1)) = \mathbb{Z}_n$ and its generator has the cocycle

$$\alpha(a, b, c) = e^{2\pi i a (b + c - \langle b + c \rangle)/n^2}. \tag{D.1.16}$$

- For D_m the dihedral group with m elements we have [213],

$$H^*(D_m, \mathbb{Z}) = \mathbb{Z}[a_2, b_2, c_3, d_4]/(2a_2, 2b_2, 2c_3, m d_4, (b_2)^2 + a_2 b_2 + (m^2/4) d_4). \tag{D.1.17}$$

In particular, $H^2(D_{2n+1}, \mathrm{U}(1)) = 0$, $H^3(D_{2n+1}, \mathrm{U}(1)) = \mathbb{Z}_{4n+2}$; $H^2(D_{2n}, \mathrm{U}(1)) = \mathbb{Z}_2 \times \mathbb{Z}_2$, $H^3(D_{2n}, \mathrm{U}(1)) = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{2n}$. The explicit generators can be found in [214].

- For Q_8, the quaternion group, we have [215]

$$H^*(Q_8, \mathbb{Z}) = \mathbb{Z}[A_2, B_2, C_4]/(2A_2, 2B_2, 8C_4, A_2^2, B_2^2, A_2B_2 - 4C_4). \tag{D.1.18}$$