On Symmetries, Anomalies and Their Generalizations

Ho Tat Lam

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Nathan Seiberg

September 2021

Abstract

This dissertation explores various generalizations of global symmetries and 't Hooft anomalies. Chapter two is based on work with Po-Shen Hsin and Nathan Seiberg [1]. It is dedicated to the study of one-form global symmetries in three and four dimensions. We investigate their physical implications, classify their 't Hooft anomalies and analyze their gauging. Chapter three is based on the work with Pranay Gorantla, Nathan Seiberg and Shu-Heng Shao [2]. It focuses on exotic theories with subsystem symmetries including theories of fractons. We reformulate these theories on a Euclidean spacetime lattice in a modified Villain formulation. This provides a rigorous treatment of the continuum theories and their singularities while preserving some of their essential properties including 't Hooft anomalies and dualities. Chapter four is based on work with Clay Córdova, Dan Freed and Nathan Seiberg [3, 4]. It extends the notion of 't Hooft anomalies to anomalies in the space of coupling constants. We demonstrate through examples in diverse dimensions that these generalized anomalies can constrain the phase diagram of the theories and their defects associated with space-dependent coupling constants.

Acknowledgements

Over the past five years in Princeton, I was fortunate to have the opportunity to interact with, collaborate with and learn from various mentors, colleagues and friends. They have helped me grow both as a theoretical physicist and as a person.

First and foremost, I'd like to express my sincerest gratitude to my advisor, Nathan Seiberg, for his patient guidance, constant encouragement and generous support. I cannot describe how much he has taught me about physics, scientific research and beyond. He has set up a role model for theoretical physicists which I will forever look up to.

I am grateful to Herman Verlinde for the opportunity to work with him during my academic adolescence and for generously sharing his time and ideas. I would also like to thank him and Sanfeng Wu for agreeing to serve on my FPO committee, and to Igor Klebanov for reading this thesis and providing constructive comments. Thanks also go to Bob Austin for supervising my experimental project and to Average Phan for his guidance and enthusiasm throughout the project, from which I learned how filamentous *E. coli* swim. Thanks also to our graduate administrator, Kate Brosowsky, for all her help throughout my time in Princeton.

I would also like to thank my wonderful collaborators: Dan Freed, for introducing me to the mathematical aspects of quantum field theory; Akash Goel, for many enjoyable late-night conversations and chess games; Pranay Gorantla, for being an unfailing comrade on the journey of exploring exotic theories; Po-Shen Hsin, for helping me understand various subtleties about symmetries, anomalies and TQFTs; Thomas Mertens, for many cheerful and interesting discussions on JT gravity; and Gustavo Joaquin Turiaci, for countless stimulating discussions across diverse topics in high-energy theory and for teaching me the importance of dedication and persistence. Special thanks to Clay Córdova and Shu-Heng Shao, who have not only been knowledgeable and reliable collaborators but also encouraging mentors and excellent friends. I would like to thank Clay for innumerable enlightening and entertaining conversations on physics and other topics, and for providing substantial help and honest career advice. I would like to thank Shu-Heng for our many collaborations, for his patient and crystal-clear explanations to all my confusions and for our endless discussions on physics and beyond. It is amazing that with him, I wrote my first and most likely also my last paper in graduate school on completely different topics.

I am grateful to Pedro Vieira and Shota Komatsu for introducing me to the fantastic world of high-energy theory and for showing me how seemingly impossible problems can be cracked through perseverance and hard work.

I am deeply indebted to my undergraduate advisor, Kwok Yip Szeto, from Hong Kong University of Science and Technology – for carefully guiding me through my first research project, for encouraging me to pursue theoretical physics as a career and for believing in me when I doubted myself. He was a wise man, an excellent teacher, a dedicated mentor and a good friend. His kindness, humor, and optimism will forever be missed.

I am thankful to my dear friends for brightening up my life here: to Ziming Ji, Wentao Fan, and Akash Goel for the weekly swimming and badminton games in our first year; to Javier Roulet and Ro Kiman for their homemade Argentinian style pizzas; to Jiaqi Jiang for being a great gym buddy; to the fourth floor dinner crew, Yiming Chen, Huan He, Jie Wang, Fang Xie, Wenli Zhao, and Yunqin Zheng for numerous fun conversations over the dinners; to Yichen Fu, Tong Gao, Ming Lyu, and Xinwei Yu, for many gatherings and games; to the members of Luo-Ye's travel agency, Jingyu Luo, Zhenbin Yang, Shuyi Wang, and Xiaowen Chen for the many exciting weekend escapes; to Xue Song for being an amazing roommate and for many homemade desserts, roasted ducks and thought-provoking conversations; to Xuelan Wu for being with me through my ups and downs and for constantly reminding me of the world outside physics. I will cherish every piece of memories with them, and I hope that all my friends have a bright future and a happy life.

Many thanks to the physicists of Jadwin Hall and the Institute for Advanced Study, and to those whom I met at conferences and schools for discussions on various topics: Ahmed Almheiri, Nathan Benjamin, Fiona Burnell, Bruno Le Floch, Mattew Heydeman, Daniel Kapec, Biao Lian, Raghu Mahajan, Márk Mezei, Vladimir Narovlansky, Kantaro Ohmori, Yifan Wang, Yizhi You, Xinan Zhou; Brad Bachu, Damon Binder, Ksenia Bulycheva, Shai Chester, Trithep Devakul, Yale Fan, Aaron Hillman, Luca Iliesiu, Chirstian Jespsen, Himashu Khanchandani, Vladimir Kirilin, Henry Lin, Alexey Milekhin, Preethi Pallegar, Fedor Popov, Grigory Tarnopolski, Junyi Zhang, Wayne Zhao; Dean Carmi, Hongbin Chen, Wei Gu, Indranil Halder, Justin Kaidi, Ethan Lake, Julio Parra-Martinez, Wolfger Peelaers, Sahand Seifnashri, Adar Sharon, Vasudev Shyam, Haoyu Sun, Zimo Sun, and Jingxiang Wu. Also, many thanks to Ilya Belopolski, Brandon Bonham, Peter Czajka, Samuel Higginbotham, Steven Li, Sarah Marie Li, Xinran Li, Zheng Ma, Diana Valverde Mendez, Sanjay Moudgalya, Seth Olsen, Justin Ripley, Yu Shen, Nana Shumiya, Corwin Shiu, Liwei Song, Wudi Wang, Zeyu Wang, Jim Wu, Bin Xu and many others for making Princeton a place filled with warmth and joy.

I would like to thank the Croucher Foundation for supporting my first three years, and the Graduate School for providing me with the Charlotte Elizabeth Procter Fellowship during my fifth year.

At the end, I'd like to express my heartfelt thanks to my families: my mom, my father and my sister for their unconditional love and support. It would have been impossible for me to achieve this stage without the sacrifice of my parents and their dedication to our education.

Dedicated to my mother 蓉冰,

and my father 英奇.

Contents

Abstract . . . iii

Acknowledgements . . . iv

1 Introduction **1**

1.1 Global symmetries and their generalizations . . . 1

1.2 't Hooft anomalies and their generalizations . . . 4

1.3 Phases of matter . . . 6

1.4 Overview and summary . . . 8

2 One-form Global Symmetries in Three and Four Dimensions **12**

2.1 Preliminary and Summary . . . 12

2.2 One-form symmetries in 3d and their gauging . . . 24

2.2.1 One-form global symmetries in 3d TQFTs . . . 24

2.2.2 The minimal Abelian TQFT $\mathcal{A}^{N,p}$ . . . 28

2.2.3 Factorization of 3d TQFTs when $\gcd(N,p)=1$ . . . 31

2.2.4 't Hooft anomaly of one-form global symmetries . . . 34

2.2.5 A generalization of the three-step gauging procedure to anomalous theories 36

2.3 Coupling to a 4d bulk . . . 39

2.3.1 The bulk coupling . . . 39

2.3.2 Gauge the one-form symmetry . . . 40

2.3.3 Interfaces between two different bulk TQFTs . . . 44

2.4 $SU(N)$ and $PSU(N)$ gauge theory in 4d . . . 50
2.4.1 $SU(N)$ gauge theory, walls and interfaces . . . 50
2.4.2 $PSU(N)$ gauge theory . . . 55
2.4.3 Interfaces in $PSU(N)$ gauge theory . . . 57
2.5 Appendix A: Definitions of Abelian anyons . . . 59
2.6 Appendix B: Jacobi symbols . . . 61
2.7 Appendix C: 2d unitary chiral RCFT for Abelian 3d TQFT . . . 61
2.8 Appendix D: Gauging a general anomaly free subgroup . . . 64
2.9 Appendix E: Two-form $\mathbb{Z}_N$ gauge theory in 4d . . . 66
2.10 Appendix F: Minimal TQFTs for general one-form symmetries . . . 71

3 A Modified Villain Formulation of Fractons and Other Exotic Theories with Subsystem Symmetries 75

3.1 Preliminary and summary . . . 75
3.1.1 Overview of continuum field theories for exotic models . . . 75
3.1.2 Modified Villain lattice models . . . 77
3.1.3 Outline . . . 81
3.2 2+1d (3d Euclidean) exotic theories . . . 83
3.2.1 ϕ-theory (XY-plaquette model) . . . 83
3.2.2 A-theory ($U(1)$ tensor gauge theory) . . . 92
3.2.3 $\mathbb{Z}_N$ tensor gauge theory . . . 97
3.3 3+1d (4d Euclidean) exotic theories with cubic symmetry . . . 103
3.3.1 ϕ-theory . . . 103
3.3.2 A-theory . . . 106
3.3.3 X-cube model . . . 109
3.4 Appendix A: Villain formulation of some classic quantum-mechanical systems . . . 115
3.4.1 Appendix A.1: Particle on a ring . . . 115
3.4.2 Appendix A.2: Noncommutative torus . . . 119

3.5 Appendix B: Modified Villain formulation of 2d Euclidean lattice theories without gauge fields . . . 122
3.5.1 Appendix B.1: 2d Euclidean XY-model . . . 122
3.5.2 Appendix B.2: 2d Euclidean $\mathbb{Z}_N$ clock model . . . 133
3.6 Appendix C: Modified Villain formulation of p-form lattice gauge theory in diverse dimensions . . . 137
3.6.1 Appendix C.1: $U(1)$ gauge theory . . . 137
3.6.2 Appendix C.2: $\mathbb{Z}_N$ gauge theory . . . 141

4 Anomalies in the Space of Coupling Constants 148
4.1 Preliminary and Summary . . . 148
4.1.1 Anomalies and Symmetries . . . 149
4.1.2 Anomalies in Parameter Space: Defects . . . 152
4.1.3 Anomalies in Parameter Space: Families of QFTs . . . 156
4.1.4 Synthesis via Anomaly Inflow . . . 159
4.1.5 An Intuitive Interpretation in Terms of -1-Form Symmetries . . . 160
4.1.6 Examples and Summary . . . 161
4.2 A Particle on a Circle . . . 163
4.2.1 Spacetime Dependent Coupling θ . . . 163
4.2.2 Coupling to Background $U(1)$ Gauge Fields . . . 166
4.2.3 The Anomaly . . . 168
4.2.4 Dynamical Consequences: Level Crossing . . . 169
4.3 Massive Fermions . . . 172
4.3.1 Fermion Quantum Mechanics . . . 173
4.3.2 Real Fermions in $3d$ . . . 176
4.3.3 Weyl Fermions in $4d$ . . . 181
4.4 QED_2 . . . 185
4.4.1 $2d$ Abelian Gauge Theory . . . 185

4.4.2 QED_2 with one charge p scalar 191
4.4.3 QED_2 with N charge 1 scalars 194
4.5 $4d$ Yang-Mills Theory 197
4.5.1 $SU(N)$ Yang-Mills Theory 197
4.5.2 Adding Adjoint Higgs Fields 200
4.5.3 Other Gauge Groups 202
4.6 $4d$ QCD 209
4.6.1 $SU(N)$ QCD 210
4.6.2 $Sp(N)$ QCD 217
4.7 Appendix A: Axions and Higher Group Symmetry 222

Chapter 1

Introduction

1.1 Global symmetries and their generalizations

Symmetry is one of the most fundamental concepts in theoretical physics. It furnishes a powerful principle for organizing different phases of matter, and provides a universally applicable tool for analyzing strongly-coupled quantum field theories.

From the modern perspective, an ordinary global symmetry[1] in quantum field theories are understood abstractly as a set of codimension one topological operators[2] whose fusion obey the group multiplication law of the symmetry group [5].

As an example, consider a d-dimensional theory with an ordinary continuous internal global symmetry G. Noether's theorem implies that there is a conserved current J^μ that obeys $\partial_\mu J^\mu = 0$. We can define a $(d-1)$-form current $J = \epsilon_{\mu_1\mu_2\cdots\mu_d} J^{\mu_1} dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_d}$ and recast the current conservation equation into the condition that J is a closed form $dJ = 0$. The conserved charges

[1]We distinguish global symmetry from gauge symmetry. Global symmetry is an intrinsic property of a quantum field theory while gauge symmetry is a redundancy in the descriptions. This is well illustrated in the context of duality. The same quantum field theory can have two equivalent descriptions whose gauge symmetries are completely different but their global symmetries must be the same.

[2]We will use operators and defects interchangeably.

are operators supported on closed $(d-1)$-dimensional manifolds $\mathcal{M}^{(d-1)}$

$$Q\left(\mathcal{M}^{(d-1)}\right) = \oint_{\mathcal{M}^{d-1}} J \ . \tag{1.1}$$

By exponentiating the charges, we construct a set of codimension one symmetry operators $U_g(\mathcal{M}^{(d-1)})$ labeled by group elements g. Two symmetry operators can fuse according to the group multiplication law

$$U_{g_1}\left(\mathcal{M}^{(d-1)}\right) U_{g_2}\left(\mathcal{M}^{(d-1)}\right) = U_{g_1 g_2}\left(\mathcal{M}^{(d-1)}\right) \tag{1.2}$$

Because J is a closed form, these symmetry operators are topological – they are invariant under small deformation of $\mathcal{M}^{(d-1)}$ unless the deformation crosses a charged operator.

The local operators transform in representations of the symmetry group G. A symmetry operator $U_g(\mathcal{M}^{(d-1)})$ acts on a local operator $\mathcal{O}_i(x)$ by surrounding its manifold $\mathcal{M}^{(d-1)}$ around the point x

$$U_g(\mathcal{M}^{(d-1)})\mathcal{O}_i(x) = R^j{}_i(g)\mathcal{O}_j(x) \ , \tag{1.3}$$

where $R^j{}_i(g)$ is the representation of the group element g.

This formulation of global symmetry in terms of symmetry operators naturally generalizes to discrete symmetries where there are no conserved currents. As we will discuss below, changing different properties of the symmetry operators lead to different generalizations of global symmetry including higher-form symmetry and subsystem system symmetry.

One generalization of global symmetry is to consider topological symmetry operators of higher codimensions. This leads to high-form symmetry [5]. A q-form global symmetry in d dimensions is generated by topological operators $U_g(\mathcal{M}^{(d-q-1)})$ supported on $(q+1)$-codimensional manifolds $\mathcal{M}^{(d-q-1)}$. These symmetry operators fuse according to group multiplication law. For $q = 0$, the q-form symmetry reduces to the ordinary symmetry so we will use zero-form symmetry and ordinary symmetry interchangeably. For $q \geq 1$, because of the topological property, the fusion

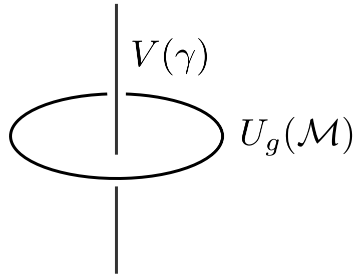


Figure 1.1: Linking the one-form symmetry operator $U_g(\mathcal{M})$ with the charged line operator $V(\gamma)$ in three dimensions. The one-form symmetry operator $U_g(\mathcal{M})$ is a line operator in three dimensions.

of the symmetry operators is commutative, and hence the symmetry group must be abelian. Operators charged under a q-form symmetry are extended operators supported on q-dimensional manifolds. For example, one-form symmetry acts on line operators, two-form symmetry acts on surface operators and so on. A q-form symmetry operator $U_g(\mathcal{M}^{(d-q-1)})$ acts on these q-dimensional charged operators when its manifold $\mathcal{M}^{(d-q-1)}$ link with them. An example of such linking for a one-form symmetry in three dimensions is illustrated in figure 1.1. More generally, higher-form symmetries of different degrees can mix and form higher-group symmetry [6–10].

One of the simplest example of theories with higher-form symmetry is $U(1)$ Maxwell theory in four dimensions with Euclidean Lagrangian

$$\mathcal{L} = \frac{1}{2g^2} F \wedge *F \ , \tag{1.4}$$

where $*F$ is the Hodge dual of the field strength $F = dA$. The theory has an electric one-form symmetry generated by a two-form current $J_E = \frac{1}{g^2} * F$, which is closed because of the equations of motion. The symmetry shifts the electric gauge fields by a flat connection. The corresponding symmetry operators are surface operators $U^E_{g=\exp(i\alpha)}(\mathcal{M}^{(2)}) = \exp(i\alpha \oint_{\mathcal{M}^{(2)}} J_E)$ and the charged operators are the Wilson lines $W = \exp(i \oint A)$. There is also a magnetic one-form symmetry generated by the two-form current $J_M = \frac{1}{2\pi} F$, which is closed due to the Bianchi identity. The symmetry operators are surface operators $U^M_{g=\exp(i\alpha)}(\mathcal{M}^{(2)}) = \exp(i\alpha \oint_{\mathcal{M}^{(2)}} J_M)$

and the charged operators are the 't Hooft lines $H = \exp(i \oint \tilde{A})$ where the magnetic gauge field $\tilde{A}$ obeys $d\tilde{A} = \frac{1}{g^2} * dA$.

Another generalization of global symmetry is to consider symmetry operators of higher codimensions that are conserved in time but are not completely topological in space. This leads to subsystem symmetry. Since the symmetry operators depend not only on their topology but also on their geometry and where they are inserted, there can be infinitely many of them in the continuum. As a result, the continuum field theories with subsystem symmetries exhibit many peculiarities such as infinite ground state degeneracy and discontinuities in physical observables.

As an example, consider a $U(1)$ subsystem symmetry with currents (J_0, J_{xy}) in 2+1 dimensions that obey a nonstandard current conservation equation [11]

$$\partial_0 J_0 = \partial_x \partial_y J_{xy} \ . \tag{1.5}$$

On a two-dimensional torus with periodic boundary conditions in both the x and the y direction, the conserved charges are line operators that extend either in the x direction or in the y direction

$$Q_y(y) = \oint dx\, J_0 \ , \quad Q_x(x) = \oint dy\, J_0 \ . \tag{1.6}$$

These charges obey the constraint $\oint dx\, Q_x(x) = \oint dy\, Q_y(y)$. The charges at different x or y are distinct charges. Hence there are infinitely many of them in the continuum. We can regularize it by introducing a lattice of finite size L_x, L_y. Then the number of independent charges is regularized to $L_x + L_y - 1$.

1.2 't Hooft anomalies and their generalizations

Ordinary global symmetry can have 't Hooft anomalies – an obstruction to gauging the symmetry. The same is true for its generalizations including higher-form symmetry and subsystem symmetry.

A useful way to characterize the 't Hooft anomalies is to couple the global symmetry to an appropriate classical background gauge field A. Depending on the symmetry, A could be a standard connection for an ordinary continuous (zero-form) global symmetry, a discrete gauge field for a discrete global symmetry, a $(q+1)$-form gauge field for a q-form symmetry, or a tensor gauge field for a subsystem symmetry. For example, the $U(1)$ subsystem symmetry (1.5) couples to a $U(1)$ tensor gauge field (A_0, A_{xy}) with a gauge symmetry $(A_0, A_{xy}) \to (A_0+\partial_0\lambda, A_{xy}+\partial_x\partial_y\lambda)$. The coupling introduces the term $A_0J_0 + A_{xy}J_{xy}$ in the Lagrangian.

The 't Hooft anomaly is a violation of the gauge invariance of the partition function $Z[A]$ under the background gauge transformation of A. We denote the gauge parameter by λ and the gauge field after the gauge transformation by A^λ. Under the background gauge transformation of λ, the partition function transforms by a phase

$$Z[A^\lambda] = Z[A]\exp\left(-i\int_{X_d}\alpha(\lambda, A)\right) , \tag{1.7}$$

where X_d is the d-dimensional spacetime. Since the partition function $Z[A]$ is subject to an ambiguity due to different regularization scheme, the 't Hooft anomaly is considered nontrivial only if it cannot be removed by classical counterterms. If the global symmetry is anomaly free, we can generate new theories by gauging the symmetry. It amounts to promoting the background gauge field A to a dynamical gauge field and summing over all gauge inequivalent configurations.

In general, a 't Hooft anomaly in d dimensions can be summarized using a $(d+1)$-dimensional classical action $\omega(A)$ of background gauge fields with the property that its partition function $\Omega[A] = \exp(i\omega(A))$ is gauge invariant on a closed manifold and transforms as

$$\Omega[A^\lambda] = \Omega[A]\exp\left(i\int_{Y_{d+1}} d\alpha(\lambda, A)\right) . \tag{1.8}$$

on an open manifold Y_{d+1}. Such actions are also referred to as invertible field theories [12].

Suppose there exists an extension of the background gauge field A from X_d to Y_{d+1} with $\partial Y_{d+1} = X_d$. We can then define a gauge invariant partition function $\tilde{Z}[A] = Z[A]\Omega[A]$ where the 't Hooft anomaly of the d-dimensional theory is canceled by the invertible field theory in the bulk Y. This is known as the anomaly inflow [13].

The 't Hooft anomaly is invariant under renormalization group flow [14]. One way to argue for this is to consider how the partition function $\tilde{Z}[A]$ evolves along the flow. The classical anomaly action ω transforms continuously under renormalization so in order to maintain the gauge invariance of $\tilde{Z}[A]$, the long distance theory on the boundary X_d must have an anomaly action that can be continuously connected to the original one. Hence the deformation class of the anomaly actions is a renormalization group invariant, which can be used to constrain renormalization group flow. In particular any theory with an anomaly action ω that is not continuously connected to the trivial action cannot flow at long distances to a trivially gapped theory with a unique vacuum and no long-range degrees of freedom.

In the orthogonal direction of generalizing 't Hooft anomaly to generalized symmetries, we can also extend the notion of 't Hooft anomaly by considering anomaly actions $\omega(A)$ that depend on both background fields and coupling constants. We refer to this generalized anomaly as anomaly in the space of coupling constants. As we will demonstrate in Chapter 4, anomaly in the space of coupling constants can be a useful tool for deducing the dynamics of a quantum field theory.

1.3 Phases of matter

Symmetry is a powerful organizing principle for thinking about phases of matter.

One criterion for distinguishing different phases is based on whether symmetries are spontaneously broken or not. For example, we can differentiate a crystalline phase and a liquid phase by the behavior of the continuous translational symmetry. The symmetry is spontaneously broken in the crystalline phase but it is preserved in the liquid phase.

Just as ordinary global symmetry, generalized global symmetries such as higher-form symmetry and subsystem symmetry can also be spontaneously broken. They lead to new criteria for classifying phases of matter.

As an example, deconfinement and confinement in a gauge theory can be rephrased as whether a one-form symmetry is spontaneously broken or not. They can be diagnosed by the behavior of large Wilson loops – whether the Willson loops obey a perimeter law or an area law. If the Wilson loops obey a perimeter law as opposed to an area law, we can redefine the loop operator by a local geometric counterterm such that it has a nonzero expectation value when it is large and correspondingly the one-form symmetry is spontaneously broken. Physically, the Wilson loops can be interpreted as the trajectories of a charged probe particle. An area law as opposed to a perimeter law means that the probe particles have a confining potential that increases with their separation and hence signals confinement in the theory. A canonical example for spontaneous one-form symmetry breaking is four-dimensional $U(1)$ Maxwell theory. The theory has two $U(1)$ one-form global symmetries which are both spontaneously broken and the associated Goldston boson is the massless photon [5].

It is more dramatic when a subsystem symmetry is spontaneously broken. For example, if a discrete subsystem symmetry is spontaneously broken, it can lead to a vast ground state degeneracy that grows with the system size. This happens for example in various fracton models including the Haah's code and the X-cube model [15–17].

Symmetry breaking however is not the only criterion for classifying phases of matter. For example, there are more refined classifications within the symmetry-preserving phases. With a given global symmetry, there are distinct gapped phases with short-ranged entanglement that cannot be smoothly connected to each other through symmetry-preserving deformations. They are known as symmetry protected topological (SPT) phases [18,19]. A characteristic of nontrivial SPT phases is that their boundary theory cannot be trivially gapped with a symmetry preserving boundary condition. This is because of the anomaly inflow from the bulk to the boundary. This suggests that SPT phases are equivalent to 't Hooft anomalies in one lower dimension [20]. Mathematically, they can both be described by invertible topological field theories [12]. Just as 't Hooft anomalies, the notion of SPT phases naturally extends to SPT phases protected by generalized symmetries such as higher-form SPT phases [7] and subsystem SPT phases [21–23].

There are also a large class of gapped topological phases that are not characterized by symmetries. They are under the name of topological order. One of the most prominent topological orders is the fractional quantum Hall states [24], which exhibit fractional statistics and topologically degenerate ground states. Mathematically, topological orders can be described by topological quantum field theory (TQFT) [25,26]. Many of them can be constructed by gauging the global symmetry of SPT phases. The same construction generalizes to higher-form symmetry. Gauging a higher-form SPT phase leads to a topological higher-form gauge theory [7]. On the other hand, if we gauge a subsystem SPT phase, we can end up with a fracton phase [17]. Some characteristics of fracton phases include massive excitations with restricted mobility and sub-extensive ground state degeneracy that grows with the system size. Because of these properties, fracton phases are different from topological orders, and it is challenging to find continuum descriptions for them.

1.4 Overview and summary

The overall theme of this dissertation is to explore generalizations of global symmetries and 't Hooft anomalies. We first explore various aspects of higher-form symmetries in three dimensions and four dimensions. We then move on to study systems with subsystem symmetries including various fracton models. Finally, we discuss an extension of 't Hooft anomalies to anomalies in the space of coupling constants. The content of each chapter is summarized below.

In chapter 2, we study three-dimensional and four-dimensional systems with a one-form global symmetry, explore their consequences, and analyze their anomalies and gauging. For simplicity, we focus on $\mathbb{Z}_N$ one-form symmetries. Generalizations to general discrete abelian groups have also been discussed.

A quantum field theory $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry has N topological symmetry lines that obey the $\mathbb{Z}_N$ fusion rule. The braiding of these lines and their spins are characterized by an integer p mod $2N$. We prove that when $\gcd(N,p)=1$, these topological lines themselves form a consistent TQFT or a modular tensor category denoted by $\mathcal{A}^{N,p}$. We further prove that when $\gcd(N,p)=1$, the quantum field theory $\mathcal{T}$ factorizes into two decoupled theories $\mathcal{T}=\mathcal{T}'\otimes\mathcal{A}^{N,p}$. If the theory $\mathcal{T}$ is a TQFT, the theory $\mathcal{T}'$ is a decoupled TQFT of lines that are neutral under the one-form symmetry.

The parameter p also characterizes the 't Hooft anomaly of the $\mathbb{Z}_N$ one-form symmetry. When $p=0 \bmod 2N$, the symmetry is anomaly free and it can be gauged. We denote the theory after gauging by $\mathcal{T}/\mathbb{Z}_N$. For theories $\mathcal{T}$ that are TQFTs, we outline a three-step procedures for gauging the symmetry. When $p\neq 0 \bmod N$, the one-form symmetry has an anomaly, which can be canceled by coupling the system to a four-dimensional SPT phase with a $\mathbb{Z}_N$ one-form symmetry. We analyze the consequences of gauging the $\mathbb{Z}_N$ one-form symmetry in this 3d-4d coupled system. After gauging, the bulk theory becomes a twisted $\mathbb{Z}_N$ two-form gauge theory which is equivalent to a $\mathbb{Z}_L$ one-form gauge theory with $L=\gcd(N,p)$. If the theory $\mathcal{T}$ is a TQFT, after gauging, the boundary theory can be described by a premodular category with L lines that braid trivially with all other lines. These L lines are the L topological lines of the bulk $\mathbb{Z}_L$ gauge theory. When $L=1$, the bulk theory is trivial and the boundary theory is a consistent TQFT (or a modular tensor category). When $L\neq 1$, the bulk theory is nontrivial and the boundary theory is not a consistent TQFT (or a modular tensor category) on its own. However, we can extract a TQFT (or a modular tensor category) from the premodular category on the boundary by identifying the L bulk lines with the trivial line. This leads to an effective 3d TQFT described by

$$\frac{\mathcal{T}\times\mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N}\ , \tag{1.9}$$

where the $\mathbb{Z}_N$ quotient represents gauging the diagonal one-form symmetry of the two theories in the numerator.

Next we apply our understanding of the one-form symmetries to four-dimensional $SU(N)$ and $PSU(N)$ gauge theories. The $PSU(N)$ gauge theory can be constructed by gauging the $\mathbb{Z}_N$ one-form symmetry of the $SU(N)$ gauge theory. Both theories have θ-parameters. For all values of θ, the dynamics of both theories are gapped and they are associated with either confinement or oblique confinement – probe quarks are confined. At low energy, the $SU(N)$ theory is trivially gapped while the $PSU(N)$ theory may include a discrete gauge theory depending on the θ-parameter.

We study these theories in backgrounds of space-dependent θ-parameters, which lead to interfaces. In the $SU(N)$ theory, the theories on the interfaces are typically not confined and they can be described by 3d TQFTs. This means that probe quarks are liberated and become anyons on the interfaces. Utilizing our understanding of one-form symmetries, we give a description to the interfaces in $PSU(N)$ gauge theory by gauging the $\mathbb{Z}_N$ one-form symmetries of the corresponding interfaces in the $SU(N)$ theory.

The materials in this chapter are based on work with Po-Shen Hsin and Nathan Seiberg [1]. They have been presented in the workshop "Geometrical Aspects of Supersymmetry" at the Simons Center for Geometry and Physics from October 22–26, 2018 and the conference "Between Topology and Quantum Field Theory" at University of Texas at Austin, from January 14–18, 2019.

In chapter 3, we study fractons and other closely-related exotic theories with subsystem symmetries. We reformulate these exotic theories on a Euclidean spacetime lattice. We first write them using the Villain approach and then modify them by suppressing topological excitations. The new lattice models are closer to the continuum than the original lattice versions. In particular, they exhibit many features of the continuum theories including emergent global symmetries, dualities and 't Hooft anomalies. Also, these new models provide a clear and rigorous formulation to the continuum theories and their singularities.

We also use this approach to review well-studied lattice models and their continuum limits. These include the XY-model, the $\mathbb{Z}_N$ clock-model, and various gauge theories in diverse dimensions. This presentation makes the role of symmetries associated with the topology of field space, duality, and various anomalies manifest.

The materials in this chapter are based on work with Pranay Gorantla, Nathan Seiberg and Shu-Heng Shao [2]. They have been presented in the workshop "New directions in topological phases: from fractons to spatial symmetries" held by the Simons Center for Geometry and Physics from May 24–28, 2021 and the conference "String Math 2021" held by IMPA, Rio de Janeiro, from June 14–18, 2021.

In chapter 4, we introduce anomalies in the space of coupling constants extending the notion of 't Hooft anomalies, analyze their dynamical implications, and demonstrate them in quantum mechanics and quantum field theories in diverse dimensions.

Same as classical background gauge fields for global symmetries, we also view coupling constants as background fields and study the theory in backgrounds of space-dependent coupling constants. In some cases, we observe that with space-dependent coupling constants, the partition function becomes no longer invariant under background gauge transformations of certain global symmetries. We interpret these phenomena as generalized anomalies involving coupling constants. Similar to ordinary 't Hooft anomalies, we can summarize these generalized anomalies using classical actions of coupling constants and background fields in one higher dimension.

Just as ordinary 't Hooft anomalies allow us to deduce dynamical consequences about the phases of the theory and its defects, the same is true for these generalized anomalies. An anomaly in the space of coupling constants implies that the infrared dynamics must be nontrivial somewhere within a family of theories labeled by coupling constants. Possible scenarios include phase transitions, conformal field theories or topological quantum field theories in the infrared. An anomaly in the space of coupling constants also implies an anomaly on the worldvolume of defects with space-dependent coupling constants.

An important class of examples that we discussed is generalized anomalies involving circle-valued θ-angles. This includes the quantum mechanics of a particle on a circle, 2d $U(1)$ gauge theory and 4d Yang-Mills theory. Another class of examples is generalized anomalies involving fermion masses in theories with fermions in various dimensions.

The materials in this chapter are based on work with Clay Córdova, Dan Freed and Nathan Seiberg [3, 4]. They have been presented in the conference "String Math 2019" at Uppsala University, from July 01–05, 2019.

Chapter 2

One-form Global Symmetries in Three and Four Dimensions

2.1 Preliminary and Summary

In this chapter we will investigate systems with one-form global symmetries in 3 and 4 dimensions. Some examples in 3d are $U(1)_N$ or $SU(N)_k$ Chern-Simons (CS) theory. They have a spontaneously broken $\mathbb{Z}_N$ one-form symmetry. An example in 4d is an $SU(N)$ gauge theory without quarks. Here the $\mathbb{Z}_N$ one-form symmetry is expected to be unbroken, which is related to the confinement of the system. If we add quarks in the fundamental representation to this theory, then the one-form symmetry is absent, and indeed the theory with quarks does not have a meaningful notion of confinement.

4d $SU(N)$ gauge theory with θ and domain walls

Of particular interest for us will be the behavior of this 4d $SU(N)$ theory with a θ-parameter. The lore is that at generic θ the system is confining and gapped with a trivial vacuum. At $\theta \in \pi\mathbb{Z}$, we have time-reversal and parity symmetries. These are unbroken at $\theta \in 2\pi\mathbb{Z}$. (For small values of N there are also other logical options [27].) But they are spontaneously broken at θ an odd multiple of π. In these cases the system has two degenerate vacua with domain walls that interpolate between them. Arguments based on anomalies in the one-form symmetry, which we will review below, suggest that the theory on the domain wall is an $SU(N)_1$ TQFT [5,27].[1]

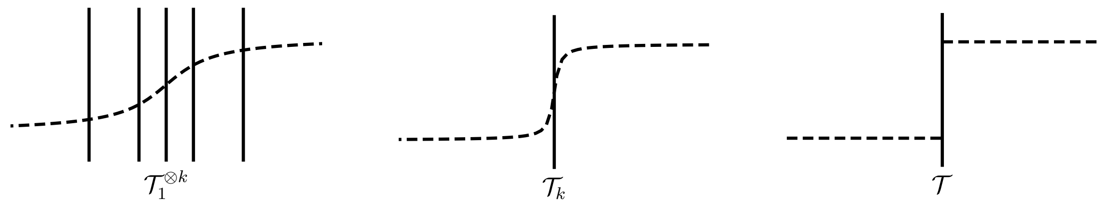
As stressed in [5,27], the transition at $\theta = \pi$ separates two distinct vacua in the following sense. On one side of the transition monopoles condense, leading to confinement, and on the other side of the transition dyons condense, leading to oblique confinement. More precisely, the transition at θ an odd multiple of π separates two distinct oblique confinement vacua. Since different dyons condense on the two sides of the domain wall, no dyon condenses on the wall. Therefore, the theory on the wall is not confining and the Wilson lines of the $SU(N)_1$ theory on the wall are world lines of unconfined probed quarks. Not only are these quarks liberated, they also have nontrivial braiding, i.e. they are anyons! Below we will give an intuitive physical argument explaining why they are anyons.

Interfaces

One of our goals is to study in detail interfaces in this theory. We let θ be a space- dependent interpolation between θ_0 to $\theta_0 + 2\pi k$. If the interpolation is over a length scale much longer than the inverse of the dynamical scale of the theory Λ, then at a generic spacetime point θ is essentially constant on the scale where confinement takes place and the vacuum is unique and varies smoothly. When θ crosses an odd multiple of π there is a domain wall separating two vacua. Therefore, the interpolation leads to k domain walls with $SU(N)_1$ on each of them [27], as illustrated in Figure 2.1a. If the interpolation is more rapid, then the TQFT $SU(N)_1 \otimes SU(N)_1 \otimes ...$ can undergo a transition to another TQFT $\mathcal{T}_k$, see Figure 2.1b. It was suggested in [27,28] that this theory is $SU(N)_k$. However, we will soon argue that there are also other logical possibilities and only a more detailed dynamical analysis can determine the right answer.

It is important that the theory on the interface is uniquely determined by the microscopic theory and by the profile of the space-dependent θ. This is to be contrasted with a sharp interface when θ is discontinuous, as illustrated in Figure 2.1c. Here we have the freedom to change the theory on the interface by adding more degrees of freedom there and to consider their dynamics.

[1] Although as spin TQFTs $SU(N)_1 \longleftrightarrow U(1)_{-N}$, we prefer to use $SU(N)_1$ because our theory is bosonic.



(a) Slow θ interpolation (b) Rapid θ interpolation (c) Sharp θ interpolation

Figure 2.1: The interfaces for different profiles of θ that interpolate from $\theta = \theta_0$ to $\theta = \theta_0 + 2\pi k$. The dashed lines are the profile of the θ parameter and the solid lines are the locations of the interfaces. In (a), there are k domain walls located at the transitions when θ crosses an odd multiple of π. The theory on each domain wall is $\mathcal{T}_1$, which we argue is $SU(N)_1$ [5]. When the θ variation is more rapid, as in (b), there is only one interface and the theory on it is $\mathcal{T}_k$. One option for that theory is $SU(N)_k$, but we will argue that other options are also possible. Finally, as in (c), θ can be discontinuous. In this case the theory on the interface $\mathcal{T}$ is not determined uniquely by the microscopic dynamics. But it is constrained by anomaly considerations.

We will not study it here. The same comments apply to a system with a boundary. As with the sharp interface, the boundary theory is constrained by anomalies, but there is a lot of freedom in adding boundary degrees of freedom.

Our main tool for analyzing the system is its $\mathbb{Z}_N$ one-form global symmetry. Related to this symmetry is an integer label p with $p \sim p + 2N$ and pN even [29, 5]. Furthermore, we have an identification in labeling the theories [29, 5, 27]

$$(\theta, p) \sim (\theta + 2\pi k, p + k(N-1)) \ . \tag{2.1}$$

One way to think about the parameter p is through coupling the $\mathbb{Z}_N$ global symmetry to a classical background two-form gauge field $\mathcal{B}_\mathcal{C}$ (the subscript $\mathcal{C}$ means that it is classical). Then, the parameter p is the coefficient of a counterterm proportional to the square of $\mathcal{B}_\mathcal{C}$ [29, 5]. This term does not affect any separated points correlation function, but it does affect contact terms and the behavior of the system with a boundary.

The key dynamical fact is that the theory confines. This means that the $\mathbb{Z}_N$ one-form symmetry is unbroken. Also, the spectrum is gapped and the low-energy dynamics is trivial

– there is not even a TQFT at long distances. The only meaningful fact that remains at low energies is the coefficient p of the counterterm of $\mathcal{B}_\mathcal{C}$, which means that the system can be in a nontrivial Symmetry Protected Topological (SPT) phase.

When we have an interface where θ changes by $2\pi k$ the two sides of the interface are typically in *different* SPT phases labeled by $p^\pm$ with

$$p^+ - p^- = k(N-1) \bmod 2N \ . \tag{2.2}$$

This means that when $p^+ \neq p^- \bmod 2N$ the theory on the interface cannot be trivial. It must have a $\mathbb{Z}_N$ one-form global symmetry with anomaly $(p^+ - p^-) \bmod 2N$.

Let us try to determine the theory on the interface. When the interface is rapid, we can shift θ on one side, as in equation (2.2), so that θ does not change across the interface, but p changes. It induces a Chern-Simons term $SU(N)_k$ on the interface. Next, as the theory becomes strongly coupled it confines and the bulk on the two sides of the interface become gapped and trivial. What happens to the $SU(N)_k$ theory on the interface? One option, which was advocated in [27], is that at least for small enough $|k|$ it is not affected by the confinement. However, the strong dynamics could change that answer.[2] But whatever the dynamics does, the one-form $\mathbb{Z}_N$ global symmetry and its anomaly $p^+ - p^-$ cannot change. Therefore, if $p^+ \neq p^- \bmod 2N$, the theory on the interface cannot be trivial, and we'll denote it by $\mathcal{T}_k$.

We start by reconsidering the special case $k = 1$. Can the UV answer $SU(N)_1$ be modified? We suggest that this cannot happen. First, as we will discuss in detail below, this particular theory is the minimal theory with a $\mathbb{Z}_N$ one-form symmetry of anomaly $N-1$. Every other TQFT with this property factorizes into $SU(N)_1$ times another TQFT, whose line operators are $\mathbb{Z}_N$ invariant. Therefore, it is natural to assume that in this case the UV answer does not change. Also, in a closely related supersymmetric theory, a string construction shows that the theory on the interface is $U(1)_{-N}$ [30], which is dual (as spin TQFT) to our answer $SU(N)_1$ [31].

As we move to higher values of k the situation is less clear. It was suggested in [27] that as

[2] We thank E. Witten for encouraging us to think about other options.

a slow interface becomes steeper, the $SU(N)_1^{\otimes k}$ TQFT can be Higgsed to the diagonal $SU(N)_k$. This would agree with the answer in the UV. However, further dynamical effects can change this answer. Since we expect the interface theory to remain non-confining, we do not anticipate monopoles to participate in this dynamics on the interface. Instead, we can consider dynamical scalar fields in the adjoint representation of $SU(N)$. Such scalar fields can arise from modes of the microscopic gluons and their presence does not break the exact $\mathbb{Z}_N$ one-form symmetry of the system. The condensation of these scalars can Higgs $SU(N)$ to various subgroups. The maximum possible Higgsing with one adjoint scalar is to the Cartan torus $U(1)^{N-1}$. In this case the $SU(N)_k$ theory becomes $U(1)^{N-1}$ with a coefficient matrix given by kK_{Cartan} with K_{Cartan} the Cartan matrix of $SU(N)$. (Note that for $k=1$ the TQFT $SU(N)_1$ is the same as this Abelian TQFT.) With more than one adjoint scalars, we can further Higgs the system all the way down to a $\mathbb{Z}_N$ gauge theory[3] with level $K=-kN(N-1)=-(p^+-p^-)N$. Below we will review in detail this TQFT and its properties.

The upshot of the discussion above is that the spontaneously broken $\mathbb{Z}_N$ one-form symmetry and its anomaly p^+-p^- restrict the TQFT on the interface $\mathcal{T}_k$, but do not uniquely determine it. For $k=1$ it is natural to assume that the correct answer is the minimal one $\mathcal{T}_1=SU(N)_1$. For higher values of k there are several natural possibilities including $SU(N)_k$, but the other options include also some Abelian TQFTs. It should be emphasized, however, that despite our inability to determine $\mathcal{T}_k$ beyond the symmetry and anomaly constraints, this theory is uniquely determined by the dynamics.

Gauging the $\mathbb{Z}_N$ one-form symmetry – 4d $PSU(N)$ gauge theory and interfaces

When the $\mathbb{Z}_N$ one-form symmetry is gauged, the microscopic 4d $SU(N)$ gauge theory becomes a $PSU(N)$ gauge theory and the macroscopic theory might no longer remain trivial [34].

[3] The $\mathbb{Z}_N$ gauge theory at level K can be expressed as the following $U(1)\times U(1)$ Chern-Simons theory [32,33,29]

$$(\mathcal{Z}_N)_K: \quad \int \left(\frac{K}{4\pi}xdx+\frac{N}{2\pi}xdy\right) . \tag{2.3}$$

For even K this is a Dijkgraaf-Witten (DW) theory [26].

Specifically, it becomes a $\mathbb{Z}_L$ gauge theory with[4]

$$L = \gcd(p, N)\ . \tag{2.4}$$

Unlike the original $SU(N)$ theory where p affects only the SPT phase, here it affects the low-energy dynamics. Now the interface is more interesting. Clearly, we have a $\mathbb{Z}_{L_\pm}$ gauge theory with $L_\pm = \gcd(p_\pm, N)$ on the two sides of the interface. But what is the resulting theory on the interface?[5]

When $L_+ = L_- = 1$ the bulk theory on the two sides is trivial and the low-energy theory is only the 3d theory on the interface and it is completely meaningful. However, when either L_+ or L_- (or both) are not equal to one, the bulk theory is not trivial and the low-energy TQFT is not three dimensional. It is four dimensional and the interface appears as a 3d defect in the 4d bulk. Therefore, it is meaningless to ask what the 3d theory on the interface is. It is not decoupled from the 4d bulk. Nevertheless, we will argue that there exists a 3d TQFT that captures many of the features of the physics along the interface. Roughly, it is a quotient of the full 4d system by the physics of the 4d bulk. We will describe this in more detail below.

One-form global symmetries in 3d and their gauging

In order to understand these TQFTs we will have to explore in more detail the one-form global symmetry, its anomaly, and its gauging in 3 and 4 dimensions. Let us start with a 3d one-form symmetry $\mathcal{A}$. The charge operators are line operators $a_{\mathbf{g}}$ labeled by a group element $\mathbf{g} \in \mathcal{A}$. The group multiplication corresponds to the fusion of the lines:

$$a_{\mathbf{g}+\mathbf{g}'} = a_{\mathbf{g}} a_{\mathbf{g}'}\ , \tag{2.5}$$

where the group multiplication of $\mathcal{A}$ is denoted by addition, and the product of two lines denotes their fusion. Each line $a_{\mathbf{g}}$ represents an Abelian anyon in the TQFT.

[4]Below we will show that on a nonspin manifold this $\mathbb{Z}_L$ gauge theory is sometimes twisted in a particular way.

[5]Note that the naive answer $PSU(N)_k$ cannot be right. For generic k this is not a consistent theory [35, 26]!

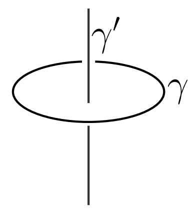


Figure 2.2: Braiding the line operators supported on the curves γ and γ'.

For simplicity we will focus on a $\mathbb{Z}_N$ one-form symmetry. The symmetry lines are a^s with

$$a^N = 1 \tag{2.6}$$

and we refer to a as the generating line. In general, this generator is not unique and some of the expressions below depend on the choice of generator.

In a TQFT with a $\mathbb{Z}_N$ one-form symmetry, each line W carries a $\mathbb{Z}_N$ charge $q(W) \in \mathbb{Z}_N$ under the symmetry, which is determined by braiding the generating line a with W (see Figure 2.2):

$$a(\gamma)W(\gamma') = W(\gamma')e^{\frac{2\pi i q(W)}{N}} \ . \tag{2.7}$$

We will show that general considerations constrain the spins of the symmetry lines to be[6]

$$h[a^s] = \frac{ps^2}{2N} \text{ mod } 1 \ , \tag{2.8}$$

for some integer $p = 0, 1, \cdots, 2N-1$ mod $2N$. Imposing (2.6) leads to

$$pN \in 2\mathbb{Z} \ . \tag{2.9}$$

The situation in spin TQFT is slightly different because such theories have a transparent spin-half line ψ. This will be discussed in detail below.

One significance of the parameter p is that it determines the $\mathbb{Z}_N$ charge $q(a) = -p \mod N$

[6]We thank Z. Komargodski and J. Gomis for a discussion about this point.

of the generating line a (see Section 2.2.1). Clearly, the symmetry can be gauged only when the symmetry lines themselves are neutral, i.e. when $q(a) = 0$. Therefore, the parameter p controls the obstruction to gauging, which is the 't Hooft anomaly.

When $p = 0$, the $\mathbb{Z}_N$ one-form symmetry is anomaly free and it can be gauged. Denoting the original TQFT by $\mathcal{T}$, we will denote the result of this gauging by the TQFT

$$\mathcal{T}' = \mathcal{T}/\mathbb{Z}_N \ . \tag{2.10}$$

When $p = N$ the generating line has spin $\frac{1}{2}$ and the gauged system $\mathcal{T}/\mathbb{Z}_N$ is a spin TQFT.[7]

There are several ways to describe the gauging procedure. From the perspective of symmetry defects, gauging a symmetry amounts to summing over all possible insertions of symmetry defects [5]. In the corresponding two-dimensional chiral algebra, gauging the one-form symmetry corresponds to extending the chiral algebra [36, 35]. For Chern-Simons theory it can sometimes be described by the quotient of the gauge group by a subgroup of the center [35, 5]. In the condensed matter literature, it is called "anyon condensation" of the Abelian anyon that corresponds to the generating line of the one-form symmetry [37].

For $p = 0$ when the symmetry generating line a has integer spin the gauging involves three steps [36, 35]:

Step 1 Discard the lines W that are not invariant under the $\mathbb{Z}_N$ one-form symmetry.

Step 2 Since a is trivial, we identify the lines W and Wa obtained by fusing with a.

Step 3 If W is a fixed point under the fusion with a, then there are N copies of W. More precisely, if s is the minimal divisor of N such that W is invariant under the fusion with a^s, then there are N/s copies of W.[8]

[7]This is the case even when the original TQFT is non-spin. In this case we can say that there is a mixed 't Hooft anomaly between the $\mathbb{Z}_N$ one-form symmetry and gravity (the bosonic Lorentz symmetry).

[8]This can be proven by iteration. Let N_1 be the highest non-trivial divisor of $N_0 = N$. Then gauging the $\mathbb{Z}_{N_0/N_1}$ subgroup generated by a^{N_1} leads to N_0/N_1 copies at each fixed point. We can continue to gauge the remaining $\mathbb{Z}_{N_1}$ symmetry by repeating the process. For the minimal divisor N_i such that W is the fixed point under the fusion with a^{N_i}, there will be $\frac{N_0}{N_1}\frac{N_1}{N_2}\cdots\frac{N_{i+1}}{N_i} = \frac{N}{N_i}$ copies of W after gauging the $\mathbb{Z}_N$ symmetry.

For even $p = N$, the generating line a has half-integer spin and then the resulting theory after gauging is a spin TQFT. As we will discuss below, this leads to the same three-step process.

When $p \neq 0, N$ the generating line a is charged under the $\mathbb{Z}_N$ symmetry and that symmetry cannot be gauged. However, a subgroup $\mathbb{Z}_L \subset \mathbb{Z}_N$ with[9]

$$L = \gcd(p, N) \tag{2.11}$$

can be gauged. It is generated by the line $\hat{a} = a^{N/L}$. Since its spin is $h = \frac{pN}{2L^2}$, its p-parameter is $\hat{p} = \frac{pN}{L} \bmod 2L$. Note that $\hat{p} = 0 \bmod L$. When $\hat{p} = 0 \bmod 2L$ we can gauge this $\mathbb{Z}_L$ subgroup as above, and when $\hat{p} = L \bmod 2L$ the resulting gauged theory is a spin TQFT. The most anomalous case has $L = 1$ and it will have particular significance below.

[9]The relation to the seemingly unrelated equation (2.4) will be clear soon.

Outline and summary of new results

In Section 2 we will discuss in detail the one-form symmetry in 3d and will prove the statements above. We will also show that for given relatively prime N and p (i.e. $L = 1$) there is a minimal TQFT with a $\mathbb{Z}_N$ one-form symmetry of anomaly p. We will denote it by $\mathcal{A}^{N,p}$. Furthermore, we will show that any TQFT $\mathcal{T}$ with such a one-form global symmetry factorizes as

$$\mathcal{T}' \otimes \mathcal{A}^{N,p} \qquad \text{for} \qquad L = \gcd(N,p) = 1 \ . \tag{2.12}$$

This means that all the lines in $\mathcal{T}'$ are $\mathbb{Z}_N$ neutral. This is quite surprising – the entire effect of the global symmetry is limited to this factor of $\mathcal{A}^{N,p}$ and the rest of the theory is not affected by it. We can also invert equation (2.12) and map the TQFT $\mathcal{T}$ to

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N,-p}}{\mathbb{Z}_N} \ . \tag{2.13}$$

When $L = N$ we have the three-step gauging procedure we discussed above that maps a TQFT $\mathcal{T}$ to $\mathcal{T}' = \mathcal{T}/\mathbb{Z}_N$ (2.10). In the other extreme of $L = 1$ we can map $\mathcal{T}$ to $\mathcal{T}'$ of (2.13). Here we simply remove the non-invariant lines, i.e. we perform only step 1 of the three steps.

In Section 2.2.5 we will generalize this procedure to generic $L = \gcd(N,p)$. We map

$$\mathcal{T} \to \mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N} = \frac{\mathcal{T}/\mathbb{Z}_L \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_{N/L}} \ . \tag{2.14}$$

The equality between these expressions will be derived in Section 2. In the map (2.14) we perform step 1 of the three-steps using $\mathbb{Z}_N$ and perform steps 2 and 3 using $\mathbb{Z}_L$. This expression coincides with (2.10) for $L = N$ and with (2.13) for $L = 1$ and generalizes them to generic L. (Depending on the details (2.14) might be a spin TQFT.)

This generalized gauging procedure has a physical interpretation, which we describe below, in terms of coupling the system to a 4d bulk gauge theory. It is also related to a more mathematical discussion in [38–41] and the discussion on the Walker-Wang lattice models in [42, 43].

In Section 3, we couple the 3d system to a 4d bulk and promote the background $\mathcal{B}_{\mathcal{C}}$ gauge fields to quantum fluctuating fields and correspondingly, we drop the subscript $\mathcal{C}$. The bulk theory becomes effectively a $\mathbb{Z}_L$ gauge theory.

As we said above, for $L = 1$ the bulk theory is trivial and therefore there is a meaningful 3d TQFT on the boundary. It cannot be $\mathcal{T}/\mathbb{Z}_N$ because the anomaly makes this quotient inconsistent. Instead, we will show that the theory on the boundary is $\mathcal{T}'$ of (2.13)

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N,-p}}{\mathbb{Z}_N} \ . \tag{2.15}$$

This equation has several complementary interpretations. First, we can say that the bulk produces a factor of our minimal theory $\mathcal{A}^{N,-p}$ on the boundary such that the combined boundary theory $\mathcal{T} \times \mathcal{A}^{N,-p}$ is anomaly free and then we can gauge the $\mathbb{Z}_N$ symmetry using the three steps above. Second, $\mathcal{T}'$ is as in (2.12), i.e. it includes only the $\mathbb{Z}_N$ invariant lines in $\mathcal{T}$. This means that it is obtained from $\mathcal{T}$ by applying only step 1 of the three-step gauging procedure above. And since $L = 1$ this leads to a consistent TQFT.

When $L \neq 1$ it is not meaningful to discuss the boundary theory, because it does not decouple from the bulk, which includes a non-trivial 4d TQFT. We could attempt to consider a 3d theory that consists only of the lines on the boundary and describes their correlation functions. We will find that these lines are the $\mathbb{Z}_N$ invariant lines from $\mathcal{T}$. This amounts to implementing step 1 of the three-step gauging procedure above. Because of the lack of decoupling from the bulk, the resulting theory is not a consistent 3d TQFT. It includes L lines that can move from the boundary to the 4d bulk and therefore they have trivial braiding with every line on the boundary. It is natural to consider a new effective theory obtained by performing a quotient by these lines.[10] In more detail, we performed step 1 of the three-step procedure above for $\mathbb{Z}_N$, and now we perform steps 2 and 3 with respect to the $\mathbb{Z}_L$ subgroup. The resulting TQFT is $\mathcal{T}'$ of

[10] This quotient is related to the discussion in [38–42].

(2.14)

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N} \tag{2.16}$$

and it is a fully consistent 3d TQFT. It captures the nontrivial correlation functions of the lines on the boundary. However, as we said above, $\mathcal{T}'$ is not "the theory on the boundary" except for $L = 1$. We will refer to it as "the effective boundary theory". We can think of the factor of $\mathcal{A}^{N/L,-p/L}$ as a 3d TQFT produced by the bulk so that the $\mathbb{Z}_N$ gauging can be performed.

We see that the 3d discussion of $\mathcal{T}'$ of (2.14) has a physical interpretation in terms of a 4d system with a boundary. We will discuss in detail the purely 3d system in Section 2 and the 4d interpretation in Section 3.

We will further generalize this discussion to interfaces between bulks with p^+ and p^-. Again, when $L^+ = L^- = 1$ there is a meaningful 3d theory on the interface. And for other values of $L^\pm$ there is only an effective description as above. It is

$$\frac{\mathcal{T} \otimes \mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,p^-/L^-}}{\mathbb{Z}_N} \ . \tag{2.17}$$

As in the case of a boundary, the two factors of $\mathcal{A}^{N/L^\pm,\mp p^\pm/L^\pm}$ can be interpreted as being produced by the bulk in the two sides such that the $\mathbb{Z}_N$ gauging can be performed on the interface.

In Section 4, we review the bulk dynamics of the $SU(N)$ and the $PSU(N)$ gauge theories and discuss their interfaces. Here we use the results in Section 3 to construct the interfaces in the $PSU(N)$ theory by gauging the one-form $\mathbb{Z}_N$ symmetry of the corresponding interfaces in $SU(N)$ theory.

In several appendices we summarize some background information and extend the analysis in the body of this chapter. Appendix A reviews the equivalence of different definitions of Abelian anyons and derives some useful facts we use in this chapter. Appendix B reviews the properties of the Jacobi symbols that appear in the central charge of the minimal Abelian TQFT $\mathcal{A}^{N,p}$. In Appendix C, we demonstrate that every Abelian TQFT corresponds to a unitary chiral RCFT.

In Appendix D, we prove the equivalence of different procedures that remove lines from a TQFT. Appendix E reviews and extends the analysis of a $\mathbb{Z}_N$ two-form gauge theory in 4d. In Appendix F, we generalize the discussion to a TQFT with an arbitrary Abelian one-form global symmetry group $\prod \mathbb{Z}_{N_I}$.

2.2 One-form symmetries in 3d and their gauging

2.2.1 One-form global symmetries in 3d TQFTs

In a 3d TQFT with a $\mathbb{Z}_N$ one-form symmetry, every line W is in some $\mathbb{Z}_N$ representation of charge $q(W)$. This means that the line transforms under a symmetry group element s by

$$a^s(\gamma)W(\gamma') = W(\gamma')e^{\frac{2\pi i s q(W)}{N}} \ , \tag{2.18}$$

where the symmetry transformation is implemented by the symmetry line a^s that braids with W as illustrated in Figure 2.2 with a the generating line of the symmetry. The charge $q(W)$ can be determined by the spins of the lines $h[W]$ [44] (for a later presentations see *e.g.* the mathematical treatment in [45] and a more physical review in [46])

$$q(W) = N\big(h[a] + h[W] - h[aW]\big) \text{ mod } N \ , \tag{2.19}$$

where aW denotes the unique line in the fusion of a and W. (The line aW is unique since a is an Abelian anyon as explained in Appendix A.)

For the special case $W = a^{s'}$, the transformation under the group element s is characterized by some integer P mod N

$$a^s(\gamma)a^{s'}(\gamma') = a^{s'}(\gamma')e^{-\frac{2\pi i s s' P}{N}} \ , \tag{2.20}$$

Using (2.19) we obtain

$$h[a^{s+s'}] - h[a^s] - h[a^{s'}] = \frac{Pss'}{N} \text{ mod } 1 \ . \tag{2.21}$$

Consider the case $s' = -s$. Since particles and their antiparticles have the same spin $h[a^s] = h[a^{-s}] \bmod 1$, and $h[1] = 0 \bmod 1$, we find two solutions with a given $P \bmod N$

$$h[a^s] = \frac{ps^2}{2N} \bmod 1, \quad p \in \{0, 1, ..., 2N-1\} \ , \tag{2.22}$$

with $p = P$ or $(P+N) \bmod 2N$.

The condition $a^N = 1$ in (2.6) requires that a^N has spin $\frac{pN}{2} = 0 \bmod 1$ and hence pN must be even. Therefore, for even N, the distinct cases are labeled by $p = 0, 1, ..., 2N-1$ and for odd N, they are labeled by $p = 0, 2, ..., 2N-2$.

Some different values of the label p can be identified using group automorphisms. For a $\mathbb{Z}_N$ one-form symmetry, this amounts to choosing a new generating line for the symmetry $\hat{a} = a^r$ with $\gcd(N, r) = 1$. The charge of a line W in the TQFT becomes $q(W)r \bmod N$. The new generating line $\hat{a}$ has spin $\frac{\hat{p}}{2N} \bmod 1$ with $\hat{p} = pr^2 \bmod 2N$ so the label p and $\hat{p} = pr^2 \bmod 2N$ should be identified.

In a spin TQFT there are new elements. These theories include a transparent spin-half line ψ. Using the language of one-form symmetries, we can say that ψ generates a $\mathbb{Z}_2$ one-from symmetry that does not act faithfully on the lines.

Consider first the case of even N. Here we can replace the generating line a with $\hat{a} = a\psi$, which also satisfies (2.6) $\hat{a}^N = 1$. The spin of $\hat{a}$ is $\frac{p}{2N} + \frac{1}{2} = \frac{p+N}{2N}$. Therefore, we can identify $p \sim p + N$. Equivalently, we can say that our system has a $\mathbb{Z}_N \otimes \mathbb{Z}_2$ one-form symmetry, where the first factor is generated either by a or by $\hat{a}$ and the second by ψ.

For odd N we could contemplate $a^N = \psi$ and therefore allow odd pN (and hence p is also odd). This means that a generates a $\mathbb{Z}_{2N}$ symmetry. Since N is odd, $\mathbb{Z}_{2N} \cong \mathbb{Z}_N \otimes \mathbb{Z}_2$. Here, the first factor is generated by $\hat{a} = a\psi$; indeed, $\hat{a}^N = 1$. The second factor is generated by ψ. The $\mathbb{Z}_N$ factor is characterized by the label $\hat{p} = (p+N) \bmod 2N$, which is even (because p and N are both odd). Therefore, without loss of generality, we can say that even in spin theories we impose that pN is even. (Alternatively, we can allow odd pN, but identify $p \sim p + N$.)

	even N	odd N
non-spin TQFT	$p=0,1,...,2N-1$	$p=0,2,...,2N-2$
spin TQFT	$p=0,1,...,N-1$	$p=0,1,...,N-1$ or equivalently $p=0,2,...,2N-2$

Table 2.1: The allowed labels p for $\mathbb{Z}_N$ one-form symmetry up to the redundancy in redefining the generators of the symmetries. A $\mathbb{Z}_N$ one-form symmetry of parameter p is generated by a line a of spin $h[a]=\frac{p}{2N}$ mod 1. For a non-spin TQFT, we need $pN\in 2\mathbb{Z}$, and $p\sim p+2N$. For a spin TQFT, we can use $pN\in\mathbb{Z}$ and $p\sim p+N$. Alternatively, we can say that in the spin case we keep the condition $pN\in 2\mathbb{Z}$ and add the identification $p\sim p+N$ only for even N.

The labels of distinct one-form symmetries for both non-spin and spin theories are summarized in Table 2.1. Recall that in addition, choosing a different generator for the $\mathbb{Z}_N$ symmetry changes p.

Examples

An example of a class of 3d TQFTs that has a $\mathbb{Z}_N$ one-form symmetry of all possible parameter $p=0,\cdots,2N-1$ mod $2N$ is the $U(1)_{pN}$ Chern-Simons theory. The symmetry lines of the $\mathbb{Z}_N$ one-form symmetry are generated by the Wilson line a of $U(1)$ charge p, and the line a^s for a general group element s has spin

$$h[a^s]=\frac{(ps)^2}{2pN}=\frac{ps^2}{2N} \mod 1\ , \tag{2.23}$$

in accordance with (2.22).

Another example is the simplest Abelian $\mathbb{Z}_N$ gauge theory in 3d, denoted by $(\mathcal{Z}_N)_0$. The theory has a $\mathbb{Z}_N\times\mathbb{Z}_N$ one-form symmetry, generated by the basic electric and magnetic lines V_E, V_M of integer spins. V_E generates a $\mathbb{Z}_N$ one-form symmetry with $p=0$ and V_M generates another $\mathbb{Z}_N$ one-form symmetry with $p=0$. However, these two lines V_E, V_M have a mutual braiding phase $e^{-2\pi i/N}$. This fact can be used to find a $\mathbb{Z}_N\subset\mathbb{Z}_N\times\mathbb{Z}_N$ of arbitrary even label p. Specifically, the line

$$b=V_E^{p/2}V_M\ , \tag{2.24}$$

generates a $\mathbb{Z}_N \subset \mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry and since its spin is $\frac{p}{2N}$ mod 1, the one-form symmetry is characterized by p.

What about the remaining lines? The line

$$c = V_E^{p/2} V_M^{-1} , \tag{2.25}$$

generates a $\mathbb{Z}_N$ one-form symmetry of even parameter $-p$ mod $2N$. However, the lines b and c satisfy

$$(bc)^{N/\gcd(N,p)} = 1 , \tag{2.26}$$

and therefore only when $\gcd(N,p) = 1$ do the two lines generate the entire $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry.

Let us study a third example. We consider $U(1)_N \otimes U(1)_{-N}$ (for N odd this is a spin TQFT) with gauge fields z and y and an action

$$\int \left(\frac{N}{4\pi} z dz - \frac{N}{4\pi} y dy \right) . \tag{2.27}$$

Writing it in terms of $x = z - y$, this action becomes

$$\int \left(\frac{N}{4\pi} x dx + \frac{N}{2\pi} x dy \right) , \tag{2.28}$$

and as in [29], it describes the $\mathbb{Z}_N$ DW theory [26] that we denote as $(\mathcal{Z}_N)_N$. It has a $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry, generated by $Z = \exp(i \oint z)$ of spin $\frac{1}{2N}$ mod 1, and $Y = \exp(i \oint y)$ of spin $-\frac{1}{2N}$ mod 1. The two lines Z and Y have trivial mutual braiding. The basic electric and magnetic lines of the DW $\mathbb{Z}_N$ gauge theory can be written as $V_E = ZY^{-1} = \exp(i \oint x)$ and $V_M = Y$. As in the previous example of $(\mathcal{Z})_0$, the line

$$b = Z^{(p+1)/2} Y^{-(p-1)/2} = V_E^{(p+1)/2} V_M , \tag{2.29}$$

generates a $\mathbb{Z}_N \subset \mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry of odd parameter $p \sim p + 2N$.

Again, we could ask about the remaining lines. The line

$$c = Z^{(p-1)/2} Y^{-(p+1)/2} = V_E^{(p-1)/2} V_M^{-1} \,. \tag{2.30}$$

generates a $\mathbb{Z}_N$ one-form symmetry of odd parameter $-p \bmod N$. As in the previous example, these lines satisfy a relation: $(bc)^{N/\gcd(N,p)} = 1$ and therefore only when $\gcd(N,p) = 1$ do the two lines b and c generate the entire $\mathbb{Z}_N \times \mathbb{Z}_N$ one-form symmetry.

Let us summarize the last two examples. A subset of the lines of $(\mathcal{Z}_N)_0$ generates a $\mathbb{Z}_N$ one-form symmetry with even parameter p and a subset of the lines of $(\mathcal{Z}_N)_N$ generates a $\mathbb{Z}_N$ one-form symmetry with odd parameter p. When $\gcd(N,p) = 1$ the remaining lines also generate a $\mathbb{Z}_N$ one-form symmetry with parameter $-p$.

We can combine these two examples more concisely using the theory $(\mathcal{Z}_N)_{-pN}$ with the action

$$\int \left(-\frac{pN}{4\pi} x dx + \frac{N}{2\pi} x dy \right) \,. \tag{2.31}$$

Here the parameter p can be identified with $p + 2$ using the redefinition $y \to y - x$ so these theories are either $(\mathcal{Z}_N)_0$ or $(\mathcal{Z}_N)_N$, and the lines b and c in $(\mathcal{Z}_N)_0$ and $(\mathcal{Z}_N)_N$ are mapped to the following lines in $(\mathcal{Z}_N)_{-pN}$

$$b = \exp(i \oint y), \quad c = \exp(ip \oint x - i \oint y) \,. \tag{2.32}$$

2.2.2 The minimal Abelian TQFT $\mathcal{A}^{N,p}$

In this section, we will show that when $\gcd(N,p) = 1$ and $pN \in 2\mathbb{Z}$ the N symmetry lines associated to a $\mathbb{Z}_N$ one-form symmetry form a consistent TQFT. We call this theory "the minimal Abelian TQFT" and denote it by $\mathcal{A}^{N,p}$. This theory was first studied in [44] and more recently in [47, 6]. Here we emphasize its one-form global symmetry and show how it appears as a sub-theory in TQFTs with a $\mathbb{Z}_N$ one-form global symmetry.[11]

Using the assumed underlying $\mathbb{Z}_N$ one-form symmetry, we can simplify the discussion in [44]. The symmetry determines the spins of the lines $h[a^s] = \frac{ps^2}{2N}$ mod 1, and their braiding leads to the following S matrix

$$S_{ss'} = \frac{1}{\sqrt{N}} \exp\Big(2\pi i \left(h[s] + h[s'] - h[ss']\right)\Big) = \frac{1}{\sqrt{N}} \exp\left(-\frac{2\pi i p}{N} ss'\right), \quad s, s' \in \{1, ..., N\} \ . \tag{2.33}$$

This matrix is unitary only when $L = \gcd(N, p) = 1$. (If $L = \gcd(N, p) \neq 1$, the line $a^{N/L}$ has trivial braiding with all the lines in the theory, so the S matrix is not unitary.)

The chiral central charge $c_N^{(p)}$ modulo 8 of the Abelian TQFT $\mathcal{A}^{N,p}$ can be computed using the following formula (see *e.g.* [48, 46])[12]

$$e^{i\frac{2\pi}{8}c_N^{(p)}} = \frac{1}{\sqrt{N}} \sum_{s=1}^{N} e^{2\pi i h[a^s]} \ . \tag{2.34}$$

The summation is a Gaussian sum with the following closed-form expression [44, 49]

$$\exp\left(i\frac{2\pi}{8}c_N^{(p)}\right) = \begin{cases} \left(\dfrac{p/2}{N}\right)\epsilon(N) & N \text{ odd}, p \text{ even} \\ \left(\dfrac{N/2}{p}\right)\epsilon(p)^{-1}\exp\left(\pi i/4\right) & N \text{ even}, p \text{ odd} \end{cases} , \tag{2.35}$$

where $\epsilon(s) = 1$ for $s = 1$ mod 4, $\epsilon(s) = i$ for $s = -1$ mod 4 and $\left(\frac{a}{b}\right)$ is the Jacobi symbol reviewed in Appendix B. The values of the chiral central charges are summarized in Table 2.2, and they are always integers.

Every Abelian TQFT can be represented by some Abelian Chern-Simons theory [50–54] (for a review see *e.g.* [55]). It is also true for $\mathcal{A}^{N,p}$. For example,[13] $\mathcal{A}^{N,1} \longleftrightarrow U(1)_N$ and

[11] A putative theory with N Abelian lines a^s with $h(a^s) = \frac{ps^2}{2N}$ is not a consistent (modular) TQFT when $\gcd(N, p) \neq 1$.

[12] The chiral central charge of a TQFT can be shifted by adding a $(E_8)_1$ theory, since it has $c = 8$ and no nontrivial lines.

[13] Typically (and perhaps always) the TQFT can be described by a Chern-Simons (CS) gauge theory and a corresponding Rational Conformal Field Theory (RCFT). In fact, there are often several distinct CS theories corresponding to the same TQFT. Then the symbol $\longleftrightarrow$ means that they are dual. It is important to stress,

Table 2.2: The chiral central charge $c_N^{(p)}$ mod 8 of the minimal Abelian theory $\mathcal{A}^{N,p}$ computed from (2.35). For each case $c_N^{(p)}$ mod 8 is one of the two possible values depending on $[N] = N$ mod 4 and $[p] = p$ mod 4. Here $\times$ means that the theories with such p and N do not exist according to the conditions that pN is even and $\gcd(N,p) = 1$.

$[N]$ \ $[p]$	0	1	2	3
0	$\times$	1,5	$\times$	3,7
1	0,4	$\times$	0,4	$\times$
2	$\times$	1,5	$\times$	3,7
3	6,2	$\times$	6,2	$\times$

$\mathcal{A}^{N,N-1} \longleftrightarrow SU(N)_1$. An alternative description of $\mathcal{A}^{N,N-1}$ is the $U(1)^{N-1}$ Chern-Simons theory with the coefficient matrix given by the Cartan matrix of $SU(N)$ (see *e.g.* [48]). The dualities also hold after taking orientation-reversal.

Similar to one-form symmetries, any two minimal Abelian TQFTs $\mathcal{A}^{N,p}$ and $\mathcal{A}^{N,pr^2}$ with $\gcd(N,r) = 1$ are related by group automorphsims.

Following the discussion of spin TQFTs in the previous subsection we can generalize the minimal theory to spin theories. Originally, we imposed $a^N = 1$ and then pN has to be even and the minimal theory is nonspin. We can make it into a spin TQFT by tensoring the almost trivial theory[14] $\{1,\psi\}$. After doing that, for odd N we can further redefine $a \to a\psi$, which makes $a^N = \psi$ and shifts $p \to p+N$ making pN odd. This way we can define a spin TQFT

$$\mathcal{A}^{N,p} \ \equiv \ \mathcal{A}^{N,p+N} \otimes \{1,\psi\} \qquad \text{for odd } pN \text{ and } \gcd(N,p) = 1 \ . \tag{2.36}$$

This is the minimal spin TQFT generated by a line of spin $\frac{p}{2N}$ mod 1.

however, that distinct RCFTs with the same TQFT are often inequivalent.

[14]The almost trivial TQFT $\{1,\psi\}$ can be represented by $SO(M)_1$ for some integer M. The dependence on M is only in the framing anomaly or equivalently in the chiral central charges $c = \frac{M}{2}$. See *e.g.* Appendix C of [56], Appendix B of [57], and also [31].

As an application, the spin TQFT $U(1)_N$ for odd N factorizes[15]

$$U(1)_N \longleftrightarrow \mathcal{A}^{N,N+1} \otimes \{1,\psi\} \ , \tag{2.37}$$

where the first factor is a nonspin minimal theory. Since $\mathcal{A}^{N,N+1} = \mathcal{A}^{N,-N+1} \longleftrightarrow SU(N)_{-1}$. This reproduces the level-rank duality $U(1)_N \longleftrightarrow SU(N)_{-1}$, which is valid only as spin TQFTs [31].[16]

2.2.3 Factorization of 3d TQFTs when $\gcd(N,p) = 1$

In this section we show that a TQFT $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry of label p such that $\gcd(N,p) = 1$ factorizes as

$$\mathcal{T} = \mathcal{A}^{N,p} \otimes \mathcal{T}' \qquad \text{when} \ \gcd(N,p) = 1\,. \tag{2.38}$$

This is quite surprising. It means that in this case all the information about the global symmetry and its action on $\mathcal{T}$ is included in a decoupled factor of the minimal theory $\mathcal{A}^{N,p}$ and $\mathcal{T}'$ is invariant under the symmetry.[17]

The theory $\mathcal{T}$ includes the $\mathbb{Z}_N$ symmetry lines a^s. When $\gcd(N,p) = 1$ these lines form the minimal theory $\mathcal{A}^{N,p}$. Next, consider any line $W \in \mathcal{T}$. Since a is Abelian, the fusion of W with a includes a single line rather than a sum of lines. (See Appendix A.) Therefore, since $\gcd(N,p) = 1$, we can always find an integer s such that the line $W' = Wa^s$ has vanishing charge $q(W') = 0 \bmod N$. Denote the set of neutral lines W' by $\mathcal{T}'$. This shows that every line $W \in \mathcal{T}$ is a product of a line $W' \in \mathcal{T}'$ and a line in $\mathcal{A}^{N,p}$. It is clear that all the conditions of a consistent TQFT are satisfied separately for $\mathcal{T}'$ and $\mathcal{A}^{N,p}$ and hence we have the factorization (2.38).

[15]We use equal sign to relate two isomorphic TQFTs. However, we used $\longleftrightarrow$ to denote two dual presentations of the same TQFT. Typically one or both of these presentations is given by a Chern-Simons gauge theory. Then the classical Chern-Simons theories are not equal (hence we do not use an equal sign), but the quantum theories are the dual.

[16]If $N = 8n$ for some integer n, the non-spin minimal Abelian TQFT satisfies $\mathcal{A}^{N,1} = \mathcal{A}^{N,N+1}$ by redefining the generating line $a \to a^{4n+1}$. Thus $U(1)_{8n} \longleftrightarrow SU(8n)_{-1}$ are dual as non-spin TQFTs in agreement with [58].

[17]If the theory $\mathcal{T}$ is a spin TQFT, then since the transparent spin-half line is invariant under any one-form symmetry, the theory $\mathcal{T}'$ also contains such a line and is a spin TQFT.

The factorization (2.38) also follows from a theorem in modular tensor category (see [40] and Theorem 3.13 in [41]). In physics language, the theorem states that if a 3d TQFT $\mathcal{T}$ has a consistent sub-theory $\mathcal{A}$, then $\mathcal{T}$ factorizes into $\mathcal{A}\otimes\mathcal{T}'$ where $\mathcal{T}'$ is another consistent TQFT that consists of all the lines in $\mathcal{T}$ that have trivial braiding with the lines in $\mathcal{A}$.[18]

Next, we use the fact that $(\mathcal{A}^{N,p}\otimes\mathcal{A}^{N,-p})/\mathbb{Z}_N$ is a trivial theory, where the quotient means gauging the anomaly free diagonal $\mathbb{Z}_N$ one-form symmetry generated by the two generating lines of the minimal Abelian TQFTs. This leads to an alternative presentation of the TQFT $\mathcal{T}'$

$$\mathcal{T}'=\frac{\mathcal{T}\otimes\mathcal{A}^{N,-p}}{\mathbb{Z}_N}~, \tag{2.39}$$

where the quotient means gauging the anomaly free diagonal $\mathbb{Z}_N$ one-form symmetry generated by the symmetry generating line a in $\mathcal{T}$ and the generating line of $\mathcal{A}^{N,-p}$.

Let us demonstrate this factorization in some examples.

The minimal Abelian TQFTs can be found as sub-theories in various examples discussed in Section 2.2.1. We start by considering $U(1)_{pN}$ when $\gcd(N,p)=1$. The theory has a $\mathbb{Z}_{pN}\cong\mathbb{Z}_N\otimes\mathbb{Z}_p$ one-form symmetry with a $\mathbb{Z}_N$ subgroup generated by a, the Wilson line of charge p, and a $\mathbb{Z}_p$ subgroup generated by b, the Wilson line of charge N. The line a and the line b each generates a minimal Abelian TQFT $\mathcal{A}^{N,p}$ and $\mathcal{A}^{p,N}$. The full theory factorizes into these minimal Abelian TQFTs[19]

$$U(1)_{pN}\;\longleftrightarrow\;\mathcal{A}^{pN,1}=\mathcal{A}^{N,p}\otimes\mathcal{A}^{p,N}\;\text{ when }\;\gcd(N,p)=1\,. \tag{2.40}$$

To show the factorization of an Abelian TQFT, it is sufficient to check the factorization in the fusion rules, the spins of the lines and the chiral central charge. The fusion rules of $U(1)_{pN}$ are the same as the group law of $\mathbb{Z}_{pN}$. When $\gcd(N,p)=1$, the group factorizes into $\mathbb{Z}_N\times\mathbb{Z}_p$ and

[18]We thank Zhenghan Wang for discussions about this point.

[19]For odd pN the full theory $U(1)_{pN}$ as well as $\mathcal{A}^{N,p}$ and $\mathcal{A}^{p,N}$ are spin TQFTs. The spin Chern-Simons theory $U(1)_{pN}$ can also factorize as $U(1)_{pN}\longleftrightarrow\mathcal{A}^{N,p+N}\otimes\mathcal{A}^{p,p+N}\otimes\{1,\psi\}$ (compare with (2.37)), where the first two factors are non-spin minimal theories.

every line in the theory can be decomposed into $W = a^s b^r$ with some unique $(s, r) \in \mathbb{Z}_N \times \mathbb{Z}_p$. The spins of the lines also factorize

$$h[W] = \frac{(ps + Nr)^2}{2pN} = \left(\frac{p}{2N}s^2 + \frac{N}{2p}r^2\right) \bmod 1 = (h[a^s] + h[b^r]) \bmod 1 \,. \tag{2.41}$$

The chiral central charge of $U(1)_{pN}$ is $c = 1$. It agrees with the sum of the chiral central charges of individual sub-theories up to a periodicity of 8

$$e^{i\frac{2\pi}{8}\left(c_N^{(p)} + c_p^{(N)}\right)} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{\pi i p j^2}{N}} \frac{1}{\sqrt{p}} \sum_{k=0}^{p-1} e^{\frac{\pi i N k^2}{p}} = \frac{1}{\sqrt{pN}} \sum_{j,k} e^{\frac{2\pi i (pj + Nk)^2}{2pN}} = e^{i\frac{2\pi}{8}} \,. \tag{2.42}$$

We conclude that $U(1)_{pN}$ factorizes into $\mathcal{A}^{N,p} \otimes \mathcal{A}^{p,N}$ when $\gcd(N, p) = 1$.

The minimal Abelian TQFT $\mathcal{A}^{N,p}$ is also a sub-theory in $(\mathcal{Z}_N)_{-pN}$ when $\gcd(N, p) = 1$. Similarly, the theory also factorizes

$$(\mathcal{Z}_N)_{-pN} \longleftrightarrow \mathcal{A}^{N,p} \otimes \mathcal{A}^{N,-p} \text{ when } \gcd(N, p) = 1 \,, \tag{2.43}$$

where $\mathcal{A}^{N,p}$ and $\mathcal{A}^{N,-p}$ are generated by the lines b and c in $(\mathcal{Z}_N)_{-pN}$ defined in (2.32).

As a consistency check, combining (2.38) and (2.39) and using the factorization property of $(\mathcal{Z}_N)_{-pN}$ in (2.43), we recover the following canonical duality [59]

$$\mathcal{T} \longleftrightarrow \frac{\mathcal{T} \otimes (\mathcal{Z}_N)_{-pN}}{\mathbb{Z}_N} \longleftrightarrow \begin{cases} \dfrac{\mathcal{T} \otimes (\mathcal{Z}_N)_0}{\mathbb{Z}_N} & \text{even } p \\ \dfrac{\mathcal{T} \otimes (\mathcal{Z}_N)_N}{\mathbb{Z}_N} & \text{odd } p \end{cases} , \tag{2.44}$$

where the quotient means gauging the anomaly free diagonal one-form symmetry generated by the line a in $\mathcal{T}$ and the line c in the $\mathbb{Z}_N$ gauge theories defined in (2.25), (2.30) and (2.32). The duality holds even when $\gcd(N, p) \neq 1$. Under the duality, the symmetry generating line a in $\mathcal{T}$ is mapped to the line b in the dual theories defined in (2.24), (2.29) and (2.32). Then the $\mathbb{Z}_N$ one-form symmetry is entirely in the $(\mathcal{Z}_N)_{-pN}$ factor.

We remark that although the 3d TQFT factorizes, the corresponding 2d RCFTs may not factorize since the unitary modular tensor category does not fully specify the 2d chiral conformal field theory [60]. For Abelian TQFTs, we provide a construction of a corresponding unitary chiral RCFT in Appendix C.

2.2.4 't Hooft anomaly of one-form global symmetries

Consider a 3d TQFT $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry of label p with the symmetry generating line a. Gauging the one-form symmetry amounts to summing over all possible insertions of the symmetry lines [5]. If the symmetry lines have non-integer spin, the partition function vanishes because of the summation. This means that the one-form symmetry has a 't Hooft anomaly unless $p = 0 \bmod 2N$. Indeed, the one-form symmetry of label $p = 0 \bmod 2N$ can be gauged following the procedure outline in Section 2.1. (When $p = N$, the theory can also be gauged as a spin TQFT by redefining the symmetry generating line using the transparent spin-half line. After gauging it becomes a spin TQFT, even though the original theory can be a non-spin theory. It reflects a mixed 't Hooft anomaly between the one-form symmetry and gravity, which we will explain in details later.)

We couple the one-form symmetry of the 3d TQFT to a classical $\mathbb{Z}_N$ two-form gauge field $\mathcal{B}_\mathcal{C} \in H^2(\mathcal{M}_4, \mathbb{Z}_N)$.[20] The anomaly of the one-form symmetry is characterized by a 4d term of the gauge field $\mathcal{B}_\mathcal{C}$ through anomaly inflow. To determine the 4d term, we use the canonical duality in (2.44) [59]

$$\mathcal{T} \longleftrightarrow \frac{\mathcal{T} \otimes (\mathcal{Z}_N)_{-pN}}{\mathbb{Z}_N} \,. \tag{2.45}$$

Under the duality, the original $\mathbb{Z}_N$ one-form symmetry in $\mathcal{T}$ is mapped to the one-form symmetry generated by line b defined in (2.32) in the dual description so the theory on the right hand side couples to the classical gauge field $\mathcal{B}_\mathcal{C}$ through the $(\mathcal{Z}_N)_{-pN}$ factor. It was shown in [29] that the

[20] The subscript $\mathcal{C}$, as in $B_\mathcal{C}$, denotes that the gauge field is classical.

anomaly of $(\mathcal{Z}_N)_{-pN}$ is cancelled by the 4d term

$$2\pi \frac{p}{2N} \int_{\mathcal{M}_4} \mathcal{P}(\mathcal{B}_{\mathcal{C}}) \ , \tag{2.46}$$

where $\mathcal{P}$ is the Pontryagin square operation (for a review see e.g. [34, 61, 10]). Therefore, the anomaly of a $\mathbb{Z}_N$ one-form symmetry of label p is characterized by the 4d term (2.46) [29, 5, 10].

The 4d term (2.46) is consistent with the $\mathbb{Z}_N$ periodicity of the $\mathcal{B}_{\mathcal{C}}$ field only for even pN. Furthermore, for $p = N$ (which is possible only for even N) it can be written as

$$\pi \int_{\mathcal{M}_4} \mathcal{P}(\mathcal{B}_{\mathcal{C}}) = \left(\pi \int_{\mathcal{M}_4} \mathcal{B}_{\mathcal{C}} \cup \mathcal{B}_{\mathcal{C}}\right) \mod 2\pi = \left(\pi \int_{\mathcal{M}_4} w_2(\mathcal{M}_4) \cup \mathcal{B}_{\mathcal{C}}\right) \mod 2\pi \ , \tag{2.47}$$

where $w_2(\mathcal{M}_4) \in H^2(\mathcal{M}_4, \mathbb{Z}_2)$ is the second Stiefel-Whitney classe of the manifold $\mathcal{M}_4$ (see *e.g.* [62,63]). Equation (2.47) follows from the identity $x \cup x = w_2(\mathcal{M}_4) \cup x$ for $x = (\mathcal{B}_{\mathcal{C}} \mod 2) \in H^2(\mathcal{M}_4, \mathbb{Z}_2)$ (on orientable manifolds). We interpret the 4d term (2.47) as a mixed 't Hooft anomaly between the one-form symmetry and gravity (fermion parity), which means that when this anomaly exists the one-form symmetry can be gauged only on spin manifolds. See also the related discussion in appendix E.

On spin manifolds, pN in (2.46) can be odd. Furthermore, (2.46) vanishes for $p = N$.

In summary, on non-spin manifolds, the anomaly is labeled by $p = 0, 1, ..., 2N - 1$ for even N and $p = 0, 2, ..., 2N - 2$ for odd N, and on a spin manifolds, the anomaly is labeled by $p = 0, 1, ..., N - 1$. This agrees with the labels of 3d $\mathbb{Z}_N$ one-form symmetries listed in Table 2.1.

The anomaly can be changed by choosing a different generating line $\hat{a} = a^r$ with $\gcd(N, r) = 1$ as explained in Section 2.2.1. It is equivalent to redefining the classical gauge field $\mathcal{B}_{\mathcal{C}}$ by a multiplication by r and the anomaly coefficient in (2.46) becomes $pr^2 \mod 2N$.

In the presence of the classical gauge field $\mathcal{B}_{\mathcal{C}}$, the line W is dressed with an open surface $e^{-\frac{2\pi i q(W)}{N} \int \mathcal{B}_{\mathcal{C}}}$ for gauge invariance and the redefinition of the classical gauge field $\mathcal{B}_{\mathcal{C}}$ rescales the charge from $q(W)$ to $q(W)r$.

An anomalous $\mathbb{Z}_N$ one-form symmetry can have anomaly free subgroups. On spin manifolds, a $\mathbb{Z}_m$ subgroup is anomaly free if the symmetry generator $\hat{a} = a^{N/m}$ has integer or half-integer spin

$$h[\hat{a}] = \frac{pN}{2m^2} \in \frac{1}{2}\mathbb{Z} \ . \tag{2.48}$$

There is always a $\mathbb{Z}_L$ subgroup with $L = \gcd(N, p)$ that satisfies this condition and hence it is anomaly free. But the $\mathbb{Z}_L$ subgroup may not be the maximal anomaly free subgroup. For $NL = r^2 t$ with some integers r, t such that t does not contain any complete-square divisors great than one, the maximal anomaly free subgroup is $\mathbb{Z}_r$. As a non-spin TQFT, a $\mathbb{Z}_m$ subgroup is anomaly free only if $h[\hat{a}] \in \mathbb{Z}$ and therefore the $\mathbb{Z}_L$ subgroup is anomaly free only for even pN/L^2.

2.2.5 A generalization of the three-step gauging procedure to anomalous theories

In this subsection, we will introduce a new operation on 3d TQFTs that generalizes the three-step gauging procedure outlined in Section 2.1. This generalized operation will appear naturally in Section 2.3, where we consider 4d theories with boundaries and interfaces.

The standard gauging procedure of an anomaly free $\mathbb{Z}_N$ one-form symmetry can be used when $p = 0$, where all the symmetry lines have integer spins. Then in step 1 we remove the non-invariant lines, in step 2 we identify lines that differ by the fusion with the symmetry lines, and in step 3 we take lines at fixed points of the identification several times.

When $p = N$ this simple process cannot be repeated because the generating line a has half-integer spin. As we said above, this can be interpreted as a mixed anomaly between the one-form symmetry and gravity. This anomaly vanishes on spin manifolds and therefore, we can gauge the symmetry and find a spin TQFT. Let us discuss it in more detail. If the original TQFT $\mathcal{T}$ is a spin theory, it has a transparent spin-half line ψ. Otherwise, we make it into a spin TQFT by tensoring the almost trivial theory $\{1, \psi\}$. Now that we have a spin TQFT we can redefine $a \to \hat{a} = a\psi$. Since $p = N$ and pN is even, this occurs only for even N and then the redefinition preserves the fact that $a^N = 1$. The redefinition shifts p to be zero. As a result, even in this case we can use the standard three-step gauging process with $\hat{a}$. The only difference is that the theory is spin.

For simplicity from this point on we will limit ourselves to spin TQFTs.

Consider a 3d spin TQFT $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry of label p such that $\gcd(N,p)=1$, the spin TQFT factorizes as discussed in Section 2.2.3

$$\mathcal{T} = \mathcal{T}' \otimes \mathcal{A}^{N,p}\,, \tag{2.49}$$

where $\mathcal{T}'$ is the 3d spin TQFT that consists of all the $\mathbb{Z}_N$ invariant lines in $\mathcal{T}$, and it can be extracted through

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N,-p}}{\mathbb{Z}_N}\,. \tag{2.50}$$

In this case, we define an operation that maps $\mathcal{T}$ to $\mathcal{T}'$. The operation discards all the $\mathbb{Z}_N$ non-invariant lines in $\mathcal{T}$. It is equivalent to applying only the step 1 of the three-step gauging procedure.

When $\gcd(N,p) \neq 1$, the $\mathbb{Z}_N$ one-form symmetry has an anomaly free $\mathbb{Z}_L$ subgroup generated by $\hat{a} = a^{N/L}$ with $L = \gcd(N,p)$. Gauging this $\mathbb{Z}_L$ subgroup produces a new spin TQFT $\mathcal{T}/\mathbb{Z}_L$. The new spin TQFT contains the original symmetry generating line a, but now it generates a $\mathbb{Z}_{N'}$ one-form symmetry ($N' = N/L$) with label $p' = p/L$. Since $\gcd(N',p') = 1$, $\mathcal{T}/\mathbb{Z}_L$ factorizes

$$\frac{\mathcal{T}}{\mathbb{Z}_L} = \left(\frac{\mathcal{T}}{\mathbb{Z}_L}\right)' \otimes \mathcal{A}^{N/L,p/L}\,, \tag{2.51}$$

where $(\mathcal{T}/\mathbb{Z}_L)'$ contains all the lines in $\mathcal{T}/\mathbb{Z}_L$ that have trivial braiding with a. We define the generalized gauging operation that maps

$$\mathcal{T} \to \mathcal{T}' \equiv \left(\frac{\mathcal{T}}{\mathbb{Z}_L}\right)' = \frac{\mathcal{T}/\mathbb{Z}_L \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_{N/L}} = \frac{\mathcal{T} \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N}\,. \tag{2.52}$$

In both presentations, the quotient in the denominator uses the symmetry generator a and the generating line of the minimal Abelian TQFT. In the second presentation, the $\mathbb{Z}_L$ subgroup of the $\mathbb{Z}_N$ quotient acts only on $\mathcal{T}$.

There are three ways to think about the map (2.52).

First, as we motivated it and as in the first presentation in (2.52), we first gauge the $\mathbb{Z}_L$ subgroup of the $\mathbb{Z}_N$ one-form symmetry and then remove the sub-theory in $\mathcal{T}/\mathbb{Z}_L$ consisting of the $\mathbb{Z}_{N/L}$ symmetry lines.

Second, since the $\mathbb{Z}_N$ symmetry is anomalous, we tensor a minimal theory $\mathcal{A}^{N/L,-p/L}$ that cancels the anomaly and then gauge the new anomaly free $\mathbb{Z}_N$ symmetry. This is clear in the second presentation in (2.52).

Third, we can perform step 1 of the three-step gauging procedure using the full $\mathbb{Z}_N$ symmetry and then perform steps 2 and 3 using only its $\mathbb{Z}_L$ subgroup:

Step 1 Select the lines invariant under the $\mathbb{Z}_N$ one-form symmetry. In particular, among the symmetry lines, only the ones associated to the $\mathbb{Z}_L$ subgroup generated by $\hat{a} = a^{N/L}$ remain.

Step 2 If $\hat{a}$ has integer spin, identify $W \sim W\hat{a}$ and if $\hat{a}$ has half-integer spin, identify $W \sim W\hat{a}\psi$.

Step 3 Take multiple copies at the fixed points of the identification.

When $p = 0, N$, the symmetry is anomaly free and the generalized gauging operation reduces to the standard gauging procedure that produces $\mathcal{T}' = \mathcal{T}/\mathbb{Z}_N$.

In general, the $\mathbb{Z}_N$ one-form symmetry can have larger anomaly free $\mathbb{Z}_m$ subgroups that contain the $\mathbb{Z}_L$ subgroup. In Appendix D we show that the same result (2.52) can be reproduced if we first gauge the $\mathbb{Z}_m$ subgroup and then apply the generalized gauging operation to the remaining theory (up to a possible transparent spin-half line, which we will ignore).

Below we will see similar operations on TQFTs, which are not minimal. Following the second presentation in (2.52), we can tensor not the minimal theory $\mathcal{A}^{N/L,-p/L}$, but other theories that cancel the anomaly, e.g.

$$\frac{\mathcal{T} \otimes \mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,p^-/L^-}}{\mathbb{Z}_N}\,, \tag{2.53}$$

where $p = p^+ - p^-$ and $L^\pm = \gcd(N, p^\pm)$. The operation adds to the theory $\mathcal{T}$ two minimal Abelian TQFTs to cancel the anomaly and then gauges the diagonal one-form symmetry. The two minimal Abelian TQFTs $\mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,-p^-/L^-}$ always have greater or equal number of lines than $\mathcal{A}^{N/L,-p/L}$ with $L = \gcd(N,p)$ and $p = p^+ - p^-$.[21] All the lines in $\mathcal{T}'$ defined in (2.52) can be identified with the lines from the original TQFT $\mathcal{T}$. In contrast, the theory (2.53) in general has additional lines.

2.3 Coupling to a 4d bulk

2.3.1 The bulk coupling

Consider a 4d symmetry protected topological (SPT) phase of $\mathbb{Z}_N$ one-form symmetry with the same action as the anomaly in (2.46)

$$2\pi \frac{p}{2N} \int_{\mathcal{M}_4} \mathcal{P}(\mathcal{B}_\mathcal{C}) \ , \tag{2.54}$$

where $\mathcal{B}_\mathcal{C} \in H^2(\mathcal{M}_4, \mathbb{Z}_N)$ is a classical $\mathbb{Z}_N$ two-form gauge field. The theory has a description, reviewed in Appendix E, in terms of a dynamical $U(1)$ one-form gauge field A and a classical $U(1)$ two-form gauge field $B_\mathcal{C}$

$$\int_{\mathcal{M}_4} \left(\frac{pN}{4\pi} B_\mathcal{C} B_\mathcal{C} + \frac{N}{2\pi} B_\mathcal{C} dA \right) \ . \tag{2.55}$$

The equation of motion of A constrains $B_\mathcal{C}$ to be a $\mathbb{Z}_N$ two-form gauge field $\frac{2\pi}{N}\mathcal{B}_\mathcal{C}$.

The theory (2.55) is invariant under a one-form gauge transformation of background fields

$$B_\mathcal{C} \to B_\mathcal{C} - d\lambda, \quad A \to A + p\lambda \,, \tag{2.56}$$

[21] $\mathcal{A}^{N/L,-p/L}$ has N/L lines and $\mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,p^-/L^-}$ has N^2/L^+L^- lines. The ratio between them is $\frac{NL}{L^+L^-} = \left(\frac{N\ell}{L^+L^-}\right)\left(\frac{L}{\ell}\right)$ with $\gcd(N, p^+, p^-) = \ell$. Since the two factors are integers, the product theory has more lines.

with λ a one-form gauge parameter.

We put the theory on a 4-manifold $\mathcal{M}_4$ with a boundary.[22] The action is gauge invariant under (2.56) up to a boundary term

$$-\int_{\partial\mathcal{M}_4}\left(\frac{pN}{4\pi}\lambda d\lambda+\frac{N}{2\pi}\lambda dA\right). \tag{2.57}$$

It can be cancelled by a theory on the boundary with a $\mathbb{Z}_N$ one-form symmetry of anomaly p, that couples to the classical gauge field $B_\mathcal{C}$. So we are going to place on the boundary an arbitrary TQFT $\mathcal{T}$ with such a symmetry and anomaly.

The coupling of the boundary TQFT $\mathcal{T}$ to the classical gauge field $B_\mathcal{C}$ has a convenient Lagrangian description using the canonical duality in (2.44) [59]

$$\mathcal{T}\quad\longleftrightarrow\quad\frac{\mathcal{T}\otimes(\mathcal{Z}_N)_{-pN}}{\mathbb{Z}_N} \tag{2.58}$$

and the Lagrangian description (2.31) of the second factor in the numerator. Then the classical gauge field $B_\mathcal{C}$ couples to the boundary theory through the $(\mathcal{Z}_N)_{-pN}$ theory

$$\int_{\partial\mathcal{M}_4}\left(-\frac{pN}{4\pi}xdx+\frac{N}{2\pi}xdy+\frac{N}{2\pi}B_\mathcal{C}y-\frac{N}{2\pi}B_\mathcal{C}A\right), \tag{2.59}$$

where the last term $B_\mathcal{C}A$ can be absorbed into the bulk action by modifying $B_\mathcal{C}dA$ to $AdB_\mathcal{C}$. Now the one-form gauge transformation (2.56) acts as

$$B_\mathcal{C}\to B_\mathcal{C}-d\lambda,\quad A\to A+p\lambda,\quad x\to x+\lambda,\quad y\to y+p\lambda\,. \tag{2.60}$$

2.3.2 Gauge the one-form symmetry

The whole system is anomaly free so there is no obstruction to gauging the one-form symmetry by turning the background gauge field $B_\mathcal{C}$ into a dynamical gauge field denoted by B. After

[22] We restrict to the 4-manifolds such that every $\mathbb{Z}_N$ two-form gauge field on the boundary can be extended to the bulk. It requires the third relative cohomology $H^3(\mathcal{M}_4,\partial\mathcal{M}_4;\mathbb{Z}_N)$ to vanish.

gauging, the bulk theory becomes a dynamical $\mathbb{Z}_N$ two-form gauge theory reviewed in Appendix E. For later convenience, we define

$$L = \gcd(N, p), \quad K = N/L\,. \tag{2.61}$$

The bulk $\mathbb{Z}_N$ two-form gauge theory is effectively a $\mathbb{Z}_L$ one-form gauge theory[23] [29, 5]. It has L genuine line operators generated by $V = \exp(iK \oint A)$ and L surface operators generated by $U = \exp(i \oint B)$. We will be interested in the effect of gauging on the boundary TQFT. For simplicity, we will limit ourselves to spin 4-manifolds.

It is important to stress that when $L \neq 1$ the bulk theory is nontrivial and hence it is meaningless to ask what the 3d theory on the boundary is. Instead, it should be thought of as part of the 4d-3d system. Nevertheless, we can discuss the physical observables such as the line operators on the boundary and their correlation functions. We will extract from the 4d-3d system an effective boundary theory that reproduces many of these observables.

Let us examine the line operators on the boundary. The bulk $\mathbb{Z}_N$ gauge theory has L line operators. When they are restricted to the boundary, they are regarded as boundary lines. But they have trivial braiding with all the boundary lines since they can smoothly move into the bulk and get un-braided. This means that unless $L = 1$ (where the bulk is trivial) the boundary lines to do not form a modular TQFT [29].

What are the other lines on the boundary? They can be constructed by fusing a line W from the 3d TQFT $\mathcal{T}$ and the bulk lines generated by $\exp(i \oint A)|$, where $|$ denotes the restriction to the boundary

$$W(\gamma) \exp\big(im \oint_\gamma A\big) \exp\big(i(mp - q(W)) \int_\Sigma B\big) \qquad \text{with} \qquad \gamma = \partial\Sigma\ , \tag{2.62}$$

where $m \sim m + N$. The coupling to B is needed for the one-form gauge symmetry. Next, we

[23]When pN/L^2 is odd, $V = \exp(iK \oint A)$ represents the worldline of a fermionic particle and the bulk theory is effectively a $\mathbb{Z}_L$ gauge theory that couples to $w_2(\mathcal{M}_4)$ of the manifold (see Appendix E).

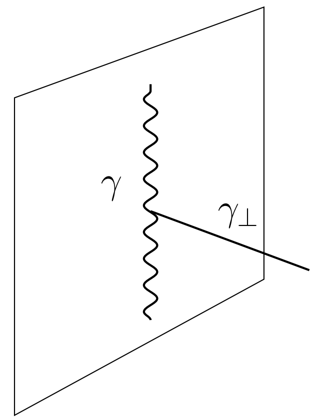


Figure 2.3: If a boundary line $W(\gamma)$ is at the fixed point of the identification using $\hat{a} = a^K$, it can form a junction by emanating a bulk line $\exp(iK \int_{\gamma_\perp} A)$.

impose that these lines are genuine line operators, i.e. independent of the choice of surface Σ. This happens when $q(W) = mp \bmod N$ [29, 5]

$$W(\gamma) \exp\left(im \oint_\gamma A\right) \qquad \text{with} \qquad q(W) = mp \bmod N \ . \tag{2.63}$$

An operator W for which we cannot solve $q(W) = mp \bmod N$ cannot be "dressed" to a physical line operator. In addition, using (2.59), the equation of motion of B on the boundary leads to

$$\exp(i \oint A)| = \exp(i \oint y) \ . \tag{2.64}$$

Now, the canonical duality (2.44) maps $\exp(i \oint y)$ to the symmetry generating line $a \in \mathcal{T}$, so $\exp(i \oint A)| = a$. Therefore all the line operators on the boundary are the $\mathbb{Z}_N$-invariant lines in $\mathcal{T}$. This means that we have performed only step 1 of the three-step gauge procedure.

Using this identification we also recognize the L symmetry lines associated to the $\mathbb{Z}_L$ subgroup generated by $\hat{a} = a^K$ as the bulk lines generated by $V = \exp(iK \oint A)$. As we said above, these lines have trivial braiding with all the lines on the boundary.

One of the main points in our discussion is that since the bulk lines are trivial in any 3d correlation functions, we find it natural to identify them with the trivial line and accordingly, identify the boundary lines $W \sim W\hat{a}$. This works when pN/L^2 is even, so that the bulk line $V| = \hat{a} = a^K$ has integer spin $h[\hat{a}] = pN/2L^2$. When pN/L^2 is odd, the bulk line V is charged under the $\mathbb{Z}_2$ fermion parity (see Appendix E), and on the boundary it is identified with $\hat{a}$ of half-integer spin. Thus, we identify $W \sim W\hat{a}\psi$. The procedure above is equivalent to quotienting by the boundary lines that can move to the bulk. This is essentially the step 2 of the gauging procedure, except that we perform it with respect to $\mathbb{Z}_L$ rather than with respect to $\mathbb{Z}_N$. As with the step 3 in the gauging procedure, the identification leads to new lines. Consider a boundary line W at the fixed point of the fusion with $\hat{a}$. It can form junctions by emanating bulk lines at some points as shown in Figure 2.3. When the bulk lines are viewed as trivial, these junctions become new boundary line operators.

We have just performed step 1 of the gauging with respect to $\mathbb{Z}_N$ and steps 2 and 3 with respect to its $\mathbb{Z}_L$ subgroup. The result is exactly $\mathcal{T}'$ defined in (2.52)

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N} . \tag{2.65}$$

We note that the identification by the bulk lines, whose correlation functions on the boundary are trivial, is similar to the procedure of the more mathematical analysis in [38–42].

In this system, the minimal theory $\mathcal{A}^{N/L,-p/L}$ can be interpreted as the 3d TQFT that the bulk theory provides to cancel the anomaly.

After gauging the $\mathbb{Z}_N$ one-form symmetry, there is an emergent dual $\mathbb{Z}_N$ one-form symmetry in the bulk and an emergent dual $\mathbb{Z}_N$ zero-form symmetry on the boundary. They are both generated by $\exp(i \oint B)$. The original system can be recovered by gauging these emergent symmetries.

In summary, starting with a general 3d TQFT $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry of anomaly p, by coupling it to the bulk (2.54) and then gauging the one-form symmetry, we find the 3d TQFT $\mathcal{T}'$ as the effective boundary theory. We emphasize again that $\mathcal{T}'$ is only an effective theory, since the boundary can only be thought of as part of the 4d-3d system when the bulk

gauging	bulk	boundary theory	effective boundary theory
none	SPT of $\mathbb{Z}_N$	$\mathcal{T}$	
$\mathbb{Z}_N$ with $L=1$	trivial	$\mathcal{T}' = \frac{\mathcal{T}\otimes\mathcal{A}^{N,-p}}{\mathbb{Z}_N}$	
$\mathbb{Z}_N$ with $L\neq 1$	$\mathbb{Z}_L$ gauge theory	not meaningful	$\mathcal{T}' = \frac{\mathcal{T}\otimes\mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N}$

Table 2.3: Gauging an SPT phase of $\mathbb{Z}_N$ one-form symmetry with a boundary supporting a 3d TQFT $\mathcal{T}$ leads to a 4d-3d system. It is not meaningful to discuss the resulting boundary theory unless the bulk is trivial. This happens when $L = \gcd(N,p) = 1$. However, we can extract an effective boundary theory that captures many of the features for any L.

theory is nontrivial. However, in the special cases when $L = 1$, the bulk theory is trivial and $\mathcal{T}'$ is the theory on the boundary.

2.3.3 Interfaces between two different bulk TQFTs

A generalization[24] is to consider interfaces between two different SPT phases of $\mathbb{Z}_N$ one-form symmetry one with coefficient p^+ and the other with p^-

$$S_{4d} = \int_{\mathcal{M}_4^-}\left(\frac{p^-N}{4\pi}B_\mathcal{C}^-B_\mathcal{C}^- + \frac{N}{2\pi}B_\mathcal{C}^-dA^-\right) + \int_{\mathcal{M}_4^+}\left(\frac{p^+N}{4\pi}B_\mathcal{C}^+B_\mathcal{C}^+ + \frac{N}{2\pi}B_\mathcal{C}^+dA^+\right) . \tag{2.66}$$

On the interface $\partial\mathcal{M}_4^+ = \partial\mathcal{M}_4^-$, we choose the boundary condition $B_\mathcal{C} = B_\mathcal{C}^+| = B_\mathcal{C}^-|$ where $|$ represents the restriction to the interface. The anomaly inflow can be cancelled by an interface theory with a $\mathbb{Z}_N$ one-form symmetry of anomaly $p = p_+ - p_-$ that couples to $B_\mathcal{C}$. As in the case of a boundary, which we discussed above, we place on the interface a 3d TQFT $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry generated by a and with anomaly p. Following the discussion of the boundary, we use the canonical duality (2.44) and couple the interface theory to $B_\mathcal{C}$ through the

[24]As above, for simplicity, we will limit ourselves to spin 4-manifolds.

$(\mathcal{Z}_N)_{-pN}$ factor

$$S_{3d} = \int_{\partial \mathcal{M}_4} \left(-\frac{pN}{4\pi} x dx + \frac{N}{2\pi} x dy + \frac{N}{2\pi} B_{\mathcal{C}} y - \frac{N}{2\pi} B_{\mathcal{C}} (A^+ - A^-) \right) . \tag{2.67}$$

The one-form gauge symmetry of the system is

$$B_{\mathcal{C}}^{\pm} \to B_{\mathcal{C}}^{\pm} - d\lambda, \quad A^{\pm} \to A^{\pm} + p^{\pm}\lambda, \quad x \to x + \lambda, \quad y \to y + p\lambda . \tag{2.68}$$

We can gauge the $\mathbb{Z}_N$ one-form symmetry in the full system, i.e. make $B_{\mathcal{C}}$ dynamical (and remove the subscript $\mathcal{C}$). For later convenience, we define

$$L^{\pm} = \gcd(N, p^{\pm}), \quad L = \gcd(L^+, L^-), \quad K^{\pm} = N/L^{\pm}, \quad K = N/L = \mathrm{lcm}(K^+, K^-) . \tag{2.69}$$

After gauging, the bulk theory becomes effectively a $\mathbb{Z}_{L^{\pm}}$ one-form gauge theory on each side. In the special cases when $L^{\pm} = 1$, the bulk theories on both sides are trivial and there is a meaningful 3d theory on the interface. Otherwise, the interface can only be thought of as coupled to the 4d TQFT.

All the line operators $\tilde{W}$ on the interface can be constructed by fusing the lines W from the original 3d TQFT $\mathcal{T}$ and the lines $\hat{V}_{\pm} = \exp(i \oint A^{\pm})|$

$$\tilde{W} = W \hat{V}_+^{m^+} \hat{V}_-^{m^-} , \qquad q(W) = (p^+ m^+ + p^- m^-) \bmod N . \tag{2.70}$$

The various factors in $\tilde{W}$ are not $\mathbb{Z}_N$ gauge invariant line operators – each of them needs to be attached to a surface with B to make them invariant. But the condition on $m^{\pm}$ means that their product $\tilde{W}$ is $\mathbb{Z}_N$ invariant and hence it is a genuine line operator. (We ignore here a possible trivial open surface $\exp(iN \int B)$ and use $m^{\pm} \sim m^{\pm} + N$.) An operator W for which we cannot solve this equation cannot be "dressed" to a physical operator.

bulk at $\mathcal{M}_4^-$	Interface	bulk at $\mathcal{M}_4^+$
one-form: $\mathbb{Z}_N \to \mathbb{Z}_{K^-}$ $U_- = \exp(i\oint B^-)$	one-form: $\mathbb{Z}_N \to 1$ $U = \exp(i\int B^+ - i\int B^-)$	one-form: $\mathbb{Z}_N \to \mathbb{Z}_{K^+}$ $U_+ = \exp(i\oint B^+)$

Table 2.4: The emergent global symmetry in a 4d system with an interface. The first row summarizes the symmetries and their spontaneous breaking. The second row presents the charge generators. In U the integral is over a closed surface that pierces the interface. $U_\pm^{L^\pm} = 1$ means that this symmetry is broken to $\mathbb{Z}_{K^\pm}$. Below we will study an effective theory on the interface by performing a quotient of the full 4d-3d system by the bulk modes. We will see that the one-form global symmetry of this effective theory is $\mathbb{Z}_{\gcd(K^+,K^-)} = \mathbb{Z}_{N/\mathrm{lcm}(L^+,L^-)}$.

Using the equation of motion of B

$$\hat{V}_+ = a\hat{V}_- \tag{2.71}$$

and that a is a special case of W, all the lines on the interface can be written as[25]

$$\tilde{W} = W\hat{V}_-^{m^-}\ , \qquad q(W) = p^- m^- \bmod N\ . \tag{2.72}$$

Let us discuss the global symmetry of the system and its breaking (Table 2.4). After gauging, the bulk theories have an emergent $\mathbb{Z}_N$ one-form symmetry. It is spontaneously broken to $\mathbb{Z}_{K^\pm}$ on the two sides. The broken $\mathbb{Z}_{L^\pm} = \mathbb{Z}_N/\mathbb{Z}_{K^\pm}$ one-form symmetry is generated by the surface operator $U_\pm = \exp(i\oint B^\pm)$ with $U_\pm^{L^\pm} = 1$. It acts on the $\mathbb{Z}_{L^\pm}$ gauge theories in the two sides.

The interface has an emergent symmetry generated by the surface operator that pierces the interface

$$U = \exp(i\int_{\Sigma^+} B^+ - i\int_{\Sigma^-} B^-) \qquad , \qquad \partial\Sigma^+ = \partial\Sigma^- \tag{2.73}$$

where $\Sigma^\pm$ are two hemispheres in the two sides of the interface. Together they form a closed surface. U acts on the interface lines (2.70) $W\hat{V}_+^{m^+}\hat{V}_-^{m^-}$ by a phase of $e^{-2\pi i(m^+ + m^-)/N}$. (As a check, this phase is invariant under the fusion with the trivial operator $a\hat{V}_+^{-1}\hat{V}_- = 1$.)

The original $\mathbb{Z}_N$ one-form symmetry acted faithfully on $\mathcal{T}$. This means that there are lines W with all possible $\mathbb{Z}_N$ charges. Therefore, for every value of $m^\pm$ we can find a line W satisfying

[25]Note that unlike the case of a boundary discussed in Section 2.3.2, where $a = \hat{V}$ was a line in the original theory $\mathcal{T}$, here the interface lines with $m^- \neq 0$ were not present in $\mathcal{T}$. Correspondingly, there are new interface lines that arise from the bulk degrees of freedom.

(2.70). After gauging this $\mathbb{Z}_N$ symmetry, the emergent $\mathbb{Z}_N$ symmetry acts with charge $-(m^+ + m^-)$. We see that it acts faithfully in the resulting TQFT. This means that this emergent $\mathbb{Z}_N$ one-form symmetry is completely broken on the interface.

There is also an emergent dual $\mathbb{Z}_N$ zero-form symmetry on the interface generated by $\exp(i \oint B|)$. All these emergent symmetries have the same origin and gauging them with appropriate counterterms recovers the original system.

We conclude that the 4d-3d system has an emergent $\mathbb{Z}_N$ one-form symmetry, which acts faithfully on the interface; i.e. it is spontaneously broken.

Effective 3d theory

Next, we imitate what we did with a boundary and construct an effective interface theory by moding out by the bulk lines

$$V_\pm = (\hat{V}_\pm)^{K^\pm} = \exp(iK^\pm \oint A^\pm)| \; . \tag{2.74}$$

Step 2 The bulk lines are trivial in all correlation functions in 3d. We identify them with the trivial lines and therefore, the interface lines $\tilde{W}$ are identified as

$$\tilde{W} \sim \tilde{W} V_- \psi^{K^- p^-/L^-} \sim \tilde{W} V_+ \psi^{K^+ p^+/L^+} = \tilde{W} a^{K^+} (\hat{V}_-)^{K^-} \psi^{K^+ p^+/L^+} \,, \tag{2.75}$$

where we used the result that $V_\pm$ has interger spin for even $K^\pm p^\pm/L^\pm$ and half integer spin for odd $K^\pm p^\pm/L^\pm$ (see Appendix E).

Step 3 A line at the fixed point of the identification using a^K with $K = \mathrm{lcm}(K^+, K^-)$ can form junctions by emanating two bulk lines V_+^{K/K^+} and V_-^{K/K^-} at the same point. These junctions become genuine line operators if the bulk lines are taken to be trivial.

As an example we consider $\mathcal{T} = (\mathcal{Z}_N)_{-pN}$ defined in (2.31). After gauging all the lines on the interface are generated by b_+ and b_-

$$b_\pm = \exp(i \oint A^\pm - ip^\pm \oint x)\,. \tag{2.76}$$

We are interested in the expectation value of a knot on the interface

$$K[\{C_i\},\{C_i'\}] = \exp\big(i\sum_i \oint_{C_i}(A^+ - p^+x) + i\sum_i \oint_{C_i'}(A^- - p^-x)\big)\,. \tag{2.77}$$

Since the path integral is quadratic it can be evaluated easily (see Appendix E for similar calculations)

$$\langle K[\{C_i\},\{C_i'\}]\rangle = \exp\left(\frac{2\pi i p^+}{N}\sum_{i<j}\ell(C_i,C_j)\right)\exp\left(-\frac{2\pi i p^-}{N}\sum_{i<j}\ell(C_i',C_j')\right)\,. \tag{2.78}$$

where $\ell(C_i, C_j)$ is the linking number between C_i and C_j. Here the result arises from contractions of $\langle A^+A^+\rangle$ and $\langle A^-A^-\rangle$. Since $(b_\pm)^{K^\pm}$ is identified with the bulk line $V_\pm = \exp(iK^\pm \oint A^\pm)$, the effective interface theory is $\mathcal{A}^{K^+,-p^+/L^+} \otimes \mathcal{A}^{K^-,p^-/L^-}$.

Using the canonical duality (2.44), the effective interface theory for a general 3d TQFT $\mathcal{T}$ is

$$\frac{\mathcal{T}\otimes\mathcal{A}^{K^+,-p^+/L^+}\otimes\mathcal{A}^{K^-,p^-/L^-}}{\mathbb{Z}_N} = \frac{\mathcal{T}/\mathbb{Z}_L\otimes\mathcal{A}^{K^+,-p^+/L^+}\otimes\mathcal{A}^{K^-,p^-/L^-}}{\mathbb{Z}_K}\,, \tag{2.79}$$

where the quotient in the first presentation means gauging the diagonal anomaly free $\mathbb{Z}_N$ one-form symmetry generated by $ab^-(b^+)^{-1} = a\exp(i\oint(px - A^+ + A^-))$.

The two minimal Abelian TQFTs $\mathcal{A}^{K^+,-p^+/L^+}$ and $\mathcal{A}^{K^-,p^-/L^-}$ can be interpreted as the 3d TQFTs that the bulk theory provides to cancel the anomaly. The sign difference in the labels comes from the different orientations of the bulk relative to the interface.

It should also be added that when we performed the quotient of the full 4d-3d system by the two bulk theories to find an effective 3d theory, we modded out by the bulk operators. This means that the effective theory captures the correlation functions of interface lines, but does not capture the correlation functions of the bulk lines and the bulk surfaces.

Let us determine the one-form global symmetry of the effective theory. Since we have modded out by some bulk lines, it is different than the $\mathbb{Z}_N$ that acts on all possible lines in the interface.

Clearly, we should focus on the surface operator U that pierces the interface (2.73). In general, it has nontrivial correlation functions with the lines in the bulk. Hence, its intersection with the interface $\partial\Sigma^+ = \partial\Sigma^-$ does not represent a genuine line operator on the interface. Since it is not included as a line operator in our effective theory, the effective theory does not have the full $\mathbb{Z}_N$ symmetry.

However, the surface operator

$$U^{\tilde{L}}\,, \qquad \text{with } \tilde{L} = \text{lcm}(L^+, L^-) \tag{2.80}$$

has trivial correlation functions with all the bulk lines and therefore we expect that it corresponds to a line operator on the interface. Indeed, it is

$$U^{\tilde{L}} = \left(\hat{V}_+^{r_+}\right)^{-\tilde{L}/L^+} \left(\hat{V}_-^{r_-}\right)^{\tilde{L}/L^-}\,, \qquad r_\pm p^\pm = L^\pm \text{ mod } N\,. \tag{2.81}$$

This line generates a $\mathbb{Z}_{N/\tilde{L}}$ subgroup of the emergent $\mathbb{Z}_N$ one-form symmetry of the full 4d-3d system.

The one-form global symmetry of the effective theory can also be obtained from (2.79). First, using the $\mathbb{Z}_K$ quotient we can express the symmetry lines as the lines in the minimal Abelian theories. Since $r_\pm p^\pm/L^\pm = 1$ mod $K^\pm$, we can choose the generating line of the minimal theories $\mathcal{A}^{K^\pm, \mp p^\pm/L^\pm}$ to be $(\hat{V}_\pm)^{r_\pm}$. Then the lines in the effective interface theory (2.79) originating from the minimal theories are

$$(\hat{V}_+^{r_+})^{m^+}(\hat{V}_-^{r_-})^{m^-}\,, \qquad m^+L^+ + m^-L^- = 0 \text{ mod } N\,, \tag{2.82}$$

with $m^\pm \sim m^\pm + K^\pm$. The condition only has solutions $(m^+, m^-) = n(\tilde{L}/L^+, -\tilde{L}/L^-)$ with integer n and hence the line (2.81) generates all the interface lines originating from the minimal theories. This means that the $\mathbb{Z}_{N/\tilde{L}}$ one-form symmetry is the largest symmetry of the effective interface theory (2.79) generated by the lines from the minimal theories.

Another way to understand this global $\mathbb{Z}_{N/\tilde{L}}$ one-form symmetry of the effective theory is the following. The full 4d-3d system realizes a spontaneously broken $\mathbb{Z}_N$ symmetry, which acts faithfully. In the bulk this symmetry is spontaneously broken to $\mathbb{Z}_{K^\pm}$, so the bulk modes realize $\mathbb{Z}_{L^\pm}$. Together, the two bulk half-spaces realize $\mathbb{Z}_{\tilde{L}} = \mathbb{Z}_{L^+} \cup \mathbb{Z}_{L^-}$. Therefore, the effective interface theory, obtained as the quotient by the bulk modes realizes $\mathbb{Z}_{N/\tilde{L}}$. Equivalently, the unbroken global one-form symmetries in the two bulks are $\mathbb{Z}_{K^\pm}$ and hence $\mathbb{Z}_{\gcd(K^+,K^-)} = \mathbb{Z}_{K^+} \cap \mathbb{Z}_{K^-}$ is unbroken throughout the two bulks. We know that the full $\mathbb{Z}_N$ symmetry is broken in the interface. Therefore, the quotient theory should realize the symmetry $\mathbb{Z}_{\gcd(K^+,K^-)} = \mathbb{Z}_{N/\tilde{L}}$.

2.4 $SU(N)$ and $PSU(N)$ gauge theory in 4d

2.4.1 $SU(N)$ gauge theory, walls and interfaces

We begin by reviewing the dynamics of 4d pure $SU(N)$ gauge theory and its domain walls and interfaces following [64, 5, 27]. The action of the theory is

$$S = -\frac{1}{4g^2} \int \mathrm{Tr}(F \wedge *F) + \frac{\theta}{8\pi^2} \int \mathrm{Tr}(F \wedge F) \,, \tag{2.83}$$

where the parameter θ is identified periodically $\theta \sim \theta + 2\pi$.

This system has a $\mathbb{Z}_N$ one-form global symmetry, which we will refer to as electric. It is generated by a surface operator

$$\mathbb{U}_{\mathbb{E}} = \exp(i \oint C) \ , \tag{2.84}$$

where C depends on the dynamical gauge fields. As expected of a charge operator, the correlation functions of the surface operator $\mathbb{U}_{\mathbb{E}}$ are topological [5]. The charged objects are Wilson lines in representations of $SU(N)$ and the $\mathbb{Z}_N$ charge is determined by the action of the center of the gauge group on the representation. We will denote the Wilson line in the fundamental representation by $\mathcal{W}$.

In addition to the Wilson lines and the charges $\mathbb{U}_{\mathbb{E}}^{r} = \exp(ir \oint C)$, the system also includes open versions of the charges

$$T(\gamma)\exp(i\int_{\Sigma} C)\,, \qquad \gamma = \partial\Sigma\ , \tag{2.85}$$

where T is the 't Hooft operator. In the $SU(N)$ theory it is not a genuine line operator and needs to be attached to an open surface operator. The 't Hooft operator is the worldline of a monopole, which is defined by being surrounded by a two-sphere with a nontrivial $PSU(N)$ bundle on it. The $SU(N)$ theory does not have such objects. They have to be attached to strings. (This is like the Dirac string of a magnetic monopole, except that it is detectable by Wilson lines, and hence it is physical.) The surface in (2.85) can be interpreted as the worldsheet of this string. This allows us to interpret the $\mathbb{Z}_N$ charge operator $\mathbb{U}_{\mathbb{E}}^{r} = \exp(ir \oint C)$ as a closed worldsheet of such strings.

It is natural to couple the global $\mathbb{Z}_N$ symmetry to background gauge fields $\mathcal{B}_{\mathcal{C}}$. Then, since the Wilson lines are charged under the symmetry, they take the form

$$\mathcal{W}(\gamma)e^{\frac{2\pi i}{N}\int_{\Sigma}\mathcal{B}_{\mathcal{C}}} \qquad , \qquad \gamma = \partial\Sigma\ . \tag{2.86}$$

One way to think about the classical background $\mathcal{B}_{\mathcal{C}}$ is that instead of summing over $SU(N)$ bundles in the path integral, we sum over $PSU(N)$ bundles $\mathcal{E}$ with fixed second Stiefel-Whitney classes $w_2(\mathcal{E}) = \mathcal{B}_{\mathcal{C}} \in H(\mathcal{M}_4, \mathbb{Z}_N)$.

Another consequence of the background field is that we can add to the action the counterterm

$$2\pi\frac{p}{2N}\int_{\mathcal{M}_4}\mathcal{P}(\mathcal{B}_{\mathcal{C}})\ . \tag{2.87}$$

In the presence of this term the θ periodicity is as in (2.1)

$$(\theta, p) \sim (\theta + 2\pi, p + N - 1) \ . \tag{2.88}$$

This lack of 2π periodicity in θ has another consequence. Because of the Witten effect [65] the open surface operators (2.85) are not invariant under $\theta \to \theta + 2\pi$. They transform as

$$T(\gamma) \exp(i \int_\Sigma C) \to \mathcal{W}(\gamma) T(\gamma) \exp(i \int_\Sigma C) \ . \tag{2.89}$$

This fact will be important below.

So far we have discussed the kinematics of the $SU(N)$ theory. Now we turn to the dynamics. At low energies the $SU(N)$ theory has a gap and it confines. This means that the $\mathbb{Z}_N$ one-form symmetry is unbroken and the charged Wilson lines (those in representations that transform nontrivially under the $\mathbb{Z}_N$ center) have an area law. Correspondingly, these Wilson lines vanish at long distances. As a result, the low-energy theory is trivial. It does not even have a TQFT. In the low-energy theory the Wilson lines $\mathcal{W}^r$ vanish and the charges $\mathbb{U}^r_{\mathbb{E}}$ are equal to one.

The dynamical objects of the system have electric and magnetic charges that are N times the basic units of the Wilson line $\mathcal{W}$ and the 't Hooft operator (with its attached surface (2.85)). Confinement means that some dynamical monopoles or dyons condense. But these are different dyons at θ and at $\theta + 2\pi$. Because of the Witten effect, their electric charges differ by N units. This means that if we have confinement at θ, we have oblique confinement at $\theta + 2\pi$. And more generally, we have different kinds of oblique confinement at these two values of θ.

At $\theta \in \pi\mathbb{Z}$, the $SU(N)$ gauge theory has a time-reversal symmetry. It is unbroken at $\theta \in 2\pi\mathbb{Z}$. At $\theta \in 2\pi\mathbb{Z} + \pi$, the theory is argued to have two degenerate vacua associated with the spontaneous symmetry breaking of the time reversal symmetry. Since the action of time reversal at these points involve a shift of θ by a multiple of 2π, the two vacua have different kinds of oblique confinement.

Let us discuss the domain walls between these two vacua. Since they have different kinds of oblique confinement in the two sides, one dyon condenses in one side and another dyon condenses in the other side. Therefore, no dyon condenses on the wall and correspondingly, the theory is not confining there. This means that the electric $\mathbb{Z}_N$ one-form symmetry is spontaneously broken on the domain wall and the fundamental Wilson loops are physical observables in the low-energy theory.

It was argued in [27] that the wall supports a nontrivial TQFT, $SU(N)_1$. This TQFT has a $\mathbb{Z}_N$ one-form symmetry with an anomaly $p = N-1$, which accounts to the different anomaly inflow from the two sides of the wall. Note that this is the minimal TQFT with these properties $\mathcal{A}^{N,N-1}$ and any other TQFT with such properties includes $SU(N)_1$ as a decoupled sector and the rest of the theory is $\mathbb{Z}_N$ invariant.

The $SU(N)$ gauge theory can also have interfaces that interpolate between θ_0 and $\theta_0 + 2\pi k$ for some integer k. The anomaly inflow requires the interfaces to support theories with a $\mathbb{Z}_N$ one-form symmetry of anomaly $p = k(N-1) \bmod 2N$. This does not uniquely specify the theories on the interfaces. However, when θ varies smoothly, the interface theory is uniquely determined by the microscopic theory and the profile of the θ-parameter. This is to be contrasted with sharp interfaces when θ is discontinuous. When θ varies smoothly and slowly with $|\nabla\theta| \ll \Lambda$, where Λ is the dynamical scale of the theory, there are k domain walls where θ crosses an odd multiple of π. Each domain wall supports an $SU(N)_1$ TQFT. When θ varies smoothly and more rapidly with $|\nabla\theta| \gg \Lambda$, the interface theory $SU(N)_1^{\otimes k}$ is argued to undergo a transition to $SU(N)_k$ theory [27, 28]. This can be understood as the Chern-Simons term induced by by the θ-term in the bulk.

However, it is possible that the strong dynamics changes the interface theory at low-energy. One logical possibility is that the dynamics Higgses $SU(N)$ using scalar fields in the adjoint representation. This preserves the $\mathbb{Z}_N$ one-form symmetry and the anomaly. The maximum possible Higgsing with one adjoint scalar is to the Cartan torus $U(1)^{N-1}$, where the $U(1)^{N-1}$

gauge fields a^I, $I = 1, \cdots, N-1$ are embedded in the $SU(N)$ gauge field a through

$$a = a^I H^I , \quad (H^I)_{ij} = \mathrm{diag}(\underbrace{0, \cdots, 0}_{I-1}, 1, -1, \underbrace{0, \cdots, 0}_{N-I-1}) . \tag{2.90}$$

In terms of these fields the $SU(N)_k$ theory becomes a $U(1)^{N-1}$ Chern-Simons theory

$$\frac{k}{4\pi}\mathrm{Tr}\left(ada - \frac{2i}{3}a^3\right) \;\to\; \frac{k}{4\pi}(K_{\mathrm{Cartan}})_{IJ}a^I da^J \tag{2.91}$$

where K_{Cartan} is the Cartan matrix of $SU(N)$

$$(K_{\mathrm{Cartan}})_{IJ} = \mathrm{Tr}(H^I H^J) = 2\delta_{I,J} - \delta_{I,J+1} - \delta_{I+1,J} . \tag{2.92}$$

For $k = 1$ this Abelian TQFT is the same as $SU(N)_1$, so this possibility is the same as the previous suggestion.

We can further Higgs $SU(N)$ all the way down to its $\mathbb{Z}_N$ center. In order to identify the TQFT of this $\mathbb{Z}_N$ gauge theory, we use a presentation of $SU(N)_k$ based on $U(N) \times U(1)$ gauge fields b and y [31]

$$\frac{k}{4\pi}\mathrm{Tr}\left(bdb - \frac{2i}{3}b^3\right) - \frac{k}{4\pi}(\mathrm{Tr}\, b)d(\mathrm{Tr}\, b) + \frac{1}{2\pi}yd(\mathrm{Tr}\, b) , \tag{2.93}$$

where the $U(1)$ field y constrains b to be a $SU(N)$ gauge field. The $\mathbb{Z}_N$ gauge field x is embedded in $U(N)$ through $b = x\mathbb{I}$. After Higgsing, the $SU(N)_k$ theory becomes a $\mathbb{Z}_N$ gauge theory $(\mathcal{Z}_N)_{-kN(N-1)} = (\mathcal{Z}_N)_{-pN}$. Alternatively, the precise $\mathbb{Z}_N$ gauge theory can be determined by matching the anomalies.

In conclusion, without a more detailed dynamical analysis we cannot uniquely determine the TQFT on the interface, so we will denote it by $\mathcal{T}_k$. The simplest case $\mathcal{T}_1$ was argued to be the minimal allowed theory, $SU(N)_1$. But for higher values of k there isn't a preferred choice and we presented several options, e.g. $SU(N)_k$ and $(\mathcal{Z}_N)_{-kN(N-1)} = (\mathcal{Z}_N)_{-pN}$. However, using the analysis in the previous sections, we can proceed without knowing exactly what $\mathcal{T}_k$ is.

Let us analyze the interface theory $\mathcal{T}_k$ in more detail. The theory has a $\mathbb{Z}_N$ one-form symmetry of anomaly $k(N-1)$, which means that the symmetry lines are anyons with a braiding phase of $e^{-2\pi i k(N-1)/N}$. These symmetry lines can be thought of as bulk charge operators generated by $\mathbb{U}_{\mathbb{E}}$ that pierce the interface. To see that, recall that because of confinement, the shape of $\mathbb{U}_{\mathbb{E}}$ in the bulk is not important (a closed surface on each side equals to one) and therefore, $\mathbb{U}_{\mathbb{E}}$, which pierces the interface is effectively a line operator on the interface. Also, $\mathbb{U}_{\mathbb{E}}$ can be interpreted as the worldsheet of a string constructed by gluing two 't Hooft lines from the two sides at the interface. So we can view $\mathbb{U}_{\mathbb{E}}$ as associated with two 't Hooft lines, T on one side of the interface and T^{-1} on the other side. Then, because of the Witten effect [65], the electric charges of these two 't Hooft lines differ by k and therefore $\mathbb{U}_{\mathbb{E}}$ that pierces the interface appears as a Wilson line with electric charge k. More precisely, it is the generator of the $\mathbb{Z}_N$ one-form global symmetry on the interface. For example, if the theory on the interface $\mathcal{T}_k$ is $SU(N)_k$, it is a Wilson line in a k index symmetric representation of $SU(N)$.

The fact that $\mathbb{U}_{\mathbb{E}}$ leads to a Wilson line on the interface shows that not only are the probe quarks on the interface liberated (because there is no confinement there), they are also anyons!

2.4.2 $PSU(N)$ gauge theory

The $PSU(N)$ gauge theory differs from the $SU(N)$ gauge theory in the global form of the gauge group. It can be constructed by gauging the electric $\mathbb{Z}_N$ one-form symmetry in the $SU(N)$ gauge theory, i.e. by making the classical background field $\mathcal{B}_{\mathcal{C}}$ dynamical (and dropping the subscript $\mathcal{C}$). Summing over $\mathcal{B}$ means that we sum over all $PSU(N)$ bundles $\mathcal{E}$. Now, the choice of the counterterm (2.87) is more significant than in the $SU(N)$ theory and the value of p affects the set of observables.

Let us discuss the operators in the theory. Since now $\mathcal{B}$ is dynamical, the Wilson loop (2.86) is no longer a genuine line operator; it depends on the surface Σ. We can consider a closed surface operator

$$\mathbb{U}_{\mathbb{M}} = \exp\left(\frac{2\pi i}{N}\oint w_2^{PSU(N)}\right) = \exp\left(\frac{2\pi i}{N}\oint \mathcal{B}\right) \ , \tag{2.94}$$

where $w_2^{PSU(N)}$ is the abbreviation for $w_2(\mathcal{E})$ (with $\mathcal{E}$ the $PSU(N)$ bundle). It is the generator of a new emergent $\mathbb{Z}_N$ one-form symmetry, which we will refer to as magnetic.

The original Wilson line is an open version of $\mathbb{U}_{\mathbb{M}}$. And just as the surface in this Wilson line can be interpreted as the worldsheet of an electric (confining) string, the closed surface operator $\mathbb{U}_{\mathbb{M}}$ can be interpreted as a closed worldsheet of such a string. (Note that in the $PSU(N)$ theory this string worldsheet is an operator in the theory.)

For $p=0$ the 't Hooft line T is a genuine line operator and we do not need to write C of (2.85). It is charged under the magnetic symmetry (2.94). Other dyonic operators of the form $T\mathcal{W}^r$ need an attached surface and they are not genuine line operators (unless $r = 0 \mod N$).

We would like to find the line operators when p is nonzero. We simplify the discussion by considering the theory on a spin manifold such that the periodicity of p is $p \sim p+N$.[26] We first keep $p=0$ and extend the range of $\theta \sim \theta + 2\pi N$. Clearly, T remains a genuine line operator as we change θ. But because of the Witten effect it acquires electric charge $-k$ as θ is shifted by $-2\pi k$. Then we restore the original θ and have nonzero $p = k(N-1)$. This means that the basic line operator has electric charge p, i.e. it is [34]

$$\hat{T}(\gamma) = T(\gamma)\mathcal{W}(\gamma)^p \ . \tag{2.95}$$

Note that this is a genuine line operator, which does not need a surface.

Another way to understand the lines (2.95) is to write them as $T\mathcal{W}^p \exp\left(i\int_\Sigma (C + \frac{2\pi p}{N}\mathcal{B})\right)$, where C comes from T (2.85) and $\frac{2\pi p}{N}\mathcal{B}$ from $\mathcal{W}$ (2.86). In the $PSU(N)$ theory with p the term in the exponent vanishes and hence this operator is independent of Σ.

[26]On an orientable non-spin manifold, the change $p \to p+N$ (with even N) produces the coupling $\pi\int w_2(\mathcal{M}_4)\cup\mathcal{B}$ (where $w_2(\mathcal{M}_4)$ is the second Stiefel-Whitney class of the 4d manifold $\mathcal{M}_4$) that is equivalent to turning on classical background field $\tilde{\mathcal{B}}_{\mathcal{C}} = Nw_2(\mathcal{M}_4)/2$ for the magnetic $\mathbb{Z}_N$ one-form symmetry generated by $\exp(\frac{2\pi i}{N}\oint\mathcal{B})$. Thus it changes the statistics of the basic 't Hooft line from a boson to a fermion and vice versa [66, 10, 67]. This does not modify the $PSU(N)$ bundle but instead gives additional weights in the path integral.

Now, let us consider the dynamics. In the $SU(N)$ theory the dyons that condense at $\theta = 2\pi k$ have the quantum numbers of $T^N \mathcal{W}^{kN}$. (Note that these dyons exist as dynamical objects regardless of the global structure of the gauge group. The global part of the group and the value of p determine the line operators in the theory.)

Let us focus on $\theta = 0$ with arbitrary p. The genuine line operators in the theory are powers of $\hat{T}$ (2.95). Some of them have area law because of the condensation and hence they vanish at low energies. Only the lines that are generated by

$$\hat{T}^K = T^K \mathcal{W}^{pK} \qquad , \qquad L = \gcd(N, p) \qquad , \qquad K = \frac{N}{L} \ , \tag{2.96}$$

are aligned with the condensed dyons and hence they have a perimeter law. These are the only nontrivial line operators in the low-energy theory.

It is clear that the magnetic $\mathbb{Z}_N$ one-form symmetry is spontaneously broken to $\mathbb{Z}_K$ and the broken elements are realized at low-energy by a $\mathbb{Z}_L$ gauge theory [34]. The operators in this $\mathbb{Z}_L$ gauge theory are generated by the basic $\mathbb{Z}_L$ Wilson line (2.96) (which is not to be confused with the microscopic $PSU(N)$ Wilson line) and its dual surface operator, which is the microscopic operator $\mathbb{U}_{\mathbb{M}}$.[27]

In conclusion, the low-energy manifestation of this spontaneous symmetry breaking of the magnetic $\mathbb{Z}_N$ one-form symmetry is the theory (2.55). And the relation between the microscopic operators in the $PSU(N)$ gauge theory and the low-energy theory is summarized in Table 2.5.

2.4.3 Interfaces in $PSU(N)$ gauge theory

Here we study an interface in the $PSU(N)$ theory. We let it interpolate smoothly between $\theta = 0$ and $\theta = 2\pi k$. As above, we can approximate it at low energies with constant $\theta = 0$ and p changing from p^+ to p^-. This is the setup we considered in the $SU(N)$ theory above, and now we simply gauge the electric $\mathbb{Z}_N$ one-form symmetry in that theory.

[27]On a nonspin manifold this $\mathbb{Z}_L$ gauge theory could be twisted, as in Appendix E.

Microscopic $PSU(N)$ gauge theory	Low energy $\mathbb{Z}_N$ two-form gauge theory
$\hat{T}^r = (T\mathcal{W}^p)^r$	0 for $r \neq 0 \mod K$
$\hat{T}^K = T^K \mathcal{W}^{pK}$	$\exp\left(iK \oint A + ipK \int B\right) \sim \exp\left(iK \oint A\right)$
$\mathbb{U}_{\mathbb{M}} = \exp\left(\frac{2\pi i}{N} \oint w_2^{PSU(N)}\right)$	$\exp\left(i \oint B\right)$

Table 2.5: The dictionary between the operators in the microscopic $PSU(N)$ gauge theory and the operators in the macroscopic $\mathbb{Z}_N$ two-form gauge theory. The line operator in the second row is the minimal line that obeys a perimeter law. It is identified with the genuine line operator in the low-energy theory (and hence we suppress the B dependent term). Here we use a continuous notation for the low-energy TQFT, which is reviewed in appendix E.

We use the definitions (2.69)

$$L^\pm = \gcd(N, p^\pm), \quad L = \gcd(L^+, L^-), \quad K^\pm = N/L^\pm, \quad K = N/L \ . \tag{2.97}$$

The low-energy dynamics of the $PSU(N)$ theory in the two sides are approximated by the $\mathbb{Z}_N$ two-form gauge theories with parameters $p^\pm$, which are equivalent to $\mathbb{Z}_{L^\pm}$ gauge theories. They describe the spontaneous breaking of the magnetic $\mathbb{Z}_N$ one-form global symmetry to $\mathbb{Z}_{K^\pm}$. Note that unlike the $SU(N)$ theory, where the two sides of the interface differed only by a counterterm for background fields, here the two sides are dynamically different.

The TQFT in the bulk and on the interface is as in Section 2.3.3, so we will not repeat its analysis in detail, except to summarize the main points.

We have already said that in the bulk the magnetic $\mathbb{Z}_N$ one-form symmetry is spontaneously broken to $\mathbb{Z}_{K^\pm}$. On the interface, since the confined line operators in the bulk become liberated, the magnetic $\mathbb{Z}_N$ one-form symmetry, generated by the surface operators piercing the interface, is completely broken. Equivalently, we have argued above that in the $SU(N)$ theory no monopole condenses on the wall and the dynamics is weakly coupled there. Therefore, the $\mathbb{Z}_N$ one-form symmetry of the $PSU(N)$ theory should also be spontaneously broken there.

When θ varies smoothly and rapidly, the interface in the $SU(N)$ gauge theory supports a TQFT $\mathcal{T}_k$. The effective interface theory on the corresponding $PSU(N)$ interface is found easily using the results in Section 2.3.3. When $L^+ = L^- = 1$ the theory on the interface is

$$\frac{\mathcal{T}_k \otimes \mathcal{A}^{N,-p^+} \otimes \mathcal{A}^{N,p^-}}{\mathbb{Z}_N} \ . \tag{2.98}$$

As in Section 2.3.3, we can interpret the two minimal theories in the numerator as produced by the bulk in the two sides, such that we can gauge an anomaly free $\mathbb{Z}_N$ one-form global symmetry.

For generic $L^\pm$ the interface couples to the $\mathbb{Z}_{L^\pm}$ gauge theory in the bulk and it is meaningless to ask what the theory on the interface is. Yet, we can identify an effective interface theory. It is

$$\frac{\mathcal{T}_k/\mathbb{Z}_L \otimes \mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,p^-/L^-}}{\mathbb{Z}_K} = \frac{\mathcal{T}_k \otimes \mathcal{A}^{N/L^+,-p^+/L^+} \otimes \mathcal{A}^{N/L^-,p^-/L^-}}{\mathbb{Z}_N} \ . \tag{2.99}$$

As an example, we argued above that the interface in the $SU(N)$ theory between $\theta = 0$ and $\theta = 2\pi$ with $p^+ = p^- = 0$ supports an $SU(N)_1$ theory. This corresponds to $\theta = 0$ with $p^+ = 0$ and $p^- = 1 - N$, and thus $L^+ = N, L^- = 1$. The effective interface theory on the corresponding $PSU(N)$ interface is trivial, since $\mathcal{A}^{N,1-N} = SU(N)_{-1}$.

2.5 Appendix A: Definitions of Abelian anyons

In this Appendix we will review some properties of Abelian anyons. There are three equivalent definitions of Abelian anyons. An anyon a in a $3d$ TQFT is called Abelian when

(1) a obeys group-law fusion, $aa^s = a^{s+1}$ for integers s with $a^0 = 1$. In particular, since the number of lines in a consistent $3d$ TQFT is finite, there exists an integer m such that $a^m = 1$.

(2) a obeys Abelian fusion rules. For any line W in the $3d$ TQFT, the fusion product aW only contains one line.

(3) the quantum dimension of a is one.

First, the definition (1) implies (3). The group-law fusion $a^m = aa \cdots a = 1$ implies $d_a^m = 1$

for the quantum dimension d_a of a. Since d_a must be a positive real number in any unitary $3d$ TQFT, we conclude $d_a = 1$.

The definition (2) implies (1) by specializing $W = a, aa, \cdots$ and defining the unique line appears in the fusion of n line a to be a^n.

Now we will show the definition (3) implies (2) by contradiction. Suppose there exists a line x that fuses with a into at least two lines that we denote by y, z:

$$a \cdot x = y + z + \cdots \ . \tag{2.100}$$

This implies

$$\bar{a} \cdot y = x + \cdots \ , \tag{2.101}$$

where $\bar{a}$ denotes the antiparticle of a, *i.e.* $a \cdot \bar{a} = 1 + \cdots$. The quantum dimensions in the fusion $u \cdot v = \sum_i w_i$ satisfy $d_u d_v = \sum_i d_{w_i}$ [46], and thus

$$d_a d_x = d_y + d_z + \cdots \ , \quad d_{\bar{a}} d_y = d_x + \cdots \quad \Rightarrow \quad d_a d_{\bar{a}} d_x \geq d_x + d_{\bar{a}} d_z > d_x \ , \tag{2.102}$$

where the last two inequalities used the property that the quantum dimensions are real and positive, and in particular the last inequality comes from the existence of the second anyon z in the fusion (2.100). Since $\bar{a}$ and a have the same quantum dimension, by definition (3) $d_a = d_{\bar{a}} = 1$. Thus the last equation in (2.102) leads to a contradiction. Therefore, any line x must fuse with a into only one line. We conclude that (3) implies (2), and since we have already shown that (1) implies (3), this means that (1) implies (2). This completes the proof that the three definitions are equivalent to one another.

2.6 Appendix B: Jacobi symbols

For any odd prime number q, the Legendre symbol is defined as

$$\left(\begin{array}{c} a \\ q \end{array}\right) = a^{\frac{q-1}{2}} \mod q = \begin{cases} 0 & a = 0 \mod q \\ 1 & a = r^2 \mod q \text{ for some integer } r \\ -1 & \text{otherwise} \end{cases} . \tag{2.103}$$

For any odd integer b with a prime factorization $b = \prod_k q_k^{\alpha_k}$, the Jacobi symbol is the generalization of the Legendre symbol defined as

$$\left(\frac{a}{b}\right) = \prod_k \left(\begin{array}{c} a \\ q_k \end{array}\right)^{\alpha_k} . \tag{2.104}$$

The Jacobi symbol obeys the following identities for odd integers a, b, c

$$\left(\frac{ab}{c}\right) = \left(\frac{a}{c}\right)\left(\frac{b}{c}\right), \quad \left(\frac{-1}{c}\right) = (-1)^{(c-1)/2} . \tag{2.105}$$

2.7 Appendix C: 2d unitary chiral RCFT for Abelian 3d TQFT

In this Appendix we will show that every Abelian 3d TQFT corresponds to a 2d unitary chiral RCFT. Such unitary CFTs are generally not unique for a given TQFT and here we construct one example of them. The unitary RCFT is characterized by an extended chiral algebra of a product of chiral algebras of free compact bosons, free complex fermions, and $SU(N)_1$ Wess-Zumino-Witten models. If the TQFT is a spin theory, then the RCFT is $\mathbb{Z}_2$-graded [26].

Every Abelian TQFT $\mathcal{A}$ can be expressed as an Abelian Chern-Simons theory [28] [50–54] (for

[28] For example, the Chern-Simons theories with gauge group of rank n including $SU(n+1)_1$, $Spin(2n)_1$ and $(E_n)_1$ can be written as $U(1)^n$ Abelian Chern-Simons theories with the coefficient matrix given by the Cartan

a review see *e.g.* [55]). Denote the $U(1)$ gauge fields by $x_0, x_1, \cdots, x_n$ for some integer n, and the Chern-Simons action is

$$\frac{k}{4\pi}x_0 dx_0 + \sum_{i=1}^{n}\left(\frac{q_{0i}}{2\pi}x_0 dx_i\right) + \mathcal{L}[x_1, \cdots, x_n] \ , \tag{2.106}$$

where k, q_{0i} are integers, and $\mathcal{L}[x_1, \cdots, x_n]$ denotes Chern-Simons terms independent of the gauge field x_0. k, q_{0i} cannot be simultaneously zero for all i, since otherwise the theory has a decoupled gapless sector described by the dual photon of x_0. If $k = 0$, there exists $q_{0i} \neq 0$ for some i, and the redefining $x_i \to x_i + x_0$ produces nonzero k. Thus we can assume k is always nonzero without loss of generality. Consider the change of variables from $x_0, x_1, \cdots, x_n$ to $y_0, y_1, \cdots, y_n$

$$x_0 = y_0 - \sum_{i=1}^{n} q_{0i} y_i, \quad x_j = k y_j, \ j = 1, \cdots, n \ . \tag{2.107}$$

The Jacobian is $|k|^n$. The theory $\mathcal{A}$ can thus be expressed as

$$\mathcal{A} = \frac{\mathcal{A}'}{\mathbb{Z}_{|k|}^n} \ , \tag{2.108}$$

where the quotient denotes gauging a one-form symmetry $\mathbb{Z}_{|k|}^n$, and $\mathcal{A}'$ is an Abelian Chern-Simons theory with $U(1)$ gauge fields $y_0, y_1, \cdots, y_n$. Substituting (2.107) into (2.106), we find the theory $\mathcal{A}'$ has the Chern-Simons action

$$\frac{k}{4\pi}y_0 dy_0 + \widetilde{\mathcal{L}}[y_1, \cdots, y_n] \ , \tag{2.109}$$

where $\widetilde{\mathcal{L}}[y_1, \cdots, y_n]$ denotes Chern-Simons terms independent of y_0. Thus $\mathcal{A}' = U(1)_k \otimes \mathcal{A}''$ for another Abelian Chern-Simons theory $\mathcal{A}''$ with gauge fields $y_1, \cdots, y_n$. By iteration, we find the

matrix of the gauge groups.

Abelian TQFT $\mathcal{A}$ can be expressed as

$$\mathcal{A} = \hat{\mathcal{A}}/\mathcal{Z}, \quad \hat{\mathcal{A}} = \prod_{i=0}^{n} U(1)_{k_i} \ , \tag{2.110}$$

where the quotient denotes gauging a one-form symmetry $\mathcal{Z}$ that is a finite Abelian group, and k_i are non-zero integers.

If all k_i are positive, then the Abelian TQFT $\mathcal{A}$ corresponds to the extended chiral algebra of a product of compact bosons in 2d (the RCFT may be $\mathbb{Z}_2$ graded).

If some of $k_i = -m_i$ is negative, the corresponding $U(1)_{-m_i}$ in $\hat{\mathcal{A}}$ can be replaced by an $SU(N)$ Chern-Simons theory at level one using the duality [29]

$$U(1)_{-m_i} \quad \longleftrightarrow \quad \begin{cases} SU(4m_i)_1/\mathbb{Z}_2 & \text{even } m_i \\ SU(m_i)_1 \otimes \{1,\psi\} & \text{odd } m_i \end{cases} \ , \tag{2.111}$$

where for even m_i the theory $U(1)_{-m_i}$ is non-spin, and we omit a trivial TQFT such as $(E_8)_1$ in the duality.

For odd m_i the theory $U(1)_{-m_i}$ is a spin theory. On the right hand side of the duality (2.111) the theory $\{1,\psi\}$ represents the almost trivial TQFT that has only two lines (of integer and half integer spins), and it includes the gravitational Chern-Simons term $-2M_i\text{CS}_g$ for some positive integer $M_i = -m_i$ mod 8. The almost trivial TQFT corresponds to M_i free complex fermions in 2d.

Thus the theory $\hat{\mathcal{A}}$ corresponds to the 2d unitary chiral RCFT ($\mathbb{Z}_2$ graded if some k_i is odd) given by the product of free compact bosons, free complex fermions, and $SU(N_i)$ Wess-Zumino-Witten models at level one with N_i given in (2.111) (or its extended chiral algebra when $k_i = -m_i$ is even and negative). The Abelian TQFT $\mathcal{A}$ then corresponds to the 2d unitary chiral RCFT given by the extended chiral algebra (2.110) of the 2d unitary chiral RCFT of $\hat{\mathcal{A}}$.

[29] For the case m_i is odd, the duality (2.111) is the level-rank duality [31].

2.8 Appendix D: Gauging a general anomaly free subgroup

In order to simplify the discussion we will assume in this appendix that all the TQFTs are spin TQFTs.

A theory $\mathcal{T}$ with a $\mathbb{Z}_N$ one-form symmetry with anomaly p can have multiple anomaly free subgroups. One of them is the $\mathbb{Z}_L$ subgroup with $L = \gcd(N, p)$. In this Appendix, we will discuss gauging a larger anomaly free symmetry $\mathbb{Z}_m$, i.e.

$$\mathbb{Z}_L \subset \mathbb{Z}_m \subset \mathbb{Z}_N \ . \tag{2.112}$$

It is anomaly free when pN/m^2 is an integer (recall that we discuss spin theories). Gauging this symmetry leads to $\mathcal{T}/\mathbb{Z}_m$, which has a $\mathbb{Z}_{N'}$ one-form symmetry of anomaly p' with

$$N' = \frac{NL}{m^2} \qquad , \qquad p' = \frac{p}{L} \ . \tag{2.113}$$

They satisfy $\gcd(N', p') = 1$. Then we can further apply the generalized gauging operation with respect to this $\mathbb{Z}_{N'}$ one-form symmetry to find

$$\frac{\mathcal{T}/\mathbb{Z}_m \otimes \mathcal{A}^{N',-p'}}{\mathbb{Z}_{N'}} \ . \tag{2.114}$$

The goal of this appendix is to show that this is the same as the answer in (2.52)

$$\frac{\mathcal{T}/\mathbb{Z}_L \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_{N/L}} = \frac{\mathcal{T} \otimes \mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N} \ . \tag{2.115}$$

Note, as a check that for $L = m$ they are trivially the same.

We will use the canonical duality in (2.44) [59]

$$\mathcal{T} \longleftrightarrow \frac{\mathcal{T} \otimes (\mathcal{Z}_N)_{-pN}}{\mathbb{Z}_N} \ . \tag{2.116}$$

The second factor in the numerator can be described by the Lagrangian (2.31)

$$\int \left(-\frac{pN}{4\pi} x dx + \frac{N}{2\pi} x dy \right) \ . \tag{2.117}$$

Its lines are generated by b and c (2.32)

$$b = \exp(i \oint y), \quad c = \exp(ip \oint x - i \oint y) \, . \tag{2.118}$$

In this dual description, the $\mathbb{Z}_N$ one-form symmetry is entirely in the $(\mathcal{Z}_N)_{-pN}$ factor and it is generated by b.

The duality allows us to only keep tack of the $(\mathcal{Z}_N)_{-pN}$ factor in various procedures (and ignore the TQFT $\mathcal{T}$).

Gauging the anomaly free $\mathbb{Z}_L$ subgroup in $(\mathcal{Z}_N)_{-pN}$ is the same as redefining x as $x' = Lx$ and viewing x' as a $U(1)$ gauge field. This leads to $(\mathcal{Z}_K)_{-p'K}$ with $K = N/L$ and $p' = p/L$. Since $\gcd(K, p') = 1$, the theory $(\mathcal{Z}_K)_{-p'K}$ factorizes (2.43)

$$(\mathcal{Z}_K)_{-p'K} = \mathcal{A}^{K,p'} \otimes \mathcal{A}^{K,-p'} \, , \tag{2.119}$$

where the first and second minimal theories are generated by b and c, respectively.

Then, gauging the anomaly free $\mathbb{Z}_m \subset \mathbb{Z}_N$ (which includes $\mathbb{Z}_L$) in $(\mathcal{Z}_N)_{-pN}$ is equivalent to gauging the anomaly free $\mathbb{Z}_{m/L}$ subgroup generated by $b^{N/m}$ in $(\mathcal{Z}_K)_{-p'K} = \mathcal{A}^{K,p'} \otimes \mathcal{A}^{K,-p'}$. Only the first minimal theory is involved in the gauging, which reduces it to $\mathcal{A}^{N',p'}$ with $N' =$

$K(L/m)^2 = NL/m^2$ and $p' = p/L$.[30] This implies that

$$\frac{\mathcal{T}}{\mathbb{Z}_m} \longleftrightarrow \left(\frac{\mathcal{T}\otimes(\mathcal{Z}_N)_{-pN}}{\mathbb{Z}_N}\right)\Big/\mathbb{Z}_m \longleftrightarrow \frac{\mathcal{T}\otimes\mathcal{A}^{N',p'}\otimes\mathcal{A}^{K,-p'}}{\mathbb{Z}_N}\,. \tag{2.120}$$

The remaining global symmetry is $\mathbb{Z}_{N'}$ and it is carried by the second factor in the numerator. Applying the generalized operation with respect to this symmetry removes this factor and leads to

$$\frac{\mathcal{T}\otimes\mathcal{A}^{N/L,-p/L}}{\mathbb{Z}_N}\,. \tag{2.121}$$

We conclude that the final theory (2.121) is the same for any choice of $\mathbb{Z}_m \supset \mathbb{Z}_L$.

2.9 Appendix E: Two-form $\mathbb{Z}_N$ gauge theory in 4d

The 4d topological $\mathbb{Z}_N$ two-form gauge theory of a gauge field $\mathcal{B}\in\mathcal{H}^2(\mathcal{M}_4,\mathbb{Z}_N)$

$$S = 2\pi\frac{p}{2N}\int\mathcal{P}(\mathcal{B})\,, \tag{2.122}$$

has a continuum description [29, 5]

$$S = \int\left(\frac{pN}{4\pi}BB + \frac{N}{2\pi}BdA\right)\,, \tag{2.123}$$

where A is a $U(1)$ one-form gauge field and B is a $U(1)$ two-form gauge field. A constrains B to be a $\mathbb{Z}_N$ two-form gauge field $B\to\frac{2\pi}{N}\mathcal{B}$. The theory has a one-form gauge symmetry

$$B\to B-d\lambda,\quad A\to A+p\lambda\,. \tag{2.124}$$

[30] More generally, $\mathcal{A}^{M,r}$ with $\gcd(M,r)=1$ is generated by a line z such that $z^M=1$ and the spin of z is $\frac{r}{2M}$. When $M=\hat{M}q^2$ with $\hat{M},q\in\mathbb{Z}$, it has a $\mathbb{Z}_q$ anomaly free subgroup generated by $z^{\hat{M}q}$. (It is anomaly free because the spin of this line is $\frac{r\hat{M}}{2}$.) The gauged theory $\mathcal{A}^{M,r}/\mathbb{Z}_q$ has $\hat{M}$ lines generated by z^q (with $(z^q)^{\hat{M}}=1$), whose spin is $\frac{r}{2\hat{M}}$. Therefore, the resulting theory is $\mathcal{A}^{M,r}/\mathbb{Z}_q=\mathcal{A}^{\hat{M},r}$.

Under the gauge transformation, the action is shifted by

$$-\int \left(\frac{pN}{4\pi} d\lambda d\lambda + \frac{N}{2\pi} d\lambda dA \right) . \tag{2.125}$$

On a closed spin manifolds it is always a multiple of 2π, but on general closed manifolds it is a multiple of 2π only when pN is even. The parameter p has an identification of $p \sim p + 2N$ on non-spin manifolds and $p \sim p + N$ on spin manifolds.

Define

$$L = \gcd(N, p), \quad K = N/L . \tag{2.126}$$

The theory has L surface operators generated by

$$U = \exp(i \oint B), \quad U^L = 1 . \tag{2.127}$$

and L genuine lines operators generated by

$$V = \exp(iK \oint_{\partial\Sigma} A + ipK \int_{\Sigma} B), \quad V^L = 1 \tag{2.128}$$

(they are genuine line operators because they do not depend on the surface Σ). These operators and their correlation functions are identical to the ones in a $\mathbb{Z}_L$ gauge theory, and they realize a $\mathbb{Z}_L = \mathbb{Z}_N/\mathbb{Z}_K$ one-form symmetry. As we will discuss below, depending on N and p this $\mathbb{Z}_L$ gauge theory could be twisted on nonspin manifolds.

This theory can arise as the low-energy approximation of a microscopic theory whose $\mathbb{Z}_N$ one-form symmetry is spontaneously broken to $\mathbb{Z}_K$. Examples of such UV theories are a $PSU(N)$ gauge theory (discussed in Section 2.4) and the Walker-Wang lattice model [42, 68, 43].

There are also open surface operators generated by

$$\exp(i \oint_{\partial\Sigma} A + ip \int_{\Sigma} B) . \tag{2.129}$$

They are genuine line operators if the surface dependence is trivial, otherwise, the surface is physical and the operators can only have contact terms. Hence, we will not include them in the list of operators.

Two special cases are particularly interesting. First, for $p = 0$ this theory is the same as an ordinary $\mathbb{Z}_N$ gauge theory. Here B implements the constraint that A is a $\mathbb{Z}_N$ one-form gauge field.

The second special case is $p = N$. On a spin manifold, it is the same as $p = 0$, i.e. it is an ordinary $\mathbb{Z}_N$ gauge theory. On a nonspin manifold, we must have $pN \in 2\mathbb{Z}$ so, $p = N$ can happen only when N is even. Then, the action (2.122) is the same as

$$\pi \int \mathcal{P}(\mathcal{B}) = \left(\pi \int w_2(\mathcal{M}_4) \cup \mathcal{B}\right) \bmod 2\pi\,, \tag{2.130}$$

where $w_2(\mathcal{M}_4)$ is the second Stiefel-Whitney class of the manifold. This fact has some interesting consequences. First, it shows that the possible added term (2.130) on nonspin manifolds for even N was already included in our labelling by $p = 0, 1, \cdots, 2N-1$. Second, it makes it manifest that on spin manifolds we can identify $p \sim p + N$. Finally, it shows that on a non-spin manifold, the theory with even $p = N$, which is an ordinary $\mathbb{Z}_N$ gauge theory on a spin manifold, becomes a $\mathbb{Z}_N$ gauge theory coupled to $w_2(\mathcal{M}_4)$ of the manifold.

In the $\mathbb{Z}_N$ gauge theory, the surface $\oint \mathcal{B}$ is the world volume of a $\mathbb{Z}_N$ magnetic string. It generates the one-form symmetry that acts on the Wilson lines in the $\mathbb{Z}_N$ gauge theory. The coupling (2.130) is thus equivalent to turning on a background gauge field for this one-form symmetry $\tilde{\mathcal{B}}_\mathcal{C} = (N/2) w_2(\mathcal{M}_4) \bmod N$. One consequence of this is that on a non-spin manifold, the basic $\mathbb{Z}_N$ Wilson line, which corresponds to the microscopic line $\oint A$, is attached to the surface $\frac{2\pi}{N} \int \tilde{\mathcal{B}}_\mathcal{C} = \pi \int w_2(\mathcal{M}_4)$. The surface represents an anomaly in the theory along the line and it implies that if we view this line as the worldline of a probe particle, this particle is a fermion [66, 10]. The conclusion is that the theory with $p = N$ for even N is a (twisted) $\mathbb{Z}_N$ gauge theory with fermionic probe particles.

Another way to see this is as follows. $w_2(\mathcal{M}_4)$ of a manifold is the obstruction to lifting the $SO(4)$ tangent bundle to an $Spin(4)$ bundle. Thus the background $\tilde{\mathcal{B}}_\mathcal{C} = (N/2)w_2(\mathcal{M}_4)$ modifies the symmetry to be

$$\frac{\mathbb{Z}_N^{\text{gauge}} \times Spin(4)}{\mathbb{Z}_2} \ . \tag{2.131}$$

The quotient identifies $\mathbb{Z}_2 \subset \mathbb{Z}_N^{\text{gauge}}$ with the $\mathbb{Z}_2$ fermion parity symmetry $(-1)^F$ of the Lorentz symmetry. Thus the $\mathbb{Z}_N$ Wilson lines in the odd-charge representations also transform under the fermion parity, and they represent fermionic probe particles.

Let us examine in more detail the path integral of the $\mathbb{Z}_N$ gauge theory coupled to fixed $w_2(\mathcal{M}_4)$ of the manifold. The path integral is performed over twisted $\mathbb{Z}_N$ gauge fields as in the symmetry (2.131), which is an extension of the bosonic Lorentz group $SO(4)$ by the $\mathbb{Z}_N$ gauge group. The twisted $\mathbb{Z}_N$ gauge field is a one-cochain a valued in $\mathbb{Z}_N$ that satisfies

$$\delta a = (N/2)w_2(\mathcal{M}_4) \text{ mod } N \ . \tag{2.132}$$

The path integral sums over all possible a with fixed $w_2(\mathcal{M}_4)$ of the manifold.

If $N/2$ is odd, $\mathbb{Z}_N \cong \mathbb{Z}_{N/2} \times \mathbb{Z}_2$ and the symmetry (2.131) is isomorphic to $\mathbb{Z}_{N/2} \times Spin(4)$. Another way to see this is that (2.132) implies $w_2(\mathcal{M}_4) = \delta a$ mod 2 by reducing both sides to mod 2. On a general manifold $w_2(\mathcal{M}_4)$ is non-trivial, and therefore the gauge field a cannot be defined everywhere. Indeed, near a surface operator insertion $\oint \mathcal{B}$ that generates the one-form symmetry, the gauge field a is not well-defined: a Wilson line of a that links with the surface transforms by its one-form charge. For a similar discussion, see [69].

Let us return to generic p. On a spin manifold the theory is the same (up to a geometric counterterm) as a $\mathbb{Z}_L$ gauge theory [5]. On a non-spin manifold the situation is more interesting. For odd N the equivalence to a $\mathbb{Z}_L$ gauge theory is still true [5]. However, for even N a new subtlety occurs, which is related to (2.130). The computation in [5] can be interpreted to mean that when both $K = N/L$ and p/L are odd (which can happen only when both N, p, and therefore also L are even), or equivalently, when pN/L^2 is odd the equivalent $\mathbb{Z}_L$ gauge theory is actually a twisted theory as mentioned above. In terms of a $\mathbb{Z}_L$ two-form gauge field, its action is

$$\pi\frac{pN}{L^2}\int w_2(\mathcal{M}_4)\cup\mathcal{B}^{(L)}\,, \tag{2.133}$$

Similarly, the basic line operator in the $\mathbb{Z}_L$ gauge theory corresponding to $\exp(i\oint KA)$ also represents a fermion when pN/L^2 is odd.

This discussion of odd pN/L^2 is consistent with our 3d analysis in Section 2.2.4, where we saw that in this case the generating line of the $\mathbb{Z}_L$ one-form symmetry is a fermion and the 3d theory has a mixed anomaly between the $\mathbb{Z}_L$ global symmetry and gravity (2.133).

Next, consider the $\mathbb{Z}_N$ two-form gauge theory on a manifold with a boundary [29, 5].[31] We choose the Dirichlet boundary condition $B| = 0$. This explicitly breaks the one-form gauge symmetry on the boundary so the line $\hat{V} = \exp(i\oint A)$ is liberated there and it satisfies

$$\begin{aligned}\langle\hat{V}(\gamma)\hat{V}(\gamma')\rangle &= \frac{1}{Z}\int DADB\exp\left(i\int\frac{pN}{4\pi}BB+\frac{N}{2\pi}BdA\right)\exp\left(i\oint_\gamma A+i\oint_{\gamma'}A\right)\\ &= \exp\left(\frac{2\pi i p}{N}\ell(\gamma,\gamma')\right)\,,\end{aligned} \tag{2.134}$$

where $\gamma, \gamma' \in \partial\mathcal{M}_4$ and $\ell(\gamma,\gamma')$ is the linking number of γ and γ'. When $L = \gcd(N,p) = 1$, the bulk theory is trivial and the N lines generated by $\hat{V}$ form the minimal Abelian TQFT $\mathcal{A}^{N,-p}$ that has a $\mathbb{Z}_N$ one-form symmetry of label p. For general L, $V = \hat{V}^K$ can smoothly move into the bulk so it has trivial braiding. Therefore the lines on the boundary do not form a modular TQFT. However, we can perform a quotient with the bulk lines generated by V to find an effective 3d TQFT $\mathcal{A}^{K,-p/L}$. If $K, p/L$ are odd, the line V has half-integer spin so from the boundary perspective, V can only be taken as ψ the transparent spin-half line and the $2K$ lines generated by $\hat{V}$ form a consistent spin TQFT $\mathcal{A}^{K,-p/L}$.

[31] Some examples were considered in [68, 43] in the context of the Walker-Wang lattice model.

2.10 Appendix F: Minimal TQFTs for general one-form symmetries

In this Appendix, we generalized the previous discussion to a general discrete one-form symmetry $\mathcal{A} = \prod \mathbb{Z}_{N_I}$.

We start with an arbitrary TQFT with one-form global symmetry $\prod \mathbb{Z}_{N_I}$ and analyze its symmetry lines, as in the introduction and in Section 2.2.1. Each $\mathbb{Z}_{N_I}$ factor is generated by a line a_I. The symmetry group means that they satisfy the mutual braiding

$$a_I^{s_I}(\gamma) a_J^{s_J}(\gamma') = a_J^{s_J}(\gamma') e^{-\frac{2\pi i s_I s_J m_{IJ}}{N_I}} \tag{2.135}$$

where γ circles around γ' as in Figure 2.2 and $m_{IJ} \in \mathbb{Z}_{N_I}$. Consistency of the mutual braiding implies $m_{IJ} N_J = m_{JI} N_I \mod N_I N_J$ and thus

$$m_{IJ} = \frac{N_I P_{IJ}}{N_{IJ}} \ , \qquad \text{with} \ \ N_{IJ} \equiv \gcd(N_I, N_J) \ \ , \ \ P_{IJ} = P_{JI} \in \mathbb{Z}. \tag{2.136}$$

This means that the spins of the symmetry lines are

$$h\left(\prod_I a_I^{s_I}\right) = \sum_{I,J} \frac{p_{IJ} s_I s_J}{2N_{IJ}} \mod 1 \, , \qquad p_{IJ} = P_{IJ} \text{ or } P_{IJ} + N_{IJ} \ . \tag{2.137}$$

The one-form symmetry $\mathcal{A} = \prod \mathbb{Z}_{N_I}$ is characterized by the symmetric integral matrix p_{IJ} that satisfies

$$p_{II} \sim p_{II} + 2N_I \quad \text{ and } \quad p_{IJ} \sim p_{IJ} + N_{IJ} \text{ for } I \neq J \ . \tag{2.138}$$

Imposing the condition $a_I^{N_I} = 1$ requires $p_{II} N_I \in 2\mathbb{Z}$. Otherwise, the theory is a spin theory.

The braiding between $V = \prod a_I^{s_I}$ and $V' = \prod a_I^{s'_I}$ is given by

$$e^{2\pi i (h[V] + h[V'] - h[VV'])} = \exp\left(-2\pi i \sum_{I,J} \frac{p_{IJ}}{N_{IJ}} s_I s'_J \right) \ . \tag{2.139}$$

It will be convenient to view the braiding as a bilinear map $\mathcal{A}\times\mathcal{A}\to U(1)$. Equivalently, it defines a linear map $M:\mathcal{A}\to\hat{\mathcal{A}}=\text{Hom}(\mathcal{A},U(1))$.

An example of a TQFT that has the one-form symmetry $\mathcal{A}=\prod\mathbb{Z}_{N_I}$ characterized by p_{IJ} is the Abelian Chern-Simons theory

$$-\sum_{I,J}\frac{p_{IJ}N_IN_J}{4\pi N_{IJ}}x^Idx^J+\sum_I\frac{N_I}{2\pi}x^Idy^I\ , \tag{2.140}$$

where the generating lines a_I are

$$a_I=\exp\left(i\oint y^I\right)\ . \tag{2.141}$$

The symmetry lines in $\mathcal{L}=\ker M$ have trivial braiding with all the symmetry lines in $\mathcal{A}$. Thus the braiding (2.139) is degenerate if and only if $\mathcal{L}$ is non-trivial. If $\mathcal{L}$ is trivial, the symmetry lines form a modular $3d$ TQFT, and we will call it the minimal Abelian TQFT for the one-form symmetry $\mathcal{A}$, denoted by $\mathcal{A}^{\{N_I\},\{p_{IJ}\}}$. An example is the $(\mathcal{Z}_N)_0$ theory that corresponds to the minimal theory with $N_1=N_2=N$, $p_{11}=p_{22}=0$ and $p_{12}=p_{21}=1$.

Next we discuss the anomaly for the one-form symmetry $\mathcal{A}$. From an argument similar to that in Section 2.2.4, the anomaly is characterized by the symmetric matrix p_{IJ}, and can be described by the following 4d term with background two-form gauge fields $\mathcal{B}_\mathcal{C}\in H^2(\mathcal{M}_4,\mathcal{A})$:

$$2\pi\int\mathcal{P}_h(\mathcal{B}_\mathcal{C})=2\pi\sum_I\frac{p_{II}}{2N_I}\int_{\mathcal{M}_4}\mathcal{P}(\mathcal{B}_\mathcal{C}^I)+\sum_{I<J}2\pi\frac{p_{IJ}}{N_{IJ}}\int_{\mathcal{M}_4}\mathcal{B}_\mathcal{C}^I\cup\mathcal{B}_\mathcal{C}^J\ , \tag{2.142}$$

where on the left hand side $\mathcal{P}_h$ is the generalized Pontryagin square with the quadratic function h that maps a line in $\mathcal{A}$ to its spin (2.137) (for a review see *e.g.* [10]). On the right hand side we express the anomaly in the basis $\{a_I\}$ for $\mathcal{A}$, and $\mathcal{B}_\mathcal{C}^I\in H^2(\mathcal{M}_4,\mathbb{Z}_{N_I})$ are the components of $\mathcal{B}_\mathcal{C}$ in this basis.

Let us use the anomaly (2.142) as the bulk action and promote the gauge field $\mathcal{B}_\mathcal{C}$ to be a dynamical gauge field $\mathcal{B}$. The theory has surfaces given by the fluxes of $\mathcal{B}$, and magnetic lines, both are described by the group $\mathcal{A}$ with the group multiplication given by the fusion of operators.

As we will see, some of the operators have trivial correlation functions, and they should not be included in the list of non-trivial operators. The equation of motion for the gauge field $\mathcal{B}$ in (2.142) implies

$$\exp\left(2\pi i \oint M(\mathcal{B})\right) = 1 \ , \tag{2.143}$$

and thus the surfaces generated by (2.143) have trivial correlation functions, while the non-trivial surfaces are described by the group $\mathcal{L} \cong \ker M$. The surfaces generated by (2.143) are described by the group $\mathcal{K} \cong \operatorname{im} M \cong \mathcal{A}/\mathcal{L}$, and the open version of them describe the line operators that have trivial correlation functions. Thus the non-trivial line operators are described by the quotient $\mathcal{L}$. The lines realize a faithful one-form symmetry $\mathcal{L}$ generated by the non-trivial surfaces. The theory can describe the spontaneous breaking of the one-form symmetry $\mathcal{A}$ generated by the surfaces to the subgroup $\mathcal{K}$ generated by the surfaces in (2.143).

Note that these $\mathcal{K}$ and $\mathcal{L}$ generalize the groups $\mathbb{Z}_K$ and $\mathbb{Z}_L$ in the case $\mathcal{A} = \mathbb{Z}_N$ that we have been discussing throughout most of this chapter.

We can also study the bulk theory in the continuum description.

$$\int_{\mathcal{M}_4} \sum_{I,J} \frac{p_{IJ} N_I N_J}{4\pi N_{IJ}} B_I B_J + \sum_I \frac{N_I}{2\pi} B_I dA_I \,, \tag{2.144}$$

in terms of $U(1)$ two-form gauge fields B_I and $U(1)$ one-form gauge fields A_I. It has a one-form gauge symmetry

$$B_I \to B_I - d\lambda_I, \quad A_I \to A_I + \sum_J \frac{p_{IJ} N_J}{N_{IJ}} \lambda_J \,. \tag{2.145}$$

Therefore the lines are attached to surfaces

$$\exp\left(i \oint_\gamma \sum s_I A_I + i \int_\Sigma \sum s_I \frac{p_{IJ} N_J}{N_{IJ}} B_J\right) \ , \quad \gamma = \partial\Sigma \,. \tag{2.146}$$

They are genuine lines, if and only if s_I is in $\mathcal{L}$, the kernel of M. Effectively, the theory becomes a one-form (ordinary) $\mathcal{L}$ gauge theory. It may couple to $w_2(\mathcal{M}_4)$ of the manifold such that the symmetry group is twisted as described in Appendix E.

On an open manifold with the choice of boundary condition $B_I| = 0$, the gauge symmetry (2.145) is completely broken on the boundary and all the bulk lines are liberated there. Their braiding is the same as (2.139) with $p_{IJ} \to -p_{IJ}$ (see Appendix E for a similar calculation). If $\mathcal{L}$ is trivial, they form a modular TQFT $\mathcal{A}^{\{N_I\},\{-p_{IJ}\}}$. Otherwise, the bulk lines associated to $\mathcal{L}$ have trivial braiding and we can only find an effective boundary theory consisting of the lines in $\mathcal{A}/\mathcal{L}$ by modding out by the bulk lines.

Alternatively, as in the main text we can consider the boundary condition $B_I| \neq 0$. To do this, we start with a 4d-3d system with an SPT phase (2.142) in the bulk and a 3d TQFT $\mathcal{T}$ on the boundary that has an anomalous one-form symmetry coupled to the classical gauge fields $(B_{\mathcal{C}})^I$, and the anomaly is cancelled by the inflow. We can then promote the gauge fields to be dynamical. When $\mathcal{L}$ is trivial, the bulk dynamics is trivial and there is a meaningful boundary theory

$$\mathcal{T}' = \frac{\mathcal{T} \otimes \mathcal{A}^{\{N_I\},\{-p_{IJ}\}}}{\prod \mathbb{Z}_{N_I}} \ , \tag{2.147}$$

It is obtained from $\mathcal{T}$ by removing all lines that are not invariant under the one-form symmetry. When $\mathcal{L}$ is non-trivial, the theory above is not modular, and we can find an effective boundary theory as a quotient by the transparent bulk lines associated to $\mathcal{L}$. The discussion can be generalized easily to interfaces.

Chapter 3

A Modified Villain Formulation of Fractons and Other Exotic Theories with Subsystem Symmetries

3.1 Preliminary and summary

The surprising discoveries of [70, 15] have stimulated exciting work on fracton models. This subject is reviewed nicely in [71, 72], which include many references to the original papers.

One of the peculiarities of these models is that their low-energy behavior does not admit a standard continuum field theory description. Finding such a description is important for two reasons. First, it will give a simple universal framework to discuss fracton phases, will organize the distinct models, and will point to new models. Second, since the field theory will inevitably be non-standard, this will teach us something new about quantum field theory.

3.1.1 Overview of continuum field theories for exotic models

There have been many progress on constructing and analyzing these exotic field theories, including theories of fractons [73–79, 11, 80–84]. The resulting theories are simple-looking, but subtle. They capture the low-energy dynamics and the behavior of massive charged particles of the underlying lattice models as probe particles.

The main features of these exotic continuum field theories are the following:

1. Unlike the underlying lattice models, which are nonlinear, the low-energy continuum actions are quadratic, i.e., the theories are free.

2. The spatial derivatives in the continuum actions are such that we should consider discontinuous and even singular field configurations and gauge transformation parameters. In fact, such discontinuities are essential in order to reproduce the microscopic lattice results.

3. Some observables, e.g., the ground state degeneracy and the spectrum of some charged states, are divergent in the continuum theory. In order to make them finite, we need to introduce a UV cutoff, i.e., a nonzero lattice spacing a. Even though these observables are divergent, the regularized versions are still meaningful.

4. Some of the continuum theories have emergent global symmetries, which are not present in the microscopic lattice models. For example, winding symmetries and magnetic symmetries, which depend on continuity of the fields, are absent on the lattice, but are present in the low-energy, continuum theory.

5. Depending on the specific microscopic description, the global symmetry of the low-energy theory can involve a quotient of the global symmetry of the lattice model. Some symmetry operators act trivially in the low-energy theory and we should quotient by them.

6. The analysis of the continuum theories leads to certain strange states that are charged under the original or the emergent symmetries with energy of order $\frac{1}{a}$. Because of the singularities and the energy of these states, this analysis appears questionable and was referred to as an "ambitious analysis."

7. The continuum models exhibit surprising dualities between seemingly unrelated models. These dualities are IR dualities, rather than exact dualities, of the underlying lattice mod-

els. They depend crucially on the precise global symmetries of the long-distance theories, including the emergent symmetries and the necessary quotients of the microscopic symmetry. These dualities also map correctly the strange charged states we mentioned above.

8. The continuum models have peculiar robustness properties. (See [11], for a general discussion of robustness in condensed-matter physics and in high-energy physics.) Some symmetry violating operators, which could have destabilized the long-distance theory, have infinitely large dimension in that theory, and therefore they are infinitely irrelevant. This comment applies both to some of the underlying symmetries of the microscopic models as well as to the emergent global symmetries.

3.1.2 Modified Villain lattice models

The purpose of this chapter is to explore further the lattice models, rather than their continuum limits. We will deform the existing lattice models in a continuous way to find new lattice models with interesting properties. In particular, despite being lattice models with nonzero lattice spacing a, they have many of the features of the continuum models we mentioned above.

Although this is not essential, we find it easier to use a discretized Euclidean spacetime lattice. Then, following Villain [85], we replace the lattice model with another model, which is close to it at weak coupling. We replace the compact fields, which take values in S^1 or $\mathbb{Z}_M$, by non-compact fields, which take values in $\mathbb{R}$ and $\mathbb{Z}$ respectively. Then, we compactify the field space by gauging an appropriate $\mathbb{Z}$ global symmetry. In most cases, this is achieved by adding certain integer-valued gauge fields.

So far, this is merely the Villain version of the original model. Then, we further modify the model by constraining the field strength of the new integer-valued gauge fields to zero. We refer to this model as the *modified Villain version* of the system.

Let us demonstrate this in the standard 2d Euclidean XY-model. (See Appendix B.1, for a more detailed discussion of this model.) The degrees of freedom are circle-valued fields ϕ on the sites of the lattice and the standard lattice action is

$$\beta \sum_{\text{link}} [1 - \cos(\Delta_\mu \phi)] \ , \tag{3.1}$$

where $\mu = x, y$ labels the directions and $\Delta_\mu \phi$ are the lattice derivatives. The standard Villain version of this action is

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 \ . \tag{3.2}$$

Here ϕ is a real-valued field and n_μ is an integer-valued field on the links. This theory has the $\mathbb{Z}$ gauge symmetry

$$\phi \sim \phi + 2\pi k \ , \qquad n_\mu \sim n_\mu + \Delta_\mu k \ , \tag{3.3}$$

where k is an integer-valued gauge parameter on the sites. Next, we deform the model further by constraining the gauge invariant field strength of the gauge field n_μ,

$$\mathcal{N} \equiv \Delta_x n_y - \Delta_y n_x \ , \tag{3.4}$$

to zero [86]. We will refer to this and similar constraints as flatness constraints. We do that by adding a Lagrange multiplier $\tilde{\phi}$, and then the full action becomes

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquette}} \tilde{\phi} \mathcal{N} \ . \tag{3.5}$$

We refer to the action (3.5) as the modified Villain version of the original action (3.1). We will analyze it in detail in Appendix B.1.

In the bulk of the chapter, we will apply this procedure to the lattice models of [87,88,17,89–92,11,80,81]. These include, in particular, the X-cube model [17]. The resulting lattice models turn out to share some of the nice features of our continuum theories, even though they are on the lattice. Comparing with the list above, these lattice models have the following features:

1. The actions are quadratic in the fields; these theories are free.

2. The fields and the gauge parameters are discontinuous on the lattice. As we take the continuum limit, they become more continuous. But some discontinuities remain. In fact, our rules in [79, 11, 80–84] about the allowed singularities in the fields and the gauge transformation parameters follow naturally from this lattice model.

3. Since these are lattice models, there is no need to introduce another regularization.

4. All the emergent symmetries of the continuum theories (except continuous translations) are exact symmetries of these lattice models. Starting with these models, there are no emergent symmetries.

5. These lattice models do not exhibit additional symmetries beyond those of the continuum models. No quotient of the microscopic global symmetry is necessary.

6. The strange charged states with energy of order $\frac{1}{a}$ of the "ambitious analysis" of the continuum theories are present in the new lattice models and they have precisely the expected properties.

7. All the surprising dualities of the continuum models are present already on the lattice. These are not IR dualities, but exact dualities. All of them follow from using the Poisson resummation formula

$$\begin{aligned}&\sum_n \exp\left[-\frac{\beta}{2}(\theta-2\pi n)^2+in\tilde{\theta}\right]\\&\qquad=\frac{1}{\sqrt{2\pi\beta}}\sum_{\tilde{n}}\exp\left[-\frac{1}{2(2\pi)^2\beta}(\tilde{\theta}-2\pi\tilde{n})^2-\frac{i\theta}{2\pi}(2\pi\tilde{n}-\tilde{\theta})\right]\ .\end{aligned}\tag{3.6}$$

8. Our new lattice models have the same global symmetry as the low-energy continuum limit. Therefore, there is no need to discuss the robustness of the low-energy theory with respect the operators violating these symmetries. The analysis of robustness with respect to symmetry-violating operators should be performed in the low-energy continuum theory and it is the same in the original models and in these new ones. We note that our

lattice theory is natural once this new symmetry is imposed. (See [11] for a discussion of naturalness and its relation to robustness.)

To summarize, we deform the original lattice models to their modified Villain versions. The new models exhibit some of the special properties of the continuum theories even without taking the continuum limit.

Furthermore, it is clear that, at least for some range of coupling constants, the previous models and the new deformed models flow to the same long-distance theories, which are described by the continuum field theories mentioned earlier.

One interesting aspect of our new lattice models is that they exhibit global symmetries with 't Hooft anomalies. For example, the model (3.5) has a global $U(1)$ momentum symmetry and a global $U(1)$ winding symmetry. These symmetries act locally ("on site"), but they still have a mixed anomaly. The anomaly arises because the Lagrangian density and even its exponential are not invariant under these two symmetries — instead, only the action, or its exponential, is invariant. See Appendix B.1, for a more detailed discussion.

We should add another clarifying comment. The original lattice model can have several different phases. The Villain version of that model has the same phases. However, this is typically not the case for the modified model. In some cases it describes one of the phases of the original model and other phases that that model does not have.

For example, as we will discuss in detail in Appendix B.1, the model (3.5) describes the large β gapless phase of the 2d XY-model (3.1) or (3.2). But instead of describing its gapped phase with small β, it describes other continuum theories there. This behavior is the same as that of the $c = 1$ conformal field theory with arbitrary radius.

Another example, which we will discuss in Appendix C.1, is the 3d $U(1)$ gauge theory. The standard lattice model and its Villain version have a gapped confining phase [93]. Our modified version of that model is gapless and is similar to the corresponding continuum gauge theory.

As we said above, some of our lattice models have global continuous symmetries with 't Hooft anomalies. This means that their long-distance behavior must be gapless. This is consistent with the fact that they are gapless even when the original lattice model is gapped.

Another perspective on these new lattice models is the following. Since our exotic continuum models involve discontinuous field configurations, their analysis can be subtle. The new lattice models can be viewed as rigorous presentations of the continuum models. In fact, as we said above, they lead to the same answers as our continuum analysis including the more subtle "ambitious analysis", thus completely justifying it.

In order to demonstrate our approach, we will use it in Appendices A, B, and C to review some well-known models. In particular, we will present lattice models of various spin systems (including the XY-model (3.1)) and gauge theories, which share many of the properties of their continuum counterparts. In addition to demonstrating our approach, some people might find that discussion helpful. It relates the condensed-matter perspective to the high-energy perspective of these theories.

3.1.3 Outline

Following [11, 80–83], Sections 3.2 and 3.3 are divided into three parts. We study an XY-type model, then the $U(1)$ gauge theory associated with the momentum symmetry of this XY-type model, and then the corresponding $\mathbb{Z}_N$ gauge theory. We present the modified Villain lattice action of each model, dualize it (if possible) using the Poisson resummation formula (3.6) for the integer-valued gauge fields, discuss the global symmetries, and take the continuum limit. All these modified Villain lattice models exhibit all the peculiarities of the corresponding continuum theories of [11, 80, 81].

Even though we do not present it here, we have performed the same analysis for the exotic 3+1d continuum theories of [83], and we found similar results for the dualities and global symmetries of these modified Villain models. In particular, we have shown that the modified Villain formulation of the $\mathbb{Z}_2$ checkerboard model [17] is exactly equivalent to two copies of the modified Villain formulation of the $\mathbb{Z}_2$ X-cube model. This equivalence can be regarded as the universal low-energy limit of the equivalence shown in the Hamiltonian formulation in [94].

In Section 3.2, we study the modified Villain formulation of the exotic 2+1d continuum theories of [11]. We start with the XY-plaquette model of [87] on a 2+1d Euclidean lattice, and present its modified Villain action. Next, we study the modified Villain formulation of the associated $U(1)$ lattice tensor gauge theory. Finally, we present two equivalent BF-type actions of the $\mathbb{Z}_N$ lattice gauge theory: one with only integer fields (integer BF-action), and another with real and integer fields (real BF-action). All these modified Villain lattice models behave exactly as the corresponding continuum theories of [11].

In Section 3.3, we study the modified Villain formulation of the exotic 3+1d continuum theories of [80, 81]. We present the modified Villain actions of the XY-plaquette model on a 3+1d Euclidean lattice, its associated $U(1)$ lattice tensor gauge theory, and the $\mathbb{Z}_N$ X-cube model. As in Section 3.2, these modified Villain models exhibit the same properties as their continuum counterparts in [81].

In three appendices we use our modified Villain formulation to review the properties of well-studied models. Some readers might find it helpful to read the appendices before reading Sections 3.2 and 3.3.

Appendix A is devoted to some classic quantum-mechanical systems. We start with the problem of particle on a ring with a θ-parameter. For $\theta \in \pi\mathbb{Z}$, our Euclidean lattice model exhibits a mixed 't Hooft anomaly between its charge conjugation symmetry and its $U(1)$ shift symmetry. We also use our Euclidean lattice formulation to study the quantum mechanics of a system whose phase space is a two-dimensional torus, a.k.a. the non-commutative torus.

In Appendix B, we discuss some famous 2d Euclidean lattice models using our modified Villain formulation. First, we study the modified Villain version of the 2d Euclidean XY-model. Unlike the standard XY-model, it has an exact winding symmetry and an exact T-duality. It is very similar to the continuum $c = 1$ conformal field theory of a compact boson. Then, we study the 2d Euclidean $\mathbb{Z}_N$ clock-model by embedding it into the XY-model.

In Appendix C, we study p-form $U(1)$ gauge theories on a d-dimensional Euclidean spacetime lattice. We discuss their duality and the role of the Polyakov mechanism for $p = d - 2$. We also study the p-form $\mathbb{Z}_N$ gauge theory. We briefly comment on the relation between $\mathbb{Z}_N$ toric code and the ordinary $\mathbb{Z}_N$ gauge theory.

3.2 2+1d (3d Euclidean) exotic theories

In this section, we describe modified Villain lattice models corresponding to the exotic 2+1d continuum theories of [11]. All lattice models discussed here are placed on a 3d Euclidean lattice with lattice spacing a, and L^μ sites in μ direction. We use integers $\hat{x}^\mu$ to label the sites along the μ direction, so that $\hat{x}^\mu \sim \hat{x}^\mu + L^\mu$.

Since the spatial lattice has a $\mathbb{Z}_4$ rotation symmetry, we will organize the fields according to the irreducible, one-dimensional representations $\mathbf{1}_n$ of $\mathbb{Z}_4$ with $n = 0, \pm 1, 2$ labeling the spin. In the discussion below, a field without any spatial index is in $\mathbf{1}_0$ and a field with the spatial indices xy is in $\mathbf{1}_2$.

3.2.1 ϕ-theory (XY-plaquette model)

We start with a Euclidean spacetime version of the XY-plaquette model of [87]. The degrees of freedom are phases $e^{i\phi}$ at every site with the action

$$\beta_0 \sum_{\tau\text{-link}} [1 - \cos(\Delta_\tau \phi)] + \beta \sum_{xy\text{-plaq}} [1 - \cos(\Delta_x \Delta_y \phi)] \ . \tag{3.7}$$

At large β_0, β, we can approximate the action by the Villain action

$$\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{xy\text{-plaq}} (\Delta_x \Delta_y \phi - 2\pi n_{xy})^2 \ , \tag{3.8}$$

with real-valued ϕ and integer-valued n_τ and n_{xy} fields on the τ-links and the xy-plaquettes, respectively. We interpret (n_τ, n_{xy}) as $\mathbb{Z}$ tensor gauge fields that make ϕ compact because of the gauge symmetry

$$\begin{aligned}
&\phi \sim \phi + 2\pi k \ , \\
&n_\tau \sim n_\tau + \Delta_\tau k \ , \\
&n_{xy} \sim n_{xy} + \Delta_x \Delta_y k \ ,
\end{aligned} \tag{3.9}$$

where k is an integer-valued gauge parameter on the sites.

We suppress the "vortices" by modifying the Villain action (3.8) as

$$\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{xy\text{-plaq}} (\Delta_x \Delta_y \phi - 2\pi n_{xy})^2 + i \sum_{\text{cube}} \phi^{xy} (\Delta_\tau n_{xy} - \Delta_x \Delta_y n_\tau) \ , \tag{3.10}$$

where ϕ^{xy} is a real Lagrange multiplier field on the cubes or dual sites of the lattice. It imposes $\Delta_\tau n_{xy} - \Delta_x \Delta_y n_\tau = 0$, which can be interpreted as vanishing field strength of the gauge field (n_τ, n_{xy}). We will refer to this and similar constraints as flatness constraints. ϕ^{xy} has a gauge symmetry

$$\phi^{xy} \sim \phi^{xy} + 2\pi k^{xy} \ , \tag{3.11}$$

where k^{xy} is an integer-valued gauge parameter on the cubes of the lattice. We will refer to (3.10) as the modified Villain version of (3.7).

Self-Duality

Using the Poisson resummation formula (3.6), we can dualize the modified Villain action (3.10) to

$$\begin{aligned}
&\frac{1}{2(2\pi)^2 \beta} \sum_{\text{dual } \tau\text{-link}} (\Delta_\tau \phi^{xy} - 2\pi n_\tau^{xy})^2 + \frac{1}{2(2\pi)^2 \beta_0} \sum_{\text{dual } xy\text{-plaq}} (\Delta_x \Delta_y \phi^{xy} - 2\pi n)^2 \\
&\qquad - i \sum_{\text{site}} \phi (\Delta_\tau n - \Delta_x \Delta_y n_\tau^{xy}) \ ,
\end{aligned} \tag{3.12}$$

where n_τ^{xy} and n are integer-valued fields on the dual τ-links and the dual xy-plaquettes respectively. We interpret (n_τ^{xy}, n) as $\mathbb{Z}$ tensor gauge fields that make ϕ^{xy} compact because of the gauge symmetry

$$\begin{aligned} \phi^{xy} &\sim \phi^{xy} + 2\pi k^{xy} \ , \\ n_\tau^{xy} &\sim n_\tau^{xy} + \Delta_\tau k^{xy} \ , \\ n &\sim n + \Delta_x \Delta_y k^{xy} \ . \end{aligned} \tag{3.13}$$

Here, the field ϕ is a Lagrange multiplier that imposes the constraint that the gauge invariant field strength of (n_τ^{xy}, n) vanishes; i.e., it is flat. Therefore, the modified Villain model (3.10) is self-dual with $\beta_0 \leftrightarrow \frac{1}{(2\pi)^2\beta}$.

Global symmetries

In all the three models, (3.7), (3.8), and (3.10), there is a $(\mathbf{1}_0, \mathbf{1}_2)$ *momentum dipole symmetry*, which acts on the fields as

$$\phi \to \phi + c^x(\hat{x}) + c^y(\hat{y}) \ , \tag{3.14}$$

where $c^i(\hat{x}^i)$ is real-valued. Due to the zero mode of the gauge symmetry (3.9), the momentum dipole symmetry is $U(1)$. Using (3.10), the components of the Noether current of the momentum dipole symmetry are

$$\begin{aligned} J_\tau &= i\beta_0(\Delta_\tau\phi - 2\pi n_\tau) = \frac{1}{2\pi}(\Delta_x\Delta_y\phi^{xy} - 2\pi n) \ , \\ J^{xy} &= i\beta(\Delta_x\Delta_y\phi - 2\pi n_{xy}) = \frac{1}{2\pi}(\Delta_\tau\phi^{xy} - 2\pi n_\tau^{xy}) \ . \end{aligned} \tag{3.15}$$

(J_τ, J^{xy}) are in the $(\mathbf{1}_0, \mathbf{1}_2)$ representations of $\mathbb{Z}_4$. They satisfy the $(\mathbf{1}_0, \mathbf{1}_2)$ dipole conservation equation

$$\Delta_\tau J_\tau = \Delta_x \Delta_y J^{xy} \ , \tag{3.16}$$

because of the equation of motion of ϕ. The momentum dipole charges are

$$\begin{aligned} Q^x(\hat{x}, \tilde{\mathcal{C}}^x) &= \sum_{\text{dual } xy\text{-plaq}\in\tilde{\mathcal{C}}^x} J_\tau + \sum_{\text{dual } \tau x\text{-plaq}\in\tilde{\mathcal{C}}^x} \Delta_x J^{xy} \ , \\ &= -\sum_{\text{dual } xy\text{-plaq}\in\tilde{\mathcal{C}}^x} n - \sum_{\text{dual } \tau x\text{-plaq}\in\tilde{\mathcal{C}}^x} \Delta_x n_\tau^{xy} \end{aligned} \tag{3.17}$$

where $\tilde{\mathcal{C}}^x$ is a strip along the dual xy- and τx-plaquettes in the τy plane at fixed $\hat{x}$. The second line can be interpreted as the Wilson "strip" operator of (n_τ^{xy}, n). Similarly, we can define $Q^y(\hat{y}, \tilde{\mathcal{C}}^y)$. When $\tilde{\mathcal{C}}^x$ and $\tilde{\mathcal{C}}^y$ are purely spatial at a fixed $\hat{\tau}$, the charges satisfy the constraint

$$\sum_{\hat{x}:\text{ fixed } \hat{\tau}} Q^x(\hat{x}) = \sum_{\hat{y}:\text{ fixed } \hat{\tau}} Q^y(\hat{y}) = \sum_{\text{dual } xy\text{-plaq: fixed } \hat{\tau}} J_\tau \ . \tag{3.18}$$

The charged momentum operators are $e^{i\phi}$.

The modified Villain model (3.10) also has a $(\mathbf{1}_2, \mathbf{1}_0)$ *winding dipole symmetry*, which acts on the fields as

$$\phi^{xy} \to \phi^{xy} + c_x^{xy}(\hat{x}) + c_y^{xy}(\hat{y}) \ , \tag{3.19}$$

where $c_i^{xy}(\hat{x}^i)$ is real-valued. By contrast, this symmetry is absent in the original lattice model (3.7) and its Villain version (3.8). Due to the zero mode of the gauge symmetry (3.13), the winding dipole symmetry is $U(1)$. The components of the Noether current of the winding dipole symmetry are

$$\begin{aligned} J_\tau^{xy} &= -\frac{i}{(2\pi)^2\beta}(\Delta_\tau\phi^{xy} - 2\pi n_\tau^{xy}) = \frac{1}{2\pi}(\Delta_x\Delta_y\phi - 2\pi n_{xy}) \ , \\ J &= -\frac{i}{(2\pi)^2\beta_0}(\Delta_x\Delta_y\phi^{xy} - 2\pi n) = \frac{1}{2\pi}(\Delta_\tau\phi - 2\pi n_\tau) \ . \end{aligned} \tag{3.20}$$

They satisfy the $(\mathbf{1}_2, \mathbf{1}_0)$ dipole conservation equation

$$\Delta_\tau J_\tau^{xy} = \Delta_x\Delta_y J \ , \tag{3.21}$$

because of the equation of motion of ϕ^{xy}. The winding dipole charges are

$$\begin{aligned} Q_x^{xy}(\hat{x}, \mathcal{C}^x) &= \sum_{xy\text{-plaq}\in\mathcal{C}^x} J_\tau^{xy} + \sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x J \ , \\ &= -\sum_{xy\text{-plaq}\in\mathcal{C}^x} n_{xy} - \sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x n_\tau \ , \end{aligned} \tag{3.22}$$

where $\mathcal{C}^x$ is a strip along the xy- and τx-plaquettes in the τy plane at fixed $\hat{x}$. The second line can be interpreted as the Wilson "strip" operator of (n_τ, n_{xy}). Similarly, we can define

$Q_y^{xy}(\hat{y}, \mathcal{C}^y)$. When $\mathcal{C}^x$ and $\mathcal{C}^y$ are purely spatial at a fixed $\hat{\tau}$, the charges satisfy the constraint

$$\sum_{\hat{x}:\ \text{fixed } \hat{\tau}} Q_x^{xy}(\hat{x}) = \sum_{\hat{y}:\ \text{fixed } \hat{\tau}} Q_y^{xy}(\hat{y}) = \sum_{xy\text{-plaq: fixed } \hat{\tau}} J_\tau^{xy} \ . \tag{3.23}$$

The charged winding operators are $e^{i\phi^{xy}}$.

There is a mixed 't Hooft anomaly between the two $U(1)$ global symmetries. One way to see this is to couple the system to the classical background gauge fields $(A_\tau, A_{xy}; N_{\tau xy})$ and $(\tilde{A}_\tau^{xy}, \tilde{A}; \tilde{N}_\tau)$ of the momentum and winding symmetries, respectively. Here $A_\tau, A_{xy}, \tilde{A}_\tau^{xy}, \tilde{A}$ are real-valued and $N_{\tau xy}, \tilde{N}_\tau$ are integer-valued. (See a similar discussion in Appendix B.1.) The action is:

$$\begin{aligned}
&\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\phi - A_\tau - 2\pi n_\tau)^2 + \frac{\beta}{2}\sum_{xy\text{-plaq}}(\Delta_x\Delta_y\phi - A_{xy} - 2\pi n_{xy})^2 \\
&\quad + i\sum_{\text{cube}}\phi^{xy}(\Delta_\tau n_{xy} - \Delta_x\Delta_y n_\tau + N_{\tau xy}) \\
&\quad - \frac{i}{2\pi}\sum_{xy\text{-plaq}}\tilde{A}_\tau^{xy}(\Delta_x\Delta_y\phi - A_{xy} - 2\pi n_{xy}) - \frac{i}{2\pi}\sum_{\tau\text{-link}}\tilde{A}(\Delta_\tau\phi - A_\tau - 2\pi n_\tau) - i\sum_{\text{site}}\tilde{N}_\tau\phi \ ,
\end{aligned} \tag{3.24}$$

with the gauge symmetry

$$\begin{aligned}
&\phi \sim \phi + \alpha + 2\pi k \ , && \phi^{xy} \sim \phi^{xy} + \tilde{\alpha}^{xy} + 2\pi k^{xy} \ , \\
&A_\tau \sim A_\tau + \Delta_\tau\alpha + 2\pi K_\tau \ , && \tilde{A}_\tau^{xy} \sim \tilde{A}_\tau^{xy} + \Delta_\tau\tilde{\alpha}^{xy} + 2\pi\tilde{K}_\tau^{xy} \ , \\
&A_{xy} \sim A_{xy} + \Delta_x\Delta_y\alpha + 2\pi K_{xy} \ , && \tilde{A} \sim \tilde{A} + \Delta_x\Delta_y\tilde{\alpha}^{xy} + 2\pi\tilde{K} \ , \\
&n_\tau \sim n_\tau + \Delta_\tau k - K_\tau \ , && \tilde{N}_\tau \sim \tilde{N}_\tau + \Delta_\tau\tilde{K} - \Delta_x\Delta_y\tilde{K}_\tau^{xy} \ . \\
&n_{xy} \sim n_{xy} + \Delta_x\Delta_y k - K_{xy} \ , && \\
&N_{\tau xy} \sim N_{\tau xy} + \Delta_\tau K_{xy} - \Delta_x\Delta_y K_\tau \ , &&
\end{aligned} \tag{3.25}$$

Here, $K_\tau, K_{xy}, \tilde{K}_\tau^{xy}, \tilde{K}$ are integers, and $\alpha, \tilde{\alpha}^{xy}$ are real. They are the classical gauge parameters of the classical background gauge fields $(A_\tau, A_{xy}; N_{\tau xy})$ and $(\tilde{A}_\tau^{xy}, \tilde{A}; \tilde{N}_\tau)$. The variation of the action under the gauge transformation is

$$-\frac{i}{2\pi}\sum_{\text{cube}}\tilde{\alpha}^{xy}(\Delta_\tau A_{xy}-\Delta_x\Delta_y A_\tau-2\pi N_{\tau xy})+i\sum_{xy\text{-plaq}}\tilde{K}^{xy}_\tau A_{xy}+i\sum_{\tau\text{-link}}\tilde{K}A_\tau\ . \tag{3.26}$$

It signals an anomaly because it cannot be cancelled by adding to the action any 2+1d local counterterms.

A convenient gauge choice

We now discuss a convenient gauge choice that sets most of the integer gauge fields to zero. We first integrate out ϕ^{xy}, which imposes the flatness condition on (n_τ, n_{xy}). We then gauge fix $n_\tau = 0$ and $n_{xy} = 0$ except for $n_\tau(L^\tau-1,\hat{x},\hat{y})$, $n_{xy}(\hat{\tau},\hat{x},L^y-1)$, and $n_{xy}(\hat{\tau},L^x-1,\hat{y})$. The remaining gauge-invariant information is in the holonomies:

$$\begin{aligned}
&n_\tau(L^\tau-1,\hat{x},\hat{y})=\bar{n}^x(\hat{x})+\bar{n}^y(\hat{y})\ ,\\
&n_{xy}(\hat{\tau},\hat{x},L^y-1)=\bar{n}^{xy}_x(\hat{x})\ ,\\
&n_{xy}(\hat{\tau},L^x-1,\hat{y})=\bar{n}^{xy}_y(\hat{y})\ ,
\end{aligned} \tag{3.27}$$

where $\bar{n}^i(\hat{x}^i)$ and $\bar{n}^{xy}_i(\hat{x}^i)$ are integer-valued. There is a gauge ambiguity in the zero modes of $\bar{n}^i(\hat{x}^i)$, while $\bar{n}^{xy}_i(\hat{x}^i)$ satisfy the constraint $\bar{n}^{xy}_x(L^x-1)=\bar{n}^{xy}_y(L^y-1)$. In total, there are $2L^x+2L^y-2$ independent integers that cannot be gauged away. The residual gauge symmetry is

$$\phi\sim\phi+2\pi w^x(\hat{x})+2\pi w^y(\hat{y})\ , \tag{3.28}$$

where $w^i(\hat{x}^i)$ is integer-valued.

Let us define a new field $\bar{\phi}$ on the sites such that in the fundamental domain

$$\bar{\phi}(\hat{\tau},\hat{x},\hat{y})=\phi(\hat{\tau},\hat{x},\hat{y})\ ,\qquad \text{for } 0\le\hat{x}^\mu<L^\mu\ , \tag{3.29}$$

and beyond the fundamental domain, it is extended via

$$
\begin{aligned}
\bar{\phi}(\hat{\tau}+L^\tau,\hat{x},\hat{y}) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\bar{n}^x(\hat{x}) - 2\pi\bar{n}^y(\hat{y}) \ , \\
\bar{\phi}(\hat{\tau},\hat{x}+L^x,\hat{y}) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\sum_{\hat{y}'=0}^{\hat{y}-1}\bar{n}^{xy}_y(\hat{y}') \ , \\
\bar{\phi}(\hat{\tau},\hat{x},\hat{y}+L^y) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\sum_{\hat{x}'=0}^{\hat{x}-1}\bar{n}^{xy}_x(\hat{x}') \ .
\end{aligned}
\tag{3.30}
$$

In particular, in the gauge (3.27), $\Delta_\tau\bar{\phi} = \Delta_\tau\phi - 2\pi n_\tau$, and $\Delta_x\Delta_y\bar{\phi} = \Delta_x\Delta_y\phi - 2\pi n_{xy}$. Although ϕ and (n_τ, n_{xy}) are single-valued, $\bar{\phi}$ can wind around the nontrivial cycles of spacetime. So, in the path integral, we should sum over nontrivial winding sectors of $\bar{\phi}$. The action (3.10) in terms of $\bar{\phi}$ is

$$
\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\bar{\phi})^2 + \frac{\beta}{2}\sum_{xy\text{-plaq}}(\Delta_x\Delta_y\bar{\phi})^2 \ . \tag{3.31}
$$

Let us discuss some charged configurations in the lattice model (3.31). We define the periodic Kronecker delta function

$$
\delta^P(\hat{x},\hat{x}_0,L^x) \equiv \sum_{I\in\mathbb{Z}}\delta_{\hat{x},\hat{x}_0-IL^x} \ , \tag{3.32}
$$

and a suitable step function $\Theta^P(\hat{x},\hat{x}_0,L^x)$ such that

$$
\Theta^P(0,\hat{x}_0,L^x) = 0 \ , \qquad \Delta_x\Theta^P(\hat{x},\hat{x}_0,L^x) = \delta^P(\hat{x},\hat{x}_0,L^x) \ . \tag{3.33}
$$

Note that this function is not periodic in $\hat{x}$. A minimal winding configuration is

$$
\bar{\phi}(\hat{\tau},\hat{x},\hat{y}) = 2\pi\left[\frac{\hat{x}}{L^x}\Theta^P(\hat{y},\hat{y}_0,L^y) + \frac{\hat{y}}{L^y}\Theta^P(\hat{x},\hat{x}_0,L^x) - \frac{\hat{x}\hat{y}}{L^xL^y}\right] \ . \tag{3.34}
$$

The most general winding configuration can be obtained by taking linear combinations with integer coefficients of (3.34) with different $\hat{x}_0,\hat{y}_0$ and adding to it a periodic function. The winding charges of (3.34) are $Q^{xy}_x(\hat{x}) = \delta^P(\hat{x},\hat{x}_0,L^x)$ and $Q^{xy}_y(\hat{y}) = \delta^P(\hat{y},\hat{y}_0,L^y)$. This configuration satisfies the equation of motion of $\bar{\phi}$, so it is a minimal action configuration with these winding charges. Its action is

$$\frac{\beta(2\pi)^2}{2} L^\tau \left(\frac{1}{L^x} + \frac{1}{L^y} - \frac{1}{L^x L^y} \right) . \tag{3.35}$$

Its Lorentzian interpretation is a winding state with energy

$$\frac{\beta(2\pi)^2}{2a} \left(\frac{1}{L^x} + \frac{1}{L^y} - \frac{1}{L^x L^y} \right) , \tag{3.36}$$

where a is the lattice spacing.

Continuum limit

In the continuum limit, we take $a \to 0$, $L^\mu \to \infty$ with fixed $\ell^\mu = aL^\mu$. In order for the limit to be nontrivial, we take the coupling constants to scale as $\beta_0 = \mu_0 a$ and $\beta = \frac{1}{\mu a}$. Then, the action becomes

$$\int d\tau dx dy \, \left[\frac{\mu_0}{2} (\partial_\tau \phi)^2 + \frac{1}{2\mu} (\partial_x \partial_y \phi)^2 \right] , \tag{3.37}$$

where we dropped the bar on ϕ. This is the Euclidean version of the 2+1d ϕ-theory of [11], which had been first introduced in [87]. (See also [95–99] for related discussions on this theory.)

The mixed 't Hooft anomaly between the momentum and winding symmetries can be seen by coupling the system to their background gauge fields (A_τ, A_{xy}) and $(\tilde{A}_\tau^{xy}, \tilde{A})$ respectively:

$$\int d\tau dx dy \, \left[\frac{\mu_0}{2} (\partial_\tau \phi - A_\tau)^2 + \frac{1}{2\mu} (\partial_x \partial_y \phi - A_{xy})^2 - \frac{i}{2\pi} \tilde{A}_\tau^{xy} (\partial_x \partial_y \phi - A_{xy}) - \frac{i}{2\pi} \tilde{A} (\partial_\tau \phi - A_\tau) \right] , \tag{3.38}$$

with gauge symmetry

$$\begin{aligned} &\phi \sim \phi + \alpha , && \phi^{xy} \sim \phi^{xy} + \tilde{\alpha}^{xy} , \\ &A_\tau \sim A_\tau + \partial_\tau \alpha , && \tilde{A}_\tau^{xy} \sim \tilde{A}_\tau^{xy} + \partial_\tau \tilde{\alpha}^{xy} , \\ &A_{xy} \sim A_{xy} + \partial_x \partial_y \alpha , && \tilde{A} \sim \tilde{A} + \partial_x \partial_y \tilde{\alpha}^{xy} . \end{aligned} \tag{3.39}$$

Here, $\alpha, \tilde{\alpha}^{xy}$ are the gauge parameters. The variation of the action under the gauge transformation is

$$-\frac{i}{2\pi}\int d\tau dx dy\ \tilde{\alpha}^{xy}(\partial_\tau A_{xy}-\partial_x\partial_y A_\tau)\ . \tag{3.40}$$

It signals an anomaly because it cannot be cancelled by adding to the action any 2+1d local counterterms. This is the continuum counterpart of the corresponding lattice expression (3.26).

We can also view the modified Villain lattice model (3.10), or its gauge fixed version (3.31), as a discretized version the continuum theory (3.37). Our analysis of this lattice model makes rigorous the various assertions in [11]. Let us discuss them in more detail.

Both the continuum theory (3.37) and the lattice theory (3.31) have real-valued fields and the periodicity in field space is implemented using the twisted boundary conditions (3.30).

One could question whether the lattice theory (3.31) with this particular sum over twisted boundary conditions is fully consistent. In the continuum, this was discussed in detail in [11,84]. On the lattice, the consistency follows from relating it to the lattice gauge theory (3.10) before the gauge fixing (3.27). Furthermore, the remaining gauge freedom (3.28) in the lattice theory (3.31) can now be interpreted as the gauge freedom of the continuum theory [11,84].

The discussion of [11] uncovered a number of surprising properties of the continuum theory (3.37), which are not present in the original microscopic theory (3.7). It has an emergent global dipole $U(1)$ winding symmetry and it is self dual. Now we see these properties already in the modified Villain lattice model (3.10). A reader who was skeptical about the continuum analysis of [11] can be reassured by seeing it derived on the lattice.

For fixed ℓ^τ and $\ell\sim\ell^x,\ell^y$, the action of the winding configuration (3.34) scales as $\ell^\tau/\mu\ell a$, which diverges as $1/a$ in the continuum limit. The configuration (3.34) gives a precise meaning to the winding configuration with infinite action in the continuum [11].[1] More generally, the classification of discontinuous configurations in the continuum theory (3.37) [11] is exactly as in the previous subsection.

In conclusion, the lattice model (3.10) flows in the continuum limit to (3.37). Conversely, the

[1]The discussion of such infinite action and infinite energy configurations was described in [11] as "ambitious." It is rigorous in the context of the modified Villain model.

lattice model (3.10), or its gauge fixed version (3.31), gives a rigorous setting for the discussion of the continuum theory (3.37) of [11].

3.2.2 A-theory ($U(1)$ tensor gauge theory)

We can gauge the $U(1)$ momentum dipole symmetry by coupling (3.10) to the $(\mathbf{1}_0, \mathbf{1}_2)$ tensor gauge fields (A_τ, A_{xy}). We will consider this system in Section 3.2.3, and restrict to the pure tensor gauge theory in this section. This pure gauge theory was discussed on the lattice and in the continuum in [11] (see also earlier work in [90, 96, 97, 100].

We place the $U(1)$ variables e^{iA_τ} and $e^{iA_{xy}}$ on τ-links and xy-plaquettes of the lattice respectively. The action for the pure $U(1)$ tensor gauge theory is

$$\gamma \sum_{\text{cube}} [1 - \cos(\Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau)] \ , \tag{3.41}$$

where A_τ and A_{xy} are circle-valued fields. It has the gauge symmetry

$$\begin{aligned} e^{iA_\tau} &\sim e^{iA_\tau + i\Delta_\tau \alpha} \ , \\ e^{iA_{xy}} &\sim e^{iA_{xy} + i\Delta_x \Delta_y \alpha} \ , \end{aligned} \tag{3.42}$$

with circle-valued α on the sites.

At large γ, we can approximate (3.41) by the Villain action

$$\frac{\gamma}{2} \sum_{\text{cube}} (\Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy})^2 \ , \tag{3.43}$$

where $n_{\tau xy}$ is an integer-valued field on the cubes. Now we view the gauge fields (A_τ, A_{xy}) and the gauge parameters α as real-valued, and the gauge symmetry (3.42) becomes

$$\begin{aligned} &A_\tau \sim A_\tau + \Delta_\tau \alpha + 2\pi k_\tau \ , \\ &A_{xy} \sim A_{xy} + \Delta_x \Delta_y \alpha + 2\pi k_{xy} \ , \\ &n_{\tau xy} \sim n_{\tau xy} + \Delta_\tau k_{xy} - \Delta_x \Delta_y k_\tau \ , \end{aligned} \tag{3.44}$$

where the gauge parameters k_τ and k_{xy} are integers on the τ-links and xy-plaquettes respectively.

We can interpret $n_{\tau xy}$ as the $\mathbb{Z}$ gauge field that makes (A_τ, A_{xy}) compact. In contrast to the XY-plaquette model, the $U(1)$ tensor gauge theory has no "vortices." So, we do not modify the Villain action (3.43) as in (3.10). Indeed, there is no local gauge-invariant field strength constructed out of the gauge field $n_{\tau xy}$.

We can also add a θ-term to the Villain action (3.43):

$$\frac{\gamma}{2} \sum_{\text{cube}} E_{xy}^2 + \frac{i\theta}{2\pi} \sum_{\text{cube}} E_{xy} \ , \tag{3.45}$$

where we defined the electric field

$$E_{xy} = \Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy} \ , \tag{3.46}$$

on the cubes. Since (A_τ, A_{xy}) is single-valued, we can write the θ-term as $-i\theta \sum_{\text{cube}} n_{\tau xy}$, which implies that the theta angle is 2π-periodic, i.e., $\theta \sim \theta + 2\pi$. Note that such a θ-term cannot be added in the original formulation (3.41), while it is straightforward and natural in the Villain version (3.43).

The quantized electric fluxes

$$\begin{aligned} &e^x(\hat{x}) = \sum_{\text{cube: fixed } \hat{x}} E_{xy} = -2\pi \sum_{\text{cube: fixed } \hat{x}} n_{\tau xy} \in 2\pi\mathbb{Z} \ , \\ &e^y(\hat{y}) = \sum_{\text{cube: fixed } \hat{y}} E_{xy} = -2\pi \sum_{\text{cube: fixed } \hat{y}} n_{\tau xy} \in 2\pi\mathbb{Z} \ , \end{aligned} \tag{3.47}$$

are associated with nontrivial holonomies of $n_{\tau xy}$ and they characterize the bundles of the tensor gauge theory. These fluxes satisfy the constraint

$$\sum_{\hat{x}} e^x(\hat{x}) = \sum_{\hat{y}} e^y(\hat{y}) = \sum_{\text{cube}} E_{xy} \ . \tag{3.48}$$

Global symmetries

The three models (3.41), (3.43), and (3.45) have an *electric tensor symmetry* that acts on the fields as

$$A_\tau \to A_\tau + \lambda_\tau \ , \qquad A_{xy} \to A_{xy} + \lambda_{xy} \ , \tag{3.49}$$

where $(\lambda_\tau, \lambda_{xy})$ is a flat, real-valued tensor gauge field (i.e., it has vanishing field strength).[2] Due to the integer-valued gauge symmetry with (k_τ, k_{xy}) (3.44), the electric tensor symmetry is $U(1)$, rather than $\mathbb{R}$. The Noether current of this electric symmetry follows from (3.45)

$$J_\tau^{xy} = -i\gamma E_{xy} + \frac{\theta}{2\pi} \ . \tag{3.50}$$

It satisfies the conservation equation and the differential condition (Gauss law)

$$\Delta_\tau J_\tau^{xy} = 0 \ , \qquad \Delta_x \Delta_y J_\tau^{xy} = 0 \ , \tag{3.51}$$

due to the equations of motion of A_{xy} and A_τ respectively. The conserved charge is

$$Q(\hat{x}, \hat{y}) = J_\tau^{xy} = Q^x(\hat{x}) + Q^y(\hat{y}) \ , \tag{3.52}$$

[2]Using the α gauge symmetry of (3.44), and the flatness of $(\lambda_\tau, \lambda_{xy})$, we can set $\lambda_\tau = c^x(\hat{x}) + c^y(\hat{y})$, and $\lambda_{xy} = c_x^{xy}(\hat{x}) + c_y^{xy}(\hat{y})$, where $c^i(\hat{x}^i)$ and $c_i^{xy}(\hat{x}^i)$ are real-valued.

where $Q^i(\hat{x}^i)$ is an integer, and the second equation follows from the Gauss law. The observables charged under the electric symmetry are the Wilson defect/operator

$$\begin{aligned} &W^\tau(\hat{x},\hat{y}) = \exp\left[i \sum_{\tau\text{-link: fixed } \hat{x},\hat{y}} A_\tau \right] \ , \\ &W^x(\hat{x},\mathcal{C}^x) = \exp\left[i \sum_{xy\text{-plaq}\in\mathcal{C}^x} A_{xy} + i \sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x A_\tau \right] \ , \end{aligned} \tag{3.53}$$

where $\mathcal{C}^x$ is a closed strip along the xy- and τx-plaquettes in the τy-plane at a fixed $\hat{x}$. Similarly, there is $W^y(\hat{y},\mathcal{C}^y)$.

Gauge-fixing and the continuum limit

Using the integer gauge freedom (3.44), we gauge fix $n_{\tau xy} = 0$, except for

$$n_{\tau xy}(L^\tau - 1, \hat{x}, L^y - 1) \equiv \bar{n}^x_{\tau xy}(\hat{x}) \ , \qquad n_{\tau xy}(L^\tau - 1, L^x - 1, \hat{y}) \equiv \bar{n}^y_{\tau xy}(\hat{y}) \ . \tag{3.54}$$

The integers $\bar{n}^i_{\tau xy}(\hat{x}^i)$ capture the only gauge-invariant information in $n_{\tau xy}$: its holonomies. They satisfy a constraint $\bar{n}^x_{\tau xy}(L^x - 1) = \bar{n}^y_{\tau xy}(L^y - 1)$. The residual gauge freedom is

$$\begin{aligned} &A_\tau \sim A_\tau + \Delta_\tau \alpha + 2\pi k_\tau \ , \\ &A_{xy} \sim A_{xy} + \Delta_x \Delta_y \alpha + 2\pi k_{xy} \ , \end{aligned} \tag{3.55}$$

where (k_τ, k_{xy}) is a flat, integer-valued tensor gauge field.

Similar to (3.30) in the ϕ-theory, we define a new tensor gauge field $(\bar{A}_\tau, \bar{A}_{xy})$ on the τ-links and xy-plaquettes such that

$$\Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau = \Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy} \ . \tag{3.56}$$

Although (A_τ, A_{xy}) and $n_{\tau xy}$ are single-valued, $(\bar{A}_\tau, \bar{A}_{xy})$ can have nontrivial monodromies around nontrivial cycles of the Euclidean spacetime. So, in the path integral, we should sum over nontrivial twisted sectors of $(\bar{A}_\tau, \bar{A}_{xy})$.

The action (3.45) in terms of $(\bar{A}_\tau, \bar{A}_{xy})$ is

$$\frac{\gamma}{2}\sum_{\text{cube}} \bar{E}_{xy}^2 + \frac{i\theta}{2\pi}\sum_{\text{cube}} \bar{E}_{xy} \ , \tag{3.57}$$

where we defined the new electric field

$$\bar{E}_{xy} = \Delta_\tau \bar{A}_{xy} - \Delta_x\Delta_y \bar{A}_\tau \ , \tag{3.58}$$

on the cubes.

In the continuum limit $a \to 0$, choosing the coupling to scale as $\gamma = \frac{2}{a^3 g_e^2}$ and the fields to scale as $\bar{A}_\tau = aA_\tau$ and $\bar{A}_{xy} = a^2 A_{xy}$,[3] the action becomes

$$\begin{aligned} &\int d\tau dx dy \ \left(\frac{1}{g_e^2}E_{xy}^2 + \frac{i\theta}{2\pi}E_{xy}\right) \ , \\ &E_{xy} = \partial_\tau A_{xy} - \partial_x\partial_y A_\tau \ . \end{aligned} \tag{3.59}$$

This is the Euclidean version of the continuum 2+1d A-theory of [11]. (See also [90,96,97,100].) The Villain model (3.45) has the same $U(1)$ electric symmetry as the continuum A-theory.

The spectrum of the lattice model consists of light states, whose action scales as a. In the continuum limit $a \to 0$ with fixed ℓ^τ, ℓ^x and ℓ^y, these light states become infinitely degenerate. The details can be found in [11].

We conclude that the lattice model (3.45) flows in the continuum limit to (3.59). Conversely, the lattice model (3.45), or its gauge fixed version (3.57), give a rigorous setting for the discussion of the continuum theory (3.59) of [11].

[3] The continuum tensor gauge fields (A_τ, A_{xy}) and their electric field defined here are not the same as the ones defined on the lattice at the beginning of this section. We hope this does not cause any confusion.

3.2.3 $\mathbb{Z}_N$ tensor gauge theory

In this subsection, we will consider the modified Villain lattice version of the 2+1d $\mathbb{Z}_N$ Ising plaquette model [89]. The modified Villain lattice model takes the form of a BF-type action, which admits two equivalent presentations. The first one, which we call the integer BF-action, uses only integer-valued fields, while the second one, which we call the real BF-action, uses both real and integer-valued fields. The real BF-action is naturally connected to the continuum $\mathbb{Z}_N$ tensor gauge theory of [11].

We can restrict the $U(1)$ variables in the $U(1)$ tensor gauge theory (3.41) to $\mathbb{Z}_N$ variables $e^{iA_\tau} = e^{\frac{2\pi i}{N} m_\tau}$ and $e^{iA_{xy}} = e^{\frac{2\pi i}{N} m_{xy}}$ with integers m_τ and m_{xy}. This leads to the $\mathbb{Z}_N$ tensor gauge theory with the action

$$\gamma \sum_{\text{cube}} \left[1 - \cos\left(\frac{2\pi}{N} (\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau) \right) \right] \ . \tag{3.60}$$

At large γ, $\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau = 0 \bmod N$ and we can replace the action by

$$\frac{2\pi i}{N} \sum_{\text{cube}} \tilde{m}^{xy} (\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau) \ , \tag{3.61}$$

where $\tilde{m}^{xy}$ is an integer-valued field on the cubes. We will refer to this presentation of the $\mathbb{Z}_N$ tensor gauge theory as the *integer BF-action*. This is analogous to the presentation (3.192) for the topological lattice $\mathbb{Z}_N$ gauge theory reviewed in Appendix C.

There is a gauge symmetry

$$\begin{aligned} m_\tau &\sim m_\tau + \Delta_\tau \ell + N k_\tau \ , \\ m_{xy} &\sim m_{xy} + \Delta_x \Delta_y \ell + N k_{xy} \ , \\ \tilde{m}^{xy} &\sim \tilde{m}^{xy} + N \tilde{k}^{xy} \ , \end{aligned} \tag{3.62}$$

where ℓ is an integer-valued field on the sites, k_τ and k_{xy} are integer-valued fields on the τ-links and xy-plaquettes respectively, and $\tilde{k}^{xy}$ is an integer-valued field on the cubes.

Global symmetries

In both models, (3.60) and (3.61), there is a $\mathbb{Z}_N$ *electric tensor symmetry*, which shifts (m_τ, m_{xy}) by a flat, integer-valued tensor gauge field. In the presentation of the model based on (3.61), the charge operator is

$$U(\hat\tau, \hat x, \hat y) = \exp\left[\frac{2\pi i}{N}\tilde m^{xy}\right] . \tag{3.63}$$

The observables charged under the electric symmetry are the Wilson defect/operator

$$\begin{aligned} W^\tau(\hat x, \hat y) &= \exp\left[\frac{2\pi i}{N}\sum_{\tau\text{-link: fixed }\hat x,\hat y} m_\tau\right] , \\ W^x(\hat x, \mathcal{C}^x) &= \exp\left[\frac{2\pi i}{N}\sum_{xy\text{-plaq}\in\mathcal{C}^x} m_{xy} + \frac{2\pi i}{N}\sum_{\tau x\text{-plaq}\in\mathcal{C}^x}\Delta_x m_\tau\right] , \end{aligned} \tag{3.64}$$

where $\mathcal{C}^x$ is a strip along the xy- and τx-plaquettes in the τy-plane at a fixed $\hat x$. Similarly, there is $W^y(\hat y, \mathcal{C}^y)$.

In the presentation of the model based on (3.61), but not in (3.60), there is also a $\mathbb{Z}_N$ magnetic dipole symmetry. The charge operators are $W^x(\hat x, \mathcal{C}^x)$ and $W^y(\hat y, \mathcal{C}^y)$, and the charged operator is $U(\hat\tau, \hat x, \hat y)$.

Ground state degeneracy

All the states of the model based on (3.61) are degenerate. The model has only ground states. Let us count them. First, we sum over the integer-valued fields m_τ and m_{xy}. They impose the following constraint on $\tilde m^{xy}$

$$\Delta_\tau \tilde m^{xy} = \Delta_x\Delta_y \tilde m^{xy} = 0 \bmod N . \tag{3.65}$$

The gauge inequivalent configurations of $\tilde m^{xy}$ are

$$\tilde m^{xy}(\hat\tau, \hat x, \hat y) = \tilde m^{xy}_x(\hat x) + \tilde m^{xy}_y(\hat y) , \tag{3.66}$$

where $\tilde{m}^{xy}_x(\hat{x})$ and $\tilde{m}^{xy}_y(\hat{y})$ are $\mathbb{Z}/N\mathbb{Z}$-valued. There is a gauge ambiguity in the zero modes of $\tilde{m}^{xy}_x(\hat{x})$ and $\tilde{m}^{xy}_y(\hat{y})$. So, in total, there are $N^{L^x+L^y-1}$ degenerate ground states.

Real BF-action and the continuum limit

The model based on the integer BF-action (3.61) has several different presentations. Here we discuss a presentation in terms of real-valued and integer-valued fields, which is closer to the continuum limit.

We start with the integer BF-action (3.61) and replace the integer-valued fields $\tilde{m}^{xy}$ and (m_τ, m_{xy}) with real-valued fields $\tilde{\phi}^{xy}$ and (A_τ, A_{xy}). In order to restrict these real-valued fields to be integer-valued, we add integer-valued Lagrange multiplier fields $n_{\tau xy}$ and $(\tilde{n}^{xy}_\tau, \tilde{n})$. Furthermore, since the gauge field (A_τ, A_{xy}) has real-valued gauge symmetry, we introduce a real-valued Stueckelberg field ϕ for that gauge symmetry. We end up with the action

$$\begin{aligned}&\frac{iN}{2\pi}\sum_{\text{cube}}\tilde{\phi}^{xy}(\Delta_\tau A_{xy} - \Delta_x\Delta_y A_\tau - 2\pi n_{\tau xy}) + iN\sum_{xy\text{-plaq}} A_{xy}\tilde{n}^{xy}_\tau \\ &\quad + iN\sum_{\tau\text{-link}} A_\tau \tilde{n} + i\sum_{\text{site}}\phi(\Delta_\tau\tilde{n} - \Delta_x\Delta_y\tilde{n}^{xy}_\tau) \ ,\end{aligned} \tag{3.67}$$

where ϕ, $\tilde{\phi}^{xy}$, A_τ and A_{xy} are real-valued fields on the sites, the dual site, the τ-links and the xy-plaquettes respectively, and $n_{\tau xy}$, $\tilde{n}^{xy}_\tau$ and $\tilde{n}$ are integer-valued fields on the cubes, the dual τ-links, and the dual xy-plaquettes, respectively.

There action (3.67) has the gauge symmetry

$$
\begin{aligned}
&\phi \sim \phi + N\alpha + 2\pi k \ , \\
&\tilde{\phi}^{xy} \sim \tilde{\phi}^{xy} + 2\pi \tilde{k}^{xy} \ , \\
&A_\tau \sim A_\tau + \Delta_\tau \alpha + 2\pi k_\tau \ , \\
&A_{xy} \sim A_{xy} + \Delta_x \Delta_y \alpha + 2\pi k_{xy} \ , \\
&\tilde{n}_\tau^{xy} \sim \tilde{n}_\tau^{xy} + \Delta_\tau \tilde{k}^{xy} \ , \\
&\tilde{n} \sim \tilde{n} + \Delta_x \Delta_y \tilde{k}^{xy} \ , \\
&n_{\tau xy} \sim n_{\tau xy} + \Delta_\tau k_{xy} - \Delta_x \Delta_y k_\tau \ .
\end{aligned}
\tag{3.68}
$$

As a check, summing over the integer-valued fields $n_{\tau xy}$, $\tilde{n}_\tau^{xy}$, and $\tilde{n}$ in (3.67) constrains

$$
\tilde{\phi}^{xy} = \frac{2\pi}{N} \tilde{m}^{xy}, \quad A_\tau - \frac{1}{N}\Delta_\tau \phi = \frac{2\pi}{N} m_\tau, \quad A_{xy} - \frac{1}{N}\Delta_x \Delta_y \phi = \frac{2\pi}{N} m_{xy} \ , \tag{3.69}
$$

where $\tilde{m}^{xy}$, m_τ and m_{xy} are integer-valued fields. Substituting them back into the action leads to (3.61).

We will refer to the presentation (3.67) of the $\mathbb{Z}_N$ tensor gauge theory as the *real BF-action*, which uses both real and integer fields. This is to be compared with the integer BF-action (3.60), which uses only integer-valued fields. These two presentations describe the same underlying lattice model, but use different sets of fields. In the real BF-action, the integer fields effectively make the real fields compact.

The real BF-action (3.67) can also be derived through Higgsing the $U(1)$ tensor gauge theory (3.45) to a $\mathbb{Z}_N$ theory using the field ϕ in (3.10). The Higgs action is

$$
\begin{aligned}
&\frac{i}{2\pi} \sum_{\tau\text{-link}} \tilde{B}(\Delta_\tau \phi - N A_\tau - 2\pi n_\tau) + \frac{i}{2\pi} \sum_{xy\text{-plaq}} \tilde{E}^{xy}(\Delta_x \Delta_y \phi - N A_{xy} - 2\pi n_{xy}) \\
&\quad - i \sum_{\text{cube}} \tilde{\phi}^{xy} \left(\Delta_\tau n_{xy} - \Delta_x \Delta_y n_\tau + N n_{\tau xy}\right) \ ,
\end{aligned}
\tag{3.70}
$$

where $\tilde{B}$ and $\tilde{E}^{xy}$ are real-valued fields on the τ-links and the xy-plaquette, which implement the Higgsing as constraints. In addition to the gauge symmetry (3.68), there is a gauge symmetry

$$\begin{aligned} n_\tau &\sim n_\tau + \Delta_\tau k - N k_\tau \ , \\ n_{xy} &\sim n_{xy} + \Delta_x \Delta_y k - N k_{xy} \ . \end{aligned} \tag{3.71}$$

Summing over the integer-valued fields n_τ and n_{xy} constrains

$$\tilde{B} - \Delta_x \Delta_y \tilde{\phi}^{xy} = -2\pi \tilde{n} \ , \quad \tilde{E}^{xy} - \Delta_\tau \tilde{\phi}^{xy} = -2\pi \tilde{n}_\tau^{xy} \ , \tag{3.72}$$

where $\tilde{n}$ and $\tilde{n}_\tau^{xy}$ are integer-valued fields. Substituting them back into the action leads to (3.67).

In a convenient gauge choice, most of the integer fields are fixed to be zero, while the remaining ones enter into the twisted boundary conditions of the real fields.

Let us make it more explicit. First, we integrate out ϕ, which imposes the constraint $\Delta_\tau \tilde{n} - \Delta_x \Delta_y \tilde{n}_\tau^{xy} = 0$. Then we can gauge fix $n_{\tau xy}$, $\tilde{n}_\tau^{xy}$ and $\tilde{n}$ to be zero almost everywhere except at

$$\begin{aligned} & n_{\tau xy}(L^\tau - 1, \hat{x}, L^y - 1) = \bar{n}_{\tau xy}^x(\hat{x}) \ , \\ & n_{\tau xy}(L^\tau - 1, L^x - 1, \hat{y}) = \bar{n}_{\tau xy}^y(\hat{y}) \ , \\ & \tilde{n}_\tau^{xy}(L^\tau - 1, \hat{x}, \hat{y}) = \bar{n}_{\tau,x}^{xy}(\hat{x}) + \bar{n}_{\tau,y}^{xy}(\hat{y}) \ , \\ & \tilde{n}(\hat{\tau}, \hat{x}, L^y - 1) = \bar{n}_{\tau x}(\hat{\tau}, \hat{x}) \ , \\ & \tilde{n}(\hat{\tau}, L^x - 1, \hat{y}) = \bar{n}_{\tau y}(\hat{\tau}, \hat{y}) \ , \end{aligned} \tag{3.73}$$

where $\bar{n}_{\tau xy}^x, \bar{n}_{\tau xy}^y, \bar{n}_{\tau,x}^{xy}, \bar{n}_{\tau,y}^{xy}, \bar{n}_{\tau x}, \bar{n}_{\tau y}$ are all integer-valued. These integers obey $\bar{n}_{\tau xy}^x(L^x - 1) = \bar{n}_{\tau xy}^y(L^y - 1)$ and $\bar{n}_{\tau x}(\hat{\tau}, L^x - 1) = \bar{n}_{\tau y}(\hat{\tau}, L^y - 1)$. The zero modes of $\bar{n}_{\tau,x}^{xy}(\hat{x})$ and $\bar{n}_{\tau,y}^{xy}(\hat{y})$ have a gauge ambiguity.

As in Sections 3.2.1 and 3.2.2, we define new fields $\bar{\phi}^{xy}$, $\bar{A}_\tau$ and $\bar{A}_{xy}$ on the sites, the τ-links, and the xy-plaquettes such that

$$\begin{aligned}
&\Delta_\tau \bar{\phi}^{xy} = \Delta_\tau \tilde{\phi}^{xy} - 2\pi \tilde{n}_\tau^{xy} \ , \\
&\Delta_x \Delta_y \bar{\phi}^{xy} = \Delta_x \Delta_y \tilde{\phi}^{xy} - 2\pi \tilde{n} \ , \\
&\Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau = \Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy} \ .
\end{aligned} \tag{3.74}$$

In contrast to the original variables that are single-valued, the new variables can have nontrivial twisted boundary conditions around the nontrivial cycles of space-time. So, in the path integral, we should sum over nontrivial twisted sectors of $\bar{\phi}^{xy}$ and $(\bar{A}_\tau, \bar{A}_{xy})$.

In terms of the new variables, the action (3.67) becomes

$$\begin{aligned}
&\frac{iN}{2\pi} \sum_{\text{cube}} \bar{\phi}^{xy} (\Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau) \\
&+ iN \sum_{\substack{xy\text{-plaq} \\ \hat{\tau} = L^\tau - 1}} \bar{A}_{xy} (\bar{n}_{\tau,x}^{xy} + \bar{n}_{\tau,y}^{xy}) \quad + iN \sum_{\substack{\tau\text{-link} \\ \hat{x} = L^x - 1}} \bar{A}_\tau \bar{n}_{\tau y} + iN \sum_{\substack{\tau\text{-link} \\ \hat{y} = L^y - 1}} \bar{A}_\tau \bar{n}_{\tau x} \ ,
\end{aligned} \tag{3.75}$$

The real BF-action of our modified Villain model is closely related to the continuum field theory. In the continuum limit, $a \to 0$, the action becomes

$$\frac{iN}{2\pi} \int d\tau dx dy \ \phi^{xy} (\partial_\tau A_{xy} - \partial_x \partial_y A_\tau) \ , \tag{3.76}$$

where we dropped the bars over the variables and rescaled them by appropriate powers of the lattice spacing a. We also omitted the boundary terms that depend on the transition functions of ϕ^{xy} and (A_τ, A_{xy}).[4] This is the Euclidean version of the 2+1d $\mathbb{Z}_N$ tensor gauge theory of [11].

We conclude that the lattice model (3.61), or equivalently (3.67), flows in the continuum limit to (3.76). Conversely, the lattice model (3.61), or equivalently (3.67), gives a rigorous setting for the discussion of the continuum theory (3.76) of [11].

[4]Such boundary terms are necessary in order to make the continuum action (3.76) well-defined. They played a crucial role in the analysis of [84].

3.3 3+1d (4d Euclidean) exotic theories with cubic symmetry

In this section, we will describe the modified Villain formulation of the exotic 3+1d continuum theories of [80, 81]. All the models are placed on a periodic 4d Euclidean lattice with lattice spacing a, and L^μ sites in the μ direction. We label the sites by integers $\hat{x}^\mu \sim \hat{x}^\mu + L^\mu$.

Since the spatial lattice has an S_4 rotation symmetry, we can organize the fields according to S_4 representations: the trivial representation $\mathbf{1}$, the sign representation $\mathbf{1}'$, a two-dimensional irreducible representation $\mathbf{2}$, the standard representation $\mathbf{3}$ and another three-dimensional irreducible representation $\mathbf{3}'$.

We will label the components of S_4 representations using $SO(3)$ vector indices i, j, k. In this section, the indices i, j, k in every expression are never equal, $i \neq j \neq k$.

We label the components of an object V in $\mathbf{3}$ of S_4 as V_i and the components of an object E in $\mathbf{3}'$ of S_4 as $E_{ij} = E_{ji}$. The labeling of the components of T in $\mathbf{2}$ of S_4 is slightly more complicated. We can label them as $T^{[ij]k} = -T^{[ji]k}$, with an identification under simultaneous shifts of $T^{[xy]z}$, $T^{[yz]x}$, $T^{[zx]y}$ by the same amount. Alternatively, we can define the combinations $T^{k(ij)} = T^{[ki]j} - T^{[jk]i}$, which are not subject to the identification. In this presentation, we have a constraint $T^{x(yz)} + T^{y(zx)} + T^{z(xy)} = 0$. We will also use $T_{k(ij)} = T_{k(ji)}$ with lower indices to label the components of $\mathbf{2}$. It has an identification under simultaneous shifts of $T_{x(yz)}$, $T_{y(zx)}$, $T_{z(xy)}$ by the same amount. Similarly, we define the combinations $T_{[ij]k} = T_{i(jk)} - T_{j(ik)}$, which are not subject to an identification, but obey the constraint $T_{[xy]z} + T_{[yz]x} + T_{[zx]y} = 0$.

3.3.1 ϕ-theory

There is a $U(1)$ variable $e^{i\phi}$ at each site of the lattice. The action is

$$\beta_0 \sum_{\tau\text{-link}} [1 - \cos(\Delta_\tau \phi)] + \beta \sum_{i<j} \sum_{ij\text{-plaq}} [1 - \cos(\Delta_i \Delta_j \phi)] \ , \tag{3.77}$$

where ϕ is circle-valued. At large β_0, β, we can approximate the action with the Villain action

$$\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\phi-2\pi n_\tau)^2+\frac{\beta}{2}\sum_{i<j}\sum_{ij\text{-plaq}}(\Delta_i\Delta_j\phi-2\pi n_{ij})^2 \ , \tag{3.78}$$

where ϕ is real and n_τ and n_{ij} are integer-valued fields on τ-links and ij-plaquettes, respectively. There is an integer gauge symmetry

$$\phi\sim\phi+2\pi p \ , \qquad n_\tau\sim n_\tau+\Delta_\tau p \ , \qquad n_{ij}\sim n_{ij}+\Delta_i\Delta_j p \ , \tag{3.79}$$

where p is an integer-valued gauge parameter on the sites. We can interpret (n_τ, n_{ij}) as $\mathbb{Z}$ tensor gauge fields that make ϕ compact.

Next, we suppress the "vortices" by modifying the Villain action as

$$\begin{aligned}&\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\phi-2\pi n_\tau)^2+\frac{\beta}{2}\sum_{i<j}\sum_{ij\text{-plaq}}(\Delta_i\Delta_j\phi-2\pi n_{ij})^2\\&+i\sum_{i<j}\sum_{\tau ij\text{-cube}}\hat{A}^{ij}(\Delta_\tau n_{ij}-\Delta_i\Delta_j n_\tau)-i\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{xyz\text{-cube}}\hat{A}_\tau^{[ij]k}(\Delta_i n_{jk}-\Delta_j n_{ik}) \ ,\end{aligned} \tag{3.80}$$

where $\hat{A}_\tau^{[ij]k}$ and $\hat{A}^{ij}$ are real-valued fields on dual τ-links and dual k-links respectively. They are Lagrange multipliers that impose the flatness constraint of (n_τ, n_{ij}). They have their own gauge symmetry

$$\begin{aligned}&\hat{A}_\tau^{[ij]k}\sim\hat{A}_\tau^{[ij]k}+\Delta_\tau\hat{\alpha}^{[ij]k}+2\pi\hat{q}_\tau^{[ij]k} \ ,\\&\hat{A}^{ij}\sim\hat{A}^{ij}+\Delta_k\hat{\alpha}^{k(ij)}+2\pi\hat{q}^{ij} \ .\end{aligned} \tag{3.81}$$

Here $\hat{\alpha}^{[ij]k}$ are real-valued fields on the dual sites, while $\hat{q}_\tau^{[ij]k}$ and $\hat{q}^{ij}$ are integers on the dual τ-links and the dual k-links, respectively.

Following similar steps in Section 3.2.1, we can integrate out the real fields $\hat{A}_\tau^{[ij]k}, \hat{A}^{ij}$ and gauge fix most of the integer fields to be zero. In this gauge choice, the continuum limit of this modified Villain model is recognized as the 3+1d ϕ-theory of [80]. See also [73, 101, 92, 102] for related discussions on this theory. Moreover, the modified Villain model has a $U(1)$ $(\mathbf{1},\mathbf{3}')$

momentum dipole symmetry and a $U(1)$ $(\mathbf{3}', \mathbf{1})$ winding dipole symmetry, which are the same as in the continuum 3+1d ϕ-theory.

Alternatively, we can apply the Poisson resummation formula (3.6) to dualize the modified Villain action (3.80) to

$$\begin{aligned} &\frac{1}{2(2\pi)^2\beta} \sum_{\substack{\text{cyclic} \\ i,j,k}} \sum_{\text{dual } \tau k\text{-plaq}} (\Delta_\tau \hat{A}^{ij} - \Delta_k \hat{A}_\tau^{k(ij)} - 2\pi \hat{n}_\tau^{ij})^2 \\ &+ \frac{1}{2(2\pi)^2\beta_0} \sum_{\text{dual } xyz\text{-cube}} \left(\sum_{i<j} \Delta_i \Delta_j \hat{A}^{ij} - 2\pi \hat{n} \right)^2 - i \sum_{\text{site}} \phi \left(\Delta_\tau \hat{n} - \sum_{i<j} \Delta_i \Delta_j \hat{n}_\tau^{ij} \right) , \end{aligned} \tag{3.82}$$

where $\hat{n}_\tau^{ij}$ and $\hat{n}$ are integer-valued fields on the dual τk-plaquettes (or ij-plaquettes) and the dual hypercubes (or sites) respectively. We interpret $(\hat{n}_\tau^{ij}, \hat{n})$ as $\mathbb{Z}$ gauge fields that make $(\hat{A}_\tau^{k(ij)}, \hat{A}^{ij})$ compact via the gauge symmetry[5]

$$\begin{aligned} &\hat{A}_\tau^{k(ij)} \sim \hat{A}_\tau^{k(ij)} + \Delta_\tau \hat{\alpha}^{k(ij)} + 2\pi \hat{q}_\tau^{k(ij)} , \\ &\hat{A}^{ij} \sim \hat{A}^{ij} + \Delta_k \hat{\alpha}^{k(ij)} + 2\pi \hat{q}^{ij} , \\ &\hat{n}_\tau^{ij} \sim \hat{n}_\tau^{ij} + \Delta_\tau \hat{q}^{ij} - \Delta_k \hat{q}_\tau^{k(ij)} , \\ &\hat{n} \sim \hat{n} + \sum_{i<j} \Delta_i \Delta_j \hat{q}^{ij} . \end{aligned} \tag{3.83}$$

The Lagrange multiplier ϕ imposes the flatness constraint of $(\hat{n}_\tau^{ij}, \hat{n})$.

Once again, following similar steps in Section 3.2.1, we can integrate out the real field ϕ and gauge fix most of the integer fields to be zero. In this gauge choice, the continuum limit of this modified Villain model is recognized as the 3+1d $\hat{A}$-theory of [80] (see also [73, 92]). Moreover, the modified Villain model has a $U(1)$ $(\mathbf{3}', \mathbf{1})$ electric dipole symmetry and a $U(1)$ $(\mathbf{1}, \mathbf{3}')$ magnetic dipole symmetry, which are the same as in the continuum 3+1d $\hat{A}$-theory. The duality maps the momentum (winding) dipole symmetry of ϕ-theory to the magnetic (electric) dipole symmetry of the $\hat{A}$-theory, exactly like in the continuum theories.

[5] $(\hat{n}_\tau^{ij}, \hat{n})$ is the $\mathbb{Z}$ version of $(\hat{C}_\tau^{ij}, \hat{C})$ of [82].

In conclusion, the modified Villain action (3.80) has the same continuum limit as the XY-plaquette action (3.77). It has all the properties of the continuum ϕ-theory of [80] including the emergent winding symmetry and the duality to the $\hat{A}$-theory. It is straightforward to check that the analysis of the singular configurations and the spectrum of charged states of the continuum theory are regularized properly by this modified Villain lattice action.

3.3.2 A-theory

There are $U(1)$ variables e^{iA_τ} and $e^{iA_{ij}}$ on the τ-links and the ij-plaquettes of the lattice, respectively. The action is

$$\gamma_0 \sum_{i<j} \sum_{\tau ij\text{-cube}} [1 - \cos(\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau)] + \gamma \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic}\\ i,j,k}} [1 - \cos(\Delta_i A_{jk} - \Delta_j A_{ik})] \ , \tag{3.84}$$

where (A_τ, A_{ij}) are circle-valued. This action has a tensor gauge symmetry

$$\begin{aligned} e^{iA_\tau} &\sim e^{iA_\tau + i\Delta_\tau \alpha} \ , \\ e^{iA_{ij}} &\sim e^{iA_{ij} + i\Delta_i \Delta_j \alpha} \ , \end{aligned} \tag{3.85}$$

with circle valued α at the sites.

At large γ_0, γ, we can approximate the action, à la Villain, as

$$\frac{\gamma_0}{2} \sum_{i<j} \sum_{\tau ij\text{-cube}} (\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau - 2\pi n_{\tau ij})^2 + \frac{\gamma}{2} \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic}\\ i,j,k}} (\Delta_i A_{jk} - \Delta_j A_{ik} - 2\pi n_{[ij]k})^2 \ , \tag{3.86}$$

where now (A_τ, A_{ij}) are real and $n_{\tau ij}$ and $n_{[ij]k}$ are integer-valued fields on the τij-cubes and the xyz-cubes respectively. The gauge symmetry (3.85) is now replaced with

$$\begin{aligned} & A_\tau \sim A_\tau + \Delta_\tau \alpha + 2\pi q_\tau \ , \\ & A_{ij} \sim A_{ij} + \Delta_i \Delta_j \alpha + 2\pi q_{ij} \ , \\ & n_{\tau ij} \sim n_{\tau ij} + \Delta_\tau q_{ij} - \Delta_i \Delta_j q_\tau \ , \\ & n_{[ij]k} \sim n_{[ij]k} + \Delta_i q_{jk} - \Delta_j q_{ik} \ . \end{aligned} \tag{3.87}$$

Here α is a real-valued field on the sites, while q_τ and q_{ij} are integer-valued fields on the τ-links and the ij-plaquettes, respectively. We interpret $(n_{\tau ij}, n_{[ij]k})$ as the $\mathbb{Z}$ gauge fields that make (A_τ, A_{ij}) compact.[6]

Next, we suppress the "vortices" by modifying the Villain action as

$$\begin{aligned} & \frac{\gamma_0}{2} \sum_{i<j} \sum_{\tau ij\text{-cube}} (\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau - 2\pi n_{\tau ij})^2 + \frac{\gamma}{2} \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic} \\ i,j,k}} (\Delta_i A_{jk} - \Delta_j A_{ik} - 2\pi n_{[ij]k})^2 \\ & \quad + i \sum_{\text{dual site}} \sum_{\substack{\text{cyclic} \\ i,j,k}} \hat{\phi}^{[ij]k} (\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik}) \ , \end{aligned} \tag{3.88}$$

where $\hat{\phi}^{[ij]k}$ is a real-valued field on the dual sites of the lattice. It is a Lagrange multiplier that imposes the flatness constraint of $(n_{\tau ij}, n_{[ij]k})$, and it has a gauge symmetry

$$\hat{\phi}^{[ij]k} \sim \hat{\phi}^{[ij]k} + 2\pi \hat{p}^{[ij]k} \ , \tag{3.89}$$

where $\hat{p}^{[ij]k}$ is an integer-valued field on the dual sites.

Following similar steps in Section 3.2.1, we can integrate out the real fields $\hat{\phi}^{[ij]k}$ and gauge fix most of the integer fields to be zero. In this gauge choice, the continuum limit of this modified Villain model is recognized as the 3+1d A-theory of [80]. See also [88,73,90,91,101,92] for related discussions on this theory. Moreover, the modified Villain model has a $U(1)$ $(\mathbf{3}', \mathbf{2})$ electric tensor symmetry and a $U(1)$ $(\mathbf{2}, \mathbf{3}')$ magnetic tensor symmetry, which are the same as in the continuum

[6] $(n_{\tau ij}, n_{[ij]k})$ is the $\mathbb{Z}$ version of $(C_\tau^{ij}, C^{[ij]k})$ of [82].

3+1d A-theory.

Alternatively, we can apply the Poisson resummation formula (3.6) to dualize the modified Villain action (3.88) to

$$\begin{aligned}&\frac{1}{6(2\pi)^2\gamma}\sum_{\text{dual }\tau\text{-link}}\sum_{\substack{\text{cyclic}\\ i,j,k}}(\Delta_\tau\hat{\phi}^{k(ij)}-2\pi\hat{n}_\tau^{k(ij)})^2+\frac{1}{2(2\pi)^2\gamma_0}\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{\text{dual }k\text{-link}}(\Delta_k\hat{\phi}^{k(ij)}-2\pi\hat{n}^{ij})^2\\&\quad+i\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{ij\text{-plaq}}A_{ij}(\Delta_\tau\hat{n}^{ij}-\Delta_k\hat{n}_\tau^{k(ij)})+i\sum_{\tau\text{-link}}A_\tau\sum_{i<j}\Delta_i\Delta_j\hat{n}^{ij}\ ,\end{aligned}\tag{3.90}$$

where $\hat{n}_\tau^{k(ij)}$ and $\hat{n}^{ij}$ are integer-valued fields on the dual τ-links and the dual k-links respectively. There is a gauge symmetry

$$\begin{aligned}&\hat{\phi}^{k(ij)}\sim\hat{\phi}^{k(ij)}+2\pi\hat{p}^{k(ij)}\ ,\\&\hat{n}_\tau^{k(ij)}\sim\hat{n}_\tau^{k(ij)}+\Delta_\tau\hat{p}^{k(ij)}\ ,\\&\hat{n}^{ij}\sim\hat{n}^{ij}+\Delta_k\hat{p}^{k(ij)}\ .\end{aligned}\tag{3.91}$$

We interpret $(\hat{n}_\tau^{k(ij)},\hat{n}^{ij})$ as $\mathbb{Z}$ gauge fields that make $\hat{\phi}^{k(ij)}$ compact. The Lagrange multipliers (A_τ,A_{ij}) impose the flatness constraint of $(\hat{n}_\tau^{k(ij)},\hat{n}^{ij})$. The dual action (3.90) is the modified Villain action of the $\hat{\phi}$-theory of [80].

Once again, following similar steps in Section 3.2.1, we can integrate out the real fields (A_τ,A_{ij}) and gauge fix most of the integer fields to be zero. In this gauge choice, the continuum limit of this modified Villain model is recognized as the 3+1d $\hat{\phi}$-theory of [80]. Moreover, the modified Villain model has a $U(1)$ $(\mathbf{2},\mathbf{3}')$ momentum tensor symmetry and a $U(1)$ $(\mathbf{3}',\mathbf{2})$ winding tensor symmetry, which are the same as in the continuum 3+1d $\hat{\phi}$-theory. The duality maps the electric (magnetic) tensor symmetry of the A-theory to the winding (momentum) tensor symmetry of $\hat{\phi}$-theory, exactly like in the continuum theories.

To summarize, the lattice A-theory (3.84) and the modified Villain action (3.88) flow to the same continuum theory – the continuum A-theory. The modified Villain action has all the features of the continuum theory. It has a magnetic symmetry and it is dual to the $\hat{\phi}$ theory. It gives a rigorous presentation of the analysis of singular field configurations and the spectrum of charged states found in [80].

3.3.3 X-cube model

In this subsection, we will start with the X-cube model in its Hamiltonian formalism and deform it to a modified Villain lattice model. The latter takes the form of a BF-type action, which admits two equivalent presentations. The first one, which we call the integer BF-action, uses only the integer fields, while the second one, which we call the real BF-action, uses both real and integer fields. The real BF-action is naturally connected to the continuum $\mathbb{Z}_N$ tensor gauge theory of [73, 81].

Review of the Hamiltonian formulation

We start with the Hamiltonian formulation of the X-cube model. On a periodic 3d lattice, there is a $\mathbb{Z}_N$ variable U and its conjugate variable V on each link. They obey $UV = e^{2\pi i/N}VU$. We label the sites by integers $\hat{s} = (\hat{x}, \hat{y}, \hat{z})$ and label the links, the plaquette and the cubes using the coordinates of their centers. The Hamiltonian of the X-cube model is [17]

$$\begin{aligned}
&H = -\beta_1 \sum_{\text{site}} (G_{\hat{s},[yz]x} + G_{\hat{s},[zx]y} + G_{\hat{s},[xy]z}) - \beta_2 \sum_{\text{cube}} L_{\hat{c}} + c.c. \ , \\
&G_{\hat{s},[yz]x} = V_{\hat{s}+(0,\frac{1}{2},0)} V^\dagger_{\hat{s}+(0,0,\frac{1}{2})} V^\dagger_{\hat{s}-(0,\frac{1}{2},0)} V_{\hat{s}-(0,0,\frac{1}{2})} \ , \\
&G_{\hat{s},[zx]y} = V^\dagger_{\hat{s}+(\frac{1}{2},0,0)} V_{\hat{s}+(0,0,\frac{1}{2})} V_{\hat{s}-(\frac{1}{2},0,0)} V^\dagger_{\hat{s}-(0,0,\frac{1}{2})} \ , \\
&G_{\hat{s},[xy]z} = V_{\hat{s}+(\frac{1}{2},0,0)} V^\dagger_{\hat{s}+(0,\frac{1}{2},0)} V^\dagger_{\hat{s}-(\frac{1}{2},0,0)} V_{\hat{s}-(0,\frac{1}{2},0)} \ , \\
&L_{\hat{c}} = U_{\hat{c}+(\frac{1}{2},\frac{1}{2},0)} U^\dagger_{\hat{c}+(-\frac{1}{2},\frac{1}{2},0)} U^\dagger_{\hat{c}+(\frac{1}{2},-\frac{1}{2},0)} U_{\hat{c}-(\frac{1}{2},\frac{1}{2},0)} \\
&\qquad U_{\hat{c}+(0,\frac{1}{2},\frac{1}{2})} U^\dagger_{\hat{c}+(0,-\frac{1}{2},\frac{1}{2})} U^\dagger_{\hat{c}+(0,\frac{1}{2},-\frac{1}{2})} U_{\hat{c}-(0,\frac{1}{2},\frac{1}{2})} \\
&\qquad U_{\hat{c}+(\frac{1}{2},0,\frac{1}{2})} U^\dagger_{\hat{c}+(-\frac{1}{2},0,\frac{1}{2})} U^\dagger_{\hat{c}+(\frac{1}{2},0,-\frac{1}{2})} U_{\hat{c}-(\frac{1}{2},0,\frac{1}{2})} \ .
\end{aligned} \tag{3.92}$$

All the terms in the Hamiltonian commute with each other. The operators $G_{\hat{s},[ij]k}$ are in the **2** of S_4 and satisfy $G_{\hat{s},[yz]x} G_{\hat{s},[zx]y} G_{\hat{s},[xy]z} = 1$.

The ground states satisfy $G_{\hat{s},[ij]k} = L_{\hat{c}} = 1$ for all $\hat{s}, \hat{c}$. There are dynamical excitations that violate only $L_{\hat{c}} = 1$ at a cube. Such excitations cannot move so they are fractons. There are also dynamical excitations that violate only $G_{\hat{s},[yz]x} = G_{\hat{s},[zx]y} = 1$ at a site. Such excitations can only move along the z direction so they are z-lineons. Similarly, there are x-lineons and y-lineons that can only move along the x and y direction, respectively. Because of the relation $G_{\hat{s},[yz]x} G_{\hat{s},[zx]y} G_{\hat{s},[xy]z} = 1$, an x-lineon, a y-lineon and a z-lineon can annihilate to the vacuum when they meet at the same point.

The X-cube model has a faithful $\mathbb{Z}_N$ $(\mathbf{3}', \mathbf{2})$ tensor symmetry and a faithful $\mathbb{Z}_N$ $(\mathbf{3}', \mathbf{1})$ dipole symmetry.[7] A typical symmetry operator of the faithful $\mathbb{Z}_N$ $(\mathbf{3}', \mathbf{2})$ tensor symmetry is the line operator $\prod_{z\text{-link: fixed } \hat{y},\hat{z}} U$. And there are similar lines along other directions. A typical symmetry operator of the faithful $\mathbb{Z}_N$ $(\mathbf{3}', \mathbf{1})$ dipole symmetry is $\prod_{\mathcal{C}^{xy}} V$ where $\mathcal{C}^{xy}$ is a closed curve along the dual links at fixed $\hat{z}_0$. Similarly, there are other symmetry operators on the other planes.

We are interested in the $\beta_1, \beta_2 \to \infty$ limit of the model. In this limit, $G_{\hat{s},[ij]k} = L_{\hat{c}} = 1$ for all $\hat{s}, \hat{c}$ and the Hilbert space is restricted to the ground states.

Integer BF-action

We now formulate the X-cube model in the $\beta_1, \beta_2 \to \infty$ limit in the Lagrangian formalism. We put the model on a periodic 4d Euclidean lattice. For each k-link, we introduce an integer-valued field $\hat{m}^{ij}$ with $i \neq j \neq k$ for the $\mathbb{Z}_N$ variable $U = \exp(\frac{2\pi i \hat{m}^{ij}}{N})$. For each dual ij-plaquette, we introduce an integer-valued field m_{ij} for the conjugate $\mathbb{Z}_N$ variable $V = \exp(\frac{2\pi i m_{ij}}{N})$.

Next, we introduce Lagrange multiplier fields to impose the constraints $G_{\hat{s},[ij]k} = L_{\hat{c}} = 1$. On each dual τ-link (or xyz-cube), we introduce an integer-valued field m_τ to impose $L_{\hat{c}} = 1$ as a constraint. On each τ-link, we introduce three integer-valued fields $\hat{m}_\tau^{[ij]k}$ to impose $G_{\hat{s},[ij]k} = 1$ as constraints. Since $G_{\hat{s},[yz]x} G_{\hat{s},[zx]y} G_{\hat{s},[xy]z} = 1$, one combination of $\hat{m}_\tau^{[ij]k}$ decouples

[7] To clarify the terminology, recall that each symmetry operator is associated with a geometrical object $\mathcal{C}$. According to [103], if the action of the operator depends only on the topology of $\mathcal{C}$, the symmetry is not faithful, while if it depends also on its geometry, the symmetry is faithful. For example, the non-relativistic q-form symmetry of [79] is faithful, while the relativistic q-form symmetry of [5] is not faithful. In [81], the faithful symmetry was referred to as "unconstrained" and the unfaithful symmetry was referred to as "constrained."

and therefore $\hat{m}_\tau^{[ij]k}$ has a gauge symmetry. Below, we will instead work with the combinations $\hat{m}_\tau^{k(ij)} = \hat{m}_\tau^{[ki]j} - \hat{m}_\tau^{[jk]i}$, which are not subject to any gauge symmetry, but are constrained to satisfy $\hat{m}_\tau^{x(yz)} + \hat{m}_\tau^{y(zx)} + \hat{m}_\tau^{z(xy)} = 0$.

In terms of these integer fields, the Euclidean lattice action for the low-energy limit of the X-cube model is

$$\frac{2\pi i}{N} \sum_{\substack{\text{cyclic} \\ i,j,k}} \sum_{\tau k\text{-plaq}} m_{ij} \left(\Delta_\tau \hat{m}^{ij} - \Delta_k \hat{m}_\tau^{k(ij)} \right) + \frac{2\pi i}{N} \sum_{xyz\text{-cube}} m_\tau \left(\sum_{i<j} \Delta_i \Delta_j \hat{m}^{ij} \right) . \tag{3.93}$$

There are gauge symmetries:

$$\begin{aligned} & m_\tau \sim m_\tau + \Delta_\tau \ell + N q_\tau \ , \\ & m_{ij} \sim m_{ij} + \Delta_i \Delta_j \ell + N q_{ij} \ , \\ & \hat{m}_\tau^{k(ij)} \sim \hat{m}_\tau^{k(ij)} + \Delta_\tau \hat{\ell}^{k(ij)} + N \hat{q}_\tau^{k(ij)} \ , \\ & \hat{m}^{ij} \sim \hat{m}^{ij} + \Delta_k \hat{\ell}^{k(ij)} + N \hat{q}^{ij} \ , \end{aligned} \tag{3.94}$$

where ℓ, $\hat{\ell}^{k(ij)}$, q_τ, q_{ij}, $\hat{q}_\tau^{k(ij)}$ and $\hat{q}^{ij}$ are integer-valued fields on the dual sites, the sites, the dual τ-links, the dual ij-plaquettes, the τ-links, and the k-links, respectively. We will refer to this presentation of the model as the integer BF-action. This is analogous to the presentation (3.192) for the topological lattice $\mathbb{Z}_N$ gauge theory reviewed in Appendix C and the presentation (3.61) of the 2+1d tensor $\mathbb{Z}_N$ tensor gauge theory.

The fields $(\hat{m}_\tau^{k(ij)}, \hat{m}^{ij})$ and (m_τ, m_{ij}) pair up into two integer-valued tensor gauge fields. Comparing with (3.82) and (3.88), we can interpret (3.93) as the $\mathbb{Z}_N$ lattice tensor gauge theory of the $\hat{A}$ gauge field or the A gauge field.

In this Lagrangian, there are no dynamical fractons and lineons. Instead, charged particles become defects of probe fractons and lineons. The probe fracton defect is

$$W^\tau(\hat{x}, \hat{y}, \hat{z}) = \exp \left[\frac{2\pi i}{N} \sum_{\text{dual } \tau\text{-link: fixed } \hat{x},\hat{y},\hat{z}} m_\tau \right] \ , \tag{3.95}$$

and the probe z-lineon defect is

$$\hat{W}^z(\hat{x},\hat{y},\mathcal{C}^z) = \exp\left[\frac{2\pi i}{N}\sum_{\tau\text{-link}\in\mathcal{C}^z}\hat{m}_\tau^{z(xy)} + \frac{2\pi i}{N}\sum_{z\text{-link}\in\mathcal{C}^z}\hat{m}^{xy}\right]~, \tag{3.96}$$

where $\mathcal{C}^z$ is a curve along the τ- and z-links in the τz-plane at fixed $\hat{x}$ and $\hat{y}$. The x- and y-lineons are defined similarly.

The $\mathbb{Z}_N$ lattice tensor gauge theory has a $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{2})$ tensor symmetry and a $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{1})$ dipole symmetry. The $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{2})$ tensor symmetry is generated by the line operator of (3.96) along a closed curve $\mathcal{C}^z$ and other similar line operators on the τx- and τy-plane. These symmetry operators are constrained by the flatness condition on $\hat{m}^{ij}$. So, the $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{2})$ tensor symmetry is unfaithful (in the sense of [103]). The charged observables are the probe fracton defect (3.95) and the Wilson observable

$$W^{xy}(\hat{z},\mathcal{C}^{xy}) = \exp\left[\frac{2\pi i}{N}\left(\sum_{\text{dual } xz\text{-plaq}\in\mathcal{C}^{xy}} m_{xz} + \sum_{\text{dual } yz\text{-plaq}\in\mathcal{C}^{xy}} m_{yz} + \sum_{\text{dual } \tau z\text{-plaq}\in\mathcal{C}^{xy}}\Delta_z m_\tau\right)\right]~, \tag{3.97}$$

where $\mathcal{C}^{xy}$ is a closed strip along the xz-, yz- and τz-plaquettes at a fixed $\hat{z}$. Similarly, there are other charged Wilson observables $W^{yz}(\hat{x},\mathcal{C}^{yz})$ and $W^{zx}(\hat{y},\mathcal{C}^{zx})$. The $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{1})$ dipole symmetry is generated by the line operator (3.95), (3.97) and similar lines operators at fixed $\hat{x}$ or $\hat{y}$. These symmetry operators are quasi-topological, i.e., they are invariant under small deformation of $\mathcal{C}^{xy}$ on the τxy-volume. So, the $\mathbb{Z}_N$ $(\mathbf{3}',\mathbf{1})$ dipole symmetry is unfaithful (in the sense of [103]). The charge operators are (3.96) and similar operators on the other planes.

Real BF-action and the continuum limit

As in Section 3.2.3, we discuss another presentation of this theory, which is closer to the continuum action.

Starting from the integer BF-action (3.93), we replace the integer-valued fields (m_τ, m_{ij}) and $(\hat{m}_\tau^{k(ij)}, \hat{m}^{ij})$ with real-valued fields (A_τ, A_{ij}) and $(\hat{A}_\tau^{k(ij)}, \hat{A}^{ij})$. We constrain them to be integer-valued using Lagrange multiplier fields $(\hat{n}_\tau^{ij}, \hat{n})$ and $(n_{\tau ij}, n_{[ij]k})$. Furthermore, since the gauge fields (A_τ, A_{ij}) and $(\hat{A}_\tau^{k(ij)}, \hat{A}^{ij})$ have real-valued gauge symmetries, we introduce Stueckelberg fields ϕ and $\hat{\phi}^{[ij]k}$ for their gauge symmetries. We end up with the action

$$\begin{aligned}
&\frac{iN}{2\pi}\sum_{\tau k\text{-plaq}} A_{ij}\left(\Delta_\tau \hat{A}^{ij} - \Delta_k \hat{A}_\tau^{k(ij)} - 2\pi \hat{n}_\tau^{ij}\right) + \frac{iN}{2\pi}\sum_{xyz\text{-cube}} A_\tau\left(\Delta_i\Delta_j \hat{A}^{ij} - 2\pi\hat{n}\right)\\
&+ iN\sum_{ij\text{-plaq}} \hat{A}^{ij} n_{\tau ij} - iN\sum_{\tau\text{-link}} \hat{A}_\tau^{[ij]k} n_{[ij]k} - i\sum_{\text{dual site}} \phi\left(\Delta_\tau \hat{n} - \Delta_i\Delta_j \hat{n}_\tau^{ij}\right)\\
&- i\sum_{\text{site}} \hat{\phi}^{[ij]k}(\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik})\ .
\end{aligned} \tag{3.98}$$

(To simplify this particular expression and (3.101), we use the convention that repeated indices i, j and i, j, k are summed over cyclically.) Here ϕ, $\hat{\phi}^{[ij]k}$, A_τ, A_{ij}, $\hat{A}_\tau^{[ij]k}$ and $\hat{A}^{ij}$ are real-valued fields on dual sites, sites, dual τ-links, dual ij-plaquettes, τ-links and k-links, respectively, and $n_{\tau ij}$, $n_{[ij]k}$, $\hat{n}_\tau^{ij}$ and $\hat{n}$ are integer-valued fields on the dual τij-cubes, the dual xyz-cubes, the τk-plaquettes, and the xyz-cubes, respectively. We will refer to this presentation as the real BF-action, which uses both the real and integer fields.

These fields have the same gauge symmetries as in (3.79), (3.83) (3.87), (3.91) except that the α and $\hat{\alpha}^{[ij]k}$ gauge symmetry also acts on ϕ and $\hat{\phi}^{[ij]k}$ as

$$\begin{aligned}
&\phi \sim \phi + N\alpha\ ,\\
&\hat{\phi}^{[ij]k} \sim \hat{\phi}^{[ij]k} + N\hat{\alpha}^{[ij]k}\ .
\end{aligned} \tag{3.99}$$

As a check, summing over $(\hat{n}_\tau^{ij}, \hat{n})$ and $(n_{\tau ij}, n_{[ij]k})$ in (3.98) constrains

$$\begin{aligned}
&\left(A_\tau - \frac{1}{N}\Delta_\tau\phi, A_{ij} - \frac{1}{N}\Delta_i\Delta_j\phi\right) = \frac{2\pi}{N}(m_\tau, m_{ij})\ ,\\
&\left(\hat{A}_\tau^{k(ij)} - \frac{1}{N}\Delta_\tau\hat{\phi}^{k(ij)}, \hat{A}^{ij} - \frac{1}{N}\Delta_k\hat{\phi}^{k(ij)}\right) = \frac{2\pi}{N}(\hat{m}_\tau^{k(ij)}, \hat{m}^{ij})\ .
\end{aligned} \tag{3.100}$$

Substituting them back to the action leads to (3.93).

The real BF-action (3.98) can also be derived through Higgsing the $U(1)$ tensor gauge theory

(3.88) to a $\mathbb{Z}_N$ theory using the field ϕ in (3.80). The Higgs action is

$$
\begin{aligned}
&\frac{i}{2\pi} \sum_{\tau\text{-link}} \hat{B}(\Delta_\tau \phi - NA_\tau - 2\pi n_\tau) + \frac{i}{2\pi} \sum_{ij\text{-plaq}} \hat{E}^{ij}(\Delta_i \Delta_j \phi - NA_{ij} - 2\pi n_{ij}) \\
&- i \sum_{\tau ij\text{-cube}} \hat{A}^{ij}(\Delta_\tau n_{ij} - \Delta_i \Delta_j n_\tau - N n_{\tau ij}) + i \sum_{xyz\text{-cube}} \hat{A}_\tau^{[ij]k}(\Delta_i n_{jk} - \Delta_j n_{ik} - N n_{[ij]k}) \\
&- i \sum_{\text{dual site}} \hat{\phi}^{[ij]k}(\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik}) \ ,
\end{aligned}
\tag{3.101}
$$

where $\hat{B}$ and $\hat{E}^{ij}$ are real-valued fields on the τ-links and the ij-plaquattes, respectively. These fields have the same gauge symmetries as in (3.79), (3.83), (3.87), (3.91), and (3.99). In addition, the fields (n_τ, n_{ij}) also transform under the (q_τ, q_{ij}) gauge symmetry

$$
\begin{aligned}
&n_\tau \sim n_\tau - N q_\tau \ , \\
&n_{ij} \sim n_{ij} - N q_{ij} \ .
\end{aligned}
\tag{3.102}
$$

Summing over the integer-valued fields (n_τ, n_{ij}) constrains

$$
\hat{B} - \sum_{i<j} \Delta_i \Delta_j \hat{A}^{ij} = -2\pi \hat{n} \ , \quad \hat{E}^{ij} - \Delta_\tau \hat{A}^{ij} + \Delta_k \hat{A}_\tau^{k(ij)} = -2\pi \hat{n}_\tau^{ij} \ , \tag{3.103}
$$

where $\hat{n}$ and $\hat{n}_\tau^{ij}$ are integer-valued fields. Substituting them back into the action leads to (3.98). Similarly, the real BF-action (3.98) can also be derived through Higgsing the $U(1)$ tensor gauge theory (3.82) to a $\mathbb{Z}_N$ theory using the field $\hat{\phi}^{[ij]k}$ in (3.90).

Let us discuss a convenient gauge choice for this lattice model. Following similar steps in Section 3.2.3 and in Appendix C.2, we first integrate out ϕ and $\hat{\phi}^{[ij]k}$, and then gauge fix most of the integers $(n_{\tau ij}, n_{[ij]k})$ and $(\hat{n}_\tau^{ij}, \hat{n})$ to zero. Next, we define new fields that are not single-valued and have transition functions. In this gauge choice, it is then straightforward to take the continuum limit of the real BF-action:

$$\frac{iN}{2\pi}\int d\tau dx dy dz \left[\sum_{\substack{\text{cyclic}\\ i,j,k}} A_{ij}\left(\partial_\tau \hat{A}^{ij} - \partial_k \hat{A}^{k(ij)}_\tau\right) + A_\tau \left(\sum_{i<j}\partial_i\partial_j \hat{A}^{ij}\right)\right] \ , \tag{3.104}$$

where we omit the terms that depend on the transition functions of these fields.[8] This is the Euclidean version of the 3+1d $\mathbb{Z}_N$ tensor gauge theory of [73, 81] which describes the low-energy limit of the X-cube model.

We conclude that the modified Villain lattice model (3.93), or equivalently (3.98), flows to the same continuum field theory (3.104) as the original X-cube model (3.92). Conversely, the modified Villain lattice model (3.93), or equivalently (3.98), gives a rigorous setting for the discussion of the continuum theory (3.104) of [73, 81].

3.4 Appendix A: Villain formulation of some classic quantum-mechanical systems

In this appendix, we review two classic quantum-mechanical systems. The various versions of the theory that we will present and the manipulations of the equations are simple warmup examples for the other models.

3.4.1 Appendix A.1: Particle on a ring

We start with the quantum mechanics of a particle on a ring parameterized by the periodic coordinate $q \sim q + 2\pi$. This problem is a classic example of the θ-parameter and its effects. We discuss it using the lattice Villain formulation.

[8]Such boundary terms are necessary in order to make the continuum action (3.104) well-defined. They played a crucial role in the analysis of [84].

The problem is characterized by the Euclidean continuum action

$$S = \oint d\tau \left(\frac{1}{2} (\partial_\tau q)^2 + \frac{i\theta}{2\pi} \partial_\tau q \right) \tag{3.105}$$

and we take the circumference of the Euclidean-time circle to be ℓ. The θ-parameter is 2π-periodic. (Here, we used the freedom to rescale τ to set the coefficient of the kinetic term to $\frac{1}{2}$.)

This system has a global $U(1)$ symmetry shifting q by a constant. And for $\theta \in \pi\mathbb{Z}$, it also has a charge conjugation symmetry $q \to -q$. These two symmetries combine to $O(2)$. As emphasized in [27], for $\theta \in (2\mathbb{Z}+1)\pi$ there is a 't Hooft anomaly stating that while the operator algebra has an $O(2)$ symmetry, the Hilbert space realizes it projectively. Related to that, this system has an anomaly in the space of coupling constants [3, 4]. We are going to reproduce these results on a Euclidean lattice.

Next, we place this theory on a Euclidean-time lattice with lattice spacing a. We label the sites by $\hat{\tau} \in \mathbb{Z}$ such that $\tau = a\hat{\tau}$ and the total number of sites is $L = \ell/a$. Then, following the Villain approach, we make the coordinate $q(\hat{\tau})$ real-valued and add an integer-valued gauge field on the links. The lattice Lagrangian and action are

$$\begin{aligned} &\mathcal{L} = \frac{1}{2a}\big(\Delta q(\hat{\tau}) - 2\pi n(\hat{\tau})\big)^2 + \frac{i\theta}{2\pi}\big(\Delta q(\hat{\tau}) - 2\pi n(\hat{\tau})\big) \ , \\ &S = \sum_{\hat{\tau}=0}^{L-1} \mathcal{L} \\ &\Delta q(\hat{\tau}) = q(\hat{\tau}+1) - q(\hat{\tau}) \ . \end{aligned} \tag{3.106}$$

This system has a $\mathbb{Z}$ gauge symmetry

$$\begin{aligned} &q(\hat{\tau}) \sim q(\hat{\tau}) + 2\pi k(\hat{\tau}) \\ &n(\hat{\tau}) \sim n(\hat{\tau}) + \Delta_\tau k(\hat{\tau}) \\ &k(\hat{\tau}) \in \mathbb{Z} \ . \end{aligned} \tag{3.107}$$

We can replace the Lagrangian in (3.106) by

$$\mathcal{L}' = \frac{1}{2a}\big(\Delta q(\hat\tau) - 2\pi n(\hat\tau)\big)^2 - i\theta n(\hat\tau) \tag{3.108}$$

without changing the action. Unlike $\mathcal{L}$, the new Lagrangian $\mathcal{L}'$ is not gauge invariant under (3.107).

The main point about (3.106) or (3.108) is the description of the θ-term using the gauge field. The integer topological charge of the continuum theory $\frac{1}{2\pi}\oint d\tau \partial_\tau q$ is described by the Wilson line of n.

As in the continuum, the global $U(1)$ symmetry acts by shifting q by a constant. It is $U(1)$ rather than $\mathbb{R}$ because its subgroup $\mathbb{Z} \subset \mathbb{R}$ is gauged. The charge conjugation operation $q \to -q$ should be combined with $n \to -n$. Unless $\theta = 0$, it is not a symmetry of the action (3.106). However, for $\theta \in \pi\mathbb{Z}$, it is a symmetry of e^{-S}.

Let us examine the charge conjugation symmetry more carefully. Its action is "on-site." However, unless $\theta = 0$, it does not leave the Lagrangian $\mathcal{L}$ or even the action S in (3.106) invariant. It does not even leave the exponential of the Lagrangian $e^{-\mathcal{L}}$ invariant. The symmetry is present for $\theta \in \pi\mathbb{Z}$ because it leaves e^{-S} invariant. This opens the door for a 't Hooft anomaly associated with this symmetry and to the related anomaly in the space of coupling constants of [3, 4].[9]

This anomaly is exactly as in the continuum discussion of [27]. It can be demonstrated by adding to (3.106) a classical $U(1)$ gauge field A

$$\mathcal{L} = \frac{1}{2a}\big(\Delta q(\hat\tau) - A(\hat\tau) - 2\pi n(\hat\tau)\big)^2 + \frac{i\theta}{2\pi}\big(\Delta q(\hat\tau) - A(\hat\tau) - 2\pi n(\hat\tau)\big) \ . \tag{3.109}$$

[9]Note that $e^{-\mathcal{L}'}$ with $\mathcal{L}'$ of (3.108) is $O(2)$ invariant for $\theta \in \pi\mathbb{Z}$, but it is not gauge invariant. This is common with anomalies. Using counterterms, we can move the problem around, but we cannot get rid of it.

To see that the gauge symmetry of A is $U(1)$ rather than $\mathbb{R}$, we note that its gauge symmetry

$$\begin{aligned}
&q(\hat{\tau}) \sim q(\hat{\tau}) + \Lambda(\hat{\tau}) + 2\pi k(\hat{\tau}) \\
&n(\hat{\tau}) \sim n(\hat{\tau}) + \Delta_\tau k(\hat{\tau}) - N(\hat{\tau}) \\
&A(\hat{\tau}) \sim A(\hat{\tau}) + \Delta\Lambda(\hat{\tau}) + 2\pi N(\hat{\tau}) \\
&k(\hat{\tau}), N(\hat{\tau}) \in \mathbb{Z}
\end{aligned} \tag{3.110}$$

includes a $\mathbb{Z}$ one-form gauge symmetry with the integer gauge parameter $N(\hat{\tau})$. Invariance under this gauge symmetry shows that the θ-term must depend on A even if we use $\mathcal{L}'$ of (3.108).[10] Now, the charge conjugation symmetry acts also on A and as a result, the θ-term is not invariant under it unless $\theta = 0$. As in [3,4], this also means that there is an anomaly in the 2π-periodicity in θ.

One way to think about this lattice model is the following. We choose the gauge $n(\hat{\tau}) = 0$ except for $n(0)$. In this gauge the Wilson line of n is given by $n(0)$, which is gauge invariant. The remaining gauge symmetry is the identification $q \sim q + 2\pi k$ with integer k independent of $\hat{\tau}$. It is convenient to redefine q to the nonperiodic (in $\hat{\tau}$) variable

$$\bar{q}(\hat{\tau}) = \begin{cases} q(\hat{\tau}) & \text{for } \hat{\tau} = 1, \cdots, L \\ q(0) + 2\pi n(0) & \text{for } \hat{\tau} = 0 \end{cases} . \tag{3.113}$$

[10] An extreme version of this system is when the lattice has only one site, i.e., $L = 1$. In that case the action becomes

$$\frac{1}{2a}\big(A(\hat{\tau}) + 2\pi n(\hat{\tau})\big)^2 - \frac{i\theta}{2\pi}\big(A(\hat{\tau}) + 2\pi n(\hat{\tau})\big) . \tag{3.111}$$

The global $U(1)$ symmetry is reflected in the fact that action is independent of q. It depends only on the integer dynamical gauge field n and the classical gauge field A. The remaining gauge symmetry is the one-form gauge symmetry

$$\begin{aligned}
&n(\hat{\tau}) \sim n(\hat{\tau}) - N(\hat{\tau}) \\
&A(\hat{\tau}) \sim A(\hat{\tau}) + 2\pi N(\hat{\tau}) \\
&N(\hat{\tau}) \in \mathbb{Z} .
\end{aligned} \tag{3.112}$$

Again, the anomaly is manifest in (3.111).

In these variables, after dropping the bar, (3.106) becomes

$$\begin{aligned}\mathcal{L} &= \frac{1}{2a}\big(\Delta q(\hat{\tau})\big)^2 + \frac{i\theta}{2\pi}\Delta q(\hat{\tau}) \ , \\ S &= \sum_{\hat{\tau}=0}^{L-1} \mathcal{L} \ .\end{aligned} \tag{3.114}$$

This can be interpreted as follows. We have a real-valued field q and we sum over twisted boundary conditions labeled by an integer $n(0)$ such that $q(\hat{\tau}+L) = q(\hat{\tau}) - 2\pi n(0)$.

In the form (3.114), it is easy to take the continuum limit. We take $a \to 0$, $L \to \infty$ with finite $\ell = La$. In this limit q becomes smooth and we recover (3.105).

3.4.2 Appendix A.2: Noncommutative torus

Next, we review the quantum mechanics of N degenerate ground states using a Euclidean lattice.

In the continuum, the theory can be described using a phase space of two circle-valued coordinates p, q with the Euclidean action

$$\frac{iN}{2\pi}\int d\tau\, p\dot{q} \ . \tag{3.115}$$

(Soon, we will make this action more precise.) Its quantization leads to N degenerate ground states. These ground states are in the minimal representation of the operator algebra

$$\begin{aligned}&UV = e^{\frac{2\pi i}{N}} VU, \\ &U = e^{ip}, \quad V = e^{iq} \ .\end{aligned} \tag{3.116}$$

Since p and q are circle-valued, i.e., $p(\tau) \sim p(\tau) + 2\pi$ and $q(\tau) \sim q(\tau) + 2\pi$, the Lagrangian in (3.115) is not well defined. There are several ways to correct it. One of them involves lifting q and p to be real-valued with transition functions at some reference point τ_*. Then, we can take the action to be [104, 3, 4] (see also [105, 106, 29, 84])[11]

$$\frac{iN}{2\pi}\int_{\tau_*}^{\tau_*+\ell} d\tau\, p\dot{q} - iNw_p(\tau_*)q(\tau_*) \ , \tag{3.117}$$

where ℓ is the period of the Euclidean time and $w_p = p(\tau_* + \ell) - p(\tau_*)$ is the winding number of p. Similarly, we define $w_q = q(\tau_* + \ell) - q(\tau_*)$ as the winding number of q. In the path integral, we sum over the integers w_p and w_q. The action is independent of the choice of τ_*, i.e., the choice of trivialization.

Note that as in (3.105), we could have added to (3.117) θ-terms for p and q. However, it is clear that they can be absorbed in shifts of q and p respectively. Therefore, without loss of generality, we can ignore them. The same comment applies to the lattice discussion below.

We now discretize the Euclidean time direction and replace it by a periodic lattice with $\tau = a\hat{\tau}$, $\hat{\tau} \in \mathbb{Z}$ and periodicity $\hat{\tau} \sim \hat{\tau} + L$. We use the Villain approach and let q and p be real-valued (as opposed to circle-valued) coordinates coupled to $\mathbb{Z}$ gauge fields n_q and n_p. The action is

$$\begin{aligned} &\frac{iN}{2\pi}\sum_{\hat{\tau}=0}^{L-1}\Big[p(\hat{\tau})\big(\Delta q(\hat{\tau}) - 2\pi n_q(\hat{\tau})\big) + 2\pi n_p(\hat{\tau})q(\hat{\tau})\Big] \ , \\ &\Delta q(\hat{\tau}) \equiv q(\hat{\tau}+1) - q(\hat{\tau}) \ . \end{aligned} \tag{3.118}$$

The fields q, n_p naturally live on the lattice sites, while p, n_q naturally live on the links. These fields are subject to gauge symmetries with integer gauge parameters k_p, k_q

$$\begin{aligned} &p(\hat{\tau}) \sim p(\hat{\tau}) + 2\pi k_p(\hat{\tau}) \ , \\ &q(\hat{\tau}) \sim q(\hat{\tau}) + 2\pi k_q(\hat{\tau}) \ , \\ &n_p(\hat{\tau}) \sim n_p(\hat{\tau}) + k_p(\hat{\tau}) - k_p(\hat{\tau}-1) \ , \\ &n_q(\hat{\tau}) \sim n_q(\hat{\tau}) + k_q(\hat{\tau}+1) - k_q(\hat{\tau}) \ . \end{aligned} \tag{3.119}$$

Note that the Lagrangian is not gauge invariant. Even the action is not gauge invariant. But

[11] The rigorous mathematical treatment uses differential cohomology [107–110] (see [111–113] and the references therein for modern developments).

e^{-S} is gauge invariant.

We can choose the gauge $n_q(\hat{\tau}) = n_p(\hat{\tau}) = 0$ except for $n_q(0), n_p(0)$. The action then becomes

$$\frac{iN}{2\pi}\sum_{\hat{\tau}=0}^{L-1} p(\hat{\tau})\Delta q(\hat{\tau}) - iNn_q(0)p(0) + iNn_p(0)q(0) \ . \tag{3.120}$$

There is a residual gauge symmetry:

$$\begin{aligned} p(\hat{\tau}) &\sim p(\hat{\tau}) + 2\pi \ , \\ q(\hat{\tau}) &\sim q(\hat{\tau}) + 2\pi \ . \end{aligned} \tag{3.121}$$

To relate the gauge fixed lattice action (3.120) to the continuum action (3.117), we define new variables $\bar{p}, \bar{q}$ on the covering space of the periodic lattice:

$$\begin{aligned} \bar{p}(\hat{\tau}) &= \begin{cases} p(\hat{\tau}) & \text{for } \hat{\tau} = 0, \cdots, L-1 \\ p(\hat{\tau}) - 2\pi n_p(0) & \text{for } \hat{\tau} = L \end{cases} \ , \\ \bar{q}(\hat{\tau}) &= \begin{cases} q(\hat{\tau}) & \text{for } \hat{\tau} = 1, \cdots, L \\ q(0) + 2\pi n_q(0) & \text{for } \hat{\tau} = 0 \end{cases} \ . \end{aligned} \tag{3.122}$$

Unlike the single-valued real fields p, q, which obey $p(0) = p(L)$, $q(0) = q(L)$, the new real fields $\bar{p}, \bar{q}$ are not single-valued on the periodic lattice; they can have non-trivial winding number $w_p = -n_p(0)$, $w_q = -n_q(0)$. In terms of the new variables, the action becomes

$$\frac{iN}{2\pi}\sum_{\hat{\tau}=0}^{L-1} \bar{p}(\hat{\tau})\Delta\bar{q}(\hat{\tau}) - iNw_p\bar{q}(0) \ , \tag{3.123}$$

In the continuum limit, this lattice action becomes (3.117).

Instead of gauge fixing the integer fields n_p, n_q, we can sum over them. This restricts the real-valued fields p, q to $p = \frac{2\pi}{N} m_p$ and $q = \frac{2\pi}{N} m_q$ with integer fields m_p, m_q. The action becomes

$$\frac{2\pi i}{N} \sum_{\hat{\tau}=1}^{L} m_p(\hat{\tau}) \Delta m_q(\hat{\tau}) \ , \tag{3.124}$$

with the following gauge symmetry making the integer fields $\mathbb{Z}_N$ variables

$$\begin{aligned} m_p(\hat{\tau}) &\sim m_p(\hat{\tau}) + N k_p(\hat{\tau}) \ , \\ m_q(\hat{\tau}) &\sim m_q(\hat{\tau}) + N k_q(\hat{\tau}) \ . \end{aligned} \tag{3.125}$$

3.5 Appendix B: Modified Villain formulation of 2d Euclidean lattice theories without gauge fields

In this appendix, we review well-known facts about some lattice models and their Villain formulation. As in the models in the bulk of the chapter, we deform the standard Villain action to another lattice action, which has special properties. In particular, it has enhanced global symmetries and it exhibits special dualities. Then, we study other models by deforming this special action.

3.5.1 Appendix B.1: 2d Euclidean XY-model

Here we study the two-dimensional Euclidean XY-model on the lattice and in the continuum limit [114, 115].

Lattice models

We place the theory on a 2d Euclidean periodic lattice, whose sites are labeled by integers $(\hat{x}, \hat{y}) \sim (\hat{x} + L^x, \hat{y}) \sim (\hat{x}, \hat{y} + L^y)$. The dynamical variables are phases $e^{i\phi}$ at each site of the lattice. The action is

$$\beta \sum_{\text{link}} [1 - \cos(\Delta_\mu \phi)] \ , \tag{3.126}$$

where $\mu = x, y$ labels the directions and $\Delta_x\phi \equiv \phi(\hat{x}+1, \hat{y}) - \phi(\hat{x}, \hat{y})$ and $\Delta_y\phi \equiv \phi(\hat{x}, \hat{y}+1) - \phi(\hat{x}, \hat{y})$ are the lattice derivatives.

At large β, we can approximate the action (3.126) by the Villain action [85]:

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 \ . \tag{3.127}$$

Here ϕ is a real-valued field and n_μ is an integer-valued field on the links. These fields satisfy periodic boundary conditions.

The fact that in the original formulation (3.126), ϕ was circle-valued rather than real-valued is related to the $\mathbb{Z}$ gauge symmetry

$$\phi \sim \phi + 2\pi k \ , \qquad n_\mu \sim n_\mu + \Delta_\mu k \ , \tag{3.128}$$

where k is an integer-valued gauge parameter on the sites. We can interpret n_μ as a $\mathbb{Z}$ gauge field, which makes ϕ compact.

The gauge invariant "field strength" of the gauge field n_μ is

$$\mathcal{N} \equiv \Delta_x n_y - \Delta_y n_x \ . \tag{3.129}$$

It can be interpreted as the local vorticity of the configurations.

We are interested is suppressing vortices. One way to do that is to add to the action (3.127) a term like

$$\kappa \sum_{\text{plaquette}} \mathcal{N}^2 \tag{3.130}$$

with positive κ. For $\kappa \to \infty$ the vortices are completely suppressed [86]. Instead of adding this term and taking this limit, we can introduce a Lagrange multiplier $\tilde{\phi}$ to impose $\mathcal{N} = 0$ as a constraint. The full action now becomes[12]

$$S = \frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 + i\sum_{\text{plaquette}}\tilde{\phi}\mathcal{N}\ , \tag{3.131}$$

where the Lagrange multiplier $\tilde{\phi}$ is a real-valued field on the plaquettes (or dual sites). It has a $\mathbb{Z}$ gauge symmetry

$$\tilde{\phi} \sim \tilde{\phi} + 2\pi\tilde{k}\ , \tag{3.132}$$

with $\tilde{k}$ is an integer-valued gauge parameter on the plaquettes.

Note that the action (3.131) is not invariant under this gauge symmetry. However, e^{-S} is gauge invariant. In fact, even the local quantity $e^{-\mathcal{L}}$, with $\mathcal{L}$ the Lagrangian density, is invariant.

The action (3.131) is the starting point of our discussion. We refer to it as the modified Villain action of the XY-model.[13]

We can restore the vortices by perturbing the modified Villain action (3.131) as

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 + i\sum_{\text{plaquette}}\tilde{\phi}\mathcal{N} - \lambda\sum_{\text{plaquette}}\cos(\tilde{\phi})\ . \tag{3.133}$$

(For simplicity of the presentation, we take $\lambda \geq 0$.) Note that the action is still invariant under the gauge symmetries (3.128) and (3.132). Integrating out $\tilde{\phi}$ gives

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 - \sum_{\text{plaquette}}\log I_{|\mathcal{N}|}(\lambda)\ , \tag{3.134}$$

where $I_k(z)$ is the modified Bessel function of the first kind. Let us compare this action with

[12] Related ideas were used in various places, including [116].

[13] Using common terminology in the condensed matter literature, one could refer to the corresponding theory as noncompact. However, we emphasize that even though the ϕ field in (3.127) and (3.131) is real-valued, i.e., noncompact, the gauge symmetry (3.128) effectively compactifies the range of ϕ. The effect of the term with $\mathcal{N}$ in (3.131) is to suppress the vortices rather than to de-compactify the target space. We will discuss it further below.

(3.130). For small $\lambda \ll 1$, we have

$$-\log I_k(\lambda) \approx \log\left[k!\left(\frac{2}{\lambda}\right)^k\right] + O(\lambda^2)\ . \tag{3.135}$$

In this case, vortices with $|\mathcal{N}| > 1$ are suppressed. For $|\mathcal{N}| = 0, 1$ we identify

$$\kappa \approx \log\frac{2}{\lambda} \gg 1\ . \tag{3.136}$$

In the other limit $\lambda \gg 1$, we have

$$-\log I_k(\lambda) \sim \frac{1}{2\lambda}k^2 + O(\lambda^{-2}) \tag{3.137}$$

where we ignored some k-independent terms that depend on λ. In this case, we can identify

$$\kappa \approx \frac{1}{2\lambda} \ll 1\ . \tag{3.138}$$

We conclude that the deformation $-\lambda\cos(\tilde{\phi})$ is mapped to $\kappa\mathcal{N}^2$, and small (large) λ corresponds to large (small) κ.

To summarize, the XY-model is usually studied using the actions (3.126) or (3.127). We added another coupling to this model (3.130). Equivalently, we can write the model as (3.133) and then the usually studied model (3.127) is obtained in the limit $\lambda \to \infty$. On the other hand, when $\lambda = 0$, this reduces to our modified Villain action (3.131) of the XY-model.

Below we will see that the modified Villain action (3.131), unlike its other lattice relatives, exhibits many properties similar to its continuum limit, including emergent global symmetries, anomalies, and self-duality.

Global symmetries

The three models, (3.126), (3.127), and (3.131) have a *momentum symmetry*, which acts as

$$\phi \to \phi + c^m \ , \tag{3.139}$$

where c^m is a real position-independent constant. Due to the zero mode of the gauge symmetry (3.128), the $2\pi\mathbb{Z}$ part of this symmetry is gauged. So the momentum symmetry is $U(1)$ rather than $\mathbb{R}$.

From (3.127) and (3.131) we find the Noether current of momentum symmetry[14]

$$J^m_\mu = -i\beta(\Delta_\mu \phi - 2\pi n_\mu) \ , \tag{3.140}$$

which is conserved because of the equation of motion of ϕ. The momentum charge is[15]

$$Q^m(\tilde{\mathcal{C}}) = \sum_{\text{dual link}\in\tilde{\mathcal{C}}} \epsilon_{\mu\nu} J^m_\nu \ , \tag{3.141}$$

where $\tilde{\mathcal{C}}$ is a curve along the dual links of the lattice. The dependence of Q^m on $\tilde{\mathcal{C}}$ is topological. The local operator $e^{i\phi}$ is charged under this symmetry.

The modified Villain action (3.131) (but not (3.126) or (3.127)) also has a *winding symmetry*, which acts as

$$\tilde\phi \to \tilde\phi + c^w \ , \tag{3.142}$$

where c^w is a real constant. Due to the zero mode of the gauge symmetry (3.132), the $2\pi\mathbb{Z}$ part of this symmetry is gauged. So the winding symmetry is also $U(1)$.

[14]The factor of i in the Euclidean signature is such that the corresponding charge is real.
[15]Here, $\epsilon_{xy} = -\epsilon_{yx} = 1$ and $\epsilon_{xx} = \epsilon_{yy} = 0$.

The Noether current of the winding symmetry is[16]

$$J^w_\mu = \frac{\epsilon_{\mu\nu}}{2\pi}(\Delta_\nu\phi - 2\pi n_\nu) \ , \tag{3.143}$$

which is conserved because of the equation of motion of $\tilde{\phi}$. It is crucial that n_μ is flat, i.e., $\mathcal{N} = 0$ and vortices are suppressed, for the Noether current to be conserved. The winding charge is

$$Q^w(\mathcal{C}) = \sum_{\text{link}\in\mathcal{C}} \epsilon_{\mu\nu} J^w_\nu = -\sum_{\text{link}\in\mathcal{C}} n_\mu \ , \tag{3.144}$$

where $\mathcal{C}$ is a curve along the links of the lattice. The last equation follows from the single-valuedness of ϕ. Hence, we can interpret $Q^w(\mathcal{C})$ as the gauge invariant Wilson line of the $\mathbb{Z}$ gauge field n_μ. It is topological due to the flatness condition of n_μ. Finally, the local operator $e^{i\tilde{\phi}}$ is charged under this symmetry.

Both the momentum symmetry (3.139) and the winding symmetry (3.142) act locally on the fields and they both leave the action (3.131) invariant. However, the Lagrangian density in (3.131) is invariant under the momentum symmetry, but not under the winding symmetry. This fact makes it possible for these symmetries to have a mixed 't Hooft anomaly, even though the two symmetries act locally ("on site").

Using "summing by parts", we can write (3.131) as

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 + i\sum_{\text{plaquette}}(n_x\Delta_y\tilde{\phi} - n_y\Delta_x\tilde{\phi}) \ . \tag{3.145}$$

In this form both the momentum symmetry (3.139) and the winding symmetry (3.142) act locally and leave the Lagrangian density invariant. How is this compatible with the anomaly? The point is that unlike (3.131), the Lagrangian density in (3.145) is not gauge invariant. As is common with anomalies, we can move the problem around, but we cannot completely avoid it.

One way to see this anomaly is by trying to couple the action (3.131) to background gauge

[16]From the action (3.131), the Noether current appears to be $J^w_\mu = -\epsilon_{\mu\nu}n_\nu$, but it is not gauge invariant. Therefore, we added to it an improvement term to construct a gauge invariant current.

fields for the momentum and winding symmetries $(A_\mu; N)$ and $(\tilde{A}_\mu; \tilde{N})$. Here $A_\mu, \tilde{A}_\mu$ are real-valued and $N, \tilde{N}$ are integer-valued. The action is

$$\begin{aligned}&\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - A_\mu - 2\pi n_\mu)^2 + i\sum_{\text{plaquette}}\tilde{\phi}(\Delta_x n_y - \Delta_y n_x + N)\\&-\frac{i}{2\pi}\sum_{\text{link}}\epsilon_{\mu\nu}\tilde{A}_\mu(\Delta_\nu\phi - A_\nu - 2\pi n_\nu) + i\sum_{site}\tilde{N}\phi\ ,\end{aligned} \tag{3.146}$$

with the gauge symmetry

$$\begin{aligned}&\phi \sim \phi + \alpha + 2\pi k\ , && \tilde{\phi} \sim \tilde{\phi} + \tilde{\alpha} + 2\pi\tilde{k}\ ,\\&A_\mu \sim A_\mu + \Delta_\mu\alpha + 2\pi K_\mu\ , && \tilde{A}_\mu \sim \tilde{A}_\mu + \Delta_\mu\tilde{\alpha} + 2\pi\tilde{K}_\mu\ ,\\&n_\mu \sim n_\mu + \Delta_\mu k - K_\mu\ , && \tilde{N} \sim \tilde{N} + \Delta_x\tilde{K}_y - \Delta_y\tilde{K}_x\ ,\\&N \sim N + \Delta_x K_y - \Delta_y K_x\ .\end{aligned} \tag{3.147}$$

Here, $K_\mu, \tilde{K}_\mu$ are integers, and $\alpha, \tilde{\alpha}$ are real. They are the gauge parameters of the background gauge fields $(A_\mu; N)$ and $(\tilde{A}_\mu; \tilde{N})$. The variation of the action under this gauge transformation is

$$-\frac{i}{2\pi}\sum_{\text{plaquette}}\tilde{\alpha}(\Delta_x A_y - \Delta_y A_x - 2\pi N) + i\sum_{\text{plaquette}}(\tilde{K}_x A_y - \tilde{K}_y A_x)\ . \tag{3.148}$$

It signals an anomaly because it cannot be cancelled by adding any 1+1d local counterterms. This expression of the anomaly is the lattice version of the familiar continuum expression $-\frac{i}{2\pi}\int dxdy\ \tilde{\alpha}(\partial_x A_y - \partial_y A_x)$.

T-Duality

Here we will demonstrate the self-duality of the modifield Villain lattice model (3.131). We start with the presentation (3.145). Using the Poisson resummation formula (3.6) for n_μ and ignoring the overall factor, we can dualize the above action to

$$\begin{aligned}&\frac{1}{2(2\pi)^2\beta}\sum_{\text{dual link}}(\Delta_\mu\tilde{\phi}-2\pi\tilde{n}_\mu)^2+i\sum_{\text{site}}\phi\tilde{\mathcal{N}}\ ,\\&\tilde{\mathcal{N}}\equiv\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x\ ,\end{aligned}\tag{3.149}$$

where $\tilde{n}_\mu$ is an integer-valued field on the dual links. The gauge symmetry of the original theory acts as

$$\tilde{\phi}\sim\tilde{\phi}+2\pi\tilde{k}\ ,\qquad \tilde{n}_\mu\sim\tilde{n}_\mu+\Delta_\mu\tilde{k},\qquad \phi\sim\phi+2\pi k\ .\tag{3.150}$$

$\tilde{n}_\mu$ can be interpreted as the $\mathbb{Z}$ gauge field associated with the gauge symmetry of $\tilde{\phi}$ and $\tilde{\mathcal{N}}$ is its field strength. Furthermore, we can interpret ϕ as a Lagrange multiplier imposing $\tilde{\mathcal{N}}=0$ as a constraint.

We conclude that the modified Villain action (3.131) is a self-dual lattice model with $\beta\leftrightarrow\frac{1}{(2\pi)^2\beta}$. Moreover, the momentum and winding currents, (3.140) and (3.143), in the dual picture are

$$J^m_\mu=\frac{\epsilon_{\mu\nu}}{2\pi}(\Delta_\nu\tilde{\phi}-2\pi\tilde{n}_\nu)\ ,\qquad J^w_\mu=-\frac{i}{(2\pi)^2\beta}(\Delta_\mu\tilde{\phi}-2\pi\tilde{n}_\mu)\ .\tag{3.151}$$

We emphasize that the lattice model (3.131) is exactly self-dual, rather than being only IR-self-dual. It has exact T-duality.

We can easily relate this discussion to the classical analysis of [114,115]. By adding the term $-\lambda\cos(\tilde{\phi})$ to the Lagrangian and taking $\lambda\to\infty$, the field $\tilde{\phi}$ is frozen at zero and we end up with Villain action (3.127). Repeating this in the dual action (3.149), we find

$$\begin{aligned}&\frac{1}{2\beta}\sum_{\text{dual link}}\tilde{n}_\mu^2+i\sum_{\text{site}}\phi\tilde{\mathcal{N}}\ ,\\&\tilde{\mathcal{N}}\equiv\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x\ .\end{aligned}\tag{3.152}$$

Locally, the Lagrange multiplier ϕ determines $\tilde{n}_\mu = \Delta_\mu q$ with an integer q.[17] We end up with

$$\frac{1}{2\beta} \sum_{\text{dual link}} (\Delta_\mu q)^2 \ , \tag{3.153}$$

which is the dual theory of [114, 115].

Gauge-fixing and the continuum limit

In the following we will pick a convenient gauge where most of the integer fields are set to zero. Following the discussion around (3.120), we integrate out $\tilde{\phi}$, which imposes the flatness condition on n_μ. Then, we gauge fix $n_\mu(\hat{x}, \hat{y}) = 0$ at all links, except $n_x(L^x-1, \hat{y})$ and $n_y(\hat{x}, L^y-1)$ (recall, $\hat{x}^\mu \sim \hat{x}^\mu + L^\mu$). The remaining information in the gauge fields n_μ is in the two integers $n_x(L^x - 1, \hat{y}) \equiv \bar{n}_x$ and $n_y(\hat{x}, L^y - 1) \equiv \bar{n}_y$, i.e., in the holonomies of n_μ around the x and y cycles. The residual gauge symmetry is

$$\phi \sim \phi + 2\pi\mathbb{Z} \ . \tag{3.154}$$

Let us define a new field $\bar{\phi}$ such that

$$\bar{\phi}(0,0) = \phi(0,0) \ , \qquad \Delta_\mu \bar{\phi} = \Delta_\mu \phi - 2\pi n_\mu \ . \tag{3.155}$$

In the gauge above, where in most of the links $n_\mu = 0$, in most of the sites $\bar{\phi} = \phi$. Then the action in terms of $\bar{\phi}$ is

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \bar{\phi})^2 \ . \tag{3.156}$$

[17] More precisely, $\tilde{\mathcal{N}} = 0$ can be solved in terms of an integer-valued field q, but q does not have to be periodic (i.e., single-vlaued on the torus). Its lack of periodicity is characterized by two integers, which are the Wilson lines of $\tilde{n}$ around two cycles of the torus. This Wilson line is the momentum charge (3.141) constructed out of the momentum current (3.151) and it is nontrivial only when q is not periodic.

Although ϕ and n_μ are single-valued fields, $\bar{\phi}$ can wind around nontrivial cycles:

$$\begin{aligned} \bar{\phi}(\hat{x}+L^x,\hat{y}) &= \bar{\phi}(\hat{x},\hat{y}) - 2\pi\bar{n}_x \ , \\ \bar{\phi}(\hat{x},\hat{y}+L^y) &= \bar{\phi}(\hat{x},\hat{y}) - 2\pi\bar{n}_y \ . \end{aligned} \tag{3.157}$$

So, in the path integral, we should sum over nontrivial winding sectors of $\bar{\phi}$.[18]

In the continuum limit $a \to 0$ such that $\ell^\mu \equiv aL^\mu$ is fixed, the action (3.156) becomes

$$\frac{\beta}{2}\int dxdy\ (\partial_\mu\phi)^2 \ , \tag{3.158}$$

where we dropped the bar on ϕ. This is the action of the 2d compact boson. Locally, this is the same as a theory of a noncompact scalar ϕ. However, here we sum over twisted boundary conditions and that makes the ϕ field compact. See the related discussion in footnote 18.

Kosterlitz-Thouless transition

In order to compare with the standard conformal field theory literature (e.g., [117, 118]), we define the radius R of the compact boson as $R = \sqrt{\pi\beta}$. The theory at radius R has momentum and winding operators with dimensions

$$(h,\bar{h}) = \left(\frac{1}{2}\left(\frac{n_m}{2R}+n_wR\right)^2, \frac{1}{2}\left(\frac{n_m}{2R}-n_wR\right)^2\right) \ , \tag{3.159}$$

where n_m, n_w are the momentum and winding charges of the operator. These operators correspond to the lattice operators $e^{i(n_m\phi+n_w\tilde{\phi})}$. T-duality exchanges the theories at radius R and $\frac{1}{2R}$. At the radius $R=\frac{1}{\sqrt{2}}$, the theory is self-dual. See Figure 3.1.

Unlike the modified Villain model (3.131), the original XY-model (3.126) and its Villain

[18]Note that the variables $\bar{\phi}$ are noncompact and we can rescale them to make the action (3.156) independent of β. Then, the compactness and the β dependence enter only through the twisted boundary conditions (3.157). One might say that therefore, the local dynamics is independent of β and the model is the same as that of a noncompact scalar. This is the rationale behind the terminology mentioned in footnote (13). This reasoning is valid when we consider the model with fixed twisted boundary conditions like (3.157). However, in our case, we sum over this twist. And this affects the set of local operators in the theory. In particular, as in (3.159), their dimensions depend on the value of $\beta = R^2/\pi$.

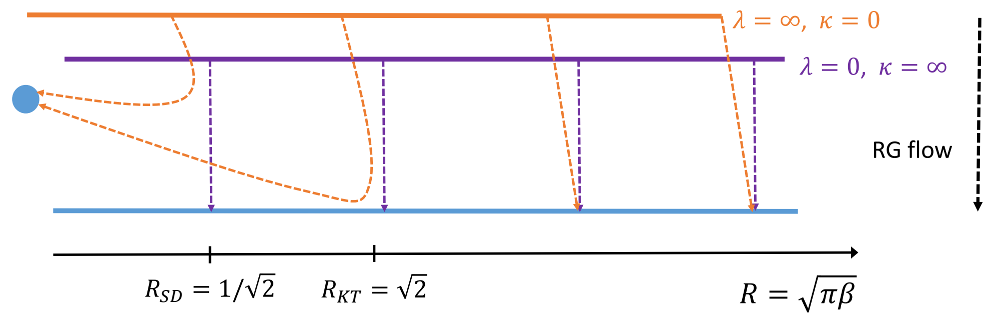


Figure 3.1: The space of coupling constants of the 2d Euclidean XY-model. The orange line corresponds to the theories based on (3.126) or (3.127), while the purple line corresponds to the modified theory (3.131). Each of them depends on the parameter $R = \sqrt{\pi\beta}$. The parameter λ (equivalently, κ) interpolates between these two lines. The theories of the purple line (3.131) are special because they have a global $U(1)$ winding symmetry and they enjoy a $R \to \frac{1}{2R}$ duality with selfduality at $R = \frac{1}{\sqrt{2}}$. The dashed lines represent the renormalization group flow, or equivalently the continuum limit. The theories of the purple line flow to the $c = 1$ compact-boson conformal field theories, which are represented by the blue line. The theories of the orange line (3.126) or (3.127) also flow to this conformal theory, provided $R \geq R_{KT} = \sqrt{2}$ (equivalently, $\beta \geq \frac{2}{\pi}$). For $R < R_{KT} = \sqrt{2}$ (equivalently, $\beta < \frac{2}{\pi}$), the theories of the orange line flow to a gapped phase, which is represented by the blue region at the left. The more generic theories with nonzero but finite λ (and κ) behave like the theories of the orange line.

counterpart (3.127) have only the momentum symmetry, but no winding symmetry. It could still happen that their long-distance theory has such an emergent winding symmetry. This happens when the winding number violating operators are irrelevant (or exactly marginal) in the IR theory. This is the case for $R \geq R_{KT} = \sqrt{2}$, or equivalently $\beta \geq \beta_{KT} = \frac{2}{\pi}$, where the subscript KT stands for Kosterlitz-Thouless. However, for smaller values of R and β the winding operators are relevant and the lattice models undergo the Kosterlitz-Thouless transition to a gapped phase. See Figure 3.1.

Finally, this reasoning implies that the qualitative behavior of the flow for finite nonzero λ is the same as the flow for infinite λ in Figure 3.1. Only for $\lambda = 0$ is the flow different (as the purple line in Figure 3.1). Also, it is straightforward to replace the deformation $\cos(\tilde{\phi})$ by $\cos(W\tilde{\phi})$ for generic integer W. This breaks the $U(1)$ winding symmetry to $\mathbb{Z}_W$. Then the flow is as from the orange curve in Figure 3.1, except that the Kosterlitz-Thouless point moves to $R = \frac{\sqrt{2}}{W}$.

3.5.2 Appendix B.2: 2d Euclidean $\mathbb{Z}_N$ clock model

Lattice models

The $\mathbb{Z}_N$ clock model [114, 119–123] can be obtained by restricting the phase variables $e^{i\phi}$ in the XY-model (3.126) to $\mathbb{Z}_N$ variables $e^{2\pi i m/N}$. More generally, this model has $\lfloor N/2 \rfloor$ nearest-neighbor couplings

$$\sum_{M=1}^{\lfloor N/2 \rfloor} J_M \sum_{\text{link}} \left[1 - \cos\left(\frac{2\pi M}{N} \Delta_\mu m \right) \right] \ . \tag{3.160}$$

where $\lfloor N/2 \rfloor$ is the integer part of $N/2$. A particular one-dimensional locus in the parameter space of $\{J_M\}$ is given by the Villain action:

$$\frac{\beta}{2} \left(\frac{2\pi}{N} \right)^2 \sum_{\text{link}} (\Delta_\mu m - N n_\mu)^2 \ . \tag{3.161}$$

The integer fields m, n_μ are subject to a gauge symmetry with integer gauge parameter k

$$\begin{aligned} m &\sim m + Nk \ , \\ n_\mu &\sim n_\mu + \Delta_\mu k \ . \end{aligned} \tag{3.162}$$

This model (3.161) can be embedded in the XY-model of Appendix B.1. In general, we can deform the action (3.131) to

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquette}} \tilde{\phi} \mathcal{N} - \lambda \sum_{\text{plaquette}} \cos(W\tilde{\phi}) - \tilde{\lambda} \sum_{\text{site}} \cos(N\phi) \ , \tag{3.163}$$

with integer N and W. The term with $\tilde{\lambda}$ breaks the $U(1)$ momentum global symmetry to $\mathbb{Z}_N$, which is generated by $\phi \to \phi + \frac{2\pi}{N}$. Similarly, the term with λ breaks the $U(1)$ winding global symmetry to $\mathbb{Z}_W$.

The most commonly analyzed case is with $W = 1$ and $\tilde{\lambda}, \lambda \to \infty$. Then, $\tilde{\phi}$ is constrained to vanish and therefore the vortices are not suppressed. Similarly, ϕ is constrained to have the values $\phi = \frac{2\pi m}{N}$, thus leading to (3.161).

Kramers-Wannier duality

It is straightforward to repeat the analysis in Appendix B.1 and to dualize (3.163) to

$$\begin{aligned}&\frac{1}{2(2\pi)^2\beta}\sum_{\text{dual link}}(\Delta_\mu\tilde{\phi}-2\pi\tilde{n}_\mu)^2+i\sum_{\text{site}}\phi\tilde{\mathcal{N}}-\lambda\sum_{\text{plaquette}}\cos(W\tilde{\phi})-\tilde{\lambda}\sum_{\text{site}}\cos(N\phi)\ ,\\&\tilde{\mathcal{N}}\equiv\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x\ ,\end{aligned} \tag{3.164}$$

where $\tilde{n}_\mu$ is an integer-valued field on the dual links. The gauge symmetry of the theory is

$$\tilde{\phi}\sim\tilde{\phi}+2\pi\tilde{k}\ ,\qquad \tilde{n}_\mu\sim\tilde{n}_\mu+\Delta_\mu\tilde{k},\qquad \phi\sim\phi+2\pi k\ . \tag{3.165}$$

We conclude that the action (3.163) is dual to a similar system with $\beta\leftrightarrow\frac{1}{(2\pi)^2\beta}$ and $N\leftrightarrow W$.

In the special case with $W=1$ and $\tilde{\lambda},\lambda\to\infty$, (3.163) is dualized to

$$\frac{1}{2\beta}\sum_{\text{dual link}}\tilde{n}_\mu^2+\frac{2\pi i}{N}\sum_{\text{site}}m(\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x) \tag{3.166}$$

with the gauge symmetry

$$m\sim m+Nk \tag{3.167}$$

with integer k. We can find it either by substituting $\phi=\frac{2\pi m}{N}$, $\tilde{\phi}=0$ in (3.164), or by directly dualizing (3.161).

We see that unlike the modified Villain action for the XY-model (3.131), this theory is not selfdual. Comparing with the general case (3.164), this follows from the fact that now $W=1$ and the duality there exchanges $W\leftrightarrow N$.

How is this consistent with the known Kramers-Wannier duality of this theory [114,119–123]?

In order to answer this question we first add integer-valued fields $\tilde{m}$ and $\hat{n}_\mu$ to the action (3.166)

$$\frac{1}{2\beta}\sum_{\text{dual link}}(\Delta_\mu\tilde{m}-N\hat{n}_\mu-\tilde{n}_\mu)^2+\frac{2\pi i}{N}\sum_{\text{site}}m(\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x)\ . \tag{3.168}$$

In addition to the gauge symmetry (3.167), this action has the gauge symmetry

$$\begin{aligned}
&\tilde{m} \sim \tilde{m} + \tilde{k} \ , \\
&\hat{n}_\mu \sim \hat{n}_\mu - \hat{q}_\mu \ , \\
&\tilde{n}_\mu \sim \tilde{n}_\mu + \Delta_\mu \tilde{k} + N\hat{q}_\mu \ .
\end{aligned} \tag{3.169}$$

Here $\tilde{k}$ is an integer zero-form gauge parameter and $\hat{q}_\mu$ is an integer one-form gauge parameter. This new action (3.168) is equivalent to (3.166), as can be seen by completely gauge fixing (3.169) by setting $\tilde{m} = \hat{n}_\mu = 0$.

Now, we can interpret (3.168) as follows. Locally, the Lagrange multiplier m sets $\tilde{n}_\mu$ to a pure gauge and we can set it to zero. Then, (3.168) is the same as the Villain form of the $\mathbb{Z}_N$ action (3.161) with the replacement $\beta \leftrightarrow \frac{N^2}{4\pi^2\beta}$. This shows that locally, the $\mathbb{Z}_N$ clock-model has Kramers-Wannier duality.

However, globally, the Lagrange multiplier m in (3.168) does not set $\tilde{n}_\mu$ to a pure gauge and it allows configurations with nontrivial holonomies $\sum_{\text{links}} n_\mu$ around closed cycles. In other words, (3.168) is not a $\mathbb{Z}_N$ clock-model but a $\mathbb{Z}_N$ clock-model coupled to a topological lattice $\mathbb{Z}_N$ gauge theory [26, 29]. The latter is described by the second term in (3.168) and will be further discussed in Appendix C.2.

We conclude that the T-duality of the underlying XY-model (3.131) leads to the Kramers-Wannier duality of the clock-model (3.161). In fact, while the T-duality is correct both locally and globally, the Kramers-Wannier duality of the clock-model is valid also globally only when a lattice topological theory is included in one side of the duality.

Long-distance limit

Here, we study the long-distance limit of the theory based on (3.163).

As in the discussion around Figure 3.1 we start with the theory with $\lambda = \tilde{\lambda} = 0$. It flows to the compact-boson theory, which is represented by the blue line in Figure 3.1. Then, for small enough λ and $\tilde{\lambda}$ we can perturb this conformal theory by these two perturbations. The momentum breaking operator $\cos(N\phi)$ is irrelevant for $R < \frac{N}{\sqrt{8}}$ and the winding breaking operator is irrelevant for $R > \frac{\sqrt{2}}{W}$. Therefore, for $NW \geq 4$ there are values of R, or equivalently of $\beta = \frac{R^2}{\pi}$, such that the compact-boson conformal field theory is robust under deformations with small λ and $\tilde{\lambda}$. This happens for

$$\begin{aligned} \frac{\sqrt{2}}{W} &\leq R \leq \frac{N}{\sqrt{8}} \\ \frac{2}{\pi W^2} &\leq \beta \leq \frac{N^2}{8\pi} \end{aligned} \tag{3.170}$$

and then the long distance theory is gapless. Note that this is consistent with the duality $\beta \leftrightarrow \frac{1}{(2\pi)^2\beta}$, which is accompanied with $N \leftrightarrow W$.

In the most studied case of $W = 1$, the long distance theory of (3.163) is given by the compact scalar CFT for $N \geq 4$. For $N = 4$ and $R = \sqrt{2}$ it is the CFT of the Kosterlitz-Thouless point. And for $N \geq 5$ and

$$\begin{aligned} \sqrt{2} &\leq R \leq \frac{N}{\sqrt{8}} \\ \frac{2}{\pi} &\leq \beta \leq \frac{N^2}{8\pi} \end{aligned} \tag{3.171}$$

it is the line of a CFT with this value of R. For other values of R the theory is gapped. Note, as a check that this is consistent with the $R \leftrightarrow \frac{N}{2R}$ duality of the local dynamics, which we discussed in Appendix B.2.

For $N = 2$ and $N = 3$ the duality determines that the theory has two gapped phases separated by a CFT at $R = 1$ and $R = \sqrt{\frac{3}{2}}$, respectively. However, these CFTs are not the CFT of the compact boson, but are of the Ising and 3-states Potts model.

We should emphasize that this discussion of the clock-model is specific to the action (3.163). For other actions, the gapless phase could be different or even absent. See the discussion in [121, 122, 124].

3.6 Appendix C: Modified Villain formulation of p-form lattice gauge theory in diverse dimensions

In this appendix, we will study p-form gauge theories on a d-dimensional Euclidean space for $p \leq d-1$ (see [125] for a review on these models). The models in Appendix A, correspond to $d=1$ and $p=0$ and perhaps do not deserve to be called gauge theories. The models in Appendix B, correspond to $d=2$ and $p=0$.

As above, the lattice spacing is a, and there are L^μ sites in the μ direction. Throughout this discussion, $A^{(p)}$ denotes a p-form field placed on the p-cells of the lattice, and $\tilde{B}^{(d-p)}$ denotes a $(d-p)$-form field placed on the dual $(d-p)$-cells.

3.6.1 Appendix C.1: $U(1)$ gauge theory

Let us place $U(1)$ variables $e^{ia^{(p)}}$ on p-cells of the d-dimensional Euclidean lattice. The standard action of this gauge field is

$$\beta \sum_{(p+1)\text{-cell}} [1 - \cos(\Delta a^{(p)})] \ , \tag{3.172}$$

where $\Delta a^{(p)}$ is a $(p+1)$-form given by the oriented sum of $a^{(p)}$ along the p-cells in the boundary of the $(p+1)$-cell, and $a^{(p)}$ is circle-valued with gauge symmetry

$$e^{ia^{(p)}} \sim e^{ia^{(p)} + i\Delta\alpha^{(p-1)}} \ , \tag{3.173}$$

where $\alpha^{(p-1)}$ is circle-valued. At large β, the action can be approximated by the Villain action [126–128]

$$\frac{\beta}{2} \sum_{(p+1)\text{-cell}} (\Delta a^{(p)} - 2\pi n^{(p+1)})^2 \ , \tag{3.174}$$

where now $a^{(p)}$ is real and $n^{(p+1)}$ is integer-valued. We can interpret $n^{(p+1)}$ as the $\mathbb{Z}$ gauge field that makes $a^{(p)}$ compact because of the gauge symmetry

$$\begin{aligned} a^{(p)} &\sim a^{(p)} + \Delta\alpha^{(p-1)} + 2\pi k^{(p)} \ , \\ n^{(p+1)} &\sim n^{(p+1)} + \Delta k^{(p)} \ . \end{aligned} \tag{3.175}$$

For $p \leq d-2$, nonzero $\Delta n^{(p+1)}$ corresponds to monopoles or vortices. They can be suppressed by modifying (3.174) to

$$\frac{\beta}{2} \sum_{(p+1)\text{-cell}} (\Delta a^{(p)} - 2\pi n^{(p+1)})^2 + i \sum_{(p+2)\text{-cell}} \tilde{a}^{(d-p-2)} \Delta n^{(p+1)} \ , \tag{3.176}$$

where $\tilde{a}^{(d-p-2)}$ is a real-valued $(d-p-2)$-form field, which acts as a Lagrange multiplier imposing the flatness constraint of $n^{(p+1)}$. We will refer to (3.176) as the modified Villain action of the $U(1)$ p-form gauge theory. In addition to (3.175), this theory also has a gauge symmetry

$$\tilde{a}^{(d-p-2)} \sim \tilde{a}^{(d-p-2)} + \Delta\tilde{\alpha}^{(d-p-3)} + 2\pi \tilde{k}^{(d-p-2)} \ , \tag{3.177}$$

where $\tilde{\alpha}^{(d-p-3)}$ is real-valued, and $\tilde{k}^{(d-p-2)}$ is integer-valued.

For $p = d-1$ we cannot write (3.176). Instead, in this case we can add another term

$$\frac{\beta}{2} \sum_{d\text{-cell}} (\Delta a^{(d-1)} - 2\pi n^{(d)})^2 + i\theta \sum_{d\text{-cell}} n^{(d)} \ . \tag{3.178}$$

This is a $U(1)$ gauge theory of a $(d-1)$-form gauge field with a θ-parameter. (Compare with (3.106) and (3.108), which corresponds to $p = 0$ and $d = 1$.) Note that this is a lattice version of the gauge theory with θ. Unlike the continuum presentation, here, the θ-term is associated with the integer-valued field. The topological charge $\sum_{d\text{-cell}} n^{(d)}$ is manifestly quantized and therefore $\theta \sim \theta + 2\pi$.

Duality

Using the Poisson resummation formula (3.6), we can dualize the modified Villain action (3.176) of a p-form gauge theory to the modified Villain action of a $(d-p-2)$-form gauge theory

$$\frac{1}{2(2\pi)^2\beta} \sum_{(p+1)\text{-cell}} (\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)})^2 + i(-1)^{d-p} \sum_{(p+1)\text{-cell}} \tilde{n}^{(d-p-1)}\Delta a^{(p)} \ , \tag{3.179}$$

where $\tilde{n}^{(d-p-1)}$ is integer-valued. We can interpret $\tilde{n}^{(d-p-1)}$ as a $\mathbb{Z}$ gauge field that makes $\tilde{a}^{(d-p-2)}$ compact because of the gauge symmetry

$$\begin{aligned} &\tilde{a}^{(d-p-2)} \sim \tilde{a}^{(d-p-2)} + \Delta\tilde{\alpha}^{(d-p-3)} + 2\pi\tilde{k}^{(d-p-2)} \ , \\ &\tilde{n}^{(d-p-1)} \sim \tilde{n}^{(d-p-1)} + \Delta\tilde{k}^{(d-p-2)} \ . \end{aligned} \tag{3.180}$$

The field $a^{(p)}$ is a Lagrange multiplier that imposes the flatness constraint of $\tilde{n}^{(d-p-1)}$. When d is even, and $p = \frac{d-2}{2}$, the model (3.176) is self-dual with $\beta \leftrightarrow \frac{1}{(2\pi)^2\beta}$.

Global symmetries

In all the three models, (3.172), (3.174), and (3.176), there is a p-form *electric symmetry* [5], which acts on the fields as

$$a^{(p)} \to a^{(p)} + \lambda^{(p)} \ , \tag{3.181}$$

where $\lambda^{(p)}$ is a real-valued, flat p-form field. Due to the gauge symmetry (3.175), the electric symmetry is $U(1)$ rather than $\mathbb{R}$. In (3.174) and (3.176), the Noether current of electric symmetry is[19]

$$J_e^{(p+1)} = i\beta(\Delta a^{(p)} - 2\pi n^{(p+1)}) = \frac{(-1)^{d-p}}{2\pi} \star (\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)}) \ , \tag{3.182}$$

which is conserved because of the equation of motion of $a^{(p)}$. The electric charge is

$$Q_e(\tilde{\mathcal{M}}^{(d-p-1)}) = \sum_{\text{dual } (d-p-1)\text{-cell} \in \tilde{\mathcal{M}}^{(d-p-1)}} \star J_e^{(p+1)} \ , \tag{3.183}$$

[19] The Hodge dual $\star A^{(p)}$ is a $(d-p)$-form field on the dual $(d-p)$-cells of the lattice.

where $\tilde{\mathcal{M}}^{(d-p-1)}$ is a codimension-$(p+1)$ submanifold along the dual $(d-p-1)$-cells of the lattice. The electrically charged objects are the Wilson observables

$$W_e(\mathcal{M}^{(p)}) = \exp\left[i \sum_{p\text{-cell}\in\mathcal{M}^{(p)}} a^{(p)}\right] , \tag{3.184}$$

where $\mathcal{M}^{(p)}$ is a dimension-p submanifold along the p-cells of the lattice.

The theory (3.176) (but not (3.172) or (3.174)) also has a $(d-p-2)$-form *magnetic symmetry* [5], which acts on the fields as

$$\tilde{a}^{(d-p-2)} \to \tilde{a}^{(d-p-2)} + \tilde{\lambda}^{(d-p-2)} , \tag{3.185}$$

where $\tilde{\lambda}^{(d-p-2)}$ is a real-valued, flat $(d-p-2)$-form. Due to the gauge symmetry (3.177), the magnetic symmetry is $U(1)$. The Noether current of magnetic symmetry is[20]

$$J_m^{(d-p-1)} = -\frac{i}{(2\pi)^2\beta} \star\star(\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)}) = \frac{(-1)^{d-p}}{2\pi} \star(\Delta a^{(p)} - 2\pi n^{(p+1)}) , \tag{3.186}$$

which is conserved because of the equation of motion of $\tilde{a}^{(d-p-2)}$. The magnetic charge is

$$Q_m(\mathcal{M}^{(p+1)}) = \sum_{(p+1)\text{-cell}\in\mathcal{M}^{(p+1)}} \star J_m^{(d-p-1)} , \tag{3.187}$$

where $\mathcal{M}^{(p+1)}$ is a dimension-$(p+1)$ submanifold along the $(p+1)$-cells of the lattice. The magnetically charged objects are the 't Hooft observables

$$W_m(\tilde{\mathcal{M}}^{(d-p-2)}) = \exp\left[i \sum_{\text{dual } (d-p-2)\text{-cell}\in\tilde{\mathcal{M}}^{(d-p-2)}} \tilde{a}^{(d-p-2)}\right] , \tag{3.188}$$

where $\tilde{\mathcal{M}}^{(d-p-2)}$ is a codimension-$(p+2)$ submanifold along the dual $(d-p-2)$-cells of the lattice.

[20] Recall that $\star\star A^{(p)} = (-1)^{p(d-p)} A^{(p)}$.

Long-distance limit

In the continuum limit, the modified Villain model (3.176) becomes a gapless continuum p-form gauge theory

$$\frac{1}{2g^2}\int d^dx\ (da^{(p)})^2\ . \tag{3.189}$$

This can be derived, as above, by choosing a convenient gauge where most of the integer-valued fields vanish and then redefining the real lattice variables appropriately.[21]

An important question is whether the lattice gauge theory (3.172), or equivalently its Villain version (3.174), flow at long distances to the same gapless theory (3.189). Unlike the modified Villain model, these two lattice models have only the electric symmetry, but no magnetic symmetry. So without fine-tuning, the long-distance theory is generically deformed by the 't Hooft operators. For the deformation to be possible, the 't Hooft operators have to be local, point-like operators. This is the case only for $p = d-2$. This is obvious in its dual version where the dual field is a scalar and the monopole operator gives it a mass. This implies that without fine-tuning a d-dimensional p-form lattice gauge theory can flow to a gapless p-form gauge theory at long distance unless $p = d-2$, in which case, the theory is generically gapped at long distance. This is the famous Polyakov mechanism [93].

We conclude that for $p = d-2$, where the standard $U(1)$ lattice gauge theory is gapped, the modification of the lattice gauge theory (3.176) keeps it massless.

3.6.2 Appendix C.2: $\mathbb{Z}_N$ gauge theory

We now describe a d-dimensional Villain $\mathbb{Z}_N$ p-form gauge theory [119, 129]. On each p-cell, there is an integer field $m^{(p)}$ and on each $(p+1)$-cell, there is an integer field $n^{(p+1)}$. The action is

$$\frac{\beta(2\pi)^2}{2N^2}\sum_{(p+1)\text{-cell}}(\Delta m^{(p)} - Nn^{(p+1)})^2\ , \tag{3.190}$$

[21] The continuum theory can also have additional θ-parameters associated with various characteristic classes of the gauge field. Our lattice formulation leads to the term $\frac{\theta}{2\pi}da^{(p)}$ for $p = d-1$, but not to the other θ-parameters.

with the integer gauge symmetry

$$\begin{aligned} &m^{(p)} \sim m^{(p)} + \Delta\ell^{(p-1)} + Nk^{(p)} \ , \\ &n^{(p+1)} \sim n^{(p+1)} + \Delta k^{(p)} \ . \end{aligned} \tag{3.191}$$

The theory has an electric $\mathbb{Z}_N$ p-form global symmetry [5], which shifts $m^{(p)}$ by a flat integer p-form field.

In the limit $\beta \to \infty$, the field strength obeys $\Delta m = 0 \bmod N$ [130, 131], and we can replace the action by

$$\frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ , \tag{3.192}$$

where $\tilde{n}^{(d-p-1)}$ is an integer-valued field with the integer gauge symmetry

$$\tilde{n}^{(d-p-1)} \sim \tilde{n}^{(d-p-1)} + \Delta\tilde{k}^{(d-p-2)} + N\tilde{q}^{(d-p-1)} \ . \tag{3.193}$$

This describes a *topological* $\mathbb{Z}_N$ *lattice gauge theory* [26, 29]. The action (3.192) is similar to the one in [29] except that the fields there are $\mathbb{Z}_N$ variables while here we use $\mathbb{Z}$ variables with $N\mathbb{Z}$ gauge symmetry.

Duality

As in Appendix B.2, we can dualize the $\mathbb{Z}_N$ p-form gauge theory (3.190) by dualizing the integer field $n^{(p+1)}$ to an integer field $\tilde{n}^{(d-p-1)}$:

$$\frac{1}{2\beta} \sum_{\text{dual } (d-p-1)\text{-cell}} (\tilde{n}^{(d-p-1)})^2 + \frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ . \tag{3.194}$$

For $p \le d-1$, we can introduce new gauge symmetries together with Stueckelberg fields, and write the action as

$$\frac{1}{2\beta} \sum_{\text{dual } (d-p-1)\text{-cell}} (\Delta\tilde{m}^{(d-p-2)} - N\hat{n}^{(d-p-1)} - \tilde{n}^{(d-p-1)})^2 + \frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ . \tag{3.195}$$

with the integer gauge symmetry

$$
\begin{aligned}
&\tilde{m}^{(d-p-2)} \sim \tilde{m}^{(d-p-2)} + \Delta\tilde{\ell}^{(d-p-3)} + \tilde{k}^{(d-p-2)} \ , \\
&\hat{n}^{(d-p-1)} \sim \hat{n}^{(d-p-1)} - \tilde{q}^{(d-p-1)} \ , \\
&\tilde{n}^{(d-p-1)} \sim \tilde{n}^{(d-p-1)} + \Delta\tilde{k}^{(d-p-2)} + N\tilde{q}^{(d-p-1)} \ , \\
&m^{(p)} \sim m^{(p)} + \Delta\ell^{(p-1)} + Nk^{(p)} \ .
\end{aligned}
\tag{3.196}
$$

The duality maps a p-form gauge theory with coefficient $\frac{2\pi^2\beta}{N^2}$ to a $(d-p-2)$-form gauge theory with coefficient $\frac{1}{2\beta}$ that couples to a topological $\mathbb{Z}_N$ $(d-p-1)$-form gauge theory. For $d=2$ and $p=0$, this reduces to the Kramers-Wannier duality of the $\mathbb{Z}_N$ clock model reviewed in Appendix B.2. The duality of the $d=3$ and $p=1$ system is the famous duality of the 3d clock model [132, 125, 133] and for $d=4$ and $p=1$ it is the famous self-duality of [132, 125, 134, 133, 119, 129].

Real BF-action and the continuum limit

This theory can be described using several different actions. Here we describe some actions using real fields that are similar to various continuum actions.

We start with the integer BF-action (3.192) and replace the integer-valued gauge fields $m^{(p)}$ and $\tilde{n}^{(d-p-1)}$ with real-valued gauge fields $a^{(p)}$ and $\tilde{b}^{(d-p-1)}$. We constrain these real-valued fields to integer values by adding integer-valued fields $\tilde{m}^{(d-p)}$ and $n^{(p+1)}$. Furthermore, since the gauge fields $a^{(p)}$ and $\tilde{b}^{(d-p-1)}$ have real-valued gauge symmetries instead of integer-valued gauge symmetries, we introduce Stueckelberg fields $\phi^{(p-1)}$ and $\tilde{\phi}^{(d-p-2)}$ for the gauge symmetries. We end up with the action

$$
\begin{aligned}
&\frac{iN}{2\pi}\sum_{p\text{-cell}} a^{(p)}\left(\Delta\tilde{b}^{(d-p-1)} - 2\pi\tilde{m}^{(d-p)}\right) + i(-1)^p N \sum_{(p+1)\text{-cell}} n^{(p+1)}\tilde{b}^{(d-p-1)} \\
&\quad - i(-1)^p \sum_{(p+1)\text{-cell}} n^{(p+1)}\Delta\tilde{\phi}^{(d-p-2)} + i\sum_{p\text{-cell}} \Delta\phi^{(p-1)}\tilde{m}^{(d-p)} \ .
\end{aligned}
\tag{3.197}
$$

We will refer to this presentation of the model as the real BF-action, which uses both real and integer fields.

As a check, summing over $\tilde{m}^{(d-p)}$ and $n^{(p+1)}$ constrains

$$a^{(p)} - \frac{1}{N}\Delta\phi^{(p-1)} = \frac{2\pi}{N}m^{(p)} \ , \qquad \tilde{b}^{(d-p-1)} - \frac{1}{N}\Delta\tilde{\phi}^{(d-p-2)} = \frac{2\pi}{N}\tilde{n}^{(d-p-1)} \ , \tag{3.198}$$

where $m^{(p)}$ and $\tilde{n}^{(d-p-1)}$ are integer-valued fields. Substituting them into (3.197), we recover the action (3.192).

The action (3.197) has the gauge symmetry

$$\begin{aligned}
&a^{(p)} \sim a^{(p)} + \Delta\alpha^{(p-1)} + 2\pi k^{(p)} \ , \\
&\tilde{b}^{(d-p-1)} \sim \tilde{b}^{(d-p-1)} + \Delta\tilde{\beta}^{(d-p-2)} + 2\pi\tilde{q}^{(d-p-1)} \ , \\
&n^{(p+1)} \sim n^{(p+1)} + \Delta k^{(p)} \ , \\
&\tilde{m}^{(d-p)} \sim \tilde{m}^{(d-p)} + \Delta\tilde{q}^{(d-p-1)} \ , \\
&\phi^{(p-1)} \sim \phi^{(p-1)} + \Delta\gamma^{p-2} + N\alpha^{(p-1)} + 2\pi k_{\phi}^{(p-1)} \ , \\
&\tilde{\phi}^{(d-p-2)} \sim \tilde{\phi}^{(d-p-2)} + \Delta\tilde{\gamma}^{(d-p-3)} + N\tilde{\beta}^{(d-p-2)} + 2\pi\tilde{q}_{\tilde{\phi}}^{(d-p-2)} \ ,
\end{aligned} \tag{3.199}$$

where $\alpha^{(p-1)}, \tilde{\beta}^{(d-p-2)}, \gamma^{(p-2)}, \tilde{\gamma}^{(d-p-3)}$ are real-valued and $k_{\phi}^{(p-1)}, \tilde{q}_{\tilde{\phi}}^{(d-p-2)}$ are integer-valued.

Another action is obtained by replacing $\tilde{m}^{(d-p)}$ by a real-valued field $\tilde{F}^{(d-p)} + \Delta\tilde{b}^{(d-p-1)}$ and adding an integer-valued field $n^{(p)}$ to constrain it. This leads to the action

$$\begin{aligned}
&\frac{i}{2\pi}\sum_{p\text{-cell}}(\Delta\phi^{(p-1)} - Na^{(p)} - 2\pi n^{(p)})\tilde{F}^{(d-p)} + i(-1)^p \sum_{(p+1)\text{-cell}}(\Delta n^{(p)} + Nn^{(p+1)})\tilde{b}^{(d-p-1)} \\
&\quad - i(-1)^p \sum_{(p+1)\text{-cell}} n^{(p+1)}\Delta\tilde{\phi}^{(d-p-2)} \ .
\end{aligned} \tag{3.200}$$

These fields have the same gauge symmetries as in (3.199). In addition, the gauge symmetries also act on $n^{(p)}$

$$n^{(p)} \sim n^{(p)} + \Delta k_{\phi}^{(p-1)} - Nk^{(p)} \ . \tag{3.201}$$

We can interpret the action (3.200) as Higgsing the $U(1)$ gauge theory of $a^{(p)}$ to a $\mathbb{Z}_N$ theory using the fields $\phi^{(p-1)}$ of charge N.

Alternatively, we can integrate out $\phi^{(p-1)}, \tilde{\phi}^{(d-p-2)}$ which constrain $n^{(p+1)}, \tilde{m}^{(d-p)}$ to be flat gauge fields. Using the gauge symmetry of $k^{(p)}, \tilde{q}^{(d-p-1)}$, we can gauge fix $n^{(p+1)}, \tilde{m}^{(d-p)}$ to be zero almost everywhere except at a few cells that capture the holonomy. The residual gauge symmetry shifts $a^{(p)}$ and $\tilde{b}^{(d-p-1)}$ by 2π multiples of flat integer gauge fields. Let us define two new fields $\bar{a}^{(p)}, \bar{\tilde{b}}^{(d-p-1)}$ such that

$$\begin{aligned} \Delta\bar{a}^{(p)} &= \Delta a^{(p)} - 2\pi n^{(p+1)} \ , \\ \Delta\bar{\tilde{b}}^{(d-p-1)} &= \Delta\tilde{b}^{(d-p-1)} - 2\pi\tilde{m}^{(d-p)} \ , \end{aligned} \tag{3.202}$$

and $\bar{a}^{(p)} = a^{(p)}$, $\bar{\tilde{b}}^{(d-p-1)} = \tilde{b}^{(d-p-1)}$ almost everywhere. Although $a^{(p)}, \tilde{b}^{(d-p-1)}$ are single-valued fields, $\bar{a}^{(p)}, \bar{\tilde{b}}^{(d-p-1)}$ can have nontrivial transition functions. In terms of the new variables, the Euclidean action is

$$\frac{iN}{2\pi}\sum_{p\text{-cell}} \bar{a}^{(p)}\Delta\bar{\tilde{b}}^{(d-p-1)} + i(-1)^p N \sum_{(p+1)\text{-cell}} n^{(p+1)}\bar{\tilde{b}}^{(d-p-1)} \ , \tag{3.203}$$

where $n^{(p+1)}$ vanishes almost everywhere except at a few $(p+1)$-cells, which encode the information in the transition function of $\bar{a}^{(p+1)}$. For $d = 1$ and $p = 0$, the action (3.203) reduces to the quantum mechanics action (3.120).

The real BF-action is closely related to the continuum field theory limit. In this gauge choice, the continuum limit is

$$\frac{iN}{2\pi}\int a^{(p)} d\tilde{b}^{(d-p-1)} \ , \tag{3.204}$$

where we dropped the bars on $a^{(p)}$ and $\tilde{b}^{(d-p-1)}$ and rescaled them by appropriate powers of the lattice spacing a. We also omitted here the terms that depend on the transition functions of $a^{(p)}$ and $\tilde{b}^{(d-p-1)}$. As in (3.117), these terms are actually essential in order to make (3.204) globally well defined. Here $a^{(p)}$ is a $U(1)$ p-form gauge field and $\tilde{b}^{(d-p-1)}$ is a $U(1)$ $(d-p-1)$-form gauge field. This is the known continuum action of the $\mathbb{Z}_N$ p-form gauge theory [32, 33, 29].

Relation to the toric code

We now review the well-known fact that the low-energy limit of the $\mathbb{Z}_N$ toric code [135] is described by the topological $\mathbb{Z}_N$ lattice gauge theory [26, 29], which in turn is given by the continuum $\mathbb{Z}_N$ gauge theory.

Consider the $\mathbb{Z}_N$ toric code on a 2d periodic square lattice. On each link, there is a $\mathbb{Z}_N$ variable U and its conjugate variable V. They obey $UV = e^{2\pi i/N} VU$ and $U^N = V^N = 1$. The Hamiltonian consists of two commuting terms G and L:

$$H_{\text{toric}} = -\beta_1 \sum_{\text{site}} G - \beta_2 \sum_{\text{plaq}} L + c.c. \ , \tag{3.205}$$

where G is an oriented product of V and $V^\dagger$ around a site and L is an oriented product of U and $U^\dagger$ around a plaquette.

The ground states satisfy $G = L = 1$ for all sites and plaquettes, while the excited states violate some of these conditions. It is common to refer to the dynamical excitations that violate only $G = 1$ at a site as the electrically-charged excitations and those that violate only $L = 1$ at a plaquette as the magnetically-charged excitations.

The toric code has a large non-relativistic electric and magnetic $\mathbb{Z}_N$ one-form symmetry (in the sense of [79]). The symmetries are generated respectively by the closed loop operator W_e made of V and $V^\dagger$, and the closed loop operator W_m made of U and $U^\dagger$. Unlike the relativistic one-form symmetry of [5], these symmetry operators are not topological, i.e., they are not invariant under small deformations.

In the $\beta_1, \beta_2 \to \infty$ limit, the Hilbert space is restricted to the ground states, which satisfy $G = L = 1$ for all sites and plaquettes. In the restricted Hilbert space, there are no electrically-charged or magnetically-charged excitations. So, the closed loop operators W_e and W_m are topological, and they generate a relativistic electric and magnetic $\mathbb{Z}_N$ symmetry, respectively.

Consider the toric code in the $\beta_1, \beta_2 \to \infty$ limit in the Lagrangian formalism on a 3d Euclidean lattice. For each spatial link along the $i = x, y$ direction, we introduce an integer field m_i for the $\mathbb{Z}_N$ variable $U = \exp(\frac{2\pi i}{N} m_i)$, and an integer field $\tilde{n}_j$ for the conjugate $\mathbb{Z}_N$ variable $V = \exp(\frac{2\pi i}{N} \epsilon^{ij} \tilde{n}_j)$. The field $\tilde{n}_j$ naturally lives on the dual links along the j direction.

To impose the constraints $G = L = 1$, we introduce two integer-valued Lagrange multiplier fields. On each τ-link, we introduce an integer field m_τ to impose $G = 1$, or equivalently $\epsilon^{ij} \Delta_i \tilde{n}_j = 0 \text{ mod } N$. On each dual τ-link (or equivalently each xy-plaquette), we introduce an integer field $\tilde{n}_\tau$ to impose $L = 1$, or equivalently $\epsilon^{ij} \Delta_i m_j = 0 \text{ mod } N$. In terms of these integer fields, the Euclidean action of the system is precisely the topological $\mathbb{Z}_N$ lattice gauge theory (3.192) with $d = 3$ and $p = 1$.

Chapter 4

Anomalies in the Space of Coupling Constants

4.1 Preliminary and Summary

't Hooft anomalies lead to powerful constraints on the dynamics and phases of quantum field theory (QFT). They also control the properties of boundaries, extended excitations like strings and domain walls, and various defects.

't Hooft anomalies do not signal an inconsistency of the theory. Instead, they show that some contact terms cannot satisfy the Ward identities of global symmetries. More generally, they are an obstruction to coupling the system to classical background gauge fields for these symmetries.

In this chapter we generalize the notion of 't Hooft anomalies to the space of coupling constants. In addition to coupling the system to classical background gauge fields, we also make the various coupling constants spacetime dependent, i.e. we view them as background fields. The generalized 't Hooft anomalies are an obstruction to making the coupling constants and the various gauge fields spacetime dependent.

As with the ordinary 't Hooft anomalies, we use these generalized anomalies to constrain the phase diagram of the theory as a function of its parameters and to learn about defects constructed by position-dependent coupling constants.

4.1.1 Anomalies and Symmetries

A useful point of view of 't Hooft anomalies is to couple a system with a global symmetry to an appropriate background gauge field A. Here A denotes a fixed classical source and leads to a partition function $Z[A]$. Depending on the context, A could be a standard background connection for an ordinary continuous (0-form) global symmetry, or an appropriate background field for more subtle concepts of symmetry such as a discrete gauge field for a discrete global symmetry, a higher-form gauge field for a higher-form symmetry [5], or a Riemannian metric for (Wick rotated) Poincaré symmetry. Additionally, the partition function may depend on discrete topological data such as a choice of spin structure in a theory with fermions or an orientation on spacetime. We will denote all this data by A.

Naively one expects that the resulting partition function $Z[A]$ should be gauge invariant under appropriate background gauge transformations. An 't Hooft anomaly is a mild violation of this expectation. Denoting a general gauge transformation with gauge parameter λ (or coordinate transformation) as $A \to A^\lambda$, the partition function $Z[A]$ is in general not gauge invariant. Instead, it transforms by a phase, which is a local functional of the gauge parameter λ and the gauge fields A

$$Z[A^\lambda] = Z[A] \exp\left(-2\pi i \int_X \alpha(\lambda, A)\right) \ , \tag{4.1}$$

where X is our d-dimensional spacetime.

The partition function $Z[A]$ is subject to a well-known ambiguity. Different regularization schemes can lead to different answers. This ambiguity can be absorbed in adding local counterterms to the action. These counterterms can depend on the dynamical fields and on background sources. This freedom in adding counterterms is the same as performing a redefinition in the space of coupling constants. A special case of such counterterms are those that multiply the unit operator, i.e. they depend only on classical backgrounds A. We refer to these terms as classical counterterms or sometimes simply as counterterms when the context is clear. An essential part of our discussion will involve such classical counterterms. The 't Hooft anomaly for the global symmetry is what remains of the phase in (4.1) after taking into account this freedom.[1]

Thus, the set of possible 't Hooft anomalies for a given global symmetry is defined by a cohomology problem of local phases consistent with the equation (4.1) modulo the variation of local functionals of the gauge field A.

It is convenient to describe anomalies using a classical, local action for the gauge fields A in $(d+1)$-spacetime dimensions. Such actions are also referred to as invertible field theories.[2] In this presentation the d-dimensional manifold X supporting the dynamical field theory is viewed as the boundary of a $(d+1)$-manifold Y, and we extend the classical gauge field sources A to the manifold Y. On Y there is a local, classical Lagrangian $-2\pi i\omega(A)$ with the property that

$$\exp\left(2\pi i\int_Y \omega(A^\lambda) - 2\pi i\int_Y \omega(A)\right) = \exp\left(2\pi i\int_X \alpha(\lambda, A)\right) \ . \tag{4.2}$$

Thus on closed $(d+1)$-manifolds the action ω defines a gauge-invariant quantity, while on manifolds with boundary it reproduces the anomaly.[3] We refer to $\omega(A)$ as the Lagrangian of the anomaly theory and we define the partition function of the anomaly theory as

$$\mathcal{A}[A] = \exp\left(2\pi i\int \omega(A)\right) \ . \tag{4.3}$$

[1] As noted in [8], one can always remove the anomalous phase by adding a d-form background field $A^{(d)}$ with a coupling $i\int_X A^{(d)}$. $A^{(d)}$ can be thought of as a background gauge field for a "$d-1$-form symmetry" that does not act on any dynamical field. (Such couplings are common in the study of branes in string theory.) Then the anomaly is removed by postulating that under gauge transformations of the background fields it transforms as $A^{(d)} \to A^{(d)} + d\lambda^{(d-1)} - 2\pi i\alpha(\lambda, A)$. The term with $\lambda^{(d-1)}$ is the standard gauge transformation of such a gauge field and the term with α, which cancels (4.1), reflects a higher-group symmetry. See e.g. [8–10] and references therein.

[2] In condensed matter physics, symmetry protected topological orders (SPTs) are also characterized at low energies by such actions. Depending on the precise definitions and context, "SPT" may be synonymous with "invertible field theory", or may instead refer to the deformation class of an invertible field theory, i.e. the equivalence class of invertible theories obtained by continuously varying parameters.

[3] In certain cases, there is no Y such that $\partial Y = X$ and A on X is extended into Y. Then, one can construct an anomaly free partition function by assuming that X is a component of the boundary of Y and Y has additional boundary components.

Using these observations, we can present another point of view on the partition function of a theory with a 't Hooft anomaly. We can introduce a modified partition function as follows:

$$\tilde{Z}[A] \equiv Z[A] \exp\left(2\pi i \int_Y \omega(A)\right) \ . \tag{4.4}$$

In (4.4), the manifold Y is again an extension of spacetime. Using the transformation law (4.1) and the definition (4.2) of ω we conclude that the partition function is exactly gauge invariant

$$\tilde{Z}[A^\lambda] = \tilde{Z}[A] \ . \tag{4.5}$$

The price we have paid is that the partition function now depends on the extension of the classical fields into the bulk. In some condensed matter applications, this added bulk Y is physical. The system X is on a boundary of a space Y in a non-trivial SPT phase. The 't Hooft anomaly of the boundary theory is provided by inflow from the nontrivial bulk Y. This is known as anomaly inflow and was first described in [13]. (See also [136].)

Although the partition functions Z and $\tilde{Z}$ are different, an essential observation is that, for $A = 0$, they encode the same correlation functions at separated points. It is this data that we view as the intrinsic defining information of a quantum field theory. However, one advantage of the presentation of the theory using $\tilde{Z}$ is that it clarifies the behavior of the anomaly under renormalization group flow.

First, such a transformation can modify the scheme used to define the theory in a continuous fashion. This means that in general, d-dimensional counterterms are modified along the flow. Second, we can also ask about the behavior of the classical anomaly action ω which resides in $(d+1)$ dimensions. If we view this term as arising from the long distance behavior of massive degrees of freedom (a choice of scheme) then along renormalization group flow we can continuously adjust the details of these heavy degrees of freedom and hence ω could evolve continuously as well. Thus, a renormalization group invariant quantity is the deformation class of the action ω, i.e. all actions that may be obtained from ω by continuous deformations.

In the applications to follow, we therefore focus on physical conclusions that depend only on the deformation class of ω. (We will also see that theories related by renormalization group flow can produce different expressions for ω in the same deformation class). In particular, any theory with an anomaly action ω that is not continuously connected to the trivial action cannot flow at long distances to a trivially gapped theory with a unique vacuum and no long-range degrees of freedom.[4] It is this feature of 't Hooft anomalies that makes them powerful tools to study the dynamics of quantum field theories.

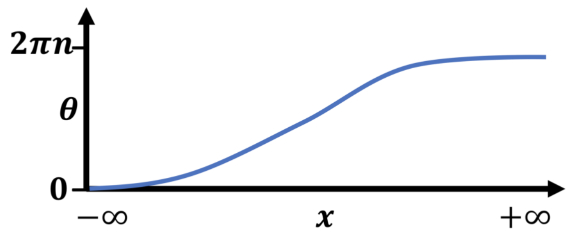
In this chapter, we generalize the notions above to the space of parameters of a QFT. We will describe how certain subtle phenomena can be viewed as a generalization of the concept of anomalies from the arena of global symmetries to this broader class of sources. In particular we will see how such anomalies of d-dimensional theories can also be summarized in terms of classical theories in $d+1$-dimensions. We will use this understanding to explore phase transitions as the parameters vary and properties of defects that are associated with spacetime dependent coupling constants.

Our analysis extends previous work on this subject in [137–141]. (For a related discussion in another context see e.g. [142].) Finally, we would like to point out that an anomaly in making certain coupling constant background superfields was discussed in [143, 144]. It would be nice to phrase these anomalies and ours in a uniform framework.

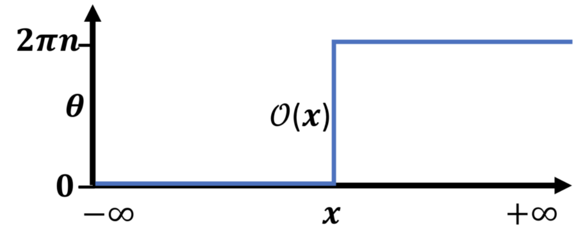
4.1.2 Anomalies in Parameter Space: Defects

Instead of phrasing the analysis above in terms of background fields, it is often convenient to formulate the discussion in terms of defects and extended operators. Indeed, an ordinary global symmetry implies the existence of codimension one operators that implement the symmetry action. This paradigm also extends to other forms of internal symmetry: for instance p-form global symmetries are encoded in extended operators of codimension $p+1$ [5]. Geometrically,

[4] A trivially gapped theory by definition has a gap in its spectrum of excitations and its long distance behavior is particularly simple. In particular, it has a single ground state on any space of finite volume. This means that it does not have even topological degrees of freedom at low energies. In this case the low-energy theory is a classical theory of the background fields also referred to as an invertible theory.



(a) Smooth Interface



(b) Sharp Interface

Figure 4.1: Interfaces defined by spatially varying coupling constant $\theta(x)$. In (a), the variation is smooth and the resulting interface dynamics are universal. In (b), the variation is abrupt. The resulting worldvolume dynamics is not universal and can be modified by coupling to degrees of freedom on the interface (schematically denoted $\mathcal{O}(x)$ above. As we will discuss, for certain special choices of $\mathcal{O}(x)$, an abrupt interface can be made completely transparent.

these symmetry defects are Poincaré dual to the flat background gauge fields described above. These extended operators have the property that they are topological: small deformations of their positions do not modify correlation functions.

Many of the implications of 't Hooft anomalies are visible when we consider correlation functions of these extended operators. In this context 't Hooft anomalies arise as mild violations (by phases) of the topological nature of the symmetry defects. This perspective points the way to a natural generalization of the concept of anomalies to the space of coupling constants in quantum field theories. We promote the parameters of a theory to be spacetime-dependent and explore the properties of the resulting topologically non-trivial extended objects.

An important example, which will occur repeatedly below, is a circle-valued parameter such as a θ-angle in gauge theory. This can be made to depend on a single spacetime coordinate x, and wind around the circle as x varies from $-\infty$ to $+\infty$ (or around a nontrivial compact cycle in spacetime). If the bulk theory is trivially gapped, i.e. does not even have topological order (as in e.g. 4d $SU(N)$ Yang-Mills theory), this leads to an effective theory in $d-1$ spacetime dimensions. Depending on the profile of the parameter variation there are several possibilities for the physics (illustrated in Figure 4.1).

- If the parameter variation is smooth, i.e. it takes place over a distance scale longer than the UV cutoff, then the resulting interface dynamics is completely determined by the UV theory. It is universal.[5] In other words, these smooth interfaces are well defined observables of the QFT. Such interfaces have been widely studied for instance recently in $4d$ QCD and related applications to $3d$ dualities [28, 145, 1, 146]. One of the main applications of our formalism is to give a systematic point of view on the worldvolume anomaly of such interfaces. In particular we will see how they may be obtained from inflow from the $d+1$-dimensional classical theory encoding the bulk anomaly in the space of parameters. We will refer to such interfaces as "smooth interfaces" or simply interfaces.

- If the parameter variation is abrupt, more precisely, if it takes place over a distance scale comparable to the UV cutoff, the dynamics on the interface depends on additional UV data. It is not universal. For instance, such a sharp interface can always be decorated by coupling it to a $(d-1)$-dimensional QFT. To illustrate the difference more explicitly, consider for instance including in the UV Lagrangian a term of the form

$$\delta\mathcal{L} = \frac{1}{\Lambda^{\Delta+1-d}}\partial^\mu\theta(x)V_\mu(x) \ , \tag{4.6}$$

where Λ is a UV cutoff scale, and V_μ is an operator of scaling dimension Δ. If Δ is sufficiently large, and the gradients $\partial^\mu\theta(x)$ are small compared to the cutoff scale then this term is irrelevant at large distances. (Dangerously irrelevant operators should be treated separately.) However, when the gradients $\partial^\mu\theta(x)$ are large terms such as (4.6) become relevant and the interface dynamics depends on their coefficients. We will refer to such interfaces as "sharp interfaces" or as defects.

- A special case of such a sharp interface is the following. If the parameter variation is completely localized and the discontinuity and $(d-1)$-dimensional theory are chosen appropriately, then the resulting interface can be made to be completely transparent. We will refer to these as "transparent interfaces." Such transparent interfaces will play a key

[5]Following standard terminology, universal properties of quantum field theories are independent of the details of the UV theory at the cutoff scale.

role below.

We should emphasize that these interfaces should be distinguished from domain walls. Domain walls are also codimension one objects. But unlike the interfaces, they are dynamical excitations. They interpolate between two degenerate ground states and can move around. In contrast, interfaces are pinned by the external variation of the parameters.

While real-valued parameters often lead to codimension one defects, complex-valued parameters are naturally associated with defects of codimension two. A characteristic example is a $4d$ Weyl fermion with a position dependent complex mass $m(x, y)$ depending on two spacetime coordinates and winding n times around infinity. This example is the essence of the phenomenon investigated in [147,148,13]. The winding mass leads to two-dimensional Majorana-Weyl fermion zero modes localized at the zeros of the mass.

Below we will explain how this example can be viewed as an instance of anomalies in the space of masses. In particular, this means that the index theorem counting zero modes can be obtained by integrating an appropriate anomaly six-form (related to the $5d$ classical anomaly theory by descent):[6]

$$\mathcal{I}_6 = \frac{1}{384\pi^2}\gamma(m) \wedge \mathrm{Tr}(R \wedge R) = \frac{1}{48}\gamma(m) \wedge p_1 \ , \tag{4.8}$$

where $\gamma(m)$ is a two-form on the mass parameter space with total integral one, and p_1 is the first Pontrjagin class of the manifold.[7] For instance, it is often natural to take $\gamma(m) = \delta^{(2)}(m)d^2m$.

One virtue of this presentation of the anomaly in the space of mass parameters is that they

[6] In general spacetime dimension d we can define an anomaly $(d+2)$-form $\mathcal{I}_{d+2}$ by the property that the anomalous variation of the action ($\alpha(\lambda, A)$ above) is computed by

$$d\alpha(\lambda, A) = \delta\omega \ , \qquad d\omega = \mathcal{I}_{d+2} \ , \tag{4.7}$$

and above δ denotes the gauge variation. However, it is also possible for the anomaly $\omega(A)$ to be non-trivial and yet nevertheless the $(d+2)$-form $\mathcal{I}_{d+2}$ vanishes.

[7] Compactifying the complex mass plane by adding the point at infinity, the parameter space is topologically a two-sphere. This means that the free $4d$ Weyl fermion gives an elementary example of a field theory with an effective two-cycle in the space of parameters. More complicated examples of such two-cycles in the parameter space involving M5 branes or electric-magnetic duality were recently discussed in [149, 139, 140] in connection with some earlier assertions in [150].

are manifestly robust under a large class of continuous deformations. For instance, we can deform the $4d$ free fermion by coupling it to any interacting theory preserving the large $|m|$ asymptotics. The anomaly (4.8) remains non-trivial and implies that in any such theory, $2d$ defects arising from position-dependent mass with winding number n have chiral central charge, $c_L - c_R = n/2$.

4.1.3 Anomalies in Parameter Space: Families of QFTs

Another significant application of our techniques is to constraining the properties of families of QFTs. A typical situation we will consider is a family of theories labelled by a parameter such that at generic values of the parameter the theory is trivially gapped. An anomaly in the space of coupling constants can then imply that somewhere in the parameter space the infrared must be non-trivial.

An illustrative example that we describe in section 4.4 is two-dimensional $U(1)$ gauge theory coupled to N scalar fields of unit charge. At long distances the theory is believed to generically have a unique ground state. However, this conclusion cannot persist for all values of the coupling constants: for at least one value $\theta_* \in [0, 2\pi)$ the infrared must be non-trivial, and hence there is a phase transition as θ is dialed through this point.

We will argue for this conclusion by carefully considering the periodicity of θ. Placing the theory on $\mathbb{R}^2$ in a topologically trivial configuration of background fields, i.e. all those necessary to consider all correlation functions of local operators in flat space, the parameter θ has periodicity 2π. However when we couple to topologically non-trivial background fields the 2π-periodicity is violated.

Specifically, this gauge theory has a $PSU(N) \cong SU(N)/\mathbb{Z}_N$ global symmetry. In the presence of general background fields A for this global symmetry group the instanton number of the dynamical $U(1)$ gauge group can fractionalize. This means that in such configurations the periodicity of θ is enlarged to $2\pi N$. This violation of the expected periodicity of θ in the presence of background fields is conceptually very similar to the general paradigm of anomalies described in section 4.1.1. As in the discussion there, we find that the 2π periodicity of θ can be restored by coupling the theory to a three-dimensional classical field theory that depends on θ. Its Lagrangian is[8]

$$\omega = \frac{1}{N}\frac{d\theta}{2\pi} \cup w_2(A) \ , \tag{4.9}$$

where $w_2(A) \in H^2(X, \mathbb{Z}_N)$ is the second Stiefel-Whitney class[9] measuring the obstruction of lifting an $PSU(N)$ bundle to an $SU(N)$ bundle. In particular, this non-trivial anomaly must be matched under renormalization group flow now applied to the family of theories labelled by $\theta \in [0, 2\pi)$. A trivially gapped theory for all θ does not match the anomaly and hence it is excluded.

We can also describe the anomaly and its consequences somewhat more physically as follows. The theories at $\theta = 0$ and $\theta = 2\pi$ differ in their coupling to background A fields by a classical function of A (a counterterm) $\frac{2\pi}{N} w_2(A)$ [27]. However, since the coefficient of this counterterm must be quantized, this difference cannot be removed by making its coefficient θ-dependent in a smooth fashion. This means that at some $\theta_* \in [0, 2\pi)$ the vacuum must become non-trivial so that the counterterm can jump discontinuously. For instance in the special case $N = 2$ with a potential leading to a $\mathbb{CP}^1$ sigma model at low energies, the theory at $\theta = \pi$ is believed to be a gapless WZW model. For larger N, the $\mathbb{CP}^{N-1}$ model is believed to have a first order transition at $\theta = \pi$ with two degenerate vacua associated to spontaneously broken charge conjugation symmetry.

This example is emblematic of our general analysis below. We discuss QFTs with two essential properties. First, parameters can change continuously between two points with the same local physics. (In this example we shift the θ-parameter by 2π.) Second, the counterterms of background fields after the change are different. (In this example, the coefficient of the counterterm $w_2(A)$ is different.) Furthermore, if the coefficient of the counterterm is quantized, this difference cannot be eliminated by making its coefficient parameter-dependent in a continuous

[8] A non-expert physicist can think of the cup product as a version of a wedge product of differential forms appropriate for cohomology classes valued in finite groups.

[9] In the mathematics literature the Stiefel-Whitney classes are defined for principal $O(N)$-bundles. The characteristic class which measures the obstruction to lifting a projective bundle to a vector bundle could be called a "Brauer class".

fashion. We interpret this as a 't Hooft anomaly in the space of parameters and other background fields.

Then, the low-energy theory must saturate that anomaly. If it is nontrivial, i.e. gapless or gapped and topological, it should have the same anomaly. And if it is gapped and trivial for generic values of the parameters, there must be a phase transition for some value of the parameters. The fact that discontinuities in counterterms require phase transitions is widely known and applied, here we see how to phrase this idea in terms of 't Hooft anomalies.[10]

The example of $U(1)$ gauge theories described above is also a good one to illustrate the relation of our discussion to previous analyses of anomalies of discrete symmetries such as time-reversal, T, and charge conjugation, $\mathcal{C}$, in these models discussed in [27,158,159]. These theories are T (and $\mathcal{C}$) invariant at the two values $\theta = 0$ and $\theta = \pi$. For even N when $\theta = \pi$, there is a mixed anomalies between T (or $\mathcal{C}$) and the $PSU(N)$ global symmetry, and hence the long-distance behavior at $\theta = \pi$ cannot be trivial in agreement with the general discussion above. For odd N the situation is more subtle. In this case there is no anomaly for θ either 0 or π, but it is not possible to write continuous counterterms as a function of θ that preserve either T or $\mathcal{C}$ in the presence of background A fields at both $\theta = 0, \pi$ [27]. (This situation was referred to in [160–162] as "a global inconsistency.") Again this implies that there must be a phase transition at some value of θ in agreement with our conclusion above.

Our conclusions agree with previous results and, significantly, extend them in new directions. Indeed, the focus of the previous analysis is on subtle aspects of discrete symmetries, while in the anomaly in the space of parameters (4.9) T and $\mathcal{C}$ play no role. This means that the anomaly in the space of parameters, and consequently our resulting dynamical conclusions, persists under T and $\mathcal{C}$-violating deformations. We illustrate this in a variety of systems below. This example is again characteristic. By isolating an anomaly in the space of parameters of a QFT we are able to see that the conclusions are robust under a large class of deformations, and apply to other

[10] For instance in the study of $3d$ dualities the total discontinuity in various Chern-Simons levels for background fields as a parameter is varied is independent of the duality frame and provides a useful consistency check on various conjectures [56, 57, 31, 58, 151–153, 59, 154–157].

theories.

4.1.4 Synthesis via Anomaly Inflow

We have now described two general physical problems of interest:

- Anomalies on the worldvolume of topologically non-trivial interfaces and defects created by position-dependent parameters.
- Discontinuous counterterms in a family of QFTs and consequences for the long-distance behavior.

One of the basic points of our analysis to follow is that the solution of these two conceptually distinct problems is unified via anomaly inflow. Indeed the same $(d+1)$-dimensional classical theory can be used as a tool to analyze both phenomena. The difference between the applications is geometric. To describe a defect, the coupling constants vary in the physical spacetime. To describe a family of QFTs the coupling constants vary in the ambient directions extending the physical spacetime.

To illustrate this essential point, let us return again to the example of two-dimensional $U(1)$ gauge theory coupled to N scalars. The same anomaly action (4.9) introduced to restore the 2π periodicity of θ in the presence of a background A field can be used to compute the worldvolume anomaly of an interface, where θ varies smoothly from 0 to $2\pi n$, for integer n. In the first application the two spacetime dimensions have nonzero $w_2(A)$ and in the second, they have nonzero $d\theta$. In the latter case we simply integrate the anomaly to find

$$\omega_{\text{interface}} = \frac{n}{N} w_2(A) \ , \tag{4.10}$$

Since the bulk physics is trivially gapped for generic θ, we can interpret the above result as the anomaly of the effective quantum mechanical degrees of freedom localized on the interface. We deduce that the ground states of this quantum mechanics are degenerate and they form a projective representation of the $PSU(N)$ symmetry (i.e. a representation of $SU(N)$) with N-ality n.

Thus we see that these two distinct physical applications are synthesized via anomaly inflow in the space of coupling constants.

4.1.5 An Intuitive Interpretation in Terms of -1-Form Symmetries

Unlike ordinary 't Hooft anomalies, our anomalies are not associated with global symmetries of the system. They describe subtleties in the interplay between global symmetries and identifications in the parameter space. However, in some cases we can make our anomalies look more like ordinary anomalies in global symmetries. The examples with the periodicity of the θ-parameter can be thought of, somewhat heuristically, as related to a "-1-form global symmetry."

The θ-parameter is coupled to the instanton number. Borrowing intuition from string theory, we can view the instantons as -1-branes. More precisely, instantons are not branes. They are not well-defined excitations in spacetime. Yet, for many purposes they can be viewed as branes. Since they are at a point in spacetime, these are -1-branes. Extending this view of instantons, we can think of them as carrying -1-form charge. Clearly, this is an abuse of language – this charge is not a well-defined operator acting on a Hilbert space.

By analogy with ordinary charges, we can view θ as a background classical "gauge field" for this -1-form symmetry. Since it is circle valued, it can have transition functions where it jumps by $2\pi\mathbb{Z}$, but its "field-strength" $d\theta$ is single-valued. Then, all our anomaly expressions are similar to ordinary anomalies, except that they involve also these kinds of "gauge fields".

We should stress, however, that as far as we understand, this intuitive picture of the anomaly associated with θ cannot be extended to other situations where the topology of the parameter space is different. For example, we do not know how to do it for the examples in section 4.3.

4.1.6 Examples and Summary

Let us now summarize the examples analyzed below. We begin in section 4.2 with the elementary example of a quantum particle moving on a circle. This system is exactly solvable and exhibits a mixed anomaly between the circle-valued parameter θ and the $U(1)$ global symmetry or its $\mathbb{Z}_N$ subgroup. This example also gives us an opportunity to illustrate the various subtleties that occur when we make parameters position dependent.

In section 4.3 we discuss theories of free fermions in various spacetime dimensions as a function of the fermion mass. We start with the pedagogical example of a complex fermion in quantum mechanics. For $3d$ fermions we discuss how the index theorem of [163] explaining the discontinuity in the Chern-Simons level in the low-energy theory as real masses are varied from $-\infty$ to ∞ encodes an anomaly. For $4d$ Weyl fermions we describe the anomaly involving the complex fermion mass and explain its relationship to previous work [147, 13, 148] on fermions with position-dependent masses.

In section 4.4 we study $2d$ $U(1)$ gauge theory coupled to charged scalar fields. This includes Coleman's original paper [164] and the $\mathbb{CP}^n$ non-linear sigma model as special cases. These theories have a circle-valued coupling θ and we describe the resulting anomaly in the space of parameters. Here we extend the recent results of [27, 158, 159], which focused on the charge conjugation ($\mathcal{C}$) symmetry of these models at $\theta = 0, \pi$ to variants of these theories without this symmetry.

In section 4.5 we study four-dimensional Yang-Mills theories with a simply-connected gauge group G and determine the anomaly (summarized in Table 4.1). Using the logic discussed above, we also determine the worldvolume anomaly of interfaces interpolating between θ and $\theta + 2\pi k$ for some integer k. The anomaly constraints on the interfaces can be satisfied by the corresponding Chern-Simons theory with level k, G_k. However, as emphasized in [1], there are other options for the theory on the interface, all with the same anomaly. These generalized anomalies are invariant under deformations that preserve the center one-form symmetries. For instance, by adding appropriate adjoint Higgs field we show that the long distance theory can

Gauge group G	One-Form Sym. $(Z(G))$	Anomaly ω
$SU(N)$	$\mathbb{Z}_N$	$\frac{N-1}{2N}\int d\theta\,\mathcal{P}(B)$
$Sp(N)$	$\mathbb{Z}_2$	$\frac{N}{4}\int d\theta\,\mathcal{P}(B)$
E_6	$\mathbb{Z}_3$	$\frac{2}{3}\int d\theta\,\mathcal{P}(B)$
E_7	$\mathbb{Z}_2$	$\frac{3}{4}\int d\theta\,\mathcal{P}(B)$
$Spin(N)\ ,\quad N$ odd	$\mathbb{Z}_2$	$\frac{1}{2}\int d\theta\,\mathcal{P}(B)$
$Spin(N)\ ,\quad N = 2 \bmod 4$	$\mathbb{Z}_4$	$\frac{N}{16}\int d\theta\,\mathcal{P}(B)$
$Spin(N)\ ,\quad N = 0 \bmod 4$	$\mathbb{Z}_2\times\mathbb{Z}_2$	$\frac{N}{16}\int d\theta\,\mathcal{P}(B_L+B_R)+\frac{1}{2}\int d\theta\,B_L\cup B_R$

Table 4.1: The anomaly involving the θ-angle in Yang-Mills theory with simply connected gauge group G. These QFTs have a one-form symmetry, which is the center of the gauge group, $Z(G)$. (In particular, the omitted groups G_2, F_4, E_8 have a trivial center and hence their corresponding Yang-Mills theories do not have a one-form symmetry.) The anomaly ω depends on the background gauge field B for these one-form symmetries. ($\mathcal{P}(B)$ is the Pontryagin square operation defined in footnote 24.)

flow to a conformal field theory or a TQFT. These long-distance theories also reproduce the same generalized anomaly.

In section 4.6 we extend our analysis to four-dimensional $SU(N)$ and $Sp(N)$ gauge theories with massive fundamental fermion matter. We will show that, depending on the number of fundamental flavors N_f, these theories have a mixed anomaly involving θ and the appropriate zero-form global symmetries. A new ingredient is that there can be nontrivial counterterms with smoothly varying coefficients which can potentially cancel the putative anomalies. For $SU(N)$ we find that the anomaly is valued in $\mathbb{Z}_L$ with $L = \gcd(N, N_f)$. In particular the anomaly is non-trivial if and only if $\gcd(N, N_f) > 1$. For $Sp(N)$ we find a $\mathbb{Z}_2$ anomaly, which is non-trivial if and only if N is odd and N_f is even. As in the pure gauge theory, the discussion of these anomalies extends previous analyses that rely on time-reversal symmetry.

We also use these generalized anomalies to constrain interfaces. These anomaly constraints can be saturated by an appropriate Chern-Simons matter theory. Our analysis extends the recent results about interfaces in $4d$ QCD in [28] and explains the relation between them and the earlier results about anomalies in $3d$ Chern-Simons-matter theory in [151].

4.2 A Particle on a Circle

In this section we begin our generalization of the notion of anomalies to families of quantum systems. We present a simple and well-known example of a particle on a circle. This theory has a coupling constant θ and, as we show, exhibits an anomaly in this parameter space.

The dynamical variable in the theory is a periodic coordinate $q \sim q + 2\pi$. The Euclidean action is:

$$\mathcal{S} = \frac{m}{2}\int d\tau \dot{q}^2 - \frac{i}{2\pi}\int d\tau \ \theta\dot{q} \ . \tag{4.11}$$

In the above, θ is a coupling constant. Since the integral of $\dot{q}$ is quantized, the effect of θ is to weight different winding sectors with a phase. So defined, the parameter θ appears to be an angular variable with $\theta \sim \theta + 2\pi$. For instance the partition function, Z, viewed as a function of θ obeys

$$Z[\theta + 2\pi] = Z[\theta] \ . \tag{4.12}$$

Our goal in the following analysis is to clarify the circumstances where this periodicity of θ is valid. A distinguished role is played by the global shift symmetry under which $q \to q + \chi$. We will see that the 2π periodicity of θ is subtle in two related ways:

- If we try to make θ a non-constant function on the circle, the shift symmetry of q can be broken.
- In certain special correlation functions, related to adding background fields for the shift symmetry, the 2π-periodicity of θ is broken.

We elucidate these points and then discuss their dynamical consequences.

4.2.1 Spacetime Dependent Coupling θ

Let us first attempt to promote the coupling constant θ to depend on Euclidean time (which we assume to be periodic). Thus, we wish to make sense of the functional integral with action

(4.11), where now $\exp(i\theta(\tau))$ is a given function to $\mathbb{S}^1$. Note that unlike the variable $q(\tau)$ in (4.11) which is summed over, the function $\exp(i\theta(\tau))$ is a fixed classical variable.

Our first task is to clarify the meaning of the integral:

$$\exp\left(\frac{i}{2\pi}\int d\tau\ \theta(\tau)\dot{q}(\tau)\right)\ . \tag{4.13}$$

In general for a periodic variable such as q or θ, the derivative is always a single-valued function. Hence when θ was a constant the integral above was well-defined. However, when we allow θ to be position dependent it may wind in spacetime and the integral requires clarification.

A systematic way to proceed is to divide the spacetime circle into patches U_i, where $i = 1, \cdots n$ labels the patches, and each patch is an open interval. (For simplicity we assume that the patches only intersect sequentially so $U_i \cap U_j$ is non-empty only when $i = j \pm 1$. We also treat i as a cyclic variable.)

In each patch we can choose a lift $\theta_i : U_i \to \mathbb{R}$. On the non-empty intersections $U_i \cap U_{i+1}$ the lifts are related as

$$\theta_i = \theta_{i+1} + 2\pi n_i\ , \qquad n_i \in \mathbb{Z}\ . \tag{4.14}$$

The collection of lifts and transition functions yields a well-defined function to the circle. However it is redundant. If we adjust the data as

$$\theta_i \to \theta_i + 2\pi m_i\ , \qquad n_i \to n_i + m_i - m_{i+1}\ , \tag{4.15}$$

we obtain the same function $\exp(i\theta(\tau))$.

We now define the integral (4.13) using the collection of lifts θ_i. In each intersection $U_i \cap U_{i+1}$ we choose a point τ_i. We then set

$$\exp\left(\frac{i}{2\pi}\int d\tau\ \theta(\tau)\dot{q}(\tau)\right) \equiv \exp\left(\frac{i}{2\pi}\sum_{i=1}^{n}\int_{\tau_{i-1}}^{\tau_i} d\tau\ \theta_i(\tau)\dot{q}(\tau) - i\sum_{i=1}^{n} n_i q(\tau_i)\right)\ . \tag{4.16}$$

It is straightforward to verify the following properties of this definition:

- If the points τ_i defining the limits of the definite integrals are changed, the answer is unmodified.

- If the lifts θ_i and transitions n_i are redefined preserving the function $\exp(i\theta(\tau))$ as in (4.15), the integral is unmodified.

- If the patches are refined, i.e. a new patch is added, the integral is unmodified.

- If we change the choice of lift of q, for instance if we replace $q(\tau_i) \to q(\tau_i) + 2\pi$, the integral is unmodified.

- In the special case of constant θ, it reduces to the obvious definition.

Thus, the prescription (4.16) allows us to explore this quantum mechanics in the presence of spacetime varying coupling constant θ. More formally, the above may be viewed as a product in differential cohomology.

It is important to stress that although the definition (4.16) privileges the points τ_i, there is no physical operator inserted at these points. Rather, (4.16) is merely a way to define an integral in the case where θ has non-trivial winding number.

More physically, the varying coupling constant $\theta(\tau)$ allows us to illustrate the general comments on interfaces from section 4.1.2. When studying a problem with varying coupling constant $\theta(\tau)$ the simplest situation is that θ varies smoothly. In this case, the long distance behavior is intrinsic to the theory. We can generalize these situations where θ varies discontinuously with a sharp jump at some point τ_*, so $\theta(\tau_* - \epsilon) = \theta(\tau_* + \epsilon) + a$ for some constant a. Such a configuration is commonly referred to as a sharp interface or defect. In this case the dynamics are not universal, as any defect may be modified by dressing it by an operator $\mathcal{O}(\tau_*)$. What the definition (4.16) shows, is that in the special case where the discontinuity a is $2\pi n$ for some integer n, and the operator $\mathcal{O}(\tau_*)$ is taken to be $\exp(-inq(\tau_*))$ then the defect is trivial.

We can now explore aspects of the physics of varying coupling constant. Of particular interest is the interplay with the global symmetries of the problem. Consider a background $\theta(\tau)$ with non-zero winding number

$$L = \frac{1}{2\pi} \oint \dot{\theta}(\tau) d\tau = \sum n_i \tag{4.17}$$

around the circle. We claim that in such a configuration the global shift symmetry is broken. To see this, we shift $q(\tau) \to q(\tau) + \chi$ where χ is a constant. We then have from (4.16)

$$\exp\left(\frac{i}{2\pi} \int d\tau\ \theta(\tau)\dot{q}(\tau)\right) \to \exp\left(\frac{i}{2\pi} \int d\tau\ \theta(\tau)\dot{q}(\tau) - iL\chi\right) . \tag{4.18}$$

Since the remainder of the action (4.11) is obviously invariant under this shift, we can promote the above to a transformation of the full partition function under shifting $q \to q + \chi$

$$Z[\theta(\tau)] \to \exp\left(-\frac{i}{2\pi} \int d\tau\ \dot{\theta}(\tau)\chi\right) Z[\theta(\tau)] . \tag{4.19}$$

Since the zero mode of q is integrated over, the above in fact means that correlation functions vanish unless the insertions are chosen to cancel this transformation. Specifically, consider inserting $\prod_j \exp(i\ell_j q(\tau_j))$ together with other operators depending only on $\dot{q}$. From (4.19) we see that these correlation functions are non-zero only if $\sum_j \ell_j = L$.

As emphasized in section 4.1.1, equation (4.19) is characteristic of phenomena typically referred to as anomalies. By activating a topologically non-trivial configuration for the background field $\theta(\tau)$, in this case a non-zero winding number, the global $U(1)$ shift symmetry is violated.

4.2.2 Coupling to Background $U(1)$ Gauge Fields

Another approach to the same problem is to study the particle on the circle, to begin with at constant coupling θ, in the presence of a background $U(1)$ gauge field $A = A_\tau d\tau$ for the global $U(1)$ shift symmetry. The action is modified to

$$\mathcal{S} = \frac{m}{2} \int d\tau (\dot{q} - A_\tau)^2 - \frac{i}{2\pi} \int d\tau\ \theta(\dot{q} - A_\tau) , \tag{4.20}$$

and is invariant under gauge transformations $q \to q + \Lambda(\tau)$ and $A_\tau \to A_\tau + \dot{\Lambda}(\tau)$. (Note that since $\Lambda(\tau)$ transforms the classical field A_τ, it should be viewed as classical, and as such it cannot be used to set the dynamical degree of freedom q to zero.) The path integral over q now yields a partition function $Z[\theta, A]$ depending on a parameter θ and a gauge field A. However, it is no longer 2π-periodic in θ. Instead:

$$Z[\theta + 2\pi, A] = Z[\theta, A] \exp\left(-i \int d\tau A_\tau(\tau)\right) \ . \tag{4.21}$$

One possible reaction to the equation above is simply that the 2π-periodicity of the parameter θ is incorrect. Instead, more precisely, when discussing the coupling to background gauge fields A we should take care to also specify the counterterms, i.e. the local functions of the background fields that may be added to the action. In this case the counterterm of interest is a one-dimensional Chern-Simons term for the background gauge field A. Thus more precisely we can write the action

$$\mathcal{S} = \frac{m}{2} \int d\tau (\dot{q} - A_\tau)^2 - \frac{i}{2\pi} \int d\tau \ \theta(\dot{q} - A_\tau) - ik \int d\tau A_\tau(\tau) \ , \tag{4.22}$$

where k must be an integer to ensure gauge invariance. Including such a term in the action we arrive at a partition function $Z[\theta, k, A]$. Then (4.21) means that

$$Z[\theta + 2\pi, k, A] = Z[\theta, k - 1, A] \ . \tag{4.23}$$

Thus, in this interpretation the true parameter space is a helix that we may view as a covering space of the circle (where θ ranges from 0 to 2π). The different values of k index the different sheets in the cover. Said differently, if we demand that two values of the coupling constant are only considered equivalent if all observables agree, including the phase of the partition function in the presence of background fields, then there is, by definition, no such thing as an anomaly in the space of coupling constants.

4.2.3 The Anomaly

There is however, an alternative point of view, which is also fruitful. Instead of enlarging the parameter space, we can retain the 2π-periodicity of θ as follows. We pick a two-manifold Y with boundary the physical quantum mechanics worldvolume of our problem. We extend the classical fields θ and A into the bulk Y. Then, we define a new partition function as

$$\tilde{Z}[\theta, k, A] = Z[\theta, k, A] \exp\left(2\pi i \omega[\theta, A]\right) = Z[\theta, k, A] \exp\left(\frac{i}{2\pi} \int_Y \theta F\right) . \tag{4.24}$$

This modified partition function now obeys the simple 2π-periodicity of θ even in the presence of background fields

$$\tilde{Z}[\theta + 2\pi, k, A] = \tilde{Z}[\theta, k, A] . \tag{4.25}$$

The price we have paid is that $\tilde{Z}$ now depends on the chosen extension of the classical fields into the bulk Y.

We can now easily extend our analysis to allow for a non-constant function $\theta(\tau)$. The integral over Y is defined similarly to (4.16). We divide the manifold Y into patches in each patch we integrate a lift of θ times the curvature F. We add to this integral a boundary contribution, which in this case is a line integral of the gauge field A weighted by the transition function $\theta_i - \theta_j$. On a closed two-manifold this results in a well-defined action independent of trivialization: moving the boundary of a patch U_1 to encompass a new region W formerly contained in U_2 leads to a new integral $\exp\left(\frac{i}{2\pi}\int_W (\theta_{U_1} - \theta_{U_2})F\right)$ together with a compensating contribution $\exp\left(-\frac{i}{2\pi}\int_{\partial W} (\theta_{U_1} - \theta_{U_2})A\right)$.

To see the interplay with the boundary action (4.16) consider now a patch U that terminates on the boundary. The result is now $U(1)$ gauge invariant even in the presence of general $\theta(\tau)$. Indeed, the Wilson lines on the edge of each bulk patch now terminate on the insertions of $\exp(-in_i q(\tau_i))$ and hence are gauge invariant. Thus, the result is a partition function $\tilde{Z}[\theta(\tau), A]$ that is a well-defined function of a circle-valued field $\exp(i\theta(\tau))$ and gauge invariant as a function of A.

For some purposes it is also useful to apply the descent procedure again to produce an anomaly three-form. Such a two-step inflow is in general possible when discussing infinite order anomalies (classified by an integer) such as those computed by one-loop diagrams in even-dimensional QFTs. In this case we find:

$$\mathcal{I}_3 = d\omega = \frac{1}{(2\pi)^2} d\theta \wedge F \ , \tag{4.26}$$

which encodes the anomaly action in (4.24).[11]

4.2.4 Dynamical Consequences: Level Crossing

We can use our improved understanding of the behavior of the theory as a function of the θ-angle to make sharp dynamical predictions about the particle on a circle. Specifically, we claim that for at least one value $\theta_* \in [0, 2\pi)$, the system must have a non-unique ground state.

To argue for this, suppose on the contrary that we have a unique ground state for all θ. We can focus on this state by scaling up all the energies in the problem. At each θ, the low-energy partition function is then that of a trivial system with a single unique state. However, a single unique state cannot produce the jump (4.21) in the one-dimensional Chern-Simons level. Here it is crucial that the coefficient of this level is quantized. In particular, this prohibits a continuously variable phase of the partition function interpolating between the value at $\theta = 0$ and $\theta = 2\pi$.

Of course, the free quantum particle on the circle is an exactly solvable system for any θ and its behavior is well-known. There is a single unique ground state for all $\theta \neq \pi$. However for $\theta = \pi$, where there is an enhanced charge conjugation symmetry, $\mathcal{C}$, acting as $q \to -q$, there are two degenerate ground states. Thus, the conclusions above are indeed correct, though the highbrow reasoning is hardly necessary. In terms of anomalies one may derive the degeneracy at $\theta = \pi$, following [27], by noting that for this value of θ, there is a mixed anomaly between $U(1)$ and $\mathcal{C}$ and hence a unique ground state is forbidden.

[11] We can also change the precise representative of the cohomology class appearing in (4.26) without modifying the essential consequences in (4.25). This means that we can replace $d\theta$ by $d(\theta + f(\theta))$ with $f(\theta)$ a 2π periodic function. In fact as we will see in section 4.3.1, the low-energy theory near $\theta = \pi$ produces a different anomaly action ω related by continuously varying the form $d\theta/2\pi$ to a periodic delta function $\delta(\theta - \pi)d\theta$.

The advantage of the more abstract arguments is that they are robust under a large number of possible deformations of this system. It is instructive to proceed in steps. We can first consider deforming the system by a potential breaking $U(1)$ to $\mathbb{Z}_N$ and preserving the other discrete symmetries

$$\begin{aligned} \mathcal{S} &= \int d\tau \left(\frac{m}{2}\dot{q}^2 - \frac{i}{2\pi}\theta\dot{q} + U(q) \right) \\ U(q) &= U(q + \frac{2\pi}{N}) = U(-q) \ , \end{aligned} \tag{4.27}$$

e.g. $U(q) = \cos(Nq)$. Here, $U(q) = U(q + \frac{2\pi}{N})$ guarantees the $\mathbb{Z}_N$ symmetry and $U(q) = U(-q)$ guarantees that the $\mathcal{C}$ and $\mathcal{T}$ symmetries are as in the problem without the potential. For even N, there is a mixed anomaly at $\theta = \pi$ between $\mathcal{C}$ and $\mathbb{Z}_N$ leading again to ground state degeneracy.[12] For odd N there is no anomaly, but it is impossible to define continuous θ-dependent counterterms to preserve $\mathcal{C}$ at both $\theta = 0$ and $\theta = \pi$ [27]. This lack of suitable counterterms was referred to as a global inconsistency in [160, 162], also implies a level crossing. In this theory our anomaly in the space of couplings persists and yields the same conclusion though it does not single out $\theta = \pi$ as special.

Finally, we can consider deformations breaking all symmetries, and in particular $\mathcal{C}$ and $\mathcal{T}$, except the $\mathbb{Z}_N$ symmetry. For instance, we can introduce a real degree of freedom X and consider the action:

$$\begin{aligned} \mathcal{S} &= \int d\tau \left(\frac{m}{2}\dot{q}^2 - \frac{i}{2\pi}\theta\dot{q} + U(q) + \frac{M}{2}\dot{X}^2 + iX\dot{q} + V(X) \right) \\ U(q) &= U(q + \frac{2\pi}{N}) \ , \end{aligned} \tag{4.28}$$

with generic $U(q)$ (subject to $\frac{2\pi}{N}$ periodicity) and $V(X)$. Note that unlike (4.27), we do not impose that $U(q)$ is even and therefore we break $\mathcal{C}$ and $\mathcal{T}$ for all θ. The free particle on a circle is obtained in the limit $U(q) \to 0$ and $M \to \infty$. The condition $U(q) = U(q + \frac{2\pi}{N})$ guarantees that the $\mathbb{Z}_N$ symmetry remains, however there is no special value of θ with enhanced symmetry. Nevertheless as we vary θ from zero to 2π the phase of the partition function changes by the

[12] The anomaly action is $\exp\left(i\pi \int C \cup K\right)$ where C is the $\mathbb{Z}_2$ charge conjugation gauge field and K is the $\mathbb{Z}_N$ gauge field. Note that this is only non-trivial when N is even. (Although charge conjugation acts non-trivially on the $\mathbb{Z}_N$ symmetry with gauge field K, this action is trivial once we reduce modulo two.)

insertion of a $\mathbb{Z}_N$ Wilson line, and thus the anomaly in the space of couplings persists. Hence if $N > 1$ we again deduce that somewhere in θ we must have level crossing for the ground state and hence ground state degeneracy.

To deduce how the anomaly action in (4.24) reduces in this more complicated situation, it is helpful to integrate the action there by parts and express it as

$$\mathcal{A}[\theta, A] = \exp\left(-i\int_Y \frac{d\theta}{2\pi}A\right) . \tag{4.29}$$

On a closed two-manifold Y this defines the same anomaly action. On a manifold Y with boundary, (4.29) and (4.24) differ by a choice of boundary counterterm (in this case θA.) In the following, we mostly use expressions such as (4.29) with the understanding that we may need to add suitable boundary terms.

Using (4.29), we can describe the anomaly action in the deformed theory (4.28) more precisely as follows. The $U(1)$ background gauge field A is now replaced by a $\mathbb{Z}_N$ gauge field K. (Our convention is that the holonomies of K are integers modulo N.) Concretely, we can embed K in a $U(1)$ gauge field A as $A = \frac{2\pi}{N}K$. Then we find that (4.29) reduces to the anomaly action

$$\mathcal{A}[\theta, K] = \exp\left(-\frac{2\pi i}{N}\int_Y \frac{d\theta}{2\pi}\cup K\right) . \tag{4.30}$$

This anomaly is non-trivial and implies the level-crossing of that system.[13]

In fact even in the general system (4.28), we can give a straightforward argument for level crossing using a canonical quantization picture. The wavefunctions of states of definite charge k mod N under the $\mathbb{Z}_N$ symmetry can be expanded in a Fourier series as

$$\psi(q, X) = \sum_{j=-\infty}^{\infty} e^{i(k+Nj)q}\mu_j(X) . \tag{4.31}$$

[13] One can also construct a direct analog of the $i\int_Y \theta F$ anomaly action even in the case of a discrete $\mathbb{Z}_N$ symmetry. To do so, we lift the $\mathbb{Z}_N$ gauge field to a $U(1)$ gauge field and evaluate the integral using the differential cohomology definition in section 4.4.1.

theory generic symmetry $\mathcal{G}$	without $\mathcal{C}$ θ-$\mathcal{G}$ anomaly	with $\mathcal{C}$ at $\theta = 0, \pi$ $\mathcal{C}$-$\mathcal{G}$ anomaly at $\theta = \pi$	 no continuous counterterms
q $\mathcal{G} = U(1)$	✓	✓	✓
q + potential (4.27) $\mathcal{G} = \mathbb{Z}_N$, $N > 1$	✓	even N ✓ odd N ✗	even N ✓ odd N ✓
q +X system (4.28) $\mathcal{G} = \mathbb{Z}_N$, $N > 1$	✓	No $\mathcal{C}$ symmetry	No $\mathcal{C}$ symmetry

Table 4.2: Summary of anomalies and existence of continuous counterterms preserving $\mathcal{C}$ ("global inconsistency") in the hierarchy of theories considered above. The left-most column shows the theory and its symmetry at generic values of θ. Without using the charge conjugation symmetry, all these theories exhibit a mixed 't Hooft anomaly in θ and $\mathcal{G}$. The anomaly implies that the theories cannot have a unique ground state for all values of $\theta \in [0, 2\pi)$. For the simpler theories there is also a charge conjugation symmetry at $\theta = \pi$, which may suffer from a 't Hooft anomaly. We have also indicated when the theories lack a continuous counterterm that can preserve $\mathcal{C}$ at both $\theta = 0, \pi$.

Let ψ above be the ground state at $\theta = 0$. We can track this state as a function of θ. The Hamiltonian is

$$H = \frac{1}{2m}\left(P_q - \frac{\theta}{2\pi} - X\right)^2 + \frac{1}{2M}P_X^2 + V(X) + U(q) \ . \tag{4.32}$$

In canonical quantization, the momentum operator is $P_q = -i\frac{d}{dq}$. From this we see that the ground state at $\theta = 2\pi$ is not $\psi(q, X)$, but rather is $e^{iq}\psi(q, X)$. Physically, this means that as we dial θ from zero to 2π, the $\mathbb{Z}_N$ charge of the ground state wavefunction increases by one unit. In particular, at some value of θ, level crossing for the ground state must occur.

4.3 Massive Fermions

In this section we consider free fermions in various spacetime dimensions as a function of their mass parameters. We will see that this gives simple examples of systems with anomalies in their parameter space. We will also see how these models can be deformed to interacting theories with the same anomaly.

4.3.1 Fermion Quantum Mechanics

Consider the quantum mechanics of a complex fermion with a real mass m. Anomalies in fermionic quantum mechanics were first discussed in [165]. The Euclidean action is

$$\mathcal{S} = \int d\tau \left(i\psi^\dagger \partial\psi + m\psi^\dagger\psi \right) \ . \tag{4.33}$$

This theory has a two-dimensional Hilbert space spanned by two energy eigenstates $|\pm\rangle$ of energy $E = \pm m/2$. On a Euclidean time circle of length β the partition function is

$$Z[m] = e^{-\beta m/2} + e^{\beta m/2} \ . \tag{4.34}$$

At vanishing mass m the theory has two degenerate ground states, while for non-zero mass, one or the other state becomes energetically favorable. As we will see, this means that this fermion quantum mechanics is identical to the theory of a particle on a circle described in section 4.2 where we have isolated the two nearly degenerate states at $\theta = \pi$. (See e.g. [166].)

Of particular interest to us is the asymptotic behavior of the theory for large $|m|$. Regardless of the sign of m we see that in this limit there is a single ground state and an infinite energy gap to the next state. Thus, the physics in these two limits is identical. Effectively we can say that the parameter space of masses is compactified to $\mathbb{S}^1$.

This simple free fermion theory has a $U(1)$ global symmetry and can be coupled to a background gauge field $A = A_\tau d\tau$, which modifies the action to

$$\mathcal{S} = \int d\tau \left(i\psi^\dagger (\partial - iA_\tau)\psi + m\psi^\dagger\psi - ikA_\tau \right) \ . \tag{4.35}$$

Here we have included in the action a counterterm depending only on A, whose coefficient k is integral. Since we transition between the two states by an action of the ψ operator, they differ in their $U(1)$ charge by one unit. Therefore the partition function including $A = A_\tau d\tau$ is (below we have shifted the Hamiltonian so that the energies are $\pm m/2$)

$$Z[m,k,A] = e^{ik\oint A}\left(e^{-\beta m/2} + e^{\beta m/2 - i\oint A}\right) . \tag{4.36}$$

Now we see that the theories at large positive and negative mass differ by a local counterterm

$$\lim_{m\to+\infty} \frac{Z[m,k,A]}{Z[-m,k,A]} = \exp\left(-i\oint A\right) . \tag{4.37}$$

Note crucially that since k is quantized, there is no way to modify the result (4.37) by adding a continuous m-dependent counterterm for the background gauge field A. This means that we can interpret (4.37) as an anomaly involving the mass parameter and the $U(1)$ global symmetry. Specifically, we define a new partition function by extending the backgrounds, in this case the gauge field A and the mass m, into a $2d$ bulk Y. We then define a new partition function by

$$\tilde{Z}[m,k,A] = Z[m,k,A]\exp\left(i\int \rho(m)F\right) , \tag{4.38}$$

where $F = dA$ is the curvature and the function $\rho(m)$ satisfies

$$\lim_{m\to-\infty} \rho(m) = 0 , \qquad \lim_{m\to+\infty} \rho(m) = 1 . \tag{4.39}$$

The modified partition function $\tilde{Z}$ now has a manifestly uniform limit as $|m|$ becomes large:

$$\lim_{m\to+\infty} \frac{\tilde{Z}[m,k,A]}{\tilde{Z}[-m,k,A]} = 1 . \tag{4.40}$$

This anomaly persists under arbitrary deformations of the theory that preserve the $U(1)$ symmetry. (For instance it persists under deformations that violate the charge conjugation symmetry, $\mathcal{C}$, which acts as $\mathcal{C}(\psi) = \psi^\dagger$.)

How shall we interpret the arbitrary function $\rho(m)$ above? One way to understand the ambiguity in the function $\rho(m)$ is that it reflects the fact that in general systems without additional symmetry there is no preferred way to parameterize the space of masses. Under a redefinition $m \to h(m)$ where $h(m)$ is any bijective function with $h(\pm\infty) = \pm\infty$ modifies the precise function ρ above but preserves the properties (4.39). This is similar to the general counterterm ambiguities that are always present when discussing anomalies, and in fact parameter redefinitions occur along renormalization group trajectories.

It is the rigid limiting behavior of the function $\rho(m)$ above that means that the deformation class of the anomaly we are describing is preserved under any continuous deformation of the theory. As in section 4.2.3, we can make the cohomological properties more manifest by applying the descent procedure again to produce an anomaly three-form (see footnote 6). In this case we find:

$$\mathcal{I}_3 = \frac{1}{2\pi} f(m) dm \wedge F \ , \tag{4.41}$$

where $f(m)dm$ is a one-form with the property that

$$\int_{-\infty}^{+\infty} f(m) dm = 1 \ . \tag{4.42}$$

In this free fermion problem, it is natural to take $f(m) = \delta(m)$, such that the anomaly is supported only at $m = 0$ where we have level crossing. This is in accord of the discussion in footnote 11. Below we will also discuss other options. Such a one-form represents a non-trivial cohomology class on the real line, once one imposes a decay condition for $|m| \to \infty$. (Here we have in mind a model such as compactly supported cohomology see e.g. [167]). The form $f(m)dm$ is non-trivial because it cannot be expressed as the derivative of any function tending to zero at $m = \pm\infty$. Alternatively as suggested above, one can compactify the real mass line to a circle in which case $f(m)dm$ represents the generator of $H^1(\mathbb{S}^1, \mathbb{Z})$. Viewed as such a cohomology class the anomaly is rigid because the integral is quantized. This feature is preserved under any continuous deformation of the theory.

4.3.2 Real Fermions in $3d$

As our first example of a quantum field theory (as opposed to a quantum mechanical theory) with an anomaly in parameter space we consider free fermions in three dimensions. We will see how some familiar properties of fermion path integrals can be reinterpreted as anomalies involving the fermion mass. We focus on the theory of a single Majorana fermion ψ, though our analysis admits straightforward extensions to fermions in general representations of global symmetry groups.

The Euclidean action of interest is

$$\mathcal{S} = \int d^3x \left(i\psi \partial\!\!\!/ \psi + im\psi\psi \right) \ , \tag{4.43}$$

where $m \in \mathbb{R}$ above is the real mass. Our goal is to understand properties of the theory as a function of the mass m. As above it is fruitful to encode these in a partition function $Z[m]$.

As in our earlier examples, we first consider the free fermion theory in flat spacetime. At non-zero m, the theory is gapped with a unique ground state and no long range topological degrees of freedom. As the mass is increased the gap above the ground state also increases and we isolate the ground state. In particular the partition function, as well as the correlation functions of all local operators, become trivial in this limit[14]

$$\lim_{m \to -\infty} Z[m] = \lim_{m \to +\infty} Z[m] = 1 \ . \tag{4.44}$$

Like the fermion quantum mechanics problem of section 4.3.1, one can interpret the above in terms of the effective topology of the parameter space. The space of masses is a real line, and we can formally compactify it to $\mathbb{S}^1$ by including $m = \infty$.

The situation is more subtle if we consider the theory on a general manifold with non-trivial metric g, and hence associated partition function $Z[m, g]$. Fixing g but scaling up the mass again

[14] The partition function $Z[m]$ is subject to an ambiguity by adding counterterms of the form $\int d^3x \ h(m)$ for any function $h(m)$. Below we assume that these terms are tuned so that (4.44) is true.

leads to a trivially gapped theory, however now the theories at large positive and negative mass differ in the phase of the partition function. Locality implies that the ratio of the two partition functions in this limit must be a well-defined classical functional of the background fields. In this case the APS index theorem [163] implies that the ratio is[15]

$$\lim_{m\to+\infty}\frac{Z[+m,g]}{Z[-m,g]}=\exp\left(i\int_X CS_{\mathrm{grav}}\right)~, \tag{4.45}$$

where CS_{grav} is the minimally consistent spin gravitational Chern-Simons term for the background metric.[16] Thus in the presence of a background metric, the identification between $m=\pm\infty$ is broken. (For early discussion of this in the physics literature, see [169, 170].)

Notice that in (4.45) we have focused only on the ratio between the partition functions. In fact since the theories at large $|m|$ are separately trivially gapped, each theory separately gives rise to a local effective action of the background metric. However in general, one may adjust the UV definition of the theory by adding such a local action for the background fields. Physically this is the ambiguity in adjusting the regularization scheme and counterterms. By considering the ratio of partition functions we remove this ambiguity. Thus while the effective gravitational Chern-Simons level is individually scheme-dependent for large positive and large negative mass, the difference between the levels is physical. (See also footnote 15.)

In fact, the jump in the gravitational Chern-Simons level (4.45) is a manifestation of the time-reversal (T) anomaly of the free fermion theory. At vanishing mass the system is time-reversal invariant, but the mass breaks this symmetry explicitly with $\mathsf{T}(m)=-m$. The gravitational Chern-Simons term is also odd under T and hence a fully time-reversal invariant quantization of

[15] As in footnote 14, below we use the freedom to tune counterterms. However, as we discuss the right-hand side of (4.45) cannot be modified by any such tuning.

[16] As usual, it is convenient to define this term by an extension to a spin four-manifold Y. Then for any integer k we have

$$\exp\left(ik\int_X CS_{\mathrm{grav}}\right)=\exp\left(2\pi ik\int_Y\frac{p_1(Y)}{48}\right)=\exp\left(\frac{ik}{192\pi}\int_Y \mathrm{Tr}(R\wedge R)\right)~, \tag{4.46}$$

where $p_1(Y)$ is the Pontrjagin class and we have used $\int_Y p_1(Y)\in 48\mathbb{Z}$ for any closed spin manifold Y. Although this term is called a gravitational 'Chern-Simons term' in the physics literature, it is not covered by the work of Chern-Simons [168]. Rather, it is an exponentiated η-invariant.

the theory in the presence of a background metric would require the effective levels at large positive and negative masses to be opposite. The jump formula (4.45) means that this is impossible to achieve by adjusting the counterterm ambiguity.

We would now like to reinterpret the jump (4.45) in terms of an anomaly involving the fermion mass viewed now as a background field. Analogous to our examples in quantum mechanics, we introduce a new partition function $\tilde{Z}[m,g]$, which depends on an extension of the mass m and metric g into a four-manifold Y with boundary X:

$$\tilde{Z}[m,g] = Z[m,g]\exp\left(-i\int_Y \rho(m) dCS_{\text{grav}}\right) = Z[m,g]\exp\left(-\frac{i}{192\pi}\int_Y \rho(m)\text{Tr}(R\wedge R)\right) \ , \tag{4.47}$$

where above $\rho(m)$ satisfies the same criteria as in the anomaly in the fermion quantum mechanics theory (4.39). (And as in the discussion there, in the free fermion theory it is natural to take $\rho(m)$ a Heaviside theta-function.) This partition function now retains the identification between $m=\pm\infty$ even in the presence of a background metric g at the expense of the extension into four dimensions.

In fact, using time-reversal symmetry we can say more about the function $\rho(m)$ above. Since m is odd under T and time-reversal changes the orientation of spacetime, demanding that the $4d$ anomaly action in (4.47) is T invariant leads to the additional constraint

$$\rho(m)+\rho(-m) = 1 \ . \tag{4.48}$$

In particular we can use this to recover the T anomaly of the theory at $m=0$: using $\rho(0)=1/2$, the anomaly becomes a familiar gravitational θ_g-angle at the non-trivial T-invariant value of $\theta_g=\pi$.

However, the virtue of viewing the anomaly (4.47) as depending only on the parameters m and g is that it is manifestly robust under T violating deformations. This means that the anomaly (4.47) has implications for a much broader class of theories. For example, consider coupling the free fermion to a real scalar field φ so the action is now

$$\mathcal{S} = \int d^3x \left(i\psi \partial\!\!\!/ \psi + (\partial\varphi)^2 + i(m+\varphi)\psi\psi + V(\varphi) \right) \ , \tag{4.49}$$

where $V(\varphi)$ is any potential. For generic $V(\varphi)$ this system does not have T symmetry. Nevertheless the arguments leading to the anomaly involving the fermion mass m and the metric g still apply. In this more general context, the constraint (4.48) does not hold, and only the general constraint (4.39) is applicable.

As in section 4.2.3, we can also apply the descent procedure again to find an anomaly five-form:

$$\mathcal{I}_5 = -\frac{1}{384\pi^2} f(m) dm \wedge \mathrm{Tr}(R \wedge R) = -\frac{1}{48} f(m) dm \wedge p_1 \ , \tag{4.50}$$

where p_1 is the first Pontrjagin class of the manifold, and $f(m)dm = d\rho(m)$ has unit total integral.

Dynamical Consequences

We now apply the anomaly (4.50) to extract general lessons about the physics. As described in section 4.1, there are broadly speaking two lessons that we can learn.

- Existence of non-trivial vacuum structure: Consider any QFT with the anomaly (4.50). Such a theory cannot have a trivially gapped vacuum (i.e. a unique ground state and an energy gap with no long-range topological degrees of freedom) for all values of the mass m. To argue for this we assume on the contrary that the theory is trivially gapped for all m. Then at long-distances $Z[m] \to 1$ for all masses m, which of course does not have the anomalous transformation required by the bulk anomaly action.

  Thus, we conclude that somewhere in the space of mass parameters the vacuum must be non-trivial. In other words, either the gap must close or a first order phase transition (leading to degenerate vacua) occurs. Of course for the free fermion this is hardly surprising since at $m = 0$ the fermion is massless. However for more general interacting systems such as that in (4.49), this conclusion is less immediate.

- Non-trivial physics on interfaces: Consider for instance a situation where for sufficiently large $|m|$ the theory is gapped. We activate a smooth space-dependent mass $m(x)$ depending on a single coordinate x which obeys $m(\pm\infty) = \pm\infty$. At low-energies in the transverse space we find an effective theory, which is necessarily non-trivial. The anomaly of this theory can be computed by integrating the anomaly action over the coordinate x. Using the property (4.42) this leads to

$$i \int_{Y_3} CS_{\text{grav}} \ , \tag{4.51}$$

  where now Y_3 is an extension of the effective two-dimensional theory. In particular, the anomaly (4.51) implies that the theory on the interface is gapless with chiral central charge fixed by the anomaly theory $c_L - c_R = 1/2$. This result is well-known in the condensed matter literature: the classical action (4.51) describes a $3d$ topological superconductor without a global symmetry, which is known to have a gapless chiral edge mode.

  Again for the free fermion this conclusion is obvious. At the special locus in x where $m = 0$, the $3d$ fermion becomes localized and leads to a massless $2d$ Majorana-Weyl fermion. However, for more general interacting systems with the same anomaly, the conclusion still holds.

In general, the basic idea encapsulated by the above example is that for any one-parameter family of generically gapped systems with symmetry $\mathcal{G}$ (either unitary internal or spacetime) we can track the long-distance $\mathcal{G}$ counterterms as a function of the parameter. The discontinuity in these counterterms as the parameter is varied from $-\infty$ to ∞ is an invariant of the family.

Such tracking of the jump in gravitational and other Chern-Simons terms in background gauge fields was a powerful consistency check on various conjectures about $3d$ dynamics and dualities [56, 57, 31, 58, 151–153, 59, 154–157]. Here, we see that this idea is formalized into an anomaly in the space of coupling constants and this consistency check is unified with standard anomaly matching.

4.3.3 Weyl Fermions in $4d$

We now consider free fermions in $4d$. We will again find mixed anomalies in the space of mass parameters and background metrics. A qualitatively new feature is that in this case the anomaly is present only if we study the full two-dimensional complex m-plane. Effectively, this means that the m-plane is a non-trivial two-cycle in parameter space. Other examples with two-cycles in parameter space are discussed in $[149, 139, 140]$. We focus below on the simplest case of a minimal free Weyl fermion. Extensions to fermions in general representations of global symmetry groups are straightforward.

Our starting point is the partition function $Z[m]$ of a free Weyl fermion ψ viewed as a function of the complex mass parameter m. The massless theory has a chiral $U(1)$ symmetry under which ψ has unit charge. A non-zero mass parameter entering the Lagrangian as $m\psi\psi + c.c$ breaks this symmetry and we can view m as a spurion of charge minus two. This means that the partition function in flat space obeys (with an appropriate choice of counterterms):

$$Z\left[e^{i\phi}m\right] = Z\left[m\right] \ . \tag{4.52}$$

In particular, the above equation holds for large $|m|$ where the theory is trivially gapped. Thus it is consistent to compactify the mass parameter space to a sphere $\mathbb{S}^2$ by viewing all masses of large absolute value as a single point.

We now couple the theory to a background metric g and reexamine the above conclusions. As in our example in $3d$, we will see that the large $|m|$ behavior of the partition function is now more subtle. Recalling that for $m = 0$ the $U(1)$ chiral symmetry participates in a mixed anomaly with the geometry, the partition function is modified under a chiral rotation as:

$$Z\left[e^{i\phi}m, g\right] = Z\left[m, g\right] \exp\left(-\frac{i\phi}{384\pi^2}\int_X \mathrm{Tr}(R\wedge R)\right) = Z\left[m, g\right] \exp\left(-i\phi\int_X \frac{p_1(X)}{48}\right) \ . \tag{4.53}$$

The dependence on the argument of m above means that the topological interpretation of the space of masses as a sphere is obstructed in the presence of a background metric.

The Anomaly

We can, however, restore the identification of the points at infinite $|m|$ by introducing a suitable bulk term. Specifically we define a new partition function as

$$\tilde{Z}[m,g] = Z[m,g]\exp\left(2\pi i\int_Y \lambda(m)\wedge\frac{p_1(Y)}{48}\right) ~, \tag{4.54}$$

where above $\lambda(m)$ is any one-form which asymptotically approaches an angular form for large $|m|$:

$$\lim_{|m|\to\infty}\lambda(m) = \frac{d\arg(m)}{2\pi} ~. \tag{4.55}$$

The partition function $\tilde{Z}[m,g]$ is then invariant under phase rotation of m for large $|m|$ and the topological interpretation of the spaces of masses as $\mathbb{S}^2$ is restored.

Observe that the anomaly (4.54) is supported by the non-trivial effective two-cycle of masses. In other words, if we restrict to any one-parameter slice of masses the anomaly trivializes. For instance along a circle of constant non-zero $|m|$ we can add to the Lagrangian a counterterm of the form change in the equation

$$\frac{i\arg(m)}{384\pi^2}\mathrm{Tr}(R\wedge R) ~, \tag{4.56}$$

and cancel the spurious transformation in (4.53). However, it is impossible to extend this counterterm to a smooth local $4d$ function of m and g on the entire two-dimensional m plane. This obstruction is the anomaly.

As in the case of the $3d$ free fermion, we can also write the anomaly by applying the descent procedure a second time to obtain an anomaly six-form. In this case it is

$$\mathcal{I}_6 = \gamma(m)\wedge\frac{p_1(Y)}{48} ~, \tag{4.57}$$

where $\gamma = d\lambda$ is a two-form with total integral one on the mass-plane. (As above, in the free fermion theory it is natural to take $\gamma(m) = \delta^{(2)}(m)d^2m$, but below we will also discuss other natural forms.) This anomaly is similar to that found in the space of marginal coupling constants in [149, 139, 140].

The fact that γ above has quantized total integral means that the anomaly is cohomologically non-trivial and hence it is preserved under continuous deformations of the theory including renormalization group flow. As with our discussion in previous sections, this also means that the same anomaly is present for more general interacting theories. For instance, analogously to (4.58) we can consider a theory with an additional real scalar φ

$$\mathcal{S} = \int d^4x \left(i\bar{\psi}\partial\!\!\!/\psi + (\partial\varphi)^2 + [(m+\varphi)\psi\psi + c.c.] + V(\varphi) \right) \ . \tag{4.58}$$

This theory still has the anomaly (4.57) and hence the consequences discussed below.

Dynamical Consequences

We now apply the anomaly (4.57) to deduce general physical consequences. As always, we can consider a family of theories labelled by m or a spacetime-dependent coupling $m(x)$.

- Non-trivial vacuum structure in codimension two: Consider the family of theories labelled by m with an anomaly (4.57). Then, in order for the anomaly to be reproduced at long distances the theory cannot be trivially gapped for all m.

  Notice that unlike the discussion in sections 4.2 and 4.3.2, the non-trivial vacuum structure need only to occur in codimension two. In particular, this is the situation for the free fermion, which is everywhere trivially gapped except at the point $m = 0$. Thus, there is a non-trivial vacuum in the m-plane, but not necessarily a phase transition.

- Non-trivial strings: We can also consider space-dependent couplings where a two-cycle in spacetime wraps the $\mathbb{S}^2$ of mass parameters. For simplicity we consider a situation where the bulk is trivially gapped for generic m. In this case the anomaly (4.57) implies that there is a non-trivial effective string in the transverse space.[17] Specifically, by integrating the anomaly polynomial we find that wrapping n times leads to an anomaly for the effective theory along the string

$$in\int_{Y_3} CS_{\rm grav}\ . \tag{4.59}$$

Thus, the $2d$ theory on the string is gapless with chiral central charge $c_L - c_R = n/2$.

In the special case of the free fermion this conclusion can be readily verified by solving the Dirac equation in a background with position dependent mass as in [147, 148, 13], where one finds that the string supports n Majorana-Weyl fermions in agreement with the general index theorems of [171, 172].

As a simple special case of these general results, consider the mass profile

$$m(r,\theta) = \alpha r e^{i\theta}\ , \tag{4.60}$$

where (r,θ) parameterize a plane in radial coordinates and the string is localized along $r = 0$. We can split the $4d$ Weyl fermion into a left-moving $2d$ fermion ψ_1 and a right-moving $2d$ fermion ψ_2. Then one can check that in the mass profile (4.60) the field ψ_2 has no normalizable solutions and ψ_1 has only one normalizable solution

$$\psi_1 = ce^{-i\pi/4}e^{-\frac{1}{2}\alpha r^2}\ , \tag{4.61}$$

with a real coefficient c. Quantizing c leads to one Majorana-Weyl fermion on the string worldvolume with chiral central charge $1/2$ as expected.

[17]Following our discussion in the introduction, these are smooth external disturbances of the system, which are universal. These are not dynamical strings. If m becomes a dynamical field, then, depending on the details of the theory, these strings could be stable dynamical objects.

4.4 QED$_2$

In this section we explore the coupling anomalies in $2d$ $U(1)$ gauge theories. These models have a θ-parameter and accordingly our analysis is similar to section 4.2.

4.4.1 $2d$ Abelian Gauge Theory

We begin with $2d$ $U(1)$ gauge theory without charged matter. The Euclidean action is:

$$\mathcal{S} = \int \frac{1}{2g^2} da \wedge *da - \frac{i}{2\pi} \int \theta da \,. \tag{4.62}$$

Since the integral of da is quantized, the transformation $\theta \to \theta + 2\pi$ does not affect the correlation functions of local operators at separated points. However, below we will show that the theories at θ and $\theta + 2\pi$ are only equivalent up to an invertible field theory.

The theory has a $U(1)$ one-form global symmetry associated to the two-form current $J \sim da$. This symmetry acts on the dynamical variable as $a \to a + \epsilon$ where ϵ is a flat connection. We can turn on a two-form background gauge field B for this symmetry leading to the action

$$\mathcal{S} = \int \frac{1}{4g^2}(da - B) \wedge *(da - B) - \frac{i}{2\pi} \int \theta(da - B) - ik \int B \,, \tag{4.63}$$

where the coefficient k of the counterterm is an integer. This action is invariant under background gauge transformation

$$a \to a + \Lambda, \quad B \to B + d\Lambda \,, \tag{4.64}$$

where Λ is a $U(1)$ one-form gauge field. As in the comment following (4.20), we cannot use the classical Λ to set the dynamical field a to zero.

In the presence of nontrivial background gauge field B, the partition function $Z[\theta, B]$ is not invariant under $\theta \to \theta + 2\pi$. Instead, it satisfies

$$\frac{Z[\theta + 2\pi, B]}{Z[\theta, B]} = \exp\left(-i \int B\right) . \tag{4.65}$$

This difference can be interpreted as an anomaly between the coupling θ and the $U(1)$ one-form global symmetry.

One can understand the anomaly more physically in terms of pair creation of probe particles, as in [164]. Adding to the action a θ term with coefficient 2π is equivalent to adding a Wilson line describing a pair of oppositely charged particles, which are created and then separated and moved to the boundary of spacetime. These particles screen the background electric field created by θ, which is the physical reason for the 2π periodicity. However, when we take into account the one-form charge, the particle pair can be detected and this gives rise to the anomaly.

Extending the backgrounds θ and B into a $3d$ bulk Y we can introduce a new partition function

$$\tilde{Z}[\theta, B] = Z[\theta, B] \exp\left(i \int \theta \, \frac{dB}{2\pi}\right) , \tag{4.66}$$

which is invariant under $\theta \to \theta + 2\pi$.

Spacetime Dependent θ

The anomaly can also be detected by promoting the coupling constant θ to be a variable function from spacetime to a circle. As in the discussion in section 4.2, our first task is to define more precisely the integral of θda (and also the integral in (4.66)).

Here, we can proceed as in section 4.2.1 and define the integral using patches. (This discussion seems more awkward than in section 4.2.1, but it is essentially the same as there.) Explicitly, we first cover spacetime by a collection of patches $\{U_I\}$. The circle-valued function θ can be lifted to real-valued functions on patches and transition functions between the patches:

$$\{\theta_I : U_I \to \mathbb{R}\} \quad \text{and} \quad \{n_{IJ} : U_I \cap U_J \to \mathbb{Z}\} , \qquad \text{with} \quad \theta_I - \theta_J = 2\pi n_{IJ} \text{ on } U_I \cap U_J . \tag{4.67}$$

This data is redundant. If we modify

$$\theta_I \to \theta_I + 2\pi m_I, \qquad n_{IJ} \to n_{IJ} + m_I - m_J\,, \tag{4.68}$$

with integer m_I, we describe the same underlying circle-valued function θ. Similarly the $U(1)$ gauge field a can be lifted into the following data

$$\{a_I : U_I \to \Omega^1(U_I)\}, \quad \{\phi_{IJ} : U_I \cap U_J \to \mathbb{R}\} \quad \text{and} \quad \{n_{IJK} : U_I \cap U_J \cap U_J \to \mathbb{Z}\}\,, \tag{4.69}$$

where $\Omega^1(U_I)$ is the space of real differential one-forms on U_I. The lifts satisfy the following consistency conditions

$$\begin{aligned} U_I \cap U_J : a_I - a_J = d\phi_{IJ}\,, \\ U_I \cap U_J \cap U_K : \phi_{JK} + \phi_{KI} + \phi_{IJ} = 2\pi n_{IJK}\,, \end{aligned} \tag{4.70}$$

and there is a redundancy coming from gauge transformation

$$a_I \to a_I + d\lambda_I, \quad \phi_{IJ} \to \phi_{IJ} + \lambda_I - \lambda_J + 2\pi m_{IJ}, \quad n_{IJK} \to n_{IJK} + m_{JK} + m_{KI} + m_{IJ}\,, \tag{4.71}$$

where λ_I are real functions on U_I and m_{IJ} are integers.

To define the integral, we need to pick a partition of spacetime into closed sets $\{\sigma_I\}$ with the properties: $\sigma_I \subset U_I$, $\sigma_{IJ} = (\partial\sigma_I \cap \partial\sigma_J) \subset U_I \cap U_J$ and $\sigma_{IJK} = (\partial\sigma_{IJ} \cap \partial\sigma_{JK} \cap \partial\sigma_{KI}) \subset U_I \cap U_J \cap U_K$. We define the integral of θda in terms of the lifted data and the partition $\{\sigma_I\}$ as

$$\begin{aligned} \exp\left(\frac{i}{2\pi}\int \theta da\right) \equiv \exp\Bigg(&\frac{i}{2\pi}\sum_I \int_{\sigma_I} \theta_I da_I \\ &-i\sum_{I<J}\int_{\sigma_{IJ}} n_{IJ} a_J + i\sum_{I<J<K} n_{IJ}\phi_{JK}|_{\sigma_{IJK}}\Bigg)\,. \end{aligned} \tag{4.72}$$

The first term in the right hand side is the naive expression. The second term is analogous to the similar term in (4.16) and the last term is needed to make the answer invariant under gauge transformations of a. One can check that this integral is independent of the choice of partitions $\{\sigma_I\}$ and the lifts of θ and A.

Similarly, the integral (4.66) should be defined more carefully when θ varies in spacetime.

In a configuration where θ has non-trivial winding number along some one-cycle the resulting integral breaks the one-form global symmetry. As an illustration, consider a simple situation where spacetime is a torus with one-cycles x and y and θ has winding number m around x and is independent of y. If we restrict to a sector with $\int_{T^2} da = 0$ then we have

$$\exp\left(\frac{i}{2\pi}\int_{T^2}\theta da\right) = \exp\left(im\oint_y a\right) \ . \tag{4.73}$$

The Wilson line on the right-hand side above is charged under the one-form symmetry (4.64) thus illustrating the breaking. One way to think about this breaking is to note that for this configuration of spacetime dependent θ nonzero correlation functions must involve an appropriate net number of Wilson lines circling the y-cycle.

As in our previous discussion, we can restore the invariance under the one-form symmetry by coupling to a bulk using the partition function $\tilde{Z}$ in (4.66). For instance, in the torus example above we can extend the background fields to a three-manifold Y which is a solid torus with the cycle y filled in to a disk D. We then evaluate the anomaly[18]

$$\exp\left(i\int_Y \theta\frac{dB}{2\pi}\right) = \exp\left(-im\int_D B\right) \ . \tag{4.74}$$

The combination of (4.73) and (4.74) is then invariant under the one-form gauge transformations (4.64).

We can also express this violation of the one-form symmetry and the anomaly (4.65) in terms

[18]The equation (4.74) is correct up to a boundary term $\exp\left(\frac{i}{2\pi}\int_{T^2}\theta B\right)$ which cancels against a similar term in the action (4.63).

of a four-form using the descent procedure

$$\mathcal{I}_4 = \frac{1}{(2\pi)^2} d\theta \wedge dB \ . \tag{4.75}$$

We remark that as discussed above, $d\theta$ can be replaced by $d(\theta + f(\theta))$ with an arbitrary 2π-periodic function $f(\theta)$.

Dynamics

The $2d$ $U(1)$ gauge theory has no local degrees of freedom – it is locally trivial. In the spirit of the 't Hooft anomaly matching the non-trivial anomaly in (4.66) or equivalently (4.75) must be reproduced by its effective description. As a result, the theory cannot be completely trivial for all values of θ. Indeed, as we will now review, it has a first order phase transition at $\theta = \pi$.

We can say more about the dynamics using charge conjugation $\mathcal{C}$, which is a symmetry when $\theta = 0$ or $\theta = \pi$. This symmetry acts as

$$\mathcal{C}(a) = -a, \quad \mathcal{C}(B) = -B \, . \tag{4.76}$$

At $\theta = \pi$, the charge conjugation symmetry is accompanied by a 2π-shift of θ and this leads to a mixed anomaly between $\mathcal{C}$ and the one-form symmetry [27, 158]. Indeed, using (4.65) we see that when $\theta = \pi$, a $\mathcal{C}$ transformation acts on the partition function as

$$Z[\pi, B] \to Z[\pi, -B] = Z[-\pi, -B] \exp\left(i \int B\right) = Z[\pi, B] \exp\left((1-2k) i \int B\right) \ , \tag{4.77}$$

and we cannot choose k to remove this transformation since k is required to be integral. This obstruction characterizes the $\mathcal{C}$ anomaly. As above, this anomaly can be written using inflow as

$$\mathcal{A}(C, B) = \exp\left(\frac{i}{2} \int_Y dB + C \cup B\right) \ , \tag{4.78}$$

where C is a $\mathbb{Z}_2$ gauge field for charge conjugation (with holonomies 0, π).[19]

The anomaly involving $\mathcal{C}$ at $\theta = \pi$ implies that the long-distance theory for this value of θ cannot be trivially gapped. This agrees with the fact that the charge conjugation symmetry $\mathcal{C}$ at $\theta = \pi$ is spontaneously broken. The $U(1)$ one-form symmetry cannot be spontaneously broken in $2d$ and the theory is gapped at long distance. Thus, the anomaly can only be saturated by the spontaneously broken charge conjugation symmetry. The anomalies and their consequences are summarized in Table 4.3.

Of course, this system is exactly solvable and this analysis of its symmetries and anomalies does not lead to any new results. However, as we will soon see, the same reasoning leads to new results in more complicated systems, which are not exactly solvable.

The second class of consequences is associated with defects where θ varies in space. Let us first place the theory on $\mathbb{S}^1 \times \mathbb{R}$ with a constant θ. The effective quantum mechanics is the particle on a circle studied in section 4.2 with $q = \oint A$ the holonomy of A along the circle. The anomaly involving θ discussed above reduces to the anomaly (4.24) between θ and the $U(1)$ global symmetry in the quantum mechanics.

Next, we also let θ vary along the $\mathbb{S}^1$ direction and insert Wilson lines $\exp(i \int k_I A(x_I))$ along the $\mathbb{R}$ direction. In Lorentzian signature, the path integral over A_t imposes the Gauss constraint

$$\partial_x F_{xt} = g^2 \left(\sum k_I \delta(x - x_I) - \frac{\partial_x \theta(x)}{2\pi} \right) , \tag{4.79}$$

and therefore $\partial_x \theta$ can be interpreted as a space-dependent background charge density. Integrating the constraint we learn that the total background charge density has to vanish

$$2\pi \sum k_I - \int \partial_x \theta = 0 . \tag{4.80}$$

This implies that the theory is not consistent on a compact space where θ has a nontrivial

[19] If $C \to Y$ is the double cover defining the $\mathbb{Z}_2$ gauge field, then because of the twisting (4.76) the characteristic class of B lies in twisted cohomology: $[B] \in H^3(Y; \mathbb{Z}_C)$. The partition function is the value of the mod 2 reduction $\overline{[B]} \in H^3(Y; \mathbb{Z}_2)$ on the fundamental class of Y.

theory symmetry $\mathcal{G}$	without $\mathcal{C}$	with $\mathcal{C}$ at $\theta = 0, \pi$	
	θ-$\mathcal{G}$ anomaly	$\mathcal{C}$-$\mathcal{G}$ anomaly at $\theta = \pi$	no smooth counterterm
$U(1)$ gauge theory $\mathcal{G} = U(1)^{(1)}$	✓	✓	✓
with 1 charge p scalar $\mathcal{G} = \mathbb{Z}_p^{(1)}$, $p > 1$	✓	even p ✓ odd p ✗	even p ✓ odd p ✓
with N charge 1 scalar $\mathcal{G} = PSU(N)^{(0)}$, $N > 1$	✓	even N ✓ odd N ✗	even N ✓ odd N ✓

Table 4.3: Summary of anomalies and existence of continuous counterterms preserving $\mathcal{C}$ ("global inconsistency") in various $2d$ theories. The superscripts of the symmetries label the q's of q-form symmetries. All these theories have a charge conjugation symmetry, $\mathcal{C}$, at $\theta = 0, \pi$. Without using the charge conjugation symmetry, all these theories exhibit a mixed anomaly involving the coupling θ and some global symmetry $\mathcal{G}$. The anomaly implies that the long distance theory cannot be trivially gapped everywhere between θ and $\theta + 2\pi$. By including $\mathcal{C}$ we see that the theories can have a mixed anomaly between $\mathcal{C}$ and some global symmetry $\mathcal{G}$ at $\theta = \pi$. Such an anomaly forbids the long distance theories to be trivially gapped at $\theta = \pi$. Even if the theories have no mixed anomaly at $\theta = \pi$, there may be no smooth counterterms that preserve $\mathcal{C}$ simultaneously at $\theta = 0$ and $\theta = \pi$. Finally, we can deform these systems and break $\mathcal{C}$. Then the results in the "without $\mathcal{C}$" column are still applicable. The only difference is that we do not know at what value of θ the transition takes place.

winding unless there are Wilson lines inserted to absorb the charge.

4.4.2 QED$_2$ with one charge p scalar

We now add to the theory a scalar of charge p. (See [173–176] for early discussion of this theory.) The Euclidean action becomes

$$\mathcal{S} = \int \frac{1}{2g^2} da \wedge * da - \frac{i}{2\pi} \int \theta da + \int d^2x \left(|D_{pa}\phi|^2 + V(|\phi|^2) \right) . \tag{4.81}$$

The charge p scalar breaks the $U(1)$ one-form symmetry to a $\mathbb{Z}_p$ one-form symmetry [5]. As before, we can activate the background gauge field $K \in H^2(X, \mathbb{Z}_p)$ for this symmetry and the modified action includes

$$\mathcal{S} \supset -\frac{i}{2\pi} \int \theta \left(da - \frac{2\pi}{p} K \right) - \frac{2\pi i k}{p} \int K \,, \tag{4.82}$$

where the coefficient of the counterterm k is an integer modulo p.

The theory has a mixed anomaly between the coupling θ and the $\mathbb{Z}_p$ one-form symmetry. The $3d$ anomaly is

$$\mathcal{A}(\theta, K) = \exp\left(-2\pi i \int \frac{d\theta}{2\pi}\frac{K}{p}\right)\,, \tag{4.83}$$

(see the discussion below (4.29) for comments on boundary terms.) The anomaly forces the long distance theory to be nontrivial for at least one-point between θ and $\theta + 2\pi$.

The same conclusion can be drawn from Hamiltonian formalism. We can decompose the Hilbert space of the theory into superselection sectors according to the $\mathbb{Z}_p$ one-form symmetry [176]

$$\mathcal{H} = \bigoplus_{n=1}^{p} \mathcal{H}_n\,. \tag{4.84}$$

Intuitively, transitions using Coleman's pair-creation mechanism [164] using the dynamical quanta can change θ by $2\pi p$ and hence they take place within one of the subspaces in (4.84). But transitions between states in different subspaces labeled by different values of n in (4.84) can take place only using probe particles. As a result, all the subspaces in (4.84) are in the same theory but time evolution preserves the subspace [176].

The Hilbert spaces at θ and $\theta+2\pi$ are isomorphic but the superselection sectors are shuffled. In particular this means that the vacuum at θ is no longer the vacuum at $\theta + 2\pi$ and therefore the long distance theory cannot be trivial everywhere between θ and $\theta + 2\pi$.

For smooth interfaces interpolating between θ and $\theta + 2\pi n$, the bulk anomaly (4.83) yields the effective anomaly of the interface

$$\mathcal{A}(K) = \exp\left(-2\pi i n \int \frac{K}{p}\right)\,. \tag{4.85}$$

Implications of $\mathcal{C}$ Symmetry

The discussion above did not make use of the charge conjugation symmetry $\mathcal{C}$ at $\theta = 0, \pi$, and as usual we can say more using this additional symmetry. $\mathcal{C}$ acts as

$$\mathcal{C}(a) = -a, \quad \mathcal{C}(\phi) = \phi^*, \quad \mathcal{C}(K) = -K\,, \tag{4.86}$$

and at $\theta = \pi$, the partition function transforms as

$$Z[\pi, K] \to Z[\pi, K] \exp\left(2\pi i \frac{1-2k}{p} \int K\right)\,. \tag{4.87}$$

Since the coefficient of the counterterm is an integer modulo p, the partition function transforms non-anomalously if there is an integer k that solves

$$2k = 1 \bmod p\,. \tag{4.88}$$

For even p, there are no solutions and the charge conjugation symmetry has a mixed anomaly with the $\mathbb{Z}_p$ one-form symmetry at $\theta = \pi$.[20] The anomaly enforces non-trivial long distance physics at $\theta = \pi$.

For odd p, the condition (4.88) can be solved by $k = \frac{p+1}{2}$ so there is no anomaly at $\theta = \pi$. We can however make a weaker statement by noticing that the counterterm that preserves charge conjugation symmetry at $\theta = 0$, has coefficient $k = 0$ and it differs from the choice of counterterm at $\theta = \pi$. Similar phenomena have been discussed in [27, 158, 160]. In [160, 162], this situation was referred to as a "global inconsistency." Concretely it means that there is no continuously varying (θ-dependent) counterterm that preserves $\mathcal{C}$ at both $\theta = 0$ and $\theta = \pi$. This again implies that the long-distance theory is non-trivial for at least one value of θ in $[0, \pi]$. This discussion is summarized in Table 4.3.

[20] The anomaly is $\mathcal{A}(C, K) = \exp\left(i\pi \int C \cup K\right)$ where C is the $\mathbb{Z}_2$ charge conjugation gauge field. Note that this is meaningful only when p is even.

All these constraints are saturated by spontaneously broken charge conjugation symmetry at $\theta = \pi$. The special value $p = 1$ deserves further comment. In this case, there is no one-form symmetry so the constraints described above no longer hold. If the scalar is very massive, the theory is effectively a pure $U(1)$ gauge theory so the theory is gapped for generic θ and the charge conjugation symmetry is spontaneously broken at $\theta = \pi$ leading to a first order phase transition. On the other hand, if the scalar condenses, the gauge field is Higgsed and the theory is trivially gapped for all θ.[21] Therefore, the line of first order phase transitions at $\theta = \pi$ must end at some intermediate value of the mass where the theory is gapless.

4.4.3 QED$_2$ with N charge 1 scalars

We now add N charge 1 scalars into the $U(1)$ gauge theory. The Euclidean action is

$$\mathcal{S} = \int \frac{1}{2g^2} da \wedge * da - \frac{i}{2\pi} \int \theta da + \int d^2x \left[\sum_{I=1}^{N} |D_a \phi_I|^2 + V\left(\sum_{I=1}^{N} |\phi_I|^2 \right) \right] . \tag{4.89}$$

If the potential $V(\sum |\phi_I|^2)$ has a minimum at $\sum |\phi_I|^2 \neq 0$ and is sufficiently steep, the above theory flows to a $\mathbb{CP}^{N-1} = \frac{U(N)}{U(N-1)\times U(1)}$ non-linear sigma model.[22]

The $U(1)$ one-form symmetry is now completely broken. Instead the theory has a $PSU(N) \cong SU(N)/\mathbb{Z}_N$ zero-form global symmetry that acts as $\phi_I \to \mathcal{G}_{IJ}\phi_J$. The reason the symmetry is $PSU(N)$ and not simply $SU(N)$ is that the $\mathbb{Z}_N$ transformation $\phi_I \to e^{2\pi i/N}\phi_I$ coincides with a $U(1)$ gauge transformation and hence acts trivially on all gauge invariant local operators.

Let us consider the system in the presence of a background gauge field A for the $PSU(N)$ global symmetry. The correlation of center of $SU(N)$ with the dynamical $U(1)$ gauge group

[21] In the limit of large scalar expectation value the smooth θ-dependence of various observables is reliably computed using instanton methods. These techniques are not reliable in the opposite limit of large mass for the scalar. And indeed, in that limit the θ-dependence is not smooth.

[22] We can easily generalize our analysis below to systems with several $U(1)$ gauge fields and various charged scalars. In that case the systems have several θ-parameters. Recently studied examples include systems that flow to $2d$ sigma-models whose target space is the flag manifold $\frac{U(N)}{U(N_1)\times\cdots\times U(N_m)}$ [177, 161, 178, 179].

means that a and A combine to a connection for the group

$$U(N) = \frac{SU(N) \times U(1)}{\mathbb{Z}_N} \ . \tag{4.90}$$

Crucially this means that in a general $PSU(N)$ background, a is no longer a $U(1)$ connection with properly quantized fluxes. Instead we have

$$\oint \frac{da}{2\pi} = \oint \frac{w_2(A)}{N} \text{ mod } 1 \,, \tag{4.91}$$

where $w_2(A) \in H^2(X, \mathbb{Z}_N)$ is the second Stiefel-Whitney class of the $PSU(N)$ bundle. Equivalently, in the presence of general $PSU(N)$ backgrounds there are fractional instantons.

Because of these fractional instantons, the partition function is no longer invariant under $\theta \to \theta + 2\pi$. Rather, θ has an extended periodicity of $2\pi N$ [27,158,159]. This represents a mixed anomaly between the 2π-periodicity of θ and the $PSU(N)$ global symmetry. The corresponding $3d$ anomaly is

$$\mathcal{A}(\theta, A) = \exp\left(2\pi i \int \frac{d\theta}{2\pi} \frac{w_2}{N}\right) \ , \tag{4.92}$$

(see the discussion below (4.29) for a comment on the boundary terms). The anomaly implies that the long distance theory cannot be trivially gapped everywhere between θ and $\theta + 2\pi$.

Like the discussion in sections 4.4.1 and 4.4.2, we can understand this anomaly physically in terms of particle pair creation following [164]. The θ-term with coefficient 2π can be screened to $\theta = 0$ by pair creation of dynamical quanta. (Note that in the discussion in section 4.4.1 we used probe particles, and in section 4.4.2 we discussed the effects of both dynamical and probe quanta.) These quanta transform projectively under $PSU(N)$ and hence the screening leads to such projective representations at the boundary of space. More mathematically, this is the meaning of the selection rule (4.91).

It is interesting to compare this discussion with the anomaly between θ-periodicity and the one-form global symmetry in section 4.4.2. The role of the background two-form $\mathbb{Z}_p$ gauge field

K there is played here by the background w_2 associated with the zero-form $PSU(N)$ global symmetry.

Implications of $\mathcal{C}$ Symmetry

We can further constrain the long distance theory using the charge conjugation symmetries $\mathcal{C}$ at $\theta = 0, \pi$, which acts as

$$\mathcal{C}(a) = -a, \quad \mathcal{C}(\phi_I) = \phi_I^*, \quad \mathcal{C}(A) = -A, \quad \mathcal{C}(w_2(A)) = -w_2(A)\,. \tag{4.93}$$

We can add to the theory a counterterm

$$\mathcal{S} \supset -2\pi i \frac{k}{N} \int w_2\,. \tag{4.94}$$

At $\theta = \pi$, the charge conjugation symmetry involves a 2π-shift of θ and it transforms the partition function as

$$Z[\pi, A] \to Z[\pi, A] \exp\left(2\pi i \frac{1-2k}{N} \int w_2\right)\,. \tag{4.95}$$

Similar to the example of QED_2 with one charge p scalar discussed in the last subsection, the above means that charge conjugation symmetry has a mixed anomaly with the $PSU(N)$ global symmetry for even N.[23] Meanwhile for odd N, there is no continuous counterterm preserving $\mathcal{C}$ at both $\theta = 0$ and $\theta = \pi$. For even N, the anomaly forces a non-trivial long distance theory at $\theta = \pi$, while for odd N we find a non-trivial theory for at least one value of θ.

These constraints agree with the common lore. For $N \geq 2$, the theory is believed to be gapped at generic θ except at $\theta = \pi$. For $N = 2$, the model at $\theta = \pi$ flows to the $SU(2)_1$ WZW model [180]. For $N > 2$, the charge conjugation symmetry is believed to be spontaneously broken at $\theta = \pi$ [181]. (The model with $N = 1$ was discussed above.)

Finally, we can consider a smooth interface between θ and $\theta + 2\pi n$. Assuming the theory

[23] The anomaly is $\exp\left(i\pi \int C \cup w_2\right)$ where C is the charge conjugation gauge field. This is meaningful only when N is even.

is gapped for generic θ, at long distances there is then an isolated quantum mechanics on the interface. The anomaly (4.92) implies that the quantum mechanical model has a non-trivial anomaly for the $PSU(N)$ global symmetry encoded by

$$\mathcal{A}(A) = \exp\left(2\pi i \int n \frac{w_2}{N}\right) . \tag{4.96}$$

This means that the ground states of the quantum mechanics are degenerate, and they form a projective representation of the $PSU(N)$ symmetry (i.e. a representation of $SU(N)$) with N-ality n. Intuitively, the interface is associated with n Φ_I quanta. But since they are strongly interacting we cannot determine their precise state except their N-ality.

4.5 $4d$ Yang-Mills Theory

In this section we compute the anomaly of $4d$ Yang-Mills theories with simply connected and simple gauge groups. We use our results to compute the anomaly on interfaces with spatially varying θ.

4.5.1 $SU(N)$ Yang-Mills Theory

We begin with the $4d$ $SU(N)$ gauge theory with the Euclidean action

$$S = -\frac{1}{4g^2}\int \mathrm{Tr}(F \wedge *F) - \frac{i\theta}{8\pi^2}\int \mathrm{Tr}(F \wedge F) . \tag{4.97}$$

Since the instanton number is quantized, the transformation $\theta \to \theta + 2\pi$ does not affect correlation functions of local operators at separated points, but it may affect more subtle observables such as contact terms involving surface operators.

The theory has a $\mathbb{Z}_N$ one-form symmetry that acts by shifting the connection by a $\mathbb{Z}_N$ connection [5]. We can turn on a background $\mathbb{Z}_N$ two-form gauge field $B \in H^2(X, \mathbb{Z}_N)$ for this one-form symmetry. In the presence of this background gauge field, the $SU(N)$ bundle is twisted into a $PSU(N)$ bundle with fixed second Stiefel-Whitney class $w_2(a) = B$ [29, 5].

The instanton number of a $PSU(N)$ bundle can be fractional. Therefore with a nontrivial background B the partition function at $\theta + 2\pi$ and θ can be different [29, 5, 27]

$$\frac{Z[\theta + 2\pi, B]}{Z[\theta, B]} = \exp\left(2\pi i \frac{N-1}{2N} \int \mathcal{P}(B)\right) , \tag{4.98}$$

where $\mathcal{P}$ is the Pontryagin square operation.[24] Thus, the theories at θ and $\theta + 2\pi$ differ by an invertible field theory, which can be detected by the contact terms of the two-dimensional symmetry operators of the $\mathbb{Z}_N$ one-form symmetry [29].

We can also add to the theory a counterterm

$$\mathcal{S} \supset -2\pi i \int_X \frac{p}{2N} \mathcal{P}(B) . \tag{4.99}$$

The coefficient p is an integer modulo $2N$ for even N and it is an even integer modulo $2N$ for odd N. The difference between θ and $\theta + 2\pi$ in (4.98) can be summarized into the following identification [29, 5, 27, 28]

$$(\theta, p) \sim (\theta + 2\pi, p + 1 - N) . \tag{4.100}$$

This means that θ has an extended periodicity of $4\pi N$ for even N and $2\pi N$ for odd N.

As discussed in detail in [3], the above phenomenon can be interpreted as a mixed anomaly between the 2π-periodicity of θ and the $\mathbb{Z}_N$ one-form symmetry. The corresponding anomaly action is

$$\mathcal{A}(\theta, B) = \exp\left(2\pi i \frac{N-1}{2N} \int \frac{d\theta}{2\pi} \mathcal{P}(B)\right) . \tag{4.101}$$

This anomaly implies that the long distance theory cannot be trivially gapped everywhere between θ and $\theta + 2\pi$.

We can further constrain the long distance theory using the time-reversal symmetry T at $\theta = 0, \pi$ following [27]. In a nontrivial background B, the time-reversal symmetry transforms

[24]For odd N, $\mathcal{P}(B) = B \cup B \in H^4(X, \mathbb{Z}_N)$. For even N, $\mathcal{P}(B) \in H^4(X, \mathbb{Z}_{2N})$ and reduces to $B \cup B$ modulo N.

<table>
<tr><td>theory
symmetry $\mathcal{G}$</td><td>without T
θ-$\mathcal{G}$ anomaly</td><td colspan="2">with T at $\theta = 0, \pi$</td></tr>
<tr><td></td><td></td><td>T-$\mathcal{G}$ anomaly at $\theta = \pi$</td><td>no smooth counterterms</td></tr>
<tr><td>$SU(N)$ gauge theory
$\mathcal{G} = \mathbb{Z}_N^{(1)}$</td><td>✓</td><td>even N ✓
odd N ✗</td><td>even N ✓
odd N ✓</td></tr>
<tr><td>with adjoint scalars
$\mathcal{G} = \mathbb{Z}_N^{(1)}$</td><td>✓</td><td>no T symmetry
in general</td><td>no T symmetry
in general</td></tr>
</table>

Table 4.4: Summary of anomalies and existence of continuous counterterms in various $4d$ theories. The superscripts of the symmetries label the q's of q-form symmetries.

the partition function as

$$Z[\pi, B] \to Z[\pi, B] \exp\left(2\pi i \frac{1 - N - 2p}{2N} \int \mathcal{P}(B)\right) . \tag{4.102}$$

A T anomaly occurs if there is no value of p such that the partition function is exactly invariant i.e. only if

$$1 - N - 2p = 0 \bmod 2N \tag{4.103}$$

has no integral solutions p. This is the case for even N, and hence for even N there must be non-trivial long distance physics at $\theta = \pi$ [27].[25] For odd N, we can solve the equation above with p even by taking

$$p = \begin{cases} \frac{1-N}{2} & N = 1 \mod 4 \ , \\ \frac{1+N}{2} & N = 3 \mod 4 \ . \end{cases} \tag{4.104}$$

Therefore, for odd N there is no T anomaly. However, for odd N the counterterm that preserves time-reversal symmetry at $\theta = 0$ has coefficient $p = 0 \bmod 2N$ and it is different from the one at $\theta = \pi$. This means that there is no continuous counterterm that preserves T at both $\theta = 0$ and π [27]. (Such reasoning was named a "global inconsistency" in [160, 162].) This again implies non-trivial long distance physics for at least one value of θ.

These results agree with the standard lore about Yang-Mills theory. For all values of θ the

[25] In this case the anomaly ω is $\frac{1}{2}\tilde{w}_1 \cup \mathcal{P}(B)$, where $\tilde{w}_1 \in H^1(Y, \tilde{\mathbb{Z}})$ denotes the natural integral uplift of the Stiefel-Whitney class w_1 with twisted integral coefficients. For further recent discussion of time-reversal anomalies in Yang-Mills theories see also [182, 183].

theory is confined (so the $\mathbb{Z}_N$ one-form symmetry is unbroken [5]) and gapped. For $\theta \neq \pi$ there is a unique vacuum. While at $\theta = \pi$ the T symmetry is spontaneously broken leading to two degenerate vacua and hence a first order phase transition.

We can also use the anomaly (4.101) to constrain the worldvolume of interfaces where θ varies. Consider a smooth interface between θ and $\theta + 2\pi k$. Assuming that the $SU(N)$ gauge theory is gapped at long distances, the interface supports an isolated $3d$ quantum field theory. The anomaly (4.101) implies that the interface theory has an anomaly associated to the $\mathbb{Z}_N$ one-form symmetry described by [5, 28, 1]

$$\mathcal{A}(B) = \exp\left(2\pi i k \int \frac{N-1}{2N}\mathcal{P}(B)\right) . \tag{4.105}$$

The anomaly can be saturated for instance, by an $SU(N)_k$ Chern-Simons theory or a $(\mathbb{Z}_N)_{-N(N-1)k}$ discrete gauge theory [1].

4.5.2 Adding Adjoint Higgs Fields

The mixed anomaly between the 2π-periodicity of θ and the $\mathbb{Z}_N$ one-form symmetry is robust under deformations that preserve the one-form symmetry. Note that such deformations generally break the time-reversal symmetry at $\theta = 0, \pi$. Below, we present two examples with different infrared behaviors that also saturate the anomaly by adding charged scalars in the adjoint representation.

As in [1], we can add one adjoint scalar to Higgs the $SU(N)$ gauge field to its Cartan torus $U(1)^{N-1}$ with gauge fields a_J. The $U(1)$ gauge fields are embedded in the $SU(N)$ gauge field through

$$a = \sum_{J=1}^{N-1} a_J T^J, \quad T^J = \text{diag}(0, \cdots, 0, \underbrace{+1}_{J\text{th entry}}, -1, 0, \cdots, 0) . \tag{4.106}$$

In the classical approximation, the low energy $U(1)^{N-1}$ gauge theory is described by the Euclidean action

$$\mathcal{S} = -\frac{1}{4g^2}\int \sum_{I,J=1}^{N-1} K_{IJ} da_I \wedge * da_J - \frac{i\theta}{8\pi^2}\int \sum_{I,J=1}^{N-1} K_{IJ} da_I \wedge da_J\,, \tag{4.107}$$

where K is the Cartan matrix of $SU(N)$. Small higher order quantum corrections renormalize the gauge coupling g and θ, but do not affect our conclusions.

The low-energy theory exhibits a spontaneously broken $U(1)^{N-1} \times U(1)^{N-1}$ one-form global symmetry. Most of it is accidental. The exact one-form symmetry is the symmetry in the UV, which is $\mathbb{Z}_N$. It acts on the infrared fields as

$$a_J \to a_J + \frac{2\pi J}{N}\epsilon\,, \tag{4.108}$$

where ϵ is a flat connection with $\mathbb{Z}_N$ holonomies. Activating the background gauge field $B \in H^2(X, \mathbb{Z}_N)$ for the one-form symmetry modifies the Euclidean action by replacing da_I with $da_I - \frac{2\pi I}{N}B$. This means that when θ is shifted by 2π, the partition function of the infrared theory transforms as

$$Z[\theta + 2\pi, B] = Z[\theta, B]\exp\left(2\pi i \frac{N-1}{2N}\int \mathcal{P}(B)\right)\,, \tag{4.109}$$

which agrees with the anomaly in the ultraviolet theory.

Note that this gapless $U(1)^{N-1}$ gauge theory reproduces the anomaly (4.101), without a phase transition.

Following [1], we can also add more adjoint scalars to Higgs the theory to a $\mathbb{Z}_N$ gauge theory. The $\mathbb{Z}_N$ gauge field c is embedded in the $SU(N)$ gauge field through (we work in continuous notation, i.e. c is a flat $U(1)$ gauge field with holonomies in $\mathbb{Z}_N$)

$$a = cT, \quad T = \text{diag}(1, \cdots, 1, -(N-1))\,. \tag{4.110}$$

The infrared theory is a topological field theory with Euclidean action

$$\mathcal{S} = \frac{iN}{2\pi}\int b \wedge dc - N(N-1)\frac{i\theta}{8\pi^2}\int dc \wedge dc\,, \tag{4.111}$$

where b is a dynamical $U(1)$ two-form gauge field and c is a dynamical $U(1)$ one-form gauge field. b acts as a Lagrange multiplier constraining c to be a $\mathbb{Z}_N$ gauge field. The equation of motion of b constrains c to be a $\mathbb{Z}_N$ gauge field that satisfies $Nc = d\phi$. The original $\mathbb{Z}_N$ one-form symmetry is spontaneously broken in the infrared. If we activate the background gauge field B for the $\mathbb{Z}_N$ one-form symmetry. The Euclidean action becomes

$$\mathcal{S} = \frac{iN}{2\pi}\int b \wedge \left(dc - \frac{2\pi}{N}B\right) - N(N-1)\frac{i\theta}{8\pi^2}\int \left(dc - \frac{2\pi}{N}B\right) \wedge \left(dc - \frac{2\pi}{N}B\right)\,. \tag{4.112}$$

As the coupling constant θ shifts by 2π, the partition function of the infrared theory transforms anomalously and agrees with the anomaly in ultraviolet theory. As in the gapless $U(1)^{N-1}$ theory discussed above, the anomaly is saturated in the IR without a phase transition.

We can also simplify the above $\mathbb{Z}_N$ gauge theory by shifting $b \to b + \frac{N-1}{4\pi}\theta dc$. The Euclidean action then becomes that of a standard $\mathbb{Z}_N$ gauge theory [32, 33]

$$\mathcal{S} = \frac{iN}{2\pi}\int b \wedge dc\,. \tag{4.113}$$

The dependance on θ now appears in the coupling of these fields to the background B and the partition function again transforms anomalously when $\theta \to \theta + 2\pi$ in agreement with (4.98) and the ultraviolet anomaly (4.101). Again, this is achieved in the IR without a phase transition.

4.5.3 Other Gauge Groups

We now discuss similar mixed anomalies involving the 2π-periodicity of θ and the center one-form symmetries in $4d$ Yang-Mills theories with other simply-connected gauge groups G. These anomalies constrain the long distance physics of these theories as well as smooth interfaces separating two regions with different θ's. As we will see, unlike the case of $SU(N)$, which we studied above, typically 2π shifts of θ do not allow us to scan all the possible values of the coefficient p of the $\mathcal{P}(B)$ counterterm.

The one-form global symmetry of any of these simply connected groups G is its center $Z(G)$. We couple it to a two-form gauge field B. This twists the gauge bundles to $G/Z(G)$ bundles with second Stiefel-Whitney (SW) class $w_2 = B$. These bundles support fractional instantons. Following [184], we will determine the relation between the fractional instantons and the background gauge fields B by evaluating the instanton number on a specific $G/Z(G)$ bundle. We will take it to be of a tensor product of various $SU(n)/\mathbb{Z}_n$ bundles, for which we already know the answer, and untwisted bundles of simply connected groups. We will generalize the discussion in [184] to non-spin manifolds.

We will discuss $Sp(N)$, $Spin(N)$, E_6 and E_7 gauge groups. The other simple Lie groups G_2, F_2, and E_8 have trivial center and therefore the corresponding gauge theories do not have similar anomalies.

$Sp(N)$ Gauge Theory

We start with a pure gauge $Sp(N)$ theory.[26] The theory has a $\mathbb{Z}_2$ one-form symmetry. We want to construct a $Sp(N)/\mathbb{Z}_2$ bundle with second SW class B. We do that by using the embedding $SU(2)^N \subset Sp(N)$ and then an $Sp(N)/\mathbb{Z}_2$ bundle is found by tensoring N $PSU(2)$ bundles each with second SW class B. Then the anomaly (4.101) implies that the $Sp(N)$ gauge theory has an anomaly

$$\mathcal{A}_{Sp(N)}(\theta, B) = \mathcal{A}_{SU(2)}(\theta, B)^N = \exp\left(2\pi i \int \frac{d\theta}{2\pi} \frac{N\mathcal{P}(B)}{4}\right) . \tag{4.114}$$

This means that a shift of θ by 2π shifts the coefficient p of the counterterm $2\pi i p \int \frac{\mathcal{P}(B)}{4}$ by N. Note that for even N not all the possible values of $p = 0, 1, 2, 3$ are scanned by shifts of θ by 2π. The anomaly becomes trivial when $N = 0$ mod 4 (on spin manifolds it is trivial when $N = 0$ mod 2).

[26] We use the notation $Sp(N) = USp(2N)$. Specifically $Sp(1) = SU(2)$ and $Sp(2) = Spin(5)$.

E_6 Gauge Theory

The theory has a $\mathbb{Z}_3$ one-form symmetry. Here we use the embedding $SU(3)^3 \subset E_6$. We can construct a $E_6/\mathbb{Z}_3$ bundle with second SW class B by tensoring an $SU(3)$ bundle, a $PSU(3)$ bundle with second SW class B, and a $PSU(3)$ bundle with second SW class $-B$. Then the anomaly (4.101) implies that the E_6 gauge theory has an anomaly

$$\mathcal{A}_{E_6}(\theta, B) = \mathcal{A}_{SU(3)}(\theta, B)^2 = \exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{2\mathcal{P}(B)}{3}\right) . \tag{4.115}$$

The anomaly is nontrivial and all possible values of p in the counterterm are scanned by shifts of θ by 2π.

E_7 Gauge Theory

The theory has a $\mathbb{Z}_2$ one-form symmetry. Here we use the embedding $SU(4)\times SU(4)\times SU(2) \subset E_7$. We can construct a $E_7/\mathbb{Z}_2$ bundles with second SW class B by tensoring an $SU(4)$ bundle, a $PSU(2)$ bundle with second Stifel-Whitney class B, and an $SU(4)/\mathbb{Z}_2$ bundle with second SW class B (which can be thought of as $SU(4)/\mathbb{Z}_4$ bundle with second SW class $2\widetilde{B}$ where the tilde denotes a lifting to a $\mathbb{Z}_4$ cochain and $2\widetilde{B}$ is independent of the lift). Then the anomaly (4.101) implies that the E_7 gauge theory has an anomaly

$$\mathcal{A}_{E_7}(\theta, B) = \mathcal{A}_{SU(2)}(\theta, B)\mathcal{A}_{SU(4)}(\theta, 2\widetilde{B}) = \exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{3\mathcal{P}(B)}{4}\right) . \tag{4.116}$$

Again, the anomaly is nontrivial and all possible values of p in the counterterm are scanned.

$Spin(N) = Spin(2n+1)$ Gauge Theory

For $N = 3$ this is the same as $SU(2)$, which was discussed above. So let us consider $N \geq 5$.

The theory has a $\mathbb{Z}_2$ one-form symmetry. Here we use the embedding $SU(2) \times SU(2) \times Spin(N-4) \subset Spin(N)$ (where the last factor is missing for $N = 5$). We can construct a $Spin(N)/\mathbb{Z}_2$ bundle with second SW class B by tensoring two $PSU(2)$ bundles each with second SW class B and a $Spin(N-4)$ bundle. Then the anomaly (4.101) implies that the $Spin(N) = Spin(2n+1)$ gauge theory has an anomaly

$$\mathcal{A}_{Spin(N)}(\theta, B) = \mathcal{A}_{SU(2)}(\theta, B)^2 = \exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{\mathcal{P}(B)}{2}\right) \quad \text{for} \quad N = 1 \bmod 2\,. \tag{4.117}$$

The anomaly is always nontrivial (but it is trivial on spin manifolds). A shift of θ by 2π shifts p by 2 and hence not all values of p are scanned by such shifts.

$Spin(N) = Spin(4n+2)$ **Gauge Theory**

For $N = 6$ this is the same as $SU(4)$ which was discussed above. So we will discuss here $N \geq 10$.

The theory has a $\mathbb{Z}_4$ one-form symmetry. We will use the embedding $Spin(6)\times Spin(4)^{n-1} \subset Spin(N)$. We can construct a $Spin(N)/\mathbb{Z}_4$ bundle with second SW class B by tensoring $(n-1)$ $SU(2)$ bundles, $(n-1)$ $PSU(2)$ bundles each with second SW class B mod 2 and a $PSU(4)$ bundle with second SW class B. Then the anomaly (4.101) implies that the $Spin(N) = Spin(4n+2)$ gauge theory has an anomaly[27]

$$\begin{aligned}\mathcal{A}_{Spin(N)}(\theta, B) =& \mathcal{A}_{SU(4)}(\theta, B)\mathcal{A}_{SU(2)}(\theta, B)^{n-1} \\ &= \exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{N\mathcal{P}(B)}{16}\right) \quad \text{for} \quad N = 2 \bmod 4\,.\end{aligned} \tag{4.118}$$

The anomaly is always nontrivial (even on spin manifolds). A shift of θ by 2π shifts p by $\frac{N}{2}$ and hence all values of p are scanned by such shifts.

If $B = 2\widehat{B}$ is even, we study $SO(N)$ bundles and the anomaly is

$$\exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{N\mathcal{P}(\widehat{B})}{4}\right) = \exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{\mathcal{P}(\widehat{B})}{2}\right)\ . \tag{4.119}$$

[27]The instanton number of a $Spin(4n+2)/\mathbb{Z}_4$ bundle with second SW class B is $\int \frac{2n+1}{4}\frac{\mathcal{P}(B)}{2}$ mod 1. On spin manifolds $\frac{\mathcal{P}(B)}{2} \in H^4(X, \mathbb{Z}_4)$ so for $N = 4n+2 = 2$ mod 8 the instanton number is $\int \frac{1}{4}\frac{\mathcal{P}(B)}{2}$ mod 1 and for $N = 4n+2 = 6$ mod 8 the instanton number is $-\int \frac{1}{4}\frac{\mathcal{P}(B)}{2}$ mod 1. For $N = 6$ mod 8, our determination of the fractional instanton number differs from [184] by a sign, which does not affect the conclusions of [184]. This sign change reverses the direction of the action of the modular T-transformation in Fig. 6 of [34] for $N = 6$ mod 8.

This is useful, e.g. when we add dynamical matter fields in the vector representation and the one-form global symmetry is only $\mathbb{Z}_2 \subset \mathbb{Z}_4$, which is coupled to $\widehat{B}$. In that case the anomaly vanishes on spin manifolds and a shift of θ by 2π shifts the coefficient p of the counterterm $2\pi ip \int \frac{\mathcal{P}(\widehat{B})}{4}$ by 2 and hence not all possible values of p are scanned. This is the same conclusion as for odd N (4.117).

$Spin(N) = Spin(4n)$ **Gauge Theory**

The theory has a $\mathbb{Z}_2^{(L)} \times \mathbb{Z}_2^{(R)}$ one-form symmetry. Here we use the embedding $SU(2)^{2n} \subset Spin(N)$.

For odd n we can construct a $Spin(N)/(\mathbb{Z}_2^{(L)} \times \mathbb{Z}_2^{(R)})$ bundle with second SW class B_L and B_R by tensoring n $PSU(2)$ bundles with second SW class B_L and n $PSU(2)$ bundles with second SW class B_R. Then the anomaly (4.101) implies that the $Spin(N) = Spin(4n)$ gauge theory for odd n has an anomaly

$$\begin{aligned}\mathcal{A}_{Spin(N)}(\theta, B_L, B_R) &= \mathcal{A}_{SU(2)}(\theta, B_L)^n \mathcal{A}_{SU(2)}(\theta, B_R)^n \\ &= \exp\left(2\pi i \int \frac{d\theta}{2\pi} \frac{N\big(\mathcal{P}(B_L) + \mathcal{P}(B_R)\big)}{16}\right) \qquad \text{for} \quad N = 4 \bmod 8\,.\end{aligned} \tag{4.120}$$

For even n we can construct the $Spin(N)/(\mathbb{Z}_2^{(L)} \times \mathbb{Z}_2^{(R)})$ bundle by tensoring an $SU(2)$ bundle, a $PSU(2)$ bundle with second SW class $B_L + B_R$, $n-1$ $PSU(2)$ bundles with second SW class B_L, and $n-1$ $PSU(2)$ bundles with second SW class B_R. Then the anomaly (4.101) implies that the $Spin(N) = Spin(4n)$ gauge theory for even n has an anomaly

$$\begin{aligned}\mathcal{A}_{Spin(N)}(\theta, B_L, B_R) &= \mathcal{A}_{SU(2)}(\theta, B_L)^{n-1} \mathcal{A}_{SU(2)}(\theta, B_R)^{n-1} \mathcal{A}_{SU(2)}(\theta, B_L + B_R) \\ &= \exp\left(2\pi i \int \frac{d\theta}{2\pi} \left(\frac{N\big(\mathcal{P}(B_L) + \mathcal{P}(B_R)\big)}{16} + \frac{B_L \cup B_R}{2}\right)\right) \qquad \text{for} \quad N = 0 \bmod 8\,.\end{aligned} \tag{4.121}$$

The two cases can be summarized as

$$\begin{aligned}&\mathcal{A}_{Spin(N)}(\theta, B_L, B_R)\\ &\quad = \exp\left(2\pi i \int \frac{d\theta}{2\pi}\left(\frac{N\mathcal{P}(B_L+B_R)}{16}+\frac{B_L\cup B_R}{2}\right)\right) \qquad \text{for} \quad N = 0 \bmod 4\,.\end{aligned} \tag{4.122}$$

The anomaly is always nontrivial (even on spin manifolds). A shift of θ by 2π shifts the coefficients of the counterterm $2\pi i p_L \int \frac{\mathcal{P}(B_L)}{4} + 2\pi i p_R \int \frac{\mathcal{P}(B_R)}{4} + 2\pi i p_{LR} \int \frac{B_L \cup B_R}{2}$ by $(p_L, p_R, p_{LR}) \to (p_L + \frac{N}{4}, p_R + \frac{N}{4}, p_{LR} + 1 + \frac{N}{4})$ and hence not all values of (p_L, p_R, p_{LR}) are scanned by such shifts.

As above, if we limit ourselves to $SO(N)$ bundles (as is the case, e.g. when we add dynamical matter fields in a vector representation), we study backgrounds with $B_L = B_R = \widehat{B}$. Then, the anomaly is

$$\exp\left(2\pi i \int \frac{d\theta}{2\pi}\frac{\mathcal{P}(\widehat{B}))}{2}\right) \tag{4.123}$$

and it vanishes on spin manifolds. A shift of θ by 2π shifts the value the coefficient p in $2\pi i p \int \frac{\mathcal{P}(\widehat{B})}{4}$ by 2 and again, not all values of p are scanned. This is the same conclusion as for odd N (4.117) and for $N = 2 \bmod 4$ (4.119).

A Check Using $3d$ TQFT or $2d$ RCFT Considerations

One way of viewing our anomaly is as the anomaly in a one-form global symmetry in the theory along interfaces separating θ and $\theta + 2\pi k$. General considerations show that in this case of a simple and semi-simple gauge group this anomaly can always be saturated by a Chern-Simons theory with gauge group G and level k. In $3d$ TQFTs, the anomaly of one-form symmetries can be determined by the spins of the lines generating the symmetry [1]. The one-form symmetries and the spins of the generating lines of various Chern-Simons theories with level 1 are summarized in Table 4.5. These results can be found by studying the $3d$ TQFT or by studying the corresponding $2d$ Kac- Moody algebra.

We can use these to check the anomaly we determined using $4d$ instantons above. When the one-form symmetry is $\mathbb{Z}_\ell$ it is generated by a line a such that $a^\ell = 1$. The coefficient of the

Gauge group G	Center $Z(G)$	Spins	Anomaly
$SU(N)$	$\mathbb{Z}_N$	$h_a = \frac{N-1}{2N}$	$\frac{N-1}{2N} \int d\theta\, \mathcal{P}(B)$
$Sp(N)$	$\mathbb{Z}_2$	$h_a = \frac{N}{4}$	$\frac{N}{4} \int d\theta\, \mathcal{P}(B)$
E_6	$\mathbb{Z}_3$	$h_a = \frac{2}{3}$	$\frac{2}{3} \int d\theta\, \mathcal{P}(B)$
E_7	$\mathbb{Z}_2$	$h_a = \frac{3}{4}$	$\frac{3}{4} \int d\theta\, \mathcal{P}(B)$
$Spin(N)$ $(N = 2n+1)$	$\mathbb{Z}_2$	$h_a = \frac{1}{2}$	$\frac{1}{2} \int d\theta\, \mathcal{P}(B)$
$Spin(N)$ $(N = 4n+2)$	$\mathbb{Z}_4$	$h_a = \frac{N}{16}$ $h_{a^2} = \frac{1}{2}$	$\frac{N}{16} \int d\theta\, \mathcal{P}(B)$
$Spin(N)$ $(N = 4n)$	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$h_a = h_b = \frac{N}{16}$ $h_{ab} = \frac{1}{2}$	$\frac{N}{16} \int d\theta\, \mathcal{P}(B_L + B_R)$ $+\frac{1}{2} \int d\theta\, B_L \cup B_R$

Table 4.5: Summary of the center one-form symmetries and the spins of the generating lines in various $3d$ Chern-Simons theories. (See the discussion in [1].) Here the gauge group is G and the level is $k = 1$, i.e. this is the TQFT G_1. When the center is $\mathbb{Z}_\ell$, the symmetry lines are $\{1, a, \cdots, a^{\ell-1}\}$ generated by the generating line a. When the center is $\mathbb{Z}_2 \times \mathbb{Z}_2$, the symmetry lines are $\{1, a, b, ab\}$. The spins of these lines are denoted by h_a, h_b and h_{ab}. In the case $Spin(N)$ with $N = 2 \bmod 4$ we also included the spin of the line a^2, which is used in the text. Note that in the context of $3d$ TQFT only the spin h modulo one is meaningful. The values in the table are those of the conformal dimensions of the corresponding Kac-Moody representation.

anomaly is the spin of the line a, h_a mod 1. Indeed, for $SU(N)$, $Sp(N)$, E_6, E_7, $Spin(2n+1)$, $Spin(4n+2)$, where the one-form symmetry is $\mathbb{Z}_\ell$ for some ℓ, the anomaly is (4.101), (4.114), (4.115), (4.116), (4.117), (4.118) respectively, in agreement with the entries in Table 4.5. In the case of $Spin(4n)$, the global symmetry is $\mathbb{Z}_2 \times \mathbb{Z}_2$ and it is generated by two lines a and b. In this case we have more kinds of anomalies. If either B_L or B_R vanishes, we can match the coefficient of $\mathcal{P}(B_R)$ and of $\mathcal{P}(B_L)$ in (4.122) with the spins of the lines. The coefficient of the mixed term can be checked by comparing the spin of the line ab with the anomaly for $B_L = B_R$.

We can also focus on the $\mathbb{Z}_2$ subgroup of the one form symmetry for $Spin(N)$ for even N that we discussed above. Its generating line, a^2 for $N = 4n+2$ or ab for $N = 4n$, has spin $\frac{1}{2}$ mod 1, which is the same as the $\mathbb{Z}_2$ generating line for $Spin(N)$ with odd N. This is consistent with the fact that the anomaly for this symmetry (4.117), (4.119), (4.123) is the same for all N.

4.6 $4d$ QCD

In this section we consider $4d$ QCD with fermions. Specifically, we will study $SU(N)$ and $Sp(N)$ with matter in the fundamental representation. This means that these theories do not have any one-form global symmetry.

Despite the absence of a one-form symmetry, these systems can still have a mixed anomaly between the θ-periodicity and its global symmetry. The reason is that even without a one-form global symmetry, twisted bundles of the dynamical gauge fields can be present with appropriate background of the gauge fields of the zero-form global symmetry.[28] These bundles do not have integer instanton numbers and hence they lead to our anomaly.

As we will see, even when all possible bundles of the dynamical fields can be present, the anomaly is not the same as in the corresponding gauge theory without matter in section 4.5. Some of that putative anomaly can be removed by adding appropriate counterterms.

[28]Many people have studied twisted bundles of the dynamical fields using a twist in the flavor to compensate it. For an early paper, see e.g. [185]. For more recent related discussions in $4d$ see [186, 187, 28, 188–192] and in $3d$ see [58, 151] and references therein.

This discussion extends the results about interfaces in $4d$ in [28] and explains the relation between them and the earlier results about anomalies in $3d$ Chern-Simons-matter theory in [151].

To briefly summarize our results, we will find that the $SU(N)$ theory with N_f fundamental quarks has a non-trivial anomaly (4.140) if and only if $L = \gcd(N, N_f) > 1$. Meanwhile the $Sp(N)$ theory with N_f fundamental quarks has a non-trivial anomaly (4.161) if and only if N is odd and N_f is even. We interpret these results in terms of the dynamics of the Chern-Simons matter theories that reside on their interfaces.

4.6.1 $SU(N)$ QCD

We begin with $4d$ $SU(N)$ QCD with N_f fermions in the fundamental representation. The Euclidean action is

$$\mathcal{S} = \int -\frac{1}{4g^2}\mathrm{Tr}(f \wedge *f) - \frac{i\theta}{8\pi^2}\mathrm{Tr}(f \wedge f) + i\overline{\psi}_I \not{D}_a \psi^I + i\overline{\widetilde{\psi}}_I \not{D}_a \widetilde{\psi}^I + (m\widetilde{\psi}_I \psi^I + c.c.)\,, \tag{4.124}$$

where f is the field strength of the $SU(N)$ gauge field a. Here we suppressed the color indices and used the standard summation convention for the flavor indices I. The theory only depends on the complex parameter $me^{i\theta/N_f}$, so without lost of generality we will take m to be a positive real parameter. Since the theory contains fermions, we will limit ourselves to spin manifolds, even though with an appropriate twist the theory can be placed on certain non-spin manifolds.

With equal masses the global symmetry of the system that acts faithfully is

$$\mathcal{G} = \frac{U(N_f)}{\mathbb{Z}_N}\ . \tag{4.125}$$

To see that, note that locally the fermions transform under

$$\mathcal{G}'_{micro} = SU(N) \times SU(N_f) \times U(1)\ . \tag{4.126}$$

where the first factor is the gauge group, the second factor is the flavor group, and the $U(1)$ is the baryon number normalized to have charge one for the fundamental quarks. However, $\mathcal{G}'_{micro}$ does not act faithfully on the quarks. The group that acts faithfully on them is

$$\mathcal{G}_{micro} = \frac{\mathcal{G}'_{micro}}{\mathbb{Z}_N \times \mathbb{Z}_{N_f}} = \frac{SU(N) \times U(N_f)}{\mathbb{Z}_N} \ . \tag{4.127}$$

Here $\mathcal{G}'_{micro}$ is represented by $(u \in SU(N), v \in SU(N_f), w \in U(1))$ and the quotient is the identification

$$(u, v, w) \sim (e^{2\pi i/N} u, v, e^{-2\pi i/N} w) \sim (u, e^{2\pi i/N_f} v, e^{-2\pi i/N_f} w) \, . \tag{4.128}$$

Finally, the global symmetry $\mathcal{G} = U(N_f)/\mathbb{Z}_N$ (4.125) is obtained by moding out $\mathcal{G}_{micro}$ by the $SU(N)$ gauge group.

Anomalies Involving θ-periodicity

In order to study the anomaly, we should couple the global symmetry $\mathcal{G} = U(N_f)/\mathbb{Z}_N$ (4.125) to background gauge fields. We will do it in steps. First, we couple the theory to $SU(N_f) \times U(1)$ background gauge fields (A, C) (the fundamental fermions have charge one under the $U(1)$). Together with the dynamical $SU(N)$ gauge fields a these gauge fields represent $\mathcal{G}'_{micro}$ (4.126).

Next, we would like to perform the quotient leading to $\mathcal{G}_{micro}$ (4.127). We do that by letting a be a $PSU(N)$ gauge field, A be a $PSU(N_f)$ gauge field, and $\widetilde{C} = KC$ with $K = \text{lcm}(N, N_f)$ be a $\widetilde{U}(1) = U(1)/\mathbb{Z}_K$ gauge field. Then, the gauge fields $(a, A, \widetilde{C})$ are correlated through

$$\oint \frac{\widetilde{F}}{2\pi} = \oint \left(\frac{N_f}{L} w_2(a) + \frac{N}{L} w_2(A) \right) \text{ mod } K \ . \tag{4.129}$$

where $\widetilde{F} = d\widetilde{C}$,

$$K = \text{lcm}(N, N_f) \, , \qquad L = \text{gcd}(N, N_f) = \frac{N N_f}{K} \tag{4.130}$$

and w_2 is the second Stiefel-Whitney class of the corresponding bundles.

In terms of these gauge fields, the background fields for $\mathcal{G} = U(N_f)/\mathbb{Z}_N$ are A and $\widetilde{C}$ in

$PSU(N_f) \times \widetilde{U}(1)$ constrained to satisfy

$$\frac{L}{N_f} \oint \frac{\widetilde{F}}{2\pi} = \frac{N}{N_f} \oint w_2(A) \bmod 1 \ . \tag{4.131}$$

Arbitrary values of these gauge fields, subject to (4.131), allow us to probe arbitrary values of $w_2(a)$ for the dynamical gauge fields. It is determined by a class $w_2^{(N)} \in H^2(X, \mathbb{Z}_N)$ of the $\mathcal{G} = U(N_f)/\mathbb{Z}_N$ gauge fields A and $\widetilde{C}$, which represents the obstruction to it being a $U(N_f)$ gauge field. Specifically,

$$\oint w_2(a) = \oint w_2^{(N)} = \left(\frac{L}{N_f} \oint \frac{\widetilde{F}}{2\pi} - \frac{N}{N_f} \oint w_2(A) \right) \bmod N \ . \tag{4.132}$$

Note that $w_2^{(N)}$ depends only on the background fields.

Now that we can use the background fields to induce arbitrary $w_2(a)$ we can repeat the analysis in section 4.5 to find that under shifting $\theta \to \theta + 2\pi$ the action is shifted

$$-\frac{2\pi i}{N} \int \frac{\mathcal{P}(w_2(a))}{2} = -\frac{2\pi i}{N} \int \frac{\mathcal{P}(w_2^{(N)})}{2} \bmod 2\pi i \, , \tag{4.133}$$

where in the last expression we expressed it in terms of the background fields A and $\widetilde{C}$ as in (4.132), showing that it is an anomaly.

We might be tempted to interpret (4.133) as a $\mathbb{Z}_N$ anomaly. However, this is not the case.

To see that, we proceed as follows. Using (4.129) and (4.132) it is straightforward to check that

$$\begin{aligned} \exp\left(-\frac{2\pi i}{N} \int \frac{\mathcal{P}(w_2^{(N)})}{2} \right) = & \exp\left(\frac{2\pi i}{L} \int \left(R \frac{\mathcal{P}(w_2^{(N)})}{2} + J w_2^{(N)} \cup w_2(A) \right) \right) \\ & \exp\left(2\pi i \int \left(-\frac{J}{K} \frac{\widetilde{F} \wedge \widetilde{F}}{8\pi^2} + \frac{NJ}{L} \frac{\mathcal{P}(w_2(A))}{2N_f} \right) \right) \end{aligned} \tag{4.134}$$

where R and J are integers satisfying

$$JN_f - RN = L \ . \tag{4.135}$$

(Different solutions of this equation for (J,R) lead to the same value in (4.134).) The significance of the apparently unmotivated expression (4.134) will be clear soon.

Given that we have background fields A and C, we can add some counterterms to the action. Two special terms are

$$\frac{i\Theta_A}{8\pi^2}\int \mathrm{Tr}(F_A\wedge F_A) + \frac{i\Theta_C}{8\pi^2}\int F_C\wedge F_C \, . \tag{4.136}$$

The normalization here is such that for an $SU(N_f)\times U(1)$ background (A,C) the coefficients Θ_A and Θ_C are 2π-periodic.

The new crucial point is that when we study the anomaly in the shift of θ we can combine this operation with continuous shifts of Θ_A and Θ_C. In other words, we can think of Θ_A and Θ_C as being θ-dependent[29]

$$\Theta_A = \Theta_A^{(0)} + n_A\theta\, , \qquad \Theta_C = \Theta_C^{(0)} + n_C\theta \ . \tag{4.137}$$

In order to preserve the 2π-periodicity of θ for $SU(N_f)\times U(1)$ background fields we take $n_A, n_C\in\mathbb{Z}$. Then, under $\theta\to\theta+2\pi$ the expression (4.136) is shifted by (recall that $\widetilde{C} = KC$)

$$2\pi i\int\left(n_A\frac{\mathrm{Tr}(F_A\wedge F_A)}{8\pi^2} + n_C\frac{\widetilde{F}\wedge\widetilde{F}}{8\pi^2K^2}\right) = 2\pi i\int\left(-n_A\frac{\mathcal{P}(w_2(A))}{2N_f} + n_C\frac{\widetilde{F}\wedge\widetilde{F}}{8\pi^2K^2}\right) \bmod 2\pi i\, . \tag{4.138}$$

Comparing this with (4.134) we see that by choosing

$$n_A = \frac{N}{L}J\, , \qquad n_C = JK \tag{4.139}$$

---

[29]We could have added to (4.136) another linearly independent counterterm $\frac{2\pi i p}{N}\int\frac{\mathcal{P}(w_2^{(N)})}{2}$. However, since its coefficient p is quantized, it cannot depend on θ as here and therefore it cannot be used to remove the variation in (4.134). This counterterm will be important in section 4.6.1.

(note that N/L is an integer) we can cancel the second factor in (4.134). This leaves us with an anomaly only because of the first factor in (4.134). As in all the examples above, it can be written as a $5d$ anomaly action

$$\mathcal{A}(\theta, A, C) = \exp\left(\frac{2\pi i}{L}\int \frac{d\theta}{2\pi}\left(R\frac{\mathcal{P}(w_2^{(N)})}{2} + Jw_2^{(N)} \cup w_2(A)\right)\right) . \tag{4.140}$$

It is crucial that unlike the variation (4.133), which appears to be a $\mathbb{Z}_N$ anomaly, this expression is only a $\mathbb{Z}_L$ anomaly.

Finally, we show that using the freedom in (4.137), we cannot remove this $\mathbb{Z}_L$ anomaly. In other words, we show that there are no integer shifts of n_A and n_C in (4.138) that can make the partition function invariant under $\theta \to \theta + 2\pi r$ with $r \neq 0$ mod L. We try to satisfy

$$\begin{aligned}\frac{r}{N}\int \frac{\mathcal{P}(w_2^{(N)})}{2} &= \left(\frac{s}{8\pi^2}\int \mathrm{Tr}(F_A \wedge F_A) + \frac{t}{8\pi^2 K^2}\int \widetilde{F} \wedge \widetilde{F}\right) \bmod 1 \\ &= \left(-\frac{s}{N_f}\int \frac{\mathcal{P}(w_2(A))}{2} + \frac{t}{8\pi^2 K^2}\int \widetilde{F} \wedge \widetilde{F}\right) \bmod 1 \,.\end{aligned} \tag{4.141}$$

with integer s and t. Clearly, we must have $t \in K\mathbb{Z}$. Then, using (4.129) it becomes

$$\left(\frac{(t - sN_f)}{N_f^2}\int \frac{\mathcal{P}(w_2(A))}{2} + \frac{t}{N^2}\int \frac{\mathcal{P}(w_2^{(N)})}{2} + \frac{t}{NN_f}\int w_2(A) \cup w_2^{(N)}\right) \bmod 1 \,. \tag{4.142}$$

Comparing with (4.141), we find that the coefficients (s, t, r) should satisfy

$$t - sN_f \in N_f^2\mathbb{Z}\,, \quad t - rN \in N^2\mathbb{Z}\,, \quad t \in NN_f\mathbb{Z}\,. \tag{4.143}$$

These conditions can be satisfied only if $r = 0$ mod L. These manipulations are identical to the discussion in section 2.2 in [151]. The reason for this relation will be clear soon.

We conclude that our theory has the anomaly (4.140). As a result, the theory is invariant only under $\theta \to \theta + 2\pi L$ and the anomaly is absent when $L = \gcd(N, N_f) = 1$.

When $L = \gcd(N, N_f) \neq 1$, the anomaly prohibits the long distance theory to be trivially gapped everywhere between θ and $\theta + 2\pi$. For small enough N_f the theory is believed to be trivially gapped at generic θ and nonzero mass. Therefore, the anomaly implies at least one phase transition when θ varies by 2π. This is consistent with [28], where it was argued for different behavior depending on N_f and the value of the mass.

The anomaly also constrains smooth interfaces between regions with different θ. Suppose the two regions have θ and $\theta + 2\pi k$ for some integer k. The theory on the interface then has an ordinary 't Hooft anomaly of the zero-form global symmetry $U(N_f)/\mathbb{Z}_N$

$$\exp\left(\frac{2\pi i k}{L}\int\left(R\frac{\mathcal{P}(w_2^{(N)})}{2} + J w_2(A)\cup w_2^{(N)}\right)\right) . \tag{4.144}$$

It is trivial when $k = 0 \bmod L$.

Clearly, this anomaly does not uniquely determined the theory on the interface (see e.g. the related discussion in [1] and the comments below). One possible choice for the theory on the interface is the $3d$ Chern-Simons-matter theory[30]

$$SU(N)_{k-N_f/2} + N_f \text{ fermions} \tag{4.145}$$

or its dual theories

$$\begin{aligned} &U(k)_{-N} + N_f \text{ scalars} \qquad && k \geq 1 \\ &U(N_f - k)_N + N_f \text{ scalars} \qquad && k < N_f \end{aligned} \tag{4.146}$$

with a $U(N_f)$ invariant scalar potential. The fact that there are two dual scalar theories for $1 \leq k < N_f$ was important in [152]. See the discussion there for more details about the validity of these dualities. All these theories have a $U(N_f)/\mathbb{Z}_N$ global symmetry with the anomaly (4.144) [151]. In deriving this anomaly the freedom to add Chern-Simons counterterms of the background gauge fields was used. These Chern-Simons counterterms can be thought of as being induced by the continuous counterterms (4.136) in the $4d$ theory. This explains the relation

[30]The special case of $k = 1$ was discussed in detail in [28]. The generalization to larger k was explored in appendix A of that paper.

between the computation of the anomaly under shifts of θ above with the computation of the anomaly in the $3d$ theory in section 2.2 in [151].

Further information about the theory along the interface can be found by considering the limits of large and small fermion masses. For $1 \leq N_f < N_{CFT}$ (with N_{CFT} the lower boundary of the conformal window) the analysis of [28] showed that for $1 \leq k < N_f$ the theories (4.145)(4.146) indeed capture the phases of the interface theory. We will not repeat this discussion here.

Implications of Time-Reversal Symmetry

As we did in the previous examples, we would like to compare our discussion using the anomaly in shift of θ to what can be derived using ordinary anomalies of global symmetries involving time-reversal symmetry T (or equivalently a CP symmetry) at $\theta = 0, \pi$.

First we discuss the possible counterterms that we can add to the theory. They are parameterized by[31]

$$\frac{\pi i s}{8\pi^2}\int \mathrm{Tr}(F_A \wedge F_A) + \frac{\pi i t}{8\pi^2 K^2}\int \widetilde{F} \wedge \widetilde{F} + \frac{2\pi i p}{N}\int \frac{\mathcal{P}(w_2^{(N)})}{2}\,. \tag{4.147}$$

All the other counterterms can be expressed as linear combinations of these three counterterms using (4.132). As in section 4.6.1, these counterterms have a redundancy which can be removed if we limit ourselves to p mod L.

Now we discuss the T-symmetry at $\theta = \pi$. In order to preserve the T-symmetry in $SU(N_f) \times U(1)$ backgrounds (as opposed to more general $U(N_f)/\mathbb{Z}_N$ backgrounds), s and t have to be integers. Under the T-symmetry, the partition function transforms by

$$Z[\theta, A, \tilde{C}] \to Z[\theta, A, \tilde{C}] \exp\left(2\pi i \int \left((1-2p)\frac{\mathcal{P}(w_2^{(N)})}{2N} - s\frac{\mathrm{Tr}(F_A \wedge F_A)}{8\pi^2} - t\frac{\widetilde{F} \wedge \widetilde{F}}{8\pi^2 K^2}\right)\right)\,. \tag{4.148}$$

Using the results in section 4.6.1, the transformations can be made non-anomalous with an appropriate choice of s and t if

$$1 - 2p = 0 \bmod L. \tag{4.149}$$

[31]The discrete counterterm $\frac{2\pi i p}{N}\int \frac{\mathcal{P}(w_2^{(N)})}{2}$ was not included in [28]. Its significance will be clear below.

theory symmetry $\mathcal{G}$	without T: θ-$\mathcal{G}$ anomaly	with T at $\theta = 0, \pi$: T-$\mathcal{G}$ anomaly at $\theta = \pi$	with T at $\theta = 0, \pi$: no continuous counterterms
$SU(N)$ QCD $\mathcal{G} = U(N_f)/\mathbb{Z}_N$	even L ✓	even L ✓	even L ✓
	odd $L \neq 1$ ✓	odd $L \neq 1$ ✗	odd $L \neq 1$ ✓
	$L = 1$ ✗	$L = 1$ ✗	$L = 1$ ✗

Table 4.6: Summary of anomalies and existence of continuous counterterms preserving time-reversal symmetry T in $4d$ QCD. Here $L = \gcd(N, N_f)$.

This equation has integer solutions for p if L is odd. Therefore, we conclude that the theory at $\theta = \pi$ has a mixed anomaly involving the time-reversal symmetry and the $U(N_f)/\mathbb{Z}_N$ zero-form symmetry only when $L = \gcd(N, N_f)$ is even. In that case, the theory at $\theta = \pi$ cannot be trivially gapped.

If $L = \gcd(N, N_f)$ is odd, the counterterms that preserve the T-symmetry at $\theta = 0$ and $\theta = \pi$ are different. In particular, we need to have $p = 0 \bmod L$ at $\theta = 0$ and $p = (L+1)/2 \bmod L$ at $\theta = \pi$. As with our various examples above, even though there is no anomaly for odd L, the fact that we need different counterterms at $\theta = 0$ and at $\theta = \pi$ can allow us to conclude that in that case the theory cannot be trivially gapped between $\theta = 0$ and $\theta = \pi$. There is an exception when $L = 1$. There we can choose $p = 0 \bmod L$ and find a continuous conterterm that preserves the T-symmetry at $\theta = 0, \pi$

$$i\theta \int \left(\frac{J}{NN_f} \frac{\widetilde{F} \wedge \widetilde{F}}{8\pi^2} + NJ \frac{\mathrm{Tr}(F_A \wedge F_A)}{8\pi^2} \right) \tag{4.150}$$

with an integer J satisfying $JN_f = 1 \bmod N$.

The existence of continuous counterterms preserving the time-reversal symmetry T at $\theta = 0, \pi$ are summarized in Table 4.6.

4.6.2 $Sp(N)$ QCD

Consider $Sp(N)$ QCD with $2N_f$ Weyl fermions in the fundamental $2N$-dimensional representation. Note that the theory is inconsistent with odd number of fermion multiplets due to a nonperturbative anomaly involving $\pi_4(Sp(N)) = \mathbb{Z}_2$ [193]. The Euclidean action includes the kinetic terms and

$$\mathcal{S} \supset -\frac{i\theta}{8\pi^2}\int \mathrm{Tr}(f\wedge f) + \int \left(m\Omega_{IJ}\widetilde{\Omega}^{ij}\psi^I_j\psi^J_j + c.c.\right)\,, \tag{4.151}$$

where Ω_{IJ} and $\widetilde{\Omega}^{ij}$ are the invariant tensors of $Sp(N_f)$ and $Sp(N)$ respectively and we used the standard summation convention for the flavor indices I, J and the color indices i, j. Note that we took equal masses m for all the quarks. Because of the chiral anomaly, the theory depends only on the complex parameter $me^{i\theta/2N_f}$, so without lost of generality we will take m to be a positive real parameter. For simplicity, we will limit ourselves to spin manifolds.

With the $Sp(N_f)$ invariant mass term the faithful global symmetry of the system is

$$\mathcal{G} = \frac{Sp(N_f)}{\mathbb{Z}_2}\,. \tag{4.152}$$

To see that, note that locally the fermions transform under

$$\mathcal{G}'_{micro} = Sp(N)\times Sp(N_f)\,, \tag{4.153}$$

where the first factor is the gauge group and the second factor is the flavor group. However the group that acts faithfully on the quarks is

$$G_{micro} = \frac{Sp(N)\times Sp(N_f)}{\mathbb{Z}_2}\,. \tag{4.154}$$

Here G'_{micro} is represented by $(u \in Sp(N), v \in Sp(N_f))$ and the quotient is the identification

$$(u, v) \sim (-u, -v)\,. \tag{4.155}$$

Finally the global symmetry $\mathcal{G} = Sp(N_f)/\mathbb{Z}_2$ (4.152) is obtained by moding out $\mathcal{G}_{micro}$ by the $Sp(N)$ gauge group.

Anomalies Involving θ-periodicity

In order to study the anomaly, we couple the global symmetry $\mathcal{G} = Sp(N_f)/\mathbb{Z}_2$ (4.152) to background gauge field. We do it in steps. First, we couple the theory to $Sp(N_f)$ gauge field A. Together with the dynamical gauge field a, these gauge fields represent $\mathcal{G}'_{micro}$ (4.153).

Next we perform the quotient leading to $\mathcal{G}_{micro}$ (4.154), This promotes a to be an $Sp(N)/\mathbb{Z}_2$ gauge field and A to be an $Sp(N_f)/\mathbb{Z}_2$ gauge field. They are correlated via

$$w_2(a) = w_2(A), \tag{4.156}$$

where w_2 is the second Stiefel-Whitney class of the corresponding bundle.

As in the case of $SU(N)$ gauge theories above, we can use the background fields to induce arbitrary $w_2(a)$. Then, using (4.114), we find that shifting $\theta \to \theta + 2\pi$, the action is shifted by

$$2\pi i \frac{N}{2} \int \frac{\mathcal{P}(w_2(a))}{2} = 2\pi i \frac{N}{2} \int \frac{\mathcal{P}(w_2(A))}{2} \bmod 2\pi i\,. \tag{4.157}$$

It is tempting to interpret this as a $\mathbb{Z}_2$ anomaly when N is odd and as no anomaly when N is even. However, we can add a continuous counterterm to the action

$$\frac{i\Theta}{8\pi^2} \int \mathrm{Tr}(F_A \wedge F_A)\,. \tag{4.158}$$

The normalization here is such that for $Sp(N_f)$ background A the coefficient Θ is 2π-periodic. We let Θ be θ-dependent

$$\Theta = \Theta^{(0)} + n\theta\,, \tag{4.159}$$

with integer n to preserve the 2π-periodicity of θ in $Sp(N_f)$ background. Then, under $\theta \to \theta + 2\pi$ the expression (4.158) is shifted by

$$2\pi i n \int \frac{\mathrm{Tr}(F_A \wedge F_A)}{8\pi^2} = 2\pi i \int \frac{nN_f}{2} \frac{\mathcal{P}(w_2(A))}{2} \bmod 2\pi i\,. \tag{4.160}$$

When N_f is odd, we can use these counterterms to cancel the shift of the action (4.157). The theory only has an anomaly when N is odd and N_f is even.

As in all the examples above, it can be written as a $5d$ action

$$\mathcal{A}(\theta, A) = \exp\left(\frac{2\pi i}{L}\int \frac{d\theta}{2\pi}\frac{\mathcal{P}(w_2(A))}{2}\right) \qquad \text{with } L = \gcd(N-1, N_f, 2)\ . \tag{4.161}$$

As a result, the theory is invariant under $\theta \to \theta + 4\pi$ when N is odd and N_f is even, and in all other cases it is invariant under $\theta \to \theta + 2\pi$.

Interfaces

The anomaly constrains smooth interfaces between regions with different θ. Suppose the two regions have θ and $\theta + 2\pi k$ for some integer k. The theory on the interface then has an ordinary 't Hooft anomaly of the zero-form global symmetry $Sp(N_f)/\mathbb{Z}_2$

$$\exp\left(2\pi i\frac{k}{L}\int \frac{\mathcal{P}(w_2)}{2}\right) \qquad \text{with} \qquad L = \gcd(N-1, N_f, 2)\,. \tag{4.162}$$

One possible choice for the theory on the interface is the $3d$ Chern-Simons-matter theory

$$Sp(N)_{k-N_f/2} + N_f \text{ fermions} \tag{4.163}$$

or its dual theory

$$\begin{aligned} &Sp(k)_{-N} + N_f \text{ scalars} && k \geq 1 \\ &Sp(N_f - k)_N + N_f \text{ scalars} && k < N_f \end{aligned} \tag{4.164}$$

with an $Sp(N_f)$ invariant scalar potential. (See [152] for more details on the validity of these dualities.) All these theories have an $Sp(N_f)/\mathbb{Z}_2$ global symmetry with the anomaly (4.162) [151].

Further information about the theory along the interface can be found by considering the limits of large and small fermion masses.

When the fermions are heavy, the $4d$ theory is effectively an $Sp(N)$ pure gauge theory and there expects to be an $Sp(N)_k$ Chern-Simons theory on the interface, or another TQFT with the same anomaly [1].

When the fermions are massless, for $1 \leq N_f < N_{CFT}$ (with N_{CFT} the lower boundary of the conformal window), the low-energy theory of the $4d$ theory is a sigma model based on $SU(2N_f)/Sp(N_f)$ [194]. The target space can be parametrized in two different ways:

$$SU(2N_f)/Sp(N_f) = \left\{ \Sigma = g\Omega g^T \,\middle|\, g \in SU(2N_f) \right\} \tag{4.165}$$

with Ω the $Sp(N_f)$-invariant tensor, or

$$SU(2N_f)/Sp(N_f) = \left\{ \Sigma \in SU(2N_f) \,\middle|\, \Sigma = -\Sigma^T \;\text{ and }\; \mathrm{Pf}(\Sigma) = 1 \right\} \tag{4.166}$$

with $\mathrm{Pf}(\Sigma)$ the Pfaffian of the anti-symmetric matrix Σ. The kinetic term is the obvious $SU(2N_f)$ invariant one. Adding a small $Sp(N_f)$-preserving mass term for the fermions in (4.151) corresponds to adding a potential to the chiral Lagrangian. The potential is proportional to

$$-m\left(e^{i\theta/2N_f}\mathrm{Tr}(\Sigma\Omega) + c.c.\right) . \tag{4.167}$$

It has a minimum at $\Sigma = e^{-2\pi i k/2N_f}\Omega$ when $\theta = 2\pi k$.

We are interested in the interfaces that interpolate between the vacuum at $\theta = 0$ and $\theta = 2\pi k$. For simplicity we restrict to the interfaces with $1 \leq k < N_f$. Following the similar analysis in [28], the interface configuration, up to symmetry transformations, is

$$\Sigma = \mathrm{diag}\left(\begin{pmatrix} 0 & e^{i\alpha_1} \\ -e^{i\alpha_1} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & e^{i\alpha_{N_f}} \\ -e^{i\alpha_{N_f}} & 0 \end{pmatrix}\right) . \tag{4.168}$$

The phases are divided into two groups $\alpha_1 = \cdots = \alpha_k$ and $\alpha_{k+1} = \cdots = \alpha_{N_f}$ that satisfy the constraint $\mathrm{Pf}(\Sigma) = \exp(i(\alpha_1 + \alpha_2 \cdots + \alpha_{N_f})) = 1$. The first group varies continuously from 0 to $2\pi(N_f - k)/N_f$ and the second group varies continuously from 0 to $-2\pi k/N_f$. The other configurations of the interface can be obtained by an $Sp(N_f)$ transformation $\Sigma \to g\Sigma g^T$. This shows that the theory along the interface is a sigma model based on the quaternionic Grassmannian

$$\mathrm{Gr}(k, N_f, \mathbb{H}) = \frac{Sp(N_f)}{Sp(k) \times Sp(N_f - k)}\,. \tag{4.169}$$

We conclude that for $1 \leq N_f < N_{CFT}$, the interfaces that interpolate between the vacuum at $\theta = 0$ and $\theta = 2\pi k$ with $1 \leq k < N_f$ has at least two phases. One is described by an $Sp(N)_k$ Chern-Simons theory and the other one is described by a nonlinear sigma model based on the quaternionic Grassmannian $\mathrm{Gr}(k, N_f, \mathbb{H})$. These two phases are captured by the theory (4.163) and its dual theory (4.164) [152].

4.7 Appendix A: Axions and Higher Group Symmetry

Throughout our analysis, we have discussed the usual presentation of anomalies via inflow. There is however another presentation of the same results by including additional higher-form gauge fields with atypical gauge transformation properties.

To carry this out for ordinary anomalies we proceed following [8]. We couple an anomalous d-dimensional field theory to a new d-form background field $A^{(d)}$ with a coupling $i\int_X A^{(d)}$. $A^{(d)}$ can be thought of as a background gauge field for a "$d-1$-form symmetry" that does not act on any dynamical field.[32] The anomaly of the d-dimensional theory is then formally removed by postulating that under gauge transformations of the background fields the new field transforms as $A^{(d)} \to A^{(d)} + d\lambda^{(d-1)} - 2\pi\alpha(\lambda, A)$ with $\alpha(\lambda, A)$ as in (4.1). The term with $\lambda^{(d-1)}$ is the standard gauge transformation of such a gauge field and the term with α, which cancels (4.1), reflects a higher-group symmetry (see e.g. [8–10] and references therein).

We can apply a similar technique to our generalized anomalies involving coupling constants. Focusing on the case of the θ-angle in $4d$ gauge theory, we couple our system to a classical

[32] Such couplings are common in the study of branes in string theory.

background three-form gauge field $A^{(3)}$ through[33]

$$\frac{i}{2\pi}\theta\left(dA^{(3)}+\frac{\pi(1-N)}{N}\mathcal{P}(B)\right)\ . \tag{4.170}$$

Now, the lack of invariance of the original system under $\theta\to\theta+2\pi$ is cancelled by this term. However, this term seems ill-defined. As in the general discussion above, this can be fixed by postulating that $A^{(3)}$ is not an ordinary three-form background field, but it transforms under the gauge transformation of B, such that the combination $F^{(4)}=dA^{(3)}+\frac{\pi(1-N)}{N}\mathcal{P}(B)$ is gauge invariant.[34] This means that the mixed anomaly between the periodicity of θ and the one-form $\mathbb{Z}_N$ global symmetry is cancelled at the cost of making the background field B participate together with $A^{(3)}$ in a higher group structure [8–10].

Note that the quantum field theory does not have a conserved current that couples to this new background gauge field $A^{(3)}$. In fact, this classical background field does not couple directly to any dynamical field. Yet, such a coupling allows us to cancel the anomaly.

The use of the background three-form gauge field $A^{(3)}$ above might seem contrived. However, when θ is a dynamical field (an axion) the treatment of the anomaly involving $A^{(3)}$ is required so that there are no bulk $5d$ terms involving dynamical fields. Moreover, in this case $A^{(3)}$ is also natural from another perspective as it couples to a conserved current for a two-form global symmetry $\frac{1}{2\pi}d\theta$ [140]. Following our rule of coupling all global symmetries to background gauge fields, in this case we must introduce $A^{(3)}$.

[33] The following discussion is similar to Appendix B of [140]. Below we explain the relation between them.

[34] A lowbrow way to think about this is to express $B=\frac{N}{2\pi}\tilde{B}$ in terms of a $U(1)$ two-form field $\tilde{B}$. Then, $\tilde{B}$ transforms under a one-form gauge transformation $\tilde{B}\to\tilde{B}+d\Lambda^{(1)}$ and $F^{(4)}=dA^{(3)}+\frac{\pi(1-N)}{N}\mathcal{P}(B)=dA^{(3)}+\frac{(1-N)N}{4\pi}\tilde{B}\wedge\tilde{B}$ is invariant provided $A^{(3)}$ transforms as

$$A^{(3)}\to A^{(3)}+d\Lambda^{(2)}+\frac{N(N-1)}{2\pi}\Lambda^{(1)}\wedge\tilde{B}+\frac{N(N-1)}{4\pi}\Lambda^{(1)}\wedge d\Lambda^{(1)}\ . \tag{4.171}$$

Bibliography

[1] P.-S. Hsin, H.T. Lam and N. Seiberg, *Comments on One-Form Global Symmetries and Their Gauging in 3d and 4d*, *SciPost Phys.* **6** (2019) 039 [1812.04716].

[2] P. Gorantla, H.T. Lam, N. Seiberg and S.-H. Shao, *A Modified Villain Formulation of Fractons and Other Exotic Theories*, 2103.01257.

[3] C. Córdova, D.S. Freed, H.T. Lam and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications I*, *SciPost Phys.* **8** (2020) 001 [1905.09315].

[4] C. Córdova, D.S. Freed, H.T. Lam and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications II*, *SciPost Phys.* **8** (2020) 002 [1905.13361].

[5] D. Gaiotto, A. Kapustin, N. Seiberg and B. Willett, *Generalized Global Symmetries*, *JHEP* **02** (2015) 172 [1412.5148].

[6] M. Barkeshli, P. Bonderson, M. Cheng and Z. Wang, *Symmetry, Defects, and Gauging of Topological Phases*, 1410.4540.

[7] A. Kapustin and R. Thorngren, *Higher symmetry and gapped phases of gauge theories*, 1309.4721.

[8] Y. Tachikawa, *On gauging finite subgroups*, 1712.09542.

[9] C. Córdova, T.T. Dumitrescu and K. Intriligator, *Exploring 2-Group Global Symmetries*, *JHEP* **02** (2019) 184 [1802.04790].

[10] F. Benini, C. Córdova and P.-S. Hsin, *On 2-Group Global Symmetries and their Anomalies*, *JHEP* **03** (2019) 118 [1803.09336].

[11] N. Seiberg and S.-H. Shao, *Exotic Symmetries, Duality, and Fractons in 2+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **10** (2021) 027 [2003.10466].

[12] D.S. Freed and M.J. Hopkins, *Reflection positivity and invertible topological phases*, arXiv:1604.06527.

[13] C.G. Callan, Jr. and J.A. Harvey, *Anomalies and Fermion Zero Modes on Strings and Domain Walls*, *Nucl. Phys.* **B250** (1985) 427.

[14] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135.

[15] J. Haah, *Local stabilizer codes in three dimensions without string logical operators*, *Phys. Rev. A* **83** (2011) 042330.

[16] S. Vijay, J. Haah and L. Fu, *A new kind of topological quantum order: A dimensional hierarchy of quasiparticles built from stationary excitations*, *Physical Review B* **92** (2015) .

[17] S. Vijay, J. Haah and L. Fu, *Fracton Topological Order, Generalized Lattice Gauge Theory and Duality*, *Phys. Rev.* **B94** (2016) 235157 [1603.04442].

[18] X. Chen, Z.-C. Gu, Z.-X. Liu and X.-G. Wen, *Symmetry protected topological orders and the group cohomology of their symmetry group*, *Physical Review B* **87** (2013) .

[19] X. Chen, Z.-C. Gu, Z.-X. Liu and X.-G. Wen, *Symmetry-protected topological orders in interacting bosonic systems*, *Science* **338** (2012) 1604 [https://science.sciencemag.org/content/338/6114/1604.full.pdf].

[20] X.-G. Wen, *Classifying gauge anomalies through symmetry-protected trivial orders and classifying gravitational anomalies through topological orders*, *Phys. Rev. D* **88** (2013) 045013.

[21] Y. You, T. Devakul, F.J. Burnell and S.L. Sondhi, *Subsystem symmetry protected topological order*, *Physical Review B* **98** (2018) .

[22] T. Devakul, D.J. Williamson and Y. You, *Classification of subsystem symmetry-protected topological phases*, *Physical Review B* **98** (2018) .

[23] T. Devakul, W. Shirley and J. Wang, *Strong planar subsystem symmetry-protected topological phases and their dual fracton orders*, *Physical Review Research* **2** (2020) .

[24] D.C. Tsui, H.L. Stormer and A.C. Gossard, *Two-dimensional magnetotransport in the extreme quantum limit*, *Phys. Rev. Lett.* **48** (1982) 1559.

[25] E. Witten, *Quantum Field Theory and the Jones Polynomial*, *Commun. Math. Phys.* **121** (1989) 351.

[26] R. Dijkgraaf and E. Witten, *Topological Gauge Theories and Group Cohomology*, *Commun. Math. Phys.* **129** (1990) 393.

[27] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, *Theta, Time Reversal, and Temperature*, *JHEP* **05** (2017) 091 [1703.00501].

[28] D. Gaiotto, Z. Komargodski and N. Seiberg, *Time-reversal breaking in QCD_4, walls, and dualities in 2 + 1 dimensions*, *JHEP* **01** (2018) 110 [1708.06806].

[29] A. Kapustin and N. Seiberg, *Coupling a QFT to a TQFT and Duality*, *JHEP* **04** (2014) 001 [1401.0740].

[30] B.S. Acharya and C. Vafa, *On domain walls of N=1 supersymmetric Yang-Mills in four-dimensions*, hep-th/0103011.

[31] P.-S. Hsin and N. Seiberg, *Level/rank Duality and Chern-Simons-Matter Theories*, *JHEP* **09** (2016) 095 [1607.07457].

[32] J.M. Maldacena, G.W. Moore and N. Seiberg, *D-brane charges in five-brane backgrounds*, *JHEP* **10** (2001) 005 [hep-th/0108152].

[33] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys. Rev.* **D83** (2011) 084019 [1011.5120].

[34] O. Aharony, N. Seiberg and Y. Tachikawa, *Reading between the lines of four-dimensional gauge theories*, *JHEP* **08** (2013) 115 [1305.0318].

[35] G.W. Moore and N. Seiberg, *Taming the Conformal Zoo*, *Phys. Lett.* **B220** (1989) 422.

[36] G.W. Moore and N. Seiberg, *Naturality in Conformal Field Theory*, *Nucl. Phys.* **B313** (1989) 16.

[37] F.A. Bais and J.K. Slingerland, *Condensate induced transitions between topologically ordered phases*, *Phys. Rev.* **B79** (2009) 045316 [0808.0627].

[38] M. Mueger, *Galois theory for braided tensor categories and the modular closure*, math/9812040.

[39] A. Bruguières, *Catégories prémodulaires, modularisations et invariants des variétés de dimension 3*, *Mathematische Annalen* **316** (2000) 215.

[40] M. Muger, *On the structure of modular categories*, *Proceedings of the London Mathematical Society* **87** (2003) 291 [arXiv:math/0201017].

[41] V. Drinfeld, S. Gelaki, D. Nikshych and V. Ostrik, *On braided fusion categories i*, *Selecta Mathematica* **16** (2010) 1.

[42] K. Walker and Z. Wang, *(3+1)-TQFTs and topological insulators*, *Frontiers of Physics* **7** (2012) 150 [1104.2632].

[43] F.J. Burnell, C.W. von Keyserlingk and S.H. Simon, *Phase transitions in three-dimensional topological lattice models with surface anyons*, *Phys. Rev.* **B88** (2013) 235120 [1303.0851].

[44] G.W. Moore and N. Seiberg, *Classical and Quantum Conformal Field Theory*, *Commun. Math. Phys.* **123** (1989) 177.

[45] B. Bakalov and A.A. Kirillov, *Lectures on Tensor Categories and Modular Functors*, vol. 21 of *University Lecture Series*, American Mathematical Soc. (2001).

[46] A. Kitaev, *Anyons in an exactly solved model and beyond*, *Annals of Physics* **321** (2006) 2 .

[47] P. Bonderson, K. Shtengel and J.K. Slingerland, *Interferometry of non-Abelian Anyons*, *Annals Phys.* **323** (2008) 2709 [0707.4206].

[48] P. Francesco, P. Di Francesco, P. Mathieu, D. Sénéchal and D. Senechal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics, Springer (1997).

[49] B.C. Berndt, R. Evans and K. Williams, *Gauss and Jacobi sums*, Canadian Mathematical Society series of monographs and advanced texts, Wiley (1998).

[50] C. Wall, *Quadratic forms on finite groups, and related topics*, *Topology* **2** (1963) 281 .

[51] C.T.C. Wall, *Quadratic Forms on Finite Groups II*, *Bulletin of the London Mathematical Society* **4** (1972) 156 [https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/blms/4.2.156].

[52] V.V. Nikulin, *Integral Symmetric Bilinear Forms and Some of Their Applications*, *Mathematics of the USSR-Izvestiya* **14** (1980) 103.

[53] D. Belov and G.W. Moore, *Classification of Abelian spin Chern-Simons theories*, hep-th/0505235.

[54] A. Kapustin and N. Saulina, *Topological boundary conditions in abelian Chern-Simons theory*, *Nucl. Phys.* **B845** (2011) 393 [1008.0654].

[55] Y. Lee and Y. Tachikawa, *A study of time reversal symmetry of abelian anyons*, *JHEP* **07** (2018) 090 [1805.02738].

[56] N. Seiberg and E. Witten, *Gapped Boundary Phases of Topological Insulators via Weak Coupling*, *PTEP* **2016** (2016) 12C101 [1602.04251].

[57] N. Seiberg, T. Senthil, C. Wang and E. Witten, *A Duality Web in 2+1 Dimensions and Condensed Matter Physics*, *Annals Phys.* **374** (2016) 395 [1606.01989].

[58] O. Aharony, F. Benini, P.-S. Hsin and N. Seiberg, *Chern-Simons-matter dualities with SO and USp gauge groups*, *JHEP* **02** (2017) 072 [1611.07874].

[59] C. Cordova, P.-S. Hsin and N. Seiberg, *Global Symmetries, Counterterms, and Duality in Chern-Simons Matter Theories with Orthogonal Gauge Groups*, *SciPost Phys.* **4** (2018) 021 [1711.10008].

[60] J.E. Tener and Z. Wang, *On classification of extremal non-holomorphic conformal field theories*, *J. Phys.* **A50** (2017) 115204 [1611.04071].

[61] A. Kapustin and R. Thorngren, *Topological Field Theory on a Lattice, Discrete Theta-Angles and Confinement*, *Adv. Theor. Math. Phys.* **18** (2014) 1233 [1308.2926].

[62] J. Milnor and J. Stasheff, *Characteristic Classes*, Annals of mathematics studies, Princeton University Press (1974).

[63] H. Lawson and M. Michelsohn, *Spin Geometry*, Princeton mathematical series, Princeton University Press (1989).

[64] M. Dierigl and A. Pritzel, *Topological Model for Domain Walls in (Super-)Yang-Mills Theories*, *Phys. Rev.* **D90** (2014) 105008 [1405.4291].

[65] E. Witten, *Dyons of Charge e theta/2 pi*, *Phys. Lett.* **B86** (1979) 283.

[66] R. Thorngren, *Framed Wilson Operators, Fermionic Strings, and Gravitational Anomaly in 4d*, *JHEP* **02** (2015) 152 [1404.4385].

[67] P.-S. Hsin and A. Turzillo, *Symmetry-enriched quantum spin liquids in (3 + 1)d*, *JHEP* **09** (2020) 022 [1904.11550].

[68] C.W. von Keyserlingk, F.J. Burnell and S.H. Simon, *Three-dimensional topological lattice models with surface anyons*, *Phys. Rev. B* **87** (2013) 045107.

[69] J. Wang, X.-G. Wen and E. Witten, *A New SU(2) Anomaly*, 1810.00844.

[70] C. Chamon, *Quantum glassiness in strongly correlated clean systems: An example of topological overprotection*, *Phys. Rev. Lett.* **94** (2005) 040402.

[71] R.M. Nandkishore and M. Hermele, *Fractons*, *Ann. Rev. Condensed Matter Phys.* **10** (2019) 295 [1803.11196].

[72] M. Pretko, X. Chen and Y. You, *Fracton Phases of Matter*, 2001.01722.

[73] K. Slagle and Y.B. Kim, *Quantum Field Theory of X-Cube Fracton Topological Order and Robust Degeneracy from Geometry*, *Phys. Rev.* **B96** (2017) 195139 [1708.04619].

[74] D. Bulmash and M. Barkeshli, *Generalized $U(1)$ Gauge Field Theories and Fractal Dynamics*, 1806.01855.

[75] W. Shirley, K. Slagle and X. Chen, *Fractional excitations in foliated fracton phases*, *Annals Phys.* **410** (2019) 167922 [1806.08625].

[76] M. Pretko, *The Fracton Gauge Principle*, *Phys. Rev.* **B98** (2018) 115134 [1807.11479].

[77] K. Slagle, D. Aasen and D. Williamson, *Foliated Field Theory and String-Membrane-Net Condensation Picture of Fracton Order*, *SciPost Phys.* **6** (2019) 043 [1812.01613].

[78] K. Slagle, *Foliated Quantum Field Theory of Fracton Order*, 2008.03852.

[79] N. Seiberg, *Field Theories With a Vector Global Symmetry*, *SciPost Phys.* **8** (2020) 050 [1909.10544].

[80] N. Seiberg and S.-H. Shao, *Exotic $U(1)$ Symmetries, Duality, and Fractons in 3+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **9** (2020) 046 [2004.00015].

[81] N. Seiberg and S.-H. Shao, *Exotic $\mathbb{Z}_N$ Symmetries, Duality, and Fractons in 3+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **10** (2021) 003 [2004.06115].

[82] P. Gorantla, H.T. Lam, N. Seiberg and S.-H. Shao, *More Exotic Field Theories in 3+1 Dimensions*, *SciPost Phys.* **9** (2020) 073 [2007.04904].

[83] P. Gorantla, H.T. Lam, N. Seiberg and S.-H. Shao, *FCC, Checkerboards, Fractons, and QFT*, 2010.16414.

[84] T. Rudelius, N. Seiberg and S.-H. Shao, *Fractons with Twisted Boundary Conditions and Their Symmetries*, 2012.11592.

[85] J. Villain, *Theory of one-dimensional and two-dimensional magnets with an easy magnetization plane. 2. The Planar, classical, two-dimensional magnet*, *J. Phys.(France)* **36** (1975) 581.

[86] D.J. Gross and I.R. Klebanov, *ONE-DIMENSIONAL STRING THEORY ON A CIRCLE*, *Nucl. Phys.* **B344** (1990) 475.

[87] A. Paramekanti, L. Balents and M.P.A. Fisher, *Ring exchange, the exciton bose liquid, and bosonization in two dimensions*, *Phys. Rev. B* **66** (2002) 054526.

[88] C. Xu and C. Wu, *Resonating plaquette phases in su(4) heisenberg antiferromagnet*, *Physical Review B* **77** (2008) .

[89] D.A. Johnston, M. Mueller and W. Janke, *Plaquette Ising models, degeneracy and scaling*, *European Physical Journal Special Topics* **226** (2017) [1612.00060].

[90] D. Bulmash and M. Barkeshli, *The Higgs Mechanism in Higher-Rank Symmetric $U(1)$ Gauge Theories*, *Phys. Rev.* **B97** (2018) 235112 [1802.10099].

[91] H. Ma, M. Hermele and X. Chen, *Fracton topological order from the Higgs and partial-confinement mechanisms of rank-two gauge theory*, *Phys. Rev.* **B98** (2018) 035111 [1802.10108].

[92] D. Radicevic, *Systematic Constructions of Fracton Theories*, 1910.06336.

[93] A. Polyakov, *Quark confinement and topology of gauge theories*, *Nuclear Physics B* **120** (1977) 429 .

[94] W. Shirley, K. Slagle and X. Chen, *Foliated fracton order in the checkerboard model*, *Phys. Rev. B* **99** (2019) 115123 [1806.08633].

[95] T. Tay and O.I. Motrunich, *Possible realization of the exciton bose liquid phase in a hard-core boson model with ring-only exchange interactions*, *Physical Review B* **83** (2011) .

[96] Y. You, Z. Bi and M. Pretko, *Emergent fractons and algebraic quantum liquid from plaquette melting transitions*, *Phys. Rev. Res.* **2** (2020) 013162 [1908.08540].

[97] Y. You, F.J. Burnell and T.L. Hughes, *Multipolar Topological Field Theories: Bridging Higher Order Topological Insulators and Fractons*, 1909.05868.

[98] A. Karch and A. Raz, *Reduced Conformal Symmetry*, 2009.12308.

[99] Y. You, J. Bibo, T.L. Hughes and F. Pollmann, *Fractonic critical point proximate to a higher-order topological insulator: How does UV blend with IR?*, 2101.01724.

[100] O. Dubinkin, A. Rasmussen and T.L. Hughes, *Higher-form Gauge Symmetries in Multipole Topological Phases*, *Annals Phys.* **422** (2020) 168297 [2007.05539].

[101] Y. You, T. Devakul, F.J. Burnell and S.L. Sondhi, *Symmetric Fracton Matter: Twisted and Enriched*, *Annals Phys.* **416** (2020) 168140 [1805.09800].

[102] A. Gromov, A. Lucas and R.M. Nandkishore, *Fracton hydrodynamics*, *Phys. Rev. Res.* **2** (2020) 033124 [2003.09429].

[103] M. Qi, L. Radzihovsky and M. Hermele, *Fracton phases via exotic higher-form symmetry-breaking*, *Annals Phys.* **424** (2021) 168360 [2010.02254].

[104] M. Bauer, G. Girardi, R. Stora and F. Thuillier, *A Class of topological actions*, *JHEP* **08** (2005) 027 [hep-th/0406221].

[105] D.S. Freed, G.W. Moore and G. Segal, *The Uncertainty of Fluxes*, *Commun. Math. Phys.* **271** (2007) 247 [hep-th/0605198].

[106] D.S. Freed, G.W. Moore and G. Segal, *Heisenberg Groups and Noncommutative Fluxes*, *Annals Phys.* **322** (2007) 236 [hep-th/0605200].

[107] J. Cheeger and J. Simons, *Differential characters and geometric invariants*, in *Geometry and topology (College Park, Md., 1983/84)*, vol. 1167 of *Lecture Notes in Math.*, (Berlin), pp. 50–80, Springer (1985), DOI.

[108] P. Deligne, *Théorie de Hodge. II*, *Inst. Hautes Études Sci. Publ. Math.* (1971) 5.

[109] P. Deligne and D.S. Freed, *Classical field theory*, in *Quantum Fields and Strings: a course for mathematicians, Vol. 1, 2 (Princeton, NJ, 1996/1997)*, (Providence, RI), pp. 137–225, Amer. Math. Soc. (1999).

[110] M.J. Hopkins and I.M. Singer, *Quadratic functions in geometry, topology, and M-theory*, *J. Diff. Geom.* **70** (2005) 329 [math/0211216].

[111] U. Bunke, *Differential cohomology*, arXiv:1208.3961.

[112] U. Schreiber, *Differential cohomology in a cohesive infinity-topos*, arXiv:1310.7930.

[113] U. Bunke, T. Nikolaus and M. Völkl, *Differential cohomology theories as sheaves of spectra*, *Journal of Homotopy and Related Structures* **11** (2014) 1 [arXiv:1311.3188].

[114] J.V. José, L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, *Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model*, *Phys. Rev. B* **16** (1977) 1217.

[115] R. Savit, *Vortices and the low-temperature structure of the $x - y$ model*, *Phys. Rev. B* **17** (1978) 1340.

[116] S. Sachdev and K. Park, *Ground states of quantum antiferromagnets in two dimensions*, *Annals of Physics* **298** (2002) 58.

[117] P.H. Ginsparg, *APPLIED CONFORMAL FIELD THEORY*, in *Les Houches Summer School in Theoretical Physics: Fields, Strings, Critical Phenomena Les Houches, France, June 28-August 5, 1988*, pp. 1–168, 1988 [hep-th/9108028].

[118] P. Di Francesco, P. Mathieu and D. Senechal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics, Springer-Verlag, New York (1997), 10.1007/978-1-4612-2256-9.

[119] S. Elitzur, R.B. Pearson and J. Shigemitsu, *Phase structure of discrete abelian spin and gauge systems*, *Phys. Rev. D* **19** (1979) 3698.

[120] L.P. Kadanoff, *Multicritical behavior at the kosterlitz-thouless critical point*, *Annals of Physics* **120** (1979) 39.

[121] J.L. Cardy, *General discrete planar models in two dimensions: Duality properties and phase diagrams*, *Journal of Physics A: Mathematical and General* **13** (1980) 1507.

[122] F.C. Alcaraz and R. Koberle, *Duality and the Phases of Z(n) Spin Systems*, *J. Phys.* **A13** (1980) L153.

[123] E. Fradkin and L.P. Kadanoff, *Disorder variables and para-fermions in two-dimensional statistical mechanics*, *Nuclear Physics B* **170** (1980) 1.

[124] V.A. Fateev and A.B. Zamolodchikov, *Parafermionic Currents in the Two-Dimensional Conformal Quantum Field Theory and Selfdual Critical Points in Z(n) Invariant Statistical Systems*, *Sov. Phys. JETP* **62** (1985) 215.

[125] R. Savit, *Duality in Field Theory and Statistical Systems*, *Rev. Mod. Phys.* **52** (1980) 453.

[126] T. Banks, R. Myerson and J. Kogut, *Phase transitions in abelian lattice gauge theories*, *Nuclear Physics B* **129** (1977) 493.

[127] B.E. Baaquie, *(2 + 1)-dimensional abelian lattice gauge theory*, *Phys. Rev. D* **16** (1977) 3040.

[128] R. Savit, *Topological excitations in u(1)-invariant theories*, *Phys. Rev. Lett.* **39** (1977) 55.

[129] A. Ukawa, P. Windey and A.H. Guth, *Dual Variables for Lattice Gauge Theories and the Phase Structure of Z(N) Systems*, *Phys. Rev. D* **21** (1980) 1013.

[130] E. Fradkin and S.H. Shenker, *Phase diagrams of lattice gauge theories with higgs fields*, *Phys. Rev. D* **19** (1979) 3682.

[131] T. Banks and E. Rabinovici, *Finite Temperature Behavior of the Lattice Abelian Higgs Model*, *Nucl. Phys. B* **160** (1979) 349.

[132] F.J. Wegner, *Duality in Generalized Ising Models and Phase Transitions Without Local Order Parameters*, *J. Math. Phys.* **12** (1971) 2259.

[133] D. Horn, M. Weinstein and S. Yankielowicz, *Hamiltonian approach to $z(n)$ lattice gauge theories*, *Phys. Rev. D* **19** (1979) 3715.

[134] E. Fradkin and L. Susskind, *Order and disorder in gauge systems and magnets*, *Phys. Rev. D* **17** (1978) 2637.

[135] A.Y. Kitaev, *Fault tolerant quantum computation by anyons*, *Annals Phys.* **303** (2003) 2 [quant-ph/9707021].

[136] L.D. Faddeev and S.L. Shatashvili, *Algebraic and Hamiltonian Methods in the Theory of Nonabelian Anomalies*, *Theor. Math. Phys.* **60** (1985) 770.

[137] R. Thorngren, *Field Theory, Family Style*, *Seminar at Simons Center for Geometry and Physics, October 2017* .

[138] R. Thorngren, *Topological Terms and Phases of Sigma Models*, 1710.02545.

[139] Y. Tachikawa and K. Yonekura, *Anomalies involving the space of couplings and the Zamolodchikov metric*, *JHEP* **12** (2017) 140 [1710.03934].

[140] N. Seiberg, Y. Tachikawa and K. Yonekura, *Anomalies of Duality Groups and Extended Conformal Manifolds*, *PTEP* **2018** (2018) 073B04 [1803.07366].

[141] A. Schwimmer and S. Theisen, *Osborn Equation and Irrelevant Operators*, 1902.04473.

[142] C. Lawrie, D. Martelli and S. Schafer-Nameki, *Theories of Class F and Anomalies*, *JHEP* **10** (2018) 090 [1806.06066].

[143] J. Gomis, Z. Komargodski, H. Ooguri, N. Seiberg and Y. Wang, *Shortening Anomalies in Supersymmetric Theories*, *JHEP* **01** (2017) 067 [1611.03101].

[144] A. Schwimmer and S. Theisen, *Moduli Anomalies and Local Terms in the Operator Product Expansion*, *JHEP* **07** (2018) 110 [1805.04202].

[145] M.M. Anber and E. Poppitz, *Domain walls in high-T SU(N) super Yang-Mills theory and QCD(adj)*, 1811.10642.

[146] V. Bashmakov, F. Benini, S. Benvenuti and M. Bertolini, *Living on the walls of super-QCD*, *SciPost Phys.* **6** (2019) 044 [1812.04645].

[147] R. Jackiw and P. Rossi, *Zero Modes of the Vortex - Fermion System*, *Nucl. Phys.* **B190** (1981) 681.

[148] E. Witten, *Superconducting Strings*, *Nucl. Phys.* **B249** (1985) 557.

[149] R. Donagi and D.R. Morrison, *Conformal field theories and compact curves in moduli spaces*, *JHEP* **05** (2018) 021 [1709.05355].

[150] J. Gomis, P.-S. Hsin, Z. Komargodski, A. Schwimmer, N. Seiberg and S. Theisen, *Anomalies, Conformal Manifolds, and Spheres*, *JHEP* **03** (2016) 022 [1509.08511].

[151] F. Benini, P.-S. Hsin and N. Seiberg, *Comments on global symmetries, anomalies, and duality in (2 + 1)d*, *JHEP* **04** (2017) 135 [1702.07035].

[152] Z. Komargodski and N. Seiberg, *A Symmetry Breaking Scenario for* QCD_3, 1706.08755.

[153] J. Gomis, Z. Komargodski and N. Seiberg, *Phases Of Adjoint* QCD_3 *And Dualities*, *SciPost Phys.* **5** (2018) 007 [1710.03258].

[154] F. Benini, *Three-dimensional dualities with bosons and fermions*, *JHEP* **02** (2018) 068 [1712.00020].

[155] C. Córdova, P.-S. Hsin and N. Seiberg, *Time-Reversal Symmetry, Anomalies, and Dualities in (2+1)d*, *SciPost Phys.* **5** (2018) 006 [1712.08639].

[156] C. Choi, D. Delmastro, J. Gomis and Z. Komargodski, *Dynamics of* QCD_3 *with Rank-Two Quarks And Duality*, 1810.07720.

[157] C. Córdova, P.-S. Hsin and K. Ohmori, *Exceptional Chern-Simons-Matter Dualities*, 1812.11705.

[158] Z. Komargodski, A. Sharon, R. Thorngren and X. Zhou, *Comments on Abelian Higgs Models and Persistent Order*, *SciPost Phys.* **6** (2019) 003 [1705.04786].

[159] Z. Komargodski, T. Sulejmanpasic and M. Ünsal, *Walls, anomalies, and deconfinement in quantum antiferromagnets*, *Phys. Rev.* **B97** (2018) 054418 [1706.05731].

[160] Y. Kikuchi and Y. Tanizaki, *Global inconsistency, 't Hooft anomaly, and level crossing in quantum mechanics*, *PTEP* **2017** (2017) 113B05 [1708.01962].

[161] Y. Tanizaki and T. Sulejmanpasic, *Anomaly and global inconsistency matching: θ-angles, $SU(3)/U(1)^2$ nonlinear sigma model, $SU(3)$ chains and its generalizations*, *Phys. Rev.* **B98** (2018) 115126 [1805.11423].

[162] A. Karasik and Z. Komargodski, *The Bi-Fundamental Gauge Theory in 3+1 Dimensions: The Vacuum Structure and a Cascade*, 1904.09551.

[163] M.F. Atiyah, V.K. Patodi and I.M. Singer, *Spectral asymmetry and Riemannian Geometry 1*, *Math. Proc. Cambridge Phil. Soc.* **77** (1975) 43.

[164] S.R. Coleman, *More About the Massive Schwinger Model*, *Annals Phys.* **101** (1976) 239.

[165] S. Elitzur, Y. Frishman, E. Rabinovici and A. Schwimmer, *Origins of Global Anomalies in Quantum Mechanics*, *Nucl. Phys.* **B273** (1986) 93.

[166] Y. Landau, *The Free Fermion Anomaly and Representations of the Pin Groups*, 1809.05078.

[167] R. Bott and L. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag (1982).

[168] S.S. Chern and J. Simons, *Characteristic forms and geometric invariants*, *Ann. of Math. (2)* **99** (1974) 48.

[169] L. Alvarez-Gaume, S. Della Pietra and G.W. Moore, *Anomalies and Odd Dimensions*, *Annals Phys.* **163** (1985) 288.

[170] J.J. van der Bij, R.D. Pisarski and S. Rao, *Topological Mass Term for Gravity Induced by Matter*, *Phys. Lett.* **B179** (1986) 87.

[171] C. Callias, *Index Theorems on Open Spaces*, *Commun. Math. Phys.* **62** (1978) 213.

[172] R. Bott and R. Seeley, *Some Remarks on the Paper of Callias*, *Commun. Math. Phys.* **62** (1978) 235.

[173] T. Pantev and E. Sharpe, *Notes on gauging noneffective group actions*, `hep-th/0502027`.

[174] T. Pantev and E. Sharpe, *GLSM's for Gerbes (and other toric stacks)*, *Adv. Theor. Math. Phys.* **10** (2006) 77 [`hep-th/0502053`].

[175] A. Caldararu, J. Distler, S. Hellerman, T. Pantev and E. Sharpe, *Non-birational twisted derived equivalences in abelian GLSMs*, *Commun. Math. Phys.* **294** (2010) 605 [`0709.3855`].

[176] N. Seiberg, *Modifying the Sum Over Topological Sectors and Constraints on Supergravity*, *JHEP* **07** (2010) 070 [`1005.0002`].

[177] M. Lajkó, K. Wamer, F. Mila and I. Affleck, *Generalization of the Haldane conjecture to SU(3) chains*, *Nucl. Phys.* **B924** (2017) 508 [`1706.06598`].

[178] K. Ohmori, N. Seiberg and S.-H. Shao, *Sigma Models on Flags*, *SciPost Phys.* **6** (2019) 017 [`1809.10604`].

[179] M. Hongo, T. Misumi and Y. Tanizaki, *Phase structure of the twisted $SU(3)/U(1)^2$ flag sigma model on $\mathbb{R} \times S^1$*, *JHEP* **02** (2019) 070 [`1812.02259`].

[180] I. Affleck and F.D.M. Haldane, *Critical Theory of Quantum Spin Chains*, *Phys. Rev.* **B36** (1987) 5291.

[181] I. Affleck, *Nonlinear sigma model at Theta = pi: Euclidean lattice formulation and solid-on-solid models*, *Phys. Rev. Lett.* **66** (1991) 2429.

[182] Z. Wan and J. Wang, *New Higher Anomalies, SU(N) Yang-Mills Gauge Theory and* $\mathbb{CP}^{\mathrm{N}-1}$ *Sigma Model*, 1812.11968.

[183] Z. Wan, J. Wang and Y. Zheng, *Quantum Yang-Mills 4d Theory and Time-Reversal Symmetric 5d Higher-Gauge TQFT: Anyonic-String/Brane Braiding Statistics to Topological Link Invariants*, 1904.00994.

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

[185] E. Cohen and C. Gomez, *A Computation of Tr(-1)**f in Supersymmetric Gauge Theories With Matter*, *Nucl. Phys.* **B223** (1983) 183.

[186] A. Cherman, S. Sen, M. Ünsal, M.L. Wagman and L.G. Yaffe, *Order parameters and color-flavor center symmetry in QCD*, *Phys. Rev. Lett.* **119** (2017) 222001 [1706.05385].

[187] H. Shimizu and K. Yonekura, *Anomaly constraints on deconfinement and chiral phase transition*, *Phys. Rev.* **D97** (2018) 105011 [1706.06104].

[188] Y. Tanizaki, T. Misumi and N. Sakai, *Circle compactification and 't Hooft anomaly*, *JHEP* **12** (2017) 056 [1710.08923].

[189] Y. Tanizaki, Y. Kikuchi, T. Misumi and N. Sakai, *Anomaly matching for the phase diagram of massless* $\mathbb{Z}_N$*-QCD*, *Phys. Rev.* **D97** (2018) 054012 [1711.10487].

[190] C. Córdova and T.T. Dumitrescu, *Candidate Phases for SU(2) Adjoint* QCD_4 *with Two Flavors from* $\mathcal{N} = 2$ *Supersymmetric Yang-Mills Theory*, 1806.09592.

[191] K. Yonekura, *Anomaly matching in QCD thermal phase transition*, 1901.08188.

[192] Y. Hidaka, Y. Hirono, M. Nitta, Y. Tanizaki and R. Yokokura, *Topological order in the color-flavor locked phase of* $(3+1)$*-dimensional* $U(N)$ *gauge-Higgs system*, 1903.06389.

[193] E. Witten, *An SU(2) Anomaly*, *Phys. Lett.* **B117** (1982) 324.

[194] E. Witten, *Current Algebra, Baryons, and Quark Confinement*, *Nucl. Phys.* **B223** (1983) 433.