Quantum Field Theory, Exotic Symmetries, and Fractons

Pranay Gorantla

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

Septermber 2023

Abstract

Quantum field theory has been extremely successful in organizing and classifying phases of matter and transitions between them. However, physicists have recently discovered a new class of phases, dubbed *fracton phases*, which do not admit a standard continuum quantum field theory description. One has to extend the framework of quantum field theory to incorporate these phases. In this thesis, I investigate various aspects of quantum field theories with exotic global symmetries, or exotic QFTs, that capture the low-energy physics of fractons and other exotic lattice models.

In chapter 2, I introduce a Euclidean spacetime lattice formulation for discretizing a continuum quantum field theory on a lattice while retaining its nice features, such as symmetries, anomalies, dualities, etc. It provides a rigorous framework for analyzing exotic QFTs by regularizing various singularities and infinities. This content is based on work with Ho Tat Lam, Nathan Seiberg, and Shu-Heng Shao [1].

In chapters 3 and 4, I use the above formulation to construct new gapless and gapped exotic lattice models, including fracton models. The novel feature of these models is that they can be defined on an arbitrary spatial graph, as opposed to a regular lattice. I also find intriguing connections between observables in these models and certain algebraic quantities associated with the graph. For instance, the ground state degeneracy of one of the models is the number of spanning trees of the graph, a fundamental measure of complexity of the graph. This content is based on works with Ho Tat Lam, Nathan Seiberg, and Shu-Heng Shao [2,3].

Acknowledgements

I am extremely grateful to my advisor, Nathan Seiberg, for his constant guidance and motivation throughout my graduate life. Without his support and encouragement, this work would not have been possible.

I would like to thank Igor Klebanov and Sanfeng Wu for agreeing to be on my thesis committee, and Herman for agreeing to read my thesis. I am also thankful to Catherine Brosowsky and Katherine Lamos at Jadwin, and Lisa Fleischer at IAS, for all their help with administrative matters.

I am glad to have worked with my amazing collaborators: Ho Tat Lam, Shu-Heng Shao, Fiona Burnell, Trithep Devakul, Ahana Chakraborty, Rajdeep Sensarma, Kunal Marwaha, and Santhoshini Velusamy. I am especially grateful to Ho Tat Lam, for being the perfect academic big brother, and Shu-Heng Shao, for all the calm and thoughtful explanations to all my confusions in physics.

I am indebted to the Physics Department at Princeton University and School of Natural Sciences at IAS for providing me with a unique atmosphere to grow as a professional physicist. I appreciate the countless valuable discussions on various topics in physics with my fellow students at Jadwin, especially Yiming Chen and Aaron Hillman. I have benefited greatly from the conversations with various postdocs at IAS, especially Abhinav Prem, Sahand Seifnashri, and Wilbur Shirley. I would also like to thank my batchmates Ben Spar, Mayer Feldman, Liz Helfenberger, Saumya Shivam, Sara Sussman, Kevin Crowley, Zijia Cheng, Thuy Vy Luu, Shuo Ma, Michael Onyszczak, and Zheyi Zhu for all the fun time at prelim lunches.

I would like to acknowledge my gratitude to all the professors and students at Tata Institute of Fundamental Research, Mumbai, in helping me with my first steps in theoretical physics. I am particularly thankful to Rajdeep Sensarma, Shiraz Minwalla, Gautam Mandal, and Kedar Damle for their guidance and inspiration. I would also like to thank Abhijit Gadde and Omkar Parrikar for hosting me over a stimulating month.

Finally, I thank my family for always encouraging me, and my friends whose support has provided me with an enjoyable environment to work in.

To my wife

and our daughter.

Contents

Abstract 3

Acknowledgements 4

1 Introduction **11**

1.1 Quantum field theory 11

1.2 Fractons 13

1.3 Exotic symmetries 14

1.4 Overview 15

1.4.1 Modified Villain formulation 16

1.4.2 Fractons on graphs: Laplacian models 18

2 A Modified Villain Formulation of Fractons and Other Exotic Theories with Subsystem Symmetries **21**

2.1 Introduction 21

2.1.1 Overview of continuum field theories for exotic models 22

2.1.2 Modified Villain lattice models 23

2.1.3 Outline 28

2.2 2+1d (3d Euclidean) exotic theories 30

2.2.1 ϕ-theory (XY-plaquette model) 31

2.2.2 A-theory ($U(1)$ tensor gauge theory) 40

2.2.3 $\mathbb{Z}_N$ tensor gauge theory 45

2.3 3+1d (4d Euclidean) exotic theories with cubic symmetry 51
2.3.1 ϕ-theory 52
2.3.2 A-theory 55
2.3.3 X-cube model 58
2.A Villain formulation of some classic quantum-mechanical systems . . . 65
2.A.1 Particle on a ring 65
2.A.2 Noncommutative torus 68
2.B Modified Villain formulation of 2d Euclidean lattice theories without gauge fields 72
2.B.1 2d Euclidean XY-model 72
2.B.2 2d Euclidean $\mathbb{Z}_N$ clock model 83
2.C Modified Villain formulation of p-form lattice gauge theory in diverse dimensions 87
2.C.1 $U(1)$ gauge theory 87
2.C.2 $\mathbb{Z}_N$ gauge theory 92

3 Fractons on Graphs and Complexity: $U(1)$ Laplacian Models 99

3.1 Introduction 99
3.2 Graph theory primer 105
3.2.1 Discrete harmonic functions and Smith decomposition of L . . 107
3.2.2 Theory of divisors and the Abel-Jacobi map on a graph 108
3.2.3 Jacobian group and complexity of a graph 110
3.3 Laplacian ϕ-theory on a graph 111
3.3.1 Self-duality 113
3.3.2 Global symmetry and the Jacobian group 113
3.3.3 GSD = Complexity 115
3.3.4 Spectrum 116
3.3.5 Robustness of GSD 118

3.3.6 Examples 120
3.4 $U(1)$ Laplacian gauge theory on a graph 124
3.4.1 Global symmetry 125
3.4.2 Spectrum 127
3.4.3 Time-like symmetry, mobility of defects, and fractons 128
3.4.4 Correspondence between divisors and fractons 132
3.4.5 Examples 133
3.A Discrete Poisson equation on a graph 134
3.B Polynomial representation of functions on square lattice 140
3.B.1 Naturalness of 2+1d Laplacian ϕ-theory 142
3.B.2 Mobility of defects in 2+1d $U(1)$ Laplacian gauge theory . . . 148

4 Fractons and Lineons on Graphs: $\mathbb{Z}_N$ Laplacian Models 154
4.1 Introduction 154
4.2 Graph theory primer 158
4.2.1 Discrete Laplacian operator Δ_L and its Smith decomposition . 160
4.2.2 Jacobian group of a graph 162
4.3 Anisotropic $\mathbb{Z}_N$ Laplacian model on a graph 162
4.3.1 Hamiltonian for the stabilizer code 163
4.3.2 Euclidean presentation 166
4.4 3+1d anisotropic $\mathbb{Z}_N$ Laplacian model on a torus 172
4.4.1 Upper bound on $\log_N$ GSD 172
4.4.2 $N = 2$ 173
4.4.3 $N = p$ prime larger than 2 175
4.A More on $\mathbb{Z}_N$-valued functions on a graph 177
4.B 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model 179
4.B.1 Ground state degeneracy 179
4.B.2 Mobility restrictions 183

4.C 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model 184
4.C.1 Ground state degeneracy 184
4.C.2 Mobility restrictions 201
4.D $\mathbb{Z}_N$ Laplacian model on a graph 204
4.D.1 Hamiltonian 205
4.D.2 Euclidean presentation 207
4.D.3 Examples 211

Bibliography **215**

Chapter 1

Introduction

1.1 Quantum field theory

Quantum field theory (QFT) is the universal language of much of modern physics, including particle physics, condensed matter physics, cosmology, gravity, etc., with applications outside physics as well, e.g., mathematics. One of the most successful applications of QFT is in describing the low-energy/long-distance properties of materials, thereby providing a framework for classification of phases of matter and transitions between them.

The underlying reason for this success is *separation of scales*, i.e., the macroscopic properties of a material are largely independent of its microscopic details. The microscopic description, or the UV theory, of a material could be a lattice model or another QFT. In any case, separation of scales ensures that the low-energy theory, or the IR theory, is captured by a scale invariant continuum (relativistic) QFT. Formally, the flow from the UV theory to the IR theory is governed by the renormalization group flow equation [4, 5].

Consider, for example, the classical (Euclidean) 2d XY model on a square lattice described by the action

$$S = -\beta \sum_{\mu\text{-links}} \cos(\Delta_\mu \phi) \ , \tag{1.1}$$

where β is a dimensionless coupling constant, $e^{i\phi}$ is a $U(1)$ variable on each site, and Δ_μ denotes the difference operator along a μ-link for $\mu = x, y$. This is an interacting theory with a $U(1)$ global symmetry that shifts ϕ by a circle-valued constant. The phase diagram of this theory is well understood [6,7]: for large β, the theory is gapless and flows to the 2d compact boson, described by the (Euclidean) Lagrangian

$$\mathcal{L} = \frac{R^2}{2\pi}(\partial_\mu \phi)^2 \ , \qquad \phi \sim \phi + 2\pi \ . \tag{1.2}$$

This is the $c = 1$ conformal field theory (CFT) at radius R.[1] (See any standard textbook on CFT, e.g., [8,9], for more details.) In contrast, for small β, the theory is trivially gapped. The transition from the gapless phase to the gapped phase is known as the Berezinskiǐ-Kosterlitz-Thouless (BKT) transition [10–12]. It is captured by the 2d compact boson as follows: the winding/vortex operators, which are irrelevant at $R \geq R_{BKT} \equiv \sqrt{2}$, become relevant at $R \leq R_{BKT}$, a phenomenon commonly referred to as proliferation of vortices.

As in the above example, it is generally believed that any gapless phase is captured by a CFT. On the other hand, any gapped phase is expected to be captured by a topological quantum field theory (TQFT). For example, the low-energy properties of the toric code [13] are described by the 2+1d topological $\mathbb{Z}_2$ gauge theory [14, 15]. Another interesting example is the fractional quantum Hall state, which is captured by a 2+1d Chern-Simons theory [16].

[1] The relation between R and β is complicated but for large β, we have $\beta \approx R^2/\pi$.

1.2 Fractons

This belief is challenged by a new class of phases known as fracton phases. (See [17,18] for reviews on this subject.) Microscopic models that flow to such phases include fracton models, such as Chamon's model [19], Haah's code [20], X-cube model [21], etc., and other exotic lattice models, such as XY plaquette model [22], etc. Some of the striking features of these models are:

- *Exotic global symmetries:* Examples include subsystem symmetries that act only on a part of the full system, such as a plane (e.g., X-cube model), or a fractal (e.g., Haah's code), and multipole symmetries, such as dipole symmetries (e.g., symmetric tensor gauge theories [23]).

- *Large ground state degeneracies:* On a periodic cubic lattice, some of these models (e.g., X-cube model) have large ground state degeneracies that grow exponentially in the linear size of the system. In some other models (e.g., Haah's code) the dependence of GSD on the linear size of the system is more complicated and often shows erratic behavior [24].

- *Restricted mobilities:* Some of these models (e.g., Chamon's model, Haah's code, X-cube model, etc.) have particle-like excitations with restricted mobilities. These include particles that are completely immobile (fractons), or can move only along a line (lineons), or a plane (planeons). In some models (e.g., X-cube model, symmetric tensor gauge theories), known as type-I models, composites of fractons can become partially or completely mobile, whereas in some other models (e.g., Haah's code), known as type-II models, any composite of fractons is immobile.

It is worth noting that both restricted mobilities and large ground state degeneracies can be interpreted as consequences of exotic global symmetries [25].

These features are fundamentally incompatible with standard continuum (relativistic) QFT. More precisely, fracton phases are characterized by a lack of separation of scales—a.k.a., *UV/IR mixing* [26], alluding to the dependence of macroscopic (IR) properties on microscopic (UV) details—and hence, do not admit a standard continuum (relativisitic) QFT description. This suggests that we need to extend the framework of QFT to accommodate them. Exotic global symmetries come in handy in this regard.

1.3 Exotic symmetries

As is often the case in physics, symmetries offer a way convenient starting point for a systematic construction and analysis of a QFT. They provide a powerful guiding principle in organizing and classifying phases, à la Landau [27]. A thorough understanding of symmetries can sometimes be sufficient to completely solve a problem, e.g., the hydrogen atom.

The modern view of ordinary global symmetries in QFT is that they are codimension-1 topological operators/defects, whose fusion is given by a group multiplication [28]. The standard example of an ordinary global symmetry is provided by a Noether current J_μ which satisfies the conservation equation $\partial^\mu J_\mu = 0$. One can reinterpret the current as a $(d-1)$-form J,[2] and then the conservation equation translates to J being closed, i.e., $dJ = 0$. Given such a form and a codimension-1 submanifold S, one can construct the charge operator $Q(S) = \oint_S J$. That this is a topological operator follows from the closedness of J. Finally, the exponential of the charge operator obeys a group multiplication law.

Recently, there have been several generalizations of ordinary global symmetries. (See [29] for a review on this subject.) In one direction, relaxing the codimension to arbitrary values gives higher-form symmetries [28]. In another direction, relaxing the

[2] Here, d is the dimension of the spacetime manifold M.

group-like fusion rule to a more general fusion rule gives non-invertible symmetries [30, 31]. Relaxing both gives a broader class of generalized global symmetries. However, in all these cases, the resulting operators are topological.

Exotic global symmetries further relax the requirement of the operators being topological. (See [32–34, 29] for reviews on exotic global symmetries in condensed matter physics.) A prototypical example of an exotic global symmetry is a subsystem symmetry. Consider the Noether currents (J_0, J_{xy}) in 2+1d which satisfy a nonstandard conservation equation [35]

$$\partial_0 J_0 = \partial_x \partial_y J_{xy} \ . \tag{1.3}$$

There are infinitely many charges, one for each line labelled by x or y, that are conserved in time:

$$Q_x(x) = \oint dy \ J_0 \ , \qquad Q_y(y) = \oint dx \ J_0 \ . \tag{1.4}$$

Each charge is associated with a symmetry that acts only on those fields along a line of fixed x or y, and hence the name (linear) subsystem symmetry. Except for the constraint

$$\oint dx \ Q_x(x) = \oint dy \ Q_y(y) \ , \tag{1.5}$$

they are independent of each other and cannot be deformed to each other. In other words, they are not topological.

1.4 Overview

In this thesis, I explore several aspects of exotic QFTs, i.e., QFTs with exotic global symmetries. I first give a novel Euclidean spacetime lattice formulation of exotic QFTs that makes their analysis rigorous while retaining their desirable features. I then consider the question of the minimal structure one needs to assume on the spacetime to realize the fracton-like physics. In particular, using the new lattice formulation, I construct exotic lattice models, including gapless and gapped fracton models, that can be defined on arbitrary spatial graphs. The main body of this thesis is divided into three chapters that are summarized below.

1.4.1 Modified Villain formulation

There are several challenges in constructing and analyzing exotic QFTs. They are typically nonrelativistic—in the example of subsystem symmetry in 2+1d in (1.3), the theory can have only a $\mathbb{Z}_4$ rotational symmetry because lines of fixed x or y are special [35]. While this may not be uncommon, there are other, more subtle issues with continuum exotic QFTs, such as discontinuous space of field configurations, infinite ground state degeneracies, states with infinite energies, etc. [35–38]. Moreover, these subtleties persist even after introducing an IR regulator, i.e., placing the theory on a compact space. One needs to introduce an UV regulator as well, i.e., place the theory on a lattice, which can be interpreted as another manifestation of UV/IR mixing in these theories.

The drawback of "discretizing" a continuum theory is that the nice features of the continuum theory are not always present on the lattice. Take the example of 2d classical XY model of (1.1) for instance. It is an interacting theory with a $U(1)$ shift symmetry without any anomaly or duality, and it flows to a gapless or a gapped phase depending on the value of β. On the other hand, the 2d compact boson of (1.2) is a free (quadratic) theory with a $U(1)$ shift symmetry and an additional $U(1)$ winding symmetry, which has a mixed 't Hooft anomaly with the former. In addition, the theory enjoys a self-duality, known as T-duality, that exchanges the two symmetries and maps the theory at radius R to the theory at radius $1/2R$. Finally, it is gapless for any R.

In Chapter 2, I introduce a new Euclidean spacetime lattice formulation of exotic QFTs with subsystem symmetries, including the X-cube model. The utility of this formulation is that it provides a rigorous framework for the "ambitious" analysis of exotic QFTs in [35–38]. The singularities of exotic QFTs, such as discontinuous field configurations, infinite degeneracies, and infinite energies, are all regularized without losing their nice features, such as emergent symmetries, anomalies, dualities, etc.

This construction is applicable to ordinary QFTs as well, as we demonstrate in several appendices at the end of the chapter. To highlight the key features of this formulation, consider again the 2d classical XY model described by the action (1.1). The standard Villain representation of this action is given by

$$S_{\text{Vill}} = \frac{\beta}{2} \sum_{\mu\text{-links}} (\Delta_\mu \phi - 2\pi n_\mu)^2 \ , \tag{1.6}$$

where n_μ are integer fields with an integer gauge symmetry. It is well known that the phase diagram of the cosine model and its Villain representation is the same: gapless for large β and gapped for small β. Consider the following modification of the Villain representation:

$$S_{\text{mod Vill}} = \frac{\beta}{2} \sum_{\mu\text{-links}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquettes}} \tilde{\phi}(\Delta_x n_y - \Delta_y n_x) \ , \tag{1.7}$$

where $\tilde{\phi}$ is a new compact scalar field which can be interpreted as a Lagrange multiplier imposing the constraint $\Delta_x n_y - \Delta_y n_x = 0$. Physically, the quantity $\Delta_x n_y - \Delta_y n_x$ measures the "vorticity" in a plaquette, and setting it to zero should be thought of as "suppressing the vortices."

As one might expect from the last sentence, the new lattice model (1.7) is always gapless at low energies. Moreover, the new lattice model has several nice features: (i) it is quadratic, and hence, free, (ii) in addition to the $U(1)$ shift symmetry, there is a $U(1)$ winding symmetry that shifts the field $\tilde{\phi}$, (iii) the two $U(1)$ symmetries have a mixed 't Hooft anomaly, and (iv) there is a self-duality that exchanges the two $U(1)$ symmetries along with the fields ϕ and $\tilde{\phi}$, and maps the coupling constant β to $1/(2\pi)^2\beta$. In other words, the new lattice model is closer to the 2d compact boson than the cosine model and its Villain representation.

The contents of this chapter are based on work with Ho Tat Lam, Nathan Seiberg, and Shu-Heng Shao [1]. They were presented at

1. *SCGP Workshop for New directions in topological phases: from fractons to spatial symmetries*, virtual, May 24-28, 2021,

2. *String Math 2021*, IMPA, Rio de Janeiro, June 14-18, 2021, and

3. *Simons Collaboration on Ultra-Quantum Matter Mini Meeting*, Harvard University and MIT, Cambridge, April 25-26, 2022.

1.4.2 Fractons on graphs: Laplacian models

QFT is an indispensable tool in obtaining (potentially, new) invariants of manifolds. Witten's work on computing the Jones polynomial of a knot in a 3d manifold using 3d Chern-Simons theory is a prime example of this [39]. A simpler example is that the GSD of an abelian Chern-Simons theory, described by a K-matrix, on a spacetime of the form $S^1 \times \Sigma_g$, where Σ_g is a compact spatial manifold of genus g, is $|\det K|^g$. It is not then surprising that one is interested in studying a QFT on more general manifolds than simply flat spacetimes.

Relatedly, ordinary lattice models can be placed on arbitrary discretizations of a manifold. For example, while the toric code is usually defined on a square lattice, it is easy to define it on an arbitrary 2d simplicial complex. More interestingly, the Ising model can be studied on an arbitrary graph, which consists of a bunch of sites/vertices connected by links/edges. This is arguably the most minimal structure that one can assume on the spatial lattice.

On the other hand, exotic QFTs do not extend in a natural way to arbitrary manifolds. When they do, one needs additional structures on the manifolds that are controlled by the exotic global symmetries. For example, a class of exotic theories with planar subsystem symmetries, known as foliated field theories, requires a foliation of the manifold [40–49]. Another class of theories with dipole symmetries, known as tensor gauge theories, can be placed on Einstein manifolds [50], or manifolds with some other structure [51, 52]. This leads us to a natural question: what is the minimal structure one needs to assume on the spacetime manifold/lattice to get fracton-like physics, and does this physics contain interesting information about the spatial manifold/lattice?

In chapter 3, I introduce two exotic lattice models on a general spatial graph Γ. The first one is a matter theory of a compact Lifshitz scalar field, while the second one is a certain rank-2 $U(1)$ gauge theory of fractons. Both lattice models are defined via the discrete Laplacian operator on the graph Γ. I unveil an intriguing correspondence between the physical observables of these lattice models and graph-theoretic quantities associated with Γ. For instance, the ground state degeneracy of the matter theory equals the number of spanning trees of Γ, which is a common measure of complexity in graph theory ("GSD = complexity"). The discrete global symmetry is identified as the Jacobian group of the graph Γ. In the gauge theory, superselection sectors of fractons are in one-to-one correspondence with the divisor classes in graph theory. In particular, under mild assumptions on the spatial graph, the fracton immobility is proven using a graph-theoretic Abel-Jacobi map.

The contents of this chapter are based on work with Ho Tat Lam and Shu-Heng Shao [2]. They were presented at *Simons Collaboration on Ultra-Quantum Matter Mini Meeting*, UT Austin, Austin, August 9-12, 2022.

In chapter 4, I introduce a $\mathbb{Z}_N$ stabilizer code that can be defined on any spatial lattice of the form $\Gamma \times C_{L_z}$, where Γ is a general graph and C_{L_z} is a cycle graph on

L_z vertices. I also present the low-energy limit of this stabilizer code as a Euclidean lattice action, which we refer to as the anisotropic $\mathbb{Z}_N$ Laplacian model. It is gapped, robust (i.e., stable under small deformations), and has lineons. Its ground state degeneracy is expressed in terms of a "mod N-reduction" of the Jacobian group of the graph Γ. In the special case when space is an $L \times L \times L_z$ cubic lattice, the logarithm of the GSD depends on L in an erratic way and grows no faster than $O(L)$, with interesting connections to a famous conjecture in number theory. I also discuss another gapped model, the $\mathbb{Z}_N$ Laplacian model, which can be defined on any graph. It has fractons and a similarly strange GSD.

The contents of this chapter are based on work with Ho Tat Lam, Nathan Seiberg, and Shu-Heng Shao [3]. (As we were finalizing this paper, [53] appeared on arXiv, which studies the same anisotropic $\mathbb{Z}_N$ Laplacian model using its stabilizer code.) They were presented at *Simons Collaboration on Ultra-Quantum Matter Annual Meeting*, Simons Foundation, New York City, January 19-20, 2023.

Chapter 2

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

2.1 Introduction

The surprising discoveries of [19,20] have stimulated exciting work on fracton models. This subject is reviewed nicely in [17,18], 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.

2.1.1 Overview of continuum field theories for exotic models

Following earlier work on such continuum field theories [54, 55, 42, 56, 45, 47], we initiated a systematic analysis of exotic field theories, including theories of fractons [57, 35–38, 58, 59]. Our 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 models. 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 [35], 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.

2.1.2 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 [60], 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. The modified Villain versions of the ordinary XY model and the $U(1)$ gauge theory have been previously constructed in [61].[1]

Let us demonstrate this in the standard 2d Euclidean XY-model. (See Appendix 2.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

$$\beta \sum_{\text{link}} [1 - \cos(\Delta_\mu \phi)] \ , \tag{2.1}$$

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

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 \ . \tag{2.2}$$

[1]We thank Z. Komargodski and T. Sulejmanpasic for pointing out this reference and related papers to us.

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

$$\phi \sim \phi + 2\pi k \ , \qquad n_\mu \sim n_\mu + \Delta_\mu k \ , \tag{2.3}$$

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,

$$\mathcal{N} \equiv \Delta_x n_y - \Delta_y n_x \ , \tag{2.4}$$

to zero [62]. 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

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquette}} \tilde{\phi} \mathcal{N} \ . \tag{2.5}$$

We refer to the action (2.5) as the modified Villain version of the original action (2.1). We will analyze it in detail in Appendix 2.B.1.

In the bulk of the chapter, we will apply this procedure to the lattice models of [22, 63, 21, 64–67, 35–37]. These include, in particular, the X-cube model [21]. 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 [57, 35–38, 58, 59] 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

$$\begin{aligned}&\sum_n \exp\left[-\frac{\beta}{2}(\theta-2\pi n)^2+in\tilde{\theta}\right]\\&\qquad=\frac{1}{\sqrt{2\pi\beta}}\sum_{\tilde{n}}\exp\left[-\frac{1}{2(2\pi)^2\beta}(\tilde{\theta}-2\pi\tilde{n})^2-\frac{i\theta}{2\pi}(2\pi\tilde{n}-\tilde{\theta})\right]\ .\end{aligned}\tag{2.6}$$

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 [35] 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 (2.5) 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. Conversely, if a global symmetry acts on-site and the Lagrangian density is invariant, it is clear that the symmetry can be gauged and there is no anomaly. See Appendix 2.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 2.B.1, the model (2.5) describes the large β gapless phase of the 2d XY-model (2.1) or (2.2). 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 2.C.1, is the 3d $U(1)$ gauge theory. The standard lattice model and its Villain version have a gapped confining phase [68]. 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 2.A, 2.B, and 2.C to review some well-known models. In particular, we will present lattice models of various spin systems (including the XY-model (2.1)) 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.

2.1.3 Outline

Following [35–38, 58], Sections 2.2 and 2.3 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 (2.6) 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 [35–37].

Even though we do not present it here, we have performed the same analysis for the exotic 3+1d continuum theories of [58], 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 [21] 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 [43].

In Section 2.2, we study the modified Villain formulation of the exotic 2+1d continuum theories of [35]. These include systems with global $U(1)$ subsystem symmetry and $U(1)$ and $\mathbb{Z}_N$ tensor gauge theories. We start with the XY-plaquette model of [22] 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 [35].

In Section 2.3, we study the modified Villain formulation of the exotic 3+1d continuum theories of [36, 37]. Again, these include systems with global $U(1)$ subsystem symmetry and $U(1)$ and $\mathbb{Z}_N$ tensor gauge theories. 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 2.2, these modified Villain models exhibit the same properties as their continuum counterparts in [37].

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 2.2 and 2.3.

Appendix 2.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 2.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 2.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.

2.2 2+1d (3d Euclidean) exotic theories

In this section, we describe modified Villain lattice models corresponding to the exotic 2+1d continuum theories of [35]. 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 x_μ to label the sites along the μ direction, so that $x_\mu \sim x_\mu + L_\mu$.

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 $\mathbf{1}_2$.

2.2.1 ϕ-theory (XY-plaquette model)

We start with a Euclidean spacetime version of the XY-plaquette model of [22]. The degrees of freedom are phases $e^{i\phi}$ at every site with the action

$$\beta_0 \sum_{\tau\text{-link}} [1 - \cos(\Delta_\tau \phi)] + \beta \sum_{xy\text{-plaq}} [1 - \cos(\Delta_x \Delta_y \phi)] \ . \tag{2.7}$$

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

$$\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{xy\text{-plaq}} (\Delta_x \Delta_y \phi - 2\pi n_{xy})^2 \ , \tag{2.8}$$

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

$$\begin{aligned} &\phi \sim \phi + 2\pi k \ , \\ &n_\tau \sim n_\tau + \Delta_\tau k \ , \\ &n_{xy} \sim n_{xy} + \Delta_x \Delta_y k \ , \end{aligned} \tag{2.9}$$

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

We suppress the "vortices" by modifying the Villain action (2.8) as

$$\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{xy\text{-plaq}} (\Delta_x \Delta_y \phi - 2\pi n_{xy})^2 + i \sum_{\text{cube}} \phi^{xy} (\Delta_\tau n_{xy} - \Delta_x \Delta_y n_\tau) \ , \tag{2.10}$$

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

$$\phi^{xy} \sim \phi^{xy} + 2\pi k^{xy} \ , \tag{2.11}$$

where k^{xy} is an integer-valued gauge parameter on the cubes of the lattice. We will refer to (2.10) as the modified Villain version of (2.7).

Self-Duality

Using the Poisson resummation formula (2.6), we can dualize the modified Villain action (2.10) to

$$\begin{aligned} \frac{1}{2(2\pi)^2\beta} \sum_{\text{dual } \tau\text{-link}} (\Delta_\tau \phi^{xy} - 2\pi n_\tau^{xy})^2 + \frac{1}{2(2\pi)^2\beta_0} \sum_{\text{dual } xy\text{-plaq}} (\Delta_x\Delta_y\phi^{xy} - 2\pi n)^2 \\ - i \sum_{\text{site}} \phi(\Delta_\tau n - \Delta_x\Delta_y n_\tau^{xy}) \ , \end{aligned} \tag{2.12}$$

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

$$\begin{aligned} &\phi^{xy} \sim \phi^{xy} + 2\pi k^{xy} \ , \\ &n_\tau^{xy} \sim n_\tau^{xy} + \Delta_\tau k^{xy} \ , \\ &n \sim n + \Delta_x\Delta_y k^{xy} \ . \end{aligned} \tag{2.13}$$

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 (2.10) is self-dual with $\beta_0 \leftrightarrow \frac{1}{(2\pi)^2\beta}$.

Global symmetries

In all the three models, (2.7), (2.8), and (2.10), there is a $(\mathbf{1}_0, \mathbf{1}_2)$ *momentum dipole symmetry*, which acts on the fields as

$$\phi \to \phi + c^x(\hat{x}) + c^y(\hat{y}) \ , \tag{2.14}$$

where $c^i(x_i)$ is real-valued. Due to the zero mode of the gauge symmetry (2.9), the momentum dipole symmetry is $U(1)$. Using (2.10), the components of the Noether current of the momentum dipole symmetry are

$$\begin{aligned} J_\tau &= i\beta_0(\Delta_\tau\phi - 2\pi n_\tau) = \frac{1}{2\pi}(\Delta_x\Delta_y\phi^{xy} - 2\pi n) \ , \\ J^{xy} &= i\beta(\Delta_x\Delta_y\phi - 2\pi n_{xy}) = \frac{1}{2\pi}(\Delta_\tau\phi^{xy} - 2\pi n_\tau^{xy}) \ . \end{aligned} \tag{2.15}$$

(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

$$\Delta_\tau J_\tau = \Delta_x\Delta_y J^{xy} \ , \tag{2.16}$$

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

$$\begin{aligned} Q^x(\hat{x}, \tilde{\mathcal{C}}^x) &= \sum_{\text{dual } xy\text{-plaq}\in\tilde{\mathcal{C}}^x} J_\tau + \sum_{\text{dual } \tau x\text{-plaq}\in\tilde{\mathcal{C}}^x} \Delta_x J^{xy} \ , \\ &= -\sum_{\text{dual } xy\text{-plaq}\in\tilde{\mathcal{C}}^x} n - \sum_{\text{dual } \tau x\text{-plaq}\in\tilde{\mathcal{C}}^x} \Delta_x n_\tau^{xy} \end{aligned} \tag{2.17}$$

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

$$\sum_{\hat{x}\text{: fixed } \hat{\tau}} Q^x(\hat{x}) = \sum_{\hat{y}\text{: fixed } \hat{\tau}} Q^y(\hat{y}) = \sum_{\text{dual } xy\text{-plaq: fixed } \hat{\tau}} J_\tau \ . \tag{2.18}$$

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

The modified Villain model (2.10) also has a $(\mathbf{1}_2, \mathbf{1}_0)$ *winding dipole symmetry*, which acts on the fields as

$$\phi^{xy} \to \phi^{xy} + c_x^{xy}(\hat{x}) + c_y^{xy}(\hat{y}) \ , \tag{2.19}$$

where $c_i^{xy}(x_i)$ is real-valued. By contrast, this symmetry is absent in the original lattice model (2.7) and its Villain version (2.8). Due to the zero mode of the gauge symmetry (2.13), the winding dipole symmetry is $U(1)$. The components of the Noether current of the winding dipole symmetry are

$$\begin{aligned} J_\tau^{xy} &= -\frac{i}{(2\pi)^2\beta}(\Delta_\tau\phi^{xy} - 2\pi n_\tau^{xy}) = \frac{1}{2\pi}(\Delta_x\Delta_y\phi - 2\pi n_{xy}) \ , \\ J &= -\frac{i}{(2\pi)^2\beta_0}(\Delta_x\Delta_y\phi^{xy} - 2\pi n) = \frac{1}{2\pi}(\Delta_\tau\phi - 2\pi n_\tau) \ . \end{aligned} \tag{2.20}$$

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

$$\Delta_\tau J_\tau^{xy} = \Delta_x\Delta_y J \ , \tag{2.21}$$

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

$$\begin{aligned} Q_x^{xy}(\hat{x}, \mathcal{C}^x) &= \sum_{xy\text{-plaq}\in\mathcal{C}^x} J_\tau^{xy} + \sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x J \ , \\ &= -\sum_{xy\text{-plaq}\in\mathcal{C}^x} n_{xy} - \sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x n_\tau \ , \end{aligned} \tag{2.22}$$

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

$$\sum_{\hat{x}:\ \text{fixed}\ \hat{\tau}} Q_x^{xy}(\hat{x}) = \sum_{\hat{y}:\ \text{fixed}\ \hat{\tau}} Q_y^{xy}(\hat{y}) = \sum_{xy\text{-plaq}:\ \text{fixed}\ \hat{\tau}} J_\tau^{xy} \ . \tag{2.23}$$

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 2.B.1.) The action is:

$$\begin{aligned}
&\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\phi - A_\tau - 2\pi n_\tau)^2 + \frac{\beta}{2}\sum_{xy\text{-plaq}}(\Delta_x\Delta_y\phi - A_{xy} - 2\pi n_{xy})^2 \\
&\quad + i\sum_{\text{cube}}\phi^{xy}(\Delta_\tau n_{xy} - \Delta_x\Delta_y n_\tau + N_{\tau xy}) \\
&\quad - \frac{i}{2\pi}\sum_{xy\text{-plaq}}\tilde{A}_\tau^{xy}(\Delta_x\Delta_y\phi - A_{xy} - 2\pi n_{xy}) - \frac{i}{2\pi}\sum_{\tau\text{-link}}\tilde{A}(\Delta_\tau\phi - A_\tau - 2\pi n_\tau) - i\sum_{\text{site}}\tilde{N}_\tau\phi \ ,
\end{aligned} \tag{2.24}$$

with the gauge symmetry

$$\begin{aligned}
&\phi \sim \phi + \alpha + 2\pi k \ , && \phi^{xy} \sim \phi^{xy} + \tilde{\alpha}^{xy} + 2\pi k^{xy} \ , \\
&A_\tau \sim A_\tau + \Delta_\tau\alpha + 2\pi K_\tau \ , && \tilde{A}_\tau^{xy} \sim \tilde{A}_\tau^{xy} + \Delta_\tau\tilde{\alpha}^{xy} + 2\pi\tilde{K}_\tau^{xy} \ , \\
&A_{xy} \sim A_{xy} + \Delta_x\Delta_y\alpha + 2\pi K_{xy} \ , && \tilde{A} \sim \tilde{A} + \Delta_x\Delta_y\tilde{\alpha}^{xy} + 2\pi\tilde{K} \ , \\
&n_\tau \sim n_\tau + \Delta_\tau k - K_\tau \ , && \tilde{N}_\tau \sim \tilde{N}_\tau + \Delta_\tau\tilde{K} - \Delta_x\Delta_y\tilde{K}_\tau^{xy} \ . \\
&n_{xy} \sim n_{xy} + \Delta_x\Delta_y k - K_{xy} \ , && \\
&N_{\tau xy} \sim N_{\tau xy} + \Delta_\tau K_{xy} - \Delta_x\Delta_y K_\tau \ , &&
\end{aligned} \tag{2.25}$$

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

$$
\begin{aligned}
&-\frac{i}{2\pi}\sum_{\text{site}}\tilde{\alpha}^{xy}(\Delta_\tau A_{xy}-\Delta_x\Delta_y A_\tau-2\pi N_{\tau xy})\\
&+i\sum_{xy\text{-plaq}}\tilde{K}^{xy}_\tau(A_{xy}+\Delta_x\Delta_y\alpha)+i\sum_{\tau\text{-link}}\tilde{K}(A_\tau+\Delta_\tau\alpha)-i\sum_{\text{site}}\tilde{N}_\tau\alpha\ .
\end{aligned}
\tag{2.26}
$$

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:

$$
\begin{aligned}
&n_\tau(L_\tau-1,\hat{x},\hat{y}) = \bar{n}^x(\hat{x})+\bar{n}^y(\hat{y})\ ,\\
&n_{xy}(\hat{\tau},\hat{x},L_y-1) = \bar{n}^{xy}_x(\hat{x})\ ,\\
&n_{xy}(\hat{\tau},L_x-1,\hat{y}) = \bar{n}^{xy}_y(\hat{y})\ ,
\end{aligned}
\tag{2.27}
$$

where $\bar{n}^i(x_i)$ and $\bar{n}^{xy}_i(x_i)$ are integer-valued. There is a gauge ambiguity in the zero modes of $\bar{n}^i(x_i)$, while $\bar{n}^{xy}_i(x_i)$ 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

$$
\phi \sim \phi + 2\pi w^x(\hat{x}) + 2\pi w^y(\hat{y})\ , \tag{2.28}
$$

where $w^i(x_i)$ is integer-valued.

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

$$\bar{\phi}(\hat{\tau},\hat{x},\hat{y}) = \phi(\hat{\tau},\hat{x},\hat{y}) \ , \qquad \text{for } 0 \leq x_\mu < L_\mu \ , \tag{2.29}$$

and beyond the fundamental domain, it is extended via

$$\begin{aligned}
\bar{\phi}(\hat{\tau}+L_\tau,\hat{x},\hat{y}) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\bar{n}^x(\hat{x}) - 2\pi\bar{n}^y(\hat{y}) \ , \\
\bar{\phi}(\hat{\tau},\hat{x}+L_x,\hat{y}) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\sum_{\hat{y}'=0}^{\hat{y}-1}\bar{n}^{xy}_y(\hat{y}') \ , \\
\bar{\phi}(\hat{\tau},\hat{x},\hat{y}+L_y) &= \bar{\phi}(\hat{\tau},\hat{x},\hat{y}) - 2\pi\sum_{\hat{x}'=0}^{\hat{x}-1}\bar{n}^{xy}_x(\hat{x}') \ .
\end{aligned} \tag{2.30}$$

In particular, in the gauge (2.27), $\Delta_\tau\bar{\phi} = \Delta_\tau\phi - 2\pi n_\tau$, 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 (2.10) in terms of $\bar{\phi}$ is

$$\frac{\beta_0}{2}\sum_{\tau\text{-link}}(\Delta_\tau\bar{\phi})^2 + \frac{\beta}{2}\sum_{xy\text{-plaq}}(\Delta_x\Delta_y\bar{\phi})^2 \ . \tag{2.31}$$

Let us discuss some charged configurations in the lattice model (2.31). We define the periodic Kronecker delta function

$$\delta^P(\hat{x},\hat{x}_0,L_x) \equiv \sum_{I\in\mathbb{Z}}\delta_{\hat{x},\hat{x}_0-IL_x} \ , \tag{2.32}$$

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

$$\Theta^P(0,\hat{x}_0,L_x) = 0 \ , \qquad \Delta_x\Theta^P(\hat{x},\hat{x}_0,L_x) = \delta^P(\hat{x},\hat{x}_0,L_x) \ . \tag{2.33}$$

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

$$\bar{\phi}(\hat{\tau},\hat{x},\hat{y}) = 2\pi\left[\frac{\hat{x}}{L_x}\Theta^P(\hat{y},\hat{y}_0,L_y) + \frac{\hat{y}}{L_y}\Theta^P(\hat{x},\hat{x}_0,L_x) - \frac{\hat{x}\hat{y}}{L_xL_y}\right] . \tag{2.34}$$

The most general winding configuration can be obtained by taking linear combinations with integer coefficients of (2.34) with different $\hat{x}_0, \hat{y}_0$ and adding to it a periodic function. The winding charges of (2.34) are $Q_x^{xy}(\hat{x}) = \delta^P(\hat{x},\hat{x}_0,L_x)$ and $Q_y^{xy}(\hat{y}) = \delta^P(\hat{y},\hat{y}_0,L_y)$. This configuration satisfies the equation of motion of $\bar{\phi}$, so it is a minimal action configuration with these winding charges. Its action is

$$\frac{\beta(2\pi)^2}{2}L_\tau\left(\frac{1}{L_x}+\frac{1}{L_y}-\frac{1}{L_xL_y}\right) . \tag{2.35}$$

Its Lorentzian interpretation is a winding state with energy

$$\frac{\beta(2\pi)^2}{2a}\left(\frac{1}{L_x}+\frac{1}{L_y}-\frac{1}{L_xL_y}\right) , \tag{2.36}$$

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

$$\int d\tau dx dy \ \left[\frac{\mu_0}{2}(\partial_\tau\phi)^2 + \frac{1}{2\mu}(\partial_x\partial_y\phi)^2\right] , \tag{2.37}$$

where we dropped the bar on ϕ. This is the Euclidean version of the 2+1d ϕ-theory of [35], which had been first introduced in [22]. (See also [69–73] 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:

$$\int d\tau dx dy \left[\frac{\mu_0}{2}(\partial_\tau\phi - A_\tau)^2 + \frac{1}{2\mu}(\partial_x\partial_y\phi - A_{xy})^2 - \frac{i}{2\pi}\tilde{A}_\tau^{xy}(\partial_x\partial_y\phi - A_{xy}) - \frac{i}{2\pi}\tilde{A}(\partial_\tau\phi - A_\tau)\right] , \tag{2.38}$$

with gauge symmetry

$$\begin{aligned} &\phi \sim \phi + \alpha \ , && \phi^{xy} \sim \phi^{xy} + \tilde{\alpha}^{xy} \ , \\ &A_\tau \sim A_\tau + \partial_\tau\alpha \ , && \tilde{A}_\tau^{xy} \sim \tilde{A}_\tau^{xy} + \partial_\tau\tilde{\alpha}^{xy} \ , \\ &A_{xy} \sim A_{xy} + \partial_x\partial_y\alpha \ , && \tilde{A} \sim \tilde{A} + \partial_x\partial_y\tilde{\alpha}^{xy} \ . \end{aligned} \tag{2.39}$$

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

$$-\frac{i}{2\pi}\int d\tau dx dy \ \tilde{\alpha}^{xy}(\partial_\tau A_{xy} - \partial_x\partial_y A_\tau) \ . \tag{2.40}$$

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 (2.26).

We can also view the modified Villain lattice model (2.10), or its gauge fixed version (2.31), as a discretized version the continuum theory (2.37). Our analysis of this lattice model makes rigorous the various assertions in [35]. Let us discuss them in more detail.

Both the continuum theory (2.37) and the lattice theory (2.31) have real-valued fields and the periodicity in field space is implemented using the twisted boundary conditions (2.30).

One could question whether the lattice theory (2.31) with this particular sum over twisted boundary conditions is fully consistent. In the continuum, this was discussed in detail in [35, 59]. On the lattice, the consistency follows from relating it to the lattice gauge theory (2.10) before the gauge fixing (2.27). Furthermore, the remaining gauge freedom (2.28) in the lattice theory (2.31) can now be interpreted as the gauge freedom of the continuum theory [35, 59].

The discussion of [35] uncovered a number of surprising properties of the continuum theory (2.37), which are not present in the original microscopic theory (2.7). 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 (2.10). A reader who was skeptical about the continuum analysis of [35] 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 (2.34) scales as $\ell_\tau/\mu\ell a$, which diverges as $1/a$ in the continuum limit. The configuration (2.34) gives a precise meaning to the winding configuration with infinite action in the continuum [35].[2] More generally, the classification of discontinuous configurations in the continuum theory (2.37) [35] is exactly as in the previous subsection.

In conclusion, the lattice model (2.10) flows in the continuum limit to (2.37). Conversely, the lattice model (2.10), or its gauge fixed version (2.31), gives a rigorous setting for the discussion of the continuum theory (2.37) of [35].

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

We can gauge the $U(1)$ momentum dipole symmetry by coupling (2.10) to the $(\mathbf{1}_0, \mathbf{1}_2)$ tensor gauge fields (A_τ, A_{xy}). We will consider this system in Section 2.2.3, 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 [35] (see also earlier work in [65, 70, 71, 74].

[2]The discussion of such infinite action and infinite energy configurations was described in [35] as "ambitious." It is rigorous in the context of the modified Villain model.

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

$$\gamma \sum_{\text{cube}} [1 - \cos(\Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau)] \ , \tag{2.41}$$

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

$$\begin{aligned} e^{iA_\tau} &\sim e^{iA_\tau + i\Delta_\tau \alpha} \ , \\ e^{iA_{xy}} &\sim e^{iA_{xy} + i\Delta_x \Delta_y \alpha} \ , \end{aligned} \tag{2.42}$$

with circle-valued α on the sites.

At large γ, we can approximate (2.41) by the Villain action

$$\frac{\gamma}{2} \sum_{\text{cube}} (\Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy})^2 \ , \tag{2.43}$$

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 (2.42) becomes

$$\begin{aligned} &A_\tau \sim A_\tau + \Delta_\tau \alpha + 2\pi k_\tau \ , \\ &A_{xy} \sim A_{xy} + \Delta_x \Delta_y \alpha + 2\pi k_{xy} \ , \\ &n_{\tau xy} \sim n_{\tau xy} + \Delta_\tau k_{xy} - \Delta_x \Delta_y k_\tau \ , \end{aligned} \tag{2.44}$$

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 (2.43) as in (2.10). 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 (2.43):

$$\frac{\gamma}{2}\sum_{\text{cube}} E_{xy}^2 + \frac{i\theta}{2\pi}\sum_{\text{cube}} E_{xy} \ , \tag{2.45}$$

where we defined the electric field

$$E_{xy} = \Delta_\tau A_{xy} - \Delta_x\Delta_y A_\tau - 2\pi n_{\tau xy} \ , \tag{2.46}$$

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 (2.41), while it is straightforward and natural in the Villain version (2.43).

The quantized electric fluxes

$$\begin{aligned} e^x(\hat{x}) &= \sum_{\text{cube: fixed } \hat{x}} E_{xy} = -2\pi \sum_{\text{cube: fixed } \hat{x}} n_{\tau xy} \in 2\pi\mathbb{Z} \ , \\ e^y(\hat{y}) &= \sum_{\text{cube: fixed } \hat{y}} E_{xy} = -2\pi \sum_{\text{cube: fixed } \hat{y}} n_{\tau xy} \in 2\pi\mathbb{Z} \ , \end{aligned} \tag{2.47}$$

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

$$\sum_{\hat{x}} e^x(\hat{x}) = \sum_{\hat{y}} e^y(\hat{y}) = \sum_{\text{cube}} E_{xy} \ . \tag{2.48}$$

Global symmetries

The three models (2.41), (2.43), and (2.45) have an *electric tensor symmetry* that acts on the fields as

$$A_\tau \to A_\tau + \lambda_\tau \ , \qquad A_{xy} \to A_{xy} + \lambda_{xy} \ , \tag{2.49}$$

where $(\lambda_\tau, \lambda_{xy})$ is a flat, real-valued tensor gauge field (i.e., it has vanishing field strength).[3] Due to the integer-valued gauge symmetry with (k_τ, k_{xy}) (2.44), the electric tensor symmetry is $U(1)$, rather than $\mathbb{R}$. The Noether current of this electric symmetry follows from (2.45)

$$J_\tau^{xy} = -i\gamma E_{xy} + \frac{\theta}{2\pi} \ . \tag{2.50}$$

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

$$\Delta_\tau J_\tau^{xy} = 0 \ , \qquad \Delta_x \Delta_y J_\tau^{xy} = 0 \ , \tag{2.51}$$

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

$$Q(\hat{x}, \hat{y}) = J_\tau^{xy} = Q^x(\hat{x}) + Q^y(\hat{y}) \ , \tag{2.52}$$

where $Q^i(x_i)$ is an integer, and the second equation follows from the Gauss law. The observables charged under the electric symmetry are the Wilson defect/operator

$$\begin{aligned} W^\tau(\hat{x}, \hat{y}) &= \exp\left[i \sum_{\tau\text{-link: fixed } \hat{x}, \hat{y}} A_\tau \right] \ , \\ W^x(\hat{x}, \mathcal{C}^x) &= \exp\left[i \sum_{xy\text{-plaq} \in \mathcal{C}^x} A_{xy} + i \sum_{\tau x\text{-plaq} \in \mathcal{C}^x} \Delta_x A_\tau \right] \ , \end{aligned} \tag{2.53}$$

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 $W^y(\hat{y}, \mathcal{C}^y)$.

[3]Using the α gauge symmetry of (2.44), 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 $c^i(x_i)$ and $c_i^{xy}(x_i)$ are real-valued.

Gauge-fixing and the continuum limit

Using the integer gauge freedom (2.44), we gauge fix $n_{\tau xy} = 0$, except for

$$n_{\tau xy}(L_\tau - 1, \hat{x}, L_y - 1) \equiv \bar{n}^x_{\tau xy}(\hat{x}) \ , \qquad n_{\tau xy}(L_\tau - 1, L_x - 1, \hat{y}) \equiv \bar{n}^y_{\tau xy}(\hat{y}) \ . \tag{2.54}$$

The integers $\bar{n}^i_{\tau xy}(x_i)$ 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

$$\begin{aligned} A_\tau &\sim A_\tau + \Delta_\tau \alpha + 2\pi k_\tau \ , \\ A_{xy} &\sim A_{xy} + \Delta_x \Delta_y \alpha + 2\pi k_{xy} \ , \end{aligned} \tag{2.55}$$

where (k_τ, k_{xy}) is a flat, integer-valued tensor gauge field.

Similar to (2.30) 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

$$\Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau = \Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy} \ . \tag{2.56}$$

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 (2.45) in terms of $(\bar{A}_\tau, \bar{A}_{xy})$ is

$$\frac{\gamma}{2} \sum_{\text{cube}} \bar{E}^2_{xy} + \frac{i\theta}{2\pi} \sum_{\text{cube}} \bar{E}_{xy} \ , \tag{2.57}$$

where we defined the new electric field

$$\bar{E}_{xy} = \Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau \ , \tag{2.58}$$

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 $\bar{A}_{xy} = a^2 A_{xy}$,[4] the action becomes

$$\begin{aligned} &\int d\tau dx dy \left(\frac{1}{g_e^2} E_{xy}^2 + \frac{i\theta}{2\pi} E_{xy} \right) , \\ &E_{xy} = \partial_\tau A_{xy} - \partial_x \partial_y A_\tau \ . \end{aligned} \tag{2.59}$$

This is the Euclidean version of the continuum 2+1d A-theory of [35]. (See also [65, 70, 71, 74].) The Villain model (2.45) 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 [35].

We conclude that the lattice model (2.45) flows in the continuum limit to (2.59). Conversely, the lattice model (2.45), or its gauge fixed version (2.57), give a rigorous setting for the discussion of the continuum theory (2.59) of [35].

2.2.3 $\mathbb{Z}_N$ tensor gauge theory

In this subsection, we will consider the modified Villain lattice version of the 2+1d $\mathbb{Z}_N$ Ising plaquette model [64]. 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 [35].

We can restrict the $U(1)$ variables in the $U(1)$ tensor gauge theory (2.41) 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

[4] 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.

the $\mathbb{Z}_N$ tensor gauge theory with the action

$$\gamma \sum_{\text{cube}} \left[1 - \cos\left(\frac{2\pi}{N} (\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau) \right) \right] \ . \tag{2.60}$$

At large γ, $\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau = 0 \bmod N$ and we can replace the action by

$$\frac{2\pi i}{N} \sum_{\text{cube}} \tilde{m}^{xy} (\Delta_\tau m_{xy} - \Delta_x \Delta_y m_\tau) \ , \tag{2.61}$$

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 (2.192) for the topological lattice $\mathbb{Z}_N$ gauge theory reviewed in Appendix 2.C.

There is a gauge symmetry

$$\begin{aligned} & m_\tau \sim m_\tau + \Delta_\tau \ell + N k_\tau \ , \\ & m_{xy} \sim m_{xy} + \Delta_x \Delta_y \ell + N k_{xy} \ , \\ & \tilde{m}^{xy} \sim \tilde{m}^{xy} + N \tilde{k}^{xy} \ , \end{aligned} \tag{2.62}$$

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, (2.60) and (2.61), 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 (2.61), the charge operator is

$$U(\hat{\tau}, \hat{x}, \hat{y}) = \exp\left[\frac{2\pi i}{N} \tilde{m}^{xy} \right] \ . \tag{2.63}$$

The observables charged under the electric symmetry are the Wilson defect/operator

$$\begin{aligned} W^\tau(\hat{x},\hat{y}) &= \exp\left[\frac{2\pi i}{N}\sum_{\tau\text{-link: fixed }\hat{x},\hat{y}} m_\tau\right] ~, \\ W^x(\hat{x},\mathcal{C}^x) &= \exp\left[\frac{2\pi i}{N}\sum_{xy\text{-plaq}\in\mathcal{C}^x} m_{xy} + \frac{2\pi i}{N}\sum_{\tau x\text{-plaq}\in\mathcal{C}^x} \Delta_x m_\tau\right] ~, \end{aligned} \tag{2.64}$$

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 (2.61), but not in (2.60), 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 (2.61) 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 $\tilde{m}^{xy}$

$$\Delta_\tau \tilde{m}^{xy} = \Delta_x\Delta_y \tilde{m}^{xy} = 0 \bmod N ~. \tag{2.65}$$

The gauge inequivalent configurations of $\tilde{m}^{xy}$ are

$$\tilde{m}^{xy}(\hat{\tau},\hat{x},\hat{y}) = \tilde{m}^{xy}_x(\hat{x}) + \tilde{m}^{xy}_y(\hat{y}) ~, \tag{2.66}$$

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 (2.61) 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 (2.61) 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 $(\tilde{n}_\tau^{xy}, \tilde{n})$. 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

$$\begin{aligned} &\frac{iN}{2\pi}\sum_{\text{cube}}\tilde{\phi}^{xy}(\Delta_\tau A_{xy} - \Delta_x\Delta_y A_\tau - 2\pi n_{\tau xy}) + iN\sum_{xy\text{-plaq}} A_{xy}\tilde{n}_\tau^{xy} \\ &\quad + iN\sum_{\tau\text{-link}} A_\tau \tilde{n} + i\sum_{\text{site}}\phi(\Delta_\tau\tilde{n} - \Delta_x\Delta_y\tilde{n}_\tau^{xy}) \ , \end{aligned} \tag{2.67}$$

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 $n_{\tau xy}$, $\tilde{n}_\tau^{xy}$ and $\tilde{n}$ are integer-valued fields on the cubes, the dual τ-links, and the dual xy-plaquettes, respectively.

There action (2.67) has the gauge symmetry

$$\begin{aligned} &\phi \sim \phi + N\alpha + 2\pi k \ , \\ &\tilde{\phi}^{xy} \sim \tilde{\phi}^{xy} + 2\pi\tilde{k}^{xy} \ , \\ &A_\tau \sim A_\tau + \Delta_\tau\alpha + 2\pi k_\tau \ , \\ &A_{xy} \sim A_{xy} + \Delta_x\Delta_y\alpha + 2\pi k_{xy} \ , \\ &\tilde{n}_\tau^{xy} \sim \tilde{n}_\tau^{xy} + \Delta_\tau\tilde{k}^{xy} \ , \\ &\tilde{n} \sim \tilde{n} + \Delta_x\Delta_y\tilde{k}^{xy} \ , \\ &n_{\tau xy} \sim n_{\tau xy} + \Delta_\tau k_{xy} - \Delta_x\Delta_y k_\tau \ . \end{aligned} \tag{2.68}$$

As a check, summing over the integer-valued fields $n_{\tau xy}$, $\tilde{n}_\tau^{xy}$, and $\tilde{n}$ in (2.67) constrains

$$\tilde{\phi}^{xy} = \frac{2\pi}{N}\tilde{m}^{xy}, \quad A_\tau - \frac{1}{N}\Delta_\tau\phi = \frac{2\pi}{N}m_\tau, \quad A_{xy} - \frac{1}{N}\Delta_x\Delta_y\phi = \frac{2\pi}{N}m_{xy} \ , \tag{2.69}$$

where $\tilde{m}^{xy}$, m_τ and m_{xy} are integer-valued fields. Substituting them back into the action leads to (2.61).

We will refer to the presentation (2.67) 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 (2.60), 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 (2.67) can also be derived through Higgsing the $U(1)$ tensor gauge theory (2.45) to a $\mathbb{Z}_N$ theory using the field ϕ in (2.10). The Higgs action is

$$\begin{aligned}
&\frac{i}{2\pi}\sum_{\tau\text{-link}}\tilde{B}(\Delta_\tau\phi - NA_\tau - 2\pi n_\tau) + \frac{i}{2\pi}\sum_{xy\text{-plaq}}\tilde{E}^{xy}(\Delta_x\Delta_y\phi - NA_{xy} - 2\pi n_{xy}) \\
&\quad - i\sum_{\text{cube}}\tilde{\phi}^{xy}\left(\Delta_\tau n_{xy} - \Delta_x\Delta_y n_\tau + Nn_{\tau xy}\right) \ ,
\end{aligned} \tag{2.70}$$

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 (2.68), there is a gauge symmetry

$$\begin{aligned}
&n_\tau \sim n_\tau + \Delta_\tau k - Nk_\tau \ , \\
&n_{xy} \sim n_{xy} + \Delta_x\Delta_y k - Nk_{xy} \ .
\end{aligned} \tag{2.71}$$

Summing over the integer-valued fields n_τ and n_{xy} constrains

$$\tilde{B} - \Delta_x\Delta_y\tilde{\phi}^{xy} = -2\pi\tilde{n} \ , \quad \tilde{E}^{xy} - \Delta_\tau\tilde{\phi}^{xy} = -2\pi\tilde{n}_\tau^{xy} \ , \tag{2.72}$$

where $\tilde{n}$ and $\tilde{n}^{xy}_\tau$ are integer-valued fields. Substituting them back into the action leads to (2.67).

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 $\Delta_\tau \tilde{n} - \Delta_x \Delta_y \tilde{n}^{xy}_\tau = 0$. Then we can gauge fix $n_{\tau xy}$, $\tilde{n}^{xy}_\tau$ and $\tilde{n}$ to be zero almost everywhere except at

$$
\begin{aligned}
&n_{\tau xy}(L_\tau - 1, \hat{x}, L_y - 1) = \bar{n}^x_{\tau xy}(\hat{x}) \ , \\
&n_{\tau xy}(L_\tau - 1, L_x - 1, \hat{y}) = \bar{n}^y_{\tau xy}(\hat{y}) \ , \\
&\tilde{n}^{xy}_\tau(L_\tau - 1, \hat{x}, \hat{y}) = \bar{n}^{xy}_{\tau,x}(\hat{x}) + \bar{n}^{xy}_{\tau,y}(\hat{y}) \ , \\
&\tilde{n}(\hat{\tau}, \hat{x}, L_y - 1) = \bar{n}_x(\hat{x}) \ , \\
&\tilde{n}(\hat{\tau}, L_x - 1, \hat{y}) = \bar{n}_y(\hat{y}) \ ,
\end{aligned}
\tag{2.73}
$$

where $\bar{n}^x_{\tau xy}, \bar{n}^y_{\tau xy}, \bar{n}^{xy}_{\tau,x}, \bar{n}^{xy}_{\tau,y}, \bar{n}_x, \bar{n}_y$ are all integer-valued. These integers obey $\bar{n}^x_{\tau xy}(L_x - 1) = \bar{n}^y_{\tau xy}(L_y - 1)$ and $\bar{n}_x(L_x - 1) = \bar{n}_y(L_y - 1)$. The zero modes of $\bar{n}^{xy}_{\tau,x}(\hat{x})$ and $\bar{n}^{xy}_{\tau,y}(\hat{y})$ have a gauge ambiguity.

As in Sections 2.2.1 and 2.2.2, 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

$$
\begin{aligned}
&\Delta_\tau \bar{\phi}^{xy} = \Delta_\tau \tilde{\phi}^{xy} - 2\pi \tilde{n}^{xy}_\tau \ , \\
&\Delta_x \Delta_y \bar{\phi}^{xy} = \Delta_x \Delta_y \tilde{\phi}^{xy} - 2\pi \tilde{n} \ , \\
&\Delta_\tau \bar{A}_{xy} - \Delta_x \Delta_y \bar{A}_\tau = \Delta_\tau A_{xy} - \Delta_x \Delta_y A_\tau - 2\pi n_{\tau xy} \ .
\end{aligned}
\tag{2.74}
$$

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 (2.67) becomes

$$\begin{aligned}&\frac{iN}{2\pi}\sum_{\text{cube}}\bar{\phi}^{xy}(\Delta_\tau \bar{A}_{xy}-\Delta_x\Delta_y\bar{A}_\tau)+iN\sum_{\substack{xy\text{-plaq}\\ \hat{\tau}=L_\tau-1}}\bar{A}_{xy}(\bar{n}^{xy}_{\tau,x}+\bar{n}^{xy}_{\tau,y})\\&+iN\sum_{\substack{\tau\text{-link}\\ \hat{x}=L_x-1}}\bar{A}_\tau\bar{n}_y+iN\sum_{\substack{\tau\text{-link}\\ \hat{y}=L_y-1}}\bar{A}_\tau\bar{n}_x-iN\sum_{\substack{\tau\text{-link}\\ \hat{x}=L_x-1\\ \hat{y}=L_y-1}}\bar{A}_\tau\bar{n}_x\ ,\end{aligned}\tag{2.75}$$

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

$$\frac{iN}{2\pi}\int d\tau dx dy\ \phi^{xy}(\partial_\tau A_{xy}-\partial_x\partial_y A_\tau)\ ,\tag{2.76}$$

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}).[5] This is the Euclidean version of the 2+1d $\mathbb{Z}_N$ tensor gauge theory of [35].

We conclude that the lattice model (2.61), or equivalently (2.67), flows in the continuum limit to (2.76). Conversely, the lattice model (2.61), or equivalently (2.67), gives a rigorous setting for the discussion of the continuum theory (2.76) of [35].

2.3 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 [36, 37]. 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 $x_\mu \sim x_\mu + L_\mu$.

[5]Such boundary terms are necessary in order to make the continuum action (2.76) well-defined. They played a crucial role in the analysis of [59].

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 $T_{[xy]z} + T_{[yz]x} + T_{[zx]y} = 0$.

2.3.1 ϕ-theory

There is a $U(1)$ variable $e^{i\phi}$ at each site of the lattice. The action is

$$\beta_0 \sum_{\tau\text{-link}} [1 - \cos(\Delta_\tau \phi)] + \beta \sum_{i<j} \sum_{ij\text{-plaq}} [1 - \cos(\Delta_i \Delta_j \phi)] \ , \tag{2.77}$$

where ϕ is circle-valued. At large β_0, β, we can approximate the action with the Villain action

$$\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{i<j} \sum_{ij\text{-plaq}} (\Delta_i \Delta_j \phi - 2\pi n_{ij})^2 \ , \tag{2.78}$$

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

$$\phi \sim \phi + 2\pi p \ , \qquad n_\tau \sim n_\tau + \Delta_\tau p \ , \qquad n_{ij} \sim n_{ij} + \Delta_i \Delta_j p \ , \tag{2.79}$$

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

$$\begin{aligned} &\frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{i<j} \sum_{ij\text{-plaq}} (\Delta_i \Delta_j \phi - 2\pi n_{ij})^2 \\ &+ i \sum_{i<j} \sum_{\tau ij\text{-cube}} \hat{A}^{ij} (\Delta_\tau n_{ij} - \Delta_i \Delta_j n_\tau) - i \sum_{\substack{\text{cyclic} \\ i,j,k}} \sum_{xyz\text{-cube}} \hat{A}_\tau^{[ij]k} (\Delta_i n_{jk} - \Delta_j n_{ik}) \ , \end{aligned} \tag{2.80}$$

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

$$\begin{aligned} &\hat{A}_\tau^{[ij]k} \sim \hat{A}_\tau^{[ij]k} + \Delta_\tau \hat{\alpha}^{[ij]k} + 2\pi \hat{q}_\tau^{[ij]k} \ , \\ &\hat{A}^{ij} \sim \hat{A}^{ij} + \Delta_k \hat{\alpha}^{k(ij)} + 2\pi \hat{q}^{ij} \ . \end{aligned} \tag{2.81}$$

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 2.2.1, 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 [36]. See also [54, 75, 67, 76] 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 (2.6) to dualize the modified Villain action (2.80) to

$$\frac{1}{2(2\pi)^2\beta} \sum_{\substack{\text{cyclic} \\ i,j,k}} \sum_{\text{dual } \tau k\text{-plaq}} (\Delta_\tau \hat{A}^{ij} - \Delta_k \hat{A}_\tau^{k(ij)} - 2\pi \hat{n}_\tau^{ij})^2$$
$$+ \frac{1}{2(2\pi)^2\beta_0} \sum_{\text{dual } xyz\text{-cube}} \left(\sum_{i<j} \Delta_i \Delta_j \hat{A}^{ij} - 2\pi \hat{n} \right)^2 - i \sum_{\text{site}} \phi \left(\Delta_\tau \hat{n} - \sum_{i<j} \Delta_i \Delta_j \hat{n}_\tau^{ij} \right) , \tag{2.82}$$

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[6]

$$\begin{aligned}
&\hat{A}_\tau^{k(ij)} \sim \hat{A}_\tau^{k(ij)} + \Delta_\tau \hat{\alpha}^{k(ij)} + 2\pi \hat{q}_\tau^{k(ij)} \ , \\
&\hat{A}^{ij} \sim \hat{A}^{ij} + \Delta_k \hat{\alpha}^{k(ij)} + 2\pi \hat{q}^{ij} \ , \\
&\hat{n}_\tau^{ij} \sim \hat{n}_\tau^{ij} + \Delta_\tau \hat{q}^{ij} - \Delta_k \hat{q}_\tau^{k(ij)} \ , \\
&\hat{n} \sim \hat{n} + \sum_{i<j} \Delta_i \Delta_j \hat{q}^{ij} \ .
\end{aligned} \tag{2.83}$$

The Lagrange multiplier ϕ imposes the flatness constraint of $(\hat{n}_\tau^{ij}, \hat{n})$.

Once again, following similar steps in Section 2.2.1, 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 [36] (see also [54, 67]). 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

[6] $(\hat{n}_\tau^{ij}, \hat{n})$ is the $\mathbb{Z}$ version of $(\hat{C}_\tau^{ij}, \hat{C})$ of [38].

(winding) dipole symmetry of ϕ-theory to the magnetic (electric) dipole symmetry of the $\hat{A}$-theory, exactly like in the continuum theories.

In conclusion, the modified Villain action (2.80) has the same continuum limit as the XY-plaquette action (2.77). It has all the properties of the continuum ϕ-theory of [36] 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.

2.3.2 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

$$\gamma_0 \sum_{i<j} \sum_{\tau ij\text{-cube}} [1-\cos(\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau)] + \gamma \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic}\\ i,j,k}} [1-\cos(\Delta_i A_{jk} - \Delta_j A_{ik})] \ , \tag{2.84}$$

where (A_τ, A_{ij}) are circle-valued. This action has a tensor gauge symmetry

$$\begin{aligned} e^{iA_\tau} &\sim e^{iA_\tau + i\Delta_\tau \alpha} \ , \\ e^{iA_{ij}} &\sim e^{iA_{ij} + i\Delta_i \Delta_j \alpha} \ , \end{aligned} \tag{2.85}$$

with circle valued α at the sites.

At large γ_0, γ, we can approximate the action, à la Villain, as

$$\frac{\gamma_0}{2} \sum_{i<j} \sum_{\tau ij\text{-cube}} (\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau - 2\pi n_{\tau ij})^2 + \frac{\gamma}{2} \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic}\\ i,j,k}} (\Delta_i A_{jk} - \Delta_j A_{ik} - 2\pi n_{[ij]k})^2 \ , \tag{2.86}$$

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 (2.85) is now replaced with

$$\begin{aligned} A_\tau &\sim A_\tau + \Delta_\tau \alpha + 2\pi q_\tau \ , \\ A_{ij} &\sim A_{ij} + \Delta_i \Delta_j \alpha + 2\pi q_{ij} \ , \\ n_{\tau ij} &\sim n_{\tau ij} + \Delta_\tau q_{ij} - \Delta_i \Delta_j q_\tau \ , \\ n_{[ij]k} &\sim n_{[ij]k} + \Delta_i q_{jk} - \Delta_j q_{ik} \ . \end{aligned} \tag{2.87}$$

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.[7]

Next, we suppress the "vortices" by modifying the Villain action as

$$\begin{aligned} &\frac{\gamma_0}{2} \sum_{i<j} \sum_{\tau ij\text{-cube}} (\Delta_\tau A_{ij} - \Delta_i \Delta_j A_\tau - 2\pi n_{\tau ij})^2 + \frac{\gamma}{2} \sum_{xyz\text{-cube}} \sum_{\substack{\text{cyclic} \\ i,j,k}} (\Delta_i A_{jk} - \Delta_j A_{ik} - 2\pi n_{[ij]k})^2 \\ &\quad + i \sum_{\text{dual site}} \sum_{\substack{\text{cyclic} \\ i,j,k}} \hat{\phi}^{[ij]k} (\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik}) \ , \end{aligned} \tag{2.88}$$

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

$$\hat{\phi}^{[ij]k} \sim \hat{\phi}^{[ij]k} + 2\pi \hat{p}^{[ij]k} \ , \tag{2.89}$$

where $\hat{p}^{[ij]k}$ is an integer-valued field on the dual sites.

Following similar steps in Section 2.2.1, 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 [36]. See also [63,54,65,66,75,67] 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 3+1d A-theory.

[7] $(n_{\tau ij}, n_{[ij]k})$ is the $\mathbb{Z}$ version of $(C_\tau^{ij}, C^{[ij]k})$ of [38].

Alternatively, we can apply the Poisson resummation formula (2.6) to dualize the modified Villain action (2.88) to

$$\frac{1}{6(2\pi)^2\gamma}\sum_{\text{dual }\tau\text{-link}}\sum_{\substack{\text{cyclic}\\ i,j,k}}(\Delta_\tau\hat{\phi}^{k(ij)}-2\pi\hat{n}_\tau^{k(ij)})^2+\frac{1}{2(2\pi)^2\gamma_0}\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{\text{dual }k\text{-link}}(\Delta_k\hat{\phi}^{k(ij)}-2\pi\hat{n}^{ij})^2$$
$$+\,i\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{ij\text{-plaq}}A_{ij}(\Delta_\tau\hat{n}^{ij}-\Delta_k\hat{n}_\tau^{k(ij)})+i\sum_{\tau\text{-link}}A_\tau\sum_{i<j}\Delta_i\Delta_j\hat{n}^{ij}\ , \tag{2.90}$$

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

$$\begin{aligned}
&\hat{\phi}^{k(ij)}\sim\hat{\phi}^{k(ij)}+2\pi\hat{p}^{k(ij)}\ ,\\
&\hat{n}_\tau^{k(ij)}\sim\hat{n}_\tau^{k(ij)}+\Delta_\tau\hat{p}^{k(ij)}\ ,\\
&\hat{n}^{ij}\sim\hat{n}^{ij}+\Delta_k\hat{p}^{k(ij)}\ .
\end{aligned} \tag{2.91}$$

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 (2.90) is the modified Villain action of the $\hat{\phi}$-theory of [36].

Once again, following similar steps in Section 2.2.1, 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 [36]. 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 (2.84) and the modified Villain action (2.88) 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 [36].

2.3.3 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 [54, 37].

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 [21]

$$\begin{aligned}
&H = -\beta_1 \sum_{\text{site}} (G_{\hat{s},[yz]x} + G_{\hat{s},[zx]y} + G_{\hat{s},[xy]z}) - \beta_2 \sum_{\text{cube}} L_{\hat{c}} + c.c. \ , \\
&G_{\hat{s},[yz]x} = V_{\hat{s}+(0,\frac{1}{2},0)} V^\dagger_{\hat{s}+(0,0,\frac{1}{2})} V^\dagger_{\hat{s}-(0,\frac{1}{2},0)} V_{\hat{s}-(0,0,\frac{1}{2})} \ , \\
&G_{\hat{s},[zx]y} = V^\dagger_{\hat{s}+(\frac{1}{2},0,0)} V_{\hat{s}+(0,0,\frac{1}{2})} V_{\hat{s}-(\frac{1}{2},0,0)} V^\dagger_{\hat{s}-(0,0,\frac{1}{2})} \ , \\
&G_{\hat{s},[xy]z} = V_{\hat{s}+(\frac{1}{2},0,0)} V^\dagger_{\hat{s}+(0,\frac{1}{2},0)} V^\dagger_{\hat{s}-(\frac{1}{2},0,0)} V_{\hat{s}-(0,\frac{1}{2},0)} \ , \\
&L_{\hat{c}} = U_{\hat{c}+(\frac{1}{2},\frac{1}{2},0)} U^\dagger_{\hat{c}+(-\frac{1}{2},\frac{1}{2},0)} U^\dagger_{\hat{c}+(\frac{1}{2},-\frac{1}{2},0)} U_{\hat{c}-(\frac{1}{2},\frac{1}{2},0)} \\
&\qquad U_{\hat{c}+(0,\frac{1}{2},\frac{1}{2})} U^\dagger_{\hat{c}+(0,-\frac{1}{2},\frac{1}{2})} U^\dagger_{\hat{c}+(0,\frac{1}{2},-\frac{1}{2})} U_{\hat{c}-(0,\frac{1}{2},\frac{1}{2})} \\
&\qquad U_{\hat{c}+(\frac{1}{2},0,\frac{1}{2})} U^\dagger_{\hat{c}+(-\frac{1}{2},0,\frac{1}{2})} U^\dagger_{\hat{c}+(\frac{1}{2},0,-\frac{1}{2})} U_{\hat{c}-(\frac{1}{2},0,\frac{1}{2})} \ .
\end{aligned} \tag{2.92}$$

All the terms in the Hamiltonian commute with each other. The operators $G_{\hat{s},[ij]k}$ are in the $\mathbf{2}$ 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.[8] 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 $U =$

[8] To clarify the terminology, recall that each symmetry operator is associated with a geometrical object $\mathcal{C}$. According to [77], 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 [57] is faithful, while the relativistic q-form symmetry of [28] is not faithful. In [37], the faithful symmetry was referred to as "unconstrained" and the unfaithful symmetry was referred to as "constrained."

$\exp(\frac{2\pi i\hat{m}^{ij}}{N})$. 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 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

$$\frac{2\pi i}{N}\sum_{\substack{\text{cyclic}\\ i,j,k}}\sum_{\tau k\text{-plaq}} m_{ij}\left(\Delta_\tau \hat{m}^{ij} - \Delta_k \hat{m}_\tau^{k(ij)}\right) + \frac{2\pi i}{N}\sum_{xyz\text{-cube}} m_\tau \left(\sum_{i<j}\Delta_i\Delta_j \hat{m}^{ij}\right) \ . \tag{2.93}$$

There are gauge symmetries:

$$\begin{aligned}
&m_\tau \sim m_\tau + \Delta_\tau \ell + N q_\tau \ , \\
&m_{ij} \sim m_{ij} + \Delta_i\Delta_j \ell + N q_{ij} \ , \\
&\hat{m}_\tau^{k(ij)} \sim \hat{m}_\tau^{k(ij)} + \Delta_\tau \hat{\ell}^{k(ij)} + N\hat{q}_\tau^{k(ij)} \ , \\
&\hat{m}^{ij} \sim \hat{m}^{ij} + \Delta_k \hat{\ell}^{k(ij)} + N\hat{q}^{ij} \ ,
\end{aligned} \tag{2.94}$$

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 (2.192) for the topological lattice $\mathbb{Z}_N$ gauge theory reviewed in Appendix 2.C and the presentation (2.61) 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 (2.82) and (2.88), we can interpret (2.93) 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

$$W^\tau(\hat{x}, \hat{y}, \hat{z}) = \exp\left[\frac{2\pi i}{N} \sum_{\text{dual } \tau\text{-link: fixed } \hat{x},\hat{y},\hat{z}} m_\tau\right] , \tag{2.95}$$

and the probe z-lineon defect is

$$\hat{W}^z(\hat{x}, \hat{y}, \mathcal{C}^z) = \exp\left[\frac{2\pi i}{N} \sum_{\tau\text{-link}\in\mathcal{C}^z} \hat{m}_\tau^{z(xy)} + \frac{2\pi i}{N} \sum_{z\text{-link}\in\mathcal{C}^z} \hat{m}^{xy}\right] , \tag{2.96}$$

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 (2.96) 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 [77]). The charged observables are the probe fracton defect (2.95) and the Wilson observable

$$W^{xy}(\hat{z}, \mathcal{C}^{xy}) = \exp\left[\frac{2\pi i}{N}\left(\sum_{\text{dual } xz\text{-plaq}\in\mathcal{C}^{xy}} m_{xz} + \sum_{\text{dual } yz\text{-plaq}\in\mathcal{C}^{xy}} m_{yz} + \sum_{\text{dual } \tau z\text{-plaq}\in\mathcal{C}^{xy}} \Delta_z m_\tau\right)\right] , \tag{2.97}$$

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 (2.95), (2.97) 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 [77]). The charge operators are (2.96) and similar operators on the other planes.

Real BF-action and the continuum limit

As in Section 2.2.3, we discuss another presentation of this theory, which is closer to the continuum action.

Starting from the integer BF-action (2.93), 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

$$\begin{aligned}
&\frac{iN}{2\pi} \sum_{\tau k\text{-plaq}} A_{ij} \left(\Delta_\tau \hat{A}^{ij} - \Delta_k \hat{A}_\tau^{k(ij)} - 2\pi \hat{n}_\tau^{ij} \right) + \frac{iN}{2\pi} \sum_{xyz\text{-cube}} A_\tau \left(\Delta_i \Delta_j \hat{A}^{ij} - 2\pi \hat{n} \right) \\
&+ iN \sum_{ij\text{-plaq}} \hat{A}^{ij} n_{\tau ij} - iN \sum_{\tau\text{-link}} \hat{A}_\tau^{[ij]k} n_{[ij]k} - i \sum_{\text{dual site}} \phi \left(\Delta_\tau \hat{n} - \Delta_i \Delta_j \hat{n}_\tau^{ij} \right) \\
&- i \sum_{\text{site}} \hat{\phi}^{[ij]k} (\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik}) \ .
\end{aligned} \tag{2.98}$$

(To simplify this particular expression and (2.101), 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 (2.79), (2.83) (2.87), (2.91) except that the α and $\hat{\alpha}^{[ij]k}$ gauge symmetry also acts on ϕ and $\hat{\phi}^{[ij]k}$ as

$$\begin{aligned} &\phi \sim \phi + N\alpha \ , \\ &\hat{\phi}^{[ij]k} \sim \hat{\phi}^{[ij]k} + N\hat{\alpha}^{[ij]k} \ . \end{aligned} \tag{2.99}$$

As a check, summing over $(\hat{n}^{ij}_\tau, \hat{n})$ and $(n_{\tau ij}, n_{[ij]k})$ in (2.98) constrains

$$\begin{aligned} &\left(A_\tau - \frac{1}{N}\Delta_\tau\phi, A_{ij} - \frac{1}{N}\Delta_i\Delta_j\phi\right) = \frac{2\pi}{N}(m_\tau, m_{ij}) \ , \\ &\left(\hat{A}^{k(ij)}_\tau - \frac{1}{N}\Delta_\tau\hat{\phi}^{k(ij)}, \hat{A}^{ij} - \frac{1}{N}\Delta_k\hat{\phi}^{k(ij)}\right) = \frac{2\pi}{N}(\hat{m}^{k(ij)}_\tau, \hat{m}^{ij}) \ . \end{aligned} \tag{2.100}$$

Substituting them back to the action leads to (2.93).

The real BF-action (2.98) can also be derived through Higgsing the $U(1)$ tensor gauge theory (2.88) to a $\mathbb{Z}_N$ theory using the field ϕ in (2.80). The Higgs action is

$$\begin{aligned} &\frac{i}{2\pi}\sum_{\tau\text{-link}} \hat{B}(\Delta_\tau\phi - NA_\tau - 2\pi n_\tau) + \frac{i}{2\pi}\sum_{ij\text{-plaq}} \hat{E}^{ij}(\Delta_i\Delta_j\phi - NA_{ij} - 2\pi n_{ij}) \\ &- i\sum_{\tau ij\text{-cube}} \hat{A}^{ij}(\Delta_\tau n_{ij} - \Delta_i\Delta_j n_\tau - Nn_{\tau ij}) + i\sum_{xyz\text{-cube}} \hat{A}^{[ij]k}_\tau(\Delta_i n_{jk} - \Delta_j n_{ik} - Nn_{[ij]k}) \\ &- i\sum_{\text{dual site}} \hat{\phi}^{[ij]k}(\Delta_\tau n_{[ij]k} - \Delta_i n_{\tau jk} + \Delta_j n_{\tau ik}) \ , \end{aligned} \tag{2.101}$$

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 (2.79), (2.83), (2.87), (2.91), and (2.99). In addition, the fields (n_τ, n_{ij}) also transform under the (q_τ, q_{ij}) gauge symmetry

$$\begin{aligned} &n_\tau \sim n_\tau - Nq_\tau \ , \\ &n_{ij} \sim n_{ij} - Nq_{ij} \ . \end{aligned} \tag{2.102}$$

Summing over the integer-valued fields (n_τ, n_{ij}) constrains

$$\hat{B} - \sum_{i<j} \Delta_i \Delta_j \hat{A}^{ij} = -2\pi\hat{n} \ , \quad \hat{E}^{ij} - \Delta_\tau \hat{A}^{ij} + \Delta_k \hat{A}_\tau^{k(ij)} = -2\pi \hat{n}_\tau^{ij} \ , \tag{2.103}$$

where $\hat{n}$ and $\hat{n}_\tau^{ij}$ are integer-valued fields. Substituting them back into the action leads to (2.98). Similarly, the real BF-action (2.98) can also be derived through Higgsing the $U(1)$ tensor gauge theory (2.82) to a $\mathbb{Z}_N$ theory using the field $\hat{\phi}^{[ij]k}$ in (2.90).

Let us discuss a convenient gauge choice for this lattice model. Following similar steps in Section 2.2.3 and in Appendix 2.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:

$$\frac{iN}{2\pi} \int d\tau dx dy dz \left[\sum_{\substack{\text{cyclic} \\ i,j,k}} A_{ij} \left(\partial_\tau \hat{A}^{ij} - \partial_k \hat{A}_\tau^{k(ij)} \right) + A_\tau \left(\sum_{i<j} \partial_i \partial_j \hat{A}^{ij} \right) \right] \ , \tag{2.104}$$

where we omit the terms that depend on the transition functions of these fields.[9] This is the Euclidean version of the 3+1d $\mathbb{Z}_N$ tensor gauge theory of [54, 37] which describes the low-energy limit of the X-cube model.

We conclude that the modified Villain lattice model (2.93), or equivalently (2.98), flows to the same continuum field theory (2.104) as the original X-cube model (2.92). Conversely, the modified Villain lattice model (2.93), or equivalently (2.98), gives a rigorous setting for the discussion of the continuum theory (2.104) of [54, 37].

[9] Such boundary terms are necessary in order to make the continuum action (2.104) well-defined. They played a crucial role in the analysis of [59].

2.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.

2.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.

The problem is characterized by the Euclidean continuum action

$$S = \oint d\tau \left(\frac{1}{2} (\partial_\tau q)^2 + \frac{i\theta}{2\pi} \partial_\tau q \right) \tag{2.105}$$

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 [78], for $\theta \in (2\mathbb{Z}+1)\pi$ there is an '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 [79, 80]. 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

$$\begin{aligned}
&\mathcal{L} = \frac{1}{2a}\big(\Delta q(\hat{\tau}) - 2\pi n(\hat{\tau})\big)^2 + \frac{i\theta}{2\pi}\big(\Delta q(\hat{\tau}) - 2\pi n(\hat{\tau})\big) \ , \\
&S = \sum_{\hat{\tau}=0}^{L-1} \mathcal{L} \\
&\Delta q(\hat{\tau}) = q(\hat{\tau}+1) - q(\hat{\tau}) \ .
\end{aligned} \tag{2.106}$$

This system has a $\mathbb{Z}$ gauge symmetry

$$\begin{aligned}
&q(\hat{\tau}) \sim q(\hat{\tau}) + 2\pi k(\hat{\tau}) \\
&n(\hat{\tau}) \sim n(\hat{\tau}) + \Delta_\tau k(\hat{\tau}) \\
&k(\hat{\tau}) \in \mathbb{Z} \ .
\end{aligned} \tag{2.107}$$

We can replace the Lagrangian in (2.106) by

$$\mathcal{L}' = \frac{1}{2a}\big(\Delta q(\hat{\tau}) - 2\pi n(\hat{\tau})\big)^2 - i\theta n(\hat{\tau}) \tag{2.108}$$

without changing the action. Unlike $\mathcal{L}$, the new Lagrangian $\mathcal{L}'$ is not gauge invariant under (2.107).

The main point about (2.106) or (2.108) 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 (2.106). 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 (2.106) 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 an 't Hooft anomaly associated with this symmetry and to the related anomaly in the space of coupling constants of [79, 80].[10]

This anomaly is exactly as in the continuum discussion of [78]. It can be demonstrated by adding to (2.106) a classical $U(1)$ gauge field A

$$\mathcal{L} = \frac{1}{2a}\big(\Delta q(\hat{\tau}) - A(\hat{\tau}) - 2\pi n(\hat{\tau})\big)^2 + \frac{i\theta}{2\pi}\big(\Delta q(\hat{\tau}) - A(\hat{\tau}) - 2\pi n(\hat{\tau})\big) \ . \tag{2.109}$$

To see that the gauge symmetry of A is $U(1)$ rather than $\mathbb{R}$, we note that its gauge symmetry

$$\begin{aligned}
&q(\hat{\tau}) \sim q(\hat{\tau}) + \Lambda(\hat{\tau}) + 2\pi k(\hat{\tau})\\
&n(\hat{\tau}) \sim n(\hat{\tau}) + \Delta_\tau k(\hat{\tau}) - N(\hat{\tau})\\
&A(\hat{\tau}) \sim A(\hat{\tau}) + \Delta\Lambda(\hat{\tau}) + 2\pi N(\hat{\tau})\\
&k(\hat{\tau}), N(\hat{\tau}) \in \mathbb{Z}
\end{aligned} \tag{2.110}$$

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 (2.108).[11] 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 [79, 80], this also means that there is an anomaly in the 2π-periodicity in θ.

[10]Note that $e^{-\mathcal{L}'}$ with $\mathcal{L}'$ of (2.108) 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.

[11]An extreme version of this system is when the lattice has only one site, i.e., $L = 1$. In that case the action becomes

$$\frac{1}{2a}\big(A(\hat{\tau}) + 2\pi n(\hat{\tau})\big)^2 - \frac{i\theta}{2\pi}\big(A(\hat{\tau}) + 2\pi n(\hat{\tau})\big) \ . \tag{2.111}$$

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

$$\begin{aligned}
&n(\hat{\tau}) \sim n(\hat{\tau}) - N(\hat{\tau})\\
&A(\hat{\tau}) \sim A(\hat{\tau}) + 2\pi N(\hat{\tau})\\
&N(\hat{\tau}) \in \mathbb{Z} \ .
\end{aligned} \tag{2.112}$$

Again, the anomaly is manifest in (2.111).

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

$$\bar{q}(\hat{\tau}) = \begin{cases} q(\hat{\tau}) & \text{for } \hat{\tau} = 1, \cdots, L \\ q(0) + 2\pi n(0) & \text{for } \hat{\tau} = 0 \end{cases} . \tag{2.113}$$

In these variables, after dropping the bar, (2.106) becomes

$$\begin{aligned} \mathcal{L} &= \frac{1}{2a}\big(\Delta q(\hat{\tau})\big)^2 + \frac{i\theta}{2\pi}\Delta q(\hat{\tau}) \ , \\ S &= \sum_{\hat{\tau}=0}^{L-1} \mathcal{L} \ . \end{aligned} \tag{2.114}$$

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 (2.114), 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 (2.105).

2.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

$$\frac{iN}{2\pi} \int d\tau \ p\dot{q} \ . \tag{2.115}$$

(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

$$\begin{aligned} &UV = e^{\frac{2\pi i}{N}} VU, \\ &U = e^{ip}, \quad V = e^{iq} \ . \end{aligned} \tag{2.116}$$

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 (2.115) 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 [81, 79, 80] (see also [82, 83, 15, 59])[12]

$$\frac{iN}{2\pi}\int_{\tau_*}^{\tau_*+\ell} d\tau \ p\dot{q} - iNw_p(\tau_*)q(\tau_*) \ , \tag{2.117}$$

where ℓ is the period of the Euclidean time and $w_p = \frac{1}{2\pi}[p(\tau_* + \ell) - p(\tau_*)]$ is the winding number of p. Similarly, we define $w_q = \frac{1}{2\pi}[q(\tau_* + \ell) - q(\tau_*)]$ 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 (2.105), we could have added to (2.117) 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

[12] The rigorous mathematical treatment uses differential cohomology [84–87] (see [88–90] and the references therein for modern developments).

fields n_q and n_p. The action is

$$\begin{aligned}
&\frac{iN}{2\pi}\sum_{\hat{\tau}=0}^{L-1}\Big[p(\hat{\tau})\big(\Delta q(\hat{\tau})-2\pi n_q(\hat{\tau})\big)+2\pi n_p(\hat{\tau})q(\hat{\tau})\Big]\ ,\\
&\Delta q(\hat{\tau})\equiv q(\hat{\tau}+1)-q(\hat{\tau})\ .
\end{aligned}\tag{2.118}$$

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

$$\begin{aligned}
&p(\hat{\tau})\sim p(\hat{\tau})+2\pi k_p(\hat{\tau})\ ,\\
&q(\hat{\tau})\sim q(\hat{\tau})+2\pi k_q(\hat{\tau})\ ,\\
&n_p(\hat{\tau})\sim n_p(\hat{\tau})+k_p(\hat{\tau})-k_p(\hat{\tau}-1)\ ,\\
&n_q(\hat{\tau})\sim n_q(\hat{\tau})+k_q(\hat{\tau}+1)-k_q(\hat{\tau})\ .
\end{aligned}\tag{2.119}$$

Note that the Lagrangian is not gauge invariant. Even the action is not gauge invariant. But 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

$$\frac{iN}{2\pi}\sum_{\hat{\tau}=0}^{L-1}p(\hat{\tau})\Delta q(\hat{\tau})-iNn_q(0)p(0)+iNn_p(0)q(0)\ .\tag{2.120}$$

There is a residual gauge symmetry:

$$\begin{aligned}
&p(\hat{\tau})\sim p(\hat{\tau})+2\pi\ ,\\
&q(\hat{\tau})\sim q(\hat{\tau})+2\pi\ .
\end{aligned}\tag{2.121}$$

To relate the gauge fixed lattice action (2.120) to the continuum action (2.117), we define new variables $\bar{p}, \bar{q}$ on the covering space of the periodic lattice:

$$\begin{aligned} \bar{p}(\hat{\tau}) &= \begin{cases} p(\hat{\tau}) & \text{for } \hat{\tau} = 0, \cdots, L-1 \\ p(\hat{\tau}) - 2\pi n_p(0) & \text{for } \hat{\tau} = L \end{cases} , \\ \bar{q}(\hat{\tau}) &= \begin{cases} q(\hat{\tau}) & \text{for } \hat{\tau} = 1, \cdots, L \\ q(0) + 2\pi n_q(0) & \text{for } \hat{\tau} = 0 \end{cases} . \end{aligned} \tag{2.122}$$

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

$$\frac{iN}{2\pi} \sum_{\hat{\tau}=0}^{L-1} \bar{p}(\hat{\tau}) \Delta \bar{q}(\hat{\tau}) - iN w_p \bar{q}(0) \ , \tag{2.123}$$

In the continuum limit, this lattice action becomes (2.117).

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

$$\frac{2\pi i}{N} \sum_{\hat{\tau}=1}^{L} m_p(\hat{\tau}) \Delta m_q(\hat{\tau}) \ , \tag{2.124}$$

with the following gauge symmetry making the integer fields $\mathbb{Z}_N$ variables

$$\begin{aligned} m_p(\hat{\tau}) &\sim m_p(\hat{\tau}) + N k_p(\hat{\tau}) \ , \\ m_q(\hat{\tau}) &\sim m_q(\hat{\tau}) + N k_q(\hat{\tau}) \ . \end{aligned} \tag{2.125}$$

2.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.

2.B.1 2d Euclidean XY-model

Here we study the two-dimensional Euclidean XY-model on the lattice and in the continuum limit [6, 7].

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

$$\beta \sum_{\text{link}} [1 - \cos(\Delta_\mu \phi)] \ , \tag{2.126}$$

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 (2.126) by the Villain action [60]:

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 \ . \tag{2.127}$$

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 (2.126), ϕ was circle-valued rather than real-valued is related to the $\mathbb{Z}$ gauge symmetry

$$\phi \sim \phi + 2\pi k \ , \qquad n_\mu \sim n_\mu + \Delta_\mu k \ , \tag{2.128}$$

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

$$\mathcal{N} \equiv \Delta_x n_y - \Delta_y n_x \ . \tag{2.129}$$

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 (2.127) a term like

$$\kappa \sum_{\text{plaquette}} \mathcal{N}^2 \tag{2.130}$$

with positive κ. For $\kappa \to \infty$ the vortices are completely suppressed [62]. 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 [61][13]

$$S = \frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquette}} \tilde{\phi} \mathcal{N} \ , \tag{2.131}$$

where the Lagrange multiplier $\tilde{\phi}$ is a real-valued field on the plaquettes (or dual sites). It has a $\mathbb{Z}$ gauge symmetry

$$\tilde{\phi} \sim \tilde{\phi} + 2\pi \tilde{k} \ , \tag{2.132}$$

with $\tilde{k}$ is an integer-valued gauge parameter on the plaquettes.

[13] Related ideas were used in various places, including [91].

Note that the action (2.131) 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 (2.131) is the starting point of our discussion. We refer to it as the modified Villain action of the XY-model.[14]

We can restore the vortices by perturbing the modified Villain action (2.131) as

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 + i\sum_{\text{plaquette}}\tilde{\phi}\mathcal{N} - \lambda\sum_{\text{plaquette}}\cos(\tilde{\phi}) \ . \tag{2.133}$$

(For simplicity of the presentation, we take $\lambda \geq 0$.) Note that the action is still invariant under the gauge symmetries (2.128) and (2.132). Integrating out $\tilde{\phi}$ gives

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi - 2\pi n_\mu)^2 - \sum_{\text{plaquette}}\log I_{|\mathcal{N}|}(\lambda) \ , \tag{2.134}$$

where $I_k(z)$ is the modified Bessel function of the first kind. Let us compare this action with (2.130). For small $\lambda \ll 1$, we have

$$-\log I_k(\lambda) \approx \log\left[k!\left(\frac{2}{\lambda}\right)^k\right] + O(\lambda^2) \ . \tag{2.135}$$

In this case, vortices with $|\mathcal{N}| > 1$ are suppressed. For $|\mathcal{N}| = 0, 1$ we identify

$$\kappa \approx \log\frac{2}{\lambda} \gg 1 \ . \tag{2.136}$$

In the other limit $\lambda \gg 1$, we have

$$-\log I_k(\lambda) \sim \frac{1}{2\lambda}k^2 + O(\lambda^{-2}) \tag{2.137}$$

[14] 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 (2.127) and (2.131) is real-valued, i.e., noncompact, the gauge symmetry (2.128) effectively compactifies the range of ϕ. The effect of the term with $\mathcal{N}$ in (2.131) is to suppress the vortices rather than to de-compactify the target space. We will discuss it further below.

where we ignored some k-independent terms that depend on λ. In this case, we can identify

$$\kappa \approx \frac{1}{2\lambda} \ll 1 \ . \tag{2.138}$$

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 (2.126) or (2.127). We added another coupling to this model (2.130). Equivalently, we can write the model as (2.133) and then the usually studied model (2.127) is obtained in the limit $\lambda \to \infty$. On the other hand, when $\lambda = 0$, this reduces to our modified Villain action (2.131) of the XY-model.

Below we will see that the modified Villain action (2.131), 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, (2.126), (2.127), and (2.131) have a *momentum symmetry*, which acts as

$$\phi \to \phi + c^m \ , \tag{2.139}$$

where c^m is a real position-independent constant. Due to the zero mode of the gauge symmetry (2.128), the $2\pi\mathbb{Z}$ part of this symmetry is gauged. So the momentum symmetry is $U(1)$ rather than $\mathbb{R}$.

From (2.127) and (2.131) we find the Noether current of momentum symmetry[15]

$$J^m_\mu = -i\beta(\Delta_\mu\phi - 2\pi n_\mu) \ , \tag{2.140}$$

[15] The factor of i in the Euclidean signature is such that the corresponding charge is real.

which is conserved because of the equation of motion of ϕ. The momentum charge is[16]

$$Q^m(\tilde{\mathcal{C}}) = \sum_{\text{dual link} \in \tilde{\mathcal{C}}} \epsilon_{\mu\nu} J^m_\nu \ , \tag{2.141}$$

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 (2.131) (but not (2.126) or (2.127)) also has a *winding symmetry*, which acts as

$$\tilde{\phi} \to \tilde{\phi} + c^w \ , \tag{2.142}$$

where c^w is a real constant. Due to the zero mode of the gauge symmetry (2.132), the $2\pi\mathbb{Z}$ part of this symmetry is gauged. So the winding symmetry is also $U(1)$.

The Noether current of the winding symmetry is[17]

$$J^w_\mu = \frac{\epsilon_{\mu\nu}}{2\pi}(\Delta_\nu \phi - 2\pi n_\nu) \ , \tag{2.143}$$

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

$$Q^w(\mathcal{C}) = \sum_{\text{link} \in \mathcal{C}} \epsilon_{\mu\nu} J^w_\nu = -\sum_{\text{link} \in \mathcal{C}} n_\mu \ , \tag{2.144}$$

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.

[16] Here, $\epsilon_{xy} = -\epsilon_{yx} = 1$ and $\epsilon_{xx} = \epsilon_{yy} = 0$.

[17] From the action (2.131), 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.

Both the momentum symmetry (2.139) and the winding symmetry (2.142) act locally on the fields and they both leave the action (2.131) invariant. However, the Lagrangian density in (2.131) 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 (2.131) as

$$\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi-2\pi n_\mu)^2+i\sum_{\text{plaquette}}(n_x\Delta_y\tilde{\phi}-n_y\Delta_x\tilde{\phi})\ . \tag{2.145}$$

In this form both the momentum symmetry (2.139) and the winding symmetry (2.142) act locally and leave the Lagrangian density invariant. How is this compatible with the anomaly? The point is that unlike (2.131), the Lagrangian density in (2.145) 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 (2.131) to background gauge 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

$$\begin{aligned}&\frac{\beta}{2}\sum_{\text{link}}(\Delta_\mu\phi-A_\mu-2\pi n_\mu)^2+i\sum_{\text{plaquette}}\tilde{\phi}(\Delta_x n_y-\Delta_y n_x+N)\\&-\frac{i}{2\pi}\sum_{\text{link}}\epsilon_{\mu\nu}\tilde{A}_\mu(\Delta_\nu\phi-A_\nu-2\pi n_\nu)+i\sum_{\text{site}}\tilde{N}\phi\ ,\end{aligned} \tag{2.146}$$

with the gauge symmetry

$$\begin{aligned}&\phi\sim\phi+\alpha+2\pi k\ , && \tilde{\phi}\sim\tilde{\phi}+\tilde{\alpha}+2\pi\tilde{k}\ ,\\&A_\mu\sim A_\mu+\Delta_\mu\alpha+2\pi K_\mu\ , && \tilde{A}_\mu\sim\tilde{A}_\mu+\Delta_\mu\tilde{\alpha}+2\pi\tilde{K}_\mu\ ,\\&n_\mu\sim n_\mu+\Delta_\mu k-K_\mu\ , && \tilde{N}\sim\tilde{N}+\Delta_x\tilde{K}_y-\Delta_y\tilde{K}_x\ ,\\&N\sim N+\Delta_x K_y-\Delta_y K_x\ .\end{aligned} \tag{2.147}$$

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

$$-\frac{i}{2\pi}\sum_{\text{plaquette}}\tilde{\alpha}(\Delta_x A_y - \Delta_y A_x - 2\pi N) + i\sum_{\text{plaquette}}(\tilde{K}_x A_y - \tilde{K}_y A_x) + i\sum_{\text{site}}(\tilde{N} + \Delta_x \tilde{K}_y - \Delta_y \tilde{K}_x)\alpha \ . \tag{2.148}$$

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 $-\frac{i}{2\pi}\int dxdy\ \tilde{\alpha}(\partial_x A_y - \partial_y A_x)$.

As a special case of this anomaly, consider the $\mathbb{Z}_N$ subgroup of the $U(1)\times U(1)$ symmetry, which is generated by $\phi \to \phi + 2\pi/N,\ \tilde{\phi} \to \tilde{\phi} + 2\pi/N$. The anomaly in this symmetry is visible in (2.148). It agrees with the general classification of $\mathbb{Z}_N$ anomalies in 1+1d bosonic systems by $H^3(\mathbb{Z}_N, U(1)) = \mathbb{Z}_N$.

T-Duality

Here we will demonstrate the self-duality of the modifield Villain lattice model (2.131). We start with the presentation (2.145). Using the Poisson resummation formula (2.6) for n_μ and ignoring the overall factor, we can dualize the above action to

$$\begin{aligned}
&\frac{1}{2(2\pi)^2\beta}\sum_{\text{dual link}}(\Delta_\mu\tilde{\phi} - 2\pi\tilde{n}_\mu)^2 + i\sum_{\text{site}}\phi\tilde{\mathcal{N}} \ , \\
&\tilde{\mathcal{N}} \equiv \Delta_x\tilde{n}_y - \Delta_y\tilde{n}_x \ ,
\end{aligned} \tag{2.149}$$

where $\tilde{n}_\mu$ is an integer-valued field on the dual links. The gauge symmetry of the original theory acts as

$$\tilde{\phi} \sim \tilde{\phi} + 2\pi\tilde{k} \ , \qquad \tilde{n}_\mu \sim \tilde{n}_\mu + \Delta_\mu\tilde{k}, \qquad \phi \sim \phi + 2\pi k \ . \tag{2.150}$$

$\tilde{n}_\mu$ 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 (2.131) is a self-dual lattice model with $\beta \leftrightarrow \frac{1}{(2\pi)^2\beta}$. Moreover, the momentum and winding currents, (2.140) and (2.143), in the dual picture are

$$J^m_\mu = \frac{\epsilon_{\mu\nu}}{2\pi}(\Delta_\nu\tilde{\phi} - 2\pi\tilde{n}_\nu) \ , \qquad J^w_\mu = -\frac{i}{(2\pi)^2\beta}(\Delta_\mu\tilde{\phi} - 2\pi\tilde{n}_\mu) \ . \tag{2.151}$$

We emphasize that the lattice model (2.131) 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 [6, 7]. 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 (2.127). Repeating this in the dual action (2.149), we find

$$\begin{aligned} &\frac{1}{2\beta}\sum_{\text{dual link}} \tilde{n}^2_\mu + i\sum_{\text{site}} \phi\tilde{\mathcal{N}} \ , \\ &\tilde{\mathcal{N}} \equiv \Delta_x\tilde{n}_y - \Delta_y\tilde{n}_x \ . \end{aligned} \tag{2.152}$$

Locally, the Lagrange multiplier ϕ determines $\tilde{n}_\mu = \Delta_\mu q$ with an integer q.[18] We end up with

$$\frac{1}{2\beta}\sum_{\text{dual link}} (\Delta_\mu q)^2 \ , \tag{2.153}$$

which is the dual theory of [6, 7].

[18]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 (2.141) constructed out of the momentum current (2.151) and it is nontrivial only when q is not periodic.

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 (2.120), 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, $x_\mu \sim x_\mu + L_\mu$). 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

$$\phi \sim \phi + 2\pi\mathbb{Z} \ . \tag{2.154}$$

Let us define a new field $\bar{\phi}$ such that

$$\bar{\phi}(0,0) = \phi(0,0) \ , \qquad \Delta_\mu \bar{\phi} = \Delta_\mu \phi - 2\pi n_\mu \ . \tag{2.155}$$

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

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \bar{\phi})^2 \ . \tag{2.156}$$

Although ϕ and n_μ are single-valued fields, $\bar{\phi}$ can wind around nontrivial cycles:

$$\begin{aligned} \bar{\phi}(\hat{x} + L_x, \hat{y}) &= \bar{\phi}(\hat{x}, \hat{y}) - 2\pi \bar{n}_x \ , \\ \bar{\phi}(\hat{x}, \hat{y} + L_y) &= \bar{\phi}(\hat{x}, \hat{y}) - 2\pi \bar{n}_y \ . \end{aligned} \tag{2.157}$$

So, in the path integral, we should sum over nontrivial winding sectors of $\bar{\phi}$.[19]

[19] Note that the variables $\bar{\phi}$ are noncompact and we can rescale them to make the action (2.156) independent of β. Then, the compactness and the β dependence enter only through the twisted boundary conditions (2.157). 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 (14). This reasoning is valid when we consider the model with fixed twisted boundary conditions like (2.157). However, in our case, we sum over this twist. And

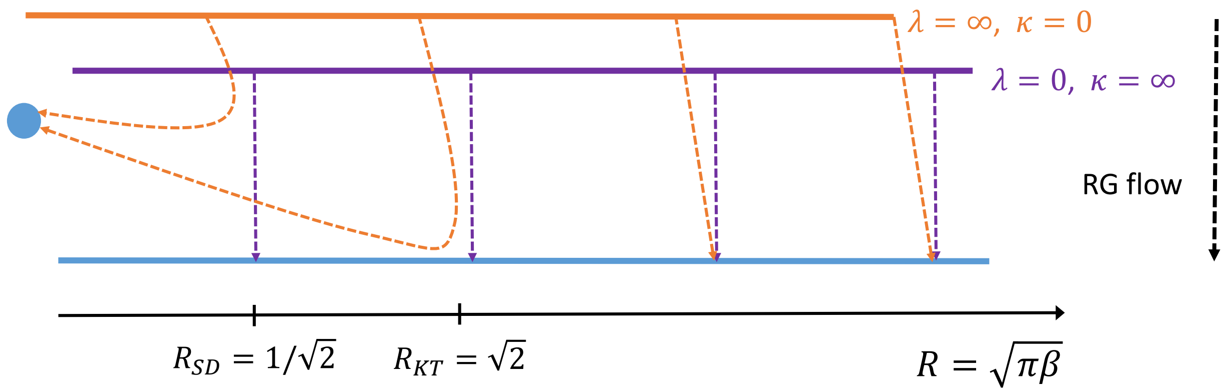


Figure 2.1: The space of coupling constants of the 2d Euclidean XY-model. The orange line corresponds to the theories based on (2.126) or (2.127), while the purple line corresponds to the modified theory (2.131). Each of them depends on the parameter $R = \sqrt{\pi\beta}$. The parameter λ (equivalently, κ) interpolates between these two lines. The theories of the purple line (2.131) 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 (2.126) or (2.127) 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.

In the continuum limit $a \to 0$ such that $\ell_\mu \equiv aL_\mu$ is fixed, the action (2.156) becomes

$$\frac{\beta}{2} \int dxdy\ (\partial_\mu \phi)^2\ , \tag{2.158}$$

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 19.

this affects the set of local operators in the theory. In particular, as in (2.159), their dimensions depend on the value of $\beta = R^2/\pi$.

Kosterlitz-Thouless transition

In order to compare with the standard conformal field theory literature (e.g., [8,9]), 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

$$(h, \bar{h}) = \left(\frac{1}{2}\left(\frac{n_m}{2R} + n_w R\right)^2, \frac{1}{2}\left(\frac{n_m}{2R} - n_w R\right)^2\right) \ , \tag{2.159}$$

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 2.1.

Unlike the modified Villain model (2.131), the original XY-model (2.126) and its Villain counterpart (2.127) 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 2.1.

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 2.1. Only for $\lambda = 0$ is the flow different (as the purple line in Figure 2.1). 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 2.1, except that the Kosterlitz-Thouless point moves to $R = \frac{\sqrt{2}}{W}$.

2.B.2 2d Euclidean $\mathbb{Z}_N$ clock model

Lattice models

The $\mathbb{Z}_N$ clock model [6, 92–96] can be obtained by restricting the phase variables $e^{i\phi}$ in the XY-model (2.126) to $\mathbb{Z}_N$ variables $e^{2\pi i m/N}$. More generally, this model has $\lfloor N/2 \rfloor$ nearest-neighbor couplings

$$\sum_{M=1}^{\lfloor N/2 \rfloor} J_M \sum_{\text{link}} \left[1 - \cos\left(\frac{2\pi M}{N} \Delta_\mu m \right) \right] \ . \tag{2.160}$$

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:

$$\frac{\beta}{2} \left(\frac{2\pi}{N} \right)^2 \sum_{\text{link}} (\Delta_\mu m - N n_\mu)^2 \ . \tag{2.161}$$

The integer fields m, n_μ are subject to a gauge symmetry with integer gauge parameter k

$$\begin{aligned} m &\sim m + Nk \ , \\ n_\mu &\sim n_\mu + \Delta_\mu k \ . \end{aligned} \tag{2.162}$$

This model (2.161) can be embedded in the XY-model of Appendix 2.B.1. In general, we can deform the action (2.131) to

$$\frac{\beta}{2} \sum_{\text{link}} (\Delta_\mu \phi - 2\pi n_\mu)^2 + i \sum_{\text{plaquette}} \tilde{\phi} \mathcal{N} - \lambda \sum_{\text{plaquette}} \cos(W\tilde{\phi}) - \tilde{\lambda} \sum_{\text{site}} \cos(N\phi) \ , \tag{2.163}$$

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 (2.161).

Kramers-Wannier duality

It is straightforward to repeat the analysis in Appendix 2.B.1 and to dualize (2.163) to

$$\begin{aligned}
&\frac{1}{2(2\pi)^2\beta} \sum_{\text{dual link}} (\Delta_\mu \tilde{\phi} - 2\pi \tilde{n}_\mu)^2 + i \sum_{\text{site}} \phi \tilde{\mathcal{N}} - \lambda \sum_{\text{plaquette}} \cos(W\tilde{\phi}) - \tilde{\lambda} \sum_{\text{site}} \cos(N\phi) \ , \\
&\tilde{\mathcal{N}} \equiv \Delta_x \tilde{n}_y - \Delta_y \tilde{n}_x \ ,
\end{aligned} \tag{2.164}$$

where $\tilde{n}_\mu$ is an integer-valued field on the dual links. The gauge symmetry of the theory is

$$\tilde{\phi} \sim \tilde{\phi} + 2\pi \tilde{k} \ , \qquad \tilde{n}_\mu \sim \tilde{n}_\mu + \Delta_\mu \tilde{k}, \qquad \phi \sim \phi + 2\pi k \ . \tag{2.165}$$

We conclude that the action (2.163) 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 $\tilde{\lambda}, \lambda \to \infty$, (2.163) is dualized to

$$\frac{1}{2\beta} \sum_{\text{dual link}} \tilde{n}_\mu^2 + \frac{2\pi i}{N} \sum_{\text{site}} m(\Delta_x \tilde{n}_y - \Delta_y \tilde{n}_x) \tag{2.166}$$

with the gauge symmetry

$$m \sim m + Nk \tag{2.167}$$

with integer k. We can find it either by substituting $\phi = \frac{2\pi m}{N}$, $\tilde{\phi} = 0$ in (2.164), or by directly dualizing (2.161).

We see that unlike the modified Villain action for the XY-model (2.131), this theory is not selfdual. Comparing with the general case (2.164), 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 [6, 92–96]?

In order to answer this question we first add integer-valued fields $\tilde{m}$ and $\hat{n}_\mu$ to the action (2.166)

$$\frac{1}{2\beta}\sum_{\text{dual link}}(\Delta_\mu\tilde{m}-N\hat{n}_\mu-\tilde{n}_\mu)^2+\frac{2\pi i}{N}\sum_{\text{site}}m(\Delta_x\tilde{n}_y-\Delta_y\tilde{n}_x)\ . \tag{2.168}$$

In addition to the gauge symmetry (2.167), this action has the gauge symmetry

$$\begin{aligned}
&\tilde{m}\sim\tilde{m}+\tilde{k}\ ,\\
&\hat{n}_\mu\sim\hat{n}_\mu-\hat{q}_\mu\ ,\\
&\tilde{n}_\mu\sim\tilde{n}_\mu+\Delta_\mu\tilde{k}+N\hat{q}_\mu\ .
\end{aligned} \tag{2.169}$$

Here $\tilde{k}$ is an integer zero-form gauge parameter and $\hat{q}_\mu$ is an integer one-form gauge parameter. This new action (2.168) is equivalent to (2.166), as can be seen by completely gauge fixing (2.169) by setting $\tilde{m}=\hat{n}_\mu=0$.

Now, we can interpret (2.168) as follows. Locally, the Lagrange multiplier m sets $\tilde{n}_\mu$ to a pure gauge and we can set it to zero. Then, (2.168) is the same as the Villain form of the $\mathbb{Z}_N$ action (2.161) 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 (2.168) 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, (2.168) 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 [14, 15]. The latter is described by the second term in (2.168) and will be further discussed in Appendix 2.C.2.

We conclude that the T-duality of the underlying XY-model (2.131) leads to the Kramers-Wannier duality of the clock-model (2.161). 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 (2.163).

As in the discussion around Figure 2.1 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 2.1. 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

$$\begin{aligned} \frac{\sqrt{2}}{W} \leq R \leq \frac{N}{\sqrt{8}} \\ \frac{2}{\pi W^2} \leq \beta \leq \frac{N^2}{8\pi} \end{aligned} \tag{2.170}$$

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 (2.163) 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

$$\begin{aligned} \sqrt{2} \leq R \leq \frac{N}{\sqrt{8}} \\ \frac{2}{\pi} \leq \beta \leq \frac{N^2}{8\pi} \end{aligned} \tag{2.171}$$

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 2.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 (2.163). For other actions, the gapless phase could be different or even absent. See the discussion in [94, 95, 97].

2.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 [98] for a review on these models). The modified Villain version of the p-form $U(1)$ gauge theory in general dimensions has been analyzed in [61]. The models in Appendix 2.A, correspond to $d = 1$ and $p = 0$ and perhaps do not deserve to be called gauge theories. The models in Appendix 2.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.

2.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

$$\beta \sum_{(p+1)\text{-cell}} [1 - \cos(\Delta a^{(p)})] \ , \tag{2.172}$$

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

$$e^{ia^{(p)}} \sim e^{ia^{(p)}+i\Delta\alpha^{(p-1)}} , \tag{2.173}$$

where $\alpha^{(p-1)}$ is circle-valued. At large β, the action can be approximated by the Villain action [99–101]

$$\frac{\beta}{2}\sum_{(p+1)\text{-cell}} (\Delta a^{(p)} - 2\pi n^{(p+1)})^2 , \tag{2.174}$$

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

$$\begin{aligned} &a^{(p)} \sim a^{(p)} + \Delta\alpha^{(p-1)} + 2\pi k^{(p)} , \\ &n^{(p+1)} \sim n^{(p+1)} + \Delta k^{(p)} . \end{aligned} \tag{2.175}$$

For $p \leq d-2$, nonzero $\Delta n^{(p+1)}$ corresponds to monopoles or vortices. They can be suppressed by modifying (2.174) to [61]

$$\frac{\beta}{2}\sum_{(p+1)\text{-cell}} (\Delta a^{(p)} - 2\pi n^{(p+1)})^2 + i \sum_{(p+2)\text{-cell}} \tilde{a}^{(d-p-2)}\Delta n^{(p+1)} , \tag{2.176}$$

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 (2.176) as the modified Villain action of the $U(1)$ p-form gauge theory. In addition to (2.175), this theory also has a gauge symmetry

$$\tilde{a}^{(d-p-2)} \sim \tilde{a}^{(d-p-2)} + \Delta\tilde{\alpha}^{(d-p-3)} + 2\pi\tilde{k}^{(d-p-2)} , \tag{2.177}$$

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 (2.176). Instead, in this case we can add another term[20]

$$\frac{\beta}{2}\sum_{d\text{-cell}}(\Delta a^{(d-1)} - 2\pi n^{(d)})^2 + i\theta \sum_{d\text{-cell}} n^{(d)} \ . \tag{2.178}$$

This is a $U(1)$ gauge theory of a $(d-1)$-form gauge field with a θ-parameter. (Compare with (2.106) and (2.108), 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 (2.6), we can dualize the modified Villain action (2.176) of a p-form gauge theory to the modified Villain action of a $(d-p-2)$-form gauge theory

$$\frac{1}{2(2\pi)^2\beta}\sum_{(p+1)\text{-cell}}(\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)})^2 + i(-1)^{d-p}\sum_{(p+1)\text{-cell}}\tilde{n}^{(d-p-1)}\Delta a^{(p)} \ , \tag{2.179}$$

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

$$\begin{aligned} \tilde{a}^{(d-p-2)} &\sim \tilde{a}^{(d-p-2)} + \Delta\tilde{\alpha}^{(d-p-3)} + 2\pi\tilde{k}^{(d-p-2)} \ , \\ \tilde{n}^{(d-p-1)} &\sim \tilde{n}^{(d-p-1)} + \Delta\tilde{k}^{(d-p-2)} \ . \end{aligned} \tag{2.180}$$

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 (2.176) is self-dual with $\beta \leftrightarrow \frac{1}{(2\pi)^2\beta}$.

[20] See, for example, [102, 61, 103–105] for discussions on the θ-angle in the Villain version of the lattice $U(1)$ gauge theory.

Global symmetries

In all the three models, (2.172), (2.174), and (2.176), there is a p-form *electric symmetry* [28], which acts on the fields as

$$a^{(p)} \to a^{(p)} + \lambda^{(p)} , \tag{2.181}$$

where $\lambda^{(p)}$ is a real-valued, flat p-form field. Due to the gauge symmetry (2.175), the electric symmetry is $U(1)$ rather than $\mathbb{R}$. In (2.174) and (2.176), the Noether current of electric symmetry is[21]

$$J_e^{(p+1)} = i\beta(\Delta a^{(p)} - 2\pi n^{(p+1)}) = \frac{(-1)^{d-p}}{2\pi} \star (\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)}) , \tag{2.182}$$

which is conserved because of the equation of motion of $a^{(p)}$. The electric charge is

$$Q_e(\tilde{\mathcal{M}}^{(d-p-1)}) = \sum_{\text{dual } (d-p-1)\text{-cell} \in \tilde{\mathcal{M}}^{(d-p-1)}} \star J_e^{(p+1)} , \tag{2.183}$$

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

$$W_e(\mathcal{M}^{(p)}) = \exp\left[i \sum_{p\text{-cell} \in \mathcal{M}^{(p)}} a^{(p)} \right] , \tag{2.184}$$

where $\mathcal{M}^{(p)}$ is a dimension-p submanifold along the p-cells of the lattice.

The theory (2.176) (but not (2.172) or (2.174)) also has a $(d-p-2)$-form *magnetic symmetry* [28], which acts on the fields as

$$\tilde{a}^{(d-p-2)} \to \tilde{a}^{(d-p-2)} + \tilde{\lambda}^{(d-p-2)} , \tag{2.185}$$

[21] The Hodge dual $\star A^{(p)}$ is a $(d-p)$-form field on the dual $(d-p)$-cells of the lattice.

where $\tilde{\lambda}^{(d-p-2)}$ is a real-valued, flat $(d-p-2)$-form. Due to the gauge symmetry (2.177), the magnetic symmetry is $U(1)$. The Noether current of magnetic symmetry is[22]

$$J_m^{(d-p-1)} = -\frac{i}{(2\pi)^2\beta} \star\star(\Delta\tilde{a}^{(d-p-2)} - 2\pi\tilde{n}^{(d-p-1)}) = \frac{(-1)^{d-p}}{2\pi} \star(\Delta a^{(p)} - 2\pi n^{(p+1)}) \ , \tag{2.186}$$

which is conserved because of the equation of motion of $\tilde{a}^{(d-p-2)}$. The magnetic charge is

$$Q_m(\mathcal{M}^{(p+1)}) = \sum_{(p+1)\text{-cell}\in\mathcal{M}^{(p+1)}} \star J_m^{(d-p-1)} \ , \tag{2.187}$$

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

$$W_m(\tilde{\mathcal{M}}^{(d-p-2)}) = \exp\left[i \sum_{\text{dual }(d-p-2)\text{-cell}\in\tilde{\mathcal{M}}^{(d-p-2)}} \tilde{a}^{(d-p-2)} \right] \ , \tag{2.188}$$

where $\tilde{\mathcal{M}}^{(d-p-2)}$ is a codimension-$(p+2)$ submanifold along the dual $(d-p-2)$-cells of the lattice.

Long-distance limit

In the continuum limit, the modified Villain model (2.176) becomes a gapless continuum p-form gauge theory

$$\frac{1}{2g^2}\int d^dx\ (da^{(p)})^2 \ . \tag{2.189}$$

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.[23]

[22] Recall that $\star\star A^{(p)} = (-1)^{p(d-p)}A^{(p)}$.

[23] 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

An important question is whether the lattice gauge theory (2.172), or equivalently its Villain version (2.174), flow at long distances to the same gapless theory (2.189). 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 [68].

We conclude that for $p = d - 2$, where the standard $U(1)$ lattice gauge theory is gapped, the modification of the lattice gauge theory (2.176) keeps it massless.

2.C.2 $\mathbb{Z}_N$ gauge theory

We now describe a d-dimensional Villain $\mathbb{Z}_N$ p-form gauge theory [92, 106]. 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

$$\frac{\beta(2\pi)^2}{2N^2} \sum_{(p+1)\text{-cell}} (\Delta m^{(p)} - N n^{(p+1)})^2 \ , \tag{2.190}$$

with the integer gauge symmetry

$$\begin{aligned} m^{(p)} &\sim m^{(p)} + \Delta \ell^{(p-1)} + N k^{(p)} \ , \\ n^{(p+1)} &\sim n^{(p+1)} + \Delta k^{(p)} \ . \end{aligned} \tag{2.191}$$

not to the other θ-parameters. For example, see [61, 103] for a discussion on the θ-parameter in the modified Villain version of the ordinary 3+1d $U(1)$ gauge theory.

The theory has an electric $\mathbb{Z}_N$ p-form global symmetry [28], which shifts $m^{(p)}$ by a flat integer p-form field.

In the limit $\beta \to \infty$, the field strength obeys $\Delta m = 0 \bmod N$ [107, 108], and we can replace the action by

$$\frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ , \tag{2.192}$$

where $\tilde{n}^{(d-p-1)}$ is an integer-valued field with the integer gauge symmetry

$$\tilde{n}^{(d-p-1)} \sim \tilde{n}^{(d-p-1)} + \Delta \tilde{k}^{(d-p-2)} + N \tilde{q}^{(d-p-1)} \ . \tag{2.193}$$

This describes a *topological* $\mathbb{Z}_N$ *lattice gauge theory* [14, 15]. The action (2.192) is similar to the one in [15] 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 2.B.2, we can dualize the $\mathbb{Z}_N$ p-form gauge theory (2.190) by dualizing the integer field $n^{(p+1)}$ to an integer field $\tilde{n}^{(d-p-1)}$:

$$\frac{1}{2\beta} \sum_{\text{dual } (d-p-1)\text{-cell}} (\tilde{n}^{(d-p-1)})^2 + \frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ . \tag{2.194}$$

For $p \leq d-1$, we can introduce new gauge symmetries together with Stueckelberg fields, and write the action as

$$\frac{1}{2\beta} \sum_{\text{dual } (d-p-1)\text{-cell}} (\Delta \tilde{m}^{(d-p-2)} - N \hat{n}^{(d-p-1)} - \tilde{n}^{(d-p-1)})^2 + \frac{2\pi i}{N} \sum_{p\text{-cell}} m^{(p)} \Delta \tilde{n}^{(d-p-1)} \ . \tag{2.195}$$

with the integer gauge symmetry

$$\begin{aligned}
&\tilde{m}^{(d-p-2)} \sim \tilde{m}^{(d-p-2)} + \Delta\tilde{\ell}^{(d-p-3)} + \tilde{k}^{(d-p-2)} \ , \\
&\hat{n}^{(d-p-1)} \sim \hat{n}^{(d-p-1)} - \tilde{q}^{(d-p-1)} \ , \\
&\tilde{n}^{(d-p-1)} \sim \tilde{n}^{(d-p-1)} + \Delta\tilde{k}^{(d-p-2)} + N\tilde{q}^{(d-p-1)} \ , \\
&m^{(p)} \sim m^{(p)} + \Delta\ell^{(p-1)} + Nk^{(p)} \ .
\end{aligned} \tag{2.196}$$

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 2.B.2. The duality of the $d=3$ and $p=1$ system is the famous duality of the 3d clock model [109, 98, 110] and for $d=4$ and $p=1$ it is the famous self-duality of [109, 98, 111, 110, 92, 106].

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 (2.192) 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

$$\begin{aligned}
&\frac{iN}{2\pi}\sum_{p\text{-cell}} a^{(p)}\left(\Delta\tilde{b}^{(d-p-1)} - 2\pi\tilde{m}^{(d-p)}\right) + i(-1)^p N \sum_{(p+1)\text{-cell}} n^{(p+1)}\tilde{b}^{(d-p-1)} \\
&\quad - i(-1)^p \sum_{(p+1)\text{-cell}} n^{(p+1)}\Delta\tilde{\phi}^{(d-p-2)} + i\sum_{p\text{-cell}} \Delta\phi^{(p-1)}\tilde{m}^{(d-p)} \ .
\end{aligned} \tag{2.197}$$

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

$$a^{(p)} - \frac{1}{N}\Delta\phi^{(p-1)} = \frac{2\pi}{N}m^{(p)} , \qquad \tilde{b}^{(d-p-1)} - \frac{1}{N}\Delta\tilde{\phi}^{(d-p-2)} = \frac{2\pi}{N}\tilde{n}^{(d-p-1)} , \tag{2.198}$$

where $m^{(p)}$ and $\tilde{n}^{(d-p-1)}$ are integer-valued fields. Substituting them into (2.197), we recover the action (2.192).

The action (2.197) has the gauge symmetry

$$\begin{aligned}
&a^{(p)} \sim a^{(p)} + \Delta\alpha^{(p-1)} + 2\pi k^{(p)} ,\\
&\tilde{b}^{(d-p-1)} \sim \tilde{b}^{(d-p-1)} + \Delta\tilde{\beta}^{(d-p-2)} + 2\pi\tilde{q}^{(d-p-1)} ,\\
&n^{(p+1)} \sim n^{(p+1)} + \Delta k^{(p)} ,\\
&\tilde{m}^{(d-p)} \sim \tilde{m}^{(d-p)} + \Delta\tilde{q}^{(d-p-1)} ,\\
&\phi^{(p-1)} \sim \phi^{(p-1)} + \Delta\gamma^{p-2} + N\alpha^{(p-1)} + 2\pi k_{\phi}^{(p-1)} ,\\
&\tilde{\phi}^{(d-p-2)} \sim \tilde{\phi}^{(d-p-2)} + \Delta\tilde{\gamma}^{(d-p-3)} + N\tilde{\beta}^{(d-p-2)} + 2\pi\tilde{q}_{\tilde{\phi}}^{(d-p-2)} ,
\end{aligned} \tag{2.199}$$

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

$$\begin{aligned}
&\frac{i}{2\pi}\sum_{p\text{-cell}}(\Delta\phi^{(p-1)} - Na^{(p)} - 2\pi n^{(p)})\tilde{F}^{(d-p)} + i(-1)^p \sum_{(p+1)\text{-cell}} (\Delta n^{(p)} + Nn^{(p+1)})\tilde{b}^{(d-p-1)}\\
&\quad - i(-1)^p \sum_{(p+1)\text{-cell}} n^{(p+1)}\Delta\tilde{\phi}^{(d-p-2)} .
\end{aligned} \tag{2.200}$$

These fields have the same gauge symmetries as in (2.199). In addition, the gauge symmetries also act on $n^{(p)}$

$$n^{(p)} \sim n^{(p)} + \Delta k_\phi^{(p-1)} - Nk^{(p)} \ . \tag{2.201}$$

We can interpret the action (2.200) 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

$$\begin{aligned} &\Delta \bar{a}^{(p)} = \Delta a^{(p)} - 2\pi n^{(p+1)} \ , \\ &\Delta \bar{\tilde{b}}^{(d-p-1)} = \Delta \tilde{b}^{(d-p-1)} - 2\pi \tilde{m}^{(d-p)} \ , \end{aligned} \tag{2.202}$$

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

$$\frac{iN}{2\pi} \sum_{p\text{-cell}} \bar{a}^{(p)} \Delta \bar{\tilde{b}}^{(d-p-1)} + i(-1)^p N \sum_{(p+1)\text{-cell}} n^{(p+1)} \bar{\tilde{b}}^{(d-p-1)} \ , \tag{2.203}$$

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 (2.203) reduces to the quantum mechanics action (2.120).

The real BF-action is closely related to the continuum field theory limit. In this gauge choice, the continuum limit is

$$\frac{iN}{2\pi}\int a^{(p)} d\tilde{b}^{(d-p-1)} , \tag{2.204}$$

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 (2.117), these terms are actually essential in order to make (2.204) 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 [112, 113, 15].

Relation to the toric code

We now review the well-known fact that the low-energy limit of the $\mathbb{Z}_N$ toric code [13] is described by the topological $\mathbb{Z}_N$ lattice gauge theory [14, 15], 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:

$$H_{\text{toric}} = -\beta_1 \sum_{\text{site}} G - \beta_2 \sum_{\text{plaq}} L + c.c. \ , \tag{2.205}$$

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 [57]). 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 [28], 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 $\epsilon^{ij} \Delta_i \tilde{n}_j = 0 \bmod N$. On each dual τ-link (or equivalently each xy-plaquette), we introduce an integer field $\tilde{n}_\tau$ to impose $L = 1$, or equivalently $\epsilon^{ij} \Delta_i m_j = 0 \bmod N$. In terms of these integer fields, the Euclidean action of the system is precisely the topological $\mathbb{Z}_N$ lattice gauge theory (2.192) with $d = 3$ and $p = 1$.

Chapter 3

Fractons on Graphs and Complexity: $U(1)$ Laplacian Models

3.1 Introduction

The past decades have seen an explosion of various exotic lattice models including gapped fracton models [19–21] (see [17, 18, 32] for reviews)[1] and gapless models [22] with subsystem symmetries. These models have various peculiar properties:

1. Exotic global symmetries such as (planar or fractal) subsystem global symmetries [22, 21], multipole global symmetries [23, 116, 56, 117], etc. (See [34, 29] for recent reviews on generalized global symmetries [28].)

2. In gapped fracton models, the logarithm of the ground state degeneracy (GSD) grows, typically subextensively, with the linear size of the system [24].

3. Massive particle-like excitations that have restricted mobility—a particle can be completely immobile, a.k.a. *fracton*, or can move only along a line, a.k.a. *lineon*, etc.

[1] The word "fracton" was also used in other different contexts, for example [114, 115].

These peculiarities do not fit into the framework of conventional continuum quantum field theory. In particular, the gapped fracton models do not admit a topological quantum field theory (TQFT) description at low energies. Instead, one has to go beyond the standard relativistic continuum field theory to describe them (see, for example, [33] for a recent review, and references therein).

While there is a plethora of exotic lattice models, in most cases, the spatial lattice is assumed to be cubic with manifest translation invariance in the three spatial directions. There is usually no obvious way to define them on a triangulation of an arbitrary spatial manifold. See [51, 50, 52] for fracton models on more general manifolds. In some examples, a *foliation* of spatial manifold is essential [40–49]. In contrast, standard lattice models, such as the Ising model, or the toric code, can be defined on an arbitrary triangulation of the spatial manifold.

What is the minimal structure we need to assume about the lattice? We need a set of vertices, which host the degrees of freedom, and a set of edges connecting them. This defines a mathematical object known as a *graph*. For example, the quantum Ising model can be defined on any spatial graph where the interaction is along the edges. In fact, lattice models on general graphs can be engineered in cold atom experiments. See [118–120] for examples. One is then naturally led to the question: *can we construct exotic lattice models such as fractons on a general graph?* (See [121] for an example in this direction.)

In this work, we propose two lattice models on an arbitrary finite, undirected graph Γ. (For simplicity, we assume that the graph is simple and connected. See Section 3.2 for the definitions of these adjectives.) We work with a Euclidean spacetime where each spatial slice is Γ (see Figure 3.1). Our results can easily be recast in a Hamiltonian formulation. One model is a matter theory based on a compact scalar field ϕ, while the other is a pure $U(1)$ gauge theory associated with the global symmetry of the matter theory. We refer to them as the *Laplacian ϕ-theory* and the $U(1)$

Laplacian gauge theory, respectively, because they are constructed using the *discrete Laplacian operator* Δ_L on a general graph Γ.

Naïvely, the Laplacian ϕ-theory can be viewed as a particular regularization of the compact scalar Lifshitz theory described by the Lagrangian

$$\mathcal{L} = \frac{\mu_0}{2}(\partial_\tau \phi)^2 + \frac{\mu}{2}(\nabla^2 \phi)^2 \ , \tag{3.1}$$

where ∇^2 is the Laplacian differential operator on the spatial manifold. On the other hand, the $U(1)$ Laplacian gauge theory can be viewed as a particular regularization of a rank-2 $U(1)$ gauge theory with gauge fields (A_τ, A) satisfying the gauge symmetry[2]

$$A_\tau \sim A_\tau + \partial_\tau \alpha \ , \qquad A \sim A + \nabla^2 \alpha \ . \tag{3.2}$$

The Lagrangian is

$$\mathcal{L} = \frac{1}{2g^2} E^2 \ , \tag{3.3}$$

where $E = \partial_\tau A - \nabla^2 A_\tau$ is the gauge invariant electric field. However, these continuum Lagrangians do not specify the systems unambiguously. In this chapter, we use the modified Villain formulation of Chapter 2 to provide a precise formulation of these systems. We will see that various physical observables (including the GSD) depend sensitively on how the space is discretized by a discrete lattice graph Γ. In particular, the discrete Laplacian difference operator Δ_L does not have a smooth continuum limit to the differential operator ∇^2.

It is important to emphasize that the Laplacian ϕ-theory is not robust. Small deformations of the short distance theory change the elaborate long distance structure that we find here. (See the discussion in [35].) On the other hand, the $U(1)$ Laplacian

[2] We call this a "rank-2" gauge theory because the gauge transformation of the spatial gauge field involves a second-order spatial derivative.

gauge theory is a robust model. In Chapter 4, we will discuss gapped robust models that are related to these models.

Since our models can be placed on any spatial graph Γ, they can be defined in general spatial dimensions. In particular, we can take the graph Γ to be a d-dimensional torus lattice for any $d \geq 1$. We will examine our models on 1d and 2d torus graphs in details below. More generally, we can place our models on a general graph where there is no clear notion of dimensionality or locality. In particular, foliation structure is not needed to define these models.

In analyzing these models, we follow the approach advocated in [57, 35–38, 58, 59, 1, 26, 122, 25], i.e., we focus on the global symmetries and pursue their consequences. Here are the highlights of these models:

1. The discrete global symmetry of the matter theory is based on the *Jacobian group* of the graph Γ, denoted as $\mathrm{Jac}(\Gamma)$, which is a well-studied finite abelian group associated with a general graph in graph theory. Relatedly, the discrete time-like global symmetry [25], which acts on defects, is also given by $\mathrm{Jac}(\Gamma)$ in the gauge theory.

2. In the matter theory, the ground state degeneracy is equal to $|\,\mathrm{Jac}(\Gamma)|$, the order of the Jacobian group. This in turn is equal to the number of *spanning trees* of Γ, which we will define in the main text. The number of spanning trees is a common measure of the *complexity* of a general graph in graph theory (see, for example, [123, 124]). Therefore, at the level of a slogan, we have

$$\mathrm{GSD} = \text{complexity}\ . \tag{3.4}$$

This is one manifestation of the UV/IR mixing phenomenon observed in many of these exotic models: certain low-energy/long-distance observables depend sensitively on the short-distance details [35, 26]. In our matter theory, the GSD equals the complexity of the discretized spatial graph, which can be thought of as a short-distance regularization. Under a mild assumption on Γ, the ground state degeneracy grows exponentially in the number of vertices of Γ [124].[3]

3. In the gauge theory, the defects, which describe the worldlines of infinitely heavy probe particles, are immobile under a mild assumption on the graph Γ. Therefore, the $U(1)$ Laplacian gauge theory is a theory of *fractons on a general graph*.[4] There is a beautiful correspondence between the physical observables for fractons and various graph-theoretic concepts. In particular, the space of superselection sectors for fractons is translated into the theory of divisors of a graph.

These features are analogous to those at the beginning of this introduction. Interestingly, they are intimately related to some well-studied properties of a graph.

This chapter is organized as follows. In Section 3.2, we collect some useful mathematical facts about graphs, and functions on their vertices. We define the Laplace difference operator Δ_L on the graph Γ, and discuss the properties of solutions to discrete Laplace and Poisson equations. We introduce the theory of divisors on a finite graph, and define the Picard group and the Jacobian group of Γ. We also define the Abel-Jacobi map on the graph and discuss its properties. We relate the Jacobian group to the Smith normal form of the Laplacian operator and the spanning trees of Γ. These results can be found in any standard textbook on graph theory such as [125, 126]. In Appendix 3.A, we solve the discrete Poisson equation on any graph using the Smith decomposition of the Laplacian matrix.

[3] One might wonder if this system is as trivial as decoupled spins on the sites of the graph because the GSD grows exponentially in N in both cases. However, as we will show below, the origins of this exponential behavior are very different. When both systems are placed on a 2d torus spatial lattice with L sites in each direction, the minimal number of generators of discrete momentum symmetry group of the Laplacian ϕ-theory grows only linearly in L, whereas it grows as L^2 in the decoupled spin system. So the large GSD of the Laplacian ϕ-theory comes from the large orders of some of the generators of $\mathrm{Jac}(\Gamma)$ rather than the number of generators.

[4] In contrast to the matter theory, the ground state of the gauge theory (which has fractons) is non-degenerate (assuming that the θ-angle is not π). See Section 3.4.2.

In Section 3.3, we introduce the Laplacian ϕ-theory on the graph Γ. It is a self-dual model with momentum and winding symmetries. The momentum (or winding) symmetry is $U(1) \times \mathrm{Jac}(\Gamma)$, where $\mathrm{Jac}(\Gamma)$ is the Jacobian group of Γ. The non-commutativity of the momentum and winding symmetries leads to a ground state degeneracy equal to $|\mathrm{Jac}(\Gamma)|$, the order of the Jacobian group. This is also equal to the number of spanning trees or the complexity of Γ. We note that this model is not robust—deforming the theory by a winding operator breaks the winding symmetry and lifts the degeneracy.

In Section 3.4, we introduce the $U(1)$ Laplacian gauge theory on the graph Γ. It is the pure gauge theory associated with the momentum symmetry of the Laplacian ϕ-theory. While the space-like global symmetry of this model is simply $U(1)$, the time-like global symmetry is $U(1) \times \mathrm{Jac}(\Gamma)$. This leads to selection rules on the mobility of defects. In particular, we prove that a defect with unit charge is completely immobile when Γ is 2-edge connected (we will define this below). In other words, a single charged particle is a fracton. We also give a complete characterization of mobility of defects for arbitrary Γ. Finally, we note an interesting correspondence between divisors in graph theory and configurations of fractons on Γ, which is summarized in Table 3.1.

In both Sections 3.3 and 3.4, we discuss two concrete examples where we place each model on spatial circle and torus respectively. While the former gives the 1+1d dipole theories analyzed in [25], the latter gives new 2+1d models with interesting properties. These new 2+1d models can be interpreted as extensions of the 1+1d dipole theories of [25] to 2+1d. In fact, there are other ways to extend the 1+1d dipole theories to 2+1d. In a separate paper [127], we compare all these 2+1d models, and discuss their relation to the 2+1d compact Lifshitz theory [128–137] and 2+1d rank-2 $U(1)$ tensor gauge theory [23, 116].

Theory of divisors	$U(1)$ Laplacian gauge theory
Graph Γ	Spatial lattice
Divisor $q \in \mathcal{F}(\Gamma, \mathbb{Z})$	Configuration of fractons with $U(1)$ time-like charges $q(i)$ at site i
Principal divisor $q \in \mathrm{im}_{\mathbb{Z}} \Delta_L$	Configuration of fractons in trivial superselection sector
Degree of q $\deg q := \sum_i q(i)$	Total $U(1)$ time-like charge of a configuration of fractons
Picard group $\mathrm{Pic}(\Gamma)$	Space of all superselection sectors, or space of all time-like charges
Jacobian group $\mathrm{Jac}(\Gamma)$	Space of superselection sectors with trivial total $U(1)$ time-like charge
Pontryagin dual of $\mathrm{Pic}(\Gamma)$ $U(1) \times \mathrm{Jac}(\Gamma)$	Time-like symmetry group

Table 3.1: The correspondence between the theory of divisors on the graph Γ, and the $U(1)$ Laplacian gauge theory on the spatial lattice Γ. The graph-theoretic objects in the left column are defined in Section 3.2, and the physical observables in the right column are discussed in Section 3.4.

In Chapter 4, we propose and analyze two gapped $\mathbb{Z}_N$ models on a graph. One of them is a fracton model, which is a Higgsed version of the $U(1)$ Laplacian gauge theory on the graph. The other is a robust lineon model. We also study these gapped models on a spatial torus, and compare them with 2+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory [65, 66, 138–140].

3.2 Graph theory primer

In this section, we collect some important mathematical facts about a finite graph, and functions on the graph valued in abelian groups such as $\mathbb{R}$, $U(1)$, $\mathbb{Z}$, or $\mathbb{Z}_N$. Most of the details can be found in a standard textbook on spectral graph theory such as [125]. The theory of divisors on finite graphs is discussed in [126].

Let Γ be a *simple* (at most one edge between any two vertices and no self-loops), *undirected* (no directed edges), *connected* (any two vertices are connected by a path) graph on N vertices. We use i to denote a vertex (or site), and $\langle i,j\rangle$ or e to denote an edge (or link).

The *adjacency matrix* A of Γ is an $\mathsf{N}\times\mathsf{N}$ symmetric matrix given by $A_{ij}=1$ if there is an edge $\langle i,j\rangle$ between vertices i and j, and $A_{ij}=0$ otherwise. The *degree* d_i of a vertex i is the number of edges incident to the vertex i. Let $D=\mathrm{diag}(d_1,\dots,d_{\mathsf{N}})$ be the degree matrix. The *Laplacian matrix* L of Γ is defined as $L:=D-A$. Note that L is symmetric.

Here are some common examples/classes of graphs:

- A *k-regular graph* is a graph where every vertex has degree k.
- A *k-edge connected graph* is a graph where removing any $k-1$ edges still leaves it connected.
- The *complete graph* on N vertices, denoted as K_{N}, is a graph that has an edge between any two vertices. Equivalently, it is the only $(\mathsf{N}-1)$-regular graph on N vertices.
- The *cycle graph* on N vertices, denoted as C_{N}, is a graph that is a cycle or loop. Equivalently, it is the only 2-regular graph on N vertices.
- A *tree* on N vertices is a graph that contains no cycle or loop. If it is connected, then it is called a *spanning tree*.

3.2.1 Discrete harmonic functions and Smith decomposition of L

Consider a function on the vertices of the graph, $f : \Gamma \to X$, where X is an abelian group. We denote the set of all such functions as $\mathcal{F}(\Gamma, X)$. Define the *discrete Laplacian* operator $\Delta_L : \mathcal{F}(\Gamma, X) \to \mathcal{F}(\Gamma, X)$ as

$$\Delta_L f(i) := \sum_j L_{ij} f(j) = d_i f(i) - \sum_{j:\langle i,j\rangle \in \Gamma} f(j) = \sum_{j:\langle i,j\rangle \in \Gamma} [f(i) - f(j)] \ . \tag{3.5}$$

The additions and subtractions here are with respect to the group multiplication of X. This is one of the most natural and universal difference operators that can be defined on any such graph Γ. The *image* of $\mathcal{F}(\Gamma, X)$ under Δ_L is denoted as $\mathrm{im}_X\, \Delta_L$. A function $f \in \mathcal{F}(\Gamma, X)$ is said to be *discrete harmonic* if it satisfies

$$\Delta_L f(i) = 0 \ , \tag{3.6}$$

where 0 is the identity element in X. We denote the set of all X-valued discrete harmonic functions as $\mathcal{H}(\Gamma, X)$, or $\ker_X \Delta_L$, the *kernel* of Δ_L.[5]

Given a $g \in \mathcal{F}(\Gamma, X)$, consider the *discrete Poisson equation*

$$\Delta_L f(i) = g(i) \ . \tag{3.7}$$

If $g = 0$, then (3.7) is called a *discrete Laplace equation.* We define the *cokernel* of Δ_L as the quotient

$$\mathrm{coker}_X\, \Delta_L := \frac{\mathcal{F}(\Gamma, X)}{\mathrm{im}_X\, \Delta_L} \ , \tag{3.8}$$

Trivially, a solution to the discrete Poisson equation (3.7) exists if and only if g is in the same equivalence class as 0 in this quotient.

[5] The space $\ker_X \Delta_L$ is also known as the *group of balanced vertex weightings* [141].

Solutions to the discrete Poisson equation (3.7), if any, can be found using the *Smith decomposition* [142] of the Laplacian matrix L [143]. The *Smith normal form* of L is given by

$$R = PLQ \ , \qquad \text{or} \qquad R_{ab} = \sum_{i,j} P_{ai} L_{ij} Q_{jb} \ , \tag{3.9}$$

where $P, Q \in GL_{\mathsf{N}}(\mathbb{Z})$, and $R = \mathrm{diag}(r_1, \ldots, r_{\mathsf{N}})$ is an $\mathsf{N} \times \mathsf{N}$ diagonal integer matrix with nonnegative diagonal entries, known as the *invariant factors* of L, such that $r_a | r_{a+1}$ (i.e., r_a divides r_{a+1}) for $a = 1, \ldots, \mathsf{N} - 1$. While R is uniquely determined by L, the matrices P and Q are not. More details on the structure of P and Q, and how to solve (3.7) using the Smith decomposition of L can be found in Appendix 3.A.

One important result that we will repeatedly use is the general solution to the $U(1)$-valued discrete Laplace equation. In Appendix 3.A (see in particular (3.77)), we show that the most general $U(1)$-valued discrete harmonic function $f \in \mathcal{H}(\Gamma, U(1))$ on a graph Γ is given by

$$f(i) = 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_a}{r_a} + c \in \mathbb{R}/2\pi\mathbb{Z} \ , \tag{3.10}$$

parametrized by a circle-valued constant c, i.e., $c \sim c + 2\pi$, and $\mathsf{N} - 1$ integers $p_a = 0, 1, \ldots, r_a - 1$. If we lift this solution to $\mathbb{R}$, then it obeys

$$\Delta_L f(i) = 2\pi \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_a \in 2\pi\mathbb{Z} \ . \tag{3.11}$$

3.2.2 Theory of divisors and the Abel-Jacobi map on a graph

There is an interesting analogy between the theory of integer-valued functions on a finite graph and the theory of divisors on a Riemann surface [144, 145, 126]. In this context, an element of $\mathcal{F}(\Gamma, \mathbb{Z})$ is known as a *divisor*, and an element of $\mathrm{im}_{\mathbb{Z}} \Delta_L$ is known as a *principal divisor*. Given a divisor q, its *degree* is defined as $\deg q := \sum_i q(i)$. Let $\mathcal{F}^k(\Gamma, \mathbb{Z})$ denote the set of all degree-k divisors. Note that any principal divisor has degree zero, so we can define the quotients:

$$\mathrm{Pic}(\Gamma) := \frac{\mathcal{F}(\Gamma,\mathbb{Z})}{\mathrm{im}_{\mathbb{Z}}\,\Delta_L} = \mathrm{coker}_{\mathbb{Z}}(\Delta_L)\ , \qquad \mathrm{Jac}(\Gamma) := \frac{\mathcal{F}^0(\Gamma,\mathbb{Z})}{\mathrm{im}_{\mathbb{Z}}\,\Delta_L}\ , \tag{3.12}$$

known as the *Picard group*, and the *Jacobian group* respectively. As the names suggest, they are groups; in fact, they are abelian groups. They are related by the split-exact sequence

$$0 \longrightarrow \mathrm{Jac}(\Gamma) \longrightarrow \mathrm{Pic}(\Gamma) \xrightarrow{\deg} \mathbb{Z} \longrightarrow 0\ . \tag{3.13}$$

The *characteristic function* of a vertex i is defined as $\chi_i(j) := \delta_{ij}$. Given two vertices $i, i' \in \Gamma$, we define the function $s_{i,i'}(j) := \chi_i(j) - \chi_{i'}(j)$. For a fixed vertex $i_0 \in \Gamma$, the *Abel-Jacobi map*, $S_{i_0} : \Gamma \to \mathrm{Jac}(\Gamma)$, is defined as

$$S_{i_0}(i) := [s_{i,i_0}]\ . \tag{3.14}$$

It is well-defined because $s_{i,i_0} \in \mathcal{F}^0(\Gamma,\mathbb{Z})$. Even though i and i_0 play similar roles in the right-hand side of this equation, they will play different roles below. This is the reason for the asymmetry between them in the left-hand side of (3.14).

The Abel-Jacobi map enjoys several nice properties:

1. It is a $\mathrm{Jac}(\Gamma)$-valued discrete harmonic function that vanishes at i_0.

2. For any abelian group X, given a group homomorphism $\psi : \mathrm{Jac}(\Gamma) \to X$, the composition $\psi \circ S_{i_0} : \Gamma \to X$ is an X-valued discrete harmonic function that vanishes at i_0. Conversely, for any X-valued discrete harmonic function f that vanishes at i_0, there is a unique group homomorphism $\psi_f : \mathrm{Jac}(\Gamma) \to X$ such that $f = \psi_f \circ S_{i_0}$. This is known as the *universal property* of the Abel-Jacobi map.

3. It is injective if and only if Γ is 2-edge connected. More generally, for any two vertices $i, i' \in \Gamma$, $S_{i_0}(i) = S_{i_0}(i')$ if and only if there is a unique path from i to i' in Γ.

These properties of the Abel-Jacobi map turn the study of discrete harmonic functions on a graph to a problem in group theory.

3.2.3 Jacobian group and complexity of a graph

The Jacobian group $\mathrm{Jac}(\Gamma)$ defined in the previous subsection is a natural finite abelian group that is associated with a general graph Γ.[6] It is closely related to the Smith normal form of the Laplacian matrix L (see Appendix 3.A, especially, around (3.88)). In particular, we have

$$\mathrm{Jac}(\Gamma) \cong \prod_{a<\mathsf{N}} \mathbb{Z}_{r_a} \ . \tag{3.15}$$

The order of $\mathrm{Jac}(\Gamma)$ can be expressed in terms of the nonzero eigenvalues of L:

$$|\,\mathrm{Jac}(\Gamma)| = \prod_{a<\mathsf{N}} r_a = \frac{\lambda_2 \cdots \lambda_{\mathsf{N}}}{\mathsf{N}} \ , \tag{3.16}$$

where $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_{\mathsf{N}}$ are the eigenvalues of L.[7] By *Kirchhoff's matrix-tree theorem* [150], this is equal to the number of *spanning trees* of Γ. Here, a spanning tree of Γ is a subgraph that is a spanning tree on the vertices of Γ. See Figure 3.1(b) for an example.

The number of spanning trees is the most fundamental and well-studied notion of *complexity* in graph theory. Intuitively, it is a measure of how "connected" Γ is.

[6] It has several different names in the graph theory literature, including the *sandpile group* [146, 147], or the *group of components* [148], or the *critical group* [149] of Γ, and it is related to the *group of bicycles* [141] of Γ. In particular, it would be interesting to understand the connection between our models and the abelian sandpile model.

[7] The eigenvalues of L are real because L is symmetric. The zero eigenvalue of L corresponds to the zero mode of the Laplacian operator Δ_L. The other eigenvalues are all positive because Γ is connected.

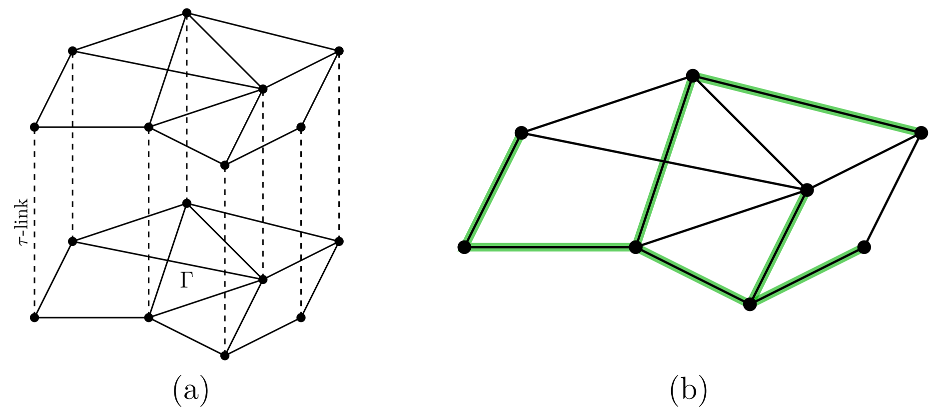


Figure 3.1: (a) The Euclidean spacetime $\mathbb{Z}_{L_\tau} \times \Gamma$: the solid lines and dots correspond to a spatial slice which is the graph Γ, and the dashed lines represent the τ-links. (b) A spanning tree of Γ associated with the highlighted (green) edges.

For example, it is easy to see that the number of spanning trees of C_{N}, the cycle graph on N vertices, is N. In contrast, an old result of Cayley states that the number of spanning trees of K_{N}, the complete graph on N vertices, is $\mathsf{N}^{\mathsf{N}-2}$ [151].[8] More generally, when Γ is k-regular, the number of spanning trees of Γ grows exponentially in N whenever $k \geq 3$ [124].

3.3 Laplacian ϕ-theory on a graph

In this section, we study the modified Villain version [1] of the Laplacian ϕ-theory on a graph Γ. It can be viewed as an extension of the 1+1d dipole ϕ-theory of [25] with Δ_x^2 on the 1d spatial lattice replaced by the discrete Laplacian operator Δ_L on the graph Γ. The Euclidean spacetime is $\mathbb{Z}_{L_\tau} \times \Gamma$, i.e., each spatial slice is Γ, and any two adjacent spatial slices are connected by τ-links at each vertex (see Figure 3.1(a)).

[8] It is not too difficult to prove this using (3.16).

The modified Villain version of the Laplacian ϕ-theory is described by the action[9]

$$\begin{aligned} S = \frac{\beta_0}{2}\sum_{\hat\tau,i}\left[\Delta_\tau\phi(\hat\tau,i) - 2\pi n_\tau(\hat\tau,i)\right]^2 + \frac{\beta}{2}\sum_{\hat\tau,i}\left[\Delta_L\phi(\hat\tau,i) - 2\pi n(\hat\tau,i)\right]^2 \\ + i\sum_{\hat\tau,i}\tilde\phi(\hat\tau,i)\left[\Delta_\tau n(\hat\tau,i) - \Delta_L n_\tau(\hat\tau,i)\right] \ , \end{aligned} \tag{3.18}$$

where ϕ is a real-valued field at each site, n is an integer gauge field at each site, n_τ is an integer gauge field on each τ-link, and $\tilde\phi$ is a real-valued Lagrange multiplier on each τ-link.[10] Recall that $\sum_i$ stands for the sum over all the sites i of the graph Γ. There is a gauge symmetry

$$\begin{aligned} &\phi(\hat\tau,i) \sim \phi(\hat\tau,i) + 2\pi k(\hat\tau,i) \ , \\ &n_\tau(\hat\tau,i) \sim n_\tau(\hat\tau,i) + \Delta_\tau k(\hat\tau,i) \ , \\ &n(\hat\tau,i) \sim n(\hat\tau,i) + \Delta_L k(\hat\tau,i) \ , \\ &\tilde\phi(\hat\tau,i) \sim \tilde\phi(\hat\tau,i) + 2\pi\tilde k(\hat\tau,i) \ , \end{aligned} \tag{3.19}$$

where k and $\tilde k$ are integer gauge parameters on the sites and τ-links, respectively. This integer gauge symmetry makes the scalar fields ϕ and $\tilde\phi$ compact. The lattice action (3.18) is a particular lattice regularization of a compact Lifshitz scalar field theory that can be defined on a general graph Γ. See Chapter 2 for similar modified Villain formulations of various standard and exotic theories of compact scalar fields.

[9] One can also study the *Laplacian XY model*, which is described by the following action:

$$S = -\beta_0\sum_{\hat\tau,i}\cos\left[\Delta_\tau\varphi(\hat\tau,i)\right] - \beta\sum_{\hat\tau,i}\cos\left[\Delta_L\varphi(\hat\tau,i)\right] \ , \tag{3.17}$$

where $\varphi(\hat\tau,i)$ is a circle-valued field on each site of the spacetime lattice. The momentum symmetry of this cosine model is the same as that of the modified Villain model (3.18).

[10] We associate the τ-link between $(\hat\tau,i)$ and $(\hat\tau+1,i)$ to the site $(\hat\tau,i)$.

3.3.1 Self-duality

Since L is symmetric, the modified Villain model (3.18) is self-dual with $\phi \leftrightarrow \tilde{\phi}$ and $\beta_0 \leftrightarrow \frac{1}{(2\pi)^2\beta}$. Indeed, using the Poisson resummation formula for the integers n_τ, n, the dual action is

$$\begin{aligned} S = &\frac{1}{2(2\pi)^2\beta}\sum_{\hat{\tau},i}\left[\Delta_\tau\tilde{\phi}(\hat{\tau},i) - 2\pi\tilde{n}_\tau(\hat{\tau},i)\right]^2 + \frac{1}{2(2\pi)^2\beta_0}\sum_{\hat{\tau},i}\left[\Delta_L\tilde{\phi}(\hat{\tau},i) - 2\pi\tilde{n}(\hat{\tau},i)\right]^2 \\ &- i\sum_{\hat{\tau},i}\phi(\hat{\tau},i)\left[\Delta_\tau\tilde{n}(\hat{\tau},i) - \Delta_L\tilde{n}_\tau(\hat{\tau},i)\right] \ , \end{aligned} \tag{3.20}$$

where $(\tilde{n}_\tau, \tilde{n})$ are integer gauge fields that make $\tilde{\phi}$ compact. Under the gauge symmetry (3.19), they transform as

$$\tilde{n}_\tau(\hat{\tau},i) \sim \tilde{n}_\tau(\hat{\tau},i) + \Delta_\tau\tilde{k}(\hat{\tau},i) \ , \qquad \tilde{n}(\hat{\tau},i) \sim \tilde{n}(\hat{\tau},i) + \Delta_L\tilde{k}(\hat{\tau},i) \ . \tag{3.21}$$

When the graph is a two-dimensional torus graph, this is related to the self-duality of the 2+1d compact Lifshitz scalar field theory discussed in [130].

3.3.2 Global symmetry and the Jacobian group

The momentum global symmetry of the action (3.18) corresponds to shifting the fields ϕ, n by

$$\begin{aligned} \phi(\hat{\tau},i) &\to \phi(\hat{\tau},i) + f(i) \ , \\ n(\hat{\tau},i) &\to n(\hat{\tau},i) + \frac{1}{2\pi}\Delta_L f(i) \ , \end{aligned} \tag{3.22}$$

where the function $f(i)$ obeys $\Delta_L f(i) \in 2\pi\mathbb{Z}$. In other words, $f(i)$ is a solution to the $U(1)$-valued discrete Laplace equation, i.e., $f \in \mathcal{H}(\Gamma, U(1))$. The simplest example of such an $f(i)$ is a constant, i.e., $f(i) = c$ with $c \sim c + 2\pi$. This corresponds to a $U(1)$ momentum symmetry.

The most general solution to the $U(1)$-valued discrete Laplace equation is given in (3.10), which leads to the following momentum symmetry

$$\begin{aligned} \phi(\hat{\tau}, i) &\to \phi(\hat{\tau}, i) + 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_a}{r_a} + c \ , \\ n(\hat{\tau}, i) &\to n(\hat{\tau}, i) + \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_a \ . \end{aligned} \tag{3.23}$$

The momentum symmetry is parametrized by a circle-valued constant c and $\mathsf{N} - 1$ integers $p_a = 0, \ldots, r_a - 1$. Here P, Q and r_a are associated with the Smith decomposition of L. Note that the shift in n is given by (3.11). The parameter p_a generates a $\mathbb{Z}_{r_a}$ discrete momentum symmetry for each $a < \mathsf{N}$. The total momentum symmetry is therefore $U(1) \times \mathrm{Jac}(\Gamma)$ where

$$\mathrm{Jac}(\Gamma) = \prod_{a<\mathsf{N}} \mathbb{Z}_{r_a} \ , \tag{3.24}$$

is the Jacobian group of the graph Γ. (See [141] for an alternative interpretation of the momentum symmetry group $U(1) \times \mathrm{Jac}(\Gamma)$.) As we will see in Section 3.3.6, when Γ is a 2d torus lattice $C_L \times C_L$, the minimal number of generators of $\mathrm{Jac}(\Gamma)$ grows only linearly in L.

In addition to the $U(1) \times \mathrm{Jac}(\Gamma)$ momentum symmetry, there is also a $U(1) \times \mathrm{Jac}(\Gamma)$ winding symmetry. This is to be expected given the self-duality of the theory. The $U(1)$ winding charge is

$$\tilde{\mathcal{Q}} = \sum_i P_{\mathsf{N}i} n(\hat{\tau}, i) = \sum_i n(\hat{\tau}, i) \ , \tag{3.25}$$

while the $\mathbb{Z}_{r_a}$ discrete winding charge is

$$\tilde{\mathcal{Q}}_a = \sum_i P_{ai} n(\hat{\tau}, i) \mod r_a \ , \tag{3.26}$$

for each $a < \mathsf{N}$. They are conserved due to the flatness of (n_τ, n) imposed by the Lagrange multiplier $\tilde{\phi}$.

3.3.3 GSD = Complexity

In this subsection, we compute the ground state degeneracy of the Laplacian ϕ-theory. To facilitate this computation, we find it convenient to first gauge fix the integer gauge fields. First, we gauge fix $n_\tau = 0$ everywhere except at $\hat{\tau} = 0$. The remaining integer gauge freedom is the set of time-independent gauge transformations $k(i)$. By the flatness condition, we also have $\Delta_\tau n = 0$, so $n(\hat{\tau}, i) = n(i)$. By the analysis in Appendix 3.A around (3.88), the honolomies of $n(i)$ with gauge parameter $k(i)$ are precisely the winding charges (3.25) and (3.26).

In this gauge, a discrete winding configuration with discrete winding charges $\tilde{\mathcal{Q}}_a = p_a \mod r_a$ for $a < \mathsf{N}$ and zero $U(1)$ winding charge, $\tilde{\mathcal{Q}} = 0$, is given by[11]

$$\phi(\hat{\tau}, i) = 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_a}{r_a} , \qquad n(\hat{\tau}, i) = \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_a . \tag{3.27}$$

This is the configuration in (3.10) with $c = 0$. There are $|\operatorname{Jac}(\Gamma)| = \prod_{a<\mathsf{N}} r_a$ such discrete winding configurations labeled by the p_a's. All these configurations have zero energy. Therefore, the ground state degeneracy is

$$\text{GSD} = |\operatorname{Jac}(\Gamma)| = \prod_{a<\mathsf{N}} r_a . \tag{3.28}$$

As mentioned in Section 3.2, this is equal to the number of spanning trees of the graph Γ, which measures how complex a graph is. It follows that, when Γ is k-regular with $k \geq 3$, the ground state degeneracy grows exponentially in the number of vertices N [124].

[11] This configuration has $\tilde{\mathcal{Q}} = 0$ because $\sum_i (P^{-1})_{ia} = 0$ for $a < \mathsf{N}$ as shown in (3.68).

Let us compare the Laplacian ϕ-theory with a decoupled spin system: a finite dimensional spin at every site with trivial Hamiltonian. In both systems, the GSD grows exponentially in the number of vertices N. How do we differentiate these systems? For simplicity, let us place both systems on a torus graph, $\Gamma = C_L \times C_L$. Then, in the Laplacian ϕ-theory, the minimal number of generators of the discrete momentum symmetry group (3.24) grows linearly in L (see Section 3.3.6), and some of these generators have very large orders. On the other hand, in the decoupled spin system, there is a generator at each site, so the symmetry group has L^2 generators each with the same fixed order. In other words, the large GSD of Laplacian ϕ-theory comes from the large orders of some of the generators of $\mathrm{Jac}(\Gamma)$ rather than the number of generators.

There is another way to derive the above ground state degeneracy. The discrete momentum and winding symmetries do not commute with each other: the shift (3.23) of n changes the discrete winding charge $\tilde{\mathcal{Q}}_a$ in (3.26) by $p_a \mod r_a$. This can be interpreted as a mixed 't Hooft anomaly between the discrete momentum and winding symmetries, and it leads to the ground state degeneracy (3.28). In fact, the entire Hilbert space is in a projective representation of $\mathrm{Jac}(\Gamma) \times \mathrm{Jac}(\Gamma)$, so every state is $|\mathrm{Jac}(\Gamma)|$-fold degenerate.

3.3.4 Spectrum

Let us now determine the spectrum of the Laplacian ϕ-theory (3.18) by working with a continuous Lorentzian time while keeping the space discrete. To do this, we first take $L_\tau \to \infty$, and gauge fix $n_\tau(\hat{\tau}, i) = 0$, so that $n(\hat{\tau}, i) = n(i)$ and $k(\hat{\tau}, i) = k(i)$ are both time-independent. We then introduce the lattice spacing a_τ in the τ-direction, and take the limit $a_\tau \to 0$ while keeping $\beta_0' = \beta_0 a_\tau$ and $\beta' = \beta / a_\tau$ fixed. Finally, we Wick rotate from Euclidean time τ to Lorentzian time t.

The equation of motion of ϕ is

$$\beta_0' \partial_0^2 \phi(t,i) + \beta' \Delta_L [\Delta_L \phi(t,i) - 2\pi n(i)] = 0 \ . \tag{3.29}$$

The general solution to this equation is

$$\begin{aligned} \phi(t,i) &= f(i) + \phi_p(i) + \sum_{\lambda \neq 0} \phi_\lambda(i) e^{i\omega_\lambda t} \ , \\ n(i) &= \frac{1}{2\pi} \Delta_L f(i) + p\delta_{i\mathsf{N}} \ , \end{aligned} \tag{3.30}$$

where $f(i) = 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_a}{r_a} + c$ is a $U(1)$-valued discrete harmonic function given by (3.10), $\phi_p(i)$ is a real-valued solution to the equation

$$\Delta_L \phi_p(i) = 2\pi p \left(\delta_{i\mathsf{N}} - \frac{1}{\mathsf{N}} \right) \ , \qquad p \in \mathbb{Z} \ , \tag{3.31}$$

$\phi_\lambda(i)$ is a real-valued eigenfunction of the Laplacian operator with eigenvalue λ, i.e., $\Delta_L \phi_\lambda(i) = \lambda \phi_\lambda(i)$, and the dispersion relation for the "plane wave modes" is[12]

$$\omega_\lambda = \sqrt{\frac{\beta'}{\beta_0'}} \, \lambda \ . \tag{3.32}$$

In other words, the plane wave spectrum is exactly the set of nonzero eigenvalues of the Laplacian operator. The smallest nonzero eigenvalue λ_2 of the Laplacian operator is known as the *spectral gap* [125], or the *algebraic connectivity* [152] of the graph. When Γ is a torus lattice in any dimension, λ_2 goes to zero with increasing number of sites. However, on a general graph, λ_2 could be finite even for large N [153]. So the plane wave spectrum could be gapped on a general graph, while it is gapless on a torus lattice.

[12]We refer to them as the "plane wave modes" because when Γ is a torus lattice in any dimension, the eigenvalues are related to the spatial momenta and the eigenfunctions form the usual Fourier basis.

The zero mode c of $f(i)$ in (3.30) is charged under the $U(1)$ momentum symmetry. After giving it a time dependence, its energy is lifted quantum mechanically to

$$E_{\text{mom}} = \frac{n^2}{2\beta_0' \mathsf{N}} \ , \tag{3.33}$$

where $n \in \mathbb{Z}$ is the $U(1)$ momentum charge. See [35] for a similar phenomenon, where a classical zero mode is lifted quantum mechanically, in another exotic model.

The rest of $f(i)$ in (3.30) is charged under the discrete winding symmetry $\text{Jac}(\Gamma)$ with charges $\tilde{\mathcal{Q}}_a = p_a \mod r_a$ for $a < \mathsf{N}$. As we saw before, the discrete winding configurations have zero energy.

Finally, $\phi_p(i)$ in (3.30) is charged under the $U(1)$ winding symmetry with charge $\tilde{\mathcal{Q}} = p$. The energy of the winding configuration $\phi_p(i)$ is

$$E_{\text{wind}} = \frac{\beta'}{2} \sum_i [\Delta_L \phi_p(i) - 2\pi p \delta_{i\mathsf{N}}]^2 = \frac{(2\pi)^2 \beta' p^2}{2\mathsf{N}} \ . \tag{3.34}$$

3.3.5 Robustness of GSD

The action (3.18) is said to be natural with respect to the global symmetry if all the relevant terms that are invariant under this symmetry are included in the action [154]. (See [35] for a recent discussion of naturalness and robustness). For example, a term that one can write on any graph is

$$-\sum_{\hat{\tau},i} \cos[\phi(\hat{\tau},i)] \ . \tag{3.35}$$

However, this term is not invariant under the $U(1)$ momentum symmetry. So if we impose the $U(1)$ momentum symmetry, it is forbidden.

A more interesting term that one can write on any graph is the usual nearest-neighbor interaction[13]

$$-\sum_{\hat{\tau},\langle i,j\rangle} \cos[\phi(\hat{\tau},i)-\phi(\hat{\tau},j)]\ . \tag{3.36}$$

This term is clearly invariant under the $U(1)$ momentum symmetry, i.e., shifts by constants. What happens if we impose the full momentum symmetry $U(1)\times \mathrm{Jac}(\Gamma)$? When Γ is a tree, the Jacobian group $\mathrm{Jac}(\Gamma)$ is trivial because $|\,\mathrm{Jac}(\Gamma)|=1$. In this case, (3.36) is invariant, and so we should add this term to the action (3.18). This does not affect the ground state degeneracy because the latter is trivial anyway. On the other hand, when Γ is not a tree, $\mathrm{Jac}(\Gamma)$ is nontrivial, which means the momentum symmetry includes non-constant shifts. So, (3.36) is forbidden.

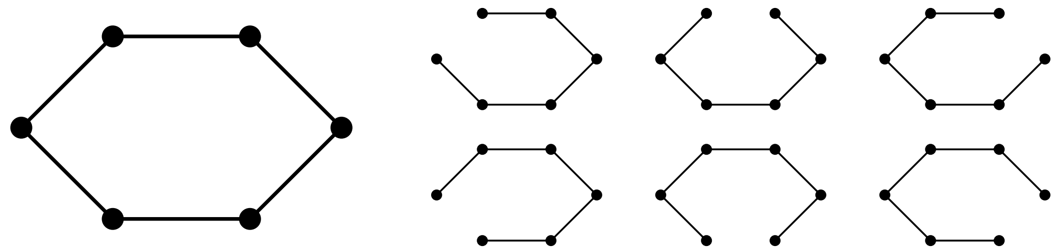
Now, consider the winding operator

$$-\sum_{\hat{\tau},i} \cos[\tilde{\phi}(\hat{\tau},i)]\ . \tag{3.37}$$

This term is invariant under the $U(1)\times\mathrm{Jac}(\Gamma)$ momentum symmetry, so if we impose only the momentum symmetry, we should add it to the action (3.18). Indeed, (3.37) is relevant because it breaks the $U(1)$ winding symmetry and lifts the ground state degeneracy. In other words, if we impose only the momentum symmetry, the winding symmetry is not robust.

It is also natural to impose both momentum and winding symmetries. In this case, (3.37) is forbidden and the ground state degeneracy cannot be lifted.

[13]This is what one would write if one were studying the standard XY model on the graph Γ. In fact, let M denote the number of edges in Γ. Let us choose an orientation for each edge in Γ arbitrarily so that we can talk about *head* and *tail* vertices of an edge. With respect to this orientation, the *(oriented) incidence matrix* B of Γ is an $\mathsf{N}\times\mathsf{M}$ matrix given by $B_{i,e}=1$ if i is the head of e, $B_{i,e}=-1$ if i is the tail of e, and $B_{i,e}=0$ otherwise. One can easily check that $BB^T=L$ independent of the choice of orientation. Therefore, (3.36) is a "lower order" term with respect to the Laplacian term.



(a) (b)

Figure 3.2: (a) The cycle graph C_6 on $L_x = 6$ vertices. (b) The six spanning trees of C_6 obtained by removing one of the six edges.

3.3.6 Examples

In this section, we discuss the Laplacian ϕ-theory on 1d and 2d spatial tori.

1+1d dipole ϕ-theory

Let Γ be a *cycle graph* C_{L_x} (see Figure 3.2), i.e., $\Gamma = C_{L_x} = \mathbb{Z}_{L_x}$, where L_x is the number of sites in the cycle. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator Δ_x^2 in the x-direction.[14] The action (3.18) becomes

$$\begin{aligned} S = \frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{\text{site}} \left(\Delta_x^2 \phi - 2\pi n_{xx}\right)^2 \\ + i \sum_{\tau\text{-link}} \tilde{\phi} \left(\Delta_\tau n_{xx} - \Delta_x^2 n_\tau\right) \ , \end{aligned} \tag{3.38}$$

after replacing n with n_{xx}. This is the modified Villain action of the 1+1d dipole ϕ-theory [25], which is the lattice regularization of the 1+1d compact Lifshitz scalar field theory.

[14] Actually, it is common to define $\Delta_x^2 f(\hat{x}) = f(\hat{x}+1) - 2f(\hat{x}) + f(\hat{x}-1)$, so $\Delta_L = -\Delta_x^2$. We will ignore this discrepancy in sign since it does not affect the rest of the discussion.

The invariant factors of L are

$$r_a = \begin{cases} 1\ , & 1 \le a < L_x - 1\ , \\ L_x\ , & a = L_x - 1\ , \\ 0\ , & a = L_x\ . \end{cases} \tag{3.39}$$

It follows that the Jacobian group for a cycle graph C_{L_x} is

$$\mathrm{Jac}(C_{L_x}) = \mathbb{Z}_{L_x}\ . \tag{3.40}$$

(3.40) is a demonstration of the fact that the large GSD comes from the large orders of some of the genrators of $\mathrm{Jac}(\Gamma)$ rather than the number of generators. Physically, (3.40) means that the discrete momentum and winding symmetries of the 1+1d dipole ϕ-theory are $\mathbb{Z}_{L_x}$.

A spanning tree of a cycle graph C_{L_x} is obtained by removing any one of the L_x edges. See Figure 3.2. Therefore, there are exactly L_x spanning trees of a cycle graph C_{L_x}. In other words, the complexity of C_{L_x} is simply L_x, which equals the ground state degeneracy of the 1+1d dipole ϕ-theory. This is in agreement with the analysis of [25]. As both the discrete global symmetry (which is $\mathrm{Jac}(C_{L_x})$) and the ground state degeneracy (which is $|\,\mathrm{Jac}(C_{L_x})|$) grow as we increase the number of lattice sites L_x, it is clear that this 1+1d model does not have an unambiguous continuum limit $L_x \to \infty$.

Incidentally, using (3.16) to compute the ground state degeneracy in terms of the eigenvalues of Δ_x^2 leads to the following identity:

$$L_x = \frac{1}{L_x} \prod_{k_x=1}^{L_x-1} 4\sin^2\left(\frac{\pi k_x}{L_x}\right)\ . \tag{3.41}$$

2+1d Laplacian ϕ-theory

Let Γ be a *torus graph*, i.e., $\Gamma = C_{L_x} \times C_{L_y} = \mathbb{Z}_{L_x} \times \mathbb{Z}_{L_y}$, where L_i is the number of sites in the i direction. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator $\Delta_x^2 + \Delta_y^2$ in the xy-plane. The action (3.18) becomes

$$\begin{aligned} S = \frac{\beta_0}{2} \sum_{\tau\text{-link}} (\Delta_\tau \phi - 2\pi n_\tau)^2 + \frac{\beta}{2} \sum_{\text{site}} \left[(\Delta_x^2 + \Delta_y^2)\phi - 2\pi n\right]^2 \\ + i \sum_{\tau\text{-link}} \tilde{\phi} \left[\Delta_\tau n - (\Delta_x^2 + \Delta_y^2) n_\tau\right] \ . \end{aligned} \tag{3.42}$$

We refer to this as the 2+1d Laplacian ϕ-theory. This is a natural lattice regularization of the 2+1d compact Lifshitz scalar field theory that can be defined on a general spatial graph Γ.

The discrete momentum and winding symmetries are determined by the Jacobian group of the torus graph. The latter is quite complicated and not known in closed form in general as a function of L_x, L_y. Below we record a few examples for small values of $L_x = L_y$:

$$\begin{aligned} &\text{Jac}(C_2 \times C_2) = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_8 \ , \\ &\text{Jac}(C_3 \times C_3) = \mathbb{Z}_6 \times \mathbb{Z}_6 \times \mathbb{Z}_{18} \times \mathbb{Z}_{18} \ , \\ &\text{Jac}(C_4 \times C_4) = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_8 \times \mathbb{Z}_{24} \times \mathbb{Z}_{24} \times \mathbb{Z}_{24} \times \mathbb{Z}_{96} \ . \end{aligned} \tag{3.43}$$

The minimal number of generators of $\text{Jac}(C_{L_x} \times C_{L_y})$ is at most the number of nontrivial spatial integer gauge fields, n's, after gauge fixing. One can gauge fix the n's to be zero everywhere except along $\hat{x} = 0, 1$, or along $\hat{y} = 0, 1$. Therefore, after gauge fixing, the number of nontrivial n's is $\min(2L_x, 2L_y)$. In other words, on a square torus graph $C_L \times C_L$, the minimal number of generators of $\text{Jac}(C_L \times C_L)$

grows at most linearly in L. Once again, this shows that the large GSD is due to the large orders of some of the generators of $\mathrm{Jac}(\Gamma)$ rather than the number of generators.

The ground state degeneracy is given by the order of the Jacobian group, GSD $= |\,\mathrm{Jac}(C_{L_x}\times C_{L_y})|$. While there is no closed form formula for the Jacobian group itself, the order of this group can be expressed in terms of the eigenvalues of the discrete Laplacian on the torus graph (see (3.16)). We therefore obtain

$$\begin{aligned}
\mathrm{GSD} &= \frac{1}{L_xL_y}\prod_{\substack{0\le k_i<L_i\\(k_x,k_y)\neq(0,0)}}\left[4\sin^2\left(\frac{\pi k_x}{L_x}\right)+4\sin^2\left(\frac{\pi k_y}{L_y}\right)\right]\\
&= L_xL_y\prod_{k_i=1}^{L_i-1}\left[4\sin^2\left(\frac{\pi k_x}{L_x}\right)+4\sin^2\left(\frac{\pi k_y}{L_y}\right)\right]\ ,
\end{aligned} \tag{3.44}$$

where we used the identity (3.41) in the second line. Since the discrete global symmetry $\mathrm{Jac}(C_{L_x}\times C_{L_y})$ and the GSD depend sensitively on the number theoretic properties of the lattice sites L_x, L_y, this regularization of the 2+1d compact Lifshitz theory does not have an unambiguous continuum limit $L_x, L_y\to\infty$.

Let us discuss the asymptotic behavior of the GSD on a torus. The GSD grows asymptotically as [155]

$$\log \mathrm{GSD}\approx\frac{L_xL_y}{\pi^2}\int_0^\pi\int_0^\pi dp_xdp_y\ \log\left[4\sin^2(p_x)+4\sin^2(p_y)\right]=\frac{4G}{\pi}L_xL_y\ , \tag{3.45}$$

where $G\approx 0.916$ is the Catalan constant. This is consistent with the intuition that the number of spanning trees of the torus graph $C_{L_x}\times C_{L_y}$, which is also equal to GSD, grows with L_x and L_y. Indeed, the result of [124] implies the ground state degeneracy of the 2+1d Laplacian ϕ-theory grows exponentially in L_xL_y.[15]

A closely related matter theory is the 2+1d dipole ϕ-theory, which is another possible regularization of the 2+1d compact Lifshitz field theory. While the 1+1d

[15] Here we have used the fact that the torus graph $\Gamma = C_{L_x}\times C_{L_y}$ is 4-regular in applying the result of [124]. In contrast, a cycle graph is only 2-regular, and indeed the GSD of the 1+1d dipole ϕ-theory grows linearly in the number of vertices.

dipole ϕ-theory is the same as the 1+1d Laplacian ϕ-theory, their 2+1d versions are very different. In a separate paper [127], we compare the 2+1d Laplacian and dipole ϕ-theories, and discuss their relation to the 2+1d compact Lifshitz theory.

3.4 $U(1)$ Laplacian gauge theory on a graph

We can gauge the momentum symmetry of the Laplacian ϕ-theory by coupling it to the gauge fields $(\mathcal{A}_\tau, \mathcal{A}; m_\tau)$. Here $\mathcal{A}_\tau$ and $\mathcal{A}$ are real-valued fields living on the τ-links and the sites, respectively, and m_τ is an integer-valued gauge field living on the τ-link. Their gauge transformations are

$$\begin{aligned}
&\mathcal{A}_\tau \sim \mathcal{A}_\tau + \Delta_\tau \alpha + 2\pi q_\tau \ , \\
&\mathcal{A} \sim \mathcal{A} + \Delta_L \alpha + 2\pi q \ , \\
&m_\tau \sim m_\tau + \Delta_\tau q - \Delta_L q_\tau \ ,
\end{aligned} \tag{3.46}$$

where α is a real-valued gauge parameter and q_τ, q are integer-valued gauge parameters.

We can leave out the matter fields and study the pure gauge theory of $(\mathcal{A}_\tau, \mathcal{A}; m_\tau)$. It is described by the following Villain action

$$S = \frac{\gamma}{2} \sum_{\hat{\tau}, i} \mathcal{E}^2 + \frac{i\theta}{2\pi} \sum_{\hat{\tau}, i} \mathcal{E} \ , \tag{3.47}$$

where $\mathcal{E} = \Delta_\tau \mathcal{A} - \Delta_L \mathcal{A}_\tau - 2\pi m_\tau$ is the gauge-invariant electric field of $(\mathcal{A}_\tau, \mathcal{A}; m_\tau)$. The θ-angle is 2π-periodic, i.e., $\theta \sim \theta + 2\pi$, because $\sum_{\hat{\tau},i} \mathcal{E} = -2\pi \sum_{\hat{\tau},i} m_\tau \in 2\pi\mathbb{Z}$. See [1] for similar Villain formulation of various standard and exotic $U(1)$ gauge theories.

The theory (3.47) can be viewed as an extension of the 1+1d rank-2 $U(1)$ gauge theory of [25] with Δ_x^2 on the 1d spatial lattice replaced by the discrete Laplacian operator Δ_L on the graph Γ.

3.4.1 Global symmetry

Here we discuss the global symmetry of the $U(1)$ Laplacian gauge theory. There are two kinds of global symmetries that we should distinguish in gauge theory. The first kind is the *space-like* global symmetry, which acts on operators and states in the Hilbert space. The second kind is the *time-like* global symmetry, which acts on defects extended in the (Euclidean) time direction. In an ordinary, relativistic gauge theory, both the space-like and time-like global symmetries are parts of the one-form global symmetry [28]. In contrast, the two global symmetries can be drastically different in non-relativistic systems. For example, even the groups for them can be different. We refer the readers to [25] for comprehensive analyses of time-like global symmetries in various standard and exotic models.

This theory has an electric symmetry that shifts

$$(\mathcal{A}_\tau, \mathcal{A}; m_\tau) \to (\mathcal{A}_\tau, \mathcal{A}; m_\tau) + (\lambda_\tau, \lambda; p_\tau) \ , \tag{3.48}$$

where $(\lambda_\tau, \lambda; p_\tau)$ is a flat $U(1)$ gauge field, i.e.,

$$\Delta_\tau \lambda - \Delta_L \lambda_\tau = 2\pi p_\tau \ . \tag{3.49}$$

The shift $(\lambda_\tau, \lambda; p_\tau)$ is subject to the gauge transformation in (3.46). Below we will use this freedom of gauge transformation to gauge-fix the shift in a particular form. Using the integer gauge parameter q, we can set $p_\tau = 0$ everywhere except at $\hat{\tau} = 0$. Similarly, we can use α to set $\lambda_\tau = 0$ everywhere except at $\hat{\tau} = 0$. Since $\sum_i p_\tau(\hat{\tau}, i) = 0$ by flatness, by the analysis in Appendix 3.A around (3.88), we can use q_τ to set

$$p_\tau(\hat{\tau}, i) = -\delta_{\hat{\tau},0} \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_{\tau a} \ , \tag{3.50}$$

where $p_{\tau a} = 0, \ldots, r_a - 1$. Now, the remaining gauge symmetry is time-independent $\alpha(i)$ and $q(i)$, and $q_\tau(\hat{\tau}, i) = \delta_{\hat{\tau},0} \bar{q}_\tau$, where $\bar{q}_\tau$ is an integer. By flatness, we have $\Delta_\tau \lambda = 0$. By the analysis in Appendix 3.A around (3.87), using $\alpha(i)$ and $q(i)$, we can set

$$\lambda(\hat{\tau}, i) = \frac{c}{\mathsf{N}} \ , \tag{3.51}$$

where $c \sim c + 2\pi$.

Since $\Delta_\tau \lambda = 0$, we have $\Delta_L \lambda_\tau + 2\pi p_\tau = 0$ at $\hat{\tau} = 0$. Using (3.10), the solution for λ_τ is

$$\lambda_\tau(\hat{\tau}, i) = \delta_{\hat{\tau},0} \left[c_\tau + 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_{\tau a}}{r_a} \right] \ , \tag{3.52}$$

where $c_\tau \sim c_\tau + 2\pi$.

The parameter c generates a $U(1)$ electric global space-like symmetry that shifts $\mathcal{A} \to \mathcal{A} + \frac{c}{\mathsf{N}}$. The operator charged under this symmetry is

$$\exp\left[i \sum_j \mathcal{A}(\hat{\tau}, j) \right] \ . \tag{3.53}$$

There are no other gauge invariant operators other than those mentioned so far. Relatedly, there is no discrete electric global symmetry.

The parameters c_τ and $p_{\tau a}$'s generate the electric global time-like symmetry that shifts

$$\begin{aligned} &\mathcal{A}_\tau(\hat{\tau}, i) \to \mathcal{A}_\tau(\hat{\tau}, i) + \delta_{\hat{\tau},0} \left[c_\tau + 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_{\tau a}}{r_a} \right] \ , \\ &m_\tau(\hat{\tau}, i) \to m_\tau(\hat{\tau}, i) - \delta_{\hat{\tau},0} \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_{\tau a} \ . \end{aligned} \tag{3.54}$$

Comparing with (3.10), we see that a general time-like global symmetry transformation is labeled by a $U(1)$-valued discrete harmonic function, which forms the group

$U(1) \times \mathrm{Jac}(\Gamma)$. This shift does not act on operators. Rather, it acts on defects that extend in the time direction. More specifically, a charged defect at site i is

$$\exp\left[i \sum_{\hat{\tau}} \mathcal{A}_\tau(\hat{\tau}, i)\right] . \tag{3.55}$$

3.4.2 Spectrum

Let us now determine the spectrum of the $U(1)$ Laplacian gauge theory (3.47) by working with continuous Lorentzian time while keeping the space discrete. To do this, we first take $L_\tau \to \infty$, and gauge fix $m_\tau(\hat{\tau}, i) = 0$. We then introduce the lattice spacing a_τ in the τ-direction, and take the limit $a_\tau \to 0$ while keeping $\gamma' = \gamma a_\tau$ fixed. We also define a scaled temporal gauge field $A_\tau = \frac{1}{a_\tau}\mathcal{A}_\tau$ and a scaled electric field $E = \partial_\tau \mathcal{A} - \Delta_L A_\tau$ while taking this limit. Finally, we Wick rotate from Euclidean time τ to Lorentzian time t.

The Gauss law (i.e., the equation of motion of A_0) gives

$$\Delta_L E(t, i) = 0 \implies E(t, i) = E(t) . \tag{3.56}$$

In the temporal gauge $A_0(t, i) = 0$, up to a time-independent gauge transformation, this equation is solved by

$$\mathcal{A}(t, i) = \frac{c(t)}{\mathsf{N}} , \tag{3.57}$$

where $c(t)$ is circle-valued, i.e., $c(t) \sim c(t) + 2\pi$. The effective action for $c(t)$ is

$$S = \int dt \, \left[\frac{\gamma'}{2\mathsf{N}}\dot{c}(t)^2 - \frac{\theta}{2\pi}\dot{c}(t)\right] . \tag{3.58}$$

The Hamiltonian is

$$H = \frac{\mathsf{N}}{2\gamma'}\left(\Pi + \frac{\theta}{2\pi}\right)^2 , \tag{3.59}$$

where Π is the conjugate momentum of $c(t)$, and $\Pi \in \mathbb{Z}$ because of the periodicity of $c(t)$. For $\theta \neq \pi$, the ground state is non-degenerate, while for $\theta = \pi$ there are two degenerate ground states.

We now discuss the robustness of the theory. All the local operators in the theory are made of the gauge-invariant electric field E. Adding them to the Lagrangian does not change the qualitative behavior of the theory. Hence, we conclude that the theory is robust.

3.4.3 Time-like symmetry, mobility of defects, and fractons

The pure $U(1)$ gauge theory has defects (3.55) describing the world lines of infinitely massive particles. The $U(1)$ time-like symmetry charges of these defects can be interpreted as the gauge charges of the massive particles. More generally, a static configuration of particles carrying gauge charge $q(i)$ at site i is represented by the following defect:

$$\exp\left[i\sum_{\hat{\tau}}\sum_{i} q(i)\mathcal{A}_\tau(\hat{\tau}, i)\right] \ , \qquad q(i) \in \mathbb{Z} \ . \tag{3.60}$$

A "move" at time $\hat{\tau}_0$ on a configuration is implemented by applying products of operators $\exp[i\mathcal{A}(\hat{\tau}_0, i)]$ at different sites. A configuration of particles can "move" to another configuration if and only if there is a gauge-invariant defect that connects the two.

Since the discrete $\mathrm{Jac}(\Gamma)$ time-like global symmetries depend on the sites, they constrain the possible shapes of defects, and therefore the mobility of the particles. Moreover, they lead to superselection sectors of defects distinguished by the time-like symmetry charges. See [25] for more discussions on time-like global symmetries.

From (3.54), we see that the discrete time-like symmetry charges of a static defect (3.55) at site i are given by $Q_{ia} \bmod r_a$ with $a = 1, 2, \cdots, \mathsf{N} - 1$. The defect (3.55) can hop from site i to site i' if and only if the time-like charges of the defects at these two positions are the same, i.e.,

$$Q_{i'a} = Q_{ia} \mod r_a \ , \qquad a = 1, \ldots, \mathsf{N} - 1 \ . \tag{3.61}$$

Then, the defect that "hops" a particle from i to i' at time $\hat{\tau} = \hat{\tau}_0$ is

$$\exp\left[i \sum_{\hat{\tau} < \hat{\tau}_0} \mathcal{A}_\tau(\hat{\tau}, i) \right] \exp\left[i \sum_{a < \mathsf{N}, j} \left(\frac{Q_{i'a} - Q_{ia}}{r_a} \right) P_{aj} \mathcal{A}(\hat{\tau}_0, j) \right] \exp\left[i \sum_{\hat{\tau} \geq \hat{\tau}_0} \mathcal{A}_\tau(\hat{\tau}, i') \right] \ . \tag{3.62}$$

We are now ready to phrase the mobility of the probe particles in terms of a graph-theoretic statement. The condition for mobility (3.61) is equivalent to the property that *all* $U(1)$-valued discrete harmonic functions (3.10) take the same value at i and i'. If the condition (3.61) is not satisfied, i.e., if there is a $U(1)$-valued discrete harmonic function that takes different values at i and i', then the particle cannot move from i to i'. If this is the case for all $i, i' \in \Gamma$, then the particle is a *fracton*.

Using the Abel-Jacobi map S_{i_0}, we can completely characterize the mobility of a particle by general properties of the graph Γ. (See Section 3.2 for the definition and properties of the Abel-Jacobi map S_{i_0}, where i_0 is some fixed vertex of the graph.)

- If there is a unique path from i to i', then $S_{i_0}(i) = S_{i_0}(i')$. By the universal property of S_{i_0}, for any $U(1)$-valued discrete harmonic function f that vanishes at i_0, there is a unique group homomorphism $\psi_f : \mathrm{Jac}(\Gamma) \to U(1)$ such that $f = \psi_f \circ S_{i_0}$. In particular, $f(i) = f(i')$ for any $U(1)$-valued discrete harmonic function $f \in \mathcal{H}(\Gamma, U(1))$, so there is no selection rule imposed by the time-like symmetry for a particle to move from i to i'. Indeed, Figure 3.3 illustrates how a particle can move from i to i' in this case.

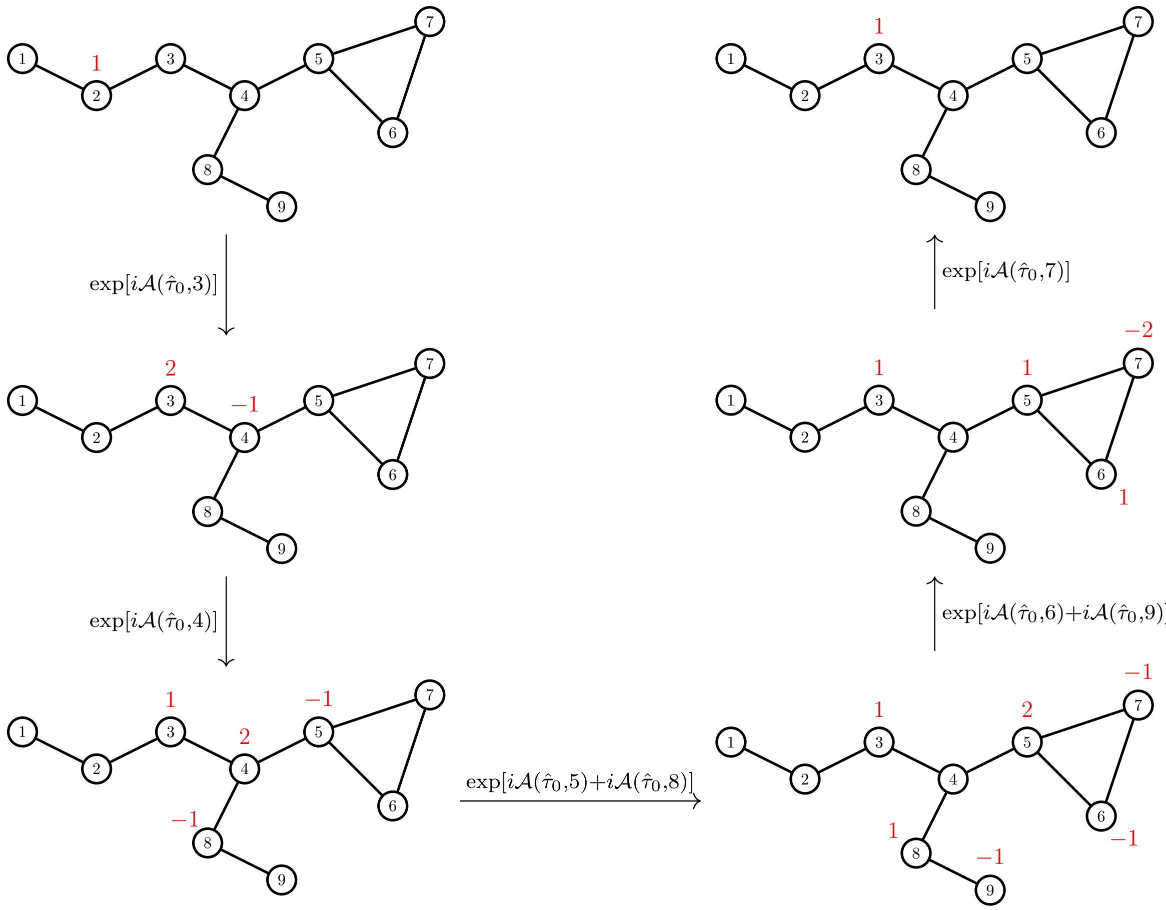


Figure 3.3: Motion of a particle between the vertices 2 and 3 that are connected by a unique path. In the graph, the numbers in black label the vertices, and the integers in red denote the charges of the particles (a zero charge is omitted). The operator $\exp[i\mathcal{A}(\hat{\tau}_0, i)]$ creates a particle of charge d_i at vertex i (where d_i is the degree of the vertex i) and particles of charge -1 at the neighbors of i at a fixed time $\hat{\tau}_0$. The above sequence of operators moves the particle from vertex 2 to vertex 3. In the correspondence between fractons and divisors, the integers in red represent values of the divisor (a zero value is omitted), and the operator $\exp[i\mathcal{A}(\hat{\tau}_0, i)]$ changes the divisor by a principal divisor.

We can also write down the defect that describes this motion. Let $i = i_0 \to i_1 \to \cdots \to i_{K-1} \to i_K = i'$ be the unique path from i to i'.[16] The following

[16]Note that the arrows do not imply that the edges are directed; they only indicate that the path is from i to i'.

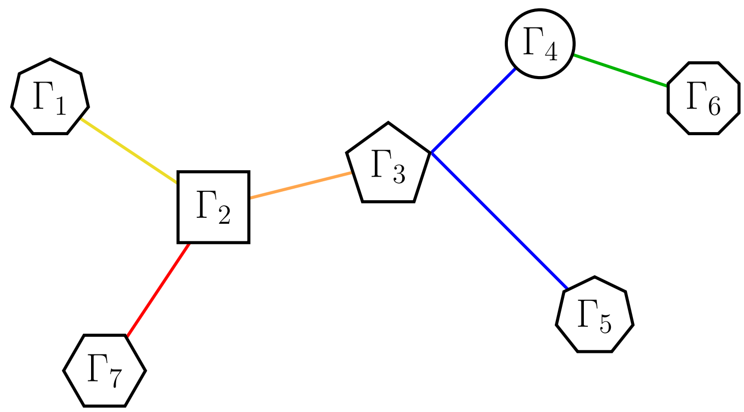


Figure 3.4: The structure of any graph Γ, where each Γ_k is a 2-edge connected graph, and each line represents a path/bridge between two vertices from adjacent Γ_k's. As explained in the main text, a particle on a bridge can move along that bridge or any other bridge that shares a vertex with that bridge but nowhere else, whereas a particle "inside" a Γ_k cannot move anywhere. In the figure, this means, particles on colored parts of Γ can move within the parts of same color.

defect describes the motion of a particle from i to i' at time $\hat{\tau} = \hat{\tau}_0$:

$$\exp\left[i \sum_{\hat{\tau}<\hat{\tau}_0} \mathcal{A}_\tau(\hat{\tau}, i)\right] \exp\left[i \sum_{k=1}^{K} \sum_{j\in\Gamma^{(k)}} \mathcal{A}(\hat{\tau}_0, j)\right] \exp\left[i \sum_{\hat{\tau}\geq\hat{\tau}_0} \mathcal{A}_\tau(\hat{\tau}, i')\right] , \tag{3.63}$$

where $\Gamma^{(k)}$ is the *connected component* of i_k in $\Gamma \setminus \langle i_{k-1}, i_k\rangle$ for $1 \leq k \leq K$.

- On the other hand, if there are at least two paths from i to i', then $S_{i_0}(i) \neq S_{i_0}(i')$. Moreover, there is a group homomorphism $\psi : \text{Jac}(\Gamma) \to U(1)$ such that $\psi(S_{i_0}(i)) \neq \psi(S_{i_0}(i'))$.[17] So, $\psi \circ S_{i_0}$ is a $U(1)$-valued discrete harmonic function that takes different values at i and i'. It follows that a particle cannot move from i to i'.

In particular, if Γ is 2-edge connected (such as a torus graph), any particle on Γ is a fracton. More generally, any Γ has the structure shown in Figure 3.4, where each

[17] Let g, h be two distinct elements in $\text{Jac}(\Gamma) = \prod_{a<\mathsf{N}} \mathbb{Z}_{r_a}$. Any element of $\text{Jac}(\Gamma)$ can be represented as an $(\mathsf{N}-1)$-tuple: $(s_1, \ldots, s_{\mathsf{N}-1})$ where $0 \leq s_a < r_a$. Say $a < \mathsf{N}$ is the first component at which g and h differ. Consider the group homomorphism $\psi : \text{Jac}(\Gamma) \to U(1)$ given by $\psi : (s_1, \ldots, s_{\mathsf{N}-1}) \mapsto e^{2\pi i s_a/r_a}$. It is clear that $\psi(g) \neq \psi(h)$.

Γ_k is a 2-edge connected graph, and the "lines" represent paths/bridges connecting two vertices from adjacent Γ_k's—for example, when each Γ_k is a single vertex, then Γ is a tree. By the last two paragraphs, a particle on a bridge can move along that bridge or any other bridge that shares a vertex with that bridge but nowhere else, whereas a particle "inside" a Γ_k cannot move anywhere.

3.4.4 Correspondence between divisors and fractons

In this subsection we unveil an intriguing correspondence between the theory of divisors in graph theory and the $U(1)$ Laplacian gauge theory of fractons. A divisor $q \in \mathcal{F}(\Gamma, \mathbb{Z})$ can be interpreted as a configuration of particles with the $U(1)$ time-like charge of the particle at a site i equal to $q(i) \in \mathbb{Z}$. More specifically, a static configuration of particles associated with a divisor q is represented by the defect (3.60). A configuration can "move" to another configuration if and only if their corresponding divisors differ by a principal divisor. In particular, a configuration associated with a principal divisor can "trivially move" by first being annihilated and then being created elsewhere.

These statements can be understood by the selection rules imposed by the time-like symmetry. A time-like superselection sector is labelled by the time-like charges of a configuration. Since the time-like symmetry group is $U(1) \times \mathrm{Jac}(\Gamma)$, the time-like charges are valued in the Pontryagin dual, i.e., $\mathbb{Z} \times \mathrm{Jac}(\Gamma)$.[18] The latter is precisely the Picard group $\mathrm{Pic}(\Gamma)$. In other words, the superselection sector of a configuration is associated with the equivalence class of a divisor in $\mathrm{Pic}(\Gamma)$. In particular, the "trivial" superselection sector (with trivial time-like charges) corresponds to the equivalence class of $\mathrm{Pic}(\Gamma)$ that consists of all the principal divisors. Therefore, a configuration can "move" to another configuration if and only if they are in the same superselection

[18]Here, the $\mathbb{Z}$ factor is the $U(1)$ time-like charge, which is associated with the degree of the divisor.

sector, i.e., if and only if their divisors differ by a principal divisor (see Figure 3.3 for an illustration). This correspondence is summarized in Table 3.1.

3.4.5 Examples

In this section, we discuss the $U(1)$ Laplacian gauge theory on 1d and 2d spatial tori.

1+1d rank-2 $U(1)$ tensor gauge theory

Let Γ be a *cycle graph*, i.e., $\Gamma = C_{L_x} = \mathbb{Z}_{L_x}$, where L_x is the number of sites in the cycle. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator Δ_x^2 in the x-direction. The action (3.47) becomes

$$S = \frac{\gamma}{2} \sum_{\tau\text{-link}} \mathcal{E}_{xx}^2 + \frac{i\theta}{2\pi} \sum_{\tau\text{-link}} \mathcal{E}_{xx} \ , \tag{3.64}$$

after replacing $\mathcal{A}$ with $\mathcal{A}_{xx}$. This is the modified Villain action of the 1+1d $U(1)$ dipole gauge theory or the rank-2 $U(1)$ tensor gauge theory [25].

The Smith decomposition of L is discussed in Section 3.3.6. It follows that the discrete electric time-like symmetry is $\mathbb{Z}_{L_x}$. Since Γ is 2-edge connected, a single particle cannot move, i.e., it is a fracton. On the other hand, dipoles can move. This is in agreement with the analysis of [25].

2+1d $U(1)$ Laplacian gauge theory

Let Γ be a *torus graph*, i.e., $\Gamma = C_{L_x} \times C_{L_y} = \mathbb{Z}_{L_x} \times \mathbb{Z}_{L_y}$, where L_i is the number of sites in the i direction. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator $\Delta_x^2 + \Delta_y^2$ in the xy-plane. The action (3.47) becomes

$$S = \frac{\gamma}{2} \sum_{\tau\text{-link}} \mathcal{E}^2 + \frac{i\theta}{2\pi} \sum_{\tau\text{-link}} \mathcal{E} \ , \tag{3.65}$$

where $\mathcal{E} = \Delta_\tau \mathcal{A} - (\Delta_x^2 + \Delta_y^2)\mathcal{A}_\tau - 2\pi m_\tau$. We refer to this as the 2+1d $U(1)$ Laplacian gauge theory. It is one of the natural higher dimensional versions of the 1+1d rank-2 $U(1)$ tensor gauge theory of Section 3.4.5.

While the Jacobian group of a 2d torus graph is not generally known, since Γ is 2-edge connected, a single particle is a fracton. In fact, in Appendix 3.B.2, we show that any finite set of particles with arbitrary charges is completely immobile (modulo sets of particles that can trivially move) on the infinite square lattice.

A closely related gauge theory is the 2+1d rank-2 $U(1)$ tensor gauge theory. While the 1+1d rank-2 $U(1)$ tensor gauge theory is the same as the 1+1d $U(1)$ Laplacian gauge theory, their 2+1d versions are very different. In a separate paper [127], we analyze both of them in detail, and compare their global properties.

3.A Discrete Poisson equation on a graph

In this appendix, we solve the discrete Poisson equation (3.7) using the Smith decomposition (3.9) of the Laplacian matrix.

Consider a function on the vertices of the graph, $f : \Gamma \to X$, where X is an abelian group. We are interested in the case where X is $\mathbb{R}$, $U(1)$, $\mathbb{Z}$, or $\mathbb{Z}_N$. Recall that the Laplacian operator Δ_L is defined as

$$\Delta_L f(i) = \sum_j L_{ij} f(j) = d_i f(i) - \sum_{j:\langle i,j\rangle \in \Gamma} f(j) = \sum_{j:\langle i,j\rangle \in \Gamma} [f(i) - f(j)] \,. \tag{3.66}$$

One is usually interested in the following questions: given a function g, is there a function f that satisfies the *discrete Poisson equation* on the graph,

$$\Delta_L f(i) = g(i) \text{ ?} \tag{3.67}$$

If yes, how many solutions are there, and what are they?

One way to answer these questions is to use the *Smith normal form* of L [143]. The Smith normal form of L is given by $R = PLQ$, where $P, Q \in GL_{\mathsf{N}}(\mathbb{Z})$, and $R = \mathrm{diag}(r_1, \ldots, r_{\mathsf{N}})$ is an $\mathsf{N} \times \mathsf{N}$ diagonal integer matrix with nonnegative diagonal entries such that $r_a | r_{a+1}$ (i.e., r_a divides r_{a+1}) for $a = 1, \ldots, \mathsf{N}-1$. While R is uniquely determined by L, the matrices P and Q are not.

Using the index notation, we can write the Smith normal form as $R_{ab} = \sum_{i,j} P_{ai} L_{ij} Q_{jb}$, where $R_{ab} = r_a \delta_{ab}$. While all the indices run from 1 to N, the indices i, j have natural interpretation as vertices of the graph Γ, whereas the indices a, b do not have an immediately obvious interpretation.

For a connected graph Γ, we have $r_a > 0$ for $a = 1, \ldots, \mathsf{N} - 1$, and $r_{\mathsf{N}} = 0$. This implies

$$\sum_i (P^{-1})_{ia} = 0 \ , \qquad \sum_i (Q^{-1})_{ai} = 0 \ , \tag{3.68}$$

for all $a < \mathsf{N}$. In fact, a convenient choice of P and Q is given by the block matrices:

$$P = \begin{pmatrix} \tilde{P} & \mathbf{0} \\ \mathbf{1}^T & 1 \end{pmatrix} \ , \qquad Q = \begin{pmatrix} \tilde{Q} & \mathbf{1} \\ \mathbf{0}^T & 1 \end{pmatrix} \ , \tag{3.69}$$

where $\tilde{P}, \tilde{Q} \in GL_{\mathsf{N}-1}(\mathbb{Z})$. It follows that

$$P^{-1} = \begin{pmatrix} \tilde{P}^{-1} & \mathbf{0} \\ -\mathbf{1}^T \tilde{P}^{-1} & 1 \end{pmatrix} \ , \qquad Q^{-1} = \begin{pmatrix} \tilde{Q}^{-1} & -\tilde{Q}^{-1}\mathbf{1} \\ \mathbf{0}^T & 1 \end{pmatrix} \ , \tag{3.70}$$

where (3.68) is manifest.

The discrete Poisson equation (3.67) simplifies in a new basis. Defining

$$\begin{aligned} f'_a &= \sum_i (Q^{-1})_{ai} f(i)\,, \\ g'_a &= \sum_i P_{ai} g(i), \end{aligned} \tag{3.71}$$

we can write (3.67) as

$$r_a f'_a = g'_a\ , \qquad a = 1, \ldots, \mathsf{N}\ . \tag{3.72}$$

Each a gives an independent equation, so we can solve for each f'_a independently. If a solution exists for all a, then a solution exists for (3.67), and vice versa.

Let us first focus on $a = \mathsf{N}$. Since $r_{\mathsf{N}} = 0$, a solution exists if and only if $g'_{\mathsf{N}} = \sum_i P_{\mathsf{N}i} g(i) = \sum_i g(i) = 0$. This condition is the same for any choice of X. If it is satisfied, then f'_{N} can take any value in X.[19] This corresponds to the zero mode (or constant mode because $Q_{i\mathsf{N}} = 1$ for all i) of the Laplacian.

Now, consider $a < \mathsf{N}$ so that $r_a > 0$. Let us analyze each X separately.

- When $X = \mathbb{R}$, there is always a unique solution

$$f'_a = \frac{1}{r_a} g'_a\ , \tag{3.73}$$

  Assuming $g'_{\mathsf{N}} = 0$, the solution in the original basis is

$$f(i) = \sum_{a<\mathsf{N}} \sum_j \frac{Q_{ia} P_{aj}}{r_a} g(j) + c\ , \tag{3.74}$$

  where c is a real constant (it is the zero mode mentioned above).

- When $X = U(1)$, the equation (3.72) can be written as

$$r_a f'_a = g'_a \mod 2\pi\ . \tag{3.75}$$

[19] Note that since $Q \in GL_{\mathsf{N}}(\mathbb{Z})$, if f takes values in X, then f' also takes values in X.

A general solution takes the form

$$f'_a = \frac{1}{r_a} g'_a + \frac{2\pi p_a}{r_a} \ , \tag{3.76}$$

where p_a is an integer. Since $X = U(1)$, the integers p_a and $p_a + r_a$ correspond to the same solution, so there are r_a inequivalent solutions associated with $p_a = 0, \ldots, r_a - 1$. Hence, assuming $g'_{\mathsf{N}} = 0 \mod 2\pi$, the general solution in the original basis is

$$f(i) = \sum_{a<\mathsf{N}} \sum_j \frac{Q_{ia} P_{aj}}{r_a} g(j) + 2\pi \sum_{a<\mathsf{N}} \frac{Q_{ia} p_a}{r_a} + c \ , \tag{3.77}$$

where $c \sim c + 2\pi$ is a circle-valued constant. If we lift f and g from $U(1)$ to $\mathbb{R}$, we have

$$\Delta_L f(i) = g(i) + 2\pi \sum_{a<\mathsf{N}} (P^{-1})_{ia} p_a \ . \tag{3.78}$$

- When $X = \mathbb{Z}$, a unique solution

$$f'_a = \frac{1}{r_a} g'_a \ , \tag{3.79}$$

exists if and only if r_a divides g'_a. So, assuming $g'_{\mathsf{N}} = 0$, and $g'_a = 0 \mod r_a$ for $a < \mathsf{N}$, the solution is

$$f(i) = \sum_{a<\mathsf{N}} \sum_j \frac{Q_{ia} P_{aj}}{r_a} g(j) + p \ , \tag{3.80}$$

where p is an integer.

- When $X = \mathbb{Z}_N$, the equation (3.72) can be written as

$$r_a f'_a = g'_a \mod N \ . \tag{3.81}$$

A solution of the form

$$f'_a = \frac{\tilde{r}_a}{\gcd(N, r_a)} g'_a + \frac{N p_a}{\gcd(N, r_a)} \ , \tag{3.82}$$

exists if and only if $\gcd(N, r_a)$ divides g'_a due to Bézout's identity. Here, $\tilde{r}_a$ is a fixed integer given by $\tilde{r}_a r_a = \gcd(N, r_a) \mod N$ (which exists by Bézout's identity), and p_a is an integer.[20] Note that $p_a \sim p_a + \gcd(N, r_a)$ because $X = \mathbb{Z}_N$, so there are $\gcd(N, r_a)$ inequivalent solutions associated with $p_a = 0, \ldots, \gcd(N, r_a) - 1$. So, assuming $g'_{\mathsf{N}} = 0 \mod N$, and $g'_a = 0 \mod \gcd(N, r_a)$ for $a < \mathsf{N}$, the general solution is

$$f(i) = \sum_{a<\mathsf{N}} \sum_j \frac{Q_{ia} \tilde{r}_a P_{aj}}{\gcd(N, r_a)} g(j) + \sum_{a<\mathsf{N}} \frac{N Q_{ia} p_a}{\gcd(N, r_a)} + p \ , \tag{3.83}$$

where p is an integer modulo N. Since $\gcd(N, r_{\mathsf{N}}) = \gcd(N, 0) = N$, and $Q_{i\mathsf{N}} = 1$, defining $p_{\mathsf{N}} = p \mod N$, we can write the above solution as

$$f(i) = \sum_{a<\mathsf{N}} \sum_j \frac{Q_{ia} \tilde{r}_a P_{aj}}{\gcd(N, r_a)} g(j) + \sum_a \frac{N Q_{ia} p_a}{\gcd(N, r_a)} \ , \tag{3.84}$$

which exists if and only if $g'_a = 0 \mod \gcd(N, r_a)$ for all a.

There is another useful perspective to the above analysis. Say g is a gauge field on Γ, and f is its gauge parameter with gauge symmetry

$$g(i) \sim g(i) - \Delta_L f(i) \ . \tag{3.85}$$

In other words, g represents an equivalence class in $\operatorname{coker}_X \Delta_L$. We are interested in the gauge invariant information in g, i.e., the *holonomies* of g. Alternatively, we can

[20] There are infinitely many choices of $\tilde{r}_a$ but they can be absorbed into p_a.

ask what part of g can be gauge-fixed to zero, i.e., if there is an f satisfying (3.67). From the above analysis, the answer is the following:

- When $X = \mathbb{R}$, the only holonomy of g is $g'_{\mathsf{N}} = \sum_i g(i)$. If $g'_{\mathsf{N}} = c$, then we can use the gauge freedom to gauge-fix g to be of the form

$$g(i) = (P^{-1})_{i\mathsf{N}} g'_{\mathsf{N}} = \delta_{i\mathsf{N}} c \ , \qquad \text{or} \qquad g(i) = \frac{c}{\mathsf{N}} \ . \tag{3.86}$$

  Here, c is a real constant.

- When $X = U(1)$, the only holonomy of g is $g'_{\mathsf{N}} = \sum_i g(i) \mod 2\pi$. If $g'_{\mathsf{N}} = c$, then we can use the gauge freedom to gauge-fix g to be of the form

$$g(i) = (P^{-1})_{i\mathsf{N}} g'_{\mathsf{N}} = \delta_{i\mathsf{N}} c \ , \qquad \text{or} \qquad g(i) = \frac{c}{\mathsf{N}} \ . \tag{3.87}$$

  Here, $c \sim c + 2\pi$ is a circle-valued constant.

- When $X = \mathbb{Z}$, the holonomies of g are $g'_{\mathsf{N}} = \sum_i g(i)$, and $g'_a = \sum_i P_{ai} g(i) \mod r_a$ for $a < \mathsf{N}$. If $g'_{\mathsf{N}} = p_{\mathsf{N}}$ and $g'_a = p_a \mod r_a$, then we can use the gauge freedom to gauge-fix g to be of the form

$$g(i) = \sum_a (P^{-1})_{ia} p_a \ . \tag{3.88}$$

  Here, p_{N} is an integer, and $p_a = 0, \ldots, r_a - 1$ for $a < \mathsf{N}$. When $p_{\mathsf{N}} = 0$, each function in (3.88) represents a unique equivalence class of $\mathrm{Jac}(\Gamma)$. Therefore,

$$\mathrm{Jac}(\Gamma) \cong \prod_{a<\mathsf{N}} \mathbb{Z}_{r_a} \ . \tag{3.89}$$

- When $X = \mathbb{Z}_N$, the holonomies of g are $g'_{\mathsf{N}} = \sum_i g(i) \mod N$, and $g'_a = \sum_i P_{ai} g(i) \mod \gcd(N, r_a)$ for $a < \mathsf{N}$. Since $r_{\mathsf{N}} = 0$, we can combine them into

$g'_a = \sum_i P_{ai} g(i) \mod \gcd(N, r_a)$ for all a. If $g'_a = p_a \mod \gcd(N, r_a)$, then we can use the gauge freedom to gauge-fix g to be of the form

$$g(i) = \sum_a (P^{-1})_{ia} p_a \ . \tag{3.90}$$

Here, $p_a = 0, \ldots, \gcd(N, r_a) - 1$ for all a.

3.B Polynomial representation of functions on square lattice

In this appendix, we develop a polynomial representation of functions on the infinite square lattice $\mathbb{Z}^2$, and use it to show the following:

1. The naturalness of the action (3.42) with respect to the global symmetry of the 2+1d Laplacian ϕ-theory of Section 3.3.6. More precisely, we show that the local operator $\prod_{i=1}^n e^{iq_i\phi(0,x_i,y_i)}$ is invariant under the momentum symmetry if and only if it can be written as $\prod_{j=1}^m e^{ir_j\Delta_L\phi(0,x_j,y_j)}$, where $q_i, r_j \in \mathbb{Z}$ and $\Delta_L := \Delta_x^2 + \Delta_y^2$ is the discrete Laplacian operator. Of course the winding operator $e^{i\tilde{\phi}}$ is invariant under the momentum symmetry, and it is relevant because it acts nontrivially on the ground states, so the action (3.42) is not natural unless we impose the winding symmetry too.

2. The immobility of a finite set of defects with arbitrary charges, unless they can be "locally annihilated," in the 2+1d $U(1)$ Laplacian gauge theory of Section 3.4.5.

The polynomial representation was originally developed in the context of translationally invariant Pauli stabilizer codes [156].

On an infinite square lattice $\mathbb{Z}^2$, any function f can be associated with a formal *Laurent power series* in the variables X, Y:

$$\hat{f}(X,Y) = \sum_{(x,y)\in\mathbb{Z}^2} f(x,y)X^{-x}Y^{-y} \ . \tag{3.91}$$

We can think of $X = e^{ip_x}$ and $Y = e^{ip_y}$ as phases with p_x and p_y momenta conjugate to x and y. Then, this definition of $\hat{f}(X,Y)$ is the discrete Fourier transform of $f(x,y)$. Related to that, X and Y are generators of lattice translations in the x and y directions:

$$X\hat{f}(X,Y) = \sum_{(x,y)\in\mathbb{Z}^2} f(x+1,y)X^{-x}Y^{-y} \ , \tag{3.92}$$

and similarly for Y. Then, the difference operator Δ_x is associated with $X-1$ because

$$(X-1)\hat{f}(X,Y) = \sum_{(x,y)\in\mathbb{Z}^2} \Delta_x f(x+\tfrac{1}{2},y)X^{-x}Y^{-y} \ . \tag{3.93}$$

Recall that $\Delta_x f(x+\frac{1}{2},y) = f(x+1,y) - f(x,y)$.

More generally, any *local* difference operator is associated with a *Laurent polynomial* $s(X,Y)$ with integer coefficients, i.e., an element of $\mathbb{Z}[X,X^{-1},Y,Y^{-1}]$ satisfying $s(1,1) = 0$. (Here, $\mathbb{Z}[X,Y,\ldots]$ is the set of polynomials in $X,Y,\ldots$ with integer coefficients, and therefore $\mathbb{Z}[X,X^{-1},Y,Y^{-1},\ldots]$ is the set of Laurent polynomials in $X,Y,\ldots$ with integer coefficients.)

For example, the discrete Laplacian operator $\Delta_L := \Delta_x^2 + \Delta_y^2$ corresponds to the Laurent polynomial

$$p(X,Y) = (X-2+X^{-1}) + (Y-2+Y^{-1}) \ . \tag{3.94}$$

We can equivalently work with

$$\tilde{p}(X,Y) = XYp(X,Y) = Y(X-1)^2 + X(Y-1)^2 \ , \tag{3.95}$$

which is simply a translated version of Δ_L. Note that $\tilde{p}(X,Y) \in \mathbb{Z}[X,Y]$, i.e., $\tilde{p}(X,Y)$ is a polynomial, whereas $p(X,Y)$ is a Laurent polynomial. Indeed, we can always translate a difference operator so that the associated Laurent polynomial is a polynomial.[21]

Let us define a *lexicographic monomial order*, $X \succ Y$, on $\mathbb{Z}[X,Y]$. This means we can compare any two monomials as follows: $X^m Y^n \succ X^k Y^l$ if $m > k$, or $m = k$ and $n > l$. Clearly, this is a total order on all monomials in X, Y. Given a nonzero polynomial, its *leading term* is the term with the largest monomial among all its terms. The corresponding coefficient and monomial are called *leading coefficient* and *leading monomial* respectively. If the leading coefficient is ± 1, the polynomial is said to be *monic*.

We say a polynomial $a(X,Y)$ is *reducible* by another polynomial $b(X,Y)$ if some term of $a(X,Y)$ is a multiple of the leading term of $b(X,Y)$. Furthermore, if $b(X,Y)$ is monic, then $a(X,Y)$ can be written uniquely as

$$a(X,Y) = c(X,Y)b(X,Y) + d(X,Y) \ , \tag{3.96}$$

where $c(X,Y)$ is the *quotient* and $d(X,Y)$ is the *remainder*, which are uniquely determined by demanding that $d(X,Y)$ is not reducible by $b(X,Y)$. This operation is known as *multivariate division* with respect to a given monomial order.

3.B.1 Naturalness of 2+1d Laplacian ϕ-theory

In this appendix, we show that the action (3.42) is natural with respect to the global symmetry of the 2+1d Laplacian ϕ-theory.

[21] In the continuum, a differential operator in space becomes in momentum space a multiplication by a polynomial in the momenta. On the lattice, we follow the interpretation of X and Y as lattice translation generators, i.e., $X = e^{ip_x}$, $Y = e^{ip_y}$, and then a difference operator becomes a polynomial in X and Y.

Usually, the notion of naturalness assumes that a certain global symmetry is imposed on the system and then all the relevant operators in the action respect this symmetry [154]. (See [35] for a more recent discussion comparing the notions of naturalness and robustness.) For this to make sense, we need some scaling property, which determines which terms in the action should be viewed as relevant, and which terms should be viewed as irrelevant. In the lattice system, without a continuum limit, there is no such obvious scaling. Instead, we show that every term that respects the symmetry can be expressed in terms of lattice derivatives acting on other terms that are already present in the action. More precisely, we will show that every term invariant under the momentum symmetry can be expressed in terms of gauge invariant functions of $\Delta_L \phi$ and lattice derivatives of them. Then, we will exclude more terms using the winding symmetry. See more details below.

In the continuum, the conclusion of this appendix is the following trivial statement. A differential operator $\mathcal{D}$ that annihilates every real harmonic function on $\mathbb{R}^2$, $f(x, y)$ includes the Laplacian as a factor. To see that, use holomorphic coordinates and write $f = g(z) + \bar{g}(\bar{z})$. Then, $\mathcal{D} f = 0$ means that $\mathcal{D}$ must include a factor of ∂_z and using the reality, it should also have a factor of $\partial_{\bar{z}}$. Therefore, $\mathcal{D} = \mathcal{D}' \partial_z \partial_{\bar{z}}$ with a differential operator $\mathcal{D}'$.

The momentum symmetry of the 2+1d Laplacian ϕ-theory on the square lattice includes shifts by real-valued discrete harmonic functions $f(x, y)$ on $\mathbb{Z}^2$. We would like to find other terms invariant under this symmetry. We look for terms depending on $D\phi$ with some difference operator D. Let $\mathcal{H}(\mathbb{Z}^2, \mathbb{R})$ be the set of all real-valued discrete harmonic functions. By definition, any $f \in \mathcal{H}(\mathbb{Z}^2, \mathbb{R})$ satisfies

$$\Delta_L f(x, y) = 0 \iff p(X, Y) \hat{f}(X, Y) = 0 \ . \tag{3.97}$$

We would like to find the condition that the difference operator D should satisfy such that $Df(x,y)=0$ for all $f\in\mathcal{H}(\mathbb{Z}^2,\mathbb{R})$.[22]

One trivial possibility is $D=D'\circ\Delta_L$ because $(D'\circ\Delta_L)f(x,y)=0$ for any operator D'. This means, we could add to the action (3.42) a term of the form

$$-\sum_{\text{site}}\cos[(D'\circ\Delta_L)\phi]\ , \tag{3.98}$$

and preserve the global symmetry. This is considered a higher-order term than Δ_L and is always compatible with the momentum global symmetry.

A more interesting possibility would be a D that cannot be written as $D'\circ\Delta_L$, and yet $Df(x,y)=0$ for all $f\in\mathcal{H}(\mathbb{Z}^2,\mathbb{R})$. Below, we will show that this is impossible. Equivalently, we show that any D that satisfies $Df(x,y)=0$ for all $f\in\mathcal{H}(\mathbb{Z}^2,\mathbb{R})$ is of higher order than Δ_L. This implies that the action (3.42) is natural with respect to the global momentum symmetry of the 2+1d Laplacian ϕ-theory.

Let us rephrase the above problem in terms of polynomials. Let $q(X,Y)$ be the Laurent polynomial associated with D. By an appropriate translation, we can write $X^mY^nq(X,Y)=\tilde{q}(X,Y)$, where $\tilde{q}(X,Y)$ is a polynomial. If there is a polynomial $\tilde{r}(X,Y)$ such that $\tilde{q}(X,Y)=\tilde{r}(X,Y)\tilde{p}(X,Y)$, then D is of higher order than Δ_L, i.e., $D=D'\circ\Delta_L$, where D' is the operator associated with $X^aY^b\tilde{r}(X,Y)$ for some $a,b\in\mathbb{Z}$.

With the above preparations, the central result of this appendix can be stated in terms of polynomials as follows: suppose $\tilde{q}(X,Y)$ is a polynomial such that

$$\tilde{q}(X,Y)\hat{f}(X,Y)=0\ ,\quad \forall f\in\mathcal{H}(\mathbb{Z}^2,\mathbb{R})\ , \tag{3.99}$$

[22] We should impose $Df=0$ rather than the weaker condition $Df(x,y)=0 \mod 2\pi$ because it should be true for $cf(x,y)$ for any $c\in\mathbb{R}$.

then $\tilde{q}(X,Y) = \tilde{r}(X,Y)\tilde{p}(X,Y)$ for some $\tilde{r}(X,Y) \in \mathbb{Z}[X,Y]$, where $\tilde{p}(X,Y)$ is the polynomial (3.95) associated with the discrete Laplacian operator Δ_L.

More specifically, since $\tilde{p}(X,Y)$ is monic with leading term X^2Y, by multivariate division with respect to lexicographic order, $\tilde{q}(X,Y)$ can be written uniquely as

$$\tilde{q}(X,Y) = X^2\alpha(X) + X\beta(Y) + \gamma(Y) + \tilde{r}(X,Y)\tilde{p}(X,Y) \ , \tag{3.100}$$

where $\alpha(X) \in \mathbb{Z}[X]$, $\beta(Y), \gamma(Y) \in \mathbb{Z}[Y]$, and $\tilde{r}(X,Y) \in \mathbb{Z}[X,Y]$. The above statement is then equivalent to showing that $\alpha(X) = \beta(Y) = \gamma(Y) = 0$ if (3.99) is obeyed, which means that D is of higher order than Δ_L.

We parameterize the polynomials as

$$\alpha(X) = \sum_{i=0}^{u} a_i X^i \ , \qquad \beta(Y) = \sum_{j=0}^{v} b_j Y^j \ , \qquad \gamma(Y) = \sum_{k=0}^{w} c_k Y^k \ , \tag{3.101}$$

with nonnegative integers u, v, and w. Then, we apply (3.100) to a specific set of discrete harmonic functions parameterized by t:[23]

$$\begin{aligned} &f_t(x,y) \equiv X_t^x Y_t^y \ , \\ &X_t \equiv -t\left(\frac{1+t}{1-t}\right) \ , \qquad Y_t \equiv t\left(\frac{1-t}{1+t}\right) \ . \end{aligned} \tag{3.102}$$

(It is easy to check that $f_t \in \mathcal{H}(\mathbb{Z}^2,\mathbb{R})$ for any $t \neq 0, \pm 1$, i.e., $\tilde{p}(X,Y)\hat{f}_t(X,Y) = 0$.) Then, using (3.100) and (3.99), we get

$$X_t^2\alpha(X_t) + X_t\beta(Y_t) + \gamma(Y_t) = 0 \ . \tag{3.103}$$

[23] It is related to the *discrete exponential function* on the square lattice $\mathbb{Z}^2$ [157, 158], which is the discrete analog the exponential function on the complex plane $\mathbb{C} \cong \mathbb{R}^2$.

Next, we turn this into a polynomial in t by multiplying it by $(1-t)^{u+2}(1+t)^{\nu}$ with $\nu = \max(v-1, w)$

$$\sum_{i=0}^{u} a_i \alpha_i(t; u, \nu) + \sum_{j=0}^{v} b_j \beta_j(t; u, \nu) + \sum_{k=0}^{w} c_k \gamma_k(t; u, \nu) = 0 \ , \tag{3.104}$$

where

$$\begin{aligned}
\alpha_i(t; u, \nu) &= (1-t)^{u+2}(1+t)^{\nu} X_t^{i+2} = (-t)^{i+2}(1+t)^{\nu+i+2}(1-t)^{u-i} \ , \\
\beta_j(t; u, \nu) &= (1-t)^{u+2}(1+t)^{\nu} X_t Y_t^{j} = -t^{j+1}(1+t)^{\nu-j+1}(1-t)^{u+j+1} \ , \\
\gamma_k(t; u, \nu) &= (1-t)^{u+2}(1+t)^{\nu} Y_t^{k} = t^{k}(1+t)^{\nu-k}(1-t)^{u+k+2} \ .
\end{aligned} \tag{3.105}$$

Since the equation (3.104) holds for any $t \neq 0, \pm 1$, the polynomial in (3.104) must vanish identically, even at $t = 0, \pm 1$. What can we say about the coefficients a_i's, b_j's, and c_k's then?

For fixed (u, v, w), we have a set $\mathcal{P}(u, v, w)$ of polynomials in t, $\{\alpha_i : i = 0, \ldots, u\} \cup \{\beta_j : j = 0, \ldots, v\} \cup \{\gamma_k : k = 0, \ldots, w\}$. We will show that these polynomials are linearly independent, and therefore, $\alpha(X) = \beta(Y) = \gamma(Y) = 0$ and $\tilde{q}(X, Y) = \tilde{r}(X, Y)\tilde{p}(X, Y)$.

First, note that for $v \leq v_0$ and $w \leq v_0 - 1$, we have $\mathcal{P}(u, v_0, w) \subseteq \mathcal{P}(u, v_0, v_0 - 1)$ and $\mathcal{P}(u, v, v_0 - 1) \subseteq \mathcal{P}(u, v_0, v_0 - 1)$ because $\max(v - 1, v_0 - 1) = \max(v_0 - 1, w)$. So it suffices to show that the polynomials in the set $\mathcal{P}(u, v_0, v_0 - 1)$ are linearly independent for all $u \geq 0$ and $v_0 \geq 1$. We proceed by induction:

- <u>Base case</u>: It is easy to check that the set $\mathcal{P}(0, 1, 0)$ is linearly independent, and hence $\mathcal{P}(0, 0, 0)$ is also linearly independent.

- Induction step: Assume that $\mathcal{P}(u, v_0, v_0 - 1)$ is linearly independent. Consider $\mathcal{P}(u + 1, v_0, v_0 - 1)$:

$$\begin{aligned}
\alpha_i(t; u+1, v_0 - 1) &= \begin{cases} (1-t)\alpha_i(t; u, v_0 - 1) \ , & \text{for } i = 0, \dots, u \ , \\ (-t)^{u+3}(1+t)^{v_0+u+2} \ , & \text{for } i = u+1 \ , \end{cases} \\
\beta_j(t; u+1, v_0 - 1) &= (1-t)\beta_j(t; u, v_0 - 1) \ , \quad \text{for } j = 0, \dots, v_0 \ , \\
\gamma_k(t; u+1, v_0 - 1) &= (1-t)\gamma_k(t; u, v_0 - 1) \ , \quad \text{for } k = 0, \dots, v_0 - 1 \ .
\end{aligned} \tag{3.106}$$

The polynomials in the first, third, and fourth lines are linearly independent by the induction hypothesis. The second line is nonzero at $t = 1$, whereas the other three lines vanish at $t = 1$, so the second line is independent of the other polynomials. Thus, $\mathcal{P}(u + 1, v_0, v_0 - 1)$ is linearly independent.

Now consider $\mathcal{P}(u, v_0 + 1, v_0)$:

$$\begin{aligned}
\alpha_i(t; u, v_0) &= (1+t)\alpha_i(t; u, v_0 - 1) \ , \quad \text{for } i = 0, \dots, u \ , \\
\beta_j(t; u, v_0) &= \begin{cases} (1+t)\beta_j(t; u, v_0 - 1) \ , & \text{for } j = 0, \dots, v_0 \ , \\ -t^{v_0+2}(1-t)^{u+v_0+2} \ , & \text{for } j = v_0 + 1, \end{cases} \\
\gamma_k(t; u, v_0) &= \begin{cases} (1+t)\gamma_k(t; u, v_0 - 1) \ , & \text{for } k = 0, \dots, v_0 - 1 \ , \\ t^{v_0}(1-t)^{u+v_0+2} \ , & \text{for } k = v_0, \end{cases}
\end{aligned} \tag{3.107}$$

The polynomials in the first, second, and fourth lines are linearly independent by the induction hypothesis. Those in the third and fifth lines are linear independent of the other polynomials because they do not vanish at $t = -1$, and of each other because they have different degrees. Thus, $\mathcal{P}(u, v_0 + 1, v_0)$ is linearly independent.

Therefore, $\mathcal{P}(u,v,w)$ is linearly independent for any (u,v,w). Since the polynomial in t in (3.104) must vanish identically, it follows that $a_i = b_j = c_k = 0$, so $\alpha(X) = \beta(Y) = \gamma(Y) = 0$. Hence, $\tilde{q}(X,Y) = \tilde{r}(X,Y)\tilde{p}(X,Y)$, which is what we set out to show.

It follows that any difference operator D (which is associated with the polynomial $\tilde{q}(X,Y)$) respecting the momentum global symmetry must be of higher order than Δ_L, i.e., $D = D' \circ \Delta_L$.

Next, we impose also the winding symmetry. This excludes terms such as $\cos\tilde{\phi}$. Using an argument similar to the one above, it is easy to see that the model is also natural with respect to the winding symmetry. One way to see that is to first dualize the theory and apply the argument above with $\phi \leftrightarrow \tilde{\phi}$. We conclude that the action (3.42) is natural if we impose its entire global symmetry.

3.B.2 Mobility of defects in 2+1d $U(1)$ Laplacian gauge theory

In this appendix, we prove the immobility of any finite set of defects with arbitrary charges (except in some trivial cases) in the 2+1d $U(1)$ Laplacian gauge theory on the infinite square lattice $\mathbb{Z}^2$.

Before proceeding, let us distinguish between two kinds of defects that capture the motion of a particle. Typically, when a particle can move between two points, there is an operator supported in a small region, e.g., a line joining the two points. However, there are also situations where the operator that moves the particle can have a support spanning $O(L_x)$ or $O(L_y)$ sites even though the two points are separated by a much smaller distance. (See [25] for a discussion and examples of both kinds of operators.) The existence of the latter kind of operators depends on the number-theoretic properties of L_i, whereas the former kind of operators exist for all L_i. In particular, only the former make sense on the infinite square lattice.

Consider the defect

$$\exp\left[i\sum_{\tau}\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)\right]\ , \tag{3.108}$$

which describes the worldline of a single particle with unit charge. Under the time-like symmetry that shifts

$$\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)\to\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)+\delta_{\tau,0}\left(\frac{2\pi m_x x}{L_x}+\frac{2\pi m_y y}{L_y}\right)\ ,\qquad m_x,m_y\in\mathbb{Z}\ , \tag{3.109}$$

the defect (3.108) acquires an (x,y)-dependent phase, so it cannot bend. In other words, the particle is completely immobile, i.e., it is a fracton.

Since the time-like symmetry in (3.109) is present also in the scalar charge theory, the same conclusion holds there. The selection rules from the time-like global symmetries give a more precise explanation of the intuitive "dipole moment conservation" discussed in [23, 116, 56]. (See [25] for a discussion.)

Next, consider the defect

$$\exp\left(i\sum_{\tau}\left[\mathcal{A}_\tau(\tau+\tfrac{1}{2},x+x_0,y+y_0)-\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)\right]\right)\ , \tag{3.110}$$

which describes the worldline of a dipole of particles with charges ±1 with separation (x_0,y_0). The shift (3.109) imposes the constraint that the defect cannot move unless the separation is held fixed. This is the only constraint in the scalar charge theory, and as long as it is met, the dipole can move. There are additional constraints in the Laplacian gauge theory. Indeed, under the time-like symmetry that shifts (for simplicity, we set $L_x=L_y=L$)

$$\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)\to\mathcal{A}_\tau(\tau+\tfrac{1}{2},x,y)+\delta_{\tau,0}\left[\frac{2\pi mxy}{L}+\frac{2\pi m'(x^2-y^2-Lx+Ly)}{2L}\right]\ , \tag{3.111}$$

where $m, m' \in \mathbb{Z}$, the defect (3.110) acquires an (x, y)-dependent phase, so it cannot bend. In other words, the dipole is also completely immobile.

More generally, consider the defect

$$\exp\left[i \sum_{\tau} \sum_{i=1}^{n} q_i \mathcal{A}_\tau(\tau + \tfrac{1}{2}, x_i, y_i)\right] . \tag{3.112}$$

which describes the world-lines of n particles labelled by $i = 1, \ldots, n$, with positions (x_i, y_i), and charges q_i. It is difficult to analyze this case in full generality on a finite lattice, so we limit ourselves to an infinite square lattice.

Under the shift of $\mathcal{A}_\tau$ by a discrete harmonic function $f \in \mathcal{H}(\mathbb{Z}^2, \mathbb{R})$ at a fixed time $\tau = 0$, the phase acquired by the defect (3.112) is $\exp\left[i \sum_{i=1}^{n} q_i f(x_i, y_i)\right]$. The defect carries trivial time-like charges (i.e., it is in the trivial superselection sector) if and only if for all discrete harmonic functions $f \in \mathcal{H}(\mathbb{Z}^2, \mathbb{R})$

$$\sum_{i=1}^{n} q_i f(x_i, y_i) = 0 \qquad \Longleftrightarrow \qquad q(X, Y)\hat{f}(X, Y) = 0 . \tag{3.113}$$

(Once again, we cannot impose the weaker condition $\sum_{i=1}^{n} q_i f(x_i, y_i) = 0 \mod 2\pi$ because this equation should be true even for $cf(x, y)$ for any $c \in \mathbb{R}$.) As we showed in Section 3.B.1, this is possible if and only if $q(X, Y) = r(X, Y)p(X, Y)$ for some Laurent polynomial $r(X, Y)$.

To see the physical meaning of being invariant under the time-like symmetry, assume that such an $r(X, Y)$ exists. Then, we can construct the following defect

$$\exp\left[i \sum_{\tau<0} \sum_{i=1}^{n} q_i \mathcal{A}_\tau(\tau + \tfrac{1}{2}, x_i, y_i)\right] \times \exp\left[-i \sum_{j=1}^{m} r_j \mathcal{A}(0, x_j, y_j)\right] . \tag{3.114}$$

Here, r_j and (x_j, y_j) are obtained from $r(X, Y) = \sum_{j=1}^{m} r_j X^{x_j} Y^{y_j}$. This defect describes annihilation of the n particles at time $\tau = 0$. To see that this defect is gauge invariant, observe that, under a gauge transformation, the exponent transforms as

$$\sum_{i=1}^{n} q_i \alpha(0, x_i, y_i) - \sum_{j=1}^{m} r_j \Delta_L \alpha(0, x_j, y_j) \ . \tag{3.115}$$

This is the coefficient of $X^0 Y^0$ term in $[q(X,Y) - r(X,Y)p(X,Y)]\hat{\alpha}(X,Y)$, and so it vanishes.[24]

In fact, we can write the defect (3.114) as

$$\begin{aligned} &\exp\left[i \sum_{\tau<0} \sum_{j=1}^{m} r_j \Delta_L \mathcal{A}_\tau(\tau + \tfrac{1}{2}, x_j, y_j)\right] \times \exp\left[-i \sum_{j=1}^{m} r_j \mathcal{A}(0, x_j, y_j)\right] \\ &= \prod_{j=1}^{m} \exp\left[i r_j \sum_{\tau<0} \Delta_L \mathcal{A}_\tau(\tau + \tfrac{1}{2}, x_j, y_j) - i r_j \mathcal{A}(0, x_j, y_j)\right] \ . \end{aligned} \tag{3.116}$$

Each factor here describes particles being annihilated "locally" because the operator that annihilates them is local.

The result in (3.116) can be understood intuitively as follows. In this special case, the collection of defects coming from the past (3.112) can be expressed as a "total spatial derivative" using the Laplacian as in the first factor in (3.116). In this form, each term with the Laplacian can end using the local operator made out of $\mathcal{A}$ in the second factor in (3.116). The result, in this case, is that the collection of defects can be annihilated by an operator at time $\tau = 0$.[25]

Let us now examine the mobility of the n particles described by the defect (3.112). We stress that we consider mobility only under the restriction that the charges of the particles and the separations between them remain fixed. (Relaxing these two restrictions can lead to more possibilities, which we will not discuss here.) Then, we

[24] Here, $\hat{\alpha}(X,Y)$ is the formal Laurent power series associated with the gauge parameter $\alpha(x,y)$.

[25] This is analogous to the following very well known fact in standard $U(1)$ gauge theories. A dipole of particles with charges ± 1 is represented by the defect $\exp(i \int d\tau\ [A_\tau(\tau,x) - A_\tau(\tau,0)]) = \exp[i \int d\tau \int_0^x dx'\ \partial_x A_\tau(\tau,x')]$. It can end at $\tau = 0$, as described by $\exp(i \int_{-\infty}^0 d\tau\ [A_\tau(\tau,x) - A_\tau(\tau,0)]) \times \exp[-i \int_0^x dx'\ A_x(0,x')]$.

say that the n particles can move by $(x_0, y_0) \neq (0,0)$ if there is a defect of the form

$$\begin{aligned}
&\exp\left[i\sum_{\tau<0}\sum_{i=1}^{n} q_i\mathcal{A}_\tau(\tau+\tfrac{1}{2}, x_i, y_i)\right] \times \exp\left[i\sum_{k=1}^{l} s_k\mathcal{A}(0, x_k, y_k)\right] \\
&\quad\times \exp\left[i\sum_{\tau\geq 0}\sum_{i=1}^{n} q_i\mathcal{A}_\tau(\tau+\tfrac{1}{2}, x_i+x_0, y_i+y_0)\right] .
\end{aligned} \tag{3.117}$$

This defect exists, i.e., it is gauge invariant, if and only if

$$(X^{x_0}Y^{y_0} - 1)q(X,Y) = s(X,Y)p(X,Y) \ , \tag{3.118}$$

where $s(X,Y) = \sum_{k=1}^{l} s_k X^{x_k} Y^{y_k}$. Equivalently, this is precisely the condition for which the time-like charges of the n particles remain unchanged after displacing them by (x_0, y_0).

If $q(X,Y)$ is a multiple of $p(X,Y)$, i.e., $q(X,Y) = r(X,Y)p(X,Y)$ for some Laurent polynomial $r(X,Y)$, then we can always choose $s(X,Y) = r(X,Y)(X^{x_0}Y^{y_0} - 1)$ so that (3.118) is satisfied. However, this is not an interesting situation because, when $q(X,Y) = r(X,Y)p(X,Y)$, the defect (3.112) has trivial time-like charges as explained around (3.113). Consequently, similar to the discussion around (3.116), this situation can be interpreted as "locally annihilating" the particles and then "locally creating" them elsewhere. For example, when $r(X,Y) = 1$, the defect (3.117) is

$$\begin{aligned}
&\exp\left[i\sum_{\tau<0}\Delta_L\mathcal{A}_\tau(\tau+\tfrac{1}{2}, 0, 0)\right] \times \exp\left[i\mathcal{A}(0, x_0, y_0) - i\mathcal{A}(0,0,0)\right] \\
&\quad\times \exp\left[i\sum_{\tau\geq 0}\Delta_L\mathcal{A}_\tau(\tau+\tfrac{1}{2}, x_0, y_0)\right] ,
\end{aligned} \tag{3.119}$$

where the operator $e^{-i\mathcal{A}(0,0,0)}$ annihilates the particles around $(0,0)$ and then the operator $e^{i\mathcal{A}(0,x_0,y_0)}$ creates them around (x_0, y_0). For more general $r(X,Y)$, the defect (3.117) is a product of defects of the form (3.119).

Can we have a defect like (3.117) when $q(X,Y)$ is not a multiple of $p(X,Y)$? Imposing (3.118), we see that this can happen if and only if $X^{x_0}Y^{y_0}-1$ shares a nontrivial factor with $p(X,Y)$. Let us show that the latter cannot happen.

First, it is easy to check that $p(X,Y)$ is monic, non-constant, and irreducible up to a monomial.[26] Let $d=\gcd(x_0,y_0)$, which is well defined because $(x_0,y_0)\neq(0,0)$. We can write

$$X^{x_0}Y^{y_0}-1=(X^{x_0'}Y^{y_0'})^d-1=(X^{x_0'}Y^{y_0'}-1)t(X,Y)\ , \tag{3.120}$$

where $x_0'=x_0/d$, $y_0'=y_0/d$, and $t(X,Y)=\sum_{c=0}^{d-1}(X^{x_0'}Y^{y_0'})^c$ is a Laurent polynomial with $t(1,1)=d\neq 0$. The last condition implies that $p(X,Y)$ cannot share a nontrivial factor with $t(X,Y)$. Since $\gcd(x_0',y_0')=1$, the factor $X^{x_0'}Y^{y_0'}-1$ is monic, non-constant, and irreducible up to a monomial [159]. So, $p(X,Y)$ cannot share a nontrivial factor with $X^{x_0'}Y^{y_0'}-1$ as well. Therefore, $p(X,Y)$ does not share a nontrivial factor with $X^{x_0}Y^{y_0}-1$.

To summarize, a finite set of charged particles cannot move in the 2+1d $U(1)$ Laplacian gauge theory on a square lattice unless they are in the trivial superselection sector, i.e., they can be "annihilated locally." We remind the reader that when we say "move", the particles retain their charges and move in a rigid way.

[26] A polynomial is said to be *irreducible* if it cannot be written as product of two polynomials, neither of which is ±1. We say a Laurent polynomial $g(X,Y)$ is *irreducible up to a monomial* if $X^aY^bg(X,Y)$ is an irreducible polynomial for some $a,b\in\mathbb{Z}$. For example, $\tilde{p}(X,Y)=XYp(X,Y)$ is an irreducible polynomial, so $p(X,Y)$ is irreducible up to a monomial.

Chapter 4

Fractons and Lineons on Graphs: $\mathbb{Z}_N$ Laplacian Models

4.1 Introduction

In recent years, there has been a rapid development in the study of exotic lattice models in condensed matter systems. Some models, known as fractons [19–21] exhibit a variety of surprising features. These include a robust ground state degeneracy (GSD) that grows sub-extensively in the system size [24], as well as particle excitations with restricted mobility. Many of these unusual properties can be understood as following from their exotic global symmetries. These symmetries are also the underlying reasons why these lattice models defy a conventional continuum limit. See [17,18,32–34,29] for reviews on these novel topological phases of matter and their exotic global symmetries.

Most of these exotic lattice models are defined on a cubic lattice, or lattices with additional structure such as foliation [40–44, 50, 45–49]. It is then natural to ask if there are exotic models that can be defined on a general lattice graph. Recently, two such lattice models, the Laplacian ϕ-theory and the $U(1)$ Laplacian gauge theory,

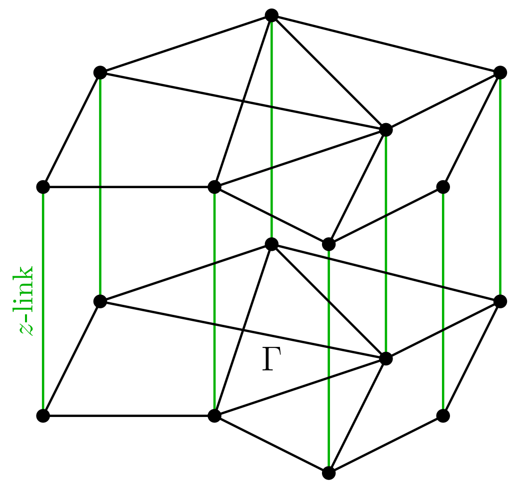


Figure 4.1: The spatial lattice $\Gamma \times C_{L_z}$: the black lines correspond to the edges of the graph Γ, and the green lines represent the z-links between two copies of Γ. Each site of the lattice is labelled as (i, z), where i denotes a vertex of the graph Γ and z denotes a vertex of C_{L_z}.

were proposed in [2,127] using the discrete Laplacian operator Δ_L. (See also [121] for a model along this line.) The former has a large GSD being the number of spanning trees of the spatial graph, which is a common measure of complexity, but it does not have fractons. The latter has defects representing immobile fracton particles, but it has no large GSD.

The $\mathbb{Z}_N$ version of the $U(1)$ Laplacian gauge theory has a large GSD and fractons, but the GSD is not robust against perturbations by local operators. This motivates us to consider a certain anisotropic generalization, which we call the *anisotropic $\mathbb{Z}_N$ Laplacian model*.[1] It has the following salient features, some of which are reminiscent of the celebrated Haah's code [20]:

- It can be placed on a spatial lattice of the form $\Gamma \times C_{L_z}$, where Γ is a general graph and C_{L_z} is a cycle graph on L_z vertices, or a 1d periodic chain with L_z sites. See Figure 4.1.

[1] The relation between the $\mathbb{Z}_N$ Laplacian model and its anisotropic uplift is analogous to that between the 2+1d $\mathbb{Z}_N$ Ising plaquette model [35] and the 3+1d anisotropic lineon model in [42,58].

- The GSD is robust[2] and is given by

$$\mathrm{GSD} = |\,\mathrm{Jac}(\Gamma, N)|^2\,, \tag{4.1}$$

  where $\mathrm{Jac}(\Gamma, N)$ is a "mod N-reduction" of the Jacobian group $\mathrm{Jac}(\Gamma)$ of Γ.

- It has lineons that can only move in the z-direction if Γ is an infinite two-dimensional square lattice.

- In the special case when the spatial lattice is a $L_x \times L_y \times L_z$ cubic lattice and when $N = p$ is prime, we have

$$\log_p \mathrm{GSD} = 2 \dim_{\mathbb{Z}_p} \frac{\mathbb{Z}_p[X, Y]}{(\,Y(X-1)^2 + X(Y-1)^2, X^{L_x} - 1, Y^{L_y} - 1\,)}\,. \tag{4.2}$$

  The definition and the explicit evaluation of this formula are discussed in Appendix 4.C. It depends on the number-theoretic properties of L_x, L_y. Interestingly, there exists a sequence of L_x, L_y going to infinity such that the $\log_p \mathrm{GSD} \sim O(L_x, L_y)$, but there is also a sequence such that $\log_p \mathrm{GSD}$ stays at order 1 if $p > 2$. See Figure 4.2.

We present this model both in terms of the low-energy limit of a stabilizer code in the Hamiltonian formalism, and in terms of a Euclidean lattice model using an integer BF action, à la Chapter 2. We compare the four Laplacian lattice models in Table 4.1.[3]

Following [57, 35–38, 58, 59, 1, 26, 122, 25, 2, 127], we focus on the exotic global symmetries of this model. The symmetries of the models on Γ (the first three models

[2] On a general graph, there is no notion of locality and therefore we cannot discuss local operators and match them between the UV and the IR theories. Consequently, the discussion of robustness is ambiguous. This is not the case on regular lattices where the usual discussion of local operators and robustness applies. In that case, the anisotropic model is robust as we will show in Section 4.3.2.

[3] In the table, we assume the θ-angle of the $U(1)$ Laplacian theory is not π, otherwise the GSD is 2.

Model	Spatial Lattice	GSD	Defects	Robust?
Laplacian ϕ-theory Chapter 3, [127]	Γ	$\lvert \mathrm{Jac}(\Gamma) \rvert$	None	No
$U(1)$ Laplacian gauge theory Chapter 3, [127]	Γ	1	Fracton	Yes
$\mathbb{Z}_N$ Laplacian model Appendix 4.D	Γ	$\lvert \mathrm{Jac}(\Gamma, N) \rvert$	Fracton	No
Anisotropic $\mathbb{Z}_N$ Laplacian model Sections 4.3 and 4.4	$\Gamma \times C_{L_z}$	$\lvert \mathrm{Jac}(\Gamma, N) \rvert^2$	Lineon	Yes

Table 4.1: The comparison of four exotic lattice models that can be defined on a general graph Γ. The model is robust if it has no relevant local operator. See Section 4.3.2 for more details on what we mean by robustness.

in Table 4.1) are not subsystem global symmetries. The symmetry operators are supported on most (or all) of the sites of Γ, rather than on a small subset of them. The precise subset depends delicately on the details of Γ. Yet, there are many such symmetries. In this sense these symmetries are generalizations of the dipole symmetries [56, 160] on cubic lattices, which are also supported on the entire lattice. The difference is that the dipole symmetries have simple dependence on the coordinates, while here the dependence on the coordinates is more complicated.

This is not the case in the anisotropic $\mathbb{Z}_N$ Laplacian model (the fourth model in Table 4.1). Here the symmetries act at fixed z and in that sense they are subsystem symmetries. (See [161,162] for other anisotropic fractal models with symmetries that act at fixed z.) In fact, as we will discuss below, at low energy these symmetries are independent of z and are similar to one-form global symmetries.[4]

[4]More generally, we can classify symmetries by the difference operators that annihilate the transformation parameters α. For example, on a regular lattice, an ordinary symmetry has $\Delta_x\alpha = \Delta_y\alpha = \Delta_z\alpha = 0$, a dipole symmetry has $\Delta_x\Delta_x\alpha = \Delta_x\Delta_y\alpha = \Delta_y\Delta_y\alpha = 0$, etc. (And of course, α can carry more indices for the various fields or directions in spacetime.). More interesting examples arise in theories associated to Haah's code [20], where the difference equations are $(\Delta_x + \Delta_y + \Delta_z)\alpha = 0$ and $[\Delta_x\Delta_y + \Delta_y\Delta_z + \Delta_z\Delta_x + 2(\Delta_x + \Delta_y + \Delta_z)]\alpha = 0$.

The rest of the chapter is organized as follows. In Section 4.2, we introduce some necessary graph theory background, including the discrete Laplace operator and the Jacobian group of a graph. In Section 4.3, we introduce the stabilizer code and the Euclidean integer BF action for the anisotropic $\mathbb{Z}_N$ Laplacian model. We derive the general expression for the GSD (4.1) and discuss the restricted mobility of the lineon defects from time-like symmetries.[5] Section 4.4 considers the special case when the spatial lattice is a cubic lattice and when N is prime. The GSD reduces to (4.2) and we discuss its asymptotic behaviors. Appendix 4.A discusses $\mathbb{Z}_N$-valued discrete harmonic functions on a general graph. Appendices 4.B and 4.C contain the detailed computation of the GSD and the mobility restrictions for the anisotropic $\mathbb{Z}_N$ Laplacian model on a cubic lattice with N a prime number. In Appendix 4.D, we study the $\mathbb{Z}_N$ Laplacian model of fractons, which is not robust.

4.2 Graph theory primer

In this section, we review some well-known facts about a finite graph, and $\mathbb{Z}_N$-valued functions on the graph. A good reference on this subject is [125]. See also [2] for more discussion on these topics in related lattice models.

Let Γ be a simple, undirected, connected graph on N vertices.[6] Here, *simple* means there is at most one edge between any two vertices and no self-loop on any vertex, *undirected* means the edges do not have any orientation, and *connected* means there is a path between any two vertices of the graph. We use i to denote a vertex (or site), and $\langle i, j\rangle$ to denote an edge (or link) of the graph. We write $\langle i, j\rangle \in \Gamma$ if there is an edge between vertices i and j in Γ.

Let d_i be the *degree* of vertex i, i.e., the number of edges incident on i. The *Laplacian matrix* L of Γ is an $\mathsf{N} \times \mathsf{N}$ symmetric matrix defined as follows: $L_{ii} = d_i$

[5] See [25] for a definition of space-like and time-like global symmetries and their applications to the ground state degeneracy and restricted mobility constraints.

[6] Note that N in $\mathbb{Z}_N$ is different from N, the number of vertices of Γ.

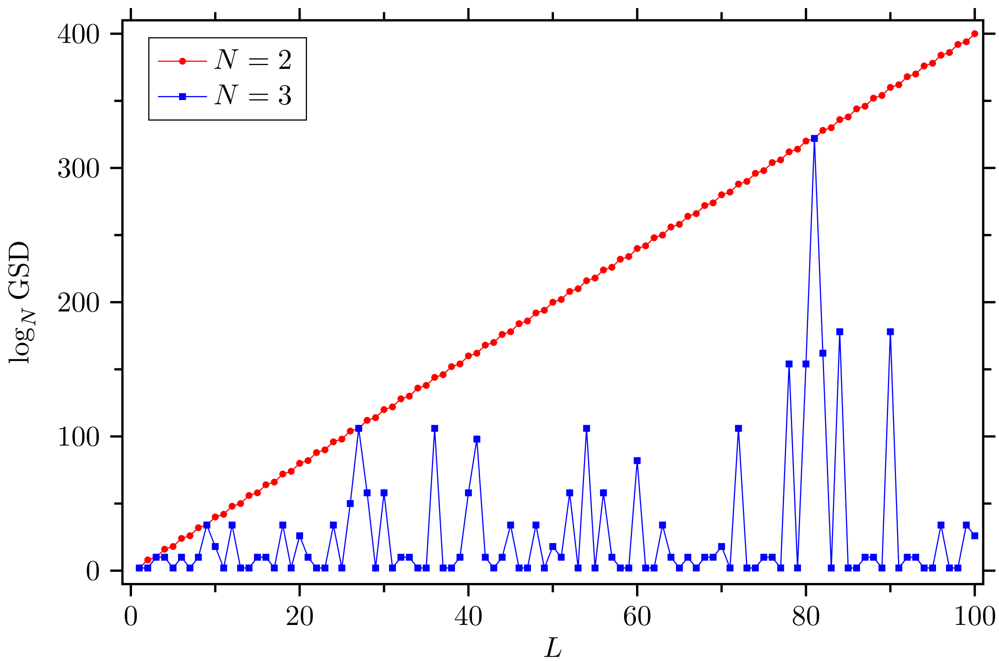


Figure 4.2: The logarithm of the ground state degeneracy $\log_N$ GSD of the 3+1d anisotropic $\mathbb{Z}_N$ Laplacian model of lineons on a cubic lattice with $L \times L \times L_z$ sites (i.e., $\Gamma = C_L \times C_L$) for $N = 2$ (red) and $N = 3$ (blue), and $1 \leq L \leq 100$. While $\log_N$ GSD grows steadily as $\sim 4L$ for $N = 2$, it behaves erratically for $N = 3$. In fact, for $N = 3$, there are infinitely many L for which it is just 2, and also infinitely many L (powers of 3) for which it is $\sim 4L$. Here, the GSD for both $N = 2, 3$ was calculated using techniques from commutative algebra explained in Appendix 4.C.1. In Section 4.4.2, we give a simpler derivation of the GSD for $N = 2$ by relating the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model to the 3+1d anisotropic $\mathbb{Z}_2$ lineon model [42,58] on a tilted lattice.

for every vertex i, $L_{ij} = -1$ if there is an edge $\langle i, j \rangle$ between vertices i and j, and $L_{ij} = 0$ otherwise.

4.2.1 Discrete Laplacian operator Δ_L and its Smith decomposition

Consider a $\mathbb{Z}_N$-valued function $f(i)$ on the vertices of the graph. We define the *discrete Laplacian* operator Δ_L as

$$\Delta_L f(i) := \sum_j L_{ij} f(j) = d_i f(i) - \sum_{j:\langle i,j\rangle\in\Gamma} f(j) = \sum_{j:\langle i,j\rangle\in\Gamma} [f(i) - f(j)] \ , \tag{4.3}$$

where the equalities are modulo N. This is one of the most natural and universal difference operators that can be defined on any such graph Γ.

We are interested in the following two questions:

1. What are all the $\mathbb{Z}_N$-valued functions $h(i)$ that satisfy the *discrete Laplacian equation*

$$\Delta_L h(i) = 0 \mod N \ ? \tag{4.4}$$

   They are known as the $\mathbb{Z}_N$-valued *discrete harmonic functions*, and we denote the set of such functions as $\mathcal{H}(\Gamma, \mathbb{Z}_N)$.[7]

2. We define an equivalence class of $\mathbb{Z}_N$-valued functions by saying that two functions $g(i)$ and $\tilde{g}(i)$ belong to the same class if there is a $\mathbb{Z}_N$-valued function $f(i)$ such that

$$\tilde{g}(i) - g(i) = \Delta_L f(i) \mod N \ . \tag{4.5}$$

   In this case, we write $\tilde{g}(i) \sim g(i)$. What are all the distinct equivalence classes under the equivalence relation "$\sim$"?

Interestingly, both questions can be answered using the *Smith decomposition* [142] of the Laplacian matrix L [143]. The *Smith normal form* of L is given by three matrices

[7] It is also known as the *group of balanced vertex weightings* [141].

R, P, and Q, such that

$$R = PLQ \ , \qquad \text{or} \qquad R_{ab} = \sum_{i,j} P_{ai} L_{ij} Q_{jb} \ , \tag{4.6}$$

where $P, Q \in GL_{\mathsf{N}}(\mathbb{Z})$, and $R = \mathrm{diag}(r_1, \ldots, r_{\mathsf{N}})$. Here, r_a's are nonnegative integers, known as the *invariant factors* of L, such that r_a divides r_{a+1} for $a = 1, \ldots, \mathsf{N} - 1$. While R is uniquely determined by L, the matrices P and Q are not. For a connected graph Γ, we have $r_a > 0$ for $a = 1, \ldots, \mathsf{N} - 1$, and $r_{\mathsf{N}} = 0$.

We state the answers to the two questions here (see Appendix 4.A for details):

1. Any $\mathbb{Z}_N$-valued discrete harmonic function takes the form

$$h(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_a}{\gcd(N, r_a)} \mod N \ , \tag{4.7}$$

   where $p_a = 0, \ldots, \gcd(N, r_a) - 1$ for $a = 1, \ldots, \mathsf{N}$. In other words, $\mathcal{H}(\Gamma, \mathbb{Z}_N)$ is isomorphic to the finite Abelian group $\prod_{a=1}^{\mathsf{N}} \mathbb{Z}_{\gcd(N, r_a)}$. Here, the group operation is simply the sum. It is well-defined because if $h_1, h_2 \in \mathcal{H}(\Gamma, \mathbb{Z}_N)$, then $h_1 + h_2 \in \mathcal{H}(\Gamma, \mathbb{Z}_N)$.

2. Any equivalence class is uniquely represented by the $\mathbb{Z}_N$-valued function

$$g(i) = \sum_{a=1}^{\mathsf{N}} p_a (Q^{-1})_{ai} \mod N \ , \tag{4.8}$$

   where $p_a = 0, \ldots, \gcd(N, r_a) - 1$ for $a = 1, \ldots, \mathsf{N}$. In other words, the set of all equivalence classes is isomorphic to the finite Abelian group $\prod_{a=1}^{\mathsf{N}} \mathbb{Z}_{\gcd(N, r_a)}$. Here, the group operation is simply the sum which is well-defined because if $\tilde{g}_1 \sim g_1$ and $\tilde{g}_2 \sim g_2$, then $\tilde{g}_1 + \tilde{g}_2 \sim g_1 + g_2$.

It follows that the number of $\mathbb{Z}_N$-valued discrete harmonic functions and the number of equivalence classes are both $\prod_{a=1}^{\mathsf{N}} \gcd(N, r_a)$, which is the order of the finite Abelian group $\prod_{a=1}^{\mathsf{N}} \mathbb{Z}_{\gcd(N,r_a)}$. We will have more to say about this group below.

4.2.2 Jacobian group of a graph

The finite Abelian group encountered above is intimately related to the *Jacobian group* $\mathrm{Jac}(\Gamma)$, which is a natural finite Abelian group associated with a general graph Γ.[8] In terms of the invariant factors of the Laplacian matrix L, we have the following isomorphism

$$\mathrm{Jac}(\Gamma) \cong \prod_{a=1}^{\mathsf{N}-1} \mathbb{Z}_{r_a} \ . \tag{4.9}$$

The order of $\mathrm{Jac}(\Gamma)$ is the most fundamental and well-studied notion of *complexity* in graph theory. What we have here is a "mod N-reduction" of the Jacobian group:

$$\mathrm{Jac}(\Gamma, N) \cong \prod_{a=1}^{\mathsf{N}} \mathbb{Z}_{\gcd(N,r_a)} \ . \tag{4.10}$$

As we will see below, the group $\mathrm{Jac}(\Gamma, N)$ plays an crucial role in the $\mathbb{Z}_N$ Laplacian models.

4.3 Anisotropic $\mathbb{Z}_N$ Laplacian model on a graph

In this section, we study a robust gapped lineon model on a spatial lattice of the form $\Gamma \times C_{L_z}$, where Γ is a simple, connected, undirected graph, and L_z is the number of sites in the z-direction (see Figure 4.1). We refer to it as the anisotropic $\mathbb{Z}_N$ Laplacian model because it is the anisotropic extension along the z-direction of the $\mathbb{Z}_N$ Laplacian model analyzed in Appendix 4.D.

[8] It has several different names in the graph theory literature, including the *sandpile group* [146], or the *group of components* [148], or the *critical group* [149] of Γ, and it is related to the *group of bicycles* [141] of Γ.

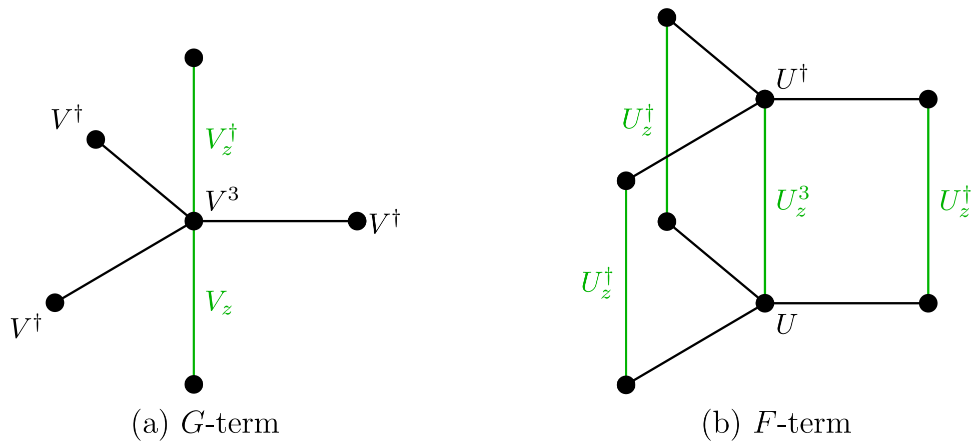


Figure 4.3: The two kinds of stabilizer terms in the Hamiltonian (4.11).

4.3.1 Hamiltonian for the stabilizer code

In the Hamiltonian formulation of the anisotropic $\mathbb{Z}_N$ Laplacian model, there are a $\mathbb{Z}_N$ variable $U(i,z)$ and its conjugate variable $V(i,z)$, i.e., $U(i,z)V(i,z) = e^{2\pi i/N}V(i,z)U(i,z)$, on every site of $\Gamma \times C_{L_z}$. There are also a $\mathbb{Z}_N$ variable $U_z(i,z+\frac{1}{2})$ and its conjugate variable $V_z(i,z+\frac{1}{2})$, i.e., $U_z(i,z+\frac{1}{2})V_z(i,z+\frac{1}{2}) = e^{2\pi i/N}V_z(i,z+\frac{1}{2})U_z(i,z+\frac{1}{2})$, on every z-link of $\Gamma \times C_{L_z}$.

The Hamiltonian is

$$H = -\gamma_1 \sum_{i,z} G(i,z) - \gamma_2 \sum_{i,z} F(i,z+\tfrac{1}{2}) + \text{h.c.}\ , \tag{4.11}$$

where

$$\begin{aligned} G(i,z) &= V_z(i,z+\tfrac{1}{2})^\dagger V_z(i,z-\tfrac{1}{2}) \prod_{j:\langle i,j\rangle\in\Gamma} V(i,z)V(j,z)^\dagger\ , \\ F(i,z+\tfrac{1}{2}) &= U(i,z+1)^\dagger U(i,z) \prod_{j:\langle i,j\rangle\in\Gamma} U_z(i,z+\tfrac{1}{2})U_z(j,z+\tfrac{1}{2})^\dagger\ . \end{aligned} \tag{4.12}$$

The two kinds of terms are shown in Figure 4.3. Since all the terms in this Hamiltonian commute with each other, it is a stabilizer code.

The model enjoys the duality transformation

$$
\begin{aligned}
&U(i,z) \to V_z(i,z+\tfrac{1}{2})\ , \qquad && U_z(i,z+\tfrac{1}{2}) \to V(i,z)\ ,\\
&V(i,z) \to U_z(i,z+\tfrac{1}{2})^\dagger\ , \qquad && V_z(i,z+\tfrac{1}{2}) \to U(i,z)^\dagger\ .
\end{aligned}
\tag{4.13}
$$

It exchanges the two kind of terms in the Hamiltonian and therefore it maps the model with (γ_1, γ_2) to the model with (γ_2, γ_1). As a result, for $\gamma_1 = \gamma_2$, this model is self-dual.

The ground states satisfy $G(i,z) = 1$ and $F(i, z+\frac{1}{2}) = 1$ for all i, z. The excited states are violations of $G = 1$ or $F = 1$, which we call electric and magnetic excitations respectively. These excitations are mobile along the z direction so they are at least z-lineons. Their mobility constraints along the graph Γ are more complicated. We postpone that discussion to Section 4.3.2.

We could also take $\gamma_1, \gamma_2 \to \infty$, in which case, the Hilbert space consists of only the ground states, and the Hamiltonian is trivial. The Euclidean presentation of this model in this limit will be discussed later in Section 4.3.2.

We are particularly interested in those operators that commute with the Hamiltonian (4.11) and act nontrivially on its ground states. They are the global symmetry operators of the model in the low energy limit, and they are also known as the logical operators of the stabilizer code. We choose a basis of these symmetry operators as follows: the electric symmetry operators are[9]

$$
\begin{aligned}
&\tilde{W}_z(a) = \prod_{i,z} V(i,z)^{(Q^{-1})_{ai}}\ , \qquad && a = 1,\dots,\mathsf{N}\ ,\\
&\tilde{W}(a; z+\tfrac{1}{2}) = \prod_{i} V_z(i,z+\tfrac{1}{2})^{\frac{N}{\gcd(N,r_a)}Q_{ia}}\ , \qquad && z = 0,\dots,L_z-1\ ,
\end{aligned}
\tag{4.14}
$$

[9]In fact, the operator $\prod_z V(i,z)$, which is local in Γ and extends in the z direction, also commutes with the Hamiltonian. Here we choose to work with the basis of $\tilde{W}_z(a)$ because the latter has a simpler commutation relation with $W(a;z)$. Similarly, $\prod_z U_z(i, z+\frac{1}{2})$ also commutes with the Hamiltonian and is local in Γ, but we choose to work in the basis of $W_z(a)$ for the same reason.

and the magnetic symmetry operators are

$$\begin{aligned} W_z(a) &= \prod_{i,z} U_z(i, z+\tfrac{1}{2})^{(Q^{-1})_{ai}} \ , \qquad a = 1, \ldots, \mathsf{N} \ , \\ W(a; z) &= \prod_{i} U(i, z)^{\frac{N}{\gcd(N, r_a)} Q_{ia}} \ , \qquad z = 0, \ldots, L_z - 1 \ , \end{aligned} \tag{4.15}$$

where Q and r_a are defined in (4.6). These operators generate a $\mathrm{Jac}(\Gamma, N)^2$ electric symmetry and a $\mathrm{Jac}(\Gamma, N)^2$ magnetic symmetry.

For each a, the four operators in (4.14) and (4.15) are all $\mathbb{Z}_{\gcd(N,r_a)}$ operators. Clearly, $W(a;z)^{\gcd(N,r_a)} = \tilde{W}(a; z+\frac{1}{2})^{\gcd(N,r_a)} = 1$. (In fact, when acting on the ground states, the operators $\tilde{W}(a; z+\frac{1}{2})$ and $W(a;z)$ are independent of z.) Moreover, the operator $W_z(a)$ satisfies $W_z(a)^{r_a} = 1$ when acting on the ground states,[10] which when combined with the obvious relation $W_z(a)^N = 1$ gives the relation $W_z(a)^{\gcd(N,r_a)} = 1$. The same conclusion holds for $\tilde{W}_z(a)$ as well.

The basis of symmetry operators defined in (4.14) and (4.15) is chosen such that they satisfy the following commutation relations:

$$W(a;z)\tilde{W}_z(b) = \exp\left[\frac{2\pi i \delta_{ab}}{\gcd(N, r_a)}\right] \tilde{W}_z(b) W(a;z) \ , \qquad a, b = 1, \ldots, \mathsf{N} \ , \tag{4.16}$$

and similarly for the other pair. So for each $a = 1, \ldots, \mathsf{N}$, there are two independent copies of $\mathbb{Z}_{\gcd(N,r_a)}$ Heisenberg algebras, leading to a ground state degeneracy of[11]

$$\mathrm{GSD} = \prod_{a=1}^{\mathsf{N}} \gcd(N, r_a)^2 = |\,\mathrm{Jac}(\Gamma, N)|^2 \ . \tag{4.17}$$

[10] This is because $W_z(a)^{r_a} = \prod_{i,z} U_z(i, z+\frac{1}{2})^{r_a (Q^{-1})_{ai}} = \prod_{i,j,z} U_z(i, z+\frac{1}{2})^{P_{aj} L_{ji}} = \prod_j \left(\prod_z U(j, z+1) U(j,z)^\dagger\right)^{P_{aj}} = 1$, where we used the facts that $RQ^{-1} = PL$, and $F = 1$ on the ground states.

[11] The power of 2 in (4.17) is related to the fact that the anisotropic $\mathbb{Z}_N$ Laplacian model is the anisotropic extension of the $\mathbb{Z}_N$ Laplacian model of Appendix 4.D, whose GSD is $|\,\mathrm{Jac}(\Gamma, N)|$ (4.133). This is similar to the relation $\mathrm{GSD}_{3+1\text{d anisotropic } \mathbb{Z}_N \text{ lineon model}} = (\mathrm{GSD}_{2+1\text{d } \mathbb{Z}_N \text{ Ising plaquette model}})^2$.

4.3.2 Euclidean presentation

We now discuss the Euclidean presentation of the anisotropic $\mathbb{Z}_N$ Laplacian model. We place the theory on a Euclidean spacetime lattice $C_{L_\tau} \times \Gamma \times C_{L_z}$, where $\Gamma \times C_{L_z}$ is the spatial slice. We use (τ, i, z) to label a site in the spacetime lattice, where i denotes a vertex of the graph Γ.

We use the integer BF formulation of Chapter 2. The integer BF-action of the anisotropic $\mathbb{Z}_N$ Laplacian model is

$$\begin{aligned} S = \frac{2\pi i}{N} \sum_{\tau,i,z} \Big(& -\tilde{m}_\tau(\tau, i, z+\tfrac{1}{2}) \left[\Delta_z m(\tau, i, z+\tfrac{1}{2}) - \Delta_L m_z(\tau, i, z+\tfrac{1}{2})\right] \\ & + \tilde{m}_z(\tau+\tfrac{1}{2}, i, z) \left[\Delta_\tau m(\tau+\tfrac{1}{2}, i, z) - \Delta_L m_\tau(\tau+\tfrac{1}{2}, i, z)\right] \\ & + \tilde{m}(\tau+\tfrac{1}{2}, i, z+\tfrac{1}{2}) \left[\Delta_\tau m_z(\tau+\tfrac{1}{2}, i, z+\tfrac{1}{2}) - \Delta_z m_\tau(\tau+\tfrac{1}{2}, i, z+\tfrac{1}{2})\right] \Big) \ , \end{aligned} \tag{4.18}$$

where the integer fields (m_τ, m, m_z) have a gauge symmetry

$$\begin{aligned} & m_\tau \sim m_\tau + \Delta_\tau k + N q_\tau \ , \\ & m \sim m + \Delta_L k + N q \ , \\ & m_z \sim m_z + \Delta_z k + N q_z \ , \end{aligned} \tag{4.19}$$

where k and (q_τ, q, q_z) are integers, and similarly for $(\tilde{m}_\tau, \tilde{m}, \tilde{m}_z)$. (Note that, when working modulo N, the second line of (4.19) is exactly the equivalence relation discussed in (4.5).)

The theory is self-dual under the map $(m_\tau, m, m_z) \to (\tilde{m}_\tau, \tilde{m}, \tilde{m}_z)$ and $(\tilde{m}_\tau, \tilde{m}, \tilde{m}_z) \to -(m_\tau, m, m_z)$.

The integer BF-action (4.18) describes the ground states of a stabilizer code given by the Hamitonian (4.11). Here, we will not elaborate on the relation between the Euclidean and Hamiltonian presentations. We refer the readers to Appendix C.2 of [1] for an analogous discussion of the relation between the 2+1d $\mathbb{Z}_N$ toric code and the 2+1d $\mathbb{Z}_N$ gauge theory in the integer BF presentation.

Ground state degeneracy

We can count the number of ground states by counting the number of solutions to the "equations of motion" of $(\tilde{m}_\tau, \tilde{m}, \tilde{m}_z)$ modulo gauge transformations:

$$\begin{aligned}
&\Delta_z m - \Delta_L m_z = 0 \mod N \ , \\
&\Delta_\tau m - \Delta_L m_\tau = 0 \mod N \ , \\
&\Delta_\tau m_z - \Delta_z m_\tau = 0 \mod N \ .
\end{aligned} \tag{4.20}$$

A gauge field (m_τ, m, m_z) that satisfies (4.20) is a flat $\mathbb{Z}_N$ gauge field. We can use the gauge freedom in k to set $m_\tau(\tau + \frac{1}{2}, i, z)|_{\tau\neq 0} = 0 \mod N$, and $m_z(\tau, i, z + \frac{1}{2})|_{z\neq 0} = 0 \mod N$. In this gauge choice, the last line of (4.20) implies that

$$\begin{aligned}
&\Delta_\tau m_z(\tau + \tfrac{1}{2}, i, z + \tfrac{1}{2})|_{z=0} = 0 \mod N \ , \\
&\Delta_z m_\tau(\tau + \tfrac{1}{2}, i, z + \tfrac{1}{2})|_{\tau=0} = 0 \mod N \ .
\end{aligned} \tag{4.21}$$

The first two lines of (4.20) then imply that

$$\begin{aligned}
&\Delta_\tau m(\tau + \tfrac{1}{2}, i, z) = 0 \mod N \ , \\
&\Delta_z m(\tau, i, z + \tfrac{1}{2}) = 0 \mod N \ ,
\end{aligned} \tag{4.22}$$

which in turn imply that

$$\begin{aligned}
&\Delta_L m_z(i, z + \tfrac{1}{2})|_{z=0} = 0 \mod N \ , \\
&\Delta_L m_\tau(\tau + \tfrac{1}{2}, i)|_{\tau=0} = 0 \mod N \ .
\end{aligned} \tag{4.23}$$

The remaining τ and z-independent gauge freedom, $m(i) \sim m(i) + \Delta_L k(i)$, is exactly the equivalence relation in (4.5). So, after gauge fixing, we can set $m(i)$ to be of the form (4.8), i.e., there are $|\operatorname{Jac}(\Gamma, N)|$ independent holonomies in $m(i)$. Since $m_z(i, z+\frac{1}{2})|_{z=0}$ and $m_\tau(\tau+\frac{1}{2}, i)|_{\tau=0}$ satisfy (4.23), which is exactly the discrete Laplace equation (4.4), they are of the form (4.7). So there are $|\operatorname{Jac}(\Gamma, N)|$ independent holonomies in both of them. Finally, the set of gauge transformations $k(\tau, i, z)$ that do not act on (m_τ, m, m_z) satisfy $\Delta_\tau k = \Delta_z k = \Delta_L k = 0 \mod N$. In other words, $k(\tau, i, z) = k(i)$ is independent of τ, z, and $k(i)$ satisfies the discrete Laplace equation (4.4). So such gauge transformations are of the form (4.7), and there are $|\operatorname{Jac}(\Gamma, N)|$ of them. Therefore, the ground state degeneracy is

$$\mathrm{GSD} = \frac{|\operatorname{Jac}(\Gamma, N)|^3}{|\operatorname{Jac}(\Gamma, N)|} = |\operatorname{Jac}(\Gamma, N)|^2 = \prod_{a=1}^{\mathsf{N}} \gcd(N, r_a)^2 \ . \tag{4.24}$$

Global symmetry

There is an electric symmetry associated with the shift of (m_τ, m, m_z) by a flat $\mathbb{Z}_N$ gauge field. By the analysis following (4.20), up to gauge transformations, the electric (space-like) symmetry acts as[12]

$$\begin{aligned}
&m(\tau, i, z) \to m(\tau, i, z) + \lambda(i) \ , && \lambda(i) = \sum_{a=1}^{\mathsf{N}} p_a (Q^{-1})_{ai} \ , \\
&m_z(\tau, i, z+\tfrac{1}{2}) \to m_z(\tau, i, z+\tfrac{1}{2}) + \delta_{z,0}\lambda_z(i) \ , && \lambda_z(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_{z,a}}{\gcd(N, r_a)} \ ,
\end{aligned} \tag{4.25}$$

where p_a and $p_{z,a}$ are both integers modulo $\gcd(N, r_a)$ for $a = 1, \ldots, \mathsf{N}$. There is also a magnetic (space-like) symmetry which acts on $\tilde{m}$ and $\tilde{m}_z$ in a similar way.

[12] These are symmetries of the action (4.18) because $\Delta_\tau m$, $\Delta_z m$, and $\Delta_\tau m_z$ are clearly unaffected by the shifts, whereas $\Delta_L m_z$ is shifted by $\delta_{z,0}\Delta_L \lambda_z(i) = 0 \mod N$ because $\lambda_z(i)$ is a $\mathbb{Z}_N$-valued discrete harmonic function (4.7).

The electric (space-like) symmetry is generated by the Wilson operators of $(\tilde{m}_\tau, \tilde{m}, \tilde{m}_z)$:

$$\begin{aligned} \tilde{W}_z(a) &= \exp\left[\frac{2\pi i}{N}\sum_{i,z}(Q^{-1})_{ai}\tilde{m}_z(\tau+\tfrac{1}{2},i,z)\right] , \\ \tilde{W}(a;z+\tfrac{1}{2}) &= \exp\left[\frac{2\pi i}{\gcd(N,r_a)}\sum_{i}\tilde{m}(\tau+\tfrac{1}{2},i,z+\tfrac{1}{2})Q_{ia}\right] , \end{aligned} \tag{4.26}$$

for $a = 1, \ldots, \mathsf{N}$ and $z = 0, \ldots, L_z - 1$. The electrically charged operators are the Wilson operators of (m_τ, m, m_z), i.e., $W(a;z)$ and $W_z(a)$. Similarly, the magnetic (space-like) symmetry is generated by $W(a;z)$ and $W_z(a)$, while the magnetically charged operators are $\tilde{W}_z(a)$ and $\tilde{W}(a;z+\frac{1}{2})$. These are the operators in (4.14) and (4.15) in the low energy limit. The commutation relation (4.16) can now be understood as a mixed 't Hooft anomaly between electric and magnetic space-like symmetries.

Time-like symmetry and lineons

The integer BF-action has defects, which extend in the time direction, such as

$$W_\tau(i,z) = \exp\left[\frac{2\pi i}{N}\sum_{\tau} m_\tau(\tau+\tfrac{1}{2},i,z)\right] . \tag{4.27}$$

This describes the world-line of an infinitely heavy particle of unit charge at position (i,z). It also represents the low energy limit of an electric excitation at position (i,z) in the stabilizer code (4.11). We can deform the defect to

$$\begin{aligned} &\exp\left[\frac{2\pi i}{N}\sum_{\tau<0} m_\tau(\tau+\tfrac{1}{2},i,z)\right]\exp\left[\frac{2\pi i}{N}\sum_{z\le z''<z'} m_z(0,i,z''+\tfrac{1}{2})\right] \\ &\quad\times\exp\left[\frac{2\pi i}{N}\sum_{\tau\ge 0} m_\tau(\tau+\tfrac{1}{2},i,z')\right] . \end{aligned} \tag{4.28}$$

This configuration describes a particle moving along the z-direction.

Next, we discuss the mobility of the particle along the graph Γ. Such a motion is constrained by the time-like global symmetry, which acts on extended defects rather than the operators or states of the Hilbert space (see [25] for more discussions on time-like global symmetries). Up to gauge transformation, the electric time-like symmetry acts as[13]

$$m_\tau(\tau + \tfrac{1}{2}, i, z) \to m_\tau(\tau + \tfrac{1}{2}, i, z) + \delta_{\tau,0}\lambda_\tau(i) \ , \qquad \lambda_\tau(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_{\tau,a}}{\gcd(N, r_a)} \ , \tag{4.29}$$

where $p_{\tau,a} = 0, \ldots, \gcd(N, r_a) - 1$. Hence, the group of electric time-like symmetry is $\mathrm{Jac}(\Gamma, N)$.

Two defects at sites (i, z) and (i', z') carry the same time-like charges, or equivalently, a particle can hop from (i, z) to (i', z'), if and only if[14]

$$Q_{ia} = Q_{i'a} \mod \gcd(N, r_a) \ , \qquad \forall a = 1, \ldots, \mathsf{N} \ . \tag{4.31}$$

In other words, the time-like charges Q_{ia} encode the superselection sector of a defect.

Similarly, there are defects of $\tilde{m}_\tau$ which represent the low energy limit of magnetic excitations of the stabilizer code (4.11). By the self-duality, similar mobility restrictions apply to the defects of $\tilde{m}_\tau$ due to a $\mathrm{Jac}(\Gamma, N)$ dual magnetic time-like symmetry.

[13] This is a symmetry of the action (4.18) because $\Delta_z m_\tau$ is clearly unaffected by the shift, and $\Delta_L m_\tau$ is shifted by $\delta_{\tau,0}\Delta_L \lambda_\tau(i) = 0 \mod N$ because $\lambda_\tau(i)$ is a $\mathbb{Z}_N$-valued discrete harmonic function (4.7).

[14] Indeed, when this condition holds, the defect that "moves" a particle from (i, z) to (i', z') at time $\tau = 0$ is given by

$$\begin{aligned} &\exp\left[\frac{2\pi i}{N}\sum_{\tau<0} m_\tau(\tau + \tfrac{1}{2}, i, z)\right] \exp\left[-\frac{2\pi i}{N}\sum_{a,j}\left(\frac{Q_{ia} - Q_{i'a}}{\gcd(N, r_a)}\right)\tilde{r}_a P_{aj} m(0, j, z)\right] \\ &\times \exp\left[\frac{2\pi i}{N}\sum_{z \le z'' < z'} m_z(0, i', z'' + \tfrac{1}{2})\right] \exp\left[\frac{2\pi i}{N}\sum_{\tau \ge 0} m_\tau(\tau + \tfrac{1}{2}, i', z')\right] \ , \end{aligned} \tag{4.30}$$

where for each a, $\tilde{r}_a$ is the integer solution of the equation $\tilde{r}_a r_a = \gcd(N, r_a) \mod N$.

While this selection rule (4.31) is not very intuitive, we will give strong mobility constraints in the special case where the spatial lattice is a cubic lattice (i.e., Γ is a 2d torus graph $C_{L_x} \times C_{L_y}$) in Section 4.4. In particular, under some mild conditions, the particles can move only along the z-direction, i.e., they are lineons.

Robustness

Let us examine the robustness of the low-energy theory. Typically, in order to address this question we should map local operators in the UV theory to local operators in the IR theory. However, if Γ is a general graph, it has no notion of locality and we cannot discuss local operators. Therefore, the usual discussion of robustness does not apply. Instead, we will restrict Γ to be a regular lattice (such as square lattice, honeycomb lattice, cubic lattice, etc.), where there is an unambiguous notion of locality and we can consider localized operators. (One might be able to extend the discussion to the case of an infinite graph Γ with some restrictions on its connectivity. We will not attempt to do it here.)

The only operators that act nontrivially on the ground states are $\tilde{W}(a;z)$ and $\tilde{W}_z(a)$ of (4.26), and similarly, $W(a;z)$ and $W_z(a)$. It is clear that $W_z(a)$ and $\tilde{W}_z(a)$ are supported over L_z sites in the z-direction, so they are not finitely supported in the infinite volume limit. Now, we show that $W(a;z)$ is not finitely supported when Γ is a regular lattice. Assume to the contrary that it is finitely supported. It generates a $\mathrm{Jac}(\Gamma, N)$ magnetic symmetry that shifts the gauge field $\tilde{m}(\tau+\frac{1}{2}, i, z')$ by $\delta_{z',z} f(i)$, where $f(i)$ is a $\mathbb{Z}_N$-valued discrete harmonic function. The support of $f(i)$ is precisely the support of the operator $W(a;z)$, so $f(i)$ is also finitely supported. However, on a regular lattice, there is no nontrivial finitely-supported discrete harmonic function.[15]

[15]For example, on a square lattice with coordinates (x, y), let $f(x, y)$ be a finitely-supported discrete harmonic function. Consider a large rectangular region R that contains the support of $f(x, y)$, i.e., $f(x, y) = 0$ for (x, y) outside R. Using the discrete Laplace equation $\Delta_L f(x, y) = 0$ at the points immediately outside R, one can show that $f(x, y) = 0$ at the points immediately inside R. By induction, $f(x, y) = 0$ everywhere inside R. This argument extends to any regular lattice. It also extends to a more general class of graphs but we do not discuss this here.

Therefore, $W(a;z)$ cannot be finitely supported. Similarly, $\tilde{W}(a;z+\frac{1}{2})$ is also not finitely supported.

Since there are no finitely-supported operators that act nontrivially in the space of ground states, the anisotropic $\mathbb{Z}_N$ Laplacian model is robust. We can deform the microscopic model with finitely-supported operators. As long as their coefficients are small enough, they map to localized deformations of the low-energy theory. However, since there are no local point-like operators acting in the low-energy theory, it cannot change.

4.4 3+1d anisotropic $\mathbb{Z}_N$ Laplacian model on a torus

In this section, we analyze the GSD and restricted mobility of the anisotropic $\mathbb{Z}_N$ Laplacian model on an $L_x \times L_y \times L_z$ cubic lattice with periodic boundary condition, i.e., Γ is a 2d torus graph $C_{L_x} \times C_{L_y}$. On $\Gamma = C_{L_x} \times C_{L_y}$, we have the following identification between the lattice points:

$$(x,y) \sim (x+L_x, y) \sim (x, y+L_y)\ . \tag{4.32}$$

Throughout this section, we use (x,y) to denote a vertex of the 2d torus graph, and reserve i to denote a vertex of a general graph Γ. Then, the discrete Laplacian operator Δ_L takes the more familiar form $\Delta_x^2 + \Delta_y^2$ in the xy-plane.

4.4.1 Upper bound on $\log_N \mathrm{GSD}$

It is clear from (4.17) that the GSD depends only on properties of Γ and is independent of L_z. Here, we give an upper bound on how fast $\log_N \mathrm{GSD}$ can grow with L_x, L_y.

Recall that the GSD of the anisotropic $\mathbb{Z}_N$ Laplacian model is $|\operatorname{Jac}(\Gamma, N)|^2$ (4.17). As we showed in Section 4.2.1, $|\operatorname{Jac}(\Gamma, N)|$ is also the number of equivalence classes under the equivalence relation "$\sim$" in (4.5). Combining these facts, we have

$$\mathrm{GSD} = \left| \frac{\{g(i)\}}{g(i) \sim g(i) + \Delta_L f(i)} \right|^2 \ , \tag{4.33}$$

where $g(i)$ and $f(i)$ are $\mathbb{Z}_N$-valued functions on the graph Γ.

When Γ is the 2d torus graph $C_{L_x} \times C_{L_y}$, interpreting the equivalence relation as a gauge symmetry, we can gauge fix $g(x, y) = 0 \mod N$ everywhere except at $x = 0, 1$, or at $y = 0, 1$. In other words, the number of sites where $g(x, y) \neq 0$ after gauge fixing is at most $2\min(L_x, L_y)$. Since $g(x, y)$ is $\mathbb{Z}_N$-valued, it follows that the number of nontrivial configurations of $g(x, y)$ is at most $N^{2\min(L_x, L_y)}$. Therefore,

$$\log_N \mathrm{GSD} \leq 4\min(L_x, L_y) \ . \tag{4.34}$$

4.4.2 $N = 2$

When $N = 2$, the stabilizer terms (4.12) simplify to

$$\begin{aligned} G(x, y, z) &= V_z(x, y, z + \tfrac{1}{2}) V_z(x, y, z - \tfrac{1}{2}) \prod_{\epsilon_x, \epsilon_y = \pm 1} V(x + \epsilon_x, y + \epsilon_y, z) \ , \\ F(x, y, z + \tfrac{1}{2}) &= U(x, y, z + 1) U(x, y, z) \prod_{\epsilon_x, \epsilon_y = \pm 1} U_z(x + \epsilon_x, y + \epsilon_y, z + \tfrac{1}{2}) \ . \end{aligned} \tag{4.35}$$

Here, the product of the four V's in G and the product of the four U_z's in F both involve only the four sites around the "$\frac{\pi}{4}$-tilted plaquette" centered at (x, y). This is illustrated in Figure 4.4. Therefore, each stabilizer term of the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model is equivalent to a stabilizer term of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model of [42, 58] on a "$\frac{\pi}{4}$-tilted" lattice.

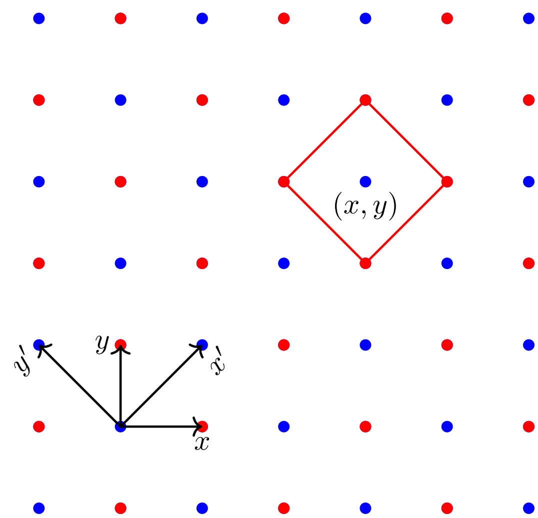


Figure 4.4: The stabilizer terms (4.35) of the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model are equivalent to those of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model of [42, 58] on a "$\frac{\pi}{4}$-tilted" lattice (the z-direction is suppressed in this figure). In particular, the product of V's in G and the product of U_z's in F involve only the four sites around the "$\frac{\pi}{4}$-tilted plaquette" centered at (x, y). Such a plaquette consists of only red sites, or only blue sites but not both. One such red tilted plaqutte is shown. The red and blue sublattices might or might not give independent copies of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model depending on the parities of L_x and L_y in the identifications (4.32). We use the coordinates (x', y') for the tilted lattice, which are related to the original coordinates (x, y) as $x' = \frac{x+y}{2}$ and $y' = \frac{x-y}{2}$. They are integers on the blue sublattice, and half-integers on the red sublattice.

Let us define the coordinates $(x', y') = (\frac{x+y}{2}, \frac{y-x}{2})$ for the tilted lattice. The tilted lattice decomposes into two sublattices: those with integral (x', y') and those with half-integral (x', y'). These are shown in blue and red in Figure 4.4. Observe that any tilted plaquette consists of sites from only one of the sublattices. Therefore, locally, there are two copies of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model, one on each sublattice. On an infinite lattice, these two copies are independent. However, the identifications (4.32) can couple them: when L_x and L_y are both even, the two sublattices are decoupled and there are two copies of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model, whereas when L_x or L_y is odd, the two sublattices are identified, so there is only one copy of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model. In all these cases, the identifications on the tilted lattice for the 3+1d anisotropic $\mathbb{Z}_2$ lineon model are given in (4.51), (4.54), and (4.58).

We present the GSD and mobility restrictions in this model for $L_x = L_y = L$ and refer the readers to Appendix 4.B on results for arbitrary L_x and L_y. The ground state degeneracy is given by

$$\mathrm{GSD} = \begin{cases} 2^{4L} \ , & L \text{ even} \ , \\ 2^{4L-2} \ , & L \text{ odd} \ . \end{cases} \tag{4.36}$$

This is in agreement with the plot for $N = 2$ in Figure 4.2, and it saturates the bound in (4.34) when L is even. Furthermore, a z-lineon cannot hop between different sites in the xy-plane. In contrast, a dipole of z-lineons separated in the $(1, \pm 1)$ direction can move in the $(1, \mp 1)$ direction in the xy-plane. These mobility restrictions follow from the relation between the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model and the 3+1d anisotropic $\mathbb{Z}_2$ lineon model on the tilted lattice.

To conclude, the $N = 2$ anisotropic Laplacian model is made out of the known anisotropic lineon model of [42, 58], with a relatively simple GSD (4.36). The next subsection discusses the anisotropic $\mathbb{Z}_p$ Laplacian model with p an odd prime, which is a genuinely new model and has a much more intricate GSD.

4.4.3 $N = p$ prime larger than 2

When $N = p$ is a prime larger than 2 we can follow [156] and use techniques from commutative algebra to compute the ground state degeneracy. We show that the GSD is given by (4.2). See Appendix 4.C.1 for the meaning and derivation of this expression.

The expression in (4.2) can be simplified in some special cases. Let $q \neq p$ be another odd prime such that p is a primitive root modulo q^m, where $m \geq 1$, i.e., p

is the generator of the multiplicative group of integers modulo q^m, denoted as $\mathbb{Z}_{q^m}^\times$.[16] Then, for $L_x = p^{k_x} q^m$ and $L_y = p^{k_y} q^m$, where $k_x, k_y, m \geq 0$, we show that

$$\log_p \mathrm{GSD} = 2\left[2p^{\min(k_x,k_y)} - \delta_{k_x,k_y}\right] \ . \tag{4.37}$$

We see that the bound (4.34) is saturated whenever $k_x \neq k_y$ and $m = 0$, i.e., there are infinitely many L_x, L_y for which $\log_p \mathrm{GSD}$ scales as $O(L_x, L_y)$. On the other hand, when $k_x = k_y = 0$, we have $\log_p \mathrm{GSD} = 2$ for any m, i.e., there are also infinitely many L_x, L_y for which $\log_p \mathrm{GSD}$ remains finite.

The last statement relies on the existence of an odd prime q such that p is a primitive root modulo q^m for all $m \geq 1$. A sufficient condition for this is that p is a primitive root modulo q^2 [163, Section 2.8]. For example, 3 is a primitive root modulo 5^2, so for $p = 3$, we can choose $q = 5$. Similarly, for $p = 5, 7$, we can choose $q = 7, 11$ respectively. In fact, one can verify numerically that for all $p \lesssim 10^9$, there is such a q. However, there is no proof of existence of such q for arbitrary p.

Interestingly, *Artin's conjecture on primitive roots* [164] states that there are infinitely many prime q such that p is a primitive root modulo q.[17] Whenever $L_x = L_y = q$ for any such q, we find that $\log_p \mathrm{GSD} = 2$. This gives another infinite family of L_x, L_y for which $\log_p \mathrm{GSD}$ remains finite. However, Artin's conjecture is still unproven, except under the assumption of the *generalized Riemann hypothesis* [165], which is also unproven.

We can apply similar techniques to determine the mobility of z-lineons in the xy-plane as well. (See Appendix 4.C.2 for more details.) There exist certain special values of L_x, L_y (e.g., $L_x = L_y = q^m$, where q is an odd prime such that p is a primitive root

[16] For any positive integer n, the set of all integers a such that $1 \leq a < n$ and $\gcd(a, n) = 1$ form a group under multiplication, known as the multiplicative group of integers modulo n, and denoted as $\mathbb{Z}_n^\times$. It is cyclic exactly when $n = 1, 2, 4, q^m$, or $2q^m$, where q is an odd prime and $m \geq 1$ [163, Section 2.8]. Whenever $\mathbb{Z}_n^\times$ is cyclic, it has a single generator, and the notion of "primitive root modulo n" is well-defined.

[17] Note that p being a primitive root modulo q does not imply that p is a primitive root modulo q^2.

modulo q^m) for which the z-lineons are completely mobile in the xy-plane. However, on an infinite square lattice, any finite set of z-lineons is completely immobile (unless they can be annihilated), assuming that their charges and the separations between them are fixed during the motion, i.e., they cannot move "rigidly."

It is surprising that the set of L_x, L_y for which $\log_p \mathrm{GSD}$ remains finite and the z-lineons are completely mobile is intimately related to well-known open problems in number theory.

4.A More on $\mathbb{Z}_N$-valued functions on a graph

In this appendix, we analyze the space of $\mathbb{Z}_N$-valued harmonic functions and the equivalence classes of $\mathbb{Z}_N$-valued functions on a general graph Γ. We use the Smith decomposition (4.6) of the Laplacian matrix L to give complete answers of these two questions mentioned in Section 4.2.1.

Recall that the Smith normal form of L is given by $R = PLQ$, where $P, Q \in GL_{\mathsf{N}}(\mathbb{Z})$, and $R = \mathrm{diag}(r_1, \ldots, r_{\mathsf{N}})$. Here, r_a's are nonnegative integers such that r_a divides r_{a+1} for $a = 1, \ldots, \mathsf{N} - 1$. While R is uniquely determined by L, the matrices P and Q are not.

In the index notation of (4.6), we have $R_{ab} = r_a \delta_{ab} = \sum_{i,j} P_{ai} L_{ij} Q_{jb}$. While all the indices here run from 1 to N, only i, j have a natural interpretation as vertices of the graph Γ.

We can now answer the first question raised in Section 4.2.1: find all the $\mathbb{Z}_N$-valued discrete harmonic functions. We first transform the $\mathbb{Z}_N$-valued function $h(i)$ to a new basis:

$$h'_a = \sum_i (Q^{-1})_{ai} h(i) \mod N \ , \tag{4.38}$$

In this basis, the discrete Laplace equation (4.4) is "diagonal":

$$r_a h'_a = 0 \mod N \ , \qquad a = 1, \ldots, \mathsf{N} \ . \tag{4.39}$$

We can solve this equation independently for each a. The most general solution is

$$h'_a = \frac{N p_a}{\gcd(N, r_a)} \mod N \ , \qquad a = 1, \ldots, \mathsf{N} \ , \tag{4.40}$$

where $p_a = 0, \ldots, \gcd(N, r_a) - 1$. Transforming back to the original basis, the most general $\mathbb{Z}_N$-valued discrete harmonic function is

$$h(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_a}{\gcd(N, r_a)} \mod N \ . \tag{4.41}$$

Let us now address the second question raised in Section 4.2.1: find all the equivalence classes under the equivalence relation "$\sim$". Since the Laplacian matrix L is symmetric, taking the transpose of $R = PLQ$ gives another Smith decomposition $R = Q^T L P^T$. Using this, we transform the $\mathbb{Z}_N$-valued function $g(i)$ to a (different) new basis

$$g''_a = \sum_i g(i) Q_{ia} \mod N \ . \tag{4.42}$$

We define $\tilde{g}''_a$ similarly for another function $\tilde{g}(i)$ in the same equivalence class. In this basis, the equivalence relation (4.5) is "diagonal":

$$\tilde{g}''_a - g''_a = r_a \hat{f}_a \mod N \ , \tag{4.43}$$

where $\hat{f}_a = \sum_i f(i) P^{-1}_{ia} \mod N$. Therefore, the equivalence class of $g(i)$ is completely determined by N congruence classes:

$$g''_a \mod \gcd(N, r_a) \ , \quad a = 1, \ldots, \mathsf{N} \ . \tag{4.44}$$

Going back to the original basis, a representative of the equivalence class "p_a mod $\gcd(N, r_a)$" is

$$g(i) = \sum_{a=1}^{\mathsf{N}} p_a (Q^{-1})_{ai} \mod N \ , \tag{4.45}$$

where $p_a = 0, \ldots, \gcd(N, r_a) - 1$ for $a = 1, \ldots, \mathsf{N}$.

4.B 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model

In this appendix, we use the relation between the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model of Section 4.4.2 and the 3+1d anisotropic $\mathbb{Z}_2$ lineon model to compute the GSD and restricted mobility of the former.

4.B.1 Ground state degeneracy

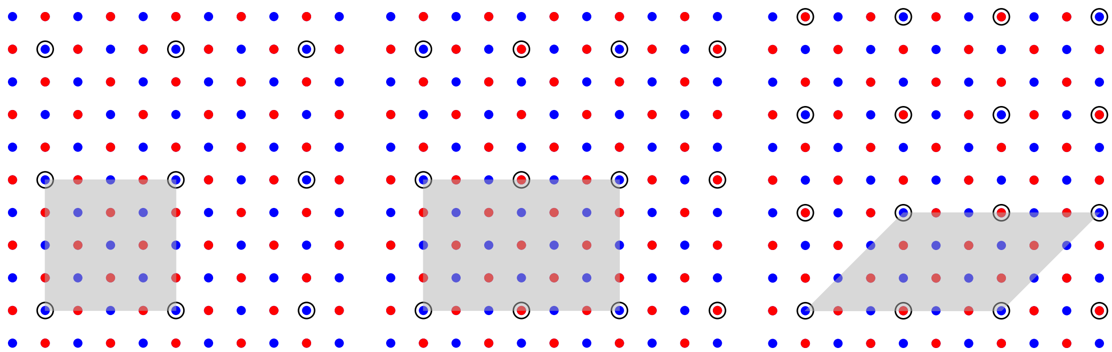
Recall that the stabilizer terms of the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model are given by (4.35), which are equivalent to those of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model of [42, 58] on a tilted lattice. Moreover, the latter is an anisotropic extension of the 2+1d $\mathbb{Z}_2$ Ising plaquette model [35] on the tilted lattice.

Now, the identifications on the original lattice are (4.32)

$$(x, y) \sim (x + L_x, y) \sim (x, y + L_y) \ , \tag{4.46}$$

In the new coordinates $(x', y') = (\frac{x+y}{2}, \frac{y-x}{2})$, the identifications on the tilted lattice take the schematic form

$$(x', y') \sim (x' + L^u_{x'}, y' + L^u_{y'}) \sim (x' + L^v_{x'}, y' + L^v_{y'}) \ . \tag{4.47}$$



(a) L_x, L_y even (b) L_x odd, L_y even (c) L_x, L_y odd

Figure 4.5: The minimal identifications on the blue sublattice of Figure 4.4 for different parities of L_x and L_y. The black circles represent the identifications (4.32), and the shaded regions represent the fundamental domains of the blue sublattice under these identifications.

The authors of [59] analyzed the 2+1d $\mathbb{Z}_2$ Ising plaquette model on a 2d spatial torus with such identifications. Their strategy was to reduce the identifications to the form

$$(x', y') \sim (x' + ML_{x'}^{\text{eff}}, y') \sim (x' + KL_{x'}^{\text{eff}}, y' + L_{y'}^{\text{eff}}) \ , \tag{4.48}$$

where $\gcd(M, K) = 1$. Then, they showed that

$$\text{GSD}_{2+1\text{d } \mathbb{Z}_2 \text{ Ising plaq}} = \gcd(2, M) \cdot 2^{L_{x'}^{\text{eff}} + L_{y'}^{\text{eff}} - 1} \ . \tag{4.49}$$

It follows that (see footnote 11)

$$\text{GSD}_{3+1\text{d aniso } \mathbb{Z}_2 \text{ lineon}} = \left[\gcd(2, M) \cdot 2^{L_{x'}^{\text{eff}} + L_{y'}^{\text{eff}} - 1}\right]^2 \ . \tag{4.50}$$

Let us use these results to compute the GSD of the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model:

- L_x, L_y even: In this case, the identifications in (4.32) do not couple the two sublattices. So, there are two independent copies of the the 3+1d anisotropic $\mathbb{Z}_2$ lineon model, and we can work with one copy at a time, say the blue sublattice in Figure 4.4. In the new coordinates (x', y'), the minimal identifications on the blue sublattice are (see Figure 4.5(a))

$$(x', y') \sim (x' + \tfrac{L_x}{2}, y' - \tfrac{L_x}{2}) \sim (x' + \tfrac{L_y}{2}, y' + \tfrac{L_y}{2}) \ . \tag{4.51}$$

Let $(\tilde{L}_x, \tilde{L}_y)$ be the integer solution of the equation $\tilde{L}_x L_x + \tilde{L}_y L_y = \gcd(L_x, L_y)$. Then, in the notation of (4.48), we have

$$ML^{\text{eff}}_{x'} = \text{lcm}(L_x, L_y) \ , \qquad KL^{\text{eff}}_{x'} = \frac{1}{2}(\tilde{L}_x L_x - \tilde{L}_y L_y) \ , \qquad L^{\text{eff}}_{y'} = \frac{1}{2}\gcd(L_x, L_y) \ , \tag{4.52}$$

and hence,

$$\text{GSD} = \left[\gcd(2, M) \cdot 2^{L^{\text{eff}}_{x'} + L^{\text{eff}}_{y'} - 1}\right]^4 \ . \tag{4.53}$$

Here, the power is 4 rather than 2 because there are two copies of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model, one on each sublattice.

- <u>L_x odd, L_y even</u>: In this case, the identifications in (4.32) couple the two sublattices so that effectively there is only one sublattice, say the blue sublattice in Figure 4.4. Hence, there is only one copy of the the 3+1d anisotropic $\mathbb{Z}_2$ lineon model. The minimal identifications on the blue sublattice are (see Figure 4.5(b))

$$(x, y) \sim (x + 2L_x, y) \sim (x, y + L_y) \ , \tag{4.54}$$

which can be written as

$$(x', y') \sim (x' + L_x, y' - L_x) \sim (x' + \tfrac{L_y}{2}, y' + \tfrac{L_y}{2}) \ , \tag{4.55}$$

in the new coordinates (x', y'). Let $(\tilde{L}_x, \tilde{L}_y)$ be the integer solution of the equation $\tilde{L}_x L_x + \tilde{L}_y(\frac{L_y}{2}) = \gcd(L_x, \frac{L_y}{2})$. Then, in the notation of (4.48), we have

$$ML^{\text{eff}}_{x'} = 2\,\text{lcm}(L_x, \tfrac{L_y}{2}) \ , \qquad KL^{\text{eff}}_{x'} = \tilde{L}_x L_x - \frac{1}{2}\tilde{L}_y L_y \ , \qquad L^{\text{eff}}_{y'} = \gcd(L_x, \frac{L_y}{2}) \ , \tag{4.56}$$

and hence,

$$\text{GSD} = \left[\gcd(2, M) \cdot 2^{L^{\text{eff}}_{x'} + L^{\text{eff}}_{y'} - 1}\right]^2 \ . \tag{4.57}$$

- $\underline{L_x, L_y \text{ odd}}$: In this case, once again, the identifications in (4.32) couple the two sublattices so that effectively there is only one sublattice, say the blue sublattice in Figure 4.4. Hence, there is only one copy of the the 3+1d anisotropic $\mathbb{Z}_2$ lineon model. The minimal identifications on the blue sublattice are (see Figure 4.5(c))

$$(x, y) \sim (x + 2L_x, y) \sim (x + L_x, y + L_y) \ , \tag{4.58}$$

which can be written as

$$(x', y') \sim (x' + L_x, y' - L_x) \sim (x' + \tfrac{L_x + L_y}{2}, y' - \tfrac{L_x - L_y}{2}) \ , \tag{4.59}$$

in the new coordinates (x', y'). Let $(\tilde{L}_x, \tilde{L}_y)$ be the integer solution of the equation $\tilde{L}_x L_x + \tilde{L}_y(\frac{L_x - L_y}{2}) = \gcd(L_x, \frac{L_x - L_y}{2})$. Then, in the notation of (4.48), we have

$$ML^{\text{eff}}_{x'} = \frac{L_x L_y}{\gcd(L_x, \frac{L_x - L_y}{2})} \ , \qquad KL^{\text{eff}}_{x'} = \tilde{L}_x L_x + \tilde{L}_y(\tfrac{L_x + L_y}{2}) \ , \qquad L^{\text{eff}}_{y'} = \gcd(L_x, \tfrac{L_x - L_y}{2}) \ , \tag{4.60}$$

and hence,

$$\text{GSD} = \left[\gcd(2, M) \cdot 2^{L^{\text{eff}}_{x'} + L^{\text{eff}}_{y'} - 1}\right]^2 \ . \tag{4.61}$$

When $L_x = L_y = L$, we have

$$L^{\text{eff}}_{x'} = L^{\text{eff}}_{y'} = \frac{L}{M} \ , \qquad K = 1 \ , \qquad M = \begin{cases} 2 \ , & L \text{ even} \ , \\ 1 \ , & L \text{ odd} \ . \end{cases} \tag{4.62}$$

Then, the above expressions for the ground state degeneracy simplify to

$$\text{GSD} = \begin{cases} 2^{4L} \ , & L \text{ even} \ , \\ 2^{4L-2} \ , & L \text{ odd} \ . \end{cases} \tag{4.63}$$

4.B.2 Mobility restrictions

Let us now discuss the mobility of the z-lineons in the xy-plane in the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model. First, note that in the 2+1d $\mathbb{Z}_2$ Ising plaquette model on a 2d spatial torus given by the identifications (4.48), the defect describing a single particle of unit charge can "hop" from (x', y') to $(x' + \gcd(2, M)L^{\text{eff}}_{x'}, y')$ [25]. This motion is nontrivial if and only if $\gcd(2, M) = 1$ and $M > 1$. It follows that the z-lineons of the 3+1d anisotropic $\mathbb{Z}_2$ lineon model, in addition to moving along the z direction, can "hop" in the xy-plane in the same way.

In particular, consider $L_x = L_y = L$. It follows from (4.62) that the two conditions $\gcd(2, M) = 1$ and $M > 1$ cannot be simultaneously satisfied for any L. Therefore, the z-lineons cannot hop in the xy-plane, but can only move in the z-direction (and hence the name lineon).

A dipole of z-lineons at (x, y) and $(x + s, y \pm s)$, where $s \in \mathbb{Z}$, can move in the $(1, \mp 1)$ direction. This follows from the motion of a dipole of fractons in the 2+1d $\mathbb{Z}_2$ Ising plaquette model in the direction orthogonal to their separation.

4.C 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model

In this appendix, we use techniques from commutative algebra to compute the ground state degeneracy and analyze the restricted mobility of the 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model of Section 4.4.3 when $p > 2$ is prime. Such techniques were used to analyze translationally invariant Pauli stabilizer codes [156]. All the mathematical facts used here can be found in standard textbooks on the subject, such as [166, 167].

4.C.1 Ground state degeneracy

Recall the relation (4.33) between the GSD of the anisotropic $\mathbb{Z}_N$ Laplacian model and the number of equivalence classes of $\mathbb{Z}_N$-valued functions under the equivalence relation "$\sim$" of (4.5). Here, we set $N = p > 2$, where p is prime,[18] and $\Gamma = C_{L_x} \times C_{L_y}$, the 2d torus graph. We first find an exact expression for the $\log_p \mathrm{GSD}$ in terms of commutative-algebraic quantities using the relation (4.33), then explain how to compute this expression in general using a *Gröbner basis*, and finally compute it explicitly for some special values of L_x, L_y. In particular, we show that there are infinitely many L_x, L_y for which $\log_p \mathrm{GSD}$ is $O(L_x, L_y)$, and also infinitely many L_x, L_y for which $\log_p \mathrm{GSD}$ is finite.

Exact expression for $\log_p \mathrm{GSD}$

Since there are $L_x L_y$ points in the $\Gamma = C_{L_x} \times C_{L_y}$ torus and since we are interested in $\mathbb{Z}_p$-valued functions on that space, it is clear that there are $p^{L_x L_y}$ such functions. As in (4.5), they fall into equivalence classes $g(x, y) \sim \tilde{g}(x, y)$ when $g(x, y) - \tilde{g}(x, y) = \Delta_L f(x, y) \mod p$. We would like to find the number of such equivalence classes.

As a first step, we give a more abstract description of these $p^{L_x L_y}$ functions. Let $\mathcal{R} = \mathbb{Z}_p[X, Y]$ be the ring of polynomials with coefficients in $\mathbb{Z}_p$, and $\mathsf{j} = (\mathsf{Q}_x, \mathsf{Q}_y)$ be

[18] Actually, all of the following discussion up to (4.99) works even for $p = 2$. The discussion after that does not work for $p = 2$ for reasons we will explain later.

the *ideal* of $\mathcal{R}$ generated by the polynomials $\mathsf{Q}_x(X,Y) = X^{L_x} - 1$ and $\mathsf{Q}_y(X,Y) = Y^{L_y} - 1$. Given two polynomials $\mathsf{F}, \mathsf{G} \in \mathcal{R}$, we write[19]

$$\mathsf{F}(X,Y) = \mathsf{G}(X,Y) \mod \mathfrak{j} \ , \tag{4.64}$$

if and only if $\mathsf{F}(X,Y) - \mathsf{G}(X,Y)$ is a polynomial in $\mathfrak{j}$. The set of equivalence classes modulo $\mathfrak{j}$ is the quotient ring $\mathcal{R}/\mathfrak{j}$.

Any equivalence class of $\mathcal{R}/\mathfrak{j}$ is represented by a unique polynomial that is a $\mathbb{Z}_p$-linear combination of the monomials $X^a Y^b$ with $0 \le a < L_x$ and $0 \le b < L_y$. (Here, we used the equivalence relations to remove higher powers of X or Y. This is a special case of a more general procedure, called complete reduction, which we will describe below.) Therefore, the number of equivalence classes is $p^{L_x L_y}$. In fact, since $\mathbb{Z}_p$ is a field, $\mathcal{R}/\mathfrak{j}$ can be thought of as a vector space over $\mathbb{Z}_p$. The above monomials form a basis of this vector space, so

$$\dim_{\mathbb{Z}_p} \mathcal{R}/\mathfrak{j} = L_x L_y \ . \tag{4.65}$$

Here, "$\dim_{\mathbb{Z}_p}$" denotes the dimension of a vector space over $\mathbb{Z}_p$.

It is convenient to represent a $\mathbb{Z}_p$-valued function $f(x,y)$ on the 2d torus graph $\Gamma = C_{L_x} \times C_{L_y}$ as a polynomial representing an equivalence class of the quotient ring $\mathcal{R}/\mathfrak{j}$ as follows:

$$\hat{f}(X,Y) = \sum_{x=0}^{L_x-1} \sum_{y=0}^{L_y-1} f(x,y) X^{L_x-x-1} Y^{L_y-y-1} \mod \mathfrak{j} \ . \tag{4.66}$$

$\hat{f}(X,Y)$ can be thought of as a lattice Fourier transform of $f(x,y)$ with $X = e^{ik_x}$ and $Y = e^{ik_y}$, which depend on the momenta k_x and k_y.

[19]We do not write the "mod p" explicitly because we are working in $\mathbb{Z}_p[X,Y]$.

Observe that, for any integer $0 \le k < L_x$, we have

$$X^k \hat{f}(X,Y) = \Bigg[\sum_{x=0}^{L_x-k-1} \sum_{y=0}^{L_y-1} f(x+k,y) X^{L_x-x-1} Y^{L_y-y-1} \\ + \sum_{x=L_x-k}^{L_x-1} \sum_{y=0}^{L_y-1} f(x-L_x+k,y) X^{L_x-x-1} Y^{L_y-y-1} \Bigg] \mod \mathfrak{j} \ , \tag{4.67}$$

so X can be interpreted as the generator of translations in the x direction.[20] The fact that translating in x by L_x takes the graph C_{L_x} back to itself is related to the trivial equation

$$X^{L_x} \hat{f}(X,Y) = \hat{f}(X,Y) \mod \mathfrak{j} \ . \tag{4.68}$$

Also, the difference operator Δ_x is associated with the polynomial $X-1$. For convenience, we define the *displaced* discrete Laplacian operator, denoted by $\tilde{\Delta}_L$, as

$$\tilde{\Delta}_L f(x,y) = (\Delta_x^2 + \Delta_y^2) f(x+1,y+1) \ . \tag{4.69}$$

(Here, we extended $f(x,y)$ to a periodic function on $\mathbb{Z}^2$.) It is associated with the polynomial

$$\tilde{\mathsf{P}}(X,Y) = Y(X-1)^2 + X(Y-1)^2 \ . \tag{4.70}$$

In general, any *local* difference operator in the xy-plane, after an appropriate displacement, is associated with a polynomial $\mathsf{S} \in \mathcal{R}$ satisfying $\mathsf{S}(1,1) = 0$.

Let $\mathfrak{i} = (\tilde{\mathsf{P}})$ be the ideal of $\mathcal{R}$ generated by the polynomial $\tilde{\mathsf{P}}(X,Y)$. Then, $\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j}) \cong (\mathfrak{i}+\mathfrak{j})/\mathfrak{j}$ is an ideal of the quotient ring $\mathcal{R}/\mathfrak{j}$.[21] In fact, it is the subspace

[20] Our functions $f(x,y)$ are defined on $L_x L_y$ points. They can be extended to periodic functions on $\mathbb{Z}^2$. Then, it is straightforward to apply k translations. The expression (4.67) corresponds to applying such k translation and then expressing the result in terms of the function $f(x,y)$ in the fundamental domain $0 \le x \le L_x - 1$, $0 \le y \le L_y - 1$. Equivalently, we can use the periodicity of f to write (4.67) as $X^k \hat{f}(X,Y) = \sum_{x=0}^{L_x-1} \sum_{y=0}^{L_y-1} f(x+k,y) X^{L_x-x-1} Y^{L_y-y-1} \mod \mathfrak{j}$.

[21] Here, $\mathfrak{i}+\mathfrak{j} = (\tilde{\mathsf{P}}, \mathsf{Q}_x, \mathsf{Q}_y)$ is the ideal of $\mathcal{R}$ generated by the three polynomials $\tilde{\mathsf{P}}(X,Y)$, $\mathsf{Q}_x(X,Y)$, and $\mathsf{Q}_y(X,Y)$. It is known as the *sum of ideals* $\mathfrak{i}$ and $\mathfrak{j}$.

of the vector space $\mathcal{R}/\mathfrak{j}$ that corresponds to the $\mathbb{Z}_p$-valued functions of the form $\tilde{\Delta}_L f(x,y)$.

It is then clear that the quotient

$$\frac{\mathcal{R}/\mathfrak{j}}{(\mathfrak{i}+\mathfrak{j})/\mathfrak{j}} \cong \mathcal{R}/(\mathfrak{i}+\mathfrak{j}) \ , \tag{4.71}$$

corresponds to the set of equivalence classes of $\mathbb{Z}_p$-valued functions under the equivalence relation "$\sim$" of (4.5) on the 2d torus graph $C_{L_x} \times C_{L_y}$. It follows that

$$\log_p \mathrm{GSD} = 2 \dim_{\mathbb{Z}_p} \mathcal{R}/(\mathfrak{i}+\mathfrak{j}) \ , \tag{4.72}$$

or more explicitly,

$$\log_p \mathrm{GSD} = 2 \dim_{\mathbb{Z}_p} \frac{\mathbb{Z}_p[X,Y]}{(Y(X-1)^2 + X(Y-1)^2, X^{L_x}-1, Y^{L_y}-1)} \ . \tag{4.73}$$

Below, we give a general procedure to compute this quantity.

Computing $\log_p \mathrm{GSD}$ using Gröbner basis

One way to compute the (vector space) dimension of $\mathcal{R}/(\mathfrak{i}+\mathfrak{j})$ is by first computing the *Gröbner basis* of the ideal $\mathfrak{i}+\mathfrak{j}$. Before defining a Gröbner basis, we need an ordering on all the monomials. For our purposes, it is sufficient to define a *lexicographic monomial ordering*: $X^m Y^n \succ X^k Y^l$ if and only if $m > k$, or $m = k$ and $n > l$. Then, for any polynomial $\mathsf{F}(X,Y)$, we define its *leading term* as the term that has the largest monomial among all the terms.

Now, we say a polynomial $\mathsf{F}(X,Y)$ is *reducible* with respect to a set of polynomials $\mathscr{G} = \{\mathsf{G}_1, \ldots, \mathsf{G}_n\}$ if some term of $\mathsf{F}(X,Y)$ is a multiple of the leading term of one of the $\mathsf{G}_i(X,Y)$'s. We say it is *irreducible* otherwise.

Given an ideal $\mathfrak{I}$ of $\mathcal{R}$ and a polynomial $\mathsf{F} \in \mathcal{R}$, one can ask what the equivalence class of $\mathsf{F}(X,Y)$ in $\mathcal{R}/\mathfrak{I}$ is. One way to answer this question is to *reduce* $\mathsf{F}(X,Y)$ with respect to a generating set $\mathscr{B}$ of $\mathfrak{I}$ repeatedly until we are left with a polynomial $\mathsf{H}(X,Y)$ that is irreducible with respect to $\mathscr{B}$. (Such a procedure is known as a *complete reduction* of $\mathsf{F}(X,Y)$ with respect to $\mathscr{B}$.) One might hope that $\mathsf{H}(X,Y)$ uniquely specifies the equivalence class of $\mathsf{F}(X,Y)$. However, for a generic $\mathscr{B}$, the $\mathsf{H}(X,Y)$ so obtained depends on the choices made in the repeated reductions, and hence, may not be unique.

A Gröbner basis $\mathscr{G} = \{\mathsf{G}_1, \ldots, \mathsf{G}_n\}$ is a special generating set of $\mathfrak{I}$ such that $\mathsf{F}(X,Y)$ can be written as

$$\mathsf{F}(X,Y) = \mathsf{H}(X,Y) + \sum_{i=1}^{n} \mathsf{H}_i(X,Y)\mathsf{G}_i(X,Y)\ , \tag{4.74}$$

where $\mathsf{H}(X,Y)$ is *uniquely* determined by the requirement that it is irreducible with respect to $\mathscr{G}$.[22] In this case, we write

$$\mathsf{F}(X,Y) = \mathsf{H}(X,Y) \mod \mathscr{G}\ . \tag{4.75}$$

It follows that there is a one-one correspondence between $\mathcal{R}/\mathfrak{I}$ and the set of all polynomials that are irreducible with respect to $\mathscr{G}$. Indeed, the set of all monomials that are irreducible with respect to $\mathscr{G}$ forms a basis of the vector space $\mathcal{R}/\mathfrak{I}$. From this we conclude that $\dim_{\mathbb{Z}_p} \mathcal{R}/\mathfrak{I}$ equals the number of monomials that are irreducible with respect to $\mathscr{G}$.

While there is an algorithm, known as *Buchberger's algorithm*, to compute a Gröbner basis of an ideal given its generators, it is not always easy to compute it

[22] Note that $\mathsf{H}_i(X,Y)$'s are not uniquely determined by this procedure. Indeed, shifting $\mathsf{H}_1(X,Y)$ by $\mathsf{G}_2(X,Y)$ and $\mathsf{H}_2(X,Y)$ by $-\mathsf{G}_1(X,Y)$ gives another complete reduction of $\mathsf{F}(X,Y)$ with respect to $\mathscr{G}$.

analytically.[23] Nonetheless, for fixed values of p, L_x, L_y, the Gröbner basis, and therefore the GSD, can be readily computed with the help of computer programs. Let us do an explicit calculation for the GSD when $p = 3$ and $L_x = L_y = 4$ as an example. Using the GroebnerBasis command with Modulus $\to$ 3 in Mathematica, we can compute the Gröbner basis for the ideal $(\tilde{\mathsf{P}}, \mathsf{Q}_x, \mathsf{Q}_y) = (Y(X-1)^2 + X(Y-1)^2, X^4 - 1, Y^4 - 1)$ in this case. We find $\mathscr{G} = \{Y^4 - 1, XY + X - Y^3 - Y^2, X^2 + Y^3 - Y^2 + Y + 1\}$ under the lexicographic ordering $X \succ Y$. The leading terms of $\mathscr{G}$ are $\{Y^4, XY, X^2\}$, and the 5 irreducible monomials with respect to $\mathscr{G}$ are $1, Y, Y^2, Y^3, X$. We conclude that $\dim_{\mathbb{Z}_3} \mathcal{R}/(\mathfrak{i}+\mathfrak{j}) = 5$, and hence $\log_3 \mathrm{GSD} = 10$. In fact, the plot of $\log_N \mathrm{GSD}$ as a function of $L_x = L_y = L$ in Figure 4.2 was obtained exactly in this way.

Fortunately, for the problem at hand, it is possible to simplify the expression (4.72) for the GSD further analytically. To prepare for the simplified expression, we first define a few things:

A field F is *algebraically closed* if any polynomial in $F[X]$ can be factorized completely into linear factors in $F[X]$. For example, $\mathbb{R}$ is not algebraically closed because the polynomial $X^2 + 1$, which is in $\mathbb{R}[X]$, cannot be factorized into linear factors in $\mathbb{R}[X]$. In contrast, it is well-known that $\mathbb{C}$ is algebraically closed. Every field, by definition, contains the multiplicative identity 1.

The *characteristic* of a field is the smallest positive integer n such that $1 + \cdots + 1 (n \text{ times}) = 0$. It is clear that $\mathbb{Z}_p$ is a field of characteristic p. By convention, the characteristics of $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all defined to be 0. It is known that the characteristic of any field is either 0 or a prime number. As we see, there are several fields with the same characteristic p.

For *finite fields*, the field is uniquely specified by the number of its elements and the characteristic. In particular, for any integer $k \geq 1$ and prime p, $\mathbb{F}_{p^k}$ denotes

[23] For the ideal $\mathfrak{j} = (\mathsf{Q}_x, \mathsf{Q}_y)$, the generating set $\{\mathsf{Q}_x, \mathsf{Q}_y\}$ is already a Gröbner basis, and moreover, the irreducible monomials are $X^x Y^y$ for $0 \leq x < L_x$ and $0 \leq y < L_y$, which form a basis of $\mathcal{R}/\mathfrak{j}$. (This fact was used in the analysis leading to (4.65).) However, for the ideal $\mathfrak{i}+\mathfrak{j} = (\tilde{\mathsf{P}}, \mathsf{Q}_x, \mathsf{Q}_y)$, the generating set $\{\tilde{\mathsf{P}}, \mathsf{Q}_x, \mathsf{Q}_y\}$ is not always a Gröbner basis.

the unique (up to field isomorphisms) finite field of order p^k and characteristic p. Furthermore, $\mathbb{F}_{p^\infty}$ denotes the unique algebraically closed field of characteristic p.[24]

Let $\mathcal{S} = \mathbb{F}_{p^\infty}[X,Y]$ be the ring of polynomials in X,Y with coefficients in $\mathbb{F}_{p^\infty}$. We use the same symbols $\mathfrak{i}$ and $\mathfrak{j}$ for the ideals of $\mathcal{S}$ generated by the polynomials $\tilde{\mathsf{P}}(X,Y)$ and $\mathsf{Q}_x(X,Y), \mathsf{Q}_y(X,Y)$ respectively. The *algebraic set* of $\mathfrak{i}+\mathfrak{j}$, denoted as $V(\mathfrak{i}+\mathfrak{j})$, is the set of distinct solutions $(X_0,Y_0) \in \mathbb{F}^2_{p^\infty}$ of the system of polynomial equations

$$X^{L_x} - 1 = Y^{L_y} - 1 = \tilde{\mathsf{P}}(X,Y) = 0 \ . \tag{4.76}$$

We will find it convenient to parameterize L_x and L_y in terms of the order of our group $\mathbb{Z}_p$ as

$$L_x = p^{k_x} L'_x \ , \qquad L_y = p^{k_y} L'_y \ , \qquad \gcd(p, L'_x) = \gcd(p, L'_y) = 1 \ . \tag{4.77}$$

Now, for each solution $(X_0,Y_0) \in V(\mathfrak{i}+\mathfrak{j})$, we define the polynomials

$$\tilde{\mathsf{Q}}_{x,X_0}(X,Y) = (X - X_0)^{p^{k_x}} \ , \qquad \tilde{\mathsf{Q}}_{y,Y_0}(X,Y) = (Y - Y_0)^{p^{k_y}} \ , \tag{4.78}$$

and the ideal $\mathfrak{i}'_{X_0,Y_0} = (\tilde{\mathsf{P}}, \tilde{\mathsf{Q}}_{x,X_0}, \tilde{\mathsf{Q}}_{y,Y_0})$ of $\mathcal{S}$.

With these preparations, the simplified expression for the GSD (4.72) is

$$\log_p \mathrm{GSD} = 2 \sum_{(X_0,Y_0)\in V(\mathfrak{i}+\mathfrak{j})} \dim_{\mathbb{F}_{p^\infty}} \mathcal{S}/\mathfrak{i}'_{X_0,Y_0} \ . \tag{4.79}$$

It is obtained using techniques from commutative algebra (see e.g., [166, 167]), that were used in [156]. For readers who are familiar with such techniques, a derivation of (4.79) is given below. Others, who are willing to accept it, can skip directly

[24] More explicitly, $\mathbb{F}_{p^\infty} = \bigcup_{n=1}^{\infty} \mathbb{F}_{p^{n!}}$. Here we used the fact that $\mathbb{F}_{p^k}$ is a subfield of $\mathbb{F}_{p^m}$ if and only if k divides m.

to Appendix 4.C.1, where we compute the GSD for some special values of L_x, L_y explicitly.

Derivation of (4.79)

Since $\mathcal{R}/(\mathfrak{i}+\mathfrak{j})$ is a finite-dimensional vector space over $\mathbb{Z}_p$, it is an *Artinian* ring,[25] and hence, it has finitely many *maximal ideals*.[26] Moreover, for an Artinian ring, it is known that

$$\mathcal{R}/(\mathfrak{i}+\mathfrak{j}) \cong \bigoplus_{\mathfrak{m}} [\mathcal{R}/(\mathfrak{i}+\mathfrak{j})]_{\mathfrak{m}} \ , \tag{4.80}$$

where the sum is over all maximal ideals of $\mathcal{R}/(\mathfrak{i}+\mathfrak{j})$, and $[\mathcal{R}/(\mathfrak{i}+\mathfrak{j})]_{\mathfrak{m}}$ denotes the *localization* of $\mathcal{R}/(\mathfrak{i}+\mathfrak{j})$ at $\mathfrak{m}$.[27] It follows that

$$\log_p \mathrm{GSD} = 2 \sum_{\mathfrak{m}} \dim_{\mathbb{Z}_p} [\mathcal{R}/(\mathfrak{i}+\mathfrak{j})]_{\mathfrak{m}} \ . \tag{4.81}$$

So, we can compute dimension for each term in the sum and then add them up. However, the maximal ideals of $\mathcal{R}/(\mathfrak{i}+\mathfrak{j})$ are a bit complicated to work with. Instead, we proceed as follows.

[25] A ring $\mathcal{R}$ is Artinian if it satisfies the *descending chain condition*, i.e., if $\mathfrak{I}_1 \supseteq \mathfrak{I}_2 \supseteq \mathfrak{I}_3 \supseteq \cdots$ is a descending chain of ideals, then there is a $k \geq 1$ such that $\mathfrak{I}_k = \mathfrak{I}_{k+1} = \mathfrak{I}_{k+2} = \cdots$. For example, the ring of integers $\mathbb{Z}$ is not Artinian because $(2) \supset (4) \supset (8) \supset \cdots$, where (n) denotes the ideal generated by the integer n. On the other hand, for any integer n, the ring of integers modulo n, $\mathbb{Z}/(n) = \mathbb{Z}_n$, is Artinian. Moreover, any ring that is also a finite dimensional vector space over a field is always Artinian, which is exactly the case here.

[26] A proper ideal is an ideal that is not the ring itself. For example, in $\mathbb{Z}$, the ideal (4) is proper because it does not contain 1. (The only ideal containing 1 is the entire ring itself.) A maximal ideal is a proper ideal that is not contained in any other proper ideal except itself. For example, in $\mathbb{Z}$, (4) is not a maximal ideal because it is contained in the proper ideal (2). The latter is maximal; in fact, the ideal (n) is maximal if and only if n is prime.

[27] Intuitively, given a multiplicatively closed subset S of a ring $\mathcal{R}$, the localization of $\mathcal{R}$ with respect to S, denoted as $S^{-1}\mathcal{R}$, means "formally adding multiplicative inverses" for all the elements of S. For example, in $\mathbb{Z}$, the subset of nonzero integers is multiplicatively closed, and localizing with respect to this set gives the rationals $\mathbb{Q}$. In any ring $\mathcal{R}$, given a maximal ideal $\mathfrak{m}$, the set $\mathcal{R} \setminus \mathfrak{m}$ is always multiplicatively closed. So we define the localization of $\mathcal{R}$ at $\mathfrak{m}$, denoted by $\mathcal{R}_{\mathfrak{m}}$, as the localization of $\mathcal{R}$ with respect to $\mathcal{R} \setminus \mathfrak{m}$.

We can replace $\mathbb{Z}_p$ by $\mathbb{F}_{p^\infty}$ in (4.72) and get the same answer for $\log_p \mathrm{GSD}$. More concretely, we have[28]

$$\dim_{\mathbb{Z}_p} \mathcal{R}/(\mathfrak{i}+\mathfrak{j}) = \dim_{\mathbb{F}_{p^\infty}} \mathcal{S}/(\mathfrak{i}+\mathfrak{j}) \ , \tag{4.82}$$

where, on the right hand side, $\mathfrak{i}$ and $\mathfrak{j}$ are ideals of $\mathcal{S}$ generated by the same polynomials as before. Since $\mathcal{S}/(\mathfrak{i}+\mathfrak{j})$ is also Artinian, we have

$$\log_p \mathrm{GSD} = 2\sum_{\mathfrak{m}} \dim_{\mathbb{F}_{p^\infty}} [\mathcal{S}/(\mathfrak{i}+\mathfrak{j})]_{\mathfrak{m}} \ , \tag{4.83}$$

where the sum is over all maximal ideals of $\mathcal{S}/(\mathfrak{i}+\mathfrak{j})$. We now characterize the maximal ideals of $\mathcal{S}/(\mathfrak{i}+\mathfrak{j})$.

By the *correspondence theorem for quotient rings*, the maximal ideals of $\mathcal{S}/(\mathfrak{i}+\mathfrak{j})$ are in one-one correspondence with the maximal ideals of $\mathcal{S}$ that contain the ideal $\mathfrak{i}+\mathfrak{j}$. Moreover, since $\mathbb{F}_{p^\infty}$ is algebraically closed, by *Hilbert's Nullstellensatz*, the maximal ideals of $\mathcal{S}$ are in one-one correspondence with ideals of the form $(X - X_0, Y - Y_0)$, where $(X_0, Y_0) \in \mathbb{F}_{p^\infty}^2$. Now, the ideal $(X - X_0, Y - Y_0)$ contains $\mathfrak{i}+\mathfrak{j}$ if and only if (X_0, Y_0) is a root of all the polynomials in $\mathfrak{i}+\mathfrak{j}$. It follows that the maximal ideals of $\mathcal{S}/(\mathfrak{i}+\mathfrak{j})$ are in one-one correspondence with ideals of $\mathcal{S}$ of the form $(X - X_0, Y - Y_0)$, where (X_0, Y_0) is a solution of the system of polynomial equations

$$X^{L_x} - 1 = Y^{L_y} - 1 = \tilde{\mathsf{P}}(X, Y) = 0 \ , \tag{4.84}$$

[28]Given a field F, one can talk about extending it to a larger field F' such that F is a subfield of F'. For example, $\mathbb{R}$ is a subfield of $\mathbb{C}$, or equivalently, $\mathbb{C}$ is an extension of $\mathbb{R}$. Now, let V be a vector space over F. One can "extend the base field" from F to F' by tensoring V with F', denoted as $V \otimes_F F'$. Then, $\dim_F V = \dim_{F'}(V \otimes_F F')$. Since $\mathbb{F}_{p^\infty}$ is an extension of $\mathbb{Z}_p$, the result in (4.82) follows.

i.e., $(X_0, Y_0) \in V(\mathfrak{i}+\mathfrak{j})$, the algebraic set of $\mathfrak{i}+\mathfrak{j}$. Then,

$$\log_p \mathrm{GSD} = 2 \sum_{(X_0,Y_0)\in V(\mathfrak{i}+\mathfrak{j})} \dim_{\mathbb{F}_{p^\infty}} [\mathcal{S}/(\mathfrak{i}+\mathfrak{j})]_{(X-X_0,Y-Y_0)} \ . \tag{4.85}$$

We can simplify it further. Let $L_i = p^{k_i} L_i'$ with $\gcd(p, L_i') = 1$. Recall that since $\mathbb{F}_{p^\infty}$ is algebraically, any polynomial in $\mathbb{F}_{p^\infty}[X]$ can be factorized completely into linear factors in $\mathbb{F}_{p^\infty}[X]$. Using this fact, we have

$$X^{L_x} - 1 = (X^{L_x'} - 1)^{p^{k_x}} = \prod_{\xi\in\mathbb{F}_{p^\infty} \ : \ \xi^{L_x'}=1} (X-\xi)^{p^{k_x}} \ , \tag{4.86}$$

That is, there are L_x' distinct ξ's in $\mathbb{F}_{p^\infty}$ that satisfy $\xi^{L_x} = 1$, each with multiplicity p^{k_x}.

Let $(X_0, Y_0) \in V(\mathfrak{i}+\mathfrak{j})$. It is clear that X_0 is one of the ξ's in (4.86). Consider the localization $\mathcal{S}_{(X-X_0,Y-Y_0)}$. Recall that every element outside the ideal $(X-X_0, Y-Y_0)$ becomes a *unit* in the localization $\mathcal{S}_{(X-X_0,Y-Y_0)}$.[29] In particular, the polynomial $\mathsf{Q}_x(X,Y) = X^{L_x} - 1$ generates the same ideal in $\mathcal{S}_{(X-X_0,Y-Y_0)}$ as the polynomial $\tilde{\mathsf{Q}}_{x,X_0}(X,Y) = (X-X_0)^{p^{k_x}}$ because the other linear factors in (4.86) associated with $\xi \neq X_0$ all have inverses.[30] Similarly, $\mathsf{Q}_y(X,Y) = Y^{L_y} - 1$ generates the same ideal as $\tilde{\mathsf{Q}}_{y,Y_0}(X,Y) = (Y-Y_0)^{p^{k_y}}$. It follows that

$$(\mathfrak{i}+\mathfrak{j})_{(X-X_0,Y-Y_0)} = (\tilde{\mathsf{P}}, \tilde{\mathsf{Q}}_{x,X_0}, \tilde{\mathsf{Q}}_{y,Y_0})_{(X-X_0,Y-Y_0)} \ , \tag{4.87}$$

in the localization $\mathcal{S}_{(X-X_0,Y-Y_0)}$. Here, we use the notation $\mathfrak{I}_\mathfrak{m}$ to denote the ideal in $\mathcal{S}_\mathfrak{m}$ generated by the image of the ideal $\mathfrak{I} \subseteq \mathcal{S}$ under the *localization map* $\mathcal{S} \to \mathcal{S}_\mathfrak{m}$, where $\mathfrak{m}$ is a maximal ideal of $\mathcal{S}$.

[29] A unit is an element of the ring that has a multiplicative inverse. In any ring, the multiplicative identity 1 is its own inverse, so it is always a unit. In particular, in $\mathbb{Z}$, ± 1 are the only units, whereas in $\mathbb{Q}$, any nonzero rational is a unit.

[30] More abstractly, if r is one of the generators of an ideal $\mathfrak{I}$ of $\mathcal{R}$, then we can replace it by ur, where u is a unit.

Defining $\mathfrak{i}'_{X_0,Y_0} = (\tilde{\mathsf{P}}, \tilde{\mathsf{Q}}_{x,X_0}, \tilde{\mathsf{Q}}_{y,Y_0})$, we then have

$$\begin{aligned}[\mathcal{S}/(\mathfrak{i}+\mathfrak{j})]_{(X-X_0,Y-Y_0)} &\cong \mathcal{S}_{(X-X_0,Y-Y_0)}/(\mathfrak{i}+\mathfrak{j})_{(X-X_0,Y-Y_0)} \\ &\cong \mathcal{S}_{(X-X_0,Y-Y_0)}/(\mathfrak{i}'_{X_0,Y_0})_{(X-X_0,Y-Y_0)} \\ &\cong (\mathcal{S}/\mathfrak{i}'_{X_0,Y_0})_{(X-X_0,Y-Y_0)} \ ,\end{aligned} \tag{4.88}$$

where in the first and third lines, we used the slogan "localization commutes with quotienting", and the second line follows from (4.87). Now, the quotient $\mathcal{S}/\mathfrak{i}'_{X_0,Y_0}$ is also Artinian, and by Hilbert's Nullstellensatz, its only maximal ideal is $(X-X_0, Y-Y_0)$. Hence, by a result similar to the one in (4.80), we have

$$(\mathcal{S}/\mathfrak{i}'_{X_0,Y_0})_{(X-X_0,Y-Y_0)} \cong \mathcal{S}/\mathfrak{i}'_{X_0,Y_0} \ , \tag{4.89}$$

and therefore,

$$\log_p \mathrm{GSD} = 2 \sum_{(X_0,Y_0)\in V(\mathfrak{i}+\mathfrak{j})} \dim_{\mathbb{F}_{p^\infty}} \mathcal{S}/\mathfrak{i}'_{X_0,Y_0} \ . \tag{4.90}$$

Computation of $\log_p \mathrm{GSD}$ for special values of L_x, L_y

Given the expression (4.79) for the GSD, all we need to do now is to compute a Gröbner basis of $\mathfrak{i}'_{X_0,Y_0}$ in $\mathcal{S}$ for each $(X_0,Y_0) \in V(\mathfrak{i}+\mathfrak{j})$. First, we note certain "symmetries" in the set $V(\mathfrak{i}+\mathfrak{j})$. Recall that $V(\mathfrak{i}+\mathfrak{j})$ is the set of $(X_0,Y_0) \in \mathbb{F}^2_{p^\infty}$ that solve the system of polynomial equations

$$X^{L'_x} - 1 = Y^{L'_y} - 1 = \tilde{\mathsf{P}}(X,Y) = 0 \ , \tag{4.91}$$

where we used the facts that $X^{L_x} - 1 = (X^{L'_x} - 1)^{p^{k_x}}$ and $Y^{L_y} - 1 = (Y^{L'_y} - 1)^{p^{k_y}}$ in $\mathbb{F}^2_{p^\infty}[X,Y]$. (Here, we used the parametrization (4.77).) Given a solution (X_0,Y_0) of

(4.91), we can generate three more solutions using the transformations

$$X_0 \to X_0^{-1}\ , \qquad Y_0 \to Y_0^{-1}\ , \tag{4.92}$$

because the equations (4.91) are invariant under these transformations. (These transformations are well defined because $X_0 \neq 0$ and $Y_0 \neq 0$.) Furthermore, if $L'_x = L'_y$, we can generate four more solutions using the exchange

$$X_0 \leftrightarrow Y_0\ . \tag{4.93}$$

It is clear that $(1,1)$ is the only solution that is invariant under all these transformations. We will exploit these transformations in our analysis below.

In general, it is hard to compute a Gröbner basis of $\mathfrak{i}'_{X_0,Y_0}$ for arbitrary $(X_0, Y_0) \in \mathbb{F}^2_{p^\infty}$ except when $(X_0, Y_0) = (1,1)$. So, below, we specialize to those values of L_x, L_y for which $(X_0, Y_0) = (1,1)$ is the only solution of (4.91). These special values of L_x, L_y contain infinite families of L_x, L_y with interesting behaviors of GSD.

1. Consider the special case $L'_x = L'_y = 1$, where (4.91) becomes

$$X - 1 = Y - 1 = \tilde{\mathsf{P}}(X,Y) = 0\ . \tag{4.94}$$

   Clearly, $(X_0, Y_0) = (1,1)$ is the only solution of (4.94). In this case, changing the variables X and Y to $\tilde{X} = X - 1$ and $\tilde{Y} = Y - 1$, we have $\mathfrak{i}'_{1,1} = (\tilde{X}^2\tilde{Y} + \tilde{X}\tilde{Y}^2 + \tilde{X}^2 + \tilde{Y}^2, \tilde{X}^{p^{k_x}}, \tilde{Y}^{p^{k_y}})$. We can assume that $k_x \geq k_y$ without loss of generality. Then, with a lexicographic monomial order on $\tilde{X}, \tilde{Y}$ with $\tilde{X} \succ \tilde{Y}$, a Gröbner basis of $\mathfrak{i}'_{1,1}$ is given by the following:

- When $k_y > 0$,

$$\begin{aligned}
\mathsf{G}_1(\tilde{X},\tilde{Y}) &= \tilde{Y}^{p^{k_y}} \ , \\
\mathsf{G}_2(\tilde{X},\tilde{Y}) &= \tilde{X}\tilde{Y}^{p^{k_y}-1}\delta_{k_x,k_y} \ , \\
\mathsf{G}_3(\tilde{X},\tilde{Y}) &= \tilde{X}^2 + (\tilde{X}+1)\tilde{Y}^2 \sum_{i=0}^{p^{k_y}-3} (-\tilde{Y})^i - \tilde{X}\tilde{Y}^{p^{k_y}-1}\delta_{k_x,k_y} \ .
\end{aligned} \tag{4.95}$$

- When $k_y = 0$,

$$\begin{aligned}
\mathsf{G}_1(\tilde{X},\tilde{Y}) &= \tilde{Y} \ , \\
\mathsf{G}_2(\tilde{X},\tilde{Y}) &= \tilde{X}\delta_{k_x,0} + \tilde{X}^2(1-\delta_{k_x,0}) \ .
\end{aligned} \tag{4.96}$$

The set of monomials that are irreducible with respect to this Gröbner basis are

$$\{\tilde{Y}^i : 0 \le i < p^{k_y}\} \cup \{\tilde{X}\tilde{Y}^j : 0 \le j < p^{k_y} - \delta_{k_x,k_y}\} \ . \tag{4.97}$$

They form a basis of $\mathcal{S}/\mathfrak{i}'_{1,1}$, so we have

$$\dim_{\mathbb{F}_{p^\infty}} \mathcal{S}/\mathfrak{i}'_{1,1} = 2p^{k_y} - \delta_{k_x,k_y} \ . \tag{4.98}$$

By (4.79), we conclude that

$$\log_p \mathrm{GSD} = 2\left(2p^{k_y} - \delta_{k_x,k_y}\right) \ . \tag{4.99}$$

2. Next, we generalize the previous special case to $L'_x = L'_y = q$ with $q > 2$ a prime such that p is a *primitive root* modulo q.[31] Then, (4.91) becomes

$$X^q - 1 = Y^q - 1 = \tilde{\mathsf{P}}(X,Y) = 0 \ . \tag{4.100}$$

[31] In other words, p is a generator of the multiplicative group of integers modulo q, denoted as $\mathbb{Z}_q^\times$. For any positive integer n, the order of the group $\mathbb{Z}_n^\times$ is known as the *Euler totient function* of n, denoted as $\varphi(n)$. It is easy to see that $\varphi(q^m) = q^m - q^{m-1}$ for any odd prime q and any $m \ge 1$.

We will show that $X_0 = Y_0 = 1$ is the only solution of (4.100). We argue by contradiction. We assume that there is a solution with $X_0 \neq 1$. Then, any other solution (X_0, Y_0) of (4.100) is obtained from a solution of the form (X_0, X_0^s) for some $1 \leq s \leq (q-1)/2$ using the transformations (4.92) and (4.93).[32] Since $X_0^q = 1$ and $X_0 \neq 1$, X_0 is a root of the *cyclotomic polynomial* $\Phi_q(X) = \sum_{j=0}^{q-1} X^j$. Moreover, since (X_0, X_0^s) satisfies $\tilde{\mathsf{P}}(X,Y) = 0$, X_0 is also a root of the polynomial $\tilde{\mathsf{P}}(X, X^s)$. We can write

$$\tilde{\mathsf{P}}(X, X^s) = X(X^s - 1)^2 + X^s(X-1)^2 = X(X-1)^2\hat{\mathsf{P}}_s(X) , \tag{4.101}$$

where

$$\hat{\mathsf{P}}_s(X) = \left(\sum_{i=0}^{s-1} X^i\right)^2 + X^{s-1} . \tag{4.102}$$

$\hat{\mathsf{P}}_s(X)$ is a nonzero polynomial in $\mathbb{Z}_p[X]$ because $\hat{\mathsf{P}}_s(0) = 1 \mod p$ for $s > 1$, and $2 \mod p$ for $s = 1$.[33] Since X_0 is a root of $\tilde{\mathsf{P}}(X, X^s)$ and $X_0 \neq 0, 1$, it must be a root of $\hat{\mathsf{P}}_s(X)$. We now use the fact that $\Phi_q(X)$ is the *minimal polynomial* of X_0 in $\mathbb{Z}_p[X]$ because p is a primitive root modulo q [168, Section 11.2.B]. This means $\Phi_q(X)$ divides $\hat{\mathsf{P}}_s(X)$. But this is impossible because

$$\deg_X \hat{\mathsf{P}}_s(X) = 2s - 2 \leq q - 3 < q - 1 = \deg_X \Phi_q(X) , \tag{4.103}$$

So, there is no such X_0. In other words, when p is a primitive root modulo q, the only solution of (4.100) is $(1, 1)$. Then, by the analysis in point 1 above, $\log_p \mathrm{GSD}$ is again given by (4.99).

[32] Since $X_0 \neq 1$, $X_0^q = 1$, and q is prime, X_0 is a primitive qth root of unity, i.e., powers of X_0 generates all the qth roots of unity. Since $Y_0^q = 1$, it is a qth root of unity, and hence, $Y_0 = X_0^s$ for some $0 \leq s < q$. But $s \neq 0$ because there is no solution of (4.100) of the form $(X_0, 1)$ for $X_0 \neq 1$. Moreover, using the transformation (4.92), (X_0, X_0^{q-s}) is also a solution, so we can restrict s to the range $1 \leq s \leq (q-1)/2$.

[33] This is not true for $p = 2$ because $\hat{\mathsf{P}}_s(X) = 0$ identically for $s = 1$, so there is always a solution of (4.91) of the form (X_0, X_0) even for $X_0 \neq 1$.

3. We now generalize as follows. Let $q > 2$ be a prime such that p is a primitive root modulo q^m for some $m \geq 2$. Set $L'_x = L'_y = q^m$, so that (4.91) becomes

$$X^{q^m} - 1 = Y^{q^m} - 1 = \tilde{\mathsf{P}}(X, Y) = 0 \ . \tag{4.104}$$

Again, we argue by contradiction that the only solution of (4.104) $(X_0, Y_0) \in \mathbb{F}^2_{p^\infty}$ is $(1, 1)$. We assume that $X_0 \neq 1$. Then, any other solution (X_0, Y_0) of (4.104) is obtained from a solution of the form (X_0, X_0^s) for some $1 \leq s \leq (q^m - 1)/2$ using the transformations (4.92) and (4.93). Actually, the range of s can be smaller than this. Let q^r be the order of X_0 for some $0 \leq r \leq m$, i.e., $X_0^{q^r} = 1$, but $X_0^{q^{r'}} \neq 1$ for any $r' < r$. Since $r = 0$ corresponds to the trivial solution $(1, 1)$, we have $r > 0$. Then, $1 \leq s \leq (q^r - 1)/2$. We consider two cases:

- $\underline{s \leq \varphi(q^r)/2}$: [See footnote 31 for the definition of $\varphi(q^r)$.] Since the order of X_0 is q^r, it is a root of the cyclotomic polynomial $\Phi_{q^r}(X) = \sum_{j=0}^{q-1} X^{jq^{r-1}}$. Since $X_0 \neq 0, 1$, it is also a root of the polynomial $\hat{\mathsf{P}}_s(X)$ in (4.102). We now use the fact that $\Phi_{q^r}(X)$ is the minimal polynomial of X_0 in $\mathbb{Z}_p[X]$ because p is a primitive root modulo q^r [168, Section 11.2.B].[34] This means $\Phi_{q^r}(X)$ must divide $\hat{\mathsf{P}}_s(X)$. But this is impossible because

$$\deg_X \hat{\mathsf{P}}_s(X) = 2s - 2 \leq \varphi(q^r) - 2 < \varphi(q^r) = \deg_X \Phi_{q^r}(X) \ . \tag{4.105}$$

  So, there is no such X_0.

- $\underline{\varphi(q^r)/2 < s \leq (q^r - 1)/2}$: Let $X_0 = Z_0^2$, where Z_0 also has order q^r. (Such a Z_0 exists because $\gcd(2, q) = 1$ for $q > 2$.) Then, the solution (X_0, Y_0) is of the form (Z_0^2, Z_0^{2s}). By the transformation (4.92), (Z_0^2, Z_0^t) is also a

[34]Here, we used the fact: p is a primitive root modulo $q^m \implies p$ is a primitive root modulo q^r for all $r \leq m$. In fact, for $m \geq 2$: p is a primitive root modulo $q^m \implies p$ is a primitive root modulo q^{m+1}. Combining these facts: p is a primitive root modulo $q^2 \implies p$ is a primitive root modulo q^m for all $m \geq 1$ [163, Section 2.8].

solution, where $t = q^r - 2s$. Then

$$\frac{\varphi(q^r)}{2} < s \leq \frac{q^r - 1}{2} \implies 1 \leq t < q^{r-1} \leq \frac{q^r - q^{r-1}}{2} = \frac{\varphi(q^r)}{2} \; , \tag{4.106}$$

where the rightmost inequality holds for $q > 2$. Clearly, Z_0 is a root of $\Phi_{q^r}(X)$. Since (Z_0^2, Z_0^t) satisfies $\tilde{\mathsf{P}}(X,Y) = 0$, Z_0 is a root of $\tilde{\mathsf{P}}(X^2, X^t)$. We can write

$$\begin{aligned} \tilde{\mathsf{P}}(X^2, X^t) &= X^2(X^t - 1)^2 + X^t(X^2 - 1)^2 \\ &= \begin{cases} X(X-1)^2 \check{\mathsf{P}}_1(X) \; , & t = 1 \; , \\ X^2(X-1)^2 \check{\mathsf{P}}_t(X) \; , & t > 1 \; , \end{cases} \end{aligned} \tag{4.107}$$

where

$$\check{\mathsf{P}}_t(X) = \begin{cases} X + (X+1)^2 \; , & t = 1 \; , \\ \left(\sum_{i=0}^{t-1} X^i\right)^2 + X^{t-2}(X+1)^2 \; , & t > 1 \; . \end{cases} \tag{4.108}$$

$\check{\mathsf{P}}_t(X)$ is a nonzero polynomial because $\check{\mathsf{P}}_t(0) = 1 \mod p$ for $t \neq 2$, and $2 \mod p$ for $t = 2$.[35] Since Z_0 is a root of $\tilde{\mathsf{P}}(X^2, X^t)$ and $Z_0 \neq 0, 1$, it is a root of $\check{\mathsf{P}}_t(X)$ as well. We now use the fact that $\Phi_{q^r}(X)$ is the minimal polynomial of Z_0 in $\mathbb{Z}_p[X]$ because p is a primitive root modulo q^r [168, Section 11.2.B]. This means $\Phi_{q^r}(X)$ must divide $\check{\mathsf{P}}_t(X)$. But this is impossible because

$$\deg_X \check{\mathsf{P}}_t(X) = t + \max(t, 2) - 2 + \delta_{t,1} \leq 2t < \varphi(q^r) = \deg_X \Phi_{q^r}(X) \; , \tag{4.109}$$

where in the third line, we used (4.106). So, there is no such Z_0.

[35] Once again, this is not true for $p = 2$ because $\check{\mathsf{P}}_t(X) = 0$ identically when $t = 2$, so there is always a solution of (4.91) of the form (Z_0^2, Z_0^2) even for $Z_0 \neq 1$.

Therefore, when p is a primitive root modulo q^m, then $(1,1)$ is the only solution of (4.104), and hence, $\log_p \mathrm{GSD}$ is still given by (4.99).

To conclude, when $L_x = p^{k_x} q^m$ and $L_y = p^{k_y} q^m$, where q is an odd prime such that p is a primitive root modulo q^m, and $k_x, k_y, m \geq 0$, the ground state degeneracy of the 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model is given by

$$\log_p \mathrm{GSD} = 2\left[2p^{\min(k_x,k_y)} - \delta_{k_x,k_y}\right] \ . \tag{4.110}$$

When $m=0$, we see that $\log_p \mathrm{GSD}$ scales as $4\min(L_x, L_y)$. This gives an infinite family of L_x, L_y for which $\log_p \mathrm{GSD}$ is $O(L_x, L_y)$.

Say q is such that p is a primitive root modulo q^2. Then, p is a primitive root modulo q^m for all $m \geq 1$ (see footnote 34). Then, for $k_x = k_y = 0$ and any $m \geq 0$, we see that $\log_p \mathrm{GSD} = 2$, a finite number. This gives an infinite family of L_x, L_y for which $\log_p \mathrm{GSD}$ remains finite.[36]

Note that the last conclusion relies on the existence of an odd prime q such that p is a primitive root modulo q^2. However, we do not know of a proof for general p. Another interesting possibility is the following. By *Artin's conjecture on primitive roots* [164],[37] there are infinitely many prime q such that p is a primitive root modulo q. (Recall from footnote 34 that this does not imply that p is a primitive root modulo q^2.) Then, choosing $L_x = L_y = q$ for all such q gives another infinite family of L_x, L_y for which $\log_p \mathrm{GSD} = 2$. However, Artin's conjecture is still unproven, except under the assumption of the *generalized Riemann hypothesis* [165], which is also unproven.

[36] In contrast, when $N=2$, $\log_2 \mathrm{GSD}$ in (4.63) always scales as $4L$ for any L. This is because, when $p=2$, the above arguments do not go through, as explained in footnotes 33 and 35.

[37] The conjecture is actually stronger: the set of such q has positive asymptotic density inside the set of all primes.

4.C.2 Mobility restrictions

We now discuss the mobility of z-lineons in the xy-plane in the 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model. The lineons are represented as defects in the low-energy theory and their motion is implemented by operators acting at fixed time.

These operators fall into two kinds. First, there are operators supported in a small region, e.g., the line joining the two points. Second, there are also situations where the operator spans over $O(L_x, L_y)$ sites. Operators of the second kind exist only for certain special values of L_x, L_y depending on some number-theoretic properties of L_x, L_y, whereas the first kind exist for all L_x, L_y. In particular, only the first kind exist on an infinite square lattice. (See the discussion in [25, 127].)

As an example of the second kind of operator, consider $L_x = L_y = q^m$, where $q > 2$ is a prime such that p is a primitive root modulo q^m. In Appendix 4.C.1, we showed that for $L_x = L_y = q^m$, where $q > 2$ is a prime such that p is a primitive root modulo q^m, the ground state degeneracy is given by $\log_p \mathrm{GSD} = 2$. It follows that $|\operatorname{Jac}(C_{q^m} \times C_{q^m}, p)| = p$, or equivalently, $\operatorname{Jac}(C_{q^m} \times C_{q^m}, p) = \mathbb{Z}_p$. Therefore, in this case, the only selection imposed by the $\operatorname{Jac}(C_{q^m} \times C_{q^m}, p)$ time-like symmetry is that the total charge of the defects is conserved modulo p. In particular, a z-lineon can move anywhere within the xy-plane when $L_x = L_y = q^m$. However, we show below that a z-lineon is completely immobile when Γ is an infinite square lattice. This means, the operator that moves a z-lineon on the 2d torus graph $C_{q^m} \times C_{q^m}$ must be of the second kind.

We now show that a z-lineon is completely immobile on an infinite square lattice. In fact, we show that any finite configuration of z-lineons is completely immobile (except in some trivial cases) as long as their charges and the separations between them are fixed during the motion, i.e., we allow only "rigid" motion. Without this restriction, the groups of lineons can move. We will not discuss this motion.

Our analysis will be similar to the analogous discussion in [127]. The main difference between them is that here our variables are in $\mathbb{Z}_p$ and therefore various properties of the polynomials will depend on p.

Consider n z-lineons, with charges q_i and positions (x_i, y_i) for $i = 1, \ldots, n$, described by the defect

$$\exp\left[\frac{2\pi i}{p}\sum_{\tau}\sum_{i=1}^{n} q_i m_\tau(\tau + \tfrac{1}{2}, x_i, y_i)\right] \tag{4.111}$$

(Since they are z-lineons, we can assume without loss of generality that they all have the same z coordinate, and omit writing it.) They can move "rigidly" by $(x_0, y_0) \neq (0,0)$ if there is a defect of the form

$$\begin{aligned} &\exp\left[\frac{2\pi i}{p}\sum_{\tau<0}\sum_{i=1}^{n} q_i m_\tau(\tau + \tfrac{1}{2}, x_i, y_i)\right] \times \exp\left[\frac{2\pi i}{p}\sum_{j=1}^{l} s_j m(0, x_j, y_j)\right] \\ &\quad \times \exp\left[\frac{2\pi i}{p}\sum_{\tau\geq 0}\sum_{i=1}^{n} q_i m_\tau(\tau + \tfrac{1}{2}, x_i + x_0, y_i + y_0)\right] . \end{aligned} \tag{4.112}$$

This defect is gauge invariant if and only if

$$\sum_{i=1}^{n} q_i \left[k(0, x_i + x_0, y_i + y_0) - k(0, x_i, y_i)\right] = \sum_{j=1}^{l} s_j (\Delta_x^2 + \Delta_y^2) k(0, x_j, y_j) \mod p \ , \tag{4.113}$$

for any integer gauge parameter k in (4.19).

Using a *formal Laurent power series*

$$\hat{k}(X, Y) = \sum_{(x,y)\in\mathbb{Z}^2} k(0, x, y) X^{-x} Y^{-y} \ , \tag{4.114}$$

associated with the gauge parameter $k(0, x, y)$, the condition (4.113) can be written as

$$(X^{x_0} Y^{y_0} - 1)\mathsf{Q}(X, Y) = \mathsf{S}(X, Y)\mathsf{P}(X, Y) \mod p \ , \tag{4.115}$$

where

$$\mathsf{P}(X,Y) = (X - 2 + X^{-1}) + (Y - 2 + Y^{-1}) \ , \tag{4.116}$$

is the *Laurent polynomial* (i.e., an element of $\mathbb{Z}_p[X, X^{-1}, Y, Y^{-1}]$) associated with the discrete Laplacian operator $\Delta_x^2 + \Delta_y^2$, and

$$\mathsf{Q}(X,Y) = \sum_{i=1}^{n} q_i X^{x_i} Y^{y_i} \ , \qquad \mathsf{S}(X,Y) = \sum_{j=1}^{l} s_j X^{x_j} Y^{y_j} \ , \tag{4.117}$$

are also Laurent polynomials. The coefficients and monomials in $\mathsf{Q}(X,Y)$ and $\mathsf{S}(X,Y)$ are obtained from the defect (4.112).

If there is a Laurent polynomial $\mathsf{R}(X,Y)$ such that $\mathsf{Q}(X,Y) = \mathsf{R}(X,Y)\mathsf{P}(X,Y)$, then (4.115) can be trivially satisfied by choosing $\mathsf{S}(X,Y) = (X^{x_0}Y^{y_0} - 1)\mathsf{R}(X,Y)$. However, in this case, the defect (4.111) can end at $\tau = 0$ as follows:

$$\exp\left[\frac{2\pi i}{p}\sum_{\tau<0}\sum_{i=1}^{n} q_i m_\tau(\tau + \tfrac{1}{2}, x_i, y_i)\right] \times \exp\left[-\frac{2\pi i}{p}\sum_{j'=1}^{l'} r_{j'} m(0, x_{j'}, y_{j'})\right] \tag{4.118}$$

where $r_{j'}$'s and $(x_{j'}, y_{j'})$'s are obtained from $\mathsf{R}(X,Y) = \sum_{j'=1}^{l'} r_{j'} X^{x_{j'}} Y^{y_{j'}}$. Therefore, in this case, the defect (4.112) describes the annihilation of the n z-lineons at their original positions and their creation at positions displaced by (x_0, y_0) at time $\tau = 0$.

A more interesting situation occurs for a defect like (4.112) when $\mathsf{Q}(X,Y)$ cannot be written as $\mathsf{R}(X,Y)\mathsf{P}(X,Y)$ for any Laurent polynomial $\mathsf{R}(X,Y)$. Then, to satisfy (4.115), $\mathsf{P}(X,Y)$ and $X^{x_0}Y^{y_0} - 1$ must share a nontrivial factor.[38] Let us show that this cannot happen. In the following, it is crucial that $(x_0, y_0) \neq (0,0)$.

[38]By *nontrivial factor* we mean a nonconstant Laurent factor that is not a Laurent monomial.

First, note that $\mathsf{P}(X,Y)$ is nonconstant and irreducible up to a monomial in $\mathbb{Z}_p[X,X^{-1},Y,Y^{-1}]$ for any odd prime p.[39] So, all we need to show is that $X^{x_0}Y^{y_0}-1$ is not a multiple of $\mathsf{P}(X,Y)$ in $\mathbb{Z}_p[X,X^{-1},Y,Y^{-1}]$.

Let p^k be the largest power of p that divides both x_0 and y_0, i.e., $x_0' = x_0/p^k$ and $y_0' = y_0/p^k$ are integers and $d = \gcd(x_0', y_0')$ is not divisible by p. Then, in $\mathbb{Z}_p[X,X^{-1},Y,Y^{-1}]$, we have

$$X^{x_0}Y^{y_0}-1 = (X^{x_0'}Y^{y_0'}-1)^{p^k} = \Big[(X^{x_0''}Y^{y_0''}-1)\mathsf{T}(X,Y)\Big]^{p^k} \ , \tag{4.119}$$

where $x_0'' = x_0'/d$, $y_0'' = y_0'/d$, and $\mathsf{T}(X,Y) = \sum_{c=0}^{d-1}(X^{x_0''}Y^{y_0''})^c$. Now, $\mathsf{T}(1,1) = d \neq 0 \mod p$, whereas $\mathsf{P}(1,1)=0$, so $\mathsf{T}(X,Y)^{p^k}$ is not a multiple of $\mathsf{P}(X,Y)$. Since $\gcd(x_0'', y_0'') = 1$, the factor $X^{x_0''}Y^{y_0''}-1$ is nonconstant and irreducible up to a monomial for any p [159]. So $(X^{x_0''}Y^{y_0''}-1)^{p^k}$ is also not a multiple of $\mathsf{P}(X,Y)$. Therefore, $X^{x_0}Y^{y_0}-1$ is not a multiple of $\mathsf{P}(X,Y)$.

To conclude, a finite set of z-lineons cannot move "rigidly" in the xy-plane in the 3+1d anisotropic $\mathbb{Z}_p$ Laplacian model, unless they can be annihilated.

4.D $\mathbb{Z}_N$ Laplacian model on a graph

In this appendix, we analyze a gapped fracton model on a simple, connected, undirected spatial graph Γ. We refer to it as the $\mathbb{Z}_N$ Laplacian model because the theory is defined using the discrete Laplacian operator Δ_L on the graph Γ. The anisotropic

[39] A polynomial in $\mathbb{Z}_p[X,Y]$ is said to be *irreducible* if it cannot be written as a product of two nonconstant polynomials. A Laurent polynomial $\mathsf{F}(X,Y)$ in $\mathbb{Z}_p[X,X^{-1},Y,Y^{-1}]$ is said to be *irreducible up to a monomial* if $X^aY^b\mathsf{F}(X,Y)$ is an irreducible polynomial for some $a,b\in\mathbb{Z}$. For example, $\mathsf{P}(X,Y)$ is irreducible up to a monomial because $\tilde{\mathsf{P}}(X,Y) = XY\mathsf{P}(X,Y)$, given by (4.70), is an irreducible polynomial. The irreducibility of $\tilde{\mathsf{P}}(X,Y)$ for any prime $p>6$ follows from [169, Corollary 3]. It is easy to verify by hand, or in Mathematica, that it is irreducible even for $p=3,5$. It is, however, not irreducible for $p=2$ because $\tilde{\mathsf{P}}(X,Y) = (XY+1)(X+Y) \mod 2$. This is one way of seeing why a dipole of z-lineons separated in $(1,\pm1)$ direction can move in $(1,\mp1)$ direction in the 3+1d anisotropic $\mathbb{Z}_2$ Laplacian model.

$\mathbb{Z}_N$ Laplacian model in Section 4.3 is an anisotropic extension of this $\mathbb{Z}_N$ Laplacian model by adding another direction.

4.D.1 Hamiltonian

In the Hamiltonian formulation of the $\mathbb{Z}_N$ Laplacian model, there are a $\mathbb{Z}_N$ variable U_i and its conjugate variable V_i, i.e. $U_iV_i = e^{2\pi i/N}V_iU_i$, on every site of the graph Γ where i labels the sites. The Hamiltonian is

$$H = -\gamma_1 \sum_i G_i + \text{h.c.} \ , \tag{4.120}$$

where

$$G_i = \prod_{j:\langle i,j\rangle\in\Gamma} V_i V_j^\dagger \ . \tag{4.121}$$

Here, $\langle i, j\rangle$ means i and j are connected by an edge in the graph Γ.

Since all the G_is commute, the ground states satisfy $G_i = 1$ for all i and the excitations are violations of $G_i = 1$. We could take the limit $\gamma_1 \to \infty$, in which case, the Hilbert space consists of only the ground states and the Hamiltonian is trivial. The Euclidean presentation of this model in this limit will be discussed later in Appendix 4.D.2.

We are particularly interested in those operators that commute with the Hamiltonian (4.120) and act nontrivially on its ground states. They are the global symmetry operators of the model in the low energy limit.

The electric symmetry operators are V_i, which trivially commute with the Hamiltonian. Since the ground states satisfy $G_i = 1$, some of these operators are equivalent when acting on the ground states. The independent symmetry operators are

$$\tilde{W}_\lambda = \prod_i V_i^{\lambda(i)} \ , \tag{4.122}$$

where $\lambda(i)$ takes the form (4.8)

$$\lambda(i) = \sum_{a=1}^{\mathsf{N}} p_a (Q^{-1})_{ai} \ , \qquad p_a \sim p_a + \gcd(N, r_a) \ . \tag{4.123}$$

Let us explain the identification on p_a. We have $p_a \sim p_a + N$ because $V_i^N = 1$. We also have $p_a \sim p_a + r_a$ because $r_a (Q^{-1})_{ai} = \sum_j L_{ij} P_{aj}$ and $\prod_i V_i^{L_{ij}} = 1$ when acting on the ground states. Combining the two identifications, we get $p_a \sim p_a + \gcd(N, r_a)$. The symmetry operators generate a $\mathrm{Jac}(\Gamma, N)$ electric symmetry.

The magnetic symmetry operators are

$$W_{\tilde{\lambda}} = \prod_i U_i^{\tilde{\lambda}(i)} \ , \tag{4.124}$$

where $\tilde{\lambda}(i)$ obeys $\Delta_L \tilde{\lambda}(i) = 0 \mod N$, and the most general solution takes the form (4.7)

$$\tilde{\lambda}(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} \tilde{p}_a}{\gcd(N, r_a)} \ , \qquad \tilde{p}_a \sim \tilde{p}_a + \gcd(N, r_a) \ . \tag{4.125}$$

The symmetry operators generate a $\mathrm{Jac}(\Gamma, N)$ magnetic symmetry.

A convenient basis of electric and magnetic space-like symmetry operators is given by

$$\begin{aligned} \tilde{W}(a) &= \prod_i V_i^{(Q^{-1})_{ai}} \ , \\ W(a) &= \prod_i U_i^{\frac{N}{\gcd(N, r_a)} Q_{ia}} \ . \end{aligned} \tag{4.126}$$

for $a = 1, \ldots, \mathsf{N}$. Both $W(a)$ and $\tilde{W}(a)$ are $\mathbb{Z}_{\gcd(N, r_a)}$ operators. They satisfy the commutation relations

$$W(a) \tilde{W}(b) = \exp\left[\frac{2\pi i \delta_{ab}}{\gcd(N, r_a)}\right] \tilde{W}(b) W(a) \ , \qquad a, b = 1, \ldots, \mathsf{N} \ . \tag{4.127}$$

For each a, there is only one b that has nontrivial commutation relation. So for each a, there is an independent $\mathbb{Z}_{\gcd(N,r_a)}$ Heisenberg algebra generated by $W(a)$ and $\tilde{W}(a)$, leading to a ground state degeneracy

$$\mathrm{GSD} = \prod_{a=1}^{\mathsf{N}} \gcd(N, r_a) = |\operatorname{Jac}(\Gamma, N)| \ . \tag{4.128}$$

4.D.2 Euclidean presentation

We now discuss the Euclidean presentation of the $\mathbb{Z}_N$ Laplacian model. We place the theory on a Euclidean spacetime lattice $C_{L_\tau} \times \Gamma$, where Γ is the spatial slice. We use (τ, i) to label a site in the spacetime lattice. The integer BF-action of the $\mathbb{Z}_N$ Laplacian theory is

$$S = \frac{2\pi i}{N} \sum_{\tau,i} \tilde{m}(\tau + \tfrac{1}{2}, i) \left[\Delta_\tau m(\tau, i) - \Delta_L m_\tau(\tau + \tfrac{1}{2}, i)\right] \ , \tag{4.129}$$

where the integer fields $\tilde{m}$ and (m_τ, m) have an integer gauge symmetry

$$\begin{aligned} &\tilde{m} \sim \tilde{m} + N\tilde{k} \ , \\ &m_\tau \sim m_\tau + \Delta_\tau k + N q_\tau \ , \\ &m \sim m + \Delta_L k + N q \ , \end{aligned} \tag{4.130}$$

where k, $\tilde{k}$, and (q_τ, q) are integers. (Note that, when working modulo N, the last line of (4.130) is precisely the equivalence relation discussed in (4.5).) The integer BF-action (4.129) describes the ground states of the Hamiltonian (4.120).

Ground state degeneracy

We can count the number of ground states by counting the number of solutions to the "equations of motion" of (m_τ, m):

$$\Delta_\tau \tilde{m} = 0 \mod N \ , \qquad \Delta_L \tilde{m} = 0 \mod N \ . \tag{4.131}$$

The first equation implies that $\tilde{m}(\tau, i)$ is independent of τ. Then, as discussed in Section 4.2.1, the general solution is

$$\tilde{m}(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_a}{\gcd(N, r_a)} \ , \qquad p_a \sim p_a + \gcd(N, r_a) \ . \tag{4.132}$$

Therefore, the ground state degeneracy is

$$\mathrm{GSD} = \prod_{a=1}^{\mathsf{N}} \gcd(N, r_a) = |\operatorname{Jac}(\Gamma, N)| \ . \tag{4.133}$$

Global symmetry

The above ground state degeneracy can also be obtained from the (space-like) global symmetry. There are electric (space-like and time-like) and magnetic (space-like) global symmetries, whose groups are both

$$\operatorname{Jac}(\Gamma, N) = \prod_{a=1}^{\mathsf{N}} \mathbb{Z}_{\gcd(N, r_a)} \ . \tag{4.134}$$

The electric global symmetry acts as

$$(m_\tau, m) \to (m_\tau, m) + (\lambda_\tau, \lambda) \ , \tag{4.135}$$

where (λ_τ, λ) is a flat $\mathbb{Z}_N$ gauge field, i.e., $\Delta_\tau \lambda - \Delta_L \lambda_\tau = 0 \mod N$. Using k, we can set $\lambda_\tau(\tau + \frac{1}{2}, i)|_{\tau \neq 0} = 0 \mod N$. Then, by flatness, we have $\Delta_\tau \lambda = 0 \mod N$. This in turn implies that $\Delta_L \lambda_\tau(\tau + \frac{1}{2}, i)|_{\tau=0} = 0 \mod N$, which is the discrete Laplace equation (4.4).

The remaining time-independent gauge freedom, $\lambda(i) \sim \lambda(i) + \Delta_L k(i)$, is precisely the equivalence relation in (4.5). Therefore, we can gauge fix $\lambda(i)$ to

$$\lambda(i) = \sum_{a=1}^{\mathsf{N}} p_a (Q^{-1})_{ai} \ , \tag{4.136}$$

where $p_a = 0, \ldots, \gcd(N, r_a) - 1$. Since $\lambda_\tau(\tau + \frac{1}{2}, i)|_{\tau=0}$ satisfies the discrete Laplace equation, the most general solution is (4.7)

$$\lambda_\tau(\tau + \tfrac{1}{2}, i)|_{\tau=0} = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} p_{\tau,a}}{\gcd(N, r_a)} \ , \tag{4.137}$$

where $p_{\tau,a} = 0, \ldots, \gcd(N, r_a) - 1$. The parameters $p_{\tau,a}$ and p_a generate the electric time-like and space-like global symmetries respectively.

The magnetic space-like global symmetry acts as

$$\tilde{m}(\tau + \tfrac{1}{2}, i) \to \tilde{m}(\tau + \tfrac{1}{2}, i) + \tilde{\lambda}(i) \ , \qquad \tilde{\lambda}(i) = \sum_{a=1}^{\mathsf{N}} \frac{N Q_{ia} \tilde{p}_a}{\gcd(N, r_a)} \ , \tag{4.138}$$

and $\tilde{p}_a = 0, \ldots, \gcd(N, r_a) - 1$.

A convenient basis of electric and magnetic space-like symmetry operators is given by

$$\begin{aligned} \tilde{W}(a) &= \exp\left[\frac{2\pi i}{N} \sum_i (Q^{-1})_{ai} \tilde{m}(\tau + \tfrac{1}{2}, i)\right] \ , \\ W(a) &= \exp\left[\frac{2\pi i}{\gcd(N, r_a)} \sum_i m(\tau, i) Q_{ia}\right] \ . \end{aligned} \tag{4.139}$$

for $a = 1, \ldots, \mathsf{N}$. These operators are the low energy counterpart of the operators in (4.126). The commutation relation (4.127) can now be understood as a mixed 't Hooft anomaly between the electric and magentic space-like symmetries.

Time-like symmetry and fractons

The $\mathbb{Z}_N$ Laplacian model has defects that extend in the time direction, such as

$$W_\tau(i) = \exp\left[\frac{2\pi i}{N}\sum_\tau m_\tau(\tau+\tfrac{1}{2},i)\right] . \tag{4.140}$$

This describes the world-line of an infinitely heavy particle of unit charge at position $i \in \Gamma$.

Below we discuss the time-like global symmetry that acts on these defects. The electric time-like symmetry acts as

$$m_\tau(\tau+\tfrac{1}{2},i) \to m_\tau(\tau+\tfrac{1}{2},i) + \delta_{\tau,0}\sum_{a=1}^{\mathsf{N}} \frac{NQ_{ia}p_{\tau,a}}{\gcd(N,r_a)} . \tag{4.141}$$

Therefore, two defects at sites i and i' carry the same time-like charges, or equivalently, a particle can hop from i to i', if and only if

$$Q_{ia} = Q_{i'a} \mod \gcd(N,r_a) , \qquad a = 1,\dots,\mathsf{N} . \tag{4.142}$$

Indeed, when this condition holds, the defect that "moves" a particle from i to i' at time $\tau = 0$ is given by

$$\begin{aligned}&\exp\left[\frac{2\pi i}{N}\sum_{\tau<0} m_\tau(\tau+\tfrac{1}{2},i)\right]\exp\left[-\frac{2\pi i}{N}\sum_{a,j}\left(\frac{Q_{ia}-Q_{i'a}}{\gcd(N,r_a)}\right)\tilde{r}_a P_{aj} m(0,j)\right]\\ &\qquad\times \exp\left[\frac{2\pi i}{N}\sum_{\tau\geq 0} m_\tau(\tau+\tfrac{1}{2},i')\right] ,\end{aligned} \tag{4.143}$$

where $\tilde{r}_a$ is an integer such that $\tilde{r}_a r_a = \gcd(N,r_a) \mod N$.

While the selection rule (4.142) is not very intuitive, it leads to strong mobility constraints in the special case where the spatial lattice is a square lattice (i.e., Γ is a

2d torus graph $C_{L_x} \times C_{L_y}$). This will be shown in Section 4.D.3. In particular, under some mild conditions, the particles are completely immobile, i.e., they are fractons.

Robustness

We now discuss the robustness of the low-energy limit of the $\mathbb{Z}_N$ Laplacian model. The only operators that act nontrivially on the ground states are $W(a)$ and $\tilde{W}(a)$ of (4.139). $W(a)$ is an extended operator with support spanning over the entire graph. In contrast, $\tilde{W}(a)$ can be written as a product of local operators of the form $e^{\frac{2\pi i}{N}\tilde{m}(\tau+\frac{1}{2},i)}$. (In (4.139) we defined $\tilde{W}(a)$ in such a way that its commutation relation (4.127) with the extended operator $W(a)$ is simple.) Since these local operators act nontrivially in the space of ground states, the low-energy limit of the model is not robust.

4.D.3 Examples

1+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory

Let Γ be a cycle graph, i.e., $\Gamma = C_{L_x}$, where L_x is the number of sites in the cycle. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator Δ_x^2 in the x-direction.

In this case, the $\mathbb{Z}_N$ Laplacian model simplifies to the 1+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory of [25]. Indeed, the diagonal entries in the Smith normal form of L are

$$r_a = \begin{cases} 1\ , & 1 \leq a < L_x - 1\ , \\ L_x\ , & a = L_x - 1\ , \\ 0\ , & a = L_x\ . \end{cases} \tag{4.144}$$

To be more concrete, we can write $R = PLQ$, where

$$R = \begin{pmatrix} I_{L_x-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & L_x & 0 \\ \mathbf{0}^T & 0 & 0 \end{pmatrix} , \qquad P = \begin{pmatrix} \tilde{P} & \mathbf{0} \\ \mathbf{1}^T & 1 \end{pmatrix} , \qquad Q = \begin{pmatrix} \tilde{Q} & \mathbf{1} \\ \mathbf{0}^T & 1 \end{pmatrix} , \tag{4.145}$$

where $\tilde{P}$ and $\tilde{Q}$ are $(L_x - 1) \times (L_x - 1)$ integer matrices given by

$$\tilde{P}_{a,x+1} = \min\{a, x+1\} , \qquad \tilde{Q}_{x+1,a} = \delta_{x+1,a} - (x+1)(1 - \delta_{x,L_x-2})\delta_{a,L_x-1} , \tag{4.146}$$

where $a = 1, \ldots, L_x$, and $x = 0, \ldots, L_x - 1$. For example, for $L_x = 5$, the 4×4 matrices $\tilde{P}$ and $\tilde{Q}$ are

$$\tilde{P} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} , \qquad \tilde{Q} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{4.147}$$

The electric (time-like and space-like) and magnetic (space-like) symmetries of the $\mathbb{Z}_N$ Laplacian gauge theory are $\mathbb{Z}_N \times \mathbb{Z}_{\gcd(N,L_x)}$. These are in agreement with the 1+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory of [25].

2+1d $\mathbb{Z}_N$ Laplacian gauge theory

Let Γ be a torus graph, i.e., $\Gamma = C_{L_x} \times C_{L_y}$, where L_x and L_y are the number of sites in the x-cycle and y-cycle. The operator Δ_L associated with the Laplacian matrix of Γ is the same as the standard Laplacian operator $\Delta_x^2 + \Delta_y^2$ on a square lattice.

In this case, the $\mathbb{Z}_N$ Laplacian model can be viewed as the $\mathbb{Z}_N$ version of the Laplacian ϕ-theory or the $\mathbb{Z}_N$ version of the $U(1)$ Laplacian gauge theory discussed in [2, 127]. Its ground state degeneracy is the square root of that of the anisotropic

$\mathbb{Z}_N$ Laplacian model, which is computed in Appendix 4.C.1. When N is a prime, the GSD of the $\mathbb{Z}_N$ Laplacian model is thus

$$\log_N \text{GSD} = \dim_{\mathbb{Z}_N} \frac{\mathbb{Z}_N[X,Y]}{(Y(X-1)^2 + X(Y-1)^2), X^{L_x}-1, Y^{L_y}-1)} \ . \tag{4.148}$$

The GSD depends on L_x, L_y in an erratic way. There exists a sequence of L_x, L_y where the $\log_N \text{GSD} \sim O(L_x, L_y)$, but there also exists a sequence where the GSD stays at order 1 if $N > 2$.

Like the GSD, the mobility of a particles depends on number-theoretic properties of L_x, L_y. Since the mobility of these particles is the same as the mobility of the z-lineons of the 3+1d anisotropic $\mathbb{Z}_N$ Laplacian model in the xy-plane, the analysis of Appendix 4.C.2 applies here. In particular, when N is an odd prime, there are infinitely many values of L_x, L_y for which a single particle is completely mobile. In contrast, on an infinite square lattice, any finite set of particles is completely immobile (unless they can be annihilated), assuming they move "rigidly."

When $N = 2$, the $\mathbb{Z}_N$ Laplacian model is equivalent to two copies of a known model, the $\mathbb{Z}_2$ Ising plaquette model [35], when both L_x and L_y are even, and only one copy when L_x or L_y is odd. Therefore, the GSD and mobility restrictions of the $\mathbb{Z}_2$ Laplacian model are relatively simple in this case, and follow from the analysis in Appendix 4.B.

Let us contrast the $\mathbb{Z}_N$ Laplacian model with the 2+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory discussed in [65,66,138–140], which is another 2+1d generalization of the 1+1d rank-2 $\mathbb{Z}_N$ tensor gauge theory. These two models differ in several aspects:

1. The $\mathbb{Z}_N$ tensor gauge theory has a ground state degeneracy of

$$N^3 \text{gcd}(N, L_x)\text{gcd}(N, L_y)\text{gcd}(N, L_x, L_y) \ . \tag{4.149}$$

In particular, the GSD of $\mathbb{Z}_N$ tensor gauge theory is always bounded by N^6, whereas there are infinitely many L_x, L_y for which $\log_N$ GSD of the $\mathbb{Z}_N$ Laplacian model scales as $O(L_x, L_y)$, at least when N is prime.

2. Relatedly, the low-energy limit of the $\mathbb{Z}_N$ tensor gauge theory is robust, whereas the low-energy limit of the $\mathbb{Z}_N$ Laplacian model is not.

3. A particle in the $\mathbb{Z}_N$ tensor gauge theory can always hop by N sites on an infinite square lattice,[40] whereas a particle in the $\mathbb{Z}_N$ Laplacian model is completely immobile on an infinite square lattice, at least when N is prime.

[40] More precisely, this is true for an electrically charged particle. Magnetically charged particles come in two flavors: x-lineons which can move anywhere in the x direction but can only hop by N sites in the y-direction, and similarly y-lineons.

Bibliography

[1] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *A modified Villain formulation of fractons and other exotic theories*, *J. Math. Phys.* **62**, 102301 (2021), [`arXiv:2103.01257`].

[2] P. Gorantla, H. T. Lam, and S.-H. Shao, *Fractons on graphs and complexity*, *Phys. Rev. B* **106**, 195139 (2022), [`arXiv:2207.08585`].

[3] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *Gapped lineon and fracton models on graphs*, *Phys. Rev. B* **107**, 125121 (2023), [`arXiv:2210.03727`].

[4] K. G. Wilson, *The renormalization group: Critical phenomena and the kondo problem*, *Rev. Mod. Phys.* **47**, 773–840 (Oct, 1975).

[5] J. Polchinski, *Renormalization and Effective Lagrangians*, *Nucl. Phys. B* **231**, 269–295 (1984).

[6] 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**, 1217–1241 (Aug, 1977).

[7] R. Savit, *Vortices and the low-temperature structure of the $x-y$ model*, *Phys. Rev. B* **17**, 1340–1350 (Feb, 1978).

[8] 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`.

[9] P. Di Francesco, P. Mathieu, and D. Senechal, *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer-Verlag, New York, 1997.

[10] V. L. Berezinskiĭ, *Destruction of Long-range Order in One-dimensional and Two-dimensional Systems having a Continuous Symmetry Group I. Classical Systems*, *Soviet Journal of Experimental and Theoretical Physics* **32**, 493 (Jan., 1971).

[11] V. L. Berezinskiĭ, *Destruction of Long-range Order in One-dimensional and Two-dimensional Systems Possessing a Continuous Symmetry Group. II.*

Quantum Systems, *Soviet Journal of Experimental and Theoretical Physics* **34**, 610 (Jan., 1972).

[12] J. M. Kosterlitz and D. J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, *J. Phys. C* **6**, 1181–1203 (1973).

[13] A. Yu. Kitaev, *Fault tolerant quantum computation by anyons*, *Annals Phys.* **303**, 2–30 (2003), [quant-ph/9707021].

[14] R. Dijkgraaf and E. Witten, *Topological Gauge Theories and Group Cohomology*, *Commun. Math. Phys.* **129**, 393 (1990).

[15] A. Kapustin and N. Seiberg, *Coupling a QFT to a TQFT and Duality*, *JHEP* **04**, 001 (2014), [arXiv:1401.0740].

[16] S. C. Zhang, T. H. Hansson, and S. Kivelson, *An effective field theory model for the fractional quantum hall effect*, *Phys. Rev. Lett.* **62**, 82–85 (1988).

[17] R. M. Nandkishore and M. Hermele, *Fractons*, *Ann. Rev. Condensed Matter Phys.* **10**, 295–313 (2019), [arXiv:1803.11196].

[18] M. Pretko, X. Chen, and Y. You, *Fracton Phases of Matter*, *Int. J. Mod. Phys. A* **35**, 2030003 (2020), [arXiv:2001.01722].

[19] C. Chamon, *Quantum glassiness in strongly correlated clean systems: An example of topological overprotection*, *Phys. Rev. Lett.* **94**, 040402 (2005), [cond-mat/0404182].

[20] J. Haah, *Local stabilizer codes in three dimensions without string logical operators*, *Phys. Rev. A* **83**, 042330 (2011), [arXiv:1101.1962].

[21] S. Vijay, J. Haah, and L. Fu, *Fracton Topological Order, Generalized Lattice Gauge Theory and Duality*, *Phys. Rev. B* **94**, 235157 (2016), [arXiv:1603.04442].

[22] A. Paramekanti, L. Balents, and M. P. A. Fisher, *Ring exchange, the exciton bose liquid, and bosonization in two dimensions*, *Phys. Rev. B* **66**, 054526 (Aug, 2002), [cond-mat/0203171].

[23] M. Pretko, *Subdimensional Particle Structure of Higher Rank U(1) Spin Liquids*, *Phys. Rev. B* **95**, 115139 (2017), [arXiv:1604.05329].

[24] J. Haah, *A degeneracy bound for homogeneous topological order*, *SciPost Phys.* **10**, 011 (2021), [arXiv:2009.13551].

[25] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *Global dipole symmetry, compact Lifshitz theory, tensor gauge theory, and fractons*, *Phys. Rev. B* **106**, 045112 (2022), [arXiv:2201.10589].

[26] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *Low-energy limit of some exotic lattice theories and UV/IR mixing*, *Phys. Rev. B* **104**, 235116 (2021), [arXiv:2108.00020].

[27] L. D. Landau, *On the theory of phase transitions. I.*, *Phys. Z. Sowjet.* **11**, 26 (1937).

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

[29] C. Cordova, T. T. Dumitrescu, K. Intriligator, and S.-H. Shao, *Snowmass White Paper: Generalized Symmetries in Quantum Field Theory and Beyond*, in *2022 Snowmass Summer Study*, 5, 2022. arXiv:2205.09545.

[30] L. Bhardwaj and Y. Tachikawa, *On finite symmetries and their gauging in two dimensions*, *JHEP* **03**, 189 (2018), [arXiv:1704.02330].

[31] C.-M. Chang, Y.-H. Lin, S.-H. Shao, Y. Wang, and X. Yin, *Topological Defect Lines and Renormalization Group Flows in Two Dimensions*, *JHEP* **01**, 026 (2019), [arXiv:1802.04445].

[32] K. T. Grosvenor, C. Hoyos, F. Peña Benitez, and P. Surówka, *Space-Dependent Symmetries and Fractons*, *Front. in Phys.* **9**, 792621 (2022), [arXiv:2112.00531].

[33] T. Brauner, S. A. Hartnoll, P. Kovtun, H. Liu, M. Mezei, A. Nicolis, R. Penco, S.-H. Shao, and D. T. Son, *Snowmass White Paper: Effective Field Theories for Condensed Matter Systems*, in *2022 Snowmass Summer Study*, 3, 2022. arXiv:2203.10110.

[34] J. McGreevy, *Generalized Symmetries in Condensed Matter*, arXiv:2204.03045.

[35] N. Seiberg and S.-H. Shao, *Exotic Symmetries, Duality, and Fractons in 2+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **10**, 027 (2021), [arXiv:2003.10466].

[36] N. Seiberg and S.-H. Shao, *Exotic $U(1)$ Symmetries, Duality, and Fractons in 3+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **9**, 046 (2020), [arXiv:2004.00015].

[37] N. Seiberg and S.-H. Shao, *Exotic $\mathbb{Z}_N$ Symmetries, Duality, and Fractons in 3+1-Dimensional Quantum Field Theory*, *SciPost Phys.* **10**, 003 (2021), [arXiv:2004.06115].

[38] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *More Exotic Field Theories in 3+1 Dimensions*, *SciPost Phys.* **9**, 073 (2020), [arXiv:2007.04904].

[39] E. Witten, *Quantum Field Theory and the Jones Polynomial*, *Commun. Math. Phys.* **121**, 351–399 (1989).

[40] K. Slagle and Y. B. Kim, *X-cube model on generic lattices: Fracton phases and geometric order*, *Phys. Rev. B* **97**, 165106 (2018), [arXiv:1712.04511].

[41] W. Shirley, K. Slagle, Z. Wang, and X. Chen, *Fracton Models on General Three-Dimensional Manifolds*, *Phys. Rev. X* **8**, 031051 (2018), [arXiv:1712.05892].

[42] W. Shirley, K. Slagle, and X. Chen, *Fractional excitations in foliated fracton phases*, *Annals Phys.* **410**, 167922 (2019), [arXiv:1806.08625].

[43] W. Shirley, K. Slagle, and X. Chen, *Foliated fracton order in the checkerboard model*, *Phys. Rev. B* **99**, 115123 (2019), [arXiv:1806.08633].

[44] W. Shirley, K. Slagle, and X. Chen, *Foliated fracton order from gauging subsystem symmetries*, *SciPost Phys.* **6**, 041 (2019), [arXiv:1806.08679].

[45] K. Slagle, D. Aasen, and D. Williamson, *Foliated Field Theory and String-Membrane-Net Condensation Picture of Fracton Order*, *SciPost Phys.* **6**, 043 (2019), [arXiv:1812.01613].

[46] W. Shirley, K. Slagle, and X. Chen, *Twisted foliated fracton phases*, *Phys. Rev. B* **102**, 115103 (2020), [arXiv:1907.09048].

[47] K. Slagle, *Foliated Quantum Field Theory of Fracton Order*, *Phys. Rev. Lett.* **126**, 101603 (2021), [arXiv:2008.03852].

[48] P.-S. Hsin and K. Slagle, *Comments on foliated gauge theories and dualities in 3+1d*, *SciPost Phys.* **11**, 032 (2021), [arXiv:2105.09363].

[49] H. Geng, S. Kachru, A. Karch, R. Nally, and B. C. Rayhaun, *Fractons and Exotic Symmetries from Branes*, *Fortsch. Phys.* **69**, 2100133 (2021), [arXiv:2108.08322].

[50] K. Slagle, A. Prem, and M. Pretko, *Symmetric Tensor Gauge Theories on Curved Spaces*, *Annals Phys.* **410**, 167910 (2019), [arXiv:1807.00827].

[51] A. Gromov, *Chiral Topological Elasticity and Fracton Order*, *Phys. Rev. Lett.* **122**, 076403 (2019), [arXiv:1712.06600].

[52] A. Jain and K. Jensen, *Fractons in curved space*, *SciPost Phys.* **12**, 142 (2022), [arXiv:2111.03973].

[53] H. Ebisu and B. Han, *Anisotropic higher rank $\mathbb{Z}_N$ topological phases on graphs*, arXiv:2209.07987.

[54] K. Slagle and Y. B. Kim, *Quantum Field Theory of X-Cube Fracton Topological Order and Robust Degeneracy from Geometry*, *Phys. Rev. B* **96**, 195139 (2017), [arXiv:1708.04619].

[55] D. Bulmash and M. Barkeshli, *Generalized $U(1)$ Gauge Field Theories and Fractal Dynamics*, arXiv:1806.01855.

[56] M. Pretko, *The Fracton Gauge Principle*, *Phys. Rev. B* **98**, 115134 (2018), [arXiv:1807.11479].

[57] N. Seiberg, *Field Theories With a Vector Global Symmetry*, *SciPost Phys.* **8**, 050 (2020), [arXiv:1909.10544].

[58] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *fcc lattice, checkerboards, fractons, and quantum field theory*, *Phys. Rev. B* **103**, 205116 (2021), [arXiv:2010.16414].

[59] T. Rudelius, N. Seiberg, and S.-H. Shao, *Fractons with Twisted Boundary Conditions and Their Symmetries*, *Phys. Rev. B* **103**, 195113 (2021), [arXiv:2012.11592].

[60] 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**, 581–590 (1975).

[61] T. Sulejmanpasic and C. Gattringer, *Abelian gauge theories on the lattice: θ-Terms and compact gauge theory with(out) monopoles*, *Nucl. Phys. B* **943**, 114616 (2019), [arXiv:1901.02637].

[62] D. J. Gross and I. R. Klebanov, *ONE-DIMENSIONAL STRING THEORY ON A CIRCLE*, *Nucl. Phys. B* **344**, 475–498 (1990).

[63] C. Xu and C. Wu, *Resonating plaquette phases in su(4) heisenberg antiferromagnet*, *Physical Review B* **77** (Apr, 2008), [arXiv:0801.0744].

[64] D. A. Johnston, M. Mueller, and W. Janke, *Plaquette Ising models, degeneracy and scaling*, *European Physical Journal Special Topics* **226** (Mar., 2017), [arXiv:1612.00060].

[65] D. Bulmash and M. Barkeshli, *The Higgs Mechanism in Higher-Rank Symmetric $U(1)$ Gauge Theories*, *Phys. Rev. B* **97**, 235112 (2018), [arXiv:1802.10099].

[66] H. Ma, M. Hermele, and X. Chen, *Fracton topological order from the Higgs and partial-confinement mechanisms of rank-two gauge theory*, *Phys. Rev. B* **98**, 035111 (2018), [arXiv:1802.10108].

[67] D. Radicevic, *Systematic Constructions of Fracton Theories*, arXiv:1910.06336.

[68] A. Polyakov, *Quark confinement and topology of gauge theories*, *Nuclear Physics B* **120**, 429 – 458 (1977).

[69] 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** (May, 2011), [arXiv:1011.0055].

[70] Y. You, Z. Bi, and M. Pretko, *Emergent fractons and algebraic quantum liquid from plaquette melting transitions*, *Phys. Rev. Res.* **2**, 013162 (2020), [arXiv:1908.08540].

[71] Y. You, F. J. Burnell, and T. L. Hughes, *Multipolar Topological Field Theories: Bridging Higher Order Topological Insulators and Fractons*, *Phys. Rev. B* **103**, 245128 (2021), [arXiv:1909.05868].

[72] A. Karch and A. Raz, *Reduced Conformal Symmetry*, *JHEP* **04**, 182 (2021), [arXiv:2009.12308].

[73] 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?*, arXiv:2101.01724.

[74] O. Dubinkin, A. Rasmussen, and T. L. Hughes, *Higher-form Gauge Symmetries in Multipole Topological Phases*, *Annals Phys.* **422**, 168297 (2020), [arXiv:2007.05539].

[75] Y. You, T. Devakul, F. J. Burnell, and S. L. Sondhi, *Symmetric Fracton Matter: Twisted and Enriched*, *Annals Phys.* **416**, 168140 (2020), [arXiv:1805.09800].

[76] A. Gromov, A. Lucas, and R. M. Nandkishore, *Fracton hydrodynamics*, *Phys. Rev. Res.* **2**, 033124 (2020), [arXiv:2003.09429].

[77] M. Qi, L. Radzihovsky, and M. Hermele, *Fracton phases via exotic higher-form symmetry-breaking*, *Annals Phys.* **424**, 168360 (2021), [arXiv:2010.02254].

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

[79] C. Cordova, D. S. Freed, H. T. Lam, and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications I*, *SciPost Phys.* **8**, 001 (2020), [arXiv:1905.09315].

[80] C. Cordova, D. S. Freed, H. T. Lam, and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications II*, *SciPost Phys.* **8**, 002 (2020), [arXiv:1905.13361].

[81] M. Bauer, G. Girardi, R. Stora, and F. Thuillier, *A Class of topological actions*, *JHEP* **08**, 027 (2005), [hep-th/0406221].

[82] D. S. Freed, G. W. Moore, and G. Segal, *The Uncertainty of Fluxes*, *Commun. Math. Phys.* **271**, 247–274 (2007), [hep-th/0605198].

[83] D. S. Freed, G. W. Moore, and G. Segal, *Heisenberg Groups and Noncommutative Fluxes*, *Annals Phys.* **322**, 236–285 (2007), [hep-th/0605200].

[84] 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.*, pp. 50–80. Springer, Berlin, 1985.

[85] P. Deligne, *Théorie de Hodge. II*, *Inst. Hautes Études Sci. Publ. Math.* 5–57 (1971).

[86] 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)*, pp. 137–225. Amer. Math. Soc., Providence, RI, 1999.

[87] M. J. Hopkins and I. M. Singer, *Quadratic functions in geometry, topology, and M-theory*, *J. Diff. Geom.* **70**, 329–452 (2005), [math/0211216].

[88] U. Bunke, *Differential cohomology*, 1208.3961. preprint.

[89] U. Schreiber, *Differential cohomology in a cohesive infinity-topos*, 1310.7930. preprint.

[90] U. Bunke, T. Nikolaus, and M. Völkl, *Differential cohomology theories as sheaves of spectra*, *Journal of Homotopy and Related Structures* **11**, 1–66 (Oct, 2014), [1311.3188].

[91] S. Sachdev and K. Park, *Ground states of quantum antiferromagnets in two dimensions*, *Annals of Physics* **298**, 58–122 (May, 2002).

[92] S. Elitzur, R. B. Pearson, and J. Shigemitsu, *Phase structure of discrete abelian spin and gauge systems*, *Phys. Rev. D* **19**, 3698–3714 (Jun, 1979).

[93] L. P. Kadanoff, *Multicritical behavior at the kosterlitz-thouless critical point*, *Annals of Physics* **120**, 39–71 (1979).

[94] J. L. Cardy, *General discrete planar models in two dimensions: Duality properties and phase diagrams*, *Journal of Physics A: Mathematical and General* **13**, 1507–1515 (apr, 1980).

[95] F. C. Alcaraz and R. Koberle, *Duality and the Phases of Z(n) Spin Systems*, *J. Phys.* **A13**, L153 (1980).

[96] E. Fradkin and L. P. Kadanoff, *Disorder variables and para-fermions in two-dimensional statistical mechanics*, *Nuclear Physics B* **170**, 1–15 (1980).

[97] 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**, 215–225 (1985). [Zh. Eksp. Teor. Fiz.89,380(1985)].

[98] R. Savit, *Duality in Field Theory and Statistical Systems*, *Rev. Mod. Phys.* **52**, 453 (1980).

[99] T. Banks, R. Myerson, and J. Kogut, *Phase transitions in abelian lattice gauge theories*, *Nuclear Physics B* **129**, 493–510 (1977).

[100] B. E. Baaquie, *(2 + 1)-dimensional abelian lattice gauge theory*, *Phys. Rev. D* **16**, 3040–3046 (Nov, 1977).

[101] R. Savit, *Topological excitations in u(1)-invariant theories*, *Phys. Rev. Lett.* **39**, 55–58 (Jul, 1977).

[102] C. Gattringer, D. Göschl, and T. Sulejmanpasic, *Dual simulation of the 2d U(1) gauge Higgs model at topological angle $\theta = \pi :$ Critical endpoint behavior*, *Nucl. Phys. B* **935**, 344–364 (2018), [`arXiv:1807.07793`].

[103] M. Anosova, C. Gattringer, D. Göschl, T. Sulejmanpasic, and P. Törek, *Topological terms in abelian lattice field theories*, *Proceedings of Science* **LATTICE2019**, 082 (2019), [`arXiv:1912.11685`].

[104] T. Sulejmanpasic, D. Göschl, and C. Gattringer, *First-Principles Simulations of 1+1D Quantum Field Theories at $\theta = \pi$ and Spin Chains*, *Phys. Rev. Lett.* **125**, 201602 (2020), [`arXiv:2007.06323`].

[105] T. Sulejmanpasic, *Ising model as a $U(1)$ lattice gauge theory with a θ-term*, *Phys. Rev. D* **103**, 034512 (2021), [`arXiv:2009.13383`].

[106] 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**, 1013 (1980).

[107] E. Fradkin and S. H. Shenker, *Phase diagrams of lattice gauge theories with higgs fields*, *Phys. Rev. D* **19**, 3682–3697 (Jun, 1979).

[108] T. Banks and E. Rabinovici, *Finite Temperature Behavior of the Lattice Abelian Higgs Model*, *Nucl. Phys. B* **160**, 349–379 (1979).

[109] F. J. Wegner, *Duality in Generalized Ising Models and Phase Transitions Without Local Order Parameters*, *J. Math. Phys.* **12**, 2259–2272 (1971).

[110] D. Horn, M. Weinstein, and S. Yankielowicz, *Hamiltonian approach to $z(n)$ lattice gauge theories*, *Phys. Rev. D* **19**, 3715–3731 (Jun, 1979).

[111] E. Fradkin and L. Susskind, *Order and disorder in gauge systems and magnets*, *Phys. Rev. D* **17**, 2637–2658 (May, 1978).

[112] J. M. Maldacena, G. W. Moore, and N. Seiberg, *D-brane charges in five-brane backgrounds*, *JHEP* **10**, 005 (2001), [hep-th/0108152].

[113] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys. Rev. D* **83**, 084019 (2011), [arXiv:1011.5120].

[114] M. Y. Khlopov, *Fractionally charged particles and quark confinement*, *JETP Letters* **33**, 162–164 (Feb 1981).

[115] Alexander, S. and Orbach, R., *Density of states on fractals : " fractons "*, *J. Physique Lett.* **43**, 625–631 (1982).

[116] M. Pretko, *Generalized Electromagnetism of Subdimensional Particles: A Spin Liquid Story*, *Phys. Rev. B* **96**, 035119 (2017), [arXiv:1606.08857].

[117] A. Gromov, *Towards classification of Fracton phases: the multipole algebra*, *Phys. Rev. X* **9**, 031035 (2019), [arXiv:1812.05104].

[118] G. Bentsen, I.-D. Potirniche, V. B. Bulchandani, T. Scaffidi, X. Cao, X.-L. Qi, M. Schleier-Smith, and E. Altman, *Integrable and Chaotic Dynamics of Spins Coupled to an Optical Cavity*, *Phys. Rev. X* **9**, 041011 (2019), [arXiv:1904.10966].

[119] G. Bentsen, T. Hashizume, A. S. Buyskikh, E. J. Davis, A. J. Daley, S. S. Gubser, and M. Schleier-Smith, *Treelike interactions and fast scrambling with cold atoms*, *Phys. Rev. Lett.* **123**, 130601 (2019), [arXiv:1905.11430].

[120] A. Periwal, E. S. Cooper, P. Kunkel, J. F. Wienand, E. J. Davis, and M. Schleier-Smith, *Programmable interactions and emergent geometry in an array of atom clouds*, *Nature* **600**, 630–635 (2021), [arXiv:2106.04070]. [Erratum: Nature 603, E29 (2022)].

[121] N. Manoj and V. B. Shenoy, *Arboreal Topological and Fracton Phases*, arXiv:2109.04259.

[122] F. J. Burnell, T. Devakul, P. Gorantla, H. T. Lam, and S.-H. Shao, *Anomaly inflow for subsystem symmetries*, *Phys. Rev. B* **106**, 085113 (2022), [arXiv:2110.09529].

[123] R. Grone and R. Merris, *A bound for the complexity of a simple graph*, *Discrete Mathematics* **69**, 97–99 (1988).

[124] N. Alon, *The number of spanning trees in regular graphs*, *Random Structures and Algorithms* **1**, 175–181 (Jan., 1990).

[125] F. Chung, *Spectral Graph Theory*, vol. 92. Published for the Conference Board of the mathematical sciences by the American Mathematical Society, 1997.

[126] S. Corry and D. Perkinson, *Divisors and Sandpiles: An Introduction to Chip-firing*. AMS Non-Series Monographs. American Mathematical Society, 2018.

[127] P. Gorantla, H. T. Lam, N. Seiberg, and S.-H. Shao, *2+1d Compact Lifshitz Theory, Tensor Gauge Theory, and Fractons*, `arXiv:2209.10030`.

[128] C. L. Henley, *Relaxation time for a dimer covering with height representation*, *Journal of Statistical Physics* **89**, 483–507 (Nov, 1997), [`cond-mat/9607222`].

[129] R. Moessner, S. L. Sondhi, and E. Fradkin, *Short-ranged resonating valence bond physics, quantum dimer models, and ising gauge theories*, *Physical Review B* **65** (Dec, 2001), [`cond-mat/0103396`].

[130] A. Vishwanath, L. Balents, and T. Senthil, *Quantum criticality and deconfinement in phase transitions between valence bond solids*, *Physical Review B* **69** (Jun, 2004), [`cond-mat/0311085`].

[131] E. Fradkin, D. A. Huse, R. Moessner, V. Oganesyan, and S. L. Sondhi, *Bipartite rokhsar–kivelson points and cantor deconfinement*, *Physical Review B* **69** (Jun, 2004), [`cond-mat/0311353`].

[132] E. Ardonne, P. Fendley, and E. Fradkin, *Topological order and conformal quantum critical points*, *Annals Phys.* **310**, 493–551 (2004), [`cond-mat/0311466`].

[133] P. Ghaemi, A. Vishwanath, and T. Senthil, *Finite-temperature properties of quantum lifshitz transitions between valence-bond solid phases: An example of local quantum criticality*, *Physical Review B* **72**, 024420 (Jul, 2005), [`cond-mat/0412409`].

[134] B. Chen and Q.-G. Huang, *Field Theory at a Lifshitz Point*, *Phys. Lett. B* **683**, 108–113 (2010), [`arXiv:0904.4565`].

[135] H. Ma and M. Pretko, *Higher-rank deconfined quantum criticality at the Lifshitz transition and the exciton Bose condensate*, *Physical Review B* **98**, 125105 (Sept., 2018), [`arXiv:1803.04980`].

[136] J.-K. Yuan, S. A. Chen, and P. Ye, *Fractonic Superfluids*, *Phys. Rev. Res.* **2**, 023267 (2020), [`arXiv:1911.02876`].

[137] E. Lake, M. Hermele, and T. Senthil, *The dipolar Bose-Hubbard model*, `arXiv:2201.04132`.

[138] Y.-T. Oh, J. Kim, E.-G. Moon, and J. H. Han, *Rank-2 toric code in two dimensions*, *Phys. Rev. B* **105**, 045128 (Jan, 2022), [`arXiv:2110.02658`].

[139] Y.-T. Oh, J. Kim, and J. H. Han, *Effective field theory of dipolar braiding statistics in two dimensions*, *Phys. Rev. B* **106**, 155150 (Oct, 2022), [`arXiv:2204.01279`].

[140] S. D. Pace and X.-G. Wen, *Position-dependent excitations and uv/ir mixing in the F_N rank-2 toric code and its low-energy effective field theory*, *Phys. Rev. B* **106**, 045145 (Jul, 2022), [`arXiv:2204.07111`].

[141] K. A. Berman, *Bicycles and spanning trees*, *SIAM Journal on Algebraic Discrete Methods* **7**, 1–12 (1986).

[142] H. J. S. Smith, *Xv. on systems of linear indeterminate equations and congruences*, *Philosophical Transactions of the Royal Society of London* **151**, 293–326 (1861).

[143] D. Lorenzini, *Smith normal form and laplacians*, *Journal of Combinatorial Theory, Series B* **98**, 1271–1300 (2008).

[144] R. Bacher, P. d. La Harpe, and T. Nagnibeda, *The lattice of integral flows and the lattice of integral cuts on a finite graph*, *Bulletin de la Société Mathématique de France* **125**, 167–198 (1997).

[145] M. Baker and S. Norine, *Riemann–roch and abel–jacobi theory on a finite graph*, *Advances in Mathematics* **215**, 766–788 (2007), [`math/0608360`].

[146] D. Dhar, *Self-organized critical state of sandpile automaton models*, *Phys. Rev. Lett.* **64**, 1613–1616 (Apr, 1990).

[147] D. Dhar, P. Ruelle, S. Sen, and D. N. Verma, *Algebraic aspects of abelian sandpile models*, *Journal of Physics A: Mathematical and General* **28**, 805 (feb, 1995).

[148] D. J. Lorenzini, *A finite group attached to the laplacian of a graph*, *Discrete Mathematics* **91**, 277–282 (1991).

[149] N. L. Biggs, *Chip-firing and the critical group of a graph*, *Journal of Algebraic Combinatorics* **9**, 25–45 (1999).

[150] G. R. Kirchhoff, *Uber die auflosung der gleichungen, auf welche man bei der untersuchung der linearen verteilung galvanischer strome gefuhrt wird*, *Ann. Phys. Chem.* **72**, 497–508 (1847).

[151] A. Cayley, *A theorem on trees*, *The Quarterly Journal of Mathematics* **23**, 376–378 (1889).

[152] M. Fiedler, *Algebraic connectivity of graphs*, *Czechoslovak Mathematical Journal* **23**, 298–305 (1973).

[153] A. Nilli, *On the second eigenvalue of a graph*, *Discrete Mathematics* **91**, 207–210 (1991).

[154] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59**, 135–157 (1980).

[155] F. Y. Wu, *Number of spanning trees on a lattice*, *Journal of Physics A: Mathematical and General* **10**, L113–L115 (jun, 1977).

[156] J. Haah, *Commuting pauli hamiltonians as maps between free modules*, *Communications in Mathematical Physics* **324**, 351–399 (oct, 2013), [`arXiv:1204.1063`].

[157] R. J. Duffin and E. L. Peterson, *The discrete analogue of a class of entire functions*, *Journal of Mathematical Analysis and Applications* **21**, 619–642 (1968).

[158] L. Lovász, *Discrete analytic functions: an exposition*, *Surveys in Differential Geometry* **9**, 241–273 (2004).

[159] S. Gao, *Absolute irreducibility of polynomials via newton polytopes*, *Journal of Algebra* **237**, 501–520 (2001).

[160] D. J. Williamson, Z. Bi, and M. Cheng, *Fractonic Matter in Symmetry-Enriched U(1) Gauge Theory*, *Phys. Rev. B* **100**, 125150 (2019), [`arXiv:1809.10275`].

[161] C. Castelnovo and C. Chamon, *Topological quantum glassiness*, *Philosophical Magazine* **92**, 304–323 (2012), [`https://doi.org/10.1080/14786435.2011.609152`].

[162] B. Yoshida, *Exotic topological order in fractal spin liquids*, *Phys. Rev. B* **88**, 125122 (2013), [`arXiv:1302.6248`].

[163] I. Niven, H. Zuckerman, and H. Montgomery, *An Introduction to the Theory of Numbers*. Wiley, 5th ed., 1991.

[164] P. Moree, *Artin's primitive root conjecture - a survey*, *Integers* **12**, 1305–1416 (2012).

[165] C. Hooley, *On Artin's conjecture*, *J. Reine Angew. Math.* **1967**, 209–220 (1967).

[166] D. S. Dummit and R. M. Foote, *Abstract Algebra*. Wiley, 3rd ed., 2004.

[167] D. Eisenbud, *Commutative Algebra: with a View Toward Algebraic Geometry*, vol. 150 of *Graduate Texts in Mathematics*. Springer New York, 2013.

[168] D. A. Cox, *Galois Theory*. John Wiley & Sons, Ltd, 2nd ed., 2012.

[169] S. Gao and V. M. Rodrigues, *Irreducibility of polynomials modulo p via newton polytopes*, *Journal of Number Theory* **101**, 32–47 (2003).