Title

Regularizability and bosonization of relativistic massless fermion

Author

Huang, Yen-Ta

Publication Date

2021

Regularizability and bosonization of relativistic massless fermion

by

Yen-Ta Huang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Dung-Hai Lee, Chair
Professor Joel Moore
Professor Michael Hutchings

Spring 2021

Regularizability and bosonization of relativistic massless fermion

Abstract

Regularizability and bosonization of relativistic massless fermion

by

Yen-Ta Huang

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Dung-Hai Lee, Chair

In this dissertation, we discuss the non-regularizability and the bosonization of massless fermions with relativistic dispersion in $1,2,3$ spatial dimensions. The non-regularizability is the root of various quantum anomalies and plays a central role in the physics of symmetry-protected topological phases. We generalized the Nielsen-Ninomiya theorem to all minimal nodal free fermion field theories protected by the time reversal, charge conservation, and charge conjugation symmetries. We prove that these massless field theories cannot be regularized on a lattice while respecting the protection symmetries. However, they can be realized on the boundaries of symmetry-protected topological phases in one higher dimension. We then generalize Witten's non-abelian bosonization of massless free fermion theories in one spatial dimension to two and three spatial dimensions. We shown the resulting boson theories share the same emergent symmetries and anomalies with the fermion theories. Moreover, we also show the boson theories possess fermion degrees of freedom, namely solitons. These bosonized models are non-linear sigma models with level-1 Wess-Zumino-Witten terms. As applications, we apply the bosonization results to the $SU(2)$ gauge theory of the Mott insulating phase of nearest-neighbor hopping Hubbard model, "bipartite-Mott insulators" in 1,2,3 spatial dimensions and twisted bilayer graphene.

To my fellow human beings

Contents

List of Figures

Acknowledgments

Like a tiny drop of dew, or a bubble floating in a stream;
Like a flash of lighting in a summer cloud,
Or a flickering lamp, an illusion, a phantom, or a dream.
So is all conditioned existence to be seen.

—Diamond Sutra

Such is the graduate program. I am nevertheless indebted to countless kindnesses and lessons from others.

My advisor Dung-Hai Lee. His straightforwardness, patience, and passion. To persistently strive for a simple and intuitive understanding of the nature of things. Not to be satisfied with “just getting the gist of it.” Not to waste time on nonsense. Being mindful of the thread of thoughts.

Lokman Tsui. Kindness, generosity and encouragement for everyone around. Always look for the positive sides of people and things. Self-reliance and cheerfulness.

Zi-Xiang Li. His altruism. Always be sincere and try to be helpful to others. Be honest with what one knows. Praise and thank others whenever appropriate.

Erik Aldape, Charles Xiong, and Tianrui Xu. To discuss physics in a playful spirit.

Anne Takizawa, Donna Sakima, Joelle Miles, and Claudia Trujillo. Always do one’s best and be ready to help others. Put oneself in other’s shoes.

Chien-I Chiang. His hospitality and meticulousness even on chores. Never give up asking seemingly simple questions.

Yi-Chuan Lu. His ability to get along with everyone. Always help others when it’s within one’s ability.

Teerawat Vilaisakulyong, Linden Chiu, Matthew Mo, Jeff Xu, James Lou, Christopher Kim, David Schneider-Joseph, Ming Curran, Yan Zhang, Geoffrey Lee, Peter de Blanc, Yuichiro Suzuki, Dennis Rhodes, and Len Sonosski. Openness in discussing different points of view. Patience in answering my recurring questions, and encouragement in various aspects.

Mei-Chun Lin. Do or do not based on justice and correctness, instead of personal convenience.

My family. Unconditional loves and supports. Respect others’ decisions.

All sentient beings on the spacetime trajectory that I do and do not remember, for their kindness in helping or testing me directly or indirectly.

Mother Nature. For the fabulous world.

Part I

Introduction

Chapter 1

Introduction: Dirac equation, anomalies, and symmetry protected topological phases

The Dirac equation [1] is a crown jewel of theoretical physics. It combines Einstein's special relativity with quantum mechanics, and predicts electron spin and the existence of position. Although it was originally proposed as the theory of electron, it turns out that all elementary matter-particles, with the exception of Higgs boson, are described by the Dirac/Majorana equation in the standard model. In condensed matter systems, it's not apparent at the first sight why relativistic massless fermions are relevant, because usually the Lorentz symmetry is broken in such systems. Therefore it is a bit of a surprise that relativistic massless fermions emerge in the low energy effective field theories in many important condensed matter systems. This include the Mott insulator, Dirac/Weyl semi-metals, magic-angle-twisted-bilayer graphene and the boundaries of topological insulators/superconductors, e.t.c. The only difference is the speed of light is replaced with the Fermi velocity. It is completely unexpected that the deep insight of Dirac finds applications in such distant corners of physics.

In the subsequent developments in quantum field theory, the concept of (quantum) anomalies played an important role in modern particle physics. Historically, the Adler-Bell-Jackiw (ABJ) anomaly [2, 3] was the first discovered, where the axial $U(1)$ symmetry of the Dirac fermion in odd space dimensions is anomalous when the fermion are coupled to the external (vector) $U(1)$ gauge field. The anomaly is manifested by the non-conservation of the axial $U(1)$ current. This is later generalized by Bardeen [4] to non-abelian symmetries. In a good sense the recent advance in condensed matter physics, namely, the discovery of symmetry-protected topological states is a revisit of anomaly in condensed matter physics. This dissertation benefits from such developments and push the forefront further.

Anomaly and regularization are inextricable. Regularization makes a low energy theory well-defined in the high energy. In the original ABJ calculation, the axial symmetry is broken when one chooses a regularization preserving the vector $U(1)$ symmetry, which is necessary for the vector $U(1)$ gauge theory (i.e., electromagnetism) to be well-defined. Nielsen-Ninomiya gave an even more concrete demonstration [5]. They showed that Weyl fermion, i.e., one chiral part of the Dirac fermion, *cannot* be regularized on a lattice by itself while preserving $U(1)$ charge conservation. This is a necessary condition for the ABJ anomaly to hold, as argued in the following. A good (on-site) continuous symmetry at lattice level can be gauged without anomaly by Peierls substitution. If a Weyl fermion can be regularized with the $U(1)$ charge conservation intact, one can stack two copies of Weyl fermions of opposite chiralities. The resulting theory has a Dirac fermion spectrum in the low energy while both the vector and axial $U(1)$ can be gauged without issue, contradicting the result of the ABJ anomaly. In short, an anomaly can show up when regularization is impossible while preserving the symmetries at all energy.

A caveat is in order at this point. For the non-regularizability, we mean that the theory cannot be regularized in the same spatial dimension. However, it is generally possible to regularize an anomalous theory on the boundary of a system in the one-higher space dimension. Take the chiral fermion as an example again. By Nielsen-Ninomiya theorem, a single chiral fermion cannot be regularized the one spatial dimension. It can nevertheless live on the boundary of a Chern insulator, e.g., on one of the boundaries of an annulus. In the presence of a time-dependent flux associated with the (vector) electromagnetic field, the charge on one boundary is not conserved by itself. The non-conservation is canceled out by the induced Hall current in the radial direction, and the total charge of the system is still conserved. We shall come back to this example in part III. The phenomenon that the anomaly on the boundary is canceled out by the higher dimensional bulk is termed *anomaly inflow* [6].

The physics of symmetry-protected topological (SPT) phases can be viewed as the reincarnation of the anomaly inflow picture. SPT phases are the generalization of Chern-insulator: a system is in an SPT phase if the Hamiltonian respects certain a symmetry group G and the bulk has a uniquely gapped symmetric ground state on any closed manifold. For the Chern insulator, the symmetry group is the charge $U(1)$ symmetry. When a non-trivial SPT phase is cut open, there are several possibilities on the boundary: 1) the boundary has spontaneous symmetry breaking, 2) establishes surface topological order, or 3) is gapless. In the present work, we shall focus on the last possibility. For fermionic SPTs, one can choose massless relativistic fermions as the representatives of the gapless boundaries [7] (in the rest of the dissertation, unless otherwise stated, "massless fermion" is always referred to massless relativistic fermion). The gaplessness of the fermion is protected by the symmetry in question. Furthermore, the massless fermion cannot be regularized in its dimension (i.e., the boundary is anomalous). The fermionic SPT phases are characterized by the anomalies of the massless fermions on the boundaries. The reverse is also true: all anomalous massless fermion can be realized on the boundary one higher dimensional SPT [8].

This naturally leads us to two questions:

1. What kind of massless fermions together with symmetry groups are non-regularizable

(anomalous)?

2. Are there other gapless theories with the same anomalies so that we can attach them to the boundaries of the same SPTs? If so, what is the relation between these gapless theories and the original massless fermions at low energy?

Each part of the present dissertation attempts to answer one of the questions above. In part II, we demonstrate that the boundaries of free fermion (non-trivial) SPTs are indeed non-regularizable, through a case-by-case study. This is consistent with the view in [7] that the non-trivialness of an SPT can be told by the absence of mass term on the boundary. In part III, we construct a series of bosonic models sharing the same symmetries and anomalies with the massless fermions. The anomalies are dictated by the level-1 Wess-Zumino-Witten (WZW) terms in the bosonic theories in $1, 2, 3$ spatial dimensions. In $(1+1)$-D, the bosonic theory reproduces the non-abelian bosonization by Witten [9]; in $(2+1)$-D and $(3+1)$-D, the solitons inherit fermion statistics from the WZW terms. This leads us to the proposal that these bosonic theories are equivalent to the original massless fermions, generalizing the non-abelian bosonizations to higher dimensions. Some applications of the bosonization follow. To make the main points clear, we generally demonstrate the ideas through simple examples in the main text and leave the detailed systematic check to the appendices.

Part II

Non-regularizability of relativistic fermion

Chapter 2

Non-regularizability of relativistic fermion

2.1 Idea of non-regularizability

The non-regularizability of massless free fermion field theories is the origin of various quantum anomalies. A famous example is the Nielsen-Ninomiya [5] theorem, namely, Weyl nodes with net chirality cannot be realized by any charge-conserved lattice model in three dimensions. However, Weyl nodes with net chirality can appear on the boundary of a 4D charge-conservation-protected topological insulator (The free fermion topological classification of this 4D topological insulator is $\mathbb{Z}$.). Another example involves Dirac cones with net vorticity in 2D. Under charge conservation and time-reversal ($T^2 = -1$) symmetries, Dirac cones with net vorticity cannot be realized by any lattice model. However, they can appear on the boundary of a 3D topological insulator. (The free fermion topological classification of such topological insulator is $\mathbb{Z}_2$.)

According to the folklore, the low energy field theory describing the boundary of on-site symmetry protected topological states (SPTs) cannot be regularized on a lattice. In other words, they can not be realized as finite-range tight-binding models where the symmetry acts on the degrees of freedom on each lattice site independently. The obstruction lies in the realization of symmetry – in the boundary dimension the on-site nature of the protection symmetry cannot be realized. This obstruction is relieved by "UV completing" the boundary degrees of freedom with the bulk degrees of freedom living in one extra spatial dimension. The bulk degrees of freedom are gapped and respect an on-site symmetry. The bulk state is called an SPT. When the boundry is one-dimensional, Ref.[10, 11] argued that the non-regularizability is manifested by the fact that the boundary theory is not modular invariant after orbifolding with respect to the protection symmetry.

If the protection symmetry is not on-site, regularization is certainly possible. A famous example is the tight-binding model of graphene. There, the two Dirac nodes are protected by the translation, charge conservation, time reversal, and inversion symmetries. Here the inversion symmetry is not on-site. In the rest of the paper we shall assume translation invariance and the term "symmetry" always refers to other on-site symmetry.

The non-regularizability discussed above lies at the heart of the physics of SPTs. It is well-known that SPTs are defined by their symmetry-protected gapless boundaries. In the following, we argue that if it were possible to realize these boundaries on a lattice, the gapless boundary modes will not be protected.

For example, in Fig. 2.1 we consider a 2D SPT having two edges. It is always possible to reconnect these edges with symmetry-respecting interactions, i.e., seal off the boundary (Fig. 2.1(a)). After the reconnection the gapless modes are removed (Fig. 2.1(b)). If it was possible to regularize the gapless boundaries on 1D lattices, one would have been able to fabricate the gapless boundaries as 1D systems (Fig. 2.1(c)). These fabricated edges can be brought around to interact with the original boundaries (via symmetry-respecting interactions) (Fig. 2.1(d)). As a result, the gapless edges can be removed (Fig. 2.1(e)), which proves that the original gapless edges are not symmetry-protected.

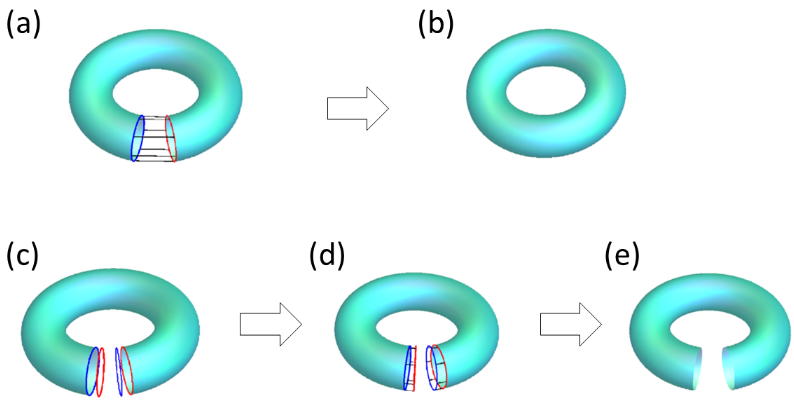


Figure 2.1: An illustration of the fact that the regularizability of the boundary Hamiltonian of an SPT implies the gapless modes are not protected. (a,b) By turning on symmetry-respecting interactions mimicking those in the bulk (the black lines) it is possible to gap out the gapless modes (red and blue circles). (c) The regularizability of the boundary SPT Hamiltonian implies it is possible to fabricate the gapless boundaries. (d,e) The fabricated boundaries can be brought to interact with the original boundaries and gap each other out.

The purpose of this paper is to prove the following folklore, namely:

Any symmetry-protected minimal nodal free-fermion field theory cannot be regularized on a lattice.

Here, "nodal free-fermion field theory" is a continuum field theory which has a gapless spectrum with a linear-dispersing gap node, characterized by a Clifford algebra, at a single time-reversal invariant momentum. Without loss of generality, we shall assume such momentum to be $\boldsymbol{k} = 0$. "Minimal" refers to the fact that the fermion field in the theory has the smallest number of components necessary to represent the symmetry transformations and the Clifford algebra. "Symmetry-protection" means there is no symmetry-allowed mass term. In this paper, we restrict ourselves to the charge conservation, time-reversal, and charge conjugation symmetries. "Lattice regularization" is the procedure which converts the continuum field theory to a finite-range tight-binding model while preserving all symmetries.

The outline of the current part is as follows. We achieve the proof by "reductio ad absurdum". In section 2.2, we assume the existence of a tight-binding Hamiltonian whose low energy limit is the field theory in question. Let the momentum space Hamiltonian of this tight-binding model be $h(\boldsymbol{k})$, we list the four constraints $h(\boldsymbol{k})$ must obey. In section 2.3 we present the $h(\boldsymbol{k})$ which has the smallest matrix size and satisfies the constraints listed in section 2.2. This is the momentum space Hamiltonian of the minimal models. In section 2.4 we lay out the symmetry protection hypothesis. In section 2.5 and A.1, we show that for each $h(\boldsymbol{k})$ obeying the constraints of sections 2.2 and 2.4 there is an associated "spectral symmetrised" counterpart, $\tilde{h}(\boldsymbol{k})$. In section 2.6 we apply the Poincaré-Hopf Theorem to $\tilde{h}(\boldsymbol{k})$, and show that it imposes a stringent constraint on the form of $\tilde{h}(\boldsymbol{k})$ at a time reversal invariant point $\boldsymbol{k}_0$ different from $\boldsymbol{k} = 0$. Section 2.7 adopts the strategy of reductio ad absurdum for the proof of non-regularizability. We complete the proof in two alternative ways. (a) When $\tilde{h}(\boldsymbol{k})$ satisfies a special condition we prove that if it obeys constraints 1-4 it must violate the symmetry-protection hypothesis. (b) For other $\tilde{h}(\boldsymbol{k})$ we prove that if it satisfies the symmetry-protection hypothesis it must have the energy gap close at $\boldsymbol{k}_0$ as well. This means it violates constraint 3 of section 2.2. The proof (b) is achieved by a case-by-case study of all nodal Hamiltonians protected by the charge conservation, time reversal, and charge conjugation symmetries. Because of the length of this proof, it is left to A.2 and A.3.

2.2 The constraints on lattice-regularized nodal Hamiltonians

In the following we assume the existence of lattice-regularized minimal SPN Hamiltonian

$$H = \sum_{\boldsymbol{k} \in BZ} \chi(-\boldsymbol{k})^T h(\boldsymbol{k}) \chi(\boldsymbol{k}), \tag{2.1}$$

and discuss the conditions it must satisfy. Here "BZ" stands for the Brillouin zone of a d-dimensional lattice. $\chi(\boldsymbol{k})$ is a Fourier transformed Majorana lattice field. We work with Majorana rather than complex fermion field because it also covers charge non-conserving (Bogoliubov-de Gennes) free fermion Hamiltonians. There are 4 constraints we require $h(\boldsymbol{k})$ to satisfy:

1. The Majorana constraint: $h^T(-\boldsymbol{k}) = -h(\boldsymbol{k})$.

2. $h(\boldsymbol{k})$ is an analytic function of $\boldsymbol{k}$ in the Brillouin zone.

3. There is an energy gap between the lower half and the upper half of the eigenvalues of $h(\boldsymbol{k})$. The energy gap between these two groups of eigenvalues exhibits a single node at $\boldsymbol{k} = 0$. Moreover

$$h(\boldsymbol{k}) \to \sum_{j=1}^{d} k_j \Gamma_j \text{ as } \boldsymbol{k} \to 0. \tag{2.2}$$

   Here $\{\Gamma_j\}$ are traceless symmetric matrices satisfying $\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}$.

4. $h(\boldsymbol{k})$ obeys the following symmetry requirement: $U_\beta^\dagger h(\boldsymbol{k}) U_\beta = h(\boldsymbol{k})$ and $A_\alpha^\dagger h(-\boldsymbol{k})^* A_\alpha = h(\boldsymbol{k})$. Here $A_\alpha,\ \alpha = 1, ..., N_A$ and $U_\beta,\ \beta = 1, ..., N_U$ are $\boldsymbol{k}$-independent orthogonal matrices representing the anti-unitary and unitary protection symmetries.

Four comments are in order:

- We assume that the Hamiltonian has translation symmetry so that we can express it in momentum space. The more general case where the translation symmetry is absent is more diffucult, and is beyond the scope of this paper. The fact that the unitary and anti-unitary symmetry matrices do not depend on $\boldsymbol{k}$ signifies that they are on-site symmetries.

- In the presence of anti-unitary symmetry, $A_\alpha^\dagger h(-\boldsymbol{k})^* A_\alpha = h(\boldsymbol{k})$ implies $A_\alpha^\dagger h(\boldsymbol{k}) A_\alpha = -h(\boldsymbol{k})$ due to constraint 1. As a result, the spectrum of $h(\boldsymbol{k})$ is symmetric about zero for each $\boldsymbol{k}$.

- In 1D we shall assume the dispersion of $h(k)$ is non-chiral. This is because for chiral Hamiltonians the constraints of continuity, Brillouin zone periodicity, and the requirement that the energy band crosses the Fermi energy only at $k = 0$ (which is the nodal condition for chiral Hamiltonians) obviously contradict one another.

- Conditions 3 and 4 impose a constraint on the minimal size of $h(\boldsymbol{k})$. We state, without proof, that the smallest such matrix for the charge conservation (unitary), time reversal (anti-unitary) and charge conjugation (unitary) symmetries has dimension $2^n \times 2^n$. Here n depends on the spatial dimension and the symmetry group.

2.3 The minimal model satisfying the constraints in section 2.2

In this section and the rest of the paper we shall focus on minimal SPN models. For these models $h(\boldsymbol{k})$ is a $2^n \times 2^n$ Hermitian matrix. Any such $2^n \times 2^n$ $h(\boldsymbol{k})$ can be constructed from linear combinations of the tensor products of n Pauli matrices. Among them $N_1 =$

$(2^{2n}+2^n)/2$ are real and symmetric and $N_2 = (2^{2n}-2^n)/2$ are imaginary and anti-symmetric, i.e.,

$$h(\boldsymbol{k}) = \sum_{i=1}^{N_1} o_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{N_2} e_j(\boldsymbol{k}) M_j^a. \tag{2.3}$$

Due to the Majorana constraint $o_i(\boldsymbol{k})$ and $e_j(\boldsymbol{k})$ are odd and even functions of $\boldsymbol{k}$, respectively. Under the action of unitary symmetries $\boldsymbol{k}$ remains unchanged. But anti-unitary symmetries send $\boldsymbol{k}$ to $-\boldsymbol{k}$. As the result, the $\{M_i^s\}$ and $\{M_j^a\}$ that can appear in Eq.(2.3) must satisfy the following equations

$$\begin{aligned} U_\beta^\dagger M_i^s U_\beta = M_i^s, \quad A_\alpha^\dagger M_i^s A_\alpha = -M_i^s \\ U_\beta^\dagger M_i^a U_\beta = M_i^a, \quad A_\alpha^\dagger M_i^a A_\alpha = -M_i^a. \end{aligned} \tag{2.4}$$

Let the number of symmetric/anti-symmetric matrices satisfying Eq.(2.4) be n_s and n_a, respectively. Thus

$$h(\boldsymbol{k}) = \sum_{i=1}^{n_s} o_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{n_a} e_j(\boldsymbol{k}) M_j^a. \tag{2.5}$$

The matrices M_i^s and M_j^a are "linear independent" with respect to the following definition of matrix inner product $\langle M_1|M_2\rangle = V_{M_1}^\dagger \cdot V_{M_2}$, where V_M is the column vector containing all matrix elements of M. In the following we shall order $\{M_i^s\}$ so that the first d of them are the Γ_i's in constraint 3. These $\{\Gamma_i\}$ satisfy the Clifford algebra $\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}$. Under the above ordering convention,

$$\begin{aligned} h(\boldsymbol{k}) \quad &= \sum_{i=1}^{n_s} o_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{n_a} e_j(\boldsymbol{k}) M_j^a \\ &= \sum_{i=1}^{d} o_i(\boldsymbol{k}) \Gamma_i + \sum_{i=d+1}^{n_s} o_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{n_a} e_i(\boldsymbol{k}) M_i^a. \end{aligned} \tag{2.6}$$

2.4 The symmetry protection hypothesis

Symmetry protection means that under the requirement of Eq.(2.4), there is no non-zero anti-symmetric matrix which anticommutes with all the Γ_i in Eq.(2.2) and Eq.(2.6).

2.5 Spectral Symmetrisation

Given a $h(\boldsymbol{k})$ satisfying constraints 1-4 in section 2.2, we can create a "spectral symmetrised" Hamiltonian satisfying the same constraints.

To perform spectral symmetrization, we first write $h(\boldsymbol{k})$ in terms of its eigenvalues and eigenvectors

$$h(\boldsymbol{k}) = U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} \tag{2.7}$$

where $D(\boldsymbol{k})$ is the diagonal matrix formed by the eigenvalues of $h(\boldsymbol{k})$ in descending order. $U_{\boldsymbol{k}}$ contains the eigenvectors. It is the unitary transformation necessary to diagonalize $h(\boldsymbol{k})$.

We first replace the upper and lower halves of the eigenvalues in $D(\boldsymbol{k})$ by their respective averages. After this replacement, $D(\boldsymbol{k})$ becomes $D'(\boldsymbol{k})$ and the Hamiltonian is given by

$$h'(\boldsymbol{k}) = U_{\boldsymbol{k}}^{\dagger} D'(\boldsymbol{k}) U_{\boldsymbol{k}}. \tag{2.8}$$

Note that in Eq.(2.8) $U_{\boldsymbol{k}}$ remains unchanged. From $h'(\boldsymbol{k})$ we define a new Hamiltonian by subtracting the average of the diagonal element $\bar{E}'(\boldsymbol{k})$ from each element of $D'(\boldsymbol{k})$ so that $D'(\boldsymbol{k}) \to \tilde{D}(\boldsymbol{k}) = D'(\boldsymbol{k}) - \bar{E}'(\boldsymbol{k}) I_n$. Here I_n represents the $2^n \times 2^n$ identity matrix. After the above two steps the Hamiltonian becomes

$$h(\boldsymbol{k}) \to \tilde{h}(\boldsymbol{k}) = U_{\boldsymbol{k}}^{\dagger} \tilde{D}(\boldsymbol{k}) U_{\boldsymbol{k}}. \tag{2.9}$$

$\tilde{h}(\boldsymbol{k})$ is the "spectral symmetrised Hamiltonian". Note that $U_{\boldsymbol{k}}$ still remains unchanged. $\tilde{h}(\boldsymbol{k})$ has the important property that

$$\tilde{h}(\boldsymbol{k})^2 \propto I_n \text{ for all } \boldsymbol{k}. \tag{2.10}$$

In A.1, we show that the spectral symmetrization does not jeopardize constraint 1-4 in section 2.2. In addition, it preserves the analyticity of $h(\boldsymbol{k})$ in the Brillouin zone region where the energy gap is non-zero. In particular, spectral symmetrization does not affect Eq.(2.2), i.e,

$$\tilde{h}(\boldsymbol{k}) \to \sum_{j=1}^{d} k_j \Gamma_j \quad \text{as } \boldsymbol{k} \to 0. \tag{2.11}$$

In Fig. 2.2 we show an example of spectral symmetrization in one dimension.

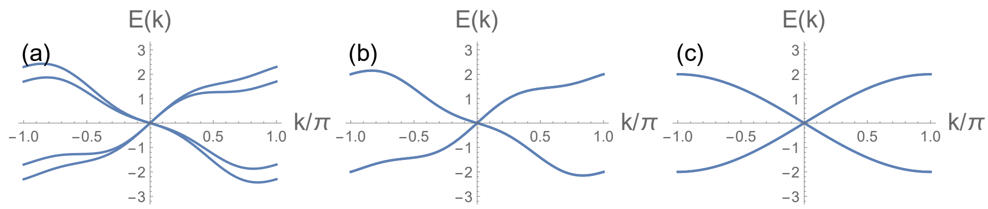


Figure 2.2: From the left to the middle panel we replaced the upper half and lower half of the eigenvalues at each k with their averages. From the middle panel to the right panel we subtracted the average of all eigenvalues from each eigenvalue at each k.

The spectral symmetrised $\tilde{h}(\boldsymbol{k})$ can also be written in the form of Eq.(2.6), *i.e.*

$$\begin{aligned}\tilde{h}(\boldsymbol{k}) &= \sum_{i=1}^{n_s} \tilde{o}_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{n_a} \tilde{e}_i(\boldsymbol{k}) M_i^a \\ &= \sum_{i=1}^{d} \tilde{o}_i(\boldsymbol{k}) \Gamma_i + \sum_{i=d+1}^{n_s} \tilde{o}_i(\boldsymbol{k}) M_i^s + \sum_{j=1}^{n_a} \tilde{e}_i(\boldsymbol{k}) M_i^a \\ &:= S(\boldsymbol{k}) + A(\boldsymbol{k}). \end{aligned} \tag{2.12}$$

Here the symmetric matrix $S(\boldsymbol{k})$ includes the first and the second sums, and the anti-symmetric matrix $A(\boldsymbol{k})$ includes the third sum.

2.6 The Poincaré-Hopf Theorem

The Poincaré-Hopf theorem (see, e.g., Ref.[12]) applies to a d-component vector function $\boldsymbol{f}(\boldsymbol{k}) = \{f_1(\boldsymbol{k}), ..., f_d(\boldsymbol{k})\}$ that vanishes at a discrete set of points $\{\boldsymbol{k}_n\}$ on a d-dimensional torus. The theorem states that the "index" of the $\boldsymbol{k} \to \boldsymbol{f}(\boldsymbol{k})$ map at each $\boldsymbol{k}_n$ must sum to zero. The meaning of the index is the following. Pick a closed ball D_n around each $\boldsymbol{k}_n$ so that $\boldsymbol{k}_n$ is the only zero of $\boldsymbol{f}(\boldsymbol{k})$ in D_n. We define the index at $\boldsymbol{k}_n$ to be the "degree" of the map from the boundary of D_n to the $(d-1)$-sphere formed by $\hat{f}(\boldsymbol{k}) = \boldsymbol{f}(\boldsymbol{k})/|\boldsymbol{f}(\boldsymbol{k})|$. For 3D the degree is the Pontryagin index of $\hat{f}(\boldsymbol{k})$, and in 2D it is the "winding number" of $\hat{f}(\boldsymbol{k})$. For 1D the degree is equal to $\left(\hat{f}(k_R) - \hat{f}(k_L)\right)/2$. Fig. 2.3 illustrates the degree 1 map for spatial dimension 1,2 and 3.

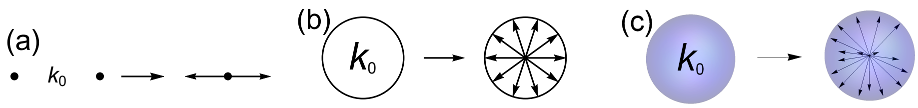


Figure 2.3: The degree 1 maps $\hat{f}(\boldsymbol{k})$ for (a) $d = 1$, (b) $d = 2$ and (c) $d = 3$.

Any zero of $\boldsymbol{f}(\boldsymbol{k})$ that has no mapping degree can be removed by infinitesimal changes. On the other hand, a zero that has non-zero mapping degree can only be shifted but not removed by infinitesimal changes. We assert, without proof, that it is always possible to deform $\tilde{h}(\boldsymbol{k})$ so that $\tilde{\boldsymbol{o}}(\boldsymbol{k}) := \{\tilde{o}_1(\boldsymbol{k}), ...\tilde{o}_d(\boldsymbol{k})\}$ only possess discrete zeros while keeping the symmetrised nature of the energy spectrum. Moreover, around each of the discrete zero $\tilde{\boldsymbol{o}}(\boldsymbol{k})$ exhibits a non-zero mapping degree.

Equation 2.11 implies $\tilde{\boldsymbol{o}}(\boldsymbol{k})$ has a degree 1 zero at $\boldsymbol{k} = 0$. Applying the Poincaré-Hopf theorem we conclude that the sum of the mapping degree in the rest of the Brillouin zone must be equal to -1. Due to the fact that $\tilde{\boldsymbol{o}}(\boldsymbol{k}) = -\tilde{\boldsymbol{o}}(-\boldsymbol{k})$, and the fact that the degree of mapping is not affected by the simultaneous sign reversal of both $\tilde{\boldsymbol{o}}$ and $\boldsymbol{k}$, we conclude that the sum of the mapping degree in the Brillouin zone excluding all time-reversal invariant $\boldsymbol{k}$ points must be an even integer. This, in turn, implies the sum of the mapping degrees across all non-zero time-reversal invariant $\boldsymbol{k}$ points must be an odd integer. (Note that by the oddness of $\tilde{\boldsymbol{o}}(\boldsymbol{k})$, it must vanish at any time reversal invariant $\boldsymbol{k}$ point.) Thus there must exist, at least, one non-zero time-reversal invariant $\boldsymbol{k}$ point, say, $\boldsymbol{k}_0$, where the mapping degree is an odd integer.

2.7 The reductio ad absurdum proof

In section 2.5 we have shown that given a lattice-regularized $h(\boldsymbol{k})$ satisfying constraints 1-4 in section 2.2 there is always a spectral symmetrised $\tilde{h}(\boldsymbol{k})$ which obeys all constraints of $h(\boldsymbol{k})$ and is lattice regularized.

In this section we complete the proof of non-regularizability via reductio ad absurdum. This proof is achieved in two alternative ways. (a) We prove that if $\tilde{\boldsymbol{o}}(\boldsymbol{k})$ satisfies a special condition (see below), and if $\tilde{h}(\boldsymbol{k})$ obeys constraints 1-4 of section 2.2, the symmetry-protection hypothesis must be violated. (b) For $\tilde{\boldsymbol{o}}(\boldsymbol{k})$ that violates the special condition, we prove that if $\tilde{h}(\boldsymbol{k})$ satisfies the symmetry-protection hypothesis its energy gap must also close at $\boldsymbol{k}_0$. This means constraint 3 of section 2.2 is violated. Because proof (b) involves a case-by-case study of all $\hat{T}, \hat{Q}, \hat{C}$ protected minimal SPN, we leave it to A.2 and A.3.

The symmetries under consideration are generated by the subsets of $\{\hat{T}, \hat{Q}, \hat{C}\}$. Here

$$\hat{Q} = i \sum_{\boldsymbol{k} \in BZ} \chi^T(\boldsymbol{k})\ Q\ \chi(\boldsymbol{k})$$

is the total charge operator. It generates the global charge U(1) gauge transformation. $\hat{T}$ and $\hat{C}$ are the generators of time reversal and charge conjugation symmetries. They act on the fermion operators according to

$$\begin{aligned} \hat{T}\chi(\boldsymbol{k})\hat{T}^{-1} &= T\chi(-\boldsymbol{k}) \\ \hat{C}\chi(\boldsymbol{k})\hat{C}^{-1} &= C\chi(\boldsymbol{k}) \end{aligned} \tag{2.13}$$

where T, Q, C are $2^n \times 2^n$ matrices. In A.2 we list the relevant T, Q, C and all symmetry allowed $\{M_i^s, i = 1, ..., n_s\}$ and $\{M_i^a, i = 1, ..., n_a\}$ in Eq.(2.12) for the minimal SPNs in spatial dimensions $1 \le d \le 3$.

Due to the spectral symmetrization condition, Eq.(2.10), the $S(\boldsymbol{k})$ and $A(\boldsymbol{k})$ in Eq.(2.12) must anticommute. This is because the square of $\tilde{h}(\boldsymbol{k})$ is

$$\tilde{h}(\boldsymbol{k})^2 = S(\boldsymbol{k})^2 + A(\boldsymbol{k})^2 + \{S(\boldsymbol{k}), A(\boldsymbol{k})\}, \tag{2.14}$$

since $\{S(\boldsymbol{k}), A(\boldsymbol{k})\}$ is an anti-symmetric matrix, while $\tilde{h}(\boldsymbol{k})^2$ is proportional to the identity matrix, it implies

$$\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0 \quad \text{for all} \quad \boldsymbol{k}. \tag{2.15}$$

Now apply Eq.(2.15) to $\boldsymbol{k}_0$. Since $\boldsymbol{k}_0$ is a time reversal invariant point $S(\boldsymbol{k}_0) = 0$, which means $\{\tilde{o}_1(\boldsymbol{k}_0), ..., \tilde{o}_d(\boldsymbol{k}_0)\} = 0$.

The simplest case to prove the contradiction is when (i) all d functions $\{\tilde{o}_1(\boldsymbol{k}_0 + \boldsymbol{q}), ..., \tilde{o}_d(\boldsymbol{k}_0 + \boldsymbol{q})\}$ vanish as the same power in $\boldsymbol{q}$ as $\boldsymbol{q} \to 0$, and (ii) all other $\tilde{o}_i(\boldsymbol{q})$, namely, $\tilde{o}_{d+1}(\boldsymbol{k}_0 + \boldsymbol{q}), ..., \tilde{o}_{n_s}(\boldsymbol{k}_0 + \boldsymbol{q})$, vanish as higher power in $\boldsymbol{q}$. Under such condition examining $\{S(\boldsymbol{k}_0 + \boldsymbol{q}), A(\boldsymbol{k}_0 + \boldsymbol{q})\} = 0$ to the lowest order in $\boldsymbol{q}$ gives us

$$\{A(\boldsymbol{k}_0), \Gamma_i\} = 0, \quad \text{for} \quad i = 1, ..., d. \tag{2.16}$$

Equation 2.16 implies $A(\boldsymbol{k}_0)$ acts like a mass term. Since $A(\boldsymbol{k}_0) = \tilde{h}(\boldsymbol{k}_0) \neq 0$ (otherwise $\tilde{h}(\boldsymbol{k})$ will have more than one gap node), this violates the symmetry-protection hypothesis. More specifically, including $A(\boldsymbol{k}_0)$ in the Hamiltonian

$$H = \int d^d x \; \chi^T(\boldsymbol{x}) \left[-i \sum_{i=1}^{d} \Gamma_i \partial_i + A(\boldsymbol{k}_0) \right] \chi(\boldsymbol{x}) \tag{2.17}$$

gaps out the node at $\boldsymbol{k} = 0$.

Under the more general condition, namely when $\{\tilde{o}_1(\boldsymbol{k}_0 + \boldsymbol{q}), ..., \tilde{o}_d(\boldsymbol{k}_0 + \boldsymbol{q})\}$ do not vanish as the same power in $\boldsymbol{q}$, and/or when $\tilde{o}_{d+1}(\boldsymbol{k}_0 + \boldsymbol{q}), ..., \tilde{o}_{n_s}(\boldsymbol{k}_0 + \boldsymbol{q})$ vanish slower than, or as slowly as, $\tilde{o}_1(\boldsymbol{k}_0 + \boldsymbol{q}), ..., \tilde{o}_d(\boldsymbol{k}_0 + \boldsymbol{q})$ the above proof does not apply.

Under such condition we adopt a different proof strategy. Instead, we assume the symmetry-protection hypothesis holds, and show that it is impossible for $\tilde{h}(\boldsymbol{k})$ to have gap node at only a single point in the Brillouin zone. This proof is achieved via a case-by-case study of all $\hat{T}, \hat{Q}, \hat{C}$ symmetry-protected minimal nodal Hamiltonians. Because of the length of the proof we leave it to A.2 and A.3.

2.8 Final discussion: the open issues

In the preceding discussions we have proven that all minimal nodal Hamiltonians protected by $\{\hat{T}, \hat{Q}, \hat{C}\}$ symmetries cannot be regularized on a lattice. Here we list some of the open issues. The first is the proof for non-minimal symmetry-protected nodal Hamiltonians. Such nodal Hamiltonians can be constructed by stacking the minimal nodal Hamiltonians together. Although it is clear that the non-regularizability of the minimal nodal Hamiltonians is a necessary condition for the non-regularizability of non-minimal symmetry-protected nodal Hamiltonians, it remains to be proven that it is a sufficient condition. The second issue concerns the assumption that in the spectral symmetrised Hamiltonian the coefficient functions in front of $\{\Gamma_1, ..., \Gamma_d\}$ exhibit isolated zeros. It remains to be proven that it is always possible to deform $\tilde{h}(\boldsymbol{k})$ so that the coefficient functions fulfill such a statement while maintaining the symmetrised spectrum. The third issue is the proof that a general symmetry-protected gapless Hamiltonian can be deformed into the single-node Hamiltonian discussed in this paper. We leave these open issues for future researches.

Part III

Bosonization of relativistic fermions

Chapter 3

Introduction of bosonization

Bosonization in $(1+1)$-D has been a very useful theoretical tool. It allows one to map a theory, where the fundamental degrees of freedom are fermionic, to a theory with bosonic degrees of freedom. Often, things that can be seen easily in one picture are difficult to see in the other. The best-known bosonization is the abelian bosonization[13, 14, 15], where fermions are solitons in the Bose field. A shortcoming of the abelian bosonization, when fermions have flavor (e.g., spin) degrees of freedom, is that the flavor symmetries are hidden. This problem was solved by Witten's non-abelian bosonization [9]. In this paper we generalize Witten's non-abelian bosonization to (2+1) and (3+1) space-time dimensions.

The limitation of our theory is that it only applies to fermions with relativistic dispersion. (However, we do not restrict the Fermi velocity to be the speed of light.) In the absence of a mass gap, such theories have Dirac-like dispersion relation. In one space dimension, massless fermions are generically relativistic at low energies. In two and three space dimensions, relativistic massless fermions have been discovered in many experimental condensed matter systems. Examples include graphene and twisted bilayer graphene, Dirac and Weyl semi-metal,...etc. Moreover, relativistic massless fermions can appear in the mean-field theory of strongly correlated systems. Such theory serves as the starting point of a more rigorous treatment. For example, the "spinon π-flux phase" mean-field theory sets the stage for a gauge theory description of the Mott insulating state of cuprates.

Another important area where relativistic massless fermions appear is at the boundary of topological insulators or superconductors, which are simple examples of symmetry-protected topological (SPT) phases. The classification of topological insulator and superconductor[7, 16] can be viewed as asking how many copies of the massless fermion theories on the boundary are required to couple together before a symmetry-allowed mass term emerges.

The remaining of this part contains two chapters: Chap. 4 for the bosonization and Chap. 5 for the applications. Each of them contains several sections, namely, 14 sections in Chap. 4 and 3 sections in Chap. 5. In each section of 4, we discuss an important step or input of the bosonization. We shall illustrate the relevant concept with examples in the lowest spatial dimension where it first appears. For higher spatial dimensions, we simply present the result while leaving the details to the appendices. Together, the 14 sections in 4 provide the readers with the idea and technical details of the bosonization. In Chap. 5 there are 3 sections, each gives an example of how this bosonization can be applied. The topics include the $SU(2)$ gauge theory of the Mott insulating phase in the cuprates, the spin effective theory in “bipartite-Mott insulators” in spatial dimensions 1, 2, 3, and the twisted bilayer graphene. Finally, the 11 appendices provide the details omitted in the main text.

Chapter 4

Bosonization

4.1 The idea

In this paper, by "Bosonization", we mean to construct bosonic theories that are equivalent to theories of massless relativistic fermions. In the rest of the paper, unless otherwise stated, "massless fermion" always refers to massless relativistic fermion. Here we stress again that "relativistic massless fermion" does not imply the Fermi velocity is the speed of light. As mentioned in the introduction, in several $(2+1)$ and $(3+1)$ dimensional condensed matter systems, relativistic massless fermions have been encountered.

We look at the massless fermion theories from two points of view. On one hand, as d (spatial) dimensional theories, the massless fermion theories have emergent symmetries and symmetry anomalies. On the other hand, the massless fermions can be realized on the boundary of $d+1$ topological insulators/superconductors where the emergent symmetries are the protection symmetries.

If the emergent symmetries are to be respected, the (massless) fermions can not develop an energy gap. But what if we introduce mass terms (or the bosonic order parameters), at the expense of breaking the emergent symmetries, then fluctuate the order parameters *smoothly* (in both space and time) until the symmetries are restored ? Since the order parameter fluctuations are smooth, we expect the fermion gap to remain intact. Under such conditions, we can integrate out the fermions to yield bosonic non-linear sigma models governing the dynamics of the order parameters. From the perspective of the boundary of topological insulators/superconductors, after integrating out the fermions, what's left are fluctuating order parameters and the non-linear sigma models. Because the protection symmetries are restored by the order parameter fluctuations, the non-linear sigma models are either gapless or possess topological order. It turns out that the non-linear sigma models have a special type of topological term: the level-one Wess-Zumino-Witten (WZW) term. Such term encodes the 't Hooft anomaly[1] associated

[1]The 't Hooft anomaly refers to the obstruction in gauging the continuous part of the emergent global symmetries. Under such conditions, once gauge field is introduced, the partition function fails to

with the boundary of topological insulators/superconductors. Due to the WZW term, the non-linear sigma models are gapless, hence potentially can be equivalent to the massless fermion theories. This equivalence is supported by the fact that the fermion and boson theories have (1) the same symmetries, (2) the same anomalies, and (3) the boson theories have fermionic solitons.

As to the question of why do we bother to bosonize? One reason is it allows us to determine the low energy physics of a non-trivial boson theory by solving the theory of free massless fermions, and often what is subtle in one picture can become clearer in the other. Of course, we will not stop at the massless free fermion theories, the goal of bosonization is to enable one to go further. This will become clear in the applications.

4.2 Emergent symmetries of the massless fermion theory

A necessary condition for two theories to be equivalent is that they have the same symmetry. Thus it is important to determine the symmetry of massless fermion theories. It turns out the symmetries of such theories are rather rich. Because the massless fermion theories are low energy *effective* theories, we shall refer to their symmetries as the *emergent symmetries*.

In the following, we shall consider massless n-flavor Dirac (or Majorana) fermion theories in spatial dimensions 1, 2, and 3. Such theories can be split into two main categories, namely, complex class and real class. A theory in complex classes can be solely written in terms of *complex* Dirac fermion fields. Moreover, in the presence of a cutoff, its Hilbert space is the *eigenspace* of certain “charge” operator Q. In the following we shall focus on the $Q = 0$ eigenspace, i.e., the “charge neutral point” in condensed matter physics. The charge operator is the generator of a (continuous) global U(1) symmetry. In contrast, a theory in the real class is expressed in terms of Majorana fermion fields. In this class, there is no requirement for a conserved charge operator.

4.2.1 Complex class

Now, as an example, let’s determine the emergent symmetry group of a one dimensional massless fermion theory. To this end let’s first consider a complex class, n-flavor, massless Dirac fermion theory described by the following action

be gauge invariant.

$$S_0 = \int dx^0 dx^1 \psi^\dagger (\partial_0 - i\Gamma_1 \partial_1)\psi \quad \text{where} \tag{4.1}$$
$$\Gamma_1 = Z I_n$$

Here I_n denotes $n \times n$ identity matrix. In the following we shall use the shorthand I, X, Y, Z, E to denote the Pauli matrix $\sigma_{0,x,y,z}, i\sigma_y$, and when two matrix symbols stand next to each other, e.g., ZI_n, it means tensor product $Z \otimes I_n$. For complex fermion field ψ, the possible unitary transformations include

$$\psi \to U \cdot \psi$$
$$\psi \to C \cdot (\psi^\dagger)^T$$

where U and C are unitary matrices. Note that as a discrete transformation (the second line of the above equations), the charge conjugation transformation does leave the $Q = 0$ eigenspace invariant [2].

One can easily show that the full emergent symmetries of the action in Eq.(4.1) are

Chiral $U(n)$ symmetry:
$$U(n)_+ \times U(n)_- : \psi \to \Big(P_+ \otimes g_+ + P_- \otimes g_-\Big)\psi \text{ where } g_\pm \in U(n)$$
Charge conjugation symmetry:
$$C : \psi \to (Z \otimes I_n)\,(\psi^\dagger)^T$$
Time reversal symmetry (anti-unitary):
$$T : \psi \to (X \otimes I_n)\,\psi \tag{4.2}$$

Here

$$P_\pm := \frac{I \pm Z}{2} \tag{4.3}$$

are the projection operators with the subscript $\pm$ denoting the "right/left" moving fermions, respectively. Note that any other anti-unitary symmetry can be written in terms of the composition of a unitary symmetry and the time reversal transformation above.

4.2.2 Real class

[2] However, we do not allow the charge conjugation operator to generate *continuous* transformations, since under such transformations ψ will go into the superposition of ψ and $\psi^\dagger$. This violates the requirement that the Hilbert space is the eigenspace of the charge operator.

Next, we consider the one-dimensional massless theory in the real class. In this case, we write the action in terms of the n-component Majorana fermion field

$$
\begin{aligned}
S_0 &= \int dx^0 dx^1\, \chi^T \left[\partial_0 - i\Gamma_1 \partial_1\right] \chi \quad \text{where} \\
\Gamma_1 &:= Z I_n
\end{aligned}
\tag{4.4}
$$

For Majorana fermion field, the possible unitary transformations are of the form

$$\chi \to O \cdot \chi$$

where O is an orthogonal matrix. The full emergent symmetries of the action in Eq.(4.4) are

$$
\begin{aligned}
&\text{Chiral } O(n) \text{ symmetry:} \\
&O(n)_+ \times O(n)_- : \chi \to \Big(P_+ \otimes g_+ + P_- \otimes g_-\Big)\chi \text{ where } g_\pm \in O(n) \\
&\text{Time reversal symmetry (anti-unitary):} \\
&T : \chi \to (X \otimes I_n)\chi.
\end{aligned}
\tag{4.5}
$$

In $D = d + 1$ space-time dimension, the massless fermion actions are

$$
\begin{aligned}
\text{Complex class:}\ \ S_0 &= \int d^D x\, \psi^\dagger \left[\partial_0 - i\sum_{i=1}^{d} \Gamma_i \partial_i\right] \psi \\
\text{Real class:}\ \ S_0 &= \int d^D x\, \chi^T \left[\partial_0 - i\sum_{i=1}^{d} \Gamma_i \partial_i\right] \chi
\end{aligned}
\tag{4.6}
$$

where ψ and χ are complex and Majorana fermion fields, respectively. In table 4.1 we summarize the emergent symmetries of massless fermion theories in 1,2 and 3 dimensions. For details see appendix B.1.

4.3 Mass terms and mass manifolds

Mass terms, or order parameters, are fermion bilinears, namely,

$$
\begin{aligned}
&\psi^\dagger M \psi, \quad \text{or} \\
&\chi^T M \chi,
\end{aligned}
\tag{4.7}
$$

which opens an energy gap when added to Eq.(4.6). To achieve that, the *hermitian* mass matrix M must anti-commute with all the gamma matrices, i.e.,

$$\{M, \Gamma_i\} = 0 \text{ for } i = 1, ..., d \tag{4.8}$$

(1 + 1)-D	Real class	Complex class
Γ_i	$Z \otimes I_n$	$Z \otimes I_n$
Emergent symmetries	$T = X \otimes I_n$ $O_+(n) \times O_-(n) : P_+ \otimes g_+ + P_- \otimes g_-$ where $g_+ \in O_+(n)$ and $g_- \in O_-(n)$	$T = X \otimes I_n$ $C = Z \otimes I_n$ $U_+(n) \times U_-(n) : P_+ \otimes g_+ + P_- \otimes g_-$ where $g_+ \in U_+(n)$ and $g_- \in U_-(n)$
(2 + 1)-D	Real class	Complex class
Γ_i	$Z \otimes I_n$, $X \otimes I_n$	$Z \otimes I_n$, $X \otimes I_n$
Emergent symmetries	$T = E \otimes I_n$ $O(n) : I \otimes g$ where $g \in O(n)$	$T = Y \otimes I_n$ $C = I \otimes I_n$ $U(n) : I \otimes g$ where $g \in U(n)$
(3 + 1)-D	Real class	Complex class
Γ_i	$ZI \otimes I_n$, $XI \otimes I_n$, $YY \otimes I_n$	$ZI \otimes I_n$, $XI \otimes I_n$, $YZ \otimes I_n$
Emergent symmetries	$T = EZ \otimes I_n$ $U(n) : II \otimes g_1 - IE \otimes g_2$ where $u = g_1 + ig_2 \in U(n)$	$T = YZ \otimes I_n$ $C = IX \otimes I_n$ $U_+(n) \times U_-(n) : IP_+ \otimes g_+ + IP_- \otimes g_-$ where $g_+ \in U_+(n)$ and $g_- \in U_-(n)$

Table 4.1: A summary of the emergent symmetries of massless fermions in (1 + 1)-D, (2 + 1)-D, and (3 + 1)-D. Here $P_\pm := (I \pm Z)/2$ as in Eq.(4.3).

We will further require that the gap is flavor independent by imposing

$$M^2 = m^2 \cdot 1 \tag{4.9}$$

Here 1 means the identity matrix of appropriate size. The mass matrices satisfying Eq.(4.8) and Eq.(4.9) form a topological space – the *mass manifold*. In the simplest case, it can be a k-dimensional sphere. In general, it is a closed k-dimensional manifold. If, in addition to Eq.(4.8), the mass terms are required to be invariant under certain unitary or anti-unitary transformations, the mass manifold will be affected. In the classification of the free fermion SPTs, it is important to know what is the homotopy group of the mass manifold [7].

In the following we give two examples in one spatial dimension, to let the readers get a feeling of what's involved in figuring out the mass manifold.

4.3.1 Complex class

Let the U(1) symmetry transforms the field field according to

$$\psi \to e^{i\theta}\psi.$$

Then all mass terms in the form

$$\psi^\dagger M^{\mathbb{C}} \psi,$$

are invariant under U(1). Here the superscript $\mathbb{C}$ is to remind us that this is a mass matrix in the complex fermion class. $M^{\mathbb{C}}$ is an $2n \times 2n$ ($2n$ is the number of component of ψ) satisfying

$$\begin{aligned} M^{\mathbb{C}} &= \left(M^{\mathbb{C}}\right)^\dagger \\ \{M^{\mathbb{C}}, \Gamma_i\} &= 0 \\ \left(M^{\mathbb{C}}\right)^2 &= m^2 I_{2n} \end{aligned}$$

Here I_{2n} is the $2n \times 2n$ identity matrix. Associated with the massless fermion action given in Eq.(4.1), the first two conditions require $M^{\mathbb{C}}$ to be of the form

$$M^{\mathbb{C}} = m \left(X \otimes H_1 + Y \otimes H_2\right) \tag{4.10}$$

where H_1 and H_2 are $n \times n$ hermitian matrices. If we define

$$Q^{\mathbb{C}} := H_1 + i H_2, \tag{4.11}$$

it can be easily shown that the third condition requires

$$Q^{\mathbb{C}} \cdot \left(Q^{\mathbb{C}}\right)^\dagger = I_n.$$

Therefore the mass manifold for one dimension, in complex class, is the topological space formed by $n \times n$ unitary matrices.

4.3.2 Real class

In this case, the mass term is the Majorana fermion bilinear

$$\chi^T M^{\mathbb{R}} \chi$$

where the matrix $M^{\mathbb{R}}$ is an anti-symmetric matrix satisfying

$$\begin{aligned} M^{\mathbb{R}} &= \left(M^{\mathbb{R}}\right)^\dagger \\ \{M^{\mathbb{R}}, \Gamma_i\} &= 0 \\ \left(M^{\mathbb{R}}\right)^2 &= m^2 I_{2n} \end{aligned}$$

The first two conditions require

$$M^{\mathbb{R}} = m\left(Y \otimes S + X \otimes (iA)\right)$$

where S and A are real symmetric and anti-symmetric matrix, respectively. If we define

$$Q^{\mathbb{R}} := S + A$$

the last condition requires

$$Q^{\mathbb{R}} \cdot \left(Q^{\mathbb{R}}\right)^T = I_n.$$

Thus, the mass manifold is the space of $n \times n$ orthogonal matrices.

In table 4.2 we summarize the mass manifolds for 1,2 and 3 dimensions. The detailed derivations are left in appendix B.2.

4.4 The symmetry anomalies of the fermionic theories

Emergent symmetries of a low-energy effective theory can be broken when a cutoff is imposed. In this section, we review the symmetry anomalies of the massless fermion theories.

4.4.1 The 't Hooft anomaly of continuous symmetry

$(1+1)$-D	Real class	Complex class
Γ_i	$Z \otimes I_n$	$Z \otimes I_n$
Mass manifold	$M = Y \otimes S + X \otimes (iA)$ where $Q^{\mathbb{R}} = S + A \in O(n)$	$M = X \otimes H_1 + Y \otimes H_2$ where $Q^{\mathbb{C}} = H_1 + iH_2 \in U(n)$
$(2+1)$-D	Real class	Complex class
Γ_i	$Z \otimes I_n$, $X \otimes I_n$	$Z \otimes I_n$, $X \otimes I_n$
Mass manifold	$M = Y \otimes S$ where $Q^{\mathbb{R}} = S \in \bigcup_{l=0}^{n} \frac{O(n)}{O(l)\times O(n-l)}$	$M = Y \otimes H$ where $Q^{\mathbb{C}} = H \in \bigcup_{l=0}^{n} \frac{U(n)}{U(l)\times U(n-l)}$
$(3+1)$-D	Real class	Complex class
Γ_i	$ZI \otimes I_n$, $XI \otimes I_n$, $YY \otimes I_n$	$ZI \otimes I_n$, $XI \otimes I_n$, $YZ \otimes I_n$
Mass manifold	$M = YX \otimes S_1 + YZ \otimes S_2$ where $Q^{\mathbb{R}} = S_1 + iS_2 \in \frac{U(n)}{O(n)}$	$M = YX \otimes H_1 + YY \otimes H_2$ where $Q^{\mathbb{C}} = H_1 + iH_2 \in U(n)$

Table 4.2: A summary of the mass manifolds for the real and complex class fermions in $(1+1)$-D, $(2+1)$-D, and $(3+1)$-D.

The emergent symmetries discussed in the section 4.2 can suffer the ‘t Hooft anomaly”. A theory is said to have the ‘t Hooft anomaly with respect to global symmetry group G if there are obstructions against gauging G [17]. In the following we shall use the $(1+1)$-D complex class to illustrate the ideas.

The simplest example is the chiral anomaly associated with the $(1 + 1)$-D complex class theory defined in Eq.(4.6). This theory has emergent global $U_+(n) \times U_-(n)$ symmetry. However, when one tries to gauge this symmetry, an anomaly is encountered. Namely, in the presence of gauge field with non-zero curvature, the theory can not be made to conserve the Noether's current associated with the *full* $U_+(n)\times U_-(n)$ symmetry.

Starting from the massless fermion theory, we can introduce the $U_+(n)\times U_-(n)$ gauge field (i.e., “gauging” $U_+(n) \times U_-(n)$) via minimal coupling. Moreover, we can define the effective gauge action after integrating out fermions,

$$
\begin{aligned}
W[A_+, A_-] &= -\ln\left[\int D\psi\, D\bar{\psi} e^{-S[\psi,\bar{\psi},A_+,A_-]}\right], \quad \text{where} \\
S[\psi, \bar{\psi}, A_+, A_-] &= \int d^2x\ \ \bar{\psi}\ \left[i\gamma^\mu \left(\partial_\mu + iP_+ \otimes A_{+,\mu} + iP_- \otimes A_{-,\mu}\right)\right]\psi. \qquad (4.12)
\end{aligned}
$$

Here $A_\pm$ are the $n \times n$ matrix value gauge fields associated with $U_\pm(n)$, and $P_\pm$ are the projection operators selecting the chiral fermion modes defined in Eq.(4.3). Adler[2], Bell, and Jackiw[3] first showed that in the presence of a diagonal (i.e., $A_+ = A_-$) $U(1)$ gauge field, the axial current is not conserved. Shortly after, this was generalized by Bardeen [4] who showed that under infinitesimal gauge transformation, W in Eq.(4.12) is not gauge invariant, namely,

$$
\begin{aligned}
\delta W &:= W[A_+ + d\epsilon_+, A_- + d\epsilon_-] - W[A_+, A_-] \\
&= -\frac{i}{4\pi}\int_{\mathcal{M}} \operatorname{tr}\left[A_+ d\epsilon_+ - A_- d\epsilon_-\right]. \qquad (4.13)
\end{aligned}
$$

This is the ‘t Hooft anomaly.

This phenomenon is also connected to the physics of SPT. In odd space dimension, this connection constitutes the so-called “anomaly in-flow picture” [6]. The most familiar example is for $n = 1$ in 1D. In this case, we can view the 1D (non-chiral) massless fermions as the edge modes of two Chern insulators stacked together, with each Chern insulator having Hall conductivity $\sigma_{xy} = \pm 1$ (see Fig. 4.1). In the presence of a time-dependent flux associated with the diagonal gauge field, there will be the electric fields in the azimuthal direction. This induces a Hall current causing the charge to flow from the outer to the inner edge on one layer, and from the inner to the outer edge on the other layer. Viewing from the edge (one-dimensional world), the chiral current $J_+ - J_-$ is not conserved. This manifests the chiral anomaly, namely gauging the diagonal U(1) symmetry breaks axial U(1) symmetry - an example of the ‘t Hooft anomaly.

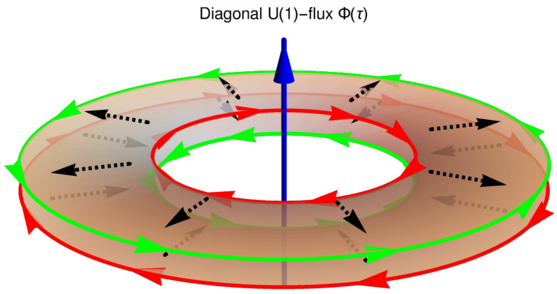


Figure 4.1: Two layers of annulus shape Chern insulators with $\sigma_{xy} = \pm 1$ stacked together. The outer edge harbors the 1D $n = 1$ non-chiral massless fermion modes. The green and red arrows represent the opposite chiralities. When a time-dependent diagonal $U(1)$ flux pierces the inner hole, the induced electric field in the azimuthal direction causes a Hall current (dashed arrows) flowing from inner to outer boundary in the top layer and from outer to inner boundary in the bottom layer. As the result, the chiral current $J_+ - J_-$ is not conserved viewed from the outer edge alone. This system is realized as the "spin Hall insulator" experimentally.

Although the $U_+(n) \times U_-(n)$ anomaly makes it impossible to gauge the whole group consistently, it's possible to gauge a subgroup of it. For example, if we only gauge the diagonal subgroup $U(n)$ within $U_+(n) \times U_-(n)$, i.e., if

$$A := A_+ = A_-$$
$$\epsilon := \epsilon_+ = \epsilon_-$$

then the two terms in Eq.(4.13) cancel out, hence the theory is anomaly free with respect to diagonal $U(n)$ subgroup.

4.4.2 A heuristic way to determine the 't Hooft anomaly

The discussions presented above require rather involved field theory calculations. However, there is a heuristic way to get the correct answer. The basis of this heuristic argument is the fact that *if a theory can be defined on a lattice with all its (continuous) symmetry, then these symmetries can be gauged without anomaly.* Again, the above statement is suggested by the SPT physics, namely, the boundary modes (which has 't Hooft anomaly) of an SPT can not be defined on a lattice in the dimension of the boundary. In the following we shall again use the $(1 + 1)$-D complex class to illustrate the ideas.

Under Wilson's regularization[18](see later), whether a theory with global symmetry group G can be defined on a lattice, is determined by whether there is a mass term that respects G [3]. Thus, a theory with the $U_+(n) \times U_-(n)$ anomaly, means no mass term is $U_+(n) \times U_-(n)$ symmetric. Again, this is the condition that the gaplessness of the boundary modes is symmetry protected.

First, we show that no mass term is allowed if $U_+(n) \times U_-(n)$ symmetry is to be respected. Under $U_+(n) \times U_-(n)$ the fermion field transform as

$$\psi \to (P_+ \otimes g_+ + P_- \otimes g_-)\,\psi \quad \text{where} \quad P_\pm = \frac{I \pm Z}{2}.$$

Under such transformation, there is, e.g., no mass term preserving the axial $U_A(1)$ generated by ZI_n. This is because according to table 4.2 the mass terms have the form

$$\psi^\dagger\,(X \otimes H_1 + Y \otimes H_2)\,\psi.$$

In fact, the anomaly is not only in the axial $U_A(1)$ part. To see that, let's consider $n > 1$. The diagonal $U(n)$ symmetry requires that both H_1 and H_2 be proportional to the identity matrix. However, such mass term would break $U_+(n)$.

Now we show that if we relax the condition to only demanding the diagonal $U(n)$ symmetry, there is a mass term. For example,

$$M_{\text{reg}} = X \otimes I_n.$$

This means that we can then write down a lattice model in momentum space using Wilson's regularization[18]

$$\hat{\mathrm{H}} = \sum_{k\in \mathrm{BZ}} \psi_k^\dagger\,[\sin k\,\Gamma_1 + (1-\cos k)M_{\text{reg}}]\,\psi_k$$

where "BZ" stands for the Brillouin zone. We can Fourier transform the above hamiltonian back to the real space which gives us a lattice tight-binding model. The diagonal $U(n)$ gauge field can then be introduced via Peierls' substitution

$$\psi_j^\dagger \psi_i \to \psi_j^\dagger e^{iA_{i,j}}\psi_i$$

for two adjacent sites i, j. Here $A_{i,j}$ is the gauge connection from site i to j.

4.4.3 Discrete global symmetry anomaly

A (global) discrete symmetry in a fermion theory can also be broken by regularization. In this subsection, we shall review the simplest example – the "parity anomaly"[20, 21]

[3] Using Wilson's regularization method [18], the existence of such a mass term is a sufficient condition for the theory to be regularizable on a lattice. However, it is more involved to show that it is the necessary condition [19].

of the $(2+1)$-D Dirac fermions in the complex class.

When the anomaly-free $U(n)$ symmetry is gauged, the low energy fermion action is given by

$$S = \int d\tau\, d^2\boldsymbol{x}\, \psi^\dagger \left[(\partial_0 + i\, I \otimes A_0) - i\Gamma_i \left(\partial_i + i\, I \otimes A_i\right)\right] \psi \qquad (4.14)$$
$$\text{where } \Gamma_1 = Z I_n, \quad \Gamma_2 = X I_n$$

Here A_μ is the $n \times n$ matrix-valued $U(n)$ gauge field. Under the global emergent symmetries listed in table 4.1, the gauged field transforms as

$$\begin{aligned} &U(n)\text{: } A_\mu \to g \cdot A_\mu \cdot g^\dagger \\ &\text{Time reversal: } A_\mu \to -\left(A_\mu\right)^* \\ &\text{Charge conjugation: } A_\mu \to -\left(A_\mu\right)^T \end{aligned} \qquad (4.15)$$

It's easy to check that the low energy action Eq.(4.14) is invariant under the combined transformation of the fermion and the gauge field.

As we saw in the preceding subsection, the condition for a symmetry to be anomaly-free is the theory can be regularized while preserving the symmetry. In the present case, to preserve $U(n)$ we need to choose a regularization that is $U(n)$ invariant. In Wilson's regularization[18] this amounts to choose a $U(n)$ invariant regularization mass. The most general mass term is given by

$$M = m\, Y \otimes H,$$

where H is an $n \times n$ hermitian matrix with $H^2 = I_n$. When acted upon by the global $U(n)$,

$$M \to (I \times g)^\dagger \cdot M \cdot (I \times g)$$

(see table 4.1). Requiring it to be invariant forces us to choose

$$M_{\text{reg}} = m\, Y \otimes I_n. \qquad (4.16)$$

Under Wilson's regularization the momentum space Hamiltonian of the massless Dirac fermion (without gauge field) read,

$$\hat{\mathrm{H}} = \sum_{\boldsymbol{k} \in BZ} \psi_{\boldsymbol{k}}^\dagger \left[\sin k_1 \Gamma_1 + \sin k_2 \Gamma_2 + (2 - \cos k_1 - \cos k_2)\, M_{\text{reg}} \right] \psi_{\boldsymbol{k}} \qquad (4.17)$$

To incorporate the gauge field, we Fourier transform the above equation back to real space and introduce the gauge field by Peierls' substitution. This is all good as far as regularizing Eq.(4.14) is concerned.

Under the action of the discrete symmetries, however

$$\begin{aligned} \text{Charge conjugation: } M_{\text{reg}} &\to -\left(I \otimes I_n\right) \cdot M_{\text{reg}}^T \cdot \left(I \otimes I_n\right) = M_{\text{reg}} \\ \text{Time reversal: } M_{\text{reg}} &\to \left(Y \otimes I_n\right) \cdot M_{\text{reg}}^* \cdot \left(Y \otimes I_n\right) = -M_{\text{reg}} \end{aligned}$$

Therefore charge conjugation is respected by the regularization, however, time-reversal symmetry is not.

It was first shown by Redlich [20, 21] that one can detect the time-reversal anomaly through the effective $U(n)$ gauge action after integrating out the fermions. We reproduce his argument in the following. In momentum space (the Brillouin zone) we have four low energy Dirac fermions, each around a time-reversal invariant $\boldsymbol{k}$ points:

$$\begin{aligned}
\boldsymbol{k} = (0,0) + \boldsymbol{q}: \quad & \hat{\mathrm{H}}_{(0,0)} \approx \sum_{\text{small } \boldsymbol{q}} \psi^\dagger_{(0,0)+\boldsymbol{q}} \left[q_1\, \Gamma_1 + q_2\, \Gamma_2\right] \psi_{(0,0)+\boldsymbol{q}} \\
\boldsymbol{k} = (\pi,0) + \boldsymbol{q}: \quad & \hat{\mathrm{H}}_{(\pi,0)} \approx \sum_{\text{small } \boldsymbol{q}} \psi^\dagger_{(\pi,0)+\boldsymbol{q}} \left[-q_1\, \Gamma_1 + q_2\, \Gamma_2 + 2m\, M_{\text{reg}}\right] \psi_{(\pi,0)+\boldsymbol{q}} \\
\boldsymbol{k} = (0,\pi) + \boldsymbol{q}: \quad & \hat{\mathrm{H}}_{(0,\pi)} \approx \sum_{\text{small } \boldsymbol{q}} \psi^\dagger_{(0,\pi)+\boldsymbol{q}} \left[q_1\, \Gamma_1 - q_2\, \Gamma_2 + 2m\, M_{\text{reg}}\right] \psi_{(0,\pi)+\boldsymbol{q}} \\
\boldsymbol{k} = (\pi,\pi) + \boldsymbol{q}: \quad & \hat{\mathrm{H}}_{(\pi,\pi)} \approx \sum_{\text{small } \boldsymbol{q}} \psi^\dagger_{(\pi,\pi)+\boldsymbol{q}} \left[-q_1\, \Gamma_1 - q_2\, \Gamma_2 + 4m\, M_{\text{reg}}\right] \psi_{(\pi,\pi)+\boldsymbol{q}}
\end{aligned} \tag{4.18}$$

Among the four, the first is massless and preserves the time-reversal symmetry. The remaining three, however, acquire a large regularization mass, which is time-reversal breaking. In the presence of the $U(n)$ gauge field, these massive Dirac fermions would each contribute a Chern-Simons effective gauge action after the fermions are integrated out[21]. In particular, for each massive fermion the effective gauge action is

$$\frac{1}{2} \times (\pm 1) \times \frac{i}{4\pi} \int AdA,$$

where the sign depends on the product of the signs in front of $q_1\Gamma_1$, $q_2\Gamma_2$, and M_{reg}. Combing them, the massive fermions contribute the following breaking effective action

$$\left(-\frac{1}{2} - \frac{1}{2} + \frac{1}{2}\right) \frac{m}{|m|} \frac{i}{4\pi} \int \mathrm{tr}\left[A\, dA + \frac{2i}{3} A^3\right] = -\frac{i}{8\pi} \int \mathrm{tr}\left[A\, dA + \frac{2i}{3} A^3\right]. \tag{4.19}$$

This is time-reversal odd, as can be explicitly shown by replacing $A_\mu \to -(A_\mu)^*$ and complex conjugating the action. As to the massless fermions near $\boldsymbol{k} = (0,0)$, based on the fact that the first line of Eq.(4.18) is time reversal invariant so should their effective gauge action. Thus after regularization, the time-reversal symmetry of Eq.(4.14) is broken! As expected, charge conjugation is not broken by the regularization. Since T is broken while C is not, based on the CPT invariance, the parity should also be broken [4].

[4] In two space dimension, the "parity" transformation P is realized by spatial reflection. Take the reflection in x-direction as an example, the fermion field transforms according to

$$\psi(\tau, x, y) \xrightarrow{P} XI_n \cdot \psi(\tau, -x, y).$$

It is easy to see the that the regularization mass defined in Eq.(4.16) changes sign under P. However, the combined CPT transformation leaves it invariant. Thus, there is no CPT anomaly. The same conclusion can be drawn by looking at the parity transformation of the effective gauge action. Under P the gauge

In table 4.3, we summarize the maximal anomaly-free continuous symmetry and the discrete symmetry that is broken after regularization. The only discrete symmetry which possesses anomaly occurs in $(2+1)$-D for the time-reversal symmetry [5]. More detailed discussions are left to appendix B.3.

4.5 Breaking the emergent symmetry by the mass terms

The mass terms discussed in the last section *necessarily* break some of the emergent symmetries in table 4.1. This is because so long as the full emergent symmetries remain unbroken, the fermions will remain massless. In the rest of this section, we use one-dimensional examples to illustrate this.

4.5.1 Complex class

The mass terms for the complex class in $(1+1)$-D can be written as

$$\psi^\dagger \left(X \otimes H_1 + Y \otimes H_2\right) \psi = \psi^\dagger \begin{bmatrix} 0 & \left(Q^{\mathbb{C}}\right)^\dagger \\ Q^{\mathbb{C}} & 0 \end{bmatrix} \psi.$$

When acted upon by the emergent symmetries in Eq.(4.2), $Q^{\mathbb{C}}$ transforms as

$$\begin{aligned} &U_+(n) \times U_-(n) : Q^{\mathbb{C}} \to g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+ \\ &\text{Charge conjugation} : Q^{\mathbb{C}} \to \left(Q^{\mathbb{C}}\right)^* \\ &\text{Time reversal} : Q^{\mathbb{C}} \to \left(Q^{\mathbb{C}}\right)^T . \end{aligned} \tag{4.20}$$

field transforms as

$$\begin{aligned} A_\tau(\tau,x,y) &\xrightarrow{P} A_\tau(\tau,-x,y) \\ A_x(\tau,x,y) &\xrightarrow{P} -A_x(\tau,-x,y) \\ A_y(\tau,x,y) &\xrightarrow{P} A_y(\tau,-x,y) \end{aligned}$$

Again, Eq.(4.19) changes sign under P, but is invariant under CPT.

[5] Note that we have made a particular choice for the anti-unitary (time reversal) and charge conjugation generators in table 4.1. This choice is not unique because a new choice can be obtained by compounding the C and the T we used with other unitary symmetries. For example, in the case of complex class in $(1+1)$-D, one can compound the T and C with a chiral $U_+(n) \times U_-(n)$ transformation. Had we done so, these discrete symmetries would also be anomalous. In $(1+1)$-D and $(3+1)$-D we specifically choose the generators of C and T to be anomaly-free after the maximal anomaly-free part of the continuous symmetry is gauged. On the other hand, it can be shown (see appendixB.3) that in $(2+1)$-D, there is no choice of T which will not be broken by a regularization that respects all continuous symmetries.

(1 + 1)-D	Real class	Complex class
Global Symmetry	**Discrete** Anti-unitary: $T^2 = +1$ **Continuous unitary** Chiral $O(n) \times O(n)$	**Discrete** Anti-unitary: $T^2 = +1$ Unitary: $C^2 = +1$ **Continuous unitary** Chiral $U(n) \times U(n)$
Aanomaly free part	Diagonal $O(n)$, T	Diagonal $U(n)$, T, C
(2 + 1)-D	Real class	Complex class
Global Symmetry	**Discrete** Anti-unitary: $T^2 = -1$ **Continuous unitary** $O(n)$	**Discrete** Anti-unitary: $T^2 = -1$ Unitary: $C^2 = +1$ **Continuous unitary** $U(n)$
Anomaly free part	$O(n)$	$U(n)$, C
(3 + 1)-D	Real class	Complex class
Global Symmetry	**Discrete** Anti-unitary: $T^2 = -1$ **Continuous unitary** $U(n)$	**Discrete** Anti-unitary: $T^2 = -1$ Unitary: $C^2 = +1$ **Continuous unitary** Chiral $U(n) \times U(n)$
Anomaly free part	$O(n)$, T	Diagonal $U(n)$, T, C

Table 4.3: The summary of the global symmetry groups and the anomaly-free parts of the symmetry groups of the massless fermions (and the bosonized non-linear sigma models) in (1 + 1)-D,(2 + 1)-D, and (3 + 1)-D.

Thus a space-time constant $Q^{\mathbb{C}}$ breaks the emergent symmetry because both g_+ and g_- can be arbitrary unitary matrices.

4.5.2 Real class

For the real class in (1 + 1)-D, the mass term can be written as

$$\chi^\dagger \left[Y \otimes S + X \otimes (iA)\right] \chi = \chi^T \begin{bmatrix} 0 & -i\left(Q^{\mathbb{R}}\right)^T \\ iQ^{\mathbb{R}} & 0 \end{bmatrix} \chi. \tag{4.21}$$

When the emergent symmetries in Eq.(4.5) acts on it $Q^{\mathbb{R}}$ transforms as

$$\begin{aligned} O_+(n) \times O_-(n) &: Q^{\mathbb{R}} \to g_-^T \cdot Q^{\mathbb{R}} \cdot g_+ \\ \text{Time reversal} &: Q^{\mathbb{R}} \to \left(Q^{\mathbb{R}}\right)^T. \end{aligned}$$

Therefore a space-time non-zero $Q^{\mathbb{R}}$ breaks the emergent symmetry because both g_+ and g_- can be arbitrary orthogonal matrices.

4.6 Restoring the emergent symmetries

So far we have seen that space-time constant $Q^{\mathbb{C}}$ or $Q^{\mathbb{R}}$ breaks the emergent symmetry. But what if $Q^{\mathbb{C}}$ and $Q^{\mathbb{R}}$ fluctuates in space-time? As in statistical mechanics, when the order parameters fluctuate, the broken symmetry can be restored. Likewise, if we fluctuate $Q^{\mathbb{C}}$ and $Q^{\mathbb{R}}$ over the appropriate mass manifold we expect the emergent symmetry to be restored.

Our approach is conceptually similar to that in Ref.[22, 23] where, on the surface of the topological insulator, the fluctuating superconducting order parameters restore the symmetries of the massless fermions. The important difference is that the required order parameter fluctuation in Ref.[22, 23] is not smooth, because it involves the proliferation of superconducting vortices. Since the structure of vortex cores is important in that approach, and such structure depends on the short-distance physics, this approach is constrained to the surface of SPTs where regularization is not an issue. In contrast, our goal is to bosonize the low energy effective theory, where the emergent symmetry is necessarily broken at short distances (due to anomaly). As the result, we restrict our order parameter to be smooth in space and time, so that they act on the low energy theory only.

But what does "appropriate mass manifold" mean? For complex class in $(1+1)$-D, $Q^{\mathbb{C}}$ needs to fluctuate over the space formed by $n \times n$ unitary matrices, or $U(n)$. Such a space is connected and has a single component. On the other hand for the real class in 1D, $Q^{\mathbb{R}}$ needs to fluctuate in the space formed by $n \times n$ orthogonal matrices, or $O(n)$. This space has two disconnected components, corresponding to $\det[Q^{\mathbb{R}}] = \pm 1$. It's only when $Q^{\mathbb{R}}$ fluctuates in both components with the equal statistical weight we can restore the emergent symmetry.

In (3+1)-D the mass manifold consists of a single component, in which $Q^{\mathbb{C},\mathbb{R}}$ fluctuate. However, in (2+1)-D the mass manifold in complex class is $\cup_{l=0}^{n} \frac{U(n)}{U(l)\times U(n-l)}$ which contains $n+1$ disconnected components. Here $Q^{\mathbb{C}}$ needs to fluctuate in the component $l = n/2$ in order to restore the time reversal symmetry[6] In real class, the mass manifold in two space dimension is $\cup_{l=0}^{n} \frac{O(n)}{O(l)\times O(n-l)}$, and $Q^{\mathbb{R}}$ needs to fluctuate in the $l = n/2$ component in order to restore the time reversal symmetry. We summarize the results for higher dimensions in table 4.4 and leave the detail in appendix B.2.

[6] Of course this requires n to be even.

(1 + 1)-D	Real class	Complex class
Symmetry transformations of $Q^{\mathbb{C},\mathbb{R}}$	$T: Q^{\mathbb{R}} \to (Q^{\mathbb{R}})^T$ $O_+(n) \times O_-(n)$: $Q^{\mathbb{R}} \to g_-^T \cdot Q^{\mathbb{R}} \cdot g_+$	$T: Q^{\mathbb{C}} \to (Q^{\mathbb{C}})^T$ $C: Q^{\mathbb{C}} \to (Q^{\mathbb{C}})^*$ $U_+(n) \times U_-(n)$: $Q^{\mathbb{C}} \to g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+$
The mass manifold required to restore the full emergent symmetries	$O(n)$	$U(n)$
(2 + 1)-D	Real class	Complex class
Symmetry transformations of $Q^{\mathbb{C},\mathbb{R}}$	$T: Q^{\mathbb{R}} \to -Q^{\mathbb{R}}$ $O(n): Q^{\mathbb{R}} \to g^T \cdot Q^{\mathbb{R}} \cdot g$	$T: Q^{\mathbb{C}} \to -(Q^{\mathbb{C}})^*$ $C: Q^{\mathbb{C}} \to (Q^{\mathbb{C}})^T$ $U(n): Q^{\mathbb{C}} \to g^\dagger \cdot Q^{\mathbb{C}} \cdot g$
The mass manifold required to restore the full emergent symmetries	$\frac{O(n)}{O(n/2)\times O(n/2)}$ for $n \in$ even	$\frac{U(n)}{U(n/2)\times U(n/2)}$ for $n \in$ even
(3 + 1)-D	Real class	Complex class
Symmetry transformations of $Q^{\mathbb{C},\mathbb{R}}$	$T: Q^{\mathbb{R}} \to (Q^{\mathbb{R}})^*$ $U(n): Q^{\mathbb{R}} \to u^T \cdot Q^{\mathbb{R}} \cdot u$	$T: Q^{\mathbb{C}} \to (Q^{\mathbb{C}})^*$ $C: Q^{\mathbb{C}} \to (Q^{\mathbb{C}})^T$ $U_+(n) \times U_-(n)$: $Q^{\mathbb{C}} \to g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+$
The mass manifold required to restore the full emergent symmetries	$\frac{U(n)}{O(n)}$	$U(n)$

Table 4.4: The summary of the symmetry transformations of $Q^{\mathbb{R},\mathbb{C}}$, and the mass manifolds in which the $Q^{\mathbb{R},\mathbb{C}}$ fluctuations can restore the full emergent symmetries.

4.7 The conditions for the effective theory being bosonic

In order to achieve bosonization, the fermions in Eq.(4.1) and Eq.(4.4) must not appear in the low energy theory. To ensure that, we need to impose some conditions on the space-time dependence of $Q^{\mathbb{C}}$ and $Q^{\mathbb{R}}$. Namely, as functions of $\boldsymbol{x}$ and τ, $Q^{\mathbb{C}}(\tau, \boldsymbol{x})$ and $Q^{\mathbb{R}}(\tau, \boldsymbol{x})$ needs to fluctuate smoothly (comparing with the length and time scale set by m). Under such conditions, the original fermions can be integrated out, yielding a non-linear sigma model for the order parameters. The idea is similar to that encountered in magnetism, where electrons form local moments. After integrating out the electrons we arrive at an effective theory – a non-linear sigma model describing the fluctuations of the local moments in space and time.

4.8 Fermion integration

In this section, using $(1+1)$-D as an example, we shall describe how to integrate out the fermions. In higher spatial dimensions we shall present the results while leaving the details in appendix B.4.

4.8.1 Complex class

The fermion action with a space-time dependent mass term reads

$$S = \int d\tau\, d\boldsymbol{x}\, \psi^\dagger \left[\partial_0 - i\Gamma_1\partial_1 + m\hat{M}(\tau,\boldsymbol{x})\right]\psi \tag{4.22}$$

where and $\Gamma_1 = ZI_n$, and

$$\{\Gamma_1, \hat{M}(t,\boldsymbol{x})\} = 0, \quad \text{and} \quad \hat{M}(\tau,\boldsymbol{x})^2 = I_{2n}. \tag{4.23}$$

The $\hat{M}(\tau,\boldsymbol{x})$ that satisfies Eq.(4.23) is given by

$$\hat{M}(\tau,\boldsymbol{x}) = m\left[X \otimes H_1(\tau,\boldsymbol{x}) + Y \otimes H_2(\tau,\boldsymbol{x})\right],$$

For smooth order parameter configurations $\hat{M}(\tau,\boldsymbol{x})$, the fermion integration can be done via gradient expansion (See [24] for example. We shall convert the action to a Lorentz invariant form and present the general formalism applicable for all spatial dimensions in appendix B.4). The resulting effective action consists of two types of terms: the non-topological and topological terms. For the non-topological term (the stiffness term) we shall keep the one with the smallest number of space-time derivatives (they are the most relevant in the renormalization group sense). The topological term is dimensionless. In $(1+1)$-D, explicit fermion integration yields (see appendix B.4 for details)

$$W[Q^{\mathbb{C}}] = \frac{1}{8\pi}\int_{\mathcal{M}} d^2x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right] - \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^3\Big], \tag{4.24}$$

where $Q^{\mathbb{C}}$ is given in Eq.(4.10) and Eq.(4.11). The first term in Eq.(4.24) is the stiffness term and the second is the Wess-Zumino-Witten (WZW) topological term. Eq.(4.24) reproduces the level-1 $U(n)$ (abbreviated as $U(n)_1$) WZW model in Witten's non-abelian bosonization [9]. Note that the symbol "tr" means tracing over the $n \times n$ portion of the matrix. (In doing fermion integration, we have already traced out the matrix part involving γ^μ's). In Eq.(4.24) $\mathcal{M}$ is the space-time manifold, and $\mathcal{B}$ is the extension of the space-time manifold $\mathcal{M}$ so that

$$\partial\mathcal{B} = \mathcal{M}.$$

In addition, $\tilde{Q}^{\mathbb{C}}(u,x)$ is an extension field of $Q^{\mathbb{C}}(x)$ so that

$$\tilde{Q}^{\mathbb{C}}(u=1,x) = Q^{\mathbb{C}}(x) \text{ and}$$
$$\tilde{Q}^{\mathbb{C}}(u=0,x) = \text{constant}$$

In the equation above, "constant" means a space-time independent matrix.

For simplicity we shall focus on the space-time manifold $\mathcal{M} = S^D$ so that $\mathcal{B}$ is a $D+1$-dimensional disk. The reason for this choice is to ensure the extension $\tilde{Q}^{\mathbb{C}}(u,x)$ exists. Because we require a smooth evolution from $Q^{\mathbb{C}}(u=0,x)$ to $\tilde{Q}^{\mathbb{C}}(u=1,x)$ (x denotes $(\tau, \boldsymbol{x})$), it means the mapping

$$Q^{\mathbb{C}} : (u=1,x) \to \text{mass manifold}$$

is homotopically equivalent to the mapping

$$Q^{\mathbb{C}} : (u=0,x) \to \text{mass manifold.}$$

Since $\tilde{Q}^{\mathbb{C}}(u=0,x) = \text{constant}$ is homotopically trivial, a necessary condition for the smooth extension to exist is

$$\pi_D(\text{mass manifold}) = 0,$$

i.e., all smooth mappings from the space-time manifold to the mass manifold are homotopically trivial. It turns out this condition is met for sufficiently large n in all spatial dimensions. We shall return to this point in appendix B.2, B.4, and B.8. For $(1+1)$-D, $\pi_2(U(n)) = 0$ for any n.

For the WZW term to be well defined, it had better not depend on the extension. When there are two different extensions on the $D+1$ dimensional disk, say one defined by $\tilde{Q}^{\mathbb{C}}_1$ on $\mathcal{B}_1$ and the other by $\tilde{Q}^{\mathbb{C}}_2$ on $\mathcal{B}_2$, the difference in the WZW term associated with these two extensions is given by

$$\Delta W_{WZW}[\tilde{Q}^{\mathbb{C}}] = -\frac{2\pi i}{24\pi^2} \int\limits_{\mathcal{B}_1 \cup (-\mathcal{B}_2)} \text{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}} \Big)^3 \Big] \tag{4.25}$$

where $-\mathcal{B}_2$ is the mirror reflection of $\mathcal{B}_2$. Since $\mathcal{B}_1 \cup (-\mathcal{B}_2) = S^{D+1}$, removing the factor $2\pi i$, Eq.(4.25) is the topological invariant associated with $\pi_{2+1}(\text{mass manifold})$. It turns out that for all relevant cases, $\pi_{D+1}(\text{mass manifold}) = \mathbb{Z}$ (see appendix B.2). In $(1+1)$-D, $\pi_3(U(n)) = \mathbb{Z}$ for $n \geq 2$ ($n=1$ corresponds to flavorless or spinless fermion where the bosonization is abelian.). The coefficient of the WZW term renders $\Delta W_{WZW} = 2\pi i \times \text{integer}$. The fact that the WZW term is $2\pi i$ times the topological invariant implies the level (k) is 1. After the exponentiation, the phase factor associated with the WZW term is well-defined.

4.8.2 Real class

The 1+1-D Majorana fermion action with a space-time dependent mass read

$$S = \int d\tau\, d\boldsymbol{x}\, \chi^T \left[\partial_0 - i\Gamma_1\partial_1 + m\hat{M}(\tau, \boldsymbol{x})\right] \chi \tag{4.26}$$

where

$$\Gamma_1 = ZI_n \text{ and } \hat{M}(\tau, \boldsymbol{x}) = [Y \otimes S + X \otimes (iA)].$$

Following the same steps discussed in the last subsection, fermion integration yields the following effective action (see appendix B.4)

$$W[Q^{\mathbb{R}}] = \frac{1}{16\pi}\int_{\mathcal{M}} d^2x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}T}\partial^\mu Q^{\mathbb{R}}\right] - \frac{2\pi i}{48\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{R}T} d\tilde{Q}^{\mathbb{R}}\right)^3\Big]. \tag{4.27}$$

Eq.(4.27) is the $O(n)_{k=1}$ WZW model. Again, $\tilde{Q}^{\mathbb{R}}(u,x)$ is extension field of $Q^{\mathbb{R}}(x)$, which exists if π_D(mass manifold) $= 0$. In $(1+1)$-D, $\pi_2(O(n)) = 0$ for $n \geq 3$. Here the difference in the WZW term associated with two different extension is the topological invariant associated with $\pi_3(O(n)) = \mathbb{Z}$ for relevant n (see appendix B.2). The coefficient of the WZW term renders $\Delta W_{WZW} = 2\pi i \times$ integer hence yields the same phase factor upon exponentiation. Again, the fact that the WZW term is $2\pi i$ times the topological invariant implies the level (k) is 1.

Thus, for both complex and real classes, the bosonization of massless fermion is the non-linear sigma model with WZW term. This reproduces Witten's non-abelian bosonization results, which was obtained using a totally different method (the current algebra).

The above bosonization scheme can be straightforwardly generalized to higher dimensions. One thing that needs some care is the fact that the homotopy group of the mass manifold depends on n. For n exceeds certain value π_{D+1}(mass manifold) $= \mathbb{Z}$. In that case fermion integration does lead to a nonlinear sigma model with $k = 1$ WZW term. However, for small n (before the "homotopy stabilization") sometimes, e.g., π_{D+1}(mass manifold) $= 0$. We shall discuss one such instance in appendix B.8. Fortunately, for the vast majority of applications n is sufficiently big so that π_{D+1}(mass manifold) $= \mathbb{Z}$.

4.9 Non-linear sigma models in $(2+1)$-D and $(3+1)$-D

As mentioned, the bosonization strategy described in the preceding section can be applied to two and three spatial dimensions. To facilitate later discussions, including the applications in $(2+1)$-D and $(3+1)$-D, the explicit form of the nonlinear sigma models in table 4.5 are given here. For briefness, we shall only include the results for sufficiently large n so that $\pi_{D+1}(\text{massmanifold}) = \mathbb{Z}$. As discussed earlier, under such conditions the non-linear sigma model possesses a WZW term.

4.9.1 Complex class in $(2+1)$-D

For Dirac fermions with n flavors in the complex class, after bosonization the sigma model matrix field (or the order parameter) lives in the space of complex Grassmannian, namely,

$$Q^{\mathbb{C}}(x) \in \frac{U(n)}{U(n/2) \times U(n/2)}.$$

This means that at any space-time point x, $Q^{\mathbb{C}}(x)$ is an $n \times n$ hermitian matrix with half of the eigenvalues $+1$, and the other half -1. One can specify $Q^{\mathbb{C}}(x)$ by the unitary matrix, $C(x)$, which renders $Q^{\mathbb{C}}(x)$ diagonalized upon similarity transformation, i.e.,

$$Q^{\mathbb{C}}(x) = C(x) \cdot \begin{pmatrix} I_{n/2} & 0 \\ 0 & -I_{n/2} \end{pmatrix} \cdot C^{\dagger}(x).$$

Obviously two different $C(x)$s related by

$$C'(x) = C(x) \cdot \begin{pmatrix} g_1(x) & 0 \\ 0 & g_2(x) \end{pmatrix},$$

where $g_1(x), g_2(x) \in U(n/2)$, will lead to identical $Q^{\mathbb{C}}(x)$. Due to this redundancy, the order parameter lives in the quotient space $\frac{U(n)}{U(n/2)\times U(n/2)}$.

Explicit fermion integration yields the following non-linear sigma model

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_3} \int_{\mathcal{M}} d^3x \, \mathrm{tr}\Big[\Big(\partial_\mu Q^{\mathbb{C}}\Big)^2\Big] - \frac{2\pi i}{256\pi^2} \int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}} \, \Big(d\tilde{Q}^{\mathbb{C}}\Big)^4\Big], \tag{4.28}$$

where λ_3 is a parameter having the dimension of length. In the limit where the short distance cutoff is zero,

$$\lambda_3 = \frac{8\pi}{m} \tag{4.29}$$

where m is the fermion energy gap.

The first term in Eq.(4.28) is the stiffness term and the second is the level-1 ($k = 1$) Wess-Zumino-Witten term. $\tilde{Q}^{\mathbb{C}}(x, u)$ is the extended field of $Q^{\mathbb{C}}(x)$, which exist because $\pi_3(\frac{U(n)}{U(n/2)\times U(n/2)}) = 0$ for $n \geq 4$. The difference in the WZW term associated with two different extensions is $2\pi i$ times the topological invariant associated with $\pi_4(\frac{U(n)}{U(n/2)\times U(n/2)}) = \mathbb{Z}$. Consequently upon exponentiation, different extensions yield the same phase factor. (To recapitulate the explanation, the readers are referred to subsection 4.8.1.)

4.9.2 Real class in $(2+1)$-D

For massless n-flavor Majorana fermions in the real class, the fluctuating order parameters $Q^{\mathbb{R}}(x)$ lives in the space of real Grassmannian, namely,

$$Q^{\mathbb{R}}(x) \in \frac{O(n)}{O(n/2) \times O(n/2)}.$$

This means that at any space-time point x, $Q^{\mathbb{R}}(x)$ is an $n \times n$ real symmetric matrix, with half of the eigenvalues $+1$, and the other half -1. One can specify $Q^{\mathbb{R}}(x)$ by the orthogonal matrix, $R(x)$, required to render $Q^{\mathbb{R}}(x)$ diagonalized, namely,

$$Q^{\mathbb{R}}(x) = R(x) \cdot \begin{pmatrix} I_{n/2} & 0 \\ 0 & -I_{n/2} \end{pmatrix} \cdot R^T(x).$$

Two different $R(x)$s related by

$$R'(x) = R(x) \cdot \begin{pmatrix} g_1(x) & 0 \\ 0 & g_2(x) \end{pmatrix},$$

where $g_1(x), g_2(x) \in O(n/2)$, will lead to identical $Q^{\mathbb{R}}(x)$. Due to this redundancy, the order parameter lives in the quotient space $\frac{O(n)}{O(n/2)\times O(n/2)}$.

Explicit fermion integration leads to the following non-linear sigma model

$$W[Q^{\mathbb{R}}] = \frac{1}{4\lambda_3} \int_{\mathcal{M}} d^3x \, \mathrm{tr}\left[\left(\partial_\mu Q^{\mathbb{R}}\right)^2\right] - \frac{2\pi i}{512\pi^2} \int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{R}} \left(d\tilde{Q}^{\mathbb{R}}\right)^4\right]. \tag{4.30}$$

Again, λ_3 has the dimension of length, and in the limit where the short-distance cutoff is zero λ_3 is given by Eq.(4.29).

The first term in Eq.(4.30) is the stiffness term and the second is the Wess-Zumino-Witten topological term of level $k = 1$. $\tilde{Q}^{\mathbb{R}}(x, u)$ is the extended field of $Q^{\mathbb{R}}(x)$, which exist because $\pi_3(\frac{O(n)}{O(n/2)\times O(n/2)}) = 0$ for $n \geq 6$. The difference in the WZW term associated with two different extensions is $2\pi i$ times the topological invariant associated with $\pi_4(\frac{O(n)}{O(n/2)\times O(n/2)}) = \mathbb{Z}$. Consequently upon exponentiation different extensions yield the same phase factor. (Again, to recapitulate the explanation, the readers are referred to subsection 4.8.2.)

4.9.3 Complex class in $(3+1)$-D

For the n-flavor massless Dirac fermions in the complex class, the fluctuating order parameters $Q^{\mathbb{C}}(x)$ lives in the space of $n \times n$ unitary matrices, namely,

$$Q^{\mathbb{C}}(x) \in U(n).$$

Explicit fermion integration leads to the following non-linear sigma model

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_4^2}\int_{\mathcal{M}} d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right] - \frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\right)^5\Big], \tag{4.31}$$

where λ_4 has the dimension of length. Using dimensional regularization λ_4 is given by

$$\frac{1}{\lambda_4} = \left[\frac{\Gamma(0^+)m^2}{8\pi^2}\right]^{1/2}, \tag{4.32}$$

signifying that λ_4 is cutoff-dependent. Here $\Gamma(0^+)$ is the gamma function evaluated at 0^+ from dimensional regularization (see appendix B.4 for the details.

The first term in Eq.(4.28) is the stiffness term and the second is the level $k = 1$ Wess-Zumino-Witten term. $\tilde{Q}^{\mathbb{C}}(x,u)$ is the extended field of $Q^{\mathbb{C}}(x)$, which exist because $\pi_4(U(n)) = 0$ for $n \geq 3$. The difference in the WZW term associated with two different extensions is $2\pi i$ times the topological invariant associated with $\pi_5(U(n)) = \mathbb{Z}$. Consequently upon exponentiation different extensions yield the same phase factor. (Again, to recapitulate the explanation, the readers are referred to subsection 4.8.1.)

4.9.4 Real class in $(3+1)$-D

For the n-flavor massless Majorana fermions in the complex class, the fluctuating order parameters $Q^{\mathbb{R}}(x)$ lives in the space of "real Lagrangian Grassmannian", namely,

$$Q^{\mathbb{R}}(x) \in \frac{U(n)}{O(n)}.$$

This means that at any space-time point x, $Q^{\mathbb{R}}(x)$ is an $n\times n$ symmetric unitary matrix. According to the Autonne decomposition (e.g., corollary 2.6.6 of [25]), any symmetric unitary matrix can be decompose into

$$Q^{\mathbb{R}}(x) = W(x)\cdot W^T(x),$$

where $W(x)$ is unitary. Hence, two different $W(x)$s related by

$$W'(x) = W(x)\cdot g(x),$$

where $g(x) \in O(n)$, will lead to identical $Q^{\mathbb{R}}(x)$. Due to this redundancy, the order parameter lives in the quotient space $\frac{U(n)}{O(n)}$.

Explicit fermion integration yields the following non-linear sigma model

$$W[Q^{\mathbb{R}}] = \frac{1}{4\lambda_4^2}\int_{\mathcal{M}} d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}}\partial^\mu Q^{\mathbb{R}\dagger}\right] - \frac{2\pi}{960\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{R}\dagger}d\tilde{Q}^{\mathbb{R}}\right)^5\Big]. \tag{4.33}$$

The first term in Eq.(4.30) is the stiffness term and the second is the level $k = 1$ Wess-Zumino-Witten topological term. $\tilde{Q}^{\mathbb{R}}(x,u)$ is the extended field of $Q^{\mathbb{R}}(x)$, which exist because $\pi_4(U(n)/O(n)) = 0$ for $n \geq 5$. The difference in the WZW term associated with two different extensions is $2\pi i$ times the topological invariant associated with $\pi_5(U(n)/O(n)) = \mathbb{Z}$. Consequently upon exponentiation different extensions yield the same phase factor. (Again, to recapitulate the explanation, the readers are referred to subsection 4.8.2.)

In table 4.5 we summarize the n values above which π_{D+1}(mass manifold) is stabilized. We shall discuss some of the small n cases which are relevant to our applications in appendix B.8.

4.10 Non-linear sigma models as the effective theories of interacting fermion models

As we have seen in section 4.9, while the coefficient in front of the stiffness term in the non-linear sigma model is dimensionless in $(1+1)$-D, those in $(2+1)$-D and $(3+1)$-D are dimensionful parameters. This begs the question of what are these parameters? and for what values of these parameters are the non-linear sigma models equivalent to the massless fermion theories? In addition, for $D = 2+1$ the mass manifold consists of more than one connected components. What kind of model can realize phases correspond to different components of the mass manifold? In the following, we answer these questions by focusing on the complex class. It is straightforward to generalize the result to the real class.

	Real class	complex class
$(1+1)$-D	$O(n)_1$ WZW term stabilized for $n \geq 3$	$U(n)_1$ WZW term stabilized for $n \geq 2$
$(2+1)$-D	$\left[\frac{O(n)}{O(n/2)\times O(n/2)}\right]_1$ WZW term stabilized for $n \geq 6$	$\left[\frac{U(n)}{U(n/2)\times U(n/2)}\right]_1$ WZW term stabilized for $n \geq 4$
$(3+1)$-D	$[U(n)/O(n)]_1$ WZW term stabilized for $n \geq 5$	$U(n)_1$ WZW term stabilized for $n \geq 3$

Table 4.5: The n values above which the π_{D+1}(mass manifold) is stabilized.

As listed in table 4.2, the mass terms correspond to $Q^{\mathbb{C}}$ are given by

$$\begin{aligned}
(1+1)\text{-D}: \quad & M[Q^{\mathbb{C}}] = X \otimes \frac{1}{2}\left[Q^{\mathbb{C}} + \left(Q^{\mathbb{C}}\right)^\dagger\right] + Y \otimes \frac{1}{2i}\left[Q^{\mathbb{C}} - \left(Q^{\mathbb{C}}\right)^\dagger\right] \\
(2+1)\text{-D}: \quad & M[Q^{\mathbb{C}}] = Y \otimes Q^{\mathbb{C}} \\
(3+1)\text{-D}: \quad & M[Q^{\mathbb{C}}] = YX \otimes \frac{1}{2}\left[Q^{\mathbb{C}} + \left(Q^{\mathbb{C}}\right)^\dagger\right] + YY \otimes \frac{1}{2i}\left[Q^{\mathbb{C}} - \left(Q^{\mathbb{C}}\right)^\dagger\right]
\end{aligned} \tag{4.34}$$

Let's consider the four-fermion interacting generated by the following inverse Hubbard-Stratonovich transformation,

$$\begin{aligned}
&\exp\left\{-S_{\mathrm{I}}\left[\psi^\dagger, \psi\right]\right\} := \\
&\int D\left[\mathcal{Q}(x)\right] \exp\left\{-\int d^D x \left[\psi^\dagger \; M[\mathcal{Q}(x)] \; \psi + \frac{1}{2\lambda_I}\mathrm{tr}\left[\mathcal{Q}(x)^\dagger \mathcal{Q}(x)\right]\right]\right\}
\end{aligned} \tag{4.35}$$

where $\mathcal{Q}(x)$ is an $n \times n$ matrix-valued function of space-time. We note that the strength of the four fermion interaction in Eq.(4.35) is proportional to λ_I.

The emergent global symmetries transform $\mathcal{Q}(x)$ in exactly the same way as $Q^{\mathbb{C}}$ (see table 4.1). This is because $\mathcal{Q}(x)$ and $Q^{\mathbb{C}}$ couple to the same fermion bi-linears. Such transformation can be absorbed by the redefinition of the integration variable $\mathcal{Q}(x)$. Therefore as long as the integration measure in Eq.(4.35) is symmetric under the symmetry transformations, S_I is invariant under the action of emergent symmetries.

When λ_I is sufficiently large, it is energetically favorable for

$$\mathrm{tr}\left[\langle \mathcal{Q}^\dagger(x)\mathcal{Q}(x)\rangle\right]$$

to acquire a non-zero expectation value. Assuming such expectation value doesn't spontaneously break the continuous symmetry[7] it must satisfy

$$\langle \mathcal{Q}^\dagger(x)\mathcal{Q}(x)\rangle \to g^\dagger \cdot \langle \mathcal{Q}^\dagger(x)\mathcal{Q}(x)\rangle \cdot g = \langle \mathcal{Q}^\dagger(x)\mathcal{Q}(x)\rangle$$

for all $g \in U(n)$ (for $(1+1)$-D and $(3+1)$-D $g \in U_+(n)$). This requires the expectation value of $\mathcal{Q}^\dagger(x)\mathcal{Q}(x)$ to be proportional identity matrix,

$$\langle \mathcal{Q}^\dagger(x)\mathcal{Q}(x)\rangle = \kappa^2 I_n$$

where κ^2 should grow monotonically with λ_I. At low energy and long wavelength, the dynamics of $\mathcal{Q}$ is governed by the Goldstone modes $Q^{\mathbb{C}}(x)$, where

$$\mathcal{Q}(x) \to \kappa Q^{\mathbb{C}}(x), \quad \text{and} \quad \left(Q^{\mathbb{C}}\right)^\dagger Q^{\mathbb{C}} = I_n.$$

The manifold in which $Q^{\mathbb{C}}(x)$ fluctuates is exactly the mass manifold given in table 4.2.

[7] The possible symmetry breaking phases are captured by the non-zero expectation value $\langle \mathcal{Q}(x)\rangle$.

The effective action governing the fluctuations of $Q^{\mathbb{C}}(x)$ is given by the results of section 4.9, where the stiffness term coefficients $\frac{1}{2\lambda_3}$ and $\frac{1}{2\lambda_4^2}$ should grow with κ^2 which, in turn, monotonically increases with λ_I. As the result, strong four-fermion interaction implies small λ_3 and λ_4^2, while weak four fermion-interaction implies large λ_3 and λ_4^2. Thus, we obtain a duality-like relation, namely, strong coupling fermion theory corresponds to weak coupling non-linear sigma model, and weak coupling fermion theory corresponds to strong coupling non-linear sigma model. Since, by dimension counting, local four-fermion interaction is an irrelevant perturbation to the massless theory in $(2+1)$- and $(3+1)$-D, we expect there is a range of large λ_3 and λ_4^2 where the non-linear sigma model is massless.

Now we come to $(2+1)$-D, where according to table 4.2 , the mass manifold has $n+1$ components, namely,

$$Q^{\mathbb{C}} \in \bigcup_{l=0}^{n} \frac{U(n)}{U(l)\times U(n-l)}.$$

(Here l corresponds to the number positive eigenvalues of $Q^{\mathbb{C}}(x)$, the readers are referred to appendix B.2 for details.) The condition that the order parameter is a smooth function of space-time confines $Q^{\mathbb{C}}(x)$ to fluctuate in one of the mass manifold components. If such fluctuation is to restore the time-reversal symmetry, it further restricts $l = n/2$ (we focus on $n =$ even). However, if we allow the possibility of spontaneous time-reversal symmetry breaking, then $Q^{\mathbb{C}}(x)$ can fluctuate in the $l \neq n/2$ mass manifold. It is interesting whether the order parameter fluctuation in the $l \neq n/2$ mass manifolds can restore the unitary part of the emergent symmetry, and if it does can the resulting phase be gapless.

4.11 Global symmetry of the non-linear sigma model

Up to this point, we have derived the non-linear sigma model. The bosonic partition function is given by

$$\mathcal{Z} = \int D[Q^{\mathbb{C},\mathbb{R}}]\; e^{-S_{\mathrm{NL}\sigma}[Q^{\mathbb{C},\mathbb{R}}]}.$$

Here $Q^{\mathbb{C},\mathbb{R}} \in$ mass manifolds, and the integration measure is defined so that at every space-time point $Q^{\mathbb{C},\mathbb{R}}$ and the symmetry transformed $Q^{\mathbb{C},\mathbb{R}}$ (see table 4.5) have the same weight.

Now, using the complex class in $(1+1)$-D as an example, we demonstrate that the non-linear sigma model in Eq.(4.24) respects the emergent symmetries of the massless free fermion theory. Under the action of the global emergent symmetries, a configuration $Q^{\mathbb{C}}(\tau,\mathbf{x})$ transforms by Eq.(4.20), namely,

$$U_+(n) \times U_-(n) : Q^{\mathbb{C}}(\tau, \mathbf{x}) \to g_-^\dagger \cdot Q^{\mathbb{C}}(\tau, \mathbf{x}) \cdot g_+$$
$$\text{Charge conjugation} : Q^{\mathbb{C}}(\tau, \mathbf{x}) \to \left(Q^{\mathbb{C}}(\tau, \mathbf{x})\right)^*$$
$$\text{Time reversal} : Q^{\mathbb{C}}(\tau, \mathbf{x}) \to \left(Q^{\mathbb{C}}(\tau, \mathbf{x})\right)^T .$$

Under the action of $U_+(n) \times U_-(n)$

$$Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \to g_+^\dagger \cdot \left(Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\right) \cdot g_+$$

Due to the cyclic invariance of trace, the similarity transformations cancel out and the action Eq.(4.24) is invariant.

Under charge conjugation, the stiffness term transforms as

$$\begin{aligned}
&-\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right)^2\right] \to -\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\left(Q^{\mathbb{C}T}\partial^\mu Q^{\mathbb{C}*}\right)^2\right] \\
=&-\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\left(\partial^\mu Q^{\mathbb{C}\dagger}Q^{\mathbb{C}}\right)\left(\partial^\mu Q^{\mathbb{C}\dagger}Q^{\mathbb{C}}\right)\right] = -\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right)\left(Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right)\right]
\end{aligned}$$

hence is invariant. Here the first equality in the second line is due to the invariance of trace under transposing, and the second equality is due to $\partial^\mu Q^{\mathbb{C}\dagger}Q^{\mathbb{C}} = -Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}$. A similar argument applies to the WZW term,

$$\begin{aligned}
&\frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\right)^3\Big] \to \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}T}d\tilde{Q}^{\mathbb{C}*}\right)^3\Big] \\
=&-\frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(d\tilde{Q}^{\mathbb{C}\dagger}\tilde{Q}^{\mathbb{C}}\right)^3\Big] = \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\right)^3\Big].
\end{aligned}$$

The extra minus sign in the second line is because transposing causes an odd number of crossing of the differential 1-forms. This negative sign is canceled out in the last term due to the odd number of negative signs arising from $dQ^{\mathbb{C}\dagger}Q^{\mathbb{C}} = -Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}}$. Therefore Eq.(4.24) is invariant under charge conjugation.

Under the action of time-reversal transformation, the stiffness term transforms as

$$\begin{aligned}
&\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right] \to \frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}*}\partial^\mu Q^{\mathbb{C}T}\right] \\
=&\frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\partial^\mu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\right] = \frac{1}{8\pi}\int_{\mathcal{M}} d^2x \, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}\dagger}\partial^\mu Q^{\mathbb{C}}\right]
\end{aligned}$$

As for the WZW term, note that the i in front becomes $-i$ due to the complex conjugation involved in the time-reversal transformation [8]. Thus the WZW term transforms as

$$\begin{aligned}
&\frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^3\Big] \to -\frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}*} d\tilde{Q}^{\mathbb{C}T}\right)^3\Big] \\
&= \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(d\tilde{Q}^{\mathbb{C}}\tilde{Q}^{\mathbb{C}\dagger}\right)\left(d\tilde{Q}^{\mathbb{C}}\tilde{Q}^{\mathbb{C}\dagger}\right)\left(d\tilde{Q}^{\mathbb{C}}\tilde{Q}^{\mathbb{C}\dagger}\right)\Big] \\
&= \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^3\Big] \qquad (4.36)
\end{aligned}$$

The disappearance of the minus sign in the second line is because transposing causes an odd number of crossings of differential 1-forms. The passing to the third line follows from the cyclic invariance of trace.

In summary, the non-linear sigma model is invariant under the action of the global emergent symmetries. The same conclusion applies to the real and complex classes non-linear sigma models in other space-time dimensions. The details is left in appendix B.4.

4.12 The symmetry anomalies of the nonlinear sigma models

A necessary condition for the bosonized non-linear sigma model to be equivalent to the massless fermion theory is that the former has the same symmetry anomalies as the original massless fermion theories. In this section, we will show this is indeed the case.

4.12.1 Gauging the non-linear sigma model and the ‘t Hooft anomalies

In table 4.3 we see that in $(1+1)$-D and $(3+1)$-D, the massless free fermion theories have the ‘t Hooft anomalies (with respect to the continuous symmetries). In this subsection, we first gauge the non-linear sigma models and then determine their ‘t Hooft anomalies.

[8]The time-reversal symmetry in Euclidean space-time requires a complex conjugation on the Boltzmann weight. It is important to check whether a term is real or complex before deciding how time-reversal transformation acts.

Again, taking the complex class (1 + 1)-D example, under an infinitesimal $U_+(n) \times U_-(n)$ transformation, $Q^{\mathbb{C}}$ and gauge fields transformed as

$$\begin{aligned} &Q^{\mathbb{C}} \to e^{-i\epsilon_-} Q^{\mathbb{C}} e^{i\epsilon_+} \\ &A_\pm \to A_\pm + d\epsilon_\pm + i[A_\pm, \epsilon_\pm] \end{aligned} \tag{4.37}$$

where we let $g_\pm = e^{i\epsilon_\pm}$ in the symmetry transformation. For the stiffness term, the usual minimal coupling guarantees the gauge invariance

$$W_{\text{stiff}}[Q^{\mathbb{C}}, A_+, A_-] = -\frac{1}{8\pi} \int_{\mathcal{M}} d^2x \, \text{tr}\left[\left(Q^{\mathbb{C}\dagger}\left(\partial_\mu Q^{\mathbb{C}} - iQ^{\mathbb{C}} A_{+,\mu} + iA_{-,\mu} Q^{\mathbb{C}}\right)\right)^2\right].$$

However, it is less clear how to gauge the WZW term. Here we follow Witten's "trial-and-error" method [26], which we shall explain in the following.

First, we determine the variation of the WZW term when $Q^{\mathbb{C}}$ undergoes space-time dependent transformation given by the first line of Eq.(4.37)

$$\begin{aligned} &\delta\Big[-\frac{i}{12\pi}\int_{\mathcal{B}} \text{tr}\left[\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right)^3\right]\Big] \\ =&\frac{1}{4\pi}\int_{\mathcal{M}} \text{tr}\left[d\epsilon_+\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right) + d\epsilon_-\left(dQ^{\mathbb{C}} Q^{\mathbb{C}\dagger}\right)\right] \end{aligned}$$

Here we remark that although writing down the action requires the extended space-time manifold $\mathcal{B}$, the variation of the action can be expressed solely in the space-time manifold $\mathcal{M}$, which is $(1+1)$-D in the example.

In an attempt to make the theory gauge invariant, we subtract a term with $d\epsilon_\pm$ replaced by $A_\pm$. Together, the gauge variant part becomes

$$\begin{aligned} &\delta\Big[-\frac{i}{12\pi}\int_{\mathcal{B}} \text{tr}\left[\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right)^3\right] - \frac{1}{4\pi}\int_{\mathcal{M}} \text{tr}\left[A_+\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right) + A_-\left(dQ^{\mathbb{C}} Q^{\mathbb{C}\dagger}\right)\right]\Big] \\ =&-\frac{i}{4\pi}\int_{\mathcal{M}} \text{tr}\left[A_+\left(d\epsilon_+ - Q^{\mathbb{C}\dagger} d\epsilon_- Q^{\mathbb{C}}\right) + A_-\left(-d\epsilon_- + Q^{\mathbb{C}} d\epsilon_+ Q^{\mathbb{C}\dagger}\right)\right] \end{aligned}$$

Last, we repeat the previous step by adding another term with $d\epsilon_\pm$ in the above equation replaced by $A_\pm$. After some work we obtain

$$\begin{aligned} &\delta\Big[-\frac{i}{12\pi}\int_{\mathcal{B}} \text{tr}\left[\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right)^3\right] - \frac{1}{4\pi}\int_{\mathcal{M}} \text{tr}\left[A_+\left(Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}\right) + A_-\left(dQ^{\mathbb{C}} Q^{\mathbb{C}\dagger}\right) + iA_+ Q^{\mathbb{C}\dagger} A_- Q^{\mathbb{C}}\right]\Big] \\ =&-\frac{i}{4\pi}\int_{\mathcal{M}} \text{tr}\left[A_+ d\epsilon_+ - A_- d\epsilon_-\right] \end{aligned}$$

Now the gauge variant part contains no $Q^{\mathbb{C}}$ anymore. Hence we cannot find any term to cancel the remaining non-gauge-invariance. This result reproduces Bardeen's result in Eq.(4.13).

In summary, the gauged WZW model is given by

$$\begin{aligned} W[Q^{\mathbb{C}}, A_+, A_-] = &-\frac{1}{8\pi}\int_{\mathcal{M}} d^2x\, \mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}\left(\partial_\mu Q^{\mathbb{C}} - iQ^{\mathbb{C}}A_{+,\mu} + iA_{-,\mu}Q^{\mathbb{C}}\right)\right)^2\right] \\ &-\frac{i}{12\pi}\int_{\mathcal{B}} \mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}}\right)^3\right] \\ &-\frac{1}{4\pi}\int_{\mathcal{M}} \mathrm{tr}\left[A_+\left(Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}}\right) + A_-\left(dQ^{\mathbb{C}}Q^{\mathbb{C}\dagger}\right) + iA_+Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}}\right]\Big]. \end{aligned}$$

Moreover, we have shown that it has the same 't Hooft anomaly for the continuous symmetry as the original massless fermion. In appendix B.6 we summarize the gauged non-linear sigma model in $d = 1, 2, 3$.

4.12.2 Discrete symmetry anomaly

In section 4.4.3, we saw that massless fermion theory has a time-reversal anomaly for the complex class in $(2+1)$-D. This anomaly originates from the massive Dirac fermion at time reversal invariant $\boldsymbol{k}$ points other than $\boldsymbol{k} = (0, 0)$ where the mass breaks time-reversal. We would like to see the same phenomenon in the nonlinear sigma model.

In the following, we focus on the complex class in $(2+1)$-D. First, let's focus on the vicinity of $\boldsymbol{k} = 0$ (under Wilson's regularization). The bosonized model is given by Eq.(4.28). Following Witten's trial-and-error method discussed in the preceding subsection (see B.6 for the detail), we obtain the following gauged nonlinear sigma model,

$$\begin{aligned} W[Q^{\mathbb{C}}, A] = &\frac{1}{2\lambda_3}\int_{\mathcal{M}} d^3x\, \mathrm{tr}\Big[\Big(\partial_\mu Q^{\mathbb{C}} + i[A_\mu, Q^{\mathbb{C}}]\Big)^2\Big] \\ &-\frac{2\pi i}{256\pi^2}\Big\{\int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\Big] \\ &+8\int_{\mathcal{M}} \mathrm{tr}\Big[iAQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2 - (AQ^{\mathbb{C}})^2 dQ^{\mathbb{C}} \\ &-\frac{i}{3}(AQ^{\mathbb{C}})^3 + iA^3Q^{\mathbb{C}} - AQ^{\mathbb{C}}F - AFQ^{\mathbb{C}}\Big]\Big\} \end{aligned} \tag{4.38}$$

This action is invariant under global symmetry transformations where the gauge field and $Q^{\mathbb{C}}$ are transformed according to Eq.(4.15) and table 4.4. This is expected, given the low energy fermion theory near $\boldsymbol{k} = 0$ respects these symmetries.

For the Dirac fermions near $\boldsymbol{k} = (\pi, 0), (0, \pi)$, and (π, π), there are time reversal breaking masses, namely, $M = 2mY \otimes I_n$ for $\boldsymbol{k} = (\pi, 0), (0, \pi)$ and $M = 4mY \otimes I_n$ for (π, π). The non-linear sigma model describes these massive fermions is again given by Eq.(4.38) except that now $l = n$ or 0. Due to the signs in front of $q_1\Gamma_1$ and $q_2\Gamma_2$ at $\boldsymbol{k} = (\pi, 0), (0, \pi)$, and (π, π) the effective mass sign for these massive fermions are given by

$$\eta_{\boldsymbol{k}} := \text{sign of } (q_1\Gamma_1) \times \text{sign of } (q_2\Gamma_2) \times \text{sign of } (m).$$

Consequently the $Q^{\mathbb{C}}$ associated with the massive fermions obeys

$$Q^{\mathbb{C}} = \eta_{\boldsymbol{k}} I_n. \tag{4.39}$$

We can thus use the gauged nonlinear sigma model in appendix B.6 to predict the Chern-Simons term due to the massive fermions at $\boldsymbol{k} = (\pi, 0), (0, \pi)$, and (π, π) by plug in Eq.(4.39). For these space-time constant $Q^{\mathbb{C}}$ we can drop all the terms with derivatives on $Q^{\mathbb{C}}$. The remaining can be combined into the Chern-Simons term. Summing the contribution from $\boldsymbol{k}$ around $(\pi, 0)$, $(0, \pi)$, and (π, π), we get

$$\begin{aligned} W_{(\pi,0)} + W_{(0,\pi)} + W_{(\pi,\pi)} &= \left(-\frac{1}{2} - \frac{1}{2} + \frac{1}{2}\right) \frac{m}{|m|} \frac{i}{4\pi} \int \text{tr}\left[A\, dA + \frac{2i}{3} A^3\right] \\ &= -\frac{i}{8\pi} \int \text{tr}\left[A\, dA + \frac{2i}{3} A^3\right] \end{aligned}$$

which agrees with Eq.(4.19).

As for other discrete symmetry anomalies, with the input of how $Q^{\mathbb{C},\mathbb{R}}$ and the gauge field transform under discrete symmetries, it's simple to show that in $(1+1)$-D and $(3+1)$-D, there is no discrete-symmetry-anomaly after gauging the anomaly-free part of the continuous symmetries. In $(2+1)$-D, gauging the continuous symmetry breaks the time-reversal symmetry as discussed in subsection 4.12.2.

In appendix B.6, we show that all the symmetry anomalies of massless fermions in table 4.3 are reproduced by the corresponding gauged nonlinear sigma models. This lends strong support to the idea that the nonlinear sigma models are equivalent to the original massless fermion theories.

4.13 Soliton of the non-linear sigma model and the Wess-Zumino-Witten term

In order for the bosonization to hold, somehow the bosonic non-linear sigma model must possess fermion degrees of freedom. In this section, we show that due to the WZW term, the solitons of the non-linear sigma model are fermions.

4.13.1 Soliton classification

Soliton is a spatial texture of the "order parameter" ($Q^{\mathbb{C},\mathbb{R}}$). Such texture represents a non-trivial mapping from the spatial space to the mass manifold, i.e., the space where the order parameter lives. In d spatial dimension, solitons are classified by the d-th homotopy group of the mass manifold, namely,

$$\pi_d\,(\text{mass manifold})\,.$$

In appendix B.2, we list the relevant homotopy groups. Since exchange statistics only make sense for spatial dimension greater than one, in the following we shall focus on $d \geq 2$. For the nonlinear sigma models considered in section 4.9, when n is sufficiently large so that there is a WZW term, the soliton classifications are $\mathbb{Z}$ for the complex classes, and are $\mathbb{Z}_2$ for the real classes, namely,

$$\begin{aligned}
&\pi_2\Big(\frac{U(n)}{U(n/2)\times U(n/2)}\Big) = \mathbb{Z} && \text{for } n \geq 4\\
&\pi_2\Big(\frac{O(n)}{O(n/2)\times O(n/2)}\Big) = \mathbb{Z}_2 && \text{for } n \geq 6\\
&\pi_3\Big(U(n)\Big) = \mathbb{Z} && \text{for } n \geq 3\\
&\pi_3\Big(\frac{U(n)}{O(n)}\Big) = \mathbb{Z}_2 && \text{for } n \geq 5
\end{aligned}$$

This means that for the complex classes, we can define a topological quantum number, namely, the "soliton charge" $Q_{\rm sol}$. When we fuse two solitons of different charges, $Q_{\rm sol}$ adds; for the real classes, on the other hand, this soliton charge is defined *mod* 2 so that two solitons with unit soliton charges can fuse into zero soliton charge.

4.13.2 Soliton charge and the conserved $U(1)$ charge Q

For the complex classes, it is natural to ask what is the relation between the soliton charge $Q_{\rm sol}$ and the conserved charge Q. The conserved charge Q is associated with a global U(1) symmetry. In $(3+1)$-D such U(1) symmetry belongs to a diagonal subgroup of $U_+(n) \times U_-(n)$. As shown in table 4.3, it is anomaly-free. For $(2+1)$-D the $U(1)$ symmetry is a subgroup of the global symmetry group $U(n)$, which is also anomaly-free according to table 4.3.

In appendix B.6 we present the gauged non-linear sigma model. In particular, by focusing on the term linear in the gauge field (associated with the anomaly-free $U(1)$

subgroup) derived from the WZW term, we can extract the $U(1)$ current. The answer is [9]

$$
\begin{aligned}
(2+1)\text{-D}: \ J^\mu &= -\frac{i}{16\pi}\epsilon^{\mu\nu\rho}\text{tr}\left[Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}}\partial_\rho Q^{\mathbb{C}}\right] \\
(3+1)\text{-D}: \ J^\mu &= -\frac{1}{24\pi^2}\epsilon^{\mu\nu\rho\sigma}\text{tr}\left[\left(Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}\right)\left(Q^{\mathbb{C}\dagger}\partial_\rho Q^{\mathbb{C}}\right)\left(Q^{\mathbb{C}\dagger}\partial_\sigma Q^{\mathbb{C}}\right)\right].
\end{aligned}
\tag{4.40}
$$

Thus the $U(1)$ charge given by

$$
\begin{aligned}
(2+1)\text{-D}: \ Q &= -\frac{i}{16\pi}\int d^2\boldsymbol{x}\,\epsilon^{ij}\text{tr}\left[Q^{\mathbb{C}}\partial_i Q^{\mathbb{C}}\partial_j Q^{\mathbb{C}}\right] \\
(3+1)\text{-D}: \ Q &= -\frac{1}{24\pi^2}\int d^3\boldsymbol{x}\,\epsilon^{ijk}\text{tr}\left[\left(Q^{\mathbb{C}\dagger}\partial_i Q^{\mathbb{C}}\right)\left(Q^{\mathbb{C}\dagger}\partial_j Q^{\mathbb{C}}\right)\left(Q^{\mathbb{C}\dagger}\partial_k Q^{\mathbb{C}}\right)\right].
\end{aligned}
$$

These are, in fact, exactly the same expression as the topological invariant corresponding to $\pi_2(\frac{U(n)}{U(n/2)\times U(n/2)}) = \mathbb{Z}$ in $(2+1)$-D and $\pi_3(U(n)) = \mathbb{Z}$ in $(3+1)$-D (see appendix B.2 for the details). Thus, for both cases

$$
Q = Q_{\text{sol}}. \tag{4.41}
$$

4.13.3 Statistics of soliton

One way to derive the statistics of soliton is to calculate the topological spin by comparing Berry's phase difference between the following two processes. In the first process, we have a static soliton. In the second process, the spatial soliton configuration is adiabatically rotated by 2π in time. Following Witten [27], we show in appendix B.7 that such Berry's phase difference is $e^{-ik\pi}$, where k is the level of the WZW term (see appendix B.7 for the details). Since all nonlinear sigma models in section 4.9 have $k = 1$ WZW term, their solitons are fermion.

4.14 A summary of bosonization

So far, we have established the fact that the fermion and boson theories have the same global symmetries and anomalies. In addition, we have shown that the solitons of the bosonic theories are fermions. All these support the equivalence between the fermion and boson theories. Now we present a brief summary of Chap. 4.

[9]The same result can be derived by fermion integration.

We begin in section 4.1 by presenting the essential idea underlying the present work. Prior to performing the fermion integration, we first identify the emergent symmetries in section 4.2, and the mass manifolds in section 4.3. For a given massless fermion theory, the mass manifold is the topological space formed by all mass terms that can fully gap out the fermions. We then work out the anomalies with respect to the emergent symmetries in section 4.4. Afterward, we introduce mass terms at the expense of breaking the emergent symmetries in section 4.5 and fluctuate the mass terms *smoothly* to regain the emergent symmetries in section 4.6. As discussed in section 4.7, the smoothness of the mass fluctuations is to ensure that the original fermions remain gapped, hence can be integrated out to yield non-linear sigma models in section 4.8 and section 4.9.[10]. The level-1 WZW term resulting from the fermion integration is checked against the prediction of homotopy groups in the appendix, which is referred to in sections 4.8 and 4.9. In section 4.10, we present local interacting fermion theories that have duality-like relationships with the bosonized non-linear sigma models. In section 4.11, we analyze the symmetries of the non-linear sigma models. A comparison with the results obtained in section 4.2 leads to the conclusion that the fermion and boson theories have the same symmetry. Using the method of reference [26] we determine the anomalies of the non-linear sigma models in section 4.12. A comparison with the results obtained in section 4.4 leads to the conclusion that the fermion and boson theories have the same anomalies. Finally, in section 4.13, we show the bosonized theories have fermionic degrees of freedom, namely the solitons of the non-linear sigma models.

[10]The procedure can be easily applied to higher dimensions, though we shall not pursue it in the present paper.

Chapter 5

Applications of bosonization

5.1 The $SU(2)$ gauge theory of the π-flux phase of the half-filled Hubbard model

5.1.1 The "spinon" representation of the half-filled Mott insulator

The paradigmatic model describing a Mott insulator is the Hubbard model in the large U limit. At half-filling, every site is occupied by one electron. Below the Mott-Hubbard gap, the active degrees of freedom are those of spins. Through Anderson's super-exchange [28], the dynamics of the spins is governed by the anti-ferromagnetic Heisenberg interaction

$$\hat{H} = \sum_{\langle ij \rangle} J_{ij} \vec{S}_i \cdot \vec{S}_j.$$

In the "spinon" treatment [29, 30] one decomposes a spin-1/2 operator into auxiliary fermion (spinon) operators

$$S_i^a = \frac{1}{2} f_{i\alpha}^\dagger \sigma_{\alpha\beta}^a f_{i\beta}, \tag{5.1}$$

and supplement it with the single occupation constraints

$$\begin{aligned} & f_{i\uparrow}^\dagger f_{i\uparrow} + f_{i\downarrow}^\dagger f_{i\downarrow} = 1 \\ & f_{i\uparrow}^\dagger f_{i\downarrow}^\dagger = 0 \\ & f_{i\downarrow} f_{i\uparrow} = 0. \end{aligned} \tag{5.2}$$

In the following we shall refer to the above constraints as the "Mott constraint". The decomposition in Eq.(5.1), where one separates the physical spin degrees of freedom into the auxiliary "spinon" degrees of freedom, is an example of the so-called "slave particle" approach.

In terms of the spinon operators the Heisenberg Hamiltonian read

$$\begin{aligned}\hat{H} =&\frac{1}{4}\sum_{\langle ij\rangle} J_{ij}\left(f^\dagger_{i\alpha}\sigma^a_{\alpha\beta}f_{i\beta}\right)\left(f^\dagger_{j\gamma}\sigma^a_{\gamma\delta}f_{j\delta}\right)\\ =&\frac{1}{4}\sum_{\langle ij\rangle} J_{ij}\left(-f^\dagger_{i\alpha}f_{i\alpha}f^\dagger_{j\beta}f_{j\beta}-2f^\dagger_{i\alpha}f_{j\alpha}f^\dagger_{j\beta}f_{i\beta}\right)\\ =&-\frac{1}{2}\sum_{\langle ij\rangle} J_{ij}\left(\frac{1}{2}f^\dagger_{i\alpha}f_{i\alpha}f^\dagger_{j\beta}f_{j\beta}+f^\dagger_{i\alpha}f_{j\alpha}f^\dagger_{j\beta}f_{i\beta}\right)\end{aligned}$$

Upon Hubbard-Stratonovich transformation, we express

$$\exp\Big\{-\int_0^\beta d\tau\Big[\sum_i f^\dagger_{i\alpha}\partial_0 f_{i\alpha}+H\Big]\Big\}=$$

$$\int D[U]\exp\Big\{-\int_0^\beta d\tau\Big[\sum_i \psi_i^\dagger\partial_0\psi_i+\sum_{\langle ij\rangle}\frac{3}{8}J_{ij}\Big(-\left(\psi_i^\dagger U_{ij}\psi_j+h.c.\right)+\frac{1}{2}\mathrm{Tr}\left[U^\dagger_{ij}U_{ij}\right]\Big)\Big]\Big\}. \tag{5.3}$$

where

$$\psi_i=\begin{pmatrix}f_{i\uparrow}\\ f^\dagger_{i\downarrow}\end{pmatrix}, U_{ij}=\begin{bmatrix}\chi^*_{ij} & \Delta_{ij}\\ \Delta^*_{ij} & -\chi_{ij}\end{bmatrix}. \tag{5.4}$$

For later convenience, we rewrite the spinon operator in terms of Majorana fermions

$$f_{i\alpha}:=F_{i,1\alpha}+iF_{i,2\alpha},$$

in terms of which, the spin operators are represented as

$$\begin{aligned}S^a_i&=\frac{1}{2}F_i^\dagger\Sigma^a F_i,\quad \text{where}\\ \Sigma^a&=\left(YX,IY,YZ\right).\end{aligned} \tag{5.5}$$

In the last line, the first and second Pauli matrices carry the Majorana and spin indices, respectively.

The spin operators in Eq.(5.5) are invariant under the following local "charge-SU(2)" transformation

$$F_i\to W_iF_i$$

where W_i is generated by

$$T^b=(XY,YI,ZY).$$

In terms of The Majorana fermion operators, the Mott constraint in Eq.(5.2) becomes

$$\begin{aligned}&f^\dagger_{i\alpha}f_{i\alpha}-1=F_i^T\left(YI\right)F_i=F_i^TT^2F_i=0\\ &\epsilon^{\alpha\beta}\left(f_{i\alpha}f_{i\beta}+f^\dagger_{i\beta}f^\dagger_{i\alpha}\right)=F_i^T\left(XY\right)F_i=F_i^TT^1F_i=0\\ &i\epsilon^{\alpha\beta}\left(f_{i\alpha}f_{i\beta}-f^\dagger_{i\beta}f^\dagger_{i\alpha}\right)=F_i^T\left(ZY\right)F_i=F_i^TT^3F_i=0\end{aligned} \tag{5.6}$$

These constraints are implemented via the Lagrange multipliers in the path integral

$$Z = \int D[F]D[U]D[a_0]\exp\left(-S\right)$$

with

$$S = \int_0^\beta d\tau \Big\{ \sum_i F_i^T \partial_0 F_i + \sum_{\langle ij \rangle} \frac{3}{8} J_{ij} \Big[F_i^T \Big(Re[\chi_{ij}] YI + i\, Im[\chi_{ij}] II + Re[\Delta_{ij}] XY - Im[\Delta_{ij}] ZY \Big) F_j + |\chi_{ij}|^2 + |\eta_{ij}|^2 \Big] + i \sum_i a_{i0}^b \left(F_i^T T^b F_i \right) \Big\}. \tag{5.7}$$

5.1.2 The π-flux phase mean-field theory and the $SU(2)$ gauge fluctuations

In treating the path integral, Eq.(5.7), one often starts from a mean-field theory where U_{ij} and a_{i0}^b are assumed to be space-time independent. To see the many possible mean-field ansatzes we refer the readers to, e.g., Ref.[30]. In the following, we shall focus on the so-called "π-flux phase mean-field theory"[31] for the nearest neighbor Heisenberg model.

The π-flux mean field theory assumes the following mean-field $\bar{U}_{ij}$ and $\bar{a}_{i0}^b$

$$\begin{aligned} &\bar{\Delta}_{ij} = 0, \quad \bar{a}_{i0}^b = 0, \\ &\bar{\chi}_{i,i+\hat{x}} = i\chi, \\ &\bar{\chi}_{i,j+\hat{y}} = i(-1)^{i_x}\chi \end{aligned} \tag{5.8}$$

where χ is a real parameter (see Fig. 5.1). This leads to the following fermion mean-field Hamiltonian,

$$\hat{H}_{\rm MF} = -\frac{3}{4} J \sum_i \Big\{ i\,\chi \left[F_{i+\hat{x}}^T \left(II\right) F_i \right] + i\left(-1\right)^{i_x} \chi \left[F_{i+\hat{y}}^T \left(II\right) F_i \right] + h.c. \Big\}$$

Because the Pauli matrices are identity in both the Majorana and spin spaces, this mean-field Hamiltonian enjoys both global spin- and charge-SU(2) symmetries generated by

$$\begin{aligned} &\text{Spin-SU(2) generators: } \Sigma^a = (YX, IY, YZ) \\ &\text{Charge-SU(2) generators: } T^a = (XY, YI, ZY)\,. \end{aligned} \tag{5.9}$$

Using the eigenvalues ± 1 of the "sub-lattice Pauli matrix" Z to label the blue and red sub-lattices in Fig. 5.1, and performing Fourier transform we obtain the following momentum-space mean-field Hamiltonian

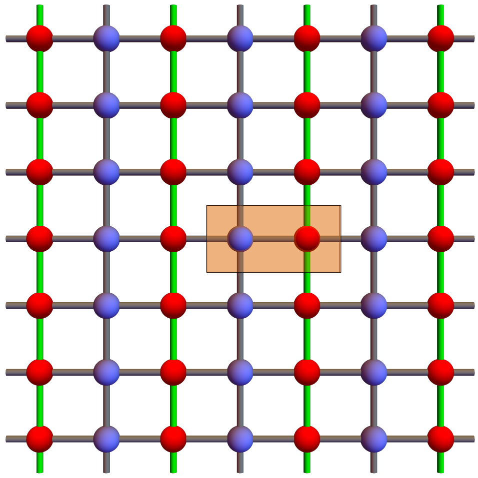


Figure 5.1: The π-flux mean-field theory. Here the black bonds represent hopping amplitude $i\chi$ in the positive x- or y-direction and the green bonds represent $-i\chi$. The unit cell is enclosed by the orange rectangle.

$$\begin{aligned}\hat{H}_{\rm MF} &= -\frac{3}{4}J\chi\sum_{\mathbf{k}} F^T_{-\mathbf{k}}\left[II\otimes\begin{bmatrix} i\left(e^{ik_2}-e^{-ik_2}\right) & -i+ie^{2ik_1} \\ i-ie^{-2ik_1} & -i\left(e^{ik_2}-e^{-ik_2}\right)\end{bmatrix}\right]F_{\mathbf{k}} \\ &= -\frac{3}{4}J\chi\sum_{\mathbf{k}} F^T_{-\mathbf{k}}\left[II\otimes\left(-\sin 2k_1\,X+(1-\cos 2k_1)\,Y-2\sin k_2\,Z\right)\right]F_{\mathbf{k}}.\end{aligned} \tag{5.10}$$

In the above equation the tensor product of Pauli matrices are ordered according to

$$\text{Majorana}\otimes\text{spin}\otimes\text{sub-lattice}.$$

In Eq.(5.10) the (halfed) Brillouin zone is

$$-\pi/2\le k_1<\pi/2,\quad -\pi\le k_2<\pi$$

and the Dirac nodes are situated at $\mathbf{k}_0=(0,0)$ and $(0,\pi)$, which are referred to as two "valleys" in the following.

Expand $\mathbf{k}=\boldsymbol{k_0}+\boldsymbol{q}$ around these two Dirac nodes, and Fourier transform (w.r.t. $\boldsymbol{q}$) back to the real space, we obtain the following low energy mean-field Hamiltonian

$$\hat{H}_{\rm MF}=\int d^2x\,\tilde{F}^T\left(-i\Gamma_i\partial_i\right)\tilde{F},$$

where

$$\begin{aligned}\Gamma_1 &= IIXI \\ \Gamma_2 &= IIZZ.\end{aligned} \tag{5.11}$$

The tensor product of four Pauli matrices in Eq.(5.11) are arranged according to

$$\text{Majorana} \otimes \text{spin} \otimes \text{sub-lattice} \otimes \text{valley}.$$

Including the sub-lattice and valley Pauli matrices the generators of the charge and spin SU(2) transformations are given by

$$\begin{aligned}&\text{Spin-SU(2) generators: } \Sigma^a = (YXII, IYII, YZII) \\ &\text{Charge-SU(2) generators: } T^a = (XYII, YIII, ZYII)\,.\end{aligned} \tag{5.12}$$

Because the local charge-SU(2) gauge degrees of freedom is a redundancy in the original half-filled Mott insulator, we expect the field theory in Eq.(5.7) to have the local charge-SU(2) symmetry. This motivates one to think the low energy theory, including fluctuations in U_{ij} and a^b_{i0}, is a charge-SU(2) gauge theory with

$$U_{ij} = \bar{U}_{ij} e^{i a_{ij}}$$

where $a_{ij} = a^b_{ij} T^b$ is the spatial component of the charge-SU(2) gauge field. According to Ref.[30, 32], because the mean-field $\bar{U}_{ij}$ commutes with the *global* charge-SU(2) transformations, the low theory is a charge-SU(2) gauge theory, with a_0 and a_{ij} playing the roles of the time and spatial components of the gauge field, respectively.

The partition function of the charge-SU(2) gauge theory reads

$$\begin{aligned}\mathcal{Z} &= \int D[\tilde{F}] D[a_\mu] e^{-S[\tilde{F}, a_\mu]} \\ S &= \int d^3x \Big\{ \tilde{F}^T \left[(\partial_0 + i a^a_0 T^a) - i \sum_{i=1}^{2} \Gamma^i (\partial_i + i a^a_i T^a) \right] \tilde{F} + \frac{1}{2g} f^2_{\mu\nu} \Big\}.\end{aligned} \tag{5.13}$$

In Eq.(5.13) the $\frac{1}{2g} f^2_{\mu\nu}$ is generated by integrating out the higher energy fermion degrees of freedom. The theory in Eq.(5.13) describes the $n = 8$ real class fermion theory coupled to a dynamic charge-SU(2) gauge field.

According to the bosonization in section 4.9.2, the bosonized theory is a gauged $\frac{O(8)}{O(4)\times O(4)}$ nonlinear sigma model with the $k = 1$ WZW term [1]. Here the charge-SU(2)

[1]Note that although for $n = 8$, the homotopy group are not yet stabilized, fermion integration still gives a WZW term. When $\mathcal{B}$ is a closed manifold, and after division by $2\pi i$, the WZW term is the topological invariant of one of the $\mathbb{Z}$ factor of the π_4.

subgroup of the fermion (emergent) global symmetry group $O(8)$ is gauged.

In the following let's *assume* that the effect of the SU(2) gauge field is to cause confinement [2]. Under such condition, the fermion-antifermion pair oder parameter (analogous to mesons in QCD) must be a charge-SU(2) singlet. Since $Q^{\mathbb{R}}$ is precisely the "meson" field, it follows that in the charge-SU(2) confined phase the finite energy $Q^{\mathbb{R}}$ are restricted to a sub-manifold of $\frac{O(8)}{O(4)\times O(4)}$ which are invariant under the charge-SU(2) transformation.[3,4]. This sub-manifold is the S^4 spanned by the following 5 mutually anti-commuting masses,

$$S^4 = \Big\{ \sum_{i=1}^{5} n_i M_i; \quad \sum_{i=1}^{5} n_i^2 = 1 \Big\}, \quad \text{where}$$
$$M_i = YXZX, IYZX, YZZX, IIYI, IIZY. \tag{5.14}$$

At this point it is important to stress that we are not implying the deconfined phase does not exist. We simply consider the scenario where the SU(2) gauge field is in the confined phase.

In order to match the gamma matrices and mass matrices convention in table 4.1 and 4.2 (based on which the non-linear sigma models in subsection 4.9 and appendix B.3 and B.6 are derived), we will make the following change the basis. We first exchange the order of the third and the fourth (i.e., sub-lattice and valley) Pauli matrices, followed by the following orthogonal transformation,

$$II \otimes \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix}$$

[2]Note that unlike the compact U(1) gauge field, here the confinement can be not caused by the proliferation of monopoles. This is based on the following homotopy argument. The $SU(2)$ gauge configurations on the space-time surface S^2 surrounding the location of the monopole are classified by the mapping classes of $S^2 \to BSU(2)$, where $BSU(2)$ is the classifying space of $SU(2)$. Using the following identity in algebraic topology,

$$[S^2, BSU(2)]_* = [\Sigma S^1, BSU(2)]_* = [S^1, SU(2)]_* = \pi_1(SU(2)) = 0$$

,it follows that there is no topologically non-trivial gauge field configuration on S^2, hence there is no monopole. Here Σ denotes "reduced suspension", and $[X_1, X_2]_*$ is the homotopy class of base-point-preserving maps $X_1 \to X_2$. Physically speaking, assuming the $SU(2)$ monopole exists, we can take the northern and southern hemispheres as the patches to define the gauge connection so that on each patch, the gauge field configuration is non-singular. On the equator, S^1, where the two patches overlap, a gauge transformation must relate the gauge fields originated from the two patches. At each point of S^1 the gauge transformation is an element in $SU(2)$. Therefore the monopole classification is given by the homotopy class of gauge transformation on the S^1, i.e., $\pi_1(SU(2))$.

[3]In addition to restricting $Q^{\mathbb{R}}$ to be invariant under charge-SU(2) transformations, the charge-SU(2) gauge fluctuations can also generate four-fermion interaction which could trigger spontaneous symmetry breaking.

[4]Our result is analogous to Witten's non-linear sigma model description of QCD in the color SU(3) confined phase[27].

In the new basis the gamma matrices and the mass terms become

$$\begin{aligned} \tilde{\Gamma}_1 &= IIIX \\ \tilde{\Gamma}_2 &= IIIZ \\ \tilde{M}_i &= YXYY, IYYY, YZYY, IIXY, IIZY \end{aligned} \tag{5.15}$$

These are consistent with the matrices shown in table table4.1 and table4.2, except a trivial exchange of the first and the last Pauli matrices. In this basis, the order parameter $Q^{\mathbb{R}}$ is defined by $\tilde{M} = m\, Q^{\mathbb{R}} \otimes Y$.

5.1.3 Antiferromagnet, Valence bond solid, and the "deconfined" quantum crtical point

For the mass manifold in Eq.(5.14), we expect the non-linear sigma model to have a WZW term because $\pi_4(S^4) = \mathbb{Z}$. Substituting

$$\begin{aligned} Q^{\mathbb{R}} &= n_i N_i \text{ where} \\ N_i &= (YXY, IYY, YZY, IIX, IIZ) \end{aligned} \tag{5.16}$$

into the non-linear sigma model given by Eq.(4.30) in subsection 4.9.2 we obtain

$$W[\hat{n}] = \frac{2}{\lambda_3} \int_{\mathcal{M}} d^3x \; (\partial_\mu n_i)^2 - \frac{2\pi i}{64\pi^2} \int_{\mathcal{B}} \epsilon^{ijklm} \tilde{n}_i \, d\tilde{n}_j \, d\tilde{n}_k \, d\tilde{n}_l \, d\tilde{n}_m. \tag{5.17}$$

This model has O(5) global symmetry generated by the pair-wise product of the matrices in $\tilde{M}_i$, which are also the generators of $O(8)$ that commutes with the charge-SU(2) generators. Hence Eq.(5.17) is often referred to as the "O(5)" non-linear sigma model in the literature [33, 34, 35, 36].

Now we address the physical meaning of the five masses given in Eq.(5.14) (or equivalently the physical meaning of $\tilde{M}_i$ in Eq.(5.15)). The first three of the masses in Eq.(5.14) correspond to the Néel order parameters, while the last two to the valence bond solid (VBS) orders. To see this, we first note that the first three masses rotate into each other under spin-SU(2),

$$\Sigma^a = (YXII, IYII, YZII)$$

while the last two are invariant.

We can also deduce the effect of translation by one lattice constant on these mass terms. In writing down the mean-field Hamiltonian we have chosen a particular charge-SU(2) gauge that explicitly breaks the symmetry associated with x-translation by one-lattice spacing. However, this is an artifact of gauge choice. The compounded transformation where the x-translation is followed by the gauge transformation which multiplies the fermion operator located on the orange rows in Fig. 5.2 by -1

$$(F_{\mathcal{I},1}, F_{\mathcal{I},2}) \xrightarrow{\hat{T}_{\hat{x}}} (-1)^{\mathcal{I}_y} \times (F_{\mathcal{I},2}, F_{\mathcal{I}+\hat{x},1})$$
$$(F_{\mathcal{I},1}, F_{\mathcal{I},2}) \xrightarrow{\hat{T}_{\hat{y}}} (F_{\mathcal{I}+\hat{y},1}, F_{\mathcal{I}+\hat{y},2})\,, \tag{5.18}$$

leaves the mean-field ansatz invariant. This is an example of "projective transformation". In Eq.(5.18) $\mathcal{I}$ label the unit cell in Fig. 5.1 , and we have omitted the Majorana and spin indices because they are unaffected by the translation.

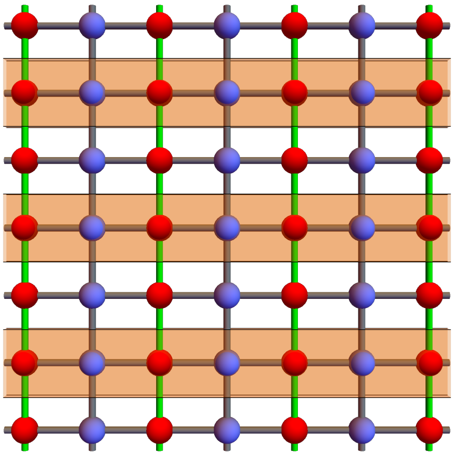


Figure 5.2: Translation by one lattice constant in the x-direction compounded with the gauge transformation which multiplies the fermion operators on sites in the orange rows by -1 leaves the mean-field Hamiltonian invariant.

In the following, we derive the effects of the "projective translation" on the fermion operator $\tilde{F}$ which is related to F via

$$F_{\mathcal{I},l} = \sum_{\text{small } \boldsymbol{q}} \left(\tilde{F}_{q,(l,1)} e^{i\boldsymbol{q}\cdot\mathcal{I}} + \tilde{F}_{q,(l,2)} e^{i((0,\pi)+\boldsymbol{q})\cdot\mathcal{I}}\right)$$

where the (l, v) are the indices for sub-lattice and valleys respectively. Doing the inverse Fourier transform, the above projective translation transforms $\tilde{F}$ according to

$$\tilde{F} \to T_{\hat{x},\hat{y}} \tilde{F}$$

where

$$T_{\hat{x}} = IIXX$$
$$T_{\hat{y}} = IIIZ. \tag{5.19}$$

Here we have put back the Majorana and spin (i.e., the first two) Pauli matrices.

Under $T_{\hat{x},\hat{y}}$ the mean-field Hamiltonian is invariant, but the first three mass terms change sign under $\hat{T}_{\hat{x}}$ and $\hat{T}_{\hat{y}}$ (as should the Néel order parameter) while the remaining two masses each breaks $\hat{T}_{\hat{x}}$ or $\hat{T}_{\hat{y}}$. These are the expected transformation properties of the VBS order parameters.

In appendix B.10 we show that the order parameters in Eq.(5.16) completely decouple from the charge-SU(2) gauge field. Thus even in the presence of such gauge field the non-linear sigma model preserves the form in Eq.(5.17). Before moving on, there is one additional thing worth mentioning, namely,

$$\pi_2(S^4) = 0.$$

Hence there is no soliton in the order parameter associated with Eq.(5.17).

In summary, we have found that after the charge-SU(2) confinement Eq.(5.17) describes the critical point between the AFM and VBS phases the so-called "deconfined quantum critical point"[33, 34, 35, 36, 37, 38]. It is also very satisfying that Eq.(5.17) captures the best-known spin-long-range ordered phase (AFM) and the spin quantum disorder phases (VBS)[39, 40, 41, 42, 43].[5]

5.2 The critical spin liquid of "bipartite Mott insulators" in $D = 1+1, 2+1$ and $3+1$.

The idea explained in the preceding section can be generalized to the insulating phase of "bipartite Mott insulators".

A bipartite Mott insulator is a Mott insulator whose lattice consists of two sub-lattices, and hoppings only occur between different sub-lattices. The nearest-neighbor spin-1/2 antiferromagnetic Heisenberg model in one spatial dimension describes the dynamics of spin degrees of freedom in a one-dimensional bipartite Mott insulator. It realizes the $SU(2)_1$ WZW non-linear sigma model, where the emergent symmetries are realized in a non-onsite (e.g., translation) fashion. This model serves as a paradigm of, e.g., quantum number fractionalization, and has profoundly influenced theoretical physics. It is natural to ask what is the generalization of this non-linear sigma model in the Mott insulating phase of higher dimensions. In the present section, we answer this question.

In a Mott insulating phase, the low energy degrees of freedom are the spins. Since the spin operators are invariant under the charge-SU(2) transformation, there are lots of

[5]Here the interaction favoring the Néel or the VBS type of symmetry breaking can be due to anisotropy terms that are omitted in our theory.

choices in fractionalizing the spin into spinons. Different choices are related by the spinon charge-SU(2) gauge transformation. The spin-spin interaction is generated by Anderson's super-exchange, the spinon mean-field theory amounts to choosing a spinon tight-binding model which reproduces the spin-spin interaction after super-exchange. Since the spin-spin interaction is independent of which charge-SU(2) gauge we choose, we shall choose the gauge so that the hoppings are purely imaginary in the following. The reason for doing so is because in such a gauge, the mean-field spinon Hamiltonian is charge-SU(2) invariant. This gauge choice exists when the Mott insulator is bipartite.

In section 5.1.2, we saw that the Mott insulating condition is imposed by the constraint that the order parameter $Q^{\mathbb{R}}$ is a charge-SU(2) singlet. In this section we show that imposing such constraints allows us to derive the spin effective theory in bipartite Mott insulators in spatial dimensions 1,2 and 3 [6].

5.2.1 (1+1)-D

For the nearest neighbor tight-binding model with real hopping in 1D, one can break the lattice into A, B sub-lattice and do the transformation $(c_j^A, c_j^B) \to (c_j^A, i\, c_j^B)$ to make the hopping purely imaginary (see figure 5.3). This leads to the lattice model

$$\begin{aligned}\hat{H} =& t\sum_k c_k^\dagger \left[I \otimes \begin{pmatrix} 0 & i + ie^{-ik} \\ -i - ie^{+ik} & 0 \end{pmatrix} \right] c_k \\ =& -t\sum_k c_k^\dagger \left[-(\sin k) IX + (1+\cos k) IY \right] c_k\end{aligned}$$

Here identity matrix I part acts on the spin. After linearizing around $k_F = \pi$, the low energy effective theory in Majorana fermion basis reads

$$H = \int dx\ \chi^T(x)\ \left[-i\Gamma_1 \partial_1 \right]\ \chi(x)$$

where

$$\Gamma_1 = IIX. \tag{5.20}$$

Here the tensor product of Pauli matrices are arranged according to

$$\text{Majorana} \otimes \text{spin} \otimes \text{sub-lattice}.$$

In the presence of Hubbard U, there is the global charge-SU(2) symmetry at half-filling. In the low energy theory, the charge-SU(2) transformation is generated by

$$\text{Charge-}SU(2) \text{ generators} : T^a = (XYI, YII, ZYI)$$

[6] Although we will not pursue it in the present paper, the discussion in the following can be generalized to the cases with larger flavor number or in higher dimensions.

On the other hand, the spin-SU(2) transformations are generated by the following charge-SU(2) invariant matrices,

$$\text{Spin-}SU(2)\text{ generators : }\Sigma^a = (YXI, IYI, YZI)$$

Following the discussion in section 4.8.2, the massless free fermion theory is equivalent to the $O(4)$ level-1 WZW model. Gauging the charge-SU(2) symmetry of the sigma model, and integrating over the gauge field, amounts to imposing the Mott insulating constraint. Assume the system is in the charge-SU(2) confined phase, only charge-SU(2) singlet order parameters (mass terms) can exist at low energies. These mass terms satisfy

$$\begin{aligned} &\{\Gamma_1, M\} = 0 \\ &[T^a, M] = 0 \\ &M^2 = I_8 \end{aligned}$$

The most general mass M satisfying the first two lines has the form

$$M = n_0\, IIY + n_1\, YXX + n_2\, IYX + n_3\, YZX \tag{5.21}$$

Among the mass terms

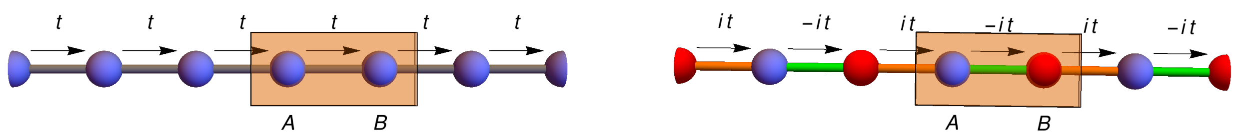
$$IIY, YXX, IYX, YZX,$$

the last three rotate into each other under the action of spin-SU(2) transformations, and the first one is invariant. They corresponds to the dimer and Néel order parameters respectively. The condition that $M^2 = I_8$ gives

$$\sum_{i=0}^{3} n_i^2 = 1.$$

The non-linear sigma model describing the fluctuations of $\hat{n}$ has a WZW term because $\pi_3(S^3) = \mathbb{Z}$, namely,

$$W[\hat{n}] = \frac{1}{4\pi}\int_{\mathcal{M}} d^2x\ (\partial_\mu \hat{n})^2 - \frac{2\pi i}{12\pi^2}\int_{\mathcal{B}} \epsilon^{ijkl}\tilde{n}_i\, d\tilde{n}_j\, d\tilde{n}_k\, d\tilde{n}_l. \tag{5.22}$$



(a) Real hopping

(b) Imaginary hopping

Figure 5.3: (a) The usual 1D nearest neighbor tight-binding with real hopping. (b) Upon the gauge transformation $(c_j^A, c_j^B) \to (c_j^A, i\, c_j^B)$, hoppings become purely imaginary with alternating sign. The hopping Hamiltonian in panel (b) is charge-SU(2) invariant.

This is the $SU(2)_1$ non-linear sigma model, known to be the effective theory of the Heisenberg spin chain[44].

5.2.2 $(2+1)$-D

In $(2+1)$-D we use the honeycomb lattice to write down the tight-binding model. The lattice vectors in the real and momentum space are

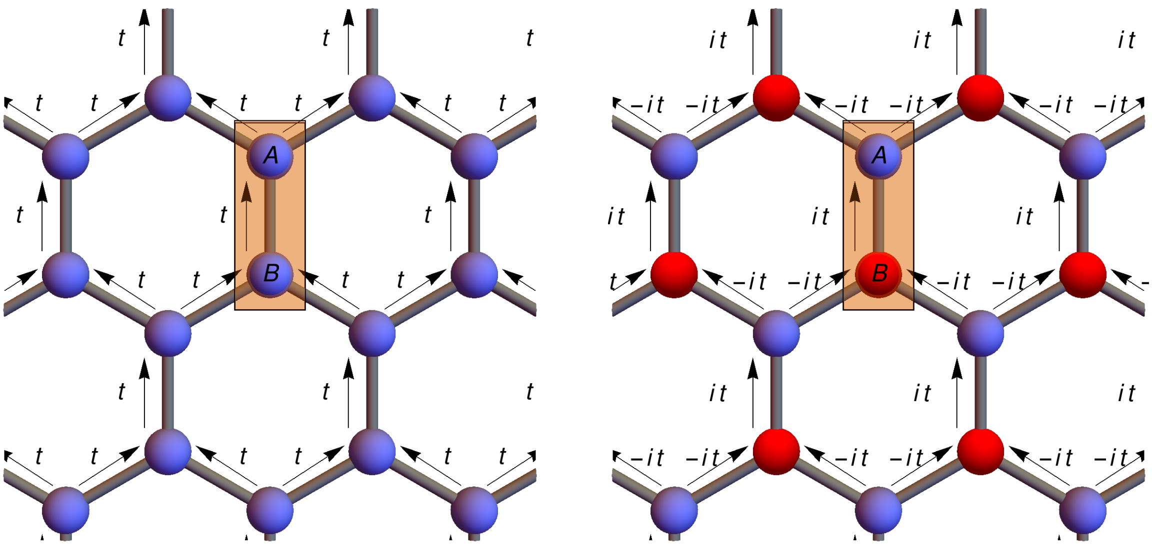
$$\boldsymbol{a}_1 = \sqrt{3}a\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \boldsymbol{a}_2 = \sqrt{3}a\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$$

and

$$\boldsymbol{b}_1 = \frac{4\pi}{3a}\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \quad \boldsymbol{b}_2 = \frac{4\pi}{3a}\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \tag{5.23}$$

respectively. In the following we perform the gauge transformation

$$(c_j^A, c_j^B) \to (c_j^A, i\, c_j^B)$$



(a) Real hopping (b) Imaginary hopping

Figure 5.4: (a) The usual nearest neighbor tight-binding model on the honeycomb lattice with real hopping. (b) Upon the gauge transformation $(c_j^A, c_j^B) \to (c_j^A, i\, c_j^B)$, hoppings become purely imaginary with alternating sign. The tight-binding Hamiltonian in panel (b) is charge-SU(2) invariant.

on the two sub-lattices, so that the nearest-neighbor hopping becomes purely imaginary (see figure 5.4). The tight-binding Hamiltonian read

$$\begin{aligned}\hat{H} =& t\sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}} \left[I \otimes \begin{pmatrix} 0 & i + ie^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + ie^{i\boldsymbol{k}\cdot\boldsymbol{a}_2} \\ -i - ie^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1} - ie^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2} & 0 \end{pmatrix} \right] c_{\boldsymbol{k}} \\ =& t\sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}} \Big[-\left(\sin(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \sin(\boldsymbol{k}\cdot\boldsymbol{a}_2)\right) IX - \left(1 + \cos(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \cos(\boldsymbol{k}\cdot\boldsymbol{a}_2)\right) IY \Big] c_{\boldsymbol{k}}\end{aligned}$$

Here the Pauli matrices are arranged according to

$$\text{spin} \otimes \text{ sub-lattice.}$$

In the Majorana fermion basis,

$$\hat{H} = t\sum_{\boldsymbol{k}} \chi^\dagger_{\boldsymbol{k}} \Big[-\left(\sin(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \sin(\boldsymbol{k}\cdot\boldsymbol{a}_2)\right) IIX - \left(1 + \cos(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \cos(\boldsymbol{k}\cdot\boldsymbol{a}_2)\right) IIY \Big] \chi_{\boldsymbol{k}} \tag{5.24}$$

Here the Pauli matrices are arranged according to

$$\text{Majorana} \otimes \text{spin} \otimes \text{sub-lattice.}$$

In the presence of repulsive Hubbard U, there is charge-SU(2) symmetry at half-filling. In the low energy theory the charge-SU(2) transformation is generated by

$$\text{Charge-SU(2) generators} : T^a = (XYI, YII, ZYI)$$

On the other hand, the spin-SU(2) transformations are generated by the following matrices,

$$\text{Spin-SU(2) generators} : \Sigma^a = (YXI, IYI, YZI)$$

Eq.(5.24) is invariant under both the charge- and spin-SU(2). In momentum space the Dirac points are located at K and $\hat{K}$ points, i.e., $\pm k_0$ where $k_0 := \frac{1}{3}(\boldsymbol{b}_1 - \boldsymbol{b}_2)$ (see Eq.(5.23)). Note that in the Majorana fermion basis, the contribution of Hamtiltonian from $\boldsymbol{k}$ and $-\boldsymbol{k}$ are the same due to the constraint $\chi^T_{-\boldsymbol{k}} = \chi^\dagger_{\boldsymbol{k}}$. This means that one can take the fermion $\chi_{\boldsymbol{k}_0+\boldsymbol{q}}$ around k_0 as the Fourier modes of complex fermion $\tilde{c}_{\boldsymbol{q}}$ and discard the other node. We then break this complex fermion $\tilde{c}$ into real fermion by $\tilde{c} = \tilde{\chi}_1 + i\tilde{\chi}_2$ (in the following we shall refer to this 1 and 2 as the "valley" indices). In this final Majorana representation, the low energy Hamiltonian reads

$$\hat{H} = \int d\boldsymbol{x}\ \tilde{\chi}^T(\boldsymbol{x})\ \left[-i\Gamma_1\partial_1 - i\Gamma_2\partial_2\right]\ \tilde{\chi}(\boldsymbol{x})$$

where $\Gamma_1 = IIIX$ and $\Gamma_2 = IIYY$. Here the Pauli matrices are arranged according to

$$\text{Majorana} \otimes \text{spin} \otimes \text{valley} \otimes \text{sub-lattice.}$$

In this basis, the symmetry generators are

$$\begin{aligned}&\text{Charge } SU(2) \text{ generators} : T^a = (XYII, YIII, ZYII)\\&\text{Spin } SU(2) \text{ generators} : \Sigma^a = (YXII, IYII, YZII).\end{aligned}$$

Following the discussions in section 4.9.2, the massless fermion theory is equivalent to the $\frac{O(8)}{O(4)\times O(4)}$ level-1 WZW model. Notice that the low energy fermion theory is identical to the π flux phase spinon mean-field theory discussed in section 5.1. Imposing the Mott constraint constraints the mass manifold. Specifically it requires the mass terms to commute with the charge-SU(2) generators. Under conditions the allowed mass terms satisfy

$$\begin{aligned}&\{\Gamma_i, M\} = 0\\&[T^a, M] = 0\\&M^2 = I_{16}\end{aligned}$$

The most general mass, $M \in \frac{O(8)}{O(4)\times O(4)}$, satisfying the first two equations has the form

$$M = n_1\, YXIZ + n_2\, IYIZ + n_3\, YZIZ + n_4\, IIXY + n_5\, IIZY$$

Similar to the discussion in section 5.1, the first three of the masses in Eq.(5.14) correspond to the Néel order parameters, while the last two to the valence bond solid (VBS) order parameters. The order parameter space forms an S^4. Plugging it into the $\frac{O(8)}{O(4)\times O(4)}$ level-1 WZW model, we arrive at the $O(5)$ WZW theory

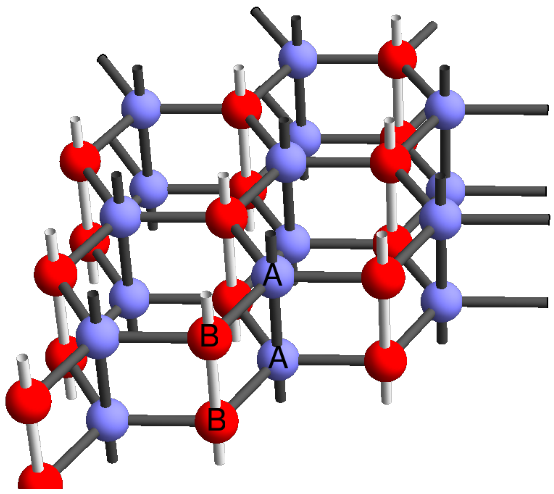
$$W[n_i] = \frac{2}{\lambda_3}\int\limits_{\mathcal{M}} d^3x\ (\partial_\mu n_i)^2 - \frac{2\pi i}{64\pi^2}\int\limits_{\mathcal{B}} \epsilon^{ijklm}\tilde{n}_i\, d\tilde{n}_j\, d\tilde{n}_k\, d\tilde{n}_l\, d\tilde{n}_m.$$

Here we note that because $\pi_2(S^4) = 0$ there is no soliton.

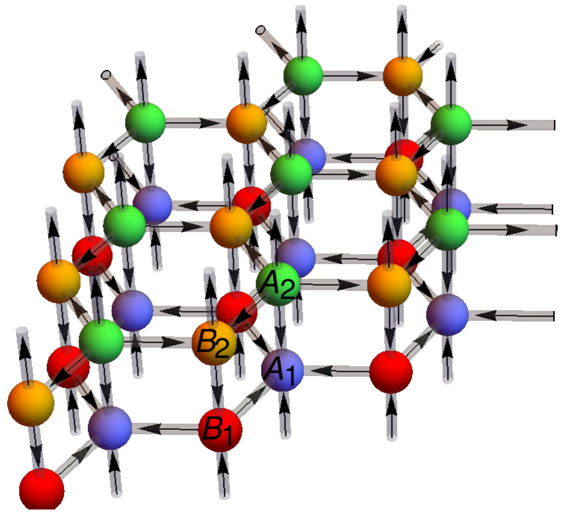
5.2.3 $(3+1)$-D

A lattice model for Dirac semi-metal

As a model for bipartite Mott insulator, we begin with a 3-dimensional tight-binding model consists of stacked honeycomb lattice. Here the lattice sites of each layer are stacked on top of those in the layer beneath. Within each layer, we have real hopping between the nearest-neighbor sites described in section 5.2.2. Between layers, the (real) hopping have the opposite sign for the A and B sub-lattice (see figure 5.5a). In order to



(a) Real hopping



(b) Imaginary hopping

Figure 5.5: (a) The tight-binding model on a stacked honeycomb lattice with real-valued nearest-neighbor hopping. Blue/red (A/B) mark the two sub-lattices of the honeycomb lattice respectively. The positive hoppings are drawn in black, while the negative hoppings in white.(b) After the gauge transformation $(c_j^{A_1}, c_j^{B_1}, c_j^{A_2}, c_j^{B_2}) \to (c_j^{A_1}, ic_j^{B_1}, ic_j^{A_2}, c_j^{B_2})$, a unit cell contains four sites. This is marked by blue/red/green/orange and labeled as $A_1/B_1/A_2/B_2$ respectively. The hoppings become purely imaginary. The direction of the arrows on the bonds label the direction of the imaginary hoppings. The tight-bonding Hamiltonian in panel (b) has charge-SU(2) symmetry.

make the hopping terms global charge-SU(2) invariant, we first enlarge the unit cell by grouping two adjacent layers to form A_1, B_1, A_2, B_2 sub-lattices as shown in figure 5.5b. We then perform the following gauge transformation,

$$(c_j^{A_1}, c_j^{B_1}, c_j^{A_2}, c_j^{B_2}) \to (c_j^{A_1}, ic_j^{B_1}, ic_j^{A_2}, c_j^{B_2}).$$

Here the lattice vectors in the real and momentum spaces are

$$\boldsymbol{a}_1 = \sqrt{3}\left(\frac{1}{2}, \frac{\sqrt{3}}{2}, 0\right), \quad \boldsymbol{a}_2 = \sqrt{3}\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0\right), \quad \boldsymbol{a}_3 = 3\,(0,0,1)$$

and

$$\boldsymbol{b}_1 = \frac{4\pi}{3}\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right), \quad \boldsymbol{b}_2 = \frac{4\pi}{3}\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right), \quad \boldsymbol{b}_3 = \frac{2\pi}{3}\,(0,0,1) \tag{5.25}$$

respectively. In the above we have assumed the magnitude of the hopping in the z-direction is the same as those within each layer. Moreover, we have tuned the lattice constant in the z-direction so that the Dirac cone is isotropic. The resulting tight-binding model reads

$$\hat{H} = i\,t \sum_{\boldsymbol{k}} c_{\boldsymbol{k}}^{\dagger} \cdot I \otimes \begin{pmatrix} 0 & S_{\mathrm{xy}}(\boldsymbol{k}) & S_{\mathrm{z}}(\boldsymbol{k}) & 0 \\ -S_{\mathrm{xy}}^{*}(\boldsymbol{k}) & 0 & 0 & S_{\mathrm{z}}(\boldsymbol{k}) \\ -S_{\mathrm{z}}^{*}(\boldsymbol{k}) & 0 & 0 & -S_{\mathrm{xy}}(\boldsymbol{k}) \\ 0 & -S_{\mathrm{z}}^{*}(\boldsymbol{k}) & S_{\mathrm{xy}}^{*}(\boldsymbol{k}) & 0 \end{pmatrix} c_{\boldsymbol{k}}$$
$$= t \sum_{\boldsymbol{k}} c_{\boldsymbol{k}}^{\dagger} I \otimes \left\{ \begin{matrix} -\left[\sin(\boldsymbol{k}\cdot a_1) + \sin(\boldsymbol{k}\cdot a_2)\right] ZX - \left[1 + \cos(\boldsymbol{k}\cdot a_1) + \cos(\boldsymbol{k}\cdot a_2)\right] ZY \\ + \sin(\boldsymbol{k}\cdot a_3) XI - \left[1 + \cos(\boldsymbol{k}\cdot \boldsymbol{a}_3) YI\right]. \end{matrix} \right\} c_{\boldsymbol{k}} \tag{5.26}$$

where the S_{xy} and S_z in Eq.(5.26) are defined as

$$S_{\mathrm{xy}}(\boldsymbol{k}) = 1 + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_2}$$
$$S_{\mathrm{z}}(\boldsymbol{k}) = 1 + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_3},$$

and the Pauli matrices are arranged according to

$$\text{spin} \otimes \text{ sub-lattice } (4\times 4).$$

It is simple to show that in the momentum space the Dirac points are located at $\pm\boldsymbol{k}_0$, where $\boldsymbol{k}_0 := \frac{1}{3}(\boldsymbol{b}_1 - \boldsymbol{b}_2)$.

Converting Eq.(5.26) into the Majorana fermion basis, the Hamiltonian reads

$$\hat{H} = t \sum_{\boldsymbol{k}} \chi_{-\boldsymbol{k}}^{T} II \otimes \left[\begin{matrix} -\left(\sin(\boldsymbol{k}\cdot a_1) + \sin(\boldsymbol{k}\cdot a_2)\right) ZX - \left(1 + \cos(\boldsymbol{k}\cdot a_1) + \cos(\boldsymbol{k}\cdot a_2)\right) ZY \\ + \sin(\boldsymbol{k}\cdot a_3) XI - \left(1 + \cos(\boldsymbol{k}\cdot \boldsymbol{a}_3) YI\right) \end{matrix} \right] \chi_{\boldsymbol{k}} \tag{5.27}$$

where the first Pauli matrix I acts in the Majorana space. The Hamiltonian in Eq.(5.27) is invariant under the global charge and spin $SU(2)$ transformations generated by

$$\begin{aligned} &\text{Charge-SU(2): } T^a = (XYII,\ YIII,\ ZYII) \\ &\text{Spin-SU(2): } \quad \Sigma^a = (YXII,\ IYII,\ YZII) \end{aligned} \tag{5.28}$$

When performing the mode expansion near $\pm\boldsymbol{k}_0$, because $\chi_{-\boldsymbol{k}}^{T} = \chi_{\boldsymbol{k}}^{\dagger}$, one can keep the complex fermion operator $\tilde{c}_{\boldsymbol{q}} = \chi_{\boldsymbol{k}_0+\boldsymbol{q}}$ while disregard the mode expansion near $-\boldsymbol{k}_0$. We subsequently break $\tilde{c}$ into real fermion operators $\tilde{c} = \tilde{\chi}_1 + i\tilde{\chi}_2$ (in the following we shall refer to this 1 and 2 as the "valley" indices). Omitting the tilde, in this final Majorana representation, the low energy theory of the Hamiltonian Eq.(5.27) is given by the $n = 8$ real class massless fermion Hamiltonian

$$\hat{H}_{\text{eff}} = \int d^3x\ \chi^T \left[-i \sum_{i=1}^{3} \Gamma_i \partial_i \right] \chi$$
$$\text{where } \Gamma_1 = IIZXI,\ \Gamma_2 = IIZYY,\ \Gamma_3 = IIXII, \tag{5.29}$$

and Eq.(5.28) is given by

$$\begin{aligned}&\text{Charge-SU(2):}\ T^a = (XYIII,\ YIIII,\ ZYIII)\\&\text{Spin-SU(2):}\quad \Sigma^a = (YXIII,\ IYIII,\ YZIII)\end{aligned}\tag{5.30}$$

In this basis, the Pauli matrices correspond to

$$\text{Majorana} \otimes \text{spin} \otimes \text{ sub-lattice } (4\times 4) \otimes \text{valley}.$$

For the gamma matrices to be in the standard basis used in table 4.1, we can do the transformation

$$\chi \to e^{i\frac{\pi}{4}IIZYI} \cdot e^{i\frac{\pi}{4}IIIXY}\chi,$$

and then switch between the third and the fifth Pauli matrices. In the new basis, Eq.(5.29) becomes

$$\begin{aligned}&\hat{H}_{\text{eff}} = \int d^3x\ \chi^T \left[-i\sum_{i=1}^{3}\Gamma_i\partial_i\right]\chi\\&\text{where } \Gamma_1 = IIIZI,\ \Gamma_2 = IIIXI,\ \Gamma_3 = IIIYY,\end{aligned}\tag{5.31}$$

while the symmetry generators in Eq.(5.30) remain unchanged. Upon bosonization, Eq.(5.31) is equivalent to the $U(8)/O(8)$ nonlinear sigma model in Eq.(4.33).

Imposing the Mott constraint

Following the discussion in section 5.2, the Mott constraint can be imposed by demanding the order parameter to be charge $SU(2)$ singlet. It is straightforward (but lengthy) to show that the following $Q^{\mathbb{R}}$ satisfies the charge-SU(2) singlet requirement

$$Q^{\mathbb{R}}(x) = e^{i\theta(x)}\left[n_0(x)\,N_0 + i\sum_{i=1}^{5}n_i(x)\,N_i\right] := e^{i\theta(x)}\mathcal{G}_S(x),\tag{5.32}$$

where

$$\begin{aligned}&N_0 = III,\ N_1 = IIZ,\ N_2 = IIX,\ N_3 = IYY,\ N_4 = YZY,\ N_5 = YXY\\&\text{and } \sum_{i=0}^{5}n_i^2 = 1, \text{i.e., } (n_0,n_1,n_2,n_3,n_4,n_5) \in S^5.\end{aligned}$$

In addition, in Eq.(5.32) $\mathcal{G}_S$ is a symmetric special unitary 8×8 matrix, namely,

$$\mathcal{G}_S(x) \in \frac{SU(8)}{O(8)}.$$

Substituting Eq.(5.32) into the bosonized nonlinear sigma model Eq.(4.33) and noting that

$$\frac{1}{i}Q^{\mathbb{R}\dagger}\partial_\mu Q^{\mathbb{R}} = \frac{1}{i}\mathcal{G}_S^\dagger\partial_\mu\mathcal{G}_S + \partial_\mu\theta,$$

the stiffness term becomes

$$\begin{aligned}
&\frac{1}{4\lambda_4^2}\int_{\mathcal{M}} d^4x \,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}}\partial^\mu Q^{\mathbb{R}\dagger}\right] \\
&= \frac{8}{4\lambda_4^2}\int_{\mathcal{M}} d^4x\ [\partial_\mu\theta\partial^\mu\theta] + \frac{1}{4\lambda_4^2}\int_{\mathcal{M}} d^4x\,\mathrm{tr}\left[\partial_\mu\mathcal{G}_S\partial^\mu\mathcal{G}_S^\dagger\right] \\
&= \frac{2}{\lambda_4^2}\int_{\mathcal{M}} d^4x\ [\partial_\mu\theta\partial^\mu\theta] + \frac{2}{\lambda_4^2}\int_{\mathcal{M}} d^4x\ \sum_{i=0}^{5}(\partial_\mu n_i)^2
\end{aligned} \tag{5.33}$$

The cross term vanishes because

$$\frac{1}{i}\mathrm{Tr}[\mathcal{G}_S^\dagger\partial_\mu\mathcal{G}_S] = 0. \tag{5.34}$$

Eq.(5.34) is due to the fact that $\mathcal{G}_S$ is a symmetric special unitary matrix hence $\in SU(n)$. As a result, the matrix part of $\frac{1}{i}\mathcal{G}_S^\dagger\partial_\mu\mathcal{G}_S$ can be decomposed into the generators $\{t^a\}$ of $su(n)$, which are traceless.

As to the WZW term it's can be shown that

$$\begin{aligned}
&-\frac{2\pi}{960\pi^3}\int\limits_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{R}\dagger}d\tilde{Q}^{\mathbb{R}}\right)^5\Big] = -\frac{2\pi}{960\pi^3}\int\limits_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{\mathcal{G}_S}^\dagger d\tilde{\mathcal{G}_S}\right)^5\Big] \\
&= -\frac{2\pi i}{120\pi^3}\int\limits_{\mathcal{B}} \epsilon^{i_1i_2i_3i_4i_5i_6}\tilde{n}_{i_1}d\tilde{n}_{i_2}d\tilde{n}_{i_3}d\tilde{n}_{i_4}d\tilde{n}_{i_5}d\tilde{n}_{i_6}.
\end{aligned} \tag{5.35}$$

(We shall prove this relation in appendix B.11.)

Putting together Eq.(5.33) and Eq.(5.35), the non-linear sigma model action is given by

$$\begin{aligned}
W[\theta,\boldsymbol{\beta}] \quad &= \frac{2}{\lambda_4^2}\int_{\mathcal{M}} d^4x\ [\partial_\mu\theta\partial^\mu\theta] + \frac{2}{\lambda_4^2}\int_{\mathcal{M}} d^4x\ \sum_{i=0}^{5}(\partial_\mu n_i)^2 \\
&-\frac{2\pi i}{120\pi^3}\int\limits_{\mathcal{B}} \epsilon^{i_1i_2i_3i_4i_5i_6}\tilde{n}_{i_1}d\tilde{n}_{i_2}d\tilde{n}_{i_3}d\tilde{n}_{i_4}d\tilde{n}_{i_5}d\tilde{n}_{i_6}.
\end{aligned} \tag{5.36}$$

Therefore unlike $(1+1)$- and $(2+1)$-D, the spin effective theory for $(3+1)$-D bipartite Mott insulator has an extra $U(1)$ mode!

Gapping out the $U(1)$ mode

In this subsection we show that there is a charge-SU(2) singlet fermion interaction term that gaps out the θ degree of freedom. For convenience, we use the basis in Eq.(5.31). The emergent symmetry is $U(n)$ which includes a subgroup $U(1)$ (not to be confused with the extra $U(1)$ mode discussed earlier) generated by

$$Q_{U(1)} = I_n \otimes IY.$$

We can use this $Q_{U(1)}$ to complexify the Majorana fermion [7] , namely,

$$\psi_i^\alpha := \frac{1}{\sqrt{2}} \left(\chi_{\alpha i 1} + i\, \chi_{\alpha i 2}\right). \tag{5.37}$$

Here the Majorana field $\chi_{\alpha i a}$ carries three indices: $\alpha = 1, 2, ..., n$ is the flavor index; i indexes the second Pauli matrix in Eq.(5.31) and $a = 1, 2$ indexes the last Pauli matrix. In terms of the complexified fermion operators the mass term reads (see table4.2)

$$\begin{aligned} &\chi^T \left[S_1 \otimes YX + S_2 \otimes YZ\right] \chi \\ &= \left[\psi_i^\alpha\, \left(i\, Y_{ij}\right) \left(S_1 + i\, S_2\right)_{\alpha\beta}\, \psi_j^\beta + h.c.\right] \\ &= \left[\psi_i^\alpha\, E_{ij} Q^{\mathbb{R}}_{\alpha\beta}\, \psi_j^\beta + h.c.\right] \end{aligned} \tag{5.38}$$

where S_1 and S_2 are symmetric real matrices.

Now we are ready to construct the desired interaction term to gap out the $U(1)$ mode in Eq.(5.32)

$$\begin{aligned} \hat{H}_{\text{int}} = -\frac{U_\theta}{2} \int d^4x \Big[&E_{i_1 j_1} E_{i_2 j_2} ... E_{i_n j_n} \left(\epsilon_{\alpha_1 \alpha_2 ... \alpha_n} \psi_{i_1}^{\alpha_1} \psi_{i_2}^{\alpha_2} ... \psi_{i_n}^{\alpha_n}\right) \\ &\times \left(\epsilon_{\beta_1 \beta_2 ... \beta_n} \psi_{j_1}^{\beta_1} \psi_{j_2}^{\beta_2} ... \psi_{j_n}^{\beta_n}\right) + \text{h.c.}\Big]. \end{aligned} \tag{5.39}$$

First we note that Eq.(5.39) is a charge-SU(2) singlet, hence is unaffected by the Mott constraint. The proof goes as follows. When acted upon by the charge-SU(2) transformation, the fermion operator in Eq.(5.37) transforms according to

$$\psi_i^\alpha \to u_\beta^\alpha \psi_i^\beta,$$

where u_β^α is the charge-SU(2) transformation matrix. Under such transformation, the term in each parenthesis of Eq.(5.39) transforms according to

$$\begin{aligned} \epsilon_{\alpha_1 \alpha_2 ... \alpha_n} \psi_{i_1}^{\alpha_1} \psi_{i_2}^{\alpha_2} ... \psi_{i_n}^{\alpha_n} &\to \epsilon_{\alpha_1 \alpha_2 ... \alpha_n} u_{\beta_1}^{\alpha_1} u_{\beta_2}^{\alpha_2} ... u_{\beta_n}^{\alpha_n} \psi_{i_1}^{\beta_1} \psi_{i_2}^{\beta_2} ... \psi_{i_n}^{\beta_n} \\ = (\det u)\, \epsilon_{\beta_1 \beta_2 ... \beta_n} \psi_{i_1}^{\beta_1} \psi_{i_2}^{\beta_2} ... \psi_{i_n}^{\beta_n} &= \epsilon_{\alpha_1 \alpha_2 ... \alpha_n} \psi_{i_1}^{\alpha_1} \psi_{i_2}^{\alpha_2} ... \psi_{i_n}^{\alpha_n}. \end{aligned}$$

[7] Note that although we have complexified the Majorana fermion using the emergent $U(1)$, this is different from the complex class because we allow the mass term to break this $U(1)$.

Therefore Eq.(5.39) is charge-SU(2) invariant.

Next, we note, upon bosonization

$$E_{ij}\psi_i^\alpha\psi_j^\beta \to Q^{\mathbb{R}}_{\alpha\beta} = (S_1 + iS_2)_{\alpha\beta}\,,$$

where

$$Q^{\mathbb{R}} \in \frac{U(8)}{O(8)}$$

is the order parameter of the nonlinear sigma model in Eq.(4.33). As the result, the action corresponds to $\hat{H}_{\text{int}}$ is

$$S_{\text{int}} = -\frac{U_\theta}{2}\int d^4x \left\{\det\left[Q^{\mathbb{R}}\right] + c.c\right\}. \tag{5.40}$$

Substituting Eq.(5.32) into Eq.(5.40) we obtain

$$S_{\text{int}} = -U_\theta \int d^4x \;\cos(8\theta). \tag{5.41}$$

Naively, it might appear that the θ-vacuum is 8-fold degenerate, corresponding to

$$\theta = \frac{2\pi l}{8} \text{ with } l = 0, 1, ...7.$$

However, this is due to a redundancy in the splitting $U(8)/O(8) \to U(1) \times SU(8)/O(8)$. The transformation

$$e^{i\theta} \to e^{i(\theta + \frac{2\pi}{8})}$$

can be absorbed by the following transformation of $\mathcal{G}_S$

$$\mathcal{G}_S \to e^{i\frac{2\pi}{8}}\mathcal{G}_S.$$

Because $\left(e^{i\frac{2\pi}{8}}\right)^8 = 1$, the transformed $\mathcal{G}_S$ still belongs to $SU(8)/O(8)$. As a result, the 8 different θ vacua should be counted as one, as long as there is no spontaneous symmetry breaking in $\mathcal{G}_S$ (i.e., when $\mathcal{G}_S(x)$ fluctuates over all possible configurations in $SU(8)/O(8)$).

In the phase that the θ degrees of freedom are gapped out, we have

$$Q^{\mathbb{R}}(x) = \left[n_0(x)\,N_0 + i\sum_{i=1}^{5} n_i(x)\,N_i\right]. \tag{5.42}$$

Among the six order parameters, the first three are spin-SU(2) singlet and the last three are spin-SU(2) triplet. The latter can be interpreted as the anti-ferromagnetic order parameters. As to the first three, they break the lattice rotation symmetry, and can be identified as the VBS order parameters. The non-linear sigma model governing the n_i degrees of freedom read

$$W[n_i] = \frac{2}{\lambda_4^2}\int_{\mathcal{M}} d^4x \sum_{i=0}^{5}(\partial_\mu n_i)^2 - \frac{2\pi i}{120\pi^3}\int_{\mathcal{B}} \epsilon^{i_1 i_2 i_3 i_4 i_5 i_6}\tilde{n}_{i_1} d\tilde{n}_{i_2} d\tilde{n}_{i_3} d\tilde{n}_{i_4} d\tilde{n}_{i_5} d\tilde{n}_{i_6}, \tag{5.43}$$

which is the S^5 (or $O(6)$) nonlinear sigma model with $k = 1$ WZW term. Note that since $\pi_3(S^5) = 0$, there is no soliton. This model is a natural generalization of the spin effective theory in $(1+1)$- and $(2+1)$-D.

5.3 Twisted bi-layer graphene

Another 2D system where relativistic electron dispersion comes into play is the twisted bilayer graphene (TBLG). When the twisting angle is close to the "magic" value, the relevant bands become very flat [45], which suggests strong correlation. Under that condition, as a function of band filling ν, a rich phase diagram emerges. This includes various insulating phases near integer filling and superconductivity when ν deviates from integer [46, 47, 48, 49, 50, 51]. In the following, we shall hold the point of view that the essence of the TBLG physics is the fact that the interaction energy overwhelms the bandwidth, which does not require the bandwidth to be zero. Therefore we restrict ourselves to twisting angles close but not exactly equal to the magic values.

In the non-interacting picture, the Fermi energy (E_F) only intersects the Dirac nodes at the charge neutral point $\nu = 0$. However, by measuring the electronic compressibility, it is recently suggested that the coincidence of E_F and Dirac nodes reappears at all integer filling factors [52, 53]. Such "Dirac revivals" is interpreted as the evidence of the unequal filling of bands induced by the polarization of the flavor (including valley and spin) degrees of freedom. Therefore the relativistic massless fermions and bosonized non-linear sigma models discussed in Part I are good starting points to address the physics of TBLG.

The real space structure of the TBLG is shown in Fig. 5.6a for a certain small but commensurate twisting angle. In Fig. 5.6b we show the associated momentum space structure. The large blue and the red hexagons are the original graphene Brillouin zones for the two layers. The small hexagons colored orange are the Brillouin zone of the Moiré superlattice. In Fig. 5.6c we blow up one of the Moiré Brillouin zones. Here K_M and K'_M labels the two valleys in the Moiré Brillouin zone, while the blue/red K and K' labels the valleys of the graphene Brillouin zone. Note that each valley of the Moiré Brillouin zone consists of two opposite valleys of the graphene Brillouin zone,

5.3.1 Charge neutral point $\nu = 0$

In the presence of inter-layer hybridization, there are eight "active" graphene-like bands. We can label these eight "flavors" by the flavor index which represents

graphene valley, Moiré valley, spin

degrees of freedom. At the charge neutral point, the Fermi level crosses the Dirac points at K_M and K'_M.

In the momentum space we expand the band dispersion around K_M and K'_M, the resulting low energy Dirac-like band structure is described by the following continuum real-space Hamiltonian

$$\hat{H} = \int d^2x \; \psi^\dagger(\boldsymbol{x}) \left(-i\Gamma_1 \partial_x - i\Gamma_2 \partial_y\right) \psi(\boldsymbol{x}), \tag{5.44}$$

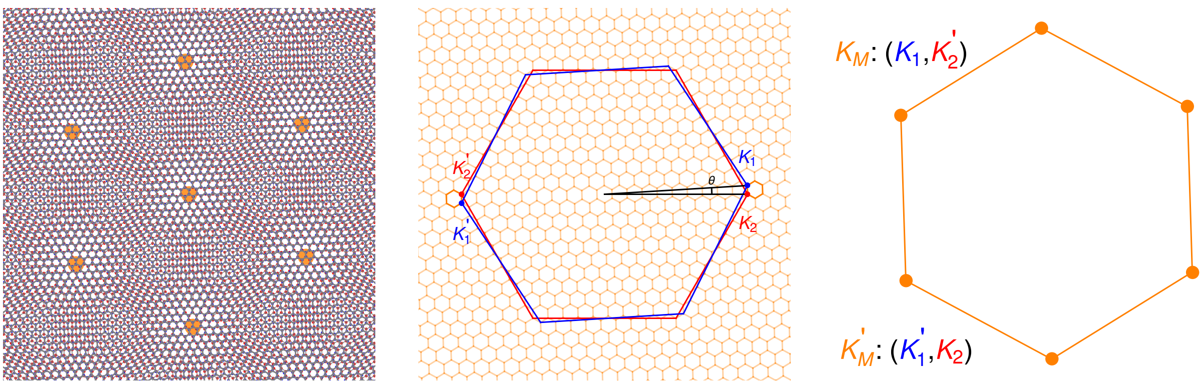
where ψ is an eight-component complex fermion field, and

$$\Gamma_1 = XZII, \quad \Gamma_2 = YIII. \tag{5.45}$$

Here the tensor product of Pauli matrices is arranged according to

sub-lattice $\otimes$ graphene valley $\otimes$ Moiré valley $\otimes$ spin.

The reason we use the complex fermion (rather than Majorana) representation in Eq.(5.44) is that at integer band fillings there is no evidence of superconductivity [48]. Therefore Eq.(5.44) belong to the complex class.



(a) Real space (b) Momentum space (c) Moiré Brillouin zone

Figure 5.6: (a) A real space picture of twisted bilayer graphene. (b) Blue and red color the Brillouin zones of the first and second layer graphene. Orange colors the Brillouin zone of the Moiré superlattice. (c) At K_M there are the K_1 of the first layer and K_2' of the first layer. At K'_M there are the K'_1 of the first layer and K_2 of the second layer.

The massless free fermion Hamiltonian in Eq.(5.44) has emergent $U(8)$ symmetry. After performing the the basis transformation

$$\psi \to e^{i\frac{\pi}{4}XIII} \cdot \left[\begin{pmatrix} I & 0 \\ 0 & Z \end{pmatrix} \otimes II\right] \psi$$

to cast the gamma matrices into the form used in table 4.1, namely,

$$\Gamma_1 = XIII, \quad \Gamma_2 = ZIII,$$

we can use our bosonization result (see appendix B.6). In the presence of the electromagnetic ($U(1)$) gauge field A, the massless fermion theory in Eq.(5.44) is equivalent to the following gauged non-linear sigma model

$$W[Q^{\mathbb{C}}, A] = \frac{1}{2\lambda_3} \int_{\mathcal{M}} d^3x \, \mathrm{tr}\Big[\left(\partial_\mu Q^{\mathbb{C}}\right)^2 \Big] - \frac{2\pi i}{256\pi^2} \Big\{ \int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}} \left(d\tilde{Q}^{\mathbb{C}}\right)^4\Big]$$
$$+8 \int_{\mathcal{M}} \mathrm{tr}\Big[iAQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2 - 2AFQ^{\mathbb{C}}\Big]\Big\}, \tag{5.46}$$

where

$$Q^{\mathbb{C}} \in \frac{U(8)}{U(4) \times U(4)}. \tag{5.47}$$

As discussed in section 4.10, there exists a local interacting fermion model which respects all emergent symmetries and the phases (which might spontaneously break the continuous or discrete symmetries) are described by the effective theories given by Eq.(5.46) but with

$$Q^{\mathbb{C}} \in \bigcup_{l=0}^{8} \frac{U(8)}{U(l) \times U(8-l)}$$

Among the last two terms of Eq.(5.46), the term linear in A_μ measures the soliton current

$$J^\mu = \frac{i}{16\pi} \epsilon^{\mu\nu\rho} \mathrm{tr}\Big[Q^{\mathbb{C}} \partial_\nu Q^{\mathbb{C}} \partial_\rho Q^{\mathbb{C}}\Big].$$

The term proportional to AF gives rise to a Chern-Simons term

$$-\frac{i}{8\pi} \int_{\mathcal{M}} \mathrm{tr}[Q^{\mathbb{C}}] AF,$$

with the corresponding Hall conductance

$$\sigma_{xy} = \frac{1}{2} \mathrm{tr}[Q^{\mathbb{C}}] = l - 4. \tag{5.48}$$

Therefore only the $l = 4$ mass manifold, $\frac{U(8)}{U(4)\times U(4)}$, has $\sigma_{xy} = 0$. Since so far there is no reported (non-zero) Hall conductivity at the charge neutral point [51], we take it as implying the relevant mass manifold is $\frac{U(8)}{U(4)\times U(4)}$.

The resulting non-linear sigma model has two phases depending on the coupling constant λ_3 in the stiffness term. For $\lambda_3 < \lambda_c$ there is a spontaneous breaking of the $U(8)$ symmetry, and the sigma model is gapped. We interpret this phase as the "symmetry-breaking insulator". For $\lambda_3 > \lambda_c$, there is a gapless phase for the non-linear sigma model, and we interpret that as the semi-metal phase. As far as we know, it is still not totally clear whether the low-temperature phase at $\nu = 0$ is a Dirac semimetal or a correlated charge insulator.

5.3.2 $\nu = \pm 1, \pm 2, \pm 3$

Experimentally a sequence of asymmetric jumps in the electronic compressibility are observed near integer filling factors[52, 53]. In Ref.[52] this is coined "Dirac fermion revivals", which is interpreted as due to "flavor polarization". In the following we shall assume this interpretation holds.

The mechanism of flavor polarization is likely due to a combination of Coulomb interaction and narrow bands, much like the occurrence of spin polarization (ferromagnetism) in narrow band metal. In the following, we shall assume the simplest flavor polarization mechanism. More complicated ones will not affect our discussions, as long as, after the polarization, the low energy spectrum forms Dirac cones and the number of active bands and the associated low-energy theory are captured correctly.

For simplicity, we shall consider $\nu \geq 0$ in the following discussion. In the cases of $\nu = 1, 2, 3$ [8], the Fermi level crossing band number is reduced to $3, 2, 1$ respectively. This can be caused by a polarization operator

$$\Delta_p \int d^2x \ \psi(\boldsymbol{x})^\dagger P \ \psi(\boldsymbol{x}) \tag{5.49}$$

where P is a hermitian matrix that commutes with Γ_i and satisfies $P^2 = I_{16}$. In addition, P needs to be identity matrix for the Moiré valley degree of freedom. This leads to the space

$$P \in \bigcup_{\nu=0}^{4} \frac{U(4)}{U(4-\nu)\times U(\nu)}, \tag{5.50}$$

Such a term will shift $4-\nu$ bands on the Moiré Brillouin zone upward and the remaining ν downward by the energy $\pm\Delta_p$. For example, $P = IZII$ is one such polarization matrix

[8] $\nu = -1, -2, 3$ can be mapped onto $\nu = 1, 2, 3$ by flipping the signs of Δ_p in Eq.(5.49) and ϵ_F in Fig. 5.7.

for $\nu = 2$ causing the polarization of graphene valleys, with half of the bands are shifted upward/downward. The resulting spectrum for each ν is schematically shown in Fig. 5.7.

After the polarization, the low energy free fermion Hamiltonian read

$$\hat{H} = \int d^2x \, \psi^\dagger(\boldsymbol{x}) \left(-i\Gamma_1^{(\nu)}\partial_x - i\Gamma_2^{(\nu)}\partial_y\right) \psi(\boldsymbol{x}), \tag{5.51}$$

where, up to a flavor basis transformation,

$$\Gamma_1^{(\nu)} = XII_{4-\nu}, \quad \Gamma_2^{(\nu)} = YII_{4-\nu}. \tag{5.52}$$

Here $I_{4-\nu}$ is the identity matrix for size $4-\nu$. The order parameter associated with the Fermi-level crossing bands is

$$Q^{\mathbb{C}} \in \bigcup_{l=0}^{8-2\nu} \frac{U(8-2\nu)}{U(l)\times U(8-2\nu-l)}, \tag{5.53}$$

and the associated non-linear sigma model reads

$$W[Q^{\mathbb{C}}, A] = \frac{1}{2\lambda_3}\int_{\mathcal{M}} d^3x \, \mathrm{tr}\Big[\left(\partial_\mu Q^{\mathbb{C}}\right)^2\Big] - \frac{2\pi i}{256\pi^2}\Big\{\int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\Big]$$
$$+8\int_{\mathcal{M}} \mathrm{tr}\Big[iAQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2 - 2AFQ^{\mathbb{C}}\Big]\Big\}. \tag{5.54}$$

where A_μ is the electromagnetic ($U(1)$) gauge field. Here, associated with each mass manifold the σ_{xy} is given by

$$\sigma_{xy} = l - (4-\nu).$$

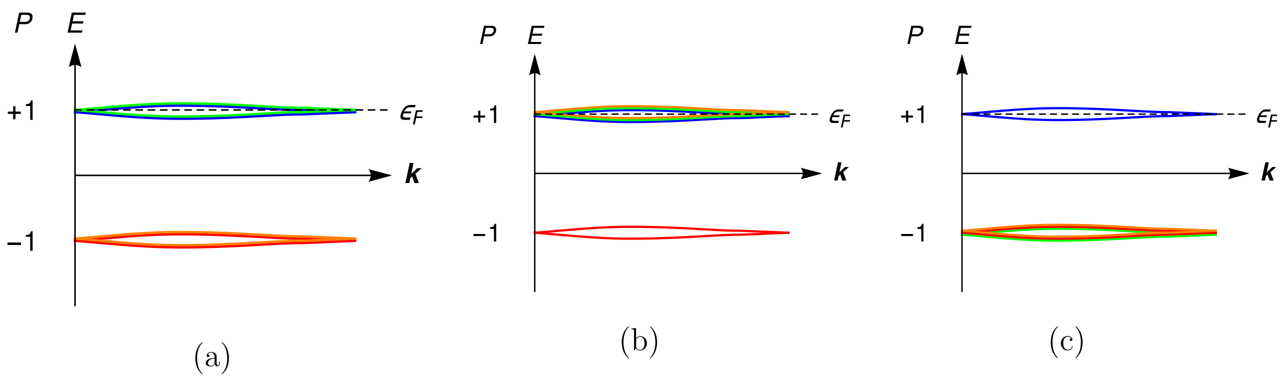


Figure 5.7: A caricature of the possible flavor polarization at (a) $\nu = 2$ and (b) $\nu = 1$ and (c) $\nu = 3$. Note that as long as the Fermi level intersects bands with the right degeneracy, the bands below the Fermi energy can overlap without changing the filling factor. For $\nu = -2, -1, -3$ we simply reflect the figures with respect to the x-axis.

First consider we $\nu = 1, 2$. Since experimentally $\sigma_{xy} = 0$ at $\nu = 1, 2$ [51] for $B = 0$, we take it as implying the relevant mass manifold is $\frac{U(8-2\nu)}{U(4-\nu)\times U(4-\nu)}$. The resulting non-linear sigma model can have two phases. One of phases occurs for $\lambda_3 < \lambda_c$, where there is a spontaneous breaking of the $U(8-2\nu)$ symmetry and the sigma model is gapped. We interpret this phase as the "symmetry-breaking correlated insulator"[9]. The other phase occurs for $\lambda_3 > \lambda_c$ where the sigma model remains gapless. We interpret that as the semi-metal phase.

For $\nu = 3$, the order parameter associated with the Fermi-level crossing bands is

$$Q^{\mathbb{C}} \in \bigcup_{l=0}^{2} \frac{U(2)}{U(l)\times U(2-l)}. \tag{5.55}$$

The $l = 2$ and $l = 0$ mass manifolds break the time-reversal symmetry and yield $\sigma_{xy} = \pm 1$ (see appendix B.8 for the details). Hence the phase corresponds to a quantum anomalous Hall state. This is consistent with the experimental observation of Ref.[51]. We stress that the non-zero σ_{xy} associated with mass manifold $l = 0$ or 2 is independent of the choice of flavor polarization P so long as it obeys Eq.(5.50).

The mass manifold associated with $l = 1$ is

$$\frac{U(2)}{U(1)\times U(1)} = S^2.$$

In that case $Q^{\mathbb{C}}$ can be replaced by a unit vector $\hat{n} \in S^2$. This leads to the bosonization of a small n case (i.e., before the WZW term is stabilized). The resulting nonlinear sigma model was first derived in Ref. [54] and reviewed in appendix B.8. The action is given by

$$W[\hat{n}] = \frac{1}{2\lambda_3'} \int_{\mathcal{M}} d^3x \ (\partial_\mu \hat{n})^2 + i\pi H[\hat{n}].$$

Here $H[\hat{n}]$ is the Hopf invariant of the $S^3 \to S^2$ mapping. In the presence of such Hopf term the solitons are fermions [55]. Depending on the value of λ_3' this non-linear sigma model can be gapless (preserving the $U(2)$ symmetry) for $\lambda_3' < \lambda_c$, or gapped (spontaneous symmetry breaking) for $\lambda_3' > \lambda_c$. In the latter case the fermionic solitons will be gapped. In either case $\sigma_{xy} = 0$. We viewed the gapped soliton phase a "correlated insulator" arising from symmetry breaking.

[9] Due to the flavor polarization, the original emergent symmetry is broken. Hence in principle, the low energy massless fermion theory can be regularized. If so there is the possibility that a Mott insulator phase exists.

Chapter 6

Conclusions of bosonization

In this part we have (non-abelian) bosonized two classes of massless fermion theories, the real and complex class, in spatial dimensions 1, 2, and 3. The boson theories are non-linear sigma models with the level-1 Wess-Zumino-Witten terms. We have also included three examples showing how to apply the bosonization results.

Of course, the goal of bosonization is not simply writing down theories equivalent to that of massless free fermions. For example, the bosonized models manifest what are the "nearby" symmetry-breaking states. These symmetry-breaking states can be reached when anisotropy terms are added to the non-linear sigma models. The bosonized theories also allow one to include the effects of strong interaction such as the charge-SU(2) confinement discussed in the first two applications. Moreover, as we have discussed, the main idea of this bosonization is inspired by the physics of topological insulators and superconductors. Indeed, the results discussed here can be applied to the boundary physics of such systems.

In this dissertation, when restoring the symmetries, we have restricted the bosonic order parameters to fluctuate smoothly. As the result, defect proliferation is not considered. In the literature, it is known that proliferation of symmetry-protected defects can lead to topological order (e.g., in Ref. [56]). However, in that case, one is restricted to the boundary of topological insulators/superconductors (or more generally symmetry-protected topological states). This is because defects are sensitive to short-distance physics, and the symmetries that protect the desired properties of defects can be broken by the regularization. Of course, unless the defects are on the boundary of an SPT, where regularization is provided by the bulk, and no symmetry breaking is necessary. An interesting question is how to reach a topological ordered state without invoking defects. These are directions that warrants more researches.

Bibliography

[1] Paul Adrien Maurice Dirac. "The quantum theory of the electron". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (Feb. 1928), pp. 610–624. ISSN: 0950-1207. DOI: 10.1098/rspa.1928.0023. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1928.0023.

[2] Stephen L. Adler. "Axial-Vector Vertex in Spinor Electrodynamics". In: *Physical Review* 177.5 (Jan. 1969), pp. 2426–2438. ISSN: 0031-899X. DOI: 10.1103/PhysRev.177.2426. URL: https://link.aps.org/doi/10.1103/PhysRev.177.2426.

[3] J. S. Bell and R. Jackiw. "A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-Model". In: *Il Nuovo Cimento A* 60.1 (Mar. 1969), pp. 47–61. ISSN: 0369-3546. DOI: 10.1007/BF02823296. URL: http://link.springer.com/10.1007/BF02823296.

[4] William A. Bardeen. "Anomalous Ward Identities in Spinor Field Theories". In: *Physical Review* 184.5 (Aug. 1969), pp. 1848–1859. ISSN: 0031-899X. DOI: 10.1103/PhysRev.184.1848. URL: https://link.aps.org/doi/10.1103/PhysRev.184.1848.

[5] H.B. Nielsen and M. Ninomiya. "A no-go theorem for regularizing chiral fermions". In: *Physics Letters B* 105.2-3 (Oct. 1981), pp. 219–223. ISSN: 03702693. DOI: 10.1016/0370-2693(81)91026-1. URL: https://linkinghub.elsevier.com/retrieve/pii/0370269381910261.

[6] C.G. Callan and J.A. Harvey. "Anomalies and fermion zero modes on strings and domain walls". In: *Nuclear Physics B* 250.1-4 (Jan. 1985), pp. 427–436. ISSN: 05503213. DOI: 10.1016/0550-3213(85)90489-4. URL: http://linkinghub.elsevier.com/retrieve/pii/0550321385904894%20https://linkinghub.elsevier.com/retrieve/pii/0550321385904894.

[7] Alexei Kitaev. "Periodic table for topological insulators and superconductors". In: *AIP Conference Proceedings* 1134.May 2009 (Jan. 2009), pp. 22–30. ISSN: 0094243X. DOI: 10.1063/1.3149495. arXiv: 0901.2686. URL: http://aip.scitation.org/doi/abs/10.1063/1.3149495%20http://arxiv.org/abs/0901.2686%20http://dx.doi.org/10.1063/1.3149495.

[8] Lokman Tsui, Yen-Ta Huang, and Dung-Hai Lee. "A holographic theory for the phase transitions between fermionic symmetry-protected topological states". In: *Nuclear Physics B* 949 (Dec. 2019), p. 114799. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2019.114799. arXiv: 1904.01544. URL: http://arxiv.org/

abs/1904.01544%20https://linkinghub.elsevier.com/retrieve/pii/S0550321319302858.

[9] Edward Witten. “Non-Abelian Bosonization in Two Dimensions”. In: *Communications in Mathematical Physics* 92.4 (Dec. 1984), pp. 455–472. ISSN: 0010-3616. DOI: 10.1007/BF01215276. URL: http://link.springer.com/10.1007/BF01215276.

[10] Shinsei Ryu and Shou-Cheng Zhang. “Interacting topological phases and modular invariance”. In: *Physical Review B* 85.24 (June 2012), p. 245132. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.85.245132. URL: https://link.aps.org/doi/10.1103/PhysRevB.85.245132.

[11] Olabode Mayodele Sule, Xiao Chen, and Shinsei Ryu. “Symmetry-protected topological phases and orbifolds: Generalized Laughlin’s argument”. In: *Physical Review B* 88.7 (Aug. 2013), p. 075125. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.88.075125. arXiv: 1305.0700. URL: https://link.aps.org/doi/10.1103/PhysRevB.88.075125.

[12] E. H. Brown and John Willard Milnor. “Topology from the Differentiable Viewpoint.” In: *The American Mathematical Monthly* 74.4 (Apr. 1967), p. 461. ISSN: 00029890. DOI: 10.2307/2314613. URL: https://www.jstor.org/stable/2314613?origin=crossref.

[13] Daniel C. Mattis and Elliott H. Lieb. “Exact Solution of a Many-Fermion System and Its Associated Boson Field”. In: *Journal of Mathematical Physics* 6.2 (Feb. 1965), pp. 304–312. ISSN: 0022-2488. DOI: 10.1063/1.1704281. URL: http://aip.scitation.org/doi/10.1063/1.1704281.

[14] Sidney Coleman. “Quantum sine-Gordon equation as the massive Thirring model”. In: *Physical Review D* 11.8 (Apr. 1975), pp. 2088–2097. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.11.2088. URL: https://link.aps.org/doi/10.1103/PhysRevD.11.2088.

[15] S. Mandelstam. “Soliton operators for the quantized sine-Gordon equation”. In: *Physical Review D* 11.10 (May 1975), pp. 3026–3030. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.11.3026. URL: https://link.aps.org/doi/10.1103/PhysRevD.11.3026.

[16] Shinsei Ryu et al. “Topological insulators and superconductors: tenfold way and dimensional hierarchy”. In: *New Journal of Physics* 12.6 (June 2010), p. 065010. ISSN: 1367-2630. DOI: 10.1088/1367-2630/12/6/065010. arXiv: 0912.2157. URL: https://iopscience.iop.org/article/10.1088/1367-2630/12/6/065010.

[17] G. ’t Hooft. “Naturalness, Chiral Symmetry, and Spontaneous Chiral Symmetry Breaking”. In: *Recent Developments in Gauge Theories*. Springer US, 1980, pp. 135–157. ISBN: 9781468475739. DOI: 10.1007/978-1-4684-7571-5_9. URL: http://link.springer.com/10.1007/978-1-4684-7571-5_9.

[18] K. G. Wilson. “Quarks and Strings on a Lattice”. In: *New Phenomena In Subnuclear Physics*. Ed. by A. Zichichi. Plenum Press, New York and London, 1977, p. 69. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[19] Yen-Ta Huang, Lokman Tsui, and Dung-Hai Lee. "The "non-regularizability" of gapless free fermion Hamiltonian protected by on-site symmetries". In: *Nuclear Physics B* 954 (May 2020), p. 115005. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2020.115005. arXiv: 1907.02522. URL: http://arxiv.org/abs/1907.02522%20https://linkinghub.elsevier.com/retrieve/pii/S0550321320300912.

[20] A. N. Redlich. "Gauge Noninvariance and Parity Nonconservation of Three-Dimensional Fermions". In: *Physical Review Letters* 52.1 (Jan. 1984), pp. 18–21. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.52.18. URL: https://link.aps.org/doi/10.1103/PhysRevLett.52.18.

[21] A. N. Redlich. "Parity violation and gauge noninvariance of the effective gauge field action in three dimensions". In: *Physical Review D* 29.10 (May 1984), pp. 2366–2374. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.29.2366. URL: https://link.aps.org/doi/10.1103/PhysRevD.29.2366.

[22] Chong Wang and T. Senthil. "Dual Dirac Liquid on the Surface of the Electron Topological Insulator". In: *Physical Review X* 5.4 (Nov. 2015), p. 041031. ISSN: 2160-3308. DOI: 10.1103/PhysRevX.5.041031. arXiv: 1505.05141. URL: https://link.aps.org/doi/10.1103/PhysRevX.5.041031.

[23] Max A. Metlitski and Ashvin Vishwanath. "Particle-vortex duality of two-dimensional Dirac fermion from electric-magnetic duality of three-dimensional topological insulators". In: *Physical Review B* 93.24 (June 2016), p. 245151. ISSN: 2469-9950. DOI: 10.1103/PhysRevB.93.245151. arXiv: 1505.05142 [cond-mat.str-el]. URL: https://link.aps.org/doi/10.1103/PhysRevB.93.245151.

[24] A.G. Abanov and P.B. Wiegmann. "Theta-terms in nonlinear sigma-models". In: *Nuclear Physics B* 570.3 (Mar. 2000), pp. 685–698. ISSN: 05503213. DOI: 10.1016/S0550-3213(99)00820-2. arXiv: 9911025 [hep-th]. URL: http://linkinghub.elsevier.com/retrieve/pii/S0550321399008202%20https://linkinghub.elsevier.com/retrieve/pii/S0550321399008202.

[25] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. 2nd. Cambridge University Press, 2012. ISBN: 978-0521548236. DOI: https://doi.org/10.1017/9781139020411.

[26] Edward Witten. "Global aspects of current algebra". In: *Nuclear Physics B* 223.2 (Aug. 1983), pp. 422–432. ISSN: 05503213. DOI: 10.1016/0550-3213(83)90063-9. URL: http://linkinghub.elsevier.com/retrieve/pii/0550321383900639%20https://linkinghub.elsevier.com/retrieve/pii/0550321383900639.

[27] Edward Witten. "Current algebra, baryons, and quark confinement". In: *Nuclear Physics B* 223.2 (Aug. 1983), pp. 433–444. ISSN: 05503213. DOI: 10.1016/0550-3213(83)90064-0. URL: https://linkinghub.elsevier.com/retrieve/pii/0550321383900640.

[28] P. W. Anderson. "Antiferromagnetism. Theory of Superexchange Interaction". In: *Physical Review* 79.2 (July 1950), pp. 350–356. ISSN: 0031-899X. DOI: 10.1103/PhysRev.79.350. URL: https://link.aps.org/doi/10.1103/PhysRev.79.350.

[29] Piers Coleman. “New approach to the mixed-valence problem”. In: *Physical Review B* 29.6 (Mar. 1984), pp. 3035–3044. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.29.3035. URL: https://link.aps.org/doi/10.1103/PhysRevB.29.3035.

[30] Patrick A. Lee, Naoto Nagaosa, and Xiao-Gang Wen. “Doping a Mott insulator: Physics of high-temperature superconductivity”. In: *Reviews of Modern Physics* 78.1 (Jan. 2006), pp. 17–85. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.78.17. arXiv: 0410445 [cond-mat]. URL: https://link.aps.org/doi/10.1103/RevModPhys.78.17.

[31] Ian Affleck and J. Brad Marston. “Large-n limit of the Heisenberg-Hubbard model: Implications for high-T_c superconductors”. In: *Physical Review B* 37.7 (Mar. 1988), pp. 3774–3777. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.37.3774. URL: https://link.aps.org/doi/10.1103/PhysRevB.37.3774.

[32] Elbio Dagotto, Eduardo Fradkin, and Adriana Moreo. “SU(2) gauge invariance and order parameters in strongly coupled electronic systems”. In: *Physical Review B* 38.4 (Aug. 1988), pp. 2926–2929. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.38.2926. URL: https://link.aps.org/doi/10.1103/PhysRevB.38.2926.

[33] T. Senthil et al. “Deconfined Quantum Critical Points”. In: *Science* 303.5663 (Mar. 2004), pp. 1490–1494. ISSN: 0036-8075. DOI: 10.1126/science.1091806. arXiv: 0311326 [cond-mat]. URL: http://arxiv.org/abs/cond-mat/0311326%0Ahttp://dx.doi.org/10.1126/science.1091806%20http://www.sciencemag.org/cgi/doi/10.1126/science.1091806%20https://www.sciencemag.org/lookup/doi/10.1126/science.1091806%20http://arxiv.org/abs/cond-mat/0311326%20http://dx.doi.o%20http://dx.doi.org/10.1126/science.1091806.

[34] Akihiro Tanaka and Xiao Hu. “Many-Body Spin Berry Phases Emerging from the π-Flux State: Competition between Antiferromagnetism and the Valence-Bond-Solid State”. In: *Physical Review Letters* 95.3 (July 2005), p. 036402. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.95.036402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.95.036402.

[35] T. Senthil and Matthew P. A. Fisher. “Competing orders, nonlinear sigma models, and topological terms in quantum magnets”. In: *Physical Review B* 74.6 (Aug. 2006), p. 064405. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.74.064405. arXiv: 0510459 [cond-mat]. URL: https://link.aps.org/doi/10.1103/PhysRevB.74.064405.

[36] Chong Wang et al. “Deconfined Quantum Critical Points: Symmetries and Dualities”. In: *Physical Review X* 7.3 (Sept. 2017), p. 031051. ISSN: 21603308. DOI: 10.1103/PhysRevX.7.031051. arXiv: 1703.02426. URL: http://arxiv.org/abs/1703.02426%20https://link.aps.org/doi/10.1103/PhysRevX.7.031051.

[37] Anders W. Sandvik. “Evidence for Deconfined Quantum Criticality in a Two-Dimensional Heisenberg Model with Four-Spin Interactions”. In: *Physical Review Letters* 98.22 (June 2007), p. 227202. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.98.227202. arXiv: 0611343 [cond-mat]. URL: https://link.aps.org/doi/10.1103/PhysRevLett.98.227202.

[38] Cenke Xu and Subir Sachdev. “Square-Lattice Algebraic Spin Liquid with SO(5) Symmetry”. In: *Physical Review Letters* 100.13 (Apr. 2008), p. 137201. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.100.137201. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.137201.

[39] F. D. M. Haldane. “O(3) Nonlinear σ Model and the Topological Distinction between Integer- and Half-Integer-Spin Antiferromagnets in Two Dimensions”. In: *Physical Review Letters* 61.8 (Aug. 1988), pp. 1029–1032. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.61.1029. URL: https://link.aps.org/doi/10.1103/PhysRevLett.61.1029.

[40] N. Read and Subir Sachdev. “Valence-Bond and Spin-Peierls Ground States of Low-Dimensional Quantum Antiferromagnets”. In: *Physical Review Letters* 62.14 (Apr. 1989), pp. 1694–1697. ISSN: 00319007. DOI: 10.1103/PhysRevLett.62.1694. URL: https://link.aps.org/doi/10.1103/PhysRevLett.62.1694.

[41] F. Figueirido et al. “Exact diagonalization of finite frustrated spin-$\frac{1}{2}$ Heisenberg models”. In: *Physical Review B* 41.7 (Mar. 1990), pp. 4619–4632. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.41.4619. URL: https://link.aps.org/doi/10.1103/PhysRevB.41.4619.

[42] Hong-Chen Jiang, Hong Yao, and Leon Balents. “Spin liquid ground state of the spin-$\frac{1}{2}$ square J_1-J_2 Heisenberg model”. In: *Physical Review B* 86.2 (July 2012), p. 024424. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.86.024424. URL: https://link.aps.org/doi/10.1103/PhysRevB.86.024424.

[43] Shou-Shu Gong et al. “Plaquette Ordered Phase and Quantum Phase Diagram in the Spin-$\frac{1}{2}$ J_1-J_2 Square Heisenberg Model”. In: *Physical Review Letters* 113.2 (July 2014), p. 027201. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.113.027201. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.027201.

[44] Ian Affleck and F. D. M. Haldane. “Critical theory of quantum spin chains”. In: *Physical Review B* 36.10 (Oct. 1987), pp. 5291–5300. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.36.5291. URL: https://link.aps.org/doi/10.1103/PhysRevB.36.5291.

[45] Rafi Bistritzer and Allan H. MacDonald. “Moiré bands in twisted double-layer graphene”. In: *Proceedings of the National Academy of Sciences* 108.30 (July 2011), pp. 12233–12237. ISSN: 0027-8424. DOI: 10.1073/pnas.1108174108. arXiv: 1009.4203. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.1108174108.

[46] Yuan Cao et al. “Correlated insulator behaviour at half-filling in magic-angle graphene superlattices”. In: *Nature* 556.7699 (Apr. 2018), pp. 80–84. ISSN: 0028-0836. DOI: 10.1038/nature26154. URL: http://dx.doi.org/10.1038/nature26154%20http://www.nature.com/articles/nature26154.

[47] Matthew Yankowitz et al. “Tuning superconductivity in twisted bilayer graphene”. In: *Science* 363.6431 (Mar. 2019), pp. 1059–1064. ISSN: 0036-8075. DOI: 10.1126/science.aav1910. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.aav1910.

[48] Xiaobo Lu et al. "Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene". In: *Nature* 574.7780 (Oct. 2019), pp. 653–657. ISSN: 0028-0836. DOI: 10.1038/s41586-019-1695-0. URL: http://dx.doi.org/10.1038/s41586-019-1695-0%20http://www.nature.com/articles/s41586-019-1695-0.

[49] Yuan Cao et al. "Unconventional superconductivity in magic-angle graphene superlattices". In: *Nature* 556.7699 (Apr. 2018), pp. 43–50. ISSN: 0028-0836. DOI: 10.1038/nature26160. arXiv: 1803.02342. URL: http://arxiv.org/abs/1803.02342%0Ahttp://dx.doi.org/10.1038/nature26160%20http://www.nature.com/articles/nature26160.

[50] Aaron L. Sharpe et al. "Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene". In: *Science* 365.6453 (Aug. 2019), pp. 605–608. ISSN: 0036-8075. DOI: 10.1126/science.aaw3780. arXiv: 1901.03520. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.aaw3780.

[51] M. Serlin et al. "Intrinsic quantized anomalous Hall effect in a Moiré heterostructure". In: *Science* 367.6480 (Feb. 2020), pp. 900–903. ISSN: 0036-8075. DOI: 10.1126/science.aay5533. arXiv: 1907.00261. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.aay5533.

[52] U. Zondiner et al. "Cascade of phase transitions and Dirac revivals in magic-angle graphene". In: *Nature* 582.7811 (June 2020), pp. 203–208. ISSN: 0028-0836. DOI: 10.1038/s41586-020-2373-y. arXiv: 1912.06150. URL: http://dx.doi.org/10.1038/s41586-020-2373-y%20http://www.nature.com/articles/s41586-020-2373-y.

[53] Dillon Wong et al. "Cascade of electronic transitions in magic-angle twisted bilayer graphene". In: *Nature* 582.7811 (June 2020), pp. 198–202. ISSN: 0028-0836. DOI: 10.1038/s41586-020-2339-0. URL: http://dx.doi.org/10.1038/s41586-020-2339-0%20http://www.nature.com/articles/s41586-020-2339-0.

[54] Alexander G Abanov. "Hopf term induced by fermions". In: *Physics Letters B* 492.3-4 (Jan. 2000), pp. 321–323. ISSN: 03702693. DOI: 10.1016/S0370-2693(00)01118-7. arXiv: 0005150 [hep-th]. URL: http://arxiv.org/abs/hep-th/0005150%20https://linkinghub.elsevier.com/retrieve/pii/S0370269300011187.

[55] Frank Wilczek and A. Zee. "Linking Numbers, Spin, and Statistics of Solitons". In: *Physical Review Letters* 51.25 (Dec. 1983), pp. 2250–2252. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.51.2250. URL: http://link.aps.org/doi/10.1103/PhysRevLett.51.2250%20https://link.aps.org/doi/10.1103/PhysRevLett.51.2250.

[56] Chong Wang, Andrew C. Potter, and T. Senthil. "Gapped symmetry preserving surface state for the electron topological insulator". In: *Physical Review B* 88.11 (Sept. 2013), p. 115137. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.88.115137. arXiv: 1306.3223. URL: https://link.aps.org/doi/10.1103/PhysRevB.88.115137.

[57] R. Bott. “THE STABLE HOMOTOPY OF THE CLASSICAL GROUPS”. In: *Proceedings of the National Academy of Sciences* 43.10 (Oct. 1957), pp. 933–935. ISSN: 0027-8424. DOI: 10.1073/pnas.43.10.933. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.43.10.933.

[58] Raoul Bott. “The Stable Homotopy of the Classical Groups”. In: *The Annals of Mathematics* 70.2 (Sept. 1959), p. 313. ISSN: 0003486X. DOI: 10.2307/1970106. URL: https://www.jstor.org/stable/1970106?origin=crossref.

[59] Allen Hatcher. *Algebraic Topology.* Cambridge University Press, 2001. ISBN: 978-0521795401.

[60] Ö. Kaymakcalan, S. Rajeev, and J. Schechter. “Non-Abelian anomaly and vector-meson decays”. In: *Physical Review D* 30.3 (Aug. 1984), pp. 594–602. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.30.594. URL: https://link.aps.org/doi/10.1103/PhysRevD.30.594.

[61] Heinz Hopf. “Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche”. In: *Mathematische Annalen* 104.1 (Dec. 1931), pp. 637–665. ISSN: 0025-5831. DOI: 10.1007/BF01457962. URL: http://link.springer.com/10.1007/BF01457962.

[62] F. D M Haldane. “Nonlinear Field Theory of Large-Spin Heisenberg Antiferromagnets: Semiclassically Quantized Solitons of the One-Dimensional Easy-Axis Néel State”. In: *Physical Review Letters* 50.15 (Apr. 1983), pp. 1153–1156. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.50.1153. URL: https://link.aps.org/doi/10.1103/PhysRevLett.50.1153.

[63] F.D.M. Haldane. “Continuum dynamics of the 1-D Heisenberg antiferromagnet: Identification with the $O(3)$ nonlinear sigma model”. In: *Physics Letters A* 93.9 (Feb. 1983), pp. 464–468. ISSN: 03759601. DOI: 10.1016/0375-9601(83)90631-X. URL: https://linkinghub.elsevier.com/retrieve/pii/037596018390631X.

[64] S. Elitzur and V.P. Nair. “Non perturbative anomalies in higher dimensions”. In: *Nuclear Physics B* 243.2 (Sept. 1984), pp. 205–211. ISSN: 05503213. DOI: 10.1016/0550-3213(84)90024-5. URL: https://linkinghub.elsevier.com/retrieve/pii/0550321384900245.

references.bib

Appendix A

Appendices of part II

A.1 The preservation of constraints 1 to 4 by the spectral symmetrization steps

In this section, we show that spectral symmetrization preserves the constraints 1 to 4. As discussed in the main text, if there is anti-unitary symmetry, the spectrum of $h(\boldsymbol{k})$ is symmetric about $E = 0$, in which case there is no need for the second step of spectral symmetrization, namely, subtracting the average of eigenenergies.

A.1.1 Constraint 1

The Majorana constraint implies the original Hamiltonian satisfies

$$h^T(-\boldsymbol{k}) = -h(\boldsymbol{k}) \tag{A.1}$$

This implies the eigenvalues at $-\boldsymbol{k}$ are the negative of the eigenvalues at $+\boldsymbol{k}$. Thus

$$D(-\boldsymbol{k}) = -W_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) W_{\boldsymbol{k}} \tag{A.2}$$

where $W_{\boldsymbol{k}}$ is the unitary transformation necessary to reorder the eigenvalues in $D(-\boldsymbol{k})$ according to descending order. The Majorana constraint of Eq.(A.1) implies

$$U_{-\boldsymbol{k}}^T D(-\boldsymbol{k}) U_{-\boldsymbol{k}}^* = -U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}}.$$

We substitute Eq.(A.2) into the above equation,

$$\begin{aligned} U_{-\boldsymbol{k}}^T W_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) W_{\boldsymbol{k}} U_{-\boldsymbol{k}}^* &= U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} \\ \Rightarrow U_{\boldsymbol{k}} U_{-\boldsymbol{k}}^T W_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) W_{\boldsymbol{k}} U_{-\boldsymbol{k}}^* U_{\boldsymbol{k}}^\dagger &= D(\boldsymbol{k}) \end{aligned} \tag{A.3}$$

The second line of the above equation can be rewritten as

$$Z_{\boldsymbol{k}} D(\boldsymbol{k}) Z_{\boldsymbol{k}}^\dagger = D(\boldsymbol{k}), \tag{A.4}$$

where the unitary matrix $Z_{\boldsymbol{k}} = U_{\boldsymbol{k}} U_{-\boldsymbol{k}}^T W_{\boldsymbol{k}}^\dagger$. In order for Eq.(A.4) to hold, $Z_{\boldsymbol{k}}$ needs to be block diagonalized where each block is spanned by the degenerate eigenvectors of $D(\boldsymbol{k})$.

Within each block, $D(\boldsymbol{k})$ is proportional to an identity matrix.

After the first step of spectral symmetrization $D(\boldsymbol{k}) \to D'(\boldsymbol{k})$. Since $D'(\boldsymbol{k})$ is still proportional to the same identity matrix in each block of $D(\boldsymbol{k})$, it follows that conjugation by $Z(\boldsymbol{k})$ still leaves it invariant, i.e.,

$$Z_{\boldsymbol{k}} D'(\boldsymbol{k}) Z_{\boldsymbol{k}}^{\dagger} = U_{\boldsymbol{k}} U_{-\boldsymbol{k}}^{T} W_{\boldsymbol{k}}^{\dagger} D'(\boldsymbol{k}) W_{\boldsymbol{k}} U_{-\boldsymbol{k}}^{*} U_{\boldsymbol{k}}^{\dagger} \quad = D'(\boldsymbol{k}). \tag{A.5}$$

Given Eq.(A.5) we can multiply the unitary matrices in the reverse order to arrive at

$$U_{-\boldsymbol{k}}^{T} D'(-\boldsymbol{k}) U_{-\boldsymbol{k}}^{*} = -U_{\boldsymbol{k}}^{\dagger} D'(\boldsymbol{k}) U_{\boldsymbol{k}},$$

which means

$$h'^{T}(-\boldsymbol{k}) \quad = -h'(\boldsymbol{k}). \tag{A.6}$$

Equation A.6 implies that the spectrum of $h'(\boldsymbol{k})$ flips sign upon the reversal of $\boldsymbol{k}$. As a result, the average of the diagonal elements $\bar{E}'(\boldsymbol{k})$ subtracted in the second step of spectral symmetrization, obeys

$$\bar{E}'(-\boldsymbol{k}) = -\bar{E}'(\boldsymbol{k}). \tag{A.7}$$

Consequently the subtracted piece $\bar{E}'(\boldsymbol{k}) I_n$ obeys the Majorana constraint, i.e.,

$$\left(\bar{E}'(-\boldsymbol{k}) I_n\right)^{T} = -\bar{E}'(\boldsymbol{k}) I_n. \tag{A.8}$$

This means if $h'(\boldsymbol{k})$ satisfies the Majorana constraint, so does $\tilde{h}(\boldsymbol{k})$ after the subtraction.

A.1.2 Constraint 2

The periodicity constraint is given by

$$h(\boldsymbol{k}) = h(\boldsymbol{k} + \boldsymbol{G}) \tag{A.9}$$

where $\boldsymbol{G}$ is any reciprocal lattice vector. This means

$$\begin{aligned} U_{\boldsymbol{k}}^{\dagger} D(\boldsymbol{k}) U_{\boldsymbol{k}} &= U_{\boldsymbol{k}+\boldsymbol{G}}^{\dagger} D(\boldsymbol{k}+\boldsymbol{G}) U_{\boldsymbol{k}+\boldsymbol{G}} \\ \Rightarrow D(\boldsymbol{k}+\boldsymbol{G}) &= U_{\boldsymbol{k}+\boldsymbol{G}} U_{\boldsymbol{k}}^{\dagger} D(\boldsymbol{k}) U_{\boldsymbol{k}} U_{\boldsymbol{k}+\boldsymbol{G}}^{\dagger}. \end{aligned}$$

Since $D(\boldsymbol{k}+\boldsymbol{G}) = D(\boldsymbol{k})$ (periodicity in Hamiltonian implies periodicity in the eigenvalues), we have

$$D(\boldsymbol{k}) = U_{\boldsymbol{k}+\boldsymbol{G}} U_{\boldsymbol{k}}^{\dagger} D(\boldsymbol{k}) U_{\boldsymbol{k}} U_{\boldsymbol{k}+\boldsymbol{G}}^{\dagger} \equiv Y_{\boldsymbol{k}} D(\boldsymbol{k}) Y_{\boldsymbol{k}}^{\dagger}. \tag{A.10}$$

Here the unitary matrix $Y_{\boldsymbol{k}} = U_{\boldsymbol{k}+\boldsymbol{G}} U_{\boldsymbol{k}}^{\dagger}$ needs to be block diagonalized where each block is spanned by the degenerate eigenvectors of $D(\boldsymbol{k})$. Within each block, $D(\boldsymbol{k})$ is proportional to an identity matrix. Since $D'(\boldsymbol{k})$ is still proportional to the same identity matrix in each block of $D(\boldsymbol{k})$, conjugation by $Y(\boldsymbol{k})$ leaves $D'(\boldsymbol{k})$ invariant. Thus

$$D'(\boldsymbol{k}) = U_{\boldsymbol{k}+\boldsymbol{G}} U_{\boldsymbol{k}}^{\dagger} D'(\boldsymbol{k}) U_{\boldsymbol{k}} U_{\boldsymbol{k}+\boldsymbol{G}}^{\dagger}.$$

Since $D'(\boldsymbol{k})$ on the LHS equals to $D'(\boldsymbol{k}+\boldsymbol{G})$, it follows that

$$D'(\boldsymbol{k}+\boldsymbol{G}) = U_{\boldsymbol{k}+\boldsymbol{G}} U_{\boldsymbol{k}}^{\dagger} D'(\boldsymbol{k}) U_{\boldsymbol{k}} U_{\boldsymbol{k}+\boldsymbol{G}}^{\dagger}.$$

Multiplying the unitary matrices in reverse order leads to

$$h'(\boldsymbol{k}) = h'(\boldsymbol{k}+\boldsymbol{G}).$$

Because $h'(\boldsymbol{k})$ satisfies the periodicity constraints, so does the average of its eigenvalues $\bar{E}'(\boldsymbol{k})$. Hence $\bar{E}'(\boldsymbol{k})I_n$, subtracted in the second step of spectral symmetrization, obeys the Brillouin zone periodicity. As as result, $\tilde{h}(\boldsymbol{k})$ satisfies the periodicity constraint.

Another important part of constraint 2 is the analytic nature of $h(\boldsymbol{k})$. In the following we show that in the $\boldsymbol{k}$ region where the spectrum is gapped, spectral symmetrization does not spoil analyticity.

Let's consider shifting $\boldsymbol{k}$ to $\boldsymbol{k}+\epsilon\hat{n}$, where $\hat{n}$ is an unit vector and ϵ is an infinitesimal. Under such infinitesimal shift

$$h(\boldsymbol{k}) \to h(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}}). \tag{A.11}$$

In the mathematics literature, e.g., theorem 1 in Chapter I (page 42) of Ref.[**Rellich1969**]), there is the following theorem.

> **Theorem** Let $h(x)$ be a finite dimensional Hermitian matrix function of a parameter x. If the polynomial expansion of $h(x)$ around $x = 0$ has a finite radius of convergence (i.e. analytic), then there exists a basis in which both the eigenvalues and the orthonormal set of eigenvectors of $h(x)$ have a convergent power series expansion within the same radius.

It is important to note that this theorem applies whether there are degeneracies in the eigenvalues of $h(0)$ or not.

Applying this theorem to our problem, the analytic nature of $h(\boldsymbol{k})$ around $\boldsymbol{k}$ implies the existence of a basis in which the eigenvalues and eigenvectors of $h(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}})$ is an analytic function of ϵ in any direction $\hat{n}$. This means we can choose a basis so that both $D(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}})$ and $U(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}})$ in

$$h(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}}) = U^{\dagger}(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}}) D(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}}) U(\boldsymbol{k}+\epsilon\hat{\boldsymbol{n}}) \tag{A.12}$$

are analytic functions of ϵ. In regions where $h(\boldsymbol{k})$ is gapped we can sort the eigenvalues into an upper half and a lower half so that no interchange of eigenvalues between the two parts take place as $\boldsymbol{k}$ moves around. Note that this is also true when there is crossing between the bands in the upper or lower halves (see Figure A.1 (a)).

Under such (no gap closure) condition, the spectral symmetrization does not change the analyticity of the Hamiltonian, because the eigenvectors are unchanged and the average of the upper/lower half of the eigenvalues as well as the average of all eigenvalues are analytic in $\boldsymbol{k}$. This is no longer true when $\boldsymbol{k}$ moves across a gap closing point (see Figure A.1 (b)). In that case there exists eigenvalues (and their associated eigenvectors) that move from the upper to the lower part (and vice versa). Under this condition although the original eigenvalues and eigenvectors are analytic in $\boldsymbol{k}$, *the sorted ones are not* (see Figure A.1 (c)).

Thus if h is analytic and gapped in a neighborhood of $\boldsymbol{k}$ the spectral symmetrised $\tilde{h}$ is analytic too. In contrast, spectral symmetrization does not maintain the analytic nature the Hamiltonian if $\boldsymbol{k}$ moves across gap nodes.

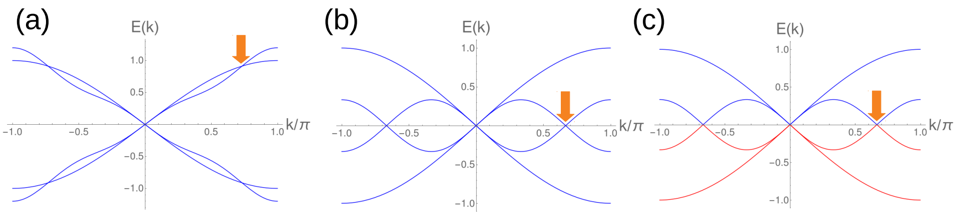


Figure A.1: Examples of band crossing at $k \neq 0$ in 1D. (a) The orange arrow points at a k point where band crossing occurs while the energy gap remains non-zero. (b) The orange arrow points at a gap-closing k point. (c) After the energy eigenvalues are sorted into upper (blue) and lower (red) halves, the eigenvalues and eigenvectors are no longer analytic across the gap-closing k point.

A.1.3 Constraint 3

The spectral symmetrization step clearly does not collapse the energy gap.

At $\boldsymbol{k} = 0$, since all eigen-energies are zero, no spectral symmetrization is necessary. Moreover, the spectral symmetrization does not change the fact that $h(\boldsymbol{k}) \to \sum_j k_j \Gamma_j$ as $\boldsymbol{k} \to 0$, because $\sum_j k_i \Gamma_j$ already satisfies the spectral symmetrization condition. Together, the above arguments imply that spectral symmetrization preserves constraint 3.

A.1.4 Constraint 4

The unitary symmetries

The unitary symmetries require

$$U_\beta^\dagger h(\boldsymbol{k}) U_\beta = h(\boldsymbol{k}),$$

which means

$$\begin{aligned} &U_\beta^\dagger U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} U_\beta = U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} \\ &\Rightarrow Q_{\boldsymbol{k}} D(\boldsymbol{k}) Q_{\boldsymbol{k}}^\dagger \equiv U_{\boldsymbol{k}} U_\beta^\dagger U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} U_\beta U_{\boldsymbol{k}}^\dagger = D(\boldsymbol{k}). \end{aligned} \tag{A.13}$$

Again, the unitary matrix $Q_{\boldsymbol{k}} = U_{\boldsymbol{k}} U_\beta^\dagger U_{\boldsymbol{k}}^\dagger$ needs to be block diagonalized where each block is spanned by the degenerate eigenvectors of $D(\boldsymbol{k})$. Within each block, $D(\boldsymbol{k})$ is proportional to an identity matrix.

After spectral symmetrization, $D'(\boldsymbol{k})$ is still proportional to the same identity matrix in each block of $D(\boldsymbol{k})$, hence conjugation by $Q(\boldsymbol{k})$ leaves $D'(\boldsymbol{k})$ invariant. Thus

$$U_{\boldsymbol{k}} U_\beta^\dagger U_{\boldsymbol{k}}^\dagger D'(\boldsymbol{k}) U_{\boldsymbol{k}} U_\beta U_{\boldsymbol{k}}^\dagger = D'(\boldsymbol{k}).$$

Multiplying the unitary matrices in reverse order leads to

$$U_\beta^\dagger h'(\boldsymbol{k}) U_\beta = h'(\boldsymbol{k}).$$

The $\bar{E}'(\boldsymbol{k}) I_n$, subtracted in the second step of spectral symmetrization, clearly satisfies the unitary symmetry constraint, namely,

$$U_\beta^\dagger \left(\bar{E}'(\boldsymbol{k}) I_n\right) U_\beta = \bar{E}'(\boldsymbol{k}) I_n.$$

As a result, the subtraction does not jeopardize the unitary symmetry.

The anti-unitary symmetries

The anti-unitary symmetries require

$$A_\alpha^\dagger h(-\boldsymbol{k})^* A_\alpha = h(\boldsymbol{k}).$$

The Majorana constraint Eq.(A.1) converts the above equation to

$$-A_\alpha^\dagger h(\boldsymbol{k}) A_\alpha = h(\boldsymbol{k}). \tag{A.14}$$

Among other things, this means the eigenvalues of $h(\boldsymbol{k})$ are in $\pm$ pairs, which means

$$D(\boldsymbol{k}) = X_{\boldsymbol{k}}^\dagger \left(-D(\boldsymbol{k})\right) X_{\boldsymbol{k}}. \tag{A.15}$$

Where $X_{\boldsymbol{k}}$ is a unitary matrix necessary to reorder the eigenvalues of $-D(\boldsymbol{k})$ in descending order. Equation A.14 implies

$$\begin{aligned} &A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} A_\alpha = -U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} \\ \Rightarrow\ &U_{\boldsymbol{k}} A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) U_{\boldsymbol{k}} A_\alpha U_{\boldsymbol{k}}^\dagger = -D(\boldsymbol{k}) \end{aligned} \tag{A.16}$$

Now we use Eq.(A.15) to convert the last line of the above equation to

$$O_{\boldsymbol{k}} D(\boldsymbol{k}) O_{\boldsymbol{k}}^\dagger \equiv U_{\boldsymbol{k}} A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger X_{\boldsymbol{k}}^\dagger D(\boldsymbol{k}) X_{\boldsymbol{k}} U_{\boldsymbol{k}} A_\alpha U_{\boldsymbol{k}}^\dagger = D(\boldsymbol{k})$$

Like before, the unitary matrix $O_{\boldsymbol{k}} = U_{\boldsymbol{k}} A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger X_{\boldsymbol{k}}^\dagger$ needs to be block diagonalized where each block is spanned by degenerate eigenvectors of $D(\boldsymbol{k})$. Within each block, $D(\boldsymbol{k})$ is proportional to an identity matrix.

After spectral symmetrization $D'(\boldsymbol{k})$ is still proportional to the same identity matrix in each block of $D(\boldsymbol{k})$, hence conjugation by $O(\boldsymbol{k})$ leaves $D'(\boldsymbol{k})$ invariant, i.e.,

$$U_{\boldsymbol{k}} A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger X_{\boldsymbol{k}}^\dagger D'(\boldsymbol{k}) X_{\boldsymbol{k}} U_{\boldsymbol{k}} A_\alpha U_{\boldsymbol{k}}^\dagger = D'(\boldsymbol{k}).$$

Since the same $X_{\boldsymbol{k}}$ can reverse the ordering of eigenvalues in $D'(\boldsymbol{k})$, the above equation turns into

$$U_{\boldsymbol{k}} A_\alpha^\dagger U_{\boldsymbol{k}}^\dagger \left(-D'(\boldsymbol{k})\right) U_{\boldsymbol{k}} A_\alpha U_{\boldsymbol{k}}^\dagger = D'(\boldsymbol{k}).$$

Multiplying the unitary matrices in reverse order leads to

$$-A_\alpha^\dagger h'(\boldsymbol{k}) A_\alpha = h'(\boldsymbol{k}). \tag{A.17}$$

Since Eq.(A.17) implies the eigenvalues of $h'(\boldsymbol{k})$ are symmetric with respect to $E = 0$, there is no subtraction step needed. Hence $\tilde{h}(\boldsymbol{k}) = h'(\boldsymbol{k})$, and

$$-A_\alpha^\dagger \tilde{h}(\boldsymbol{k}) A_\alpha = \tilde{h}(\boldsymbol{k}) \Rightarrow A_\alpha^\dagger \tilde{h}(-\boldsymbol{k})^* A_\alpha = \tilde{h}(\boldsymbol{k}).$$

A.2 Impossibility for the gap of $\tilde{h}(\boldsymbol{k})$ to close at only a single point in the Brillouin zone

In this appendix, we prove that the symmetry protection constraint plus constraints 1,2,4 and Eq.(2.2) in section 2.2 lead to the violation of the single gap node assumption in constraint 3. More specifically, we prove that under the conditions described above $A(\boldsymbol{k}_0) = 0$. Here $A(\boldsymbol{k})$ is defined in Eq.(2.12) and $\boldsymbol{k}_0$ is the non-zero time reversal invariant point discussed in the main text. Since $S(\boldsymbol{k}_0) = 0$ this implies $\tilde{h}(\boldsymbol{k}_0) = 0$. This violates the statement that energy gap closes only at $\boldsymbol{k} = 0$. This proof addresses the generic situations discussed in section 2.7 of the main text.

For each symmetry group generated by a subset of $\{\hat{T}, \hat{Q}, \hat{C}\}$ we focus on the minimal models where the number of components, n_0, in $\chi(\boldsymbol{k})$ is the minimum. This is the minimal number of components required to realize a particular SPN. Under this condition the dimension of all associated matrices is $n_0 \times n_0$. In other words, n_0 is the minimum integer for which there exists n_0-by-n_0 matrices representing the available symmetries and $\Gamma_1, ..., \Gamma_d$ (d is the spatial dimension). Here the $\{\Gamma_i\}$ obey the symmetry requirement (constraint 4 in section 2.2) and satisfy the Clifford algebra $\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}$.

For each spatial dimension d, we will go through all the symmetry groups G which gives rise to an SPN. (These groups protect non-trivial SPT's in $d+1$ dimensions.) For each (d, G) we write down the number n_0, the symmetry matrices, and the most general form of $S(\boldsymbol{k})$ and $A(\boldsymbol{k})$ allowed by symmetry.

To characterize each symmetry group we shall use the short hand

$$G^{\pm}([\ ]_{\pm}, [\ ]_{\pm}, [\ ]_{\pm}).$$

Between the square brackets we insert T, Q or C (the maximal number of symbols in the argument of G is 3). The subscript of the symbols, when present, denotes whether the matrix representing the $\hat{T}, \hat{Q}, \hat{C}$ squares to identity or minus identity. The superscript on G specifies whether the time reversal matrix T commutes ($+$) or anticommutes ($-$) with the charge conjugation matrix C. The matrix Q always anticommutes with T and C, and always squares to minus identity. Hence we do not bother to attach a subscript to Q, nor do we need to specify the commutator between Q and T, C. To simplify the notation we shall abbreviate the Pauli matrices $\sigma_0, \sigma_x, \sigma_y, \sigma_z, i\sigma_y$ as I, X, Y, Z, E, respectively. When two Pauli matrices appear next to each other it means tensor product. For example EX means $i\sigma_y \otimes \sigma_x$.

The proof is based on the following facts.

1. After spectral symmetrization, the Hamiltonian is given by Eq.(2.12) in the main text, where $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ and $S(\boldsymbol{k})^2 + A(\boldsymbol{k})^2 \propto I_n$.

2. As shown in A.1.2 the spectral symmetrization preserves the analytic nature of $h(\boldsymbol{k})$ in regions of $\boldsymbol{k}$ where the spectrum of $h(\boldsymbol{k})$ is fully gapped. Hence in the gapped region of $h(\boldsymbol{k})$, $\tilde{h}(\boldsymbol{k})$ and the coefficient functions $\tilde{o}_i(\boldsymbol{k})$ and $\tilde{e}_j(\boldsymbol{k})$ in Eq.(2.12) are analytic.

3. The Poincaré -Hopf theorem implies the mapping degree of

$$\{\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), ..., \tilde{o}_d(\boldsymbol{k})\}$$

   is odd around, at least, one other time-reversal invariant point $\boldsymbol{k}_0 \neq 0$.

In addition, for the ease of later discussions, we define the curves $\{\mathcal{C}_i,\ i = 1, ..., d\}$ near $\boldsymbol{k}_0$ as follows.

Definition 1. *Given $i \in \{1, \ldots, d\}$, let's consider the map $\boldsymbol{q} \to (\tilde{o}_1(\boldsymbol{k}_0 + \boldsymbol{q}), \ldots, \tilde{o}_d(\boldsymbol{k}_0 + \boldsymbol{q}))$ from any circle of radius $|\boldsymbol{q}| = r > 0$. Due to the non-zero degree of this map there must exist, at least, one point $\boldsymbol{q}$ on the circle such that $\tilde{o}_j(\boldsymbol{k}_0 + \boldsymbol{q}) = 0$ for $j \neq i$ and $\tilde{o}_i(\boldsymbol{k}_0 + \boldsymbol{q}) > 0$. Let's select such a point. Because the coefficient functions are continuous we can connect the points for different r into a curve $\mathcal{C}_i$ which approaches the point $\boldsymbol{k}_0$ as $r \to 0$.*

A.2.1 1D SPNs

$G(\emptyset)$**,** or equivalently $G(C_+)$ after block-diagonalizing C

$$G(\emptyset), \quad n_0 = 1, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k) \\ \hline A(k) & 0 \end{array}$$

As mentioned in the main text, this is a chiral SPN. It is not regularizable because the continuity and the Brillouin zone periodicity contradict with each other.

$G(T_-)$**,** or equivalently $G^+(T_-, C_+)$ after block-diagonalizing C

$$T = E, \quad n_0 = 2, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)X + \tilde{o}_2(k)Z \\ \hline A(k) & 0 \end{array}$$

Since there is no $A(k) \implies A(k_0) = 0$.

$G(C_-)$**,** or equivalently $G(Q)$ after identifying C with Q

$$C = E, \quad n_0 = 2, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)I \\ \hline A(k) & \tilde{e}_1(k)Y \end{array}$$

Since $S(\boldsymbol{k}) \propto I$ this is a chiral SPN. It is not regularizable because the continuity and the Brillouin zone periodicity contradict with each other.

$G^-(T_+, C_+)$

$$T = Z, C = X, \quad n_0 = 2, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)X \\ \hline A(k) & 0 \end{array}$$

Since there is no $A(k) \implies A(k_0) = 0$.

$G^{-}(T_{-}, C_{+})$

$$T = E, C = Z, \quad , n_0 = 2, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)Z \\ \hline A(k) & 0 \end{array}$$

Since there is no $A(k) \implies A(k_0) = 0$.

$G^{-}(T_{-}, C_{-})$, or equivalently $G(Q, T_{-})$ after identifying C with Q

$$T = ZE, C = EI, \quad n_0 = 4, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)YY + \tilde{o}_2(k)IX + \tilde{o}_3(k)IZ \\ \hline A(k) & \tilde{e}_1(k)YI \end{array}$$

$\{S(k), A(k)\} = 0$ implies

$$\begin{cases} \tilde{o}_1(k)\tilde{e}_1(k) = 0 \\ \tilde{o}_2(k)\tilde{e}_1(k) = 0 \\ \tilde{o}_3(k)\tilde{e}_1(k) = 0 \end{cases}$$

Because the mapping degree of $\tilde{o}_1(k)$ is odd in the neighborhood of $k = k_0$, it requires $\tilde{o}_1(k)$ to be non-zero when k is in the neighborhood but not equal to k_0. This implies $\tilde{e}_1(k) = 0$ in the neighborhood of k_0. The continuity of $\tilde{e}_1(k)$ implies $\tilde{e}_1(k_0) = 0$, which in turn implies $A(k_0) = 0$.

$G(Q, C_{+})$

$$Q = E, C = Z, \quad n_0 = 2, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)I \\ \hline A(k) & 0 \end{array}$$

Since $S(\boldsymbol{k}) \propto I$ this is a chiral SPN. It is not regularizable because the continuity and the Brillouin zone periodicity contradict with each other.

$G(Q, C_{-})$

$$Q = EI, C = ZE, \quad n_0 = 4, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)II \\ \hline A(k) & \tilde{e}_1(k)YX + \tilde{e}_2(k)YZ + \tilde{e}_3(k)IY \end{array}$$

Since $S(\boldsymbol{k}) \propto II$ this is a chiral SPN. It is not regularizable because the continuity and the Brillouin zone periodicity contradict with each other.

$G^+(Q, T_-, C_+)$ **or equivalently** $G^-(Q, T_-, C_+)$ **after identifying** C **with** QC

$$Q = EI, T = ZE, C = XX, \quad n_0 = 4, \quad \begin{array}{c|c} S(k) & \tilde{o}_1(k)YY + \tilde{o}_2(k)IX \\ \hline A(k) & 0 \end{array}$$

Since there is no $A(k) \implies A(k_0) = 0$.

A.2.2 2D SPNs

$G(T_-)$**,** or equivalently $G^+(T_-, C_+)$ after block-diagonalizing C

$$T = E, \quad n_0 = 2, \quad \begin{array}{c|c} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})X + \tilde{o}_2(\boldsymbol{k})Z \\ \hline A(\boldsymbol{k}) & 0 \end{array}$$

Since there is no $A(k) \implies A(k_0) = 0$.

$G^+(T_+, C_-)$ **or equivalently,** $G^+(T_-, C_-)$ **after identifying** T_- **with** T_+C_-

$$T = ZI, C = ZE, \quad n_0 = 4, \quad \begin{array}{c|c} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})XX + \tilde{o}_2(\boldsymbol{k})XZ \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YX + \tilde{e}_2(\boldsymbol{k})YZ \end{array}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = \tilde{o}_1(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k})$$

We examine the above equation in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$.

On the curve $\mathcal{C}_1$ defined in A.2 with $d = 2$, for any $r = |\boldsymbol{q}| \neq 0$,

$$0 = \tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = \tilde{o}_1(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) \tag{A.18}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_2(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_2(\boldsymbol{k})$ we conclude $\tilde{e}_2(\boldsymbol{k}) = 0$ at $r = 0$. In other words $\tilde{e}_2(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$. This will lead to $\tilde{e}_1(\boldsymbol{k}_0) = 0$. Combining the above results, we obtain $A(\boldsymbol{k}_0) = 0$.

$G^-(T_-, C_-)$**,** or equivalently $G(Q, T_-)$ after identifying C with Q

$$T = ZE, C = EI, \quad n_0 = 4, \quad \begin{array}{c|c} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YY + \tilde{o}_2(\boldsymbol{k})IX + \tilde{o}_3(\boldsymbol{k})IZ \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YI \end{array}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \end{cases}$$

Let's focus on the first two equations.

On the curve $\mathcal{C}_1$ defined in A.2 with $d = 2$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \tag{A.19}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_1(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_1(\boldsymbol{k})$ we conclude $\tilde{e}_1(\boldsymbol{k}) = 0$ at $r = 0$. In other words $\tilde{e}_1(\boldsymbol{k}_0) = 0$. This means $A(\boldsymbol{k}_0) = 0$.

$G^+(Q, T_+, C_-)$**,** or equivalently $G^-(Q, T_+, C_-)$ after identifying C with QC

$Q = EII, T = ZII, C = ZEI, \quad n_0 = 8,$

$S(\boldsymbol{k})$	$\tilde{o}_1(\boldsymbol{k})YXY + \tilde{o}_2(\boldsymbol{k})YZY$
$A(\boldsymbol{k})$	$\tilde{e}_1(\boldsymbol{k})YXX + \tilde{e}_2(\boldsymbol{k})YXZ + \tilde{e}_3(\boldsymbol{k})YXI+$ $\tilde{e}_4(\boldsymbol{k})YZX + \tilde{e}_5(\boldsymbol{k})YZZ + \tilde{e}_6(\boldsymbol{k})YZI$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_5(\boldsymbol{k}) - \tilde{o}_2(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_6(\boldsymbol{k}) = 0 \end{cases}$$

We examine the above equation in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 2$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_5(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \end{cases}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_5(\boldsymbol{k}) = \tilde{e}_4(\boldsymbol{k}) = \tilde{e}_3(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_{3,4,5}(\boldsymbol{k})$ we conclude $\tilde{e}_5(\boldsymbol{k}) = \tilde{e}_4(\boldsymbol{k}) = \tilde{e}_3(\boldsymbol{k}) = 0$ at $r = 0$, or in other words, $\tilde{e}_5(\boldsymbol{k}_0) = \tilde{e}_4(\boldsymbol{k}_0) = \tilde{e}_3(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$, which will lead to $\tilde{e}_1(\boldsymbol{k}_0) = \tilde{e}_2(\boldsymbol{k}_0) = \tilde{e}_6(\boldsymbol{k}_0) = 0$. Combining these results we conclude $A(\boldsymbol{k}_0) = 0$.

$G^+(Q, T_-, C_+)$**,** or equivalently $G^-(Q, T_-, C_+)$ after identifying C with QC

$Q = EI, T = ZE, C = ZI, \quad n_0 = 4,$

$S(\boldsymbol{k})$	$\tilde{o}_1(\boldsymbol{k})IX + \tilde{o}_2(\boldsymbol{k})IZ$
$A(\boldsymbol{k})$	0

Since there is no $A(k) \implies A(k_0) = 0$.

$G^{+}(Q, T_{-}, C_{-})$, or equivalently $G^{-}(Q, T_{-}, C_{-})$ after identifying C with QC

$$Q = EII, T = ZEI, C = ZIE, \quad n_0 = 8,$$
$$\begin{array}{r|l} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YYX + \tilde{o}_2(\boldsymbol{k})YYZ + \tilde{o}_3(\boldsymbol{k})IXI + \tilde{o}_4(\boldsymbol{k})IZI \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YIX + \tilde{e}_2(\boldsymbol{k})YIZ + \tilde{e}_3(\boldsymbol{k})IXY + \tilde{e}_4(\boldsymbol{k})IZY \end{array}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) - \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \\ \tilde{o}_4(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) - \tilde{o}_2(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_4(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \end{cases}$$

Let's focus on the first three equations. We examine these equations in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 2$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \end{cases}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_1(\boldsymbol{k}) = \tilde{e}_3(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_{1,3}(\boldsymbol{k})$ we conclude $\tilde{e}_1(\boldsymbol{k}) = \tilde{e}_3(\boldsymbol{k}) = 0$ at $r = 0$. In other words $\tilde{e}_1(\boldsymbol{k}_0) = \tilde{e}_3(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$, which will lead to $\tilde{e}_2(\boldsymbol{k}_0) = \tilde{e}_4(\boldsymbol{k}_0) = 0$. Combining these results we conclude $A(\boldsymbol{k}_0) = 0$.

A.2.3 3D SPNs

$G(C_{-})$, or equivalently $G(Q)$ after identifying C with Q

$$C = EI, \quad n_0 = 4, \quad \begin{array}{r|c} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YY + \tilde{o}_2(\boldsymbol{k})IX + \tilde{o}_3(\boldsymbol{k})IZ \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YX + \tilde{e}_2(\boldsymbol{k})YZ + \tilde{e}_3(\boldsymbol{k})YI + \tilde{e}_4(\boldsymbol{k})IY \end{array}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_2(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) + \tilde{o}_3(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) + \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \end{cases}$$

We examine the above equation in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 3$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \end{cases}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_3(\boldsymbol{k}) = \tilde{e}_4(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_{3,4}(\boldsymbol{k})$ we conclude $\tilde{e}_3(\boldsymbol{k}) = \tilde{e}_4(\boldsymbol{k}) = 0$ at $r = 0$, or in other words, $\tilde{e}_3(\boldsymbol{k}_0) = \tilde{e}_4(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$ and $\mathcal{C}_3$, which will lead to $\tilde{e}_1(\boldsymbol{k}_0) = \tilde{e}_2(\boldsymbol{k}_0) = 0$. Combining these results, one gets $A(\boldsymbol{k}_0) = 0$.

$G^-(T_+, C_-)$**,** or equivalently $G(Q, T_+)$ after identifying C with Q

$$\begin{array}{c|c} & C = EII, T = ZII, \quad n_0 = 8, \\ S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YXY + \tilde{o}_2(\boldsymbol{k})YYI + \tilde{o}_3(\boldsymbol{k})YZY \\ & +\tilde{o}_4(\boldsymbol{k})YYX + \tilde{o}_5(\boldsymbol{k})YYZ + \tilde{o}_6(\boldsymbol{k})YIY \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YXX + \tilde{e}_2(\boldsymbol{k})YXZ + \tilde{e}_3(\boldsymbol{k})YXI \\ & +\tilde{e}_4(\boldsymbol{k})YYY + \tilde{e}_5(\boldsymbol{k})YZX + \tilde{e}_6(\boldsymbol{k})YZZ \\ & +\tilde{e}_7(\boldsymbol{k})YZI + \tilde{e}_8(\boldsymbol{k})YIX + \tilde{e}_9(\boldsymbol{k})YIZ + \tilde{e}_{10}(\boldsymbol{k})YII \end{array} \tag{A.20}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{e}_{10}(\boldsymbol{k}) \begin{bmatrix} \tilde{o}_4(\boldsymbol{k}) \\ \tilde{o}_5(\boldsymbol{k}) \\ \tilde{o}_6(\boldsymbol{k}) \end{bmatrix} = \begin{bmatrix} -\tilde{e}_6(\boldsymbol{k}) & -\tilde{e}_8(\boldsymbol{k}) & \tilde{e}_2(\boldsymbol{k}) \\ \tilde{e}_5(\boldsymbol{k}) & -\tilde{e}_9(\boldsymbol{k}) & -\tilde{e}_1(\boldsymbol{k}) \\ -\tilde{e}_3(\boldsymbol{k}) & -\tilde{e}_4(\boldsymbol{k}) & -\tilde{e}_7(\boldsymbol{k}) \end{bmatrix} \cdot \begin{bmatrix} \tilde{o}_1(\boldsymbol{k}) \\ \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) \end{bmatrix} \\ \tilde{e}_{10}(\boldsymbol{k}) \begin{bmatrix} \tilde{o}_1(\boldsymbol{k}) \\ \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) \end{bmatrix} = \begin{bmatrix} -\tilde{e}_6(\boldsymbol{k}) & \tilde{e}_5(\boldsymbol{k}) & -\tilde{e}_3(\boldsymbol{k}) \\ -\tilde{e}_8(\boldsymbol{k}) & -\tilde{e}_9(\boldsymbol{k}) & -\tilde{e}_4(\boldsymbol{k}) \\ \tilde{e}_2(\boldsymbol{k}) & -\tilde{e}_1(\boldsymbol{k}) & -\tilde{e}_7(\boldsymbol{k}) \end{bmatrix} \cdot \begin{bmatrix} \tilde{o}_4(\boldsymbol{k}) \\ \tilde{o}_5(\boldsymbol{k}) \\ \tilde{o}_6(\boldsymbol{k}) \end{bmatrix} \end{cases} \tag{A.21}$$

It's straightforward to check that the above equations imply

$$\left[\tilde{o}_4^2(\boldsymbol{k}) + \tilde{o}_5^2(\boldsymbol{k}) + \tilde{o}_6^2(\boldsymbol{k}) - \tilde{o}_1^2(\boldsymbol{k}) - \tilde{o}_2^2(\boldsymbol{k}) - \tilde{o}_3^2(\boldsymbol{k})\right] \tilde{e}_{10}(\boldsymbol{k}) = 0$$

The solutions are

$$\tilde{e}_{10}(\boldsymbol{k}) = 0 \text{ or } \left[\tilde{o}_4^2(\boldsymbol{k}) + \tilde{o}_5^2(\boldsymbol{k}) + \tilde{o}_6^2(\boldsymbol{k}) - \tilde{o}_1^2(\boldsymbol{k}) - \tilde{o}_2^2(\boldsymbol{k}) - \tilde{o}_3^2(\boldsymbol{k})\right] = 0$$

In the following we prove that $\tilde{e}_{10}(\boldsymbol{k})$ must vanish.

The spectral symmetrised Hamiltonian $\tilde{h}(\boldsymbol{k})$ satisfies $\tilde{h}^2(\boldsymbol{k}) = [S(\boldsymbol{k}) + A(\boldsymbol{k})]^2 = w^2(\boldsymbol{k})III$. We may assume $w(\boldsymbol{k}) > 0$ without loss of generality. In the following we show that $\tilde{e}_{10}(\boldsymbol{k})$ must take one of the following values

$$\{w(\boldsymbol{k}), w(\boldsymbol{k})/2, 0, -w(\boldsymbol{k})/2, -w(\boldsymbol{k})\}$$

for each $\boldsymbol{k}$. We first observe that according to Eq.(A.20) all tensor products in $S(\boldsymbol{k})$ and $A(\boldsymbol{k})$ contain Y as the first factor. Therefore we can factor it out and write $\tilde{h}(\boldsymbol{k}) = Y \otimes g(\boldsymbol{k})$ where $g(\boldsymbol{k})$ is a 4×4 Hermitian matrix function. Next, we express $g(\boldsymbol{k})$ in terms of its eigenbasis, i.e., $g(\boldsymbol{k}) = U(\boldsymbol{k})\Lambda(\boldsymbol{k})U^{-1}(\boldsymbol{k})$ where $U(\boldsymbol{k})$ is the basis transformation matrix and $\Lambda(\boldsymbol{k})$ is the diagonal matrix containing the eigenvalues. Under this basis $\tilde{h}^2(\boldsymbol{k}) = I \otimes U(\boldsymbol{k})\Lambda^2(\boldsymbol{k})U^{-1}(\boldsymbol{k})$. Since the spectral symmetrization condition requires $\tilde{h}^2(\boldsymbol{k}) = w^2(\boldsymbol{k})III$, it follows that

$$U(\boldsymbol{k})\Lambda^2(\boldsymbol{k})U^{-1}(\boldsymbol{k}) = w^2(\boldsymbol{k})II.$$

This implies the eigenvalues of $\Lambda^2(\boldsymbol{k})$ are four-fold degenerate and are equal to $w^2(\boldsymbol{k})$. Thus the diagonal elements of $\Lambda(\boldsymbol{k})$ are $\pm w(\boldsymbol{k})$. According to Eq.(A.20)

$$\tilde{e}_{10}(\boldsymbol{k}) = \frac{1}{8}\mathrm{Tr}[(YII)\tilde{h}(\boldsymbol{k})] = \frac{1}{4}\mathrm{Tr}[\Lambda(\boldsymbol{k})].$$

Because the the diagonal elements of $\Lambda(\boldsymbol{k})$ are $\pm w(\boldsymbol{k})$, $\tilde{e}_{10}(\boldsymbol{k})$ must be equal to one of the five possible values

$$\{w(\boldsymbol{k}), w(\boldsymbol{k})/2, 0, -w(\boldsymbol{k})/2, -w(\boldsymbol{k})\} \tag{A.22}$$

for each $\boldsymbol{k}$.

Moreover, because $\tilde{e}_{10}(\boldsymbol{k})$ is a analytic function of $\boldsymbol{k}$ and $w(\boldsymbol{k}) > 0$ away from $\boldsymbol{k} = 0$, $\tilde{e}_{10}(\boldsymbol{k})$ can not "switch track", i.e., it must be equal to one of above five possible functions throughout the Brillouin zone, away from $\boldsymbol{k} = 0$.

Since $\tilde{h}(\boldsymbol{k}) \to \sum_{j=1}^{d} k_j \Gamma_j$ as $\boldsymbol{k} \to 0$, it follows that $w(\boldsymbol{k}) \to |\boldsymbol{k}|$ as $\boldsymbol{k} \to 0$. On the other hand, since $\tilde{e}_{10}(\boldsymbol{k})$ is an even function of $\boldsymbol{k}$, it must vanishes as an even power in $\boldsymbol{k}$ as $\boldsymbol{k} \to 0$, hence

$$|\tilde{e}_{10}(\boldsymbol{k})| << w(\boldsymbol{k}) \;\text{ as }\; \boldsymbol{k} \to 0. \tag{A.23}$$

The only choice in Eq.(A.22) that is consistent with Eq.(A.23) is

$$\tilde{e}_{10}(\boldsymbol{k}) = 0. \tag{A.24}$$

Now we may set $\tilde{e}_{10}(\boldsymbol{k}) = 0$ in the first three equations of Eq.(A.21) and examine these equations in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 3$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\begin{cases} -\tilde{o}_1(\boldsymbol{k})\tilde{e}_6(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_5(\boldsymbol{k}) = 0 \\ -\tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \end{cases}$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_3(\boldsymbol{k}) = \tilde{e}_5(\boldsymbol{k}) = \tilde{e}_6(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_{3,5,6}(\boldsymbol{k})$ we conclude $\tilde{e}_3(\boldsymbol{k}) = \tilde{e}_5(\boldsymbol{k}) = \tilde{e}_6(\boldsymbol{k}) = 0$ at $r = 0$, or in other words, $\tilde{e}_3(\boldsymbol{k}_0) = \tilde{e}_5(\boldsymbol{k}_0) = \tilde{e}_6(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$ and $\mathcal{C}_3$, which will lead to $\tilde{e}_4(\boldsymbol{k}_0) = \tilde{e}_8(\boldsymbol{k}_0) = \tilde{e}_9(\boldsymbol{k}_0) = 0$ and $\tilde{e}_1(\boldsymbol{k}_0) = \tilde{e}_2(\boldsymbol{k}_0) = \tilde{e}_7(\boldsymbol{k}_0) = 0$. Combining these results, one gets $A(\boldsymbol{k}_0) = 0$.

$G^-(T_-, C_-)$**,** or equivalently $G(Q, T_-)$ after identifying C with Q

$$C = EI, T = ZE, \quad n_0 = 4, \quad \begin{array}{c|c} S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YY + \tilde{o}_2(\boldsymbol{k})IX + \tilde{o}_3(\boldsymbol{k})IZ \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YI \end{array}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_2(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \end{cases}$$

We examine the above equations in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 3$, for any $r = |\boldsymbol{q}| \neq 0$,

$$\tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0$$

Because $\tilde{o}_1(\boldsymbol{k}) > 0$ it implies $\tilde{e}_1(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_1(\boldsymbol{k})$ we conclude $\tilde{e}_1(\boldsymbol{k}) = 0$ at $r = 0$, or in other words, $\tilde{e}_1(\boldsymbol{k}_0) = 0$. This implies $A(\boldsymbol{k}_0) = 0$.

$G(Q, C_-)$

$$\begin{array}{c|c} \multicolumn{2}{l}{Q = EII, C = ZEI, \quad n_0 = 8,} \\ S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YXY + \tilde{o}_2(\boldsymbol{k})YZY + \tilde{o}_3(\boldsymbol{k})IYY \\ & +\tilde{o}_4(\boldsymbol{k})IIX + \tilde{o}_5(\boldsymbol{k})IIZ \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YXX + \tilde{e}_2(\boldsymbol{k})YXZ + \tilde{e}_3(\boldsymbol{k})YXI \\ & +\tilde{e}_4(\boldsymbol{k})YZX + \tilde{e}_5(\boldsymbol{k})YZZ + \tilde{e}_6(\boldsymbol{k})YZI \\ & +\tilde{e}_7(\boldsymbol{k})IYX + \tilde{e}_8(\boldsymbol{k})IYZ + \tilde{e}_9(\boldsymbol{k})IYI + \tilde{e}_{10}(\boldsymbol{k})IIY \end{array} \tag{A.25}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_6(\boldsymbol{k}) + \tilde{o}_3(\boldsymbol{k})\tilde{e}_9(\boldsymbol{k}) = 0 \\ \begin{bmatrix} 0 & -\tilde{o}_3(\boldsymbol{k}) & \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) & 0 & -\tilde{o}_1(\boldsymbol{k}) \\ -\tilde{o}_2(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k}) & 0 \end{bmatrix} \begin{bmatrix} \tilde{e}_1(\boldsymbol{k}) \\ \tilde{e}_4(\boldsymbol{k}) \\ \tilde{e}_7(\boldsymbol{k}) \end{bmatrix} = \tilde{o}_5(\boldsymbol{k}) \begin{bmatrix} \tilde{e}_3(\boldsymbol{k}) \\ \tilde{e}_6(\boldsymbol{k}) \\ \tilde{e}_9(\boldsymbol{k}) \end{bmatrix} \\ \begin{bmatrix} 0 & -\tilde{o}_3(\boldsymbol{k}) & \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) & 0 & -\tilde{o}_1(\boldsymbol{k}) \\ -\tilde{o}_2(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k}) & 0 \end{bmatrix} \begin{bmatrix} \tilde{e}_2(\boldsymbol{k}) \\ \tilde{e}_5(\boldsymbol{k}) \\ \tilde{e}_8(\boldsymbol{k}) \end{bmatrix} = -\tilde{o}_4(\boldsymbol{k}) \begin{bmatrix} \tilde{e}_3(\boldsymbol{k}) \\ \tilde{e}_6(\boldsymbol{k}) \\ \tilde{e}_9(\boldsymbol{k}) \end{bmatrix} \\ \begin{bmatrix} \tilde{o}_1(\boldsymbol{k}) \\ \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) \end{bmatrix} \tilde{e}_{10}(\boldsymbol{k}) = \tilde{o}_4(\boldsymbol{k}) \begin{bmatrix} \tilde{e}_1(\boldsymbol{k}) \\ \tilde{e}_4(\boldsymbol{k}) \\ \tilde{e}_7(\boldsymbol{k}) \end{bmatrix} + \tilde{o}_5(\boldsymbol{k}) \begin{bmatrix} \tilde{e}_2(\boldsymbol{k}) \\ \tilde{e}_5(\boldsymbol{k}) \\ \tilde{e}_8(\boldsymbol{k}) \end{bmatrix} \end{cases} \tag{A.26}$$

We examine the above equations in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 3$, for any $r = |\boldsymbol{q}| \neq 0$, the first equation gives

$$\tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0$$

which implies $\tilde{e}_3(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_3(\boldsymbol{k})$ we conclude $\tilde{e}_3(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$ and $\mathcal{C}_3$, which lead to $\tilde{e}_6(\boldsymbol{k}_0) = \tilde{e}_9(\boldsymbol{k}_0) = 0$.

By theorem 1 of A.3, for any radius $|\boldsymbol{q}| = r$, we can find a non-self-intersecting closed loop γ_5, such that (i) $\tilde{o}_5(\boldsymbol{k}) = 0$ for $\boldsymbol{k} \in \gamma_5$, (ii) γ_5 splits the sphere $|\boldsymbol{q}| = r$ into two equal-area regions, and (iii) the antipodal point of any $\boldsymbol{k} \in \gamma_5$ is also on γ_5. Such γ_5 loops for different radius r form a surface S_5 which can be arbitrarily close to $r = 0$ (i.e. $\boldsymbol{k}_0$). On S_5 the second to the fourth lines of Eq.(A.26) gives

$$\begin{bmatrix} 0 & -\tilde{o}_3(\boldsymbol{k}) & \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) & 0 & -\tilde{o}_1(\boldsymbol{k}) \\ -\tilde{o}_2(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k}) & 0 \end{bmatrix} \begin{bmatrix} \tilde{e}_1(\boldsymbol{k}) \\ \tilde{e}_4(\boldsymbol{k}) \\ \tilde{e}_7(\boldsymbol{k}) \end{bmatrix} = 0$$

Note that the 3×3 matrix on the left hand side is rank 2 as long as $\tilde{o}_1(\boldsymbol{k})^2 + \tilde{o}_2(\boldsymbol{k})^2 + \tilde{o}_3(\boldsymbol{k})^2 \neq 0$, which is true for in the neighborhood of $\boldsymbol{k}_0$. This gives the general solution

$$\begin{bmatrix} \tilde{e}_1(\boldsymbol{k}) \\ \tilde{e}_4(\boldsymbol{k}) \\ \tilde{e}_7(\boldsymbol{k}) \end{bmatrix} = a(\boldsymbol{k}) \begin{bmatrix} \tilde{o}_1(\boldsymbol{k}) \\ \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) \end{bmatrix} \tag{A.27}$$

Note that as $\boldsymbol{k} \to \boldsymbol{k}_0$ we can have the following two possibilities: (i) $a(\boldsymbol{k})$ is non-singular, in which case $(\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k})) \to 0$ as $\boldsymbol{k} \to \boldsymbol{k}_0$, or (ii) $a(\boldsymbol{k})$ diverges and it compensates for the vanishing magnitude of $(\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k}))$.

We first consider possibility (ii). In this case as $\boldsymbol{k} \to \boldsymbol{k}_0$, $(\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k}))$ can be non-zero. However, its direction must be parallel (or antiparallel) to

$$\hat{n}(\boldsymbol{k}) = (\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k}))/|(\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k}))|.$$

Let's look at the pair of antipodal points on a γ_5 loop at an infinitesimal radius $|\boldsymbol{q}| = r$. By continuity of $(\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k}))$ and $\hat{n}(\boldsymbol{k})$ must change continuously on γ_5. This implies $\hat{n}(\boldsymbol{k}) \cdot (\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k}))$ changes continuously on γ_5. Since $\hat{n}(\boldsymbol{k})$ is odd and $\tilde{e}_{1,4,7}(\boldsymbol{k})$ are even, $\hat{n}(\boldsymbol{k}) \cdot (\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k}))$ has opposite sign among antipodal points on γ_5. Thus it must vanish at some intermediate point $\boldsymbol{k}'$ on γ_5. Since $(\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k})) \parallel \hat{n}(\boldsymbol{k})$ on γ_5 by (A.27), thus $(\tilde{e}_1(\boldsymbol{k}'), \tilde{e}_4(\boldsymbol{k}'), \tilde{e}_7(\boldsymbol{k}')) = \boldsymbol{0}$. By connecting such point for different r, we arrive at a continuous path on which $(\tilde{e}_1(\boldsymbol{k}), \tilde{e}_4(\boldsymbol{k}), \tilde{e}_7(\boldsymbol{k})) = \boldsymbol{0}$. By continuity we have

$$(\tilde{e}_1(\boldsymbol{k}_0), \tilde{e}_4(\boldsymbol{k}_0), \tilde{e}_7(\boldsymbol{k}_0)) = 0$$

We can repeat the same arguments for the surface corresponds to $\tilde{o}_4(\boldsymbol{k}) = 0$. This lead to $(\tilde{e}_2(\boldsymbol{k}_0), \tilde{e}_5(\boldsymbol{k}_0), \tilde{e}_8(\boldsymbol{k}_0)) = \boldsymbol{0}$.

Moreover, by the theorem 2 of A.3, on the sphere correspond to any $r = |\boldsymbol{q}|$, one can find a point $\boldsymbol{k}$ such that both $\tilde{o}_4(\boldsymbol{k})$ and $\tilde{o}_5(\boldsymbol{k})$ are zero. Such points for different r form a curve which approaches $\boldsymbol{k}_0$ as $r \to 0$. On the curve, the last of Eq.(A.26) gives

$$\begin{bmatrix} \tilde{o}_1(\boldsymbol{k}) \\ \tilde{o}_2(\boldsymbol{k}) \\ \tilde{o}_3(\boldsymbol{k}) \end{bmatrix} \tilde{e}_{10}(\boldsymbol{k}) = 0$$

Since on this curve since $\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k})$ cannot simultaneously be zero, it follows that $\tilde{e}_{10}(\boldsymbol{k}) = 0$ on the curve. Due to the continuity of $\tilde{e}_{10}(\boldsymbol{k})$ we conclude that $\tilde{e}_{10}(\boldsymbol{k}_0) = 0$. Combining all of the above results, we conclude $A(\boldsymbol{k}_0) = 0$.

$G^+(Q, T_-, C_-)$

$$\begin{array}{c|l} \multicolumn{2}{l}{Q = EII, T = ZEI, C = ZIE, \quad n_0 = 8,} \\ S(\boldsymbol{k}) & \tilde{o}_1(\boldsymbol{k})YYX + \tilde{o}_2(\boldsymbol{k})YYZ + \tilde{o}_3(\boldsymbol{k})IXI + \tilde{o}_4(\boldsymbol{k})IZI \\ \hline A(\boldsymbol{k}) & \tilde{e}_1(\boldsymbol{k})YIX + \tilde{e}_2(\boldsymbol{k})YIZ + \tilde{e}_3(\boldsymbol{k})IXY + \tilde{e}_4(\boldsymbol{k})IZY \end{array} \tag{A.28}$$

Here $\{S(\boldsymbol{k}), A(\boldsymbol{k})\} = 0$ implies

$$\begin{cases} \tilde{o}_2(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) + \tilde{o}_3(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) - \tilde{o}_3(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) + \tilde{o}_2(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_2(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) - \tilde{o}_4(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_4(\boldsymbol{k})\tilde{e}_2(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) + \tilde{o}_4(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \end{cases} \tag{A.29}$$

We examine the above equations in the neighborhood of $\boldsymbol{k}_0$ by expanding $\boldsymbol{k} = \boldsymbol{k}_0 + \boldsymbol{q}$. On the curve $\mathcal{C}_1$ defined in A.2 with $d = 3$, for any $r = |\boldsymbol{q}| \neq 0$, the second and third lines of Eq.(A.29) give

$$\begin{cases} \tilde{o}_1(\boldsymbol{k})\tilde{e}_4(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_1(\boldsymbol{k}) = 0 \end{cases}$$

which implies $\tilde{e}_1(\boldsymbol{k}) = \tilde{e}_4(\boldsymbol{k}) = 0$. By the continuity of $\tilde{e}_{1,4}(\boldsymbol{k})$ we conclude $\tilde{e}_1(\boldsymbol{k}_0) = \tilde{e}_4(\boldsymbol{k}_0) = 0$. We can repeat this argument by looking at $\mathcal{C}_2$ and $\mathcal{C}_3$, which will lead to $\tilde{e}_2(\boldsymbol{k}_0) = 0$.

It remains to prove that $\tilde{e}_3(\boldsymbol{k}_0) = 0$. By the theorem 1 in A.3, for any radius $r = |\boldsymbol{k}|$ one can find a non-self-intersecting closed loop γ_4 such that $\tilde{o}_4(\boldsymbol{k}_0) = 0$. As a function of r all such loops span surface which approach $\boldsymbol{k}_0$ as $r \to 0$. Everywhere on the surface, the 4-6 lines of Eq.(A.29) give

$$\begin{cases} \tilde{o}_2(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_1(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \\ \tilde{o}_3(\boldsymbol{k})\tilde{e}_3(\boldsymbol{k}) = 0 \end{cases}$$

Because $(\tilde{o}_1(\boldsymbol{k}), \tilde{o}_2(\boldsymbol{k}), \tilde{o}_3(\boldsymbol{k}))$ has non-trivial mapping degree around k_0, they cannot be simultaneously zero. It follows that $\tilde{e}_3(\boldsymbol{k}) = 0$ everywhere on the surface. By the continuity of $\tilde{e}_3(\boldsymbol{k})$, we conclude that $\tilde{e}_3(\boldsymbol{k}_0) = 0$. Combining these results, one gets $A(\boldsymbol{k}_0) = 0$.

A.3 Odd continuous functions on S^2

In this appendix, we prove some properties for odd continuous functions obeying $o(-\boldsymbol{q}) = -o(\boldsymbol{q})$, on a two-sphere S^2 formed by $|\boldsymbol{q}| =$ constant.

A.3.1 Theorem 1

Theorem 1 For any continuous odd function $o(\boldsymbol{q})$ defined on a sphere formed by $|\boldsymbol{q}| = r$, there exists a non-self intersecting closed loop γ_o on the sphere, such that (i) $o(\boldsymbol{q}) = 0$ for $\boldsymbol{q} \in \gamma_o$, (ii) the curve separates the sphere into two equal-area regions, and (iii) the antipodal point of any point $\boldsymbol{q}$ on the loop also belongs to the loop.

Proof: We will prove it by explicitly constructing γ_o. If $o(\boldsymbol{q}) = 0$ everywhere on the sphere, any arbitrary great circle on S^2 can be used for γ_o. Thus the non-trivial case must have at least one point, $\boldsymbol{q}_*$, such that $o(\boldsymbol{q}_*) \neq 0$. Without loss of generality, let's assume $o(\boldsymbol{q}_*) > 0$. Due to the oddness, $o(-\boldsymbol{q}_*) < 0$. Now consider a geodesic (or a great arc) connecting $\boldsymbol{q}_*$ and $-\boldsymbol{q}_*$. Owing to the continuity of $o(\boldsymbol{q})$, the function must change sign an odd number of times as the geodesic is traversed. The points at which the sign changes take place must correspond to $o(\boldsymbol{q}) = 0$. They can either be discrete points or form a continuous segment on the great arc. In either case we can choose a middle point (which can either be the mid point of the middle zero-segment, or just the mid point among the discrete points where the sign change takes place). We then rotate the great arc through the whole 2π angle. As a function of angle, the aforementioned mid points span the loop γ_o. The loop can not self-intersect because we only choose a single point on every great arc.

Moreover, due to the oddness of $o(\boldsymbol{q})$, the mid point $\boldsymbol{q}_m$ chosen for a given great arc must be antipodal to $-\boldsymbol{q}_m$ chosen on the complementary great arc (a great arc and its complementary form a great circle). This guarantees that the loop γ_o will separate the sphere into two regions with equal areas. By construction, the antipodal point of any point $\boldsymbol{q}$ on the γ_o is also on the loop. Q.E.D..

A.3.2 Theorem 2

Theorem 2 For any two continuous odd functions $o_1(\boldsymbol{q}), o_2(\boldsymbol{q})$ defined on a sphere $|\boldsymbol{q}| = r$, there exists at least a point $\boldsymbol{q}_{**}$ such that $o_1(\boldsymbol{q}_{**}) = o_2(\boldsymbol{q}_{**}) = 0$.

Proof: Assuming the opposite, namely, there is no point $\boldsymbol{q}$ at which $o_1(\boldsymbol{q}) = o_2(\boldsymbol{q}) = 0$. For these two functions $o_1(\boldsymbol{q}), o_2(\boldsymbol{q})$, we can use theorem 1 to find the non-self-intersecting closed loops γ_1 and γ_2 which separately divide the sphere into two equal-area regions, and $o_1(\boldsymbol{q}) = 0$ for $\boldsymbol{q} \in \gamma_1$ and $o_2(\boldsymbol{q}) = 0$ for $\boldsymbol{q} \in \gamma_2$. γ_1 and γ_2 must not intersect each other, otherwise the intersection will satisfy $o_1(\boldsymbol{q}) = o_2(\boldsymbol{q}) = 0$. Thus, one loop must be totally enclosed by the other loop, which contradicts the statement that they separately split the sphere into two equal-area regions. Q.E.D.

Appendix B

Appendices of part III

B.1 The emergent symmetries for $(2+1)$ and $(3+1)$-D

In this appendix we derive the emergent symmetries of the massless fermion theory (see table 4.1) for spatial dimension $d = 2, 3$ (for $d = 1$ the result has already been discussed in sections 4.2).

B.1.1 Complex class in $(2+1)$-D

In the complex fermion representation, the minimal size of the gamma matrices in two spatial dimensions is 2×2. If the fermion has n flavors, modulo a basis transformation, we have

$$\begin{aligned} &S_0 = \int d^3x\, \psi^\dagger(\partial_0 - i\sum_{i=1}^{2}\Gamma_i\partial_i)\psi \quad \text{where} \\ &\Gamma_1 = ZI_n, \quad \Gamma_2 = XI_n. \end{aligned} \tag{B.1}$$

It's easy to see that the full emergent symmetries include $U(n)$ transformations in the flavor degrees of freedom. In addition, there are discrete symmetries, namely, charge conjugation and time-reversal symmetries. To summarize, Eq.(B.1) is invariant under

$$\begin{aligned} &U(n) \text{ symmetry :} \\ &U(n) : \psi \to (I \otimes g)\, \psi \quad \text{ where } g \in U(n) \\ &\text{Charge conjugation symmetry :} \\ &C : \psi \to (I \otimes I_n)\, (\psi^\dagger)^T \\ &\text{Time reversal symmetry :} \\ &T : \psi \to (Y \otimes I_n)\, \psi. \end{aligned} \tag{B.2}$$

B.1.2 Real class in $(2+1)$-D

In the Majorana fermion representation, the minimal size of the gamma matrices in two spatial dimensions, is 2×2. If the fermion has n flavors, modulo a basis transformation, we have

$$\begin{aligned} & S_0 = \int d^3x\, \chi^T(\partial_0 - i\sum_{i=1}^{2}\Gamma_i\partial_i)\chi \quad \text{where} \\ & \Gamma_1 = ZI_n, \quad \Gamma_2 = XI_n. \end{aligned} \tag{B.3}$$

It's easy to see that the full emergent symmetries include $O(n)$ transformations in the flavor degrees of freedom. In addition, there is time reversal symmetry. To summarize, Eq.(B.3) is invariant under

$$\begin{aligned} & O(n) \text{ symmetry :} \\ & O(n) : \chi \to (I \otimes g)\,\chi \quad \text{ where } g \in O(n) \\ & \text{Time reversal symmetry :} \\ & T : \chi \to (E \otimes I_n)\,\chi. \end{aligned} \tag{B.4}$$

B.1.3 Complex class in $(3+1)$-D

In the complex fermion representation, the minimal size of the gamma matrices in three spatial dimensions is 4×4. If the fermion has n flavors, modulo a basis transformation, we have

$$\begin{aligned} & S_0 = \int d^4x\, \psi^\dagger(\partial_0 - i\sum_{i=1}^{3}\Gamma_i\partial_i)\psi \quad \text{where} \\ & \Gamma_1 = ZII_n, \quad \Gamma_2 = XII_n, \quad \Gamma_3 = YZI_n. \end{aligned} \tag{B.5}$$

Similar to the $(1+1)$-D case, the chirality matrix

$$\Gamma_5 := -i\Gamma_1\Gamma_2\Gamma_3 = IZI_n$$

commutes with the gamma matrices. As a result, the full emergent include chiral $U(n)$ transformations, namely, $U_+(n) \times U_-(n)$ (see below). In addition, there are discrete symmetries, namely, charge conjugation, and time-reversal symmetries. To summarize, Eq.(B.5) is invariant under

$$\begin{aligned} & \text{Chiral } U(n) \text{ symmetry :} \\ & U_+(n) \times U_-(n) : \psi \to (IP_+ \otimes g_+ + IP_- \otimes g_-)\,\psi \quad \text{ where } g_\pm \in U_\pm(n) \\ & \text{Charge conjugation symmetry :} \\ & C : \psi \to (IX \otimes I_n)\,(\psi^\dagger)^T \\ & \text{Time reversal symmetry :} \\ & T : \psi \to (YZ \otimes I_n)\,\psi, \end{aligned} \tag{B.6}$$

where

$$P_\pm := \frac{I \pm Z}{2}.$$

B.1.4 Real class in $(3+1)$-D

In the Majorana fermion representation, the minimal size of the gamma matrices in three spatial dimensions is 4×4. If the fermion has n flavors, modulo a basis transformation, we have

$$S_0 = \int d^4x\, \chi^T(\partial_0 - i\sum_{i=1}^{3}\Gamma_i\partial_i)\chi \quad \text{where}$$
$$\Gamma_1 = ZII_n, \quad \Gamma_2 = XII_n, \quad \Gamma_3 = YYI_n. \tag{B.7}$$

Although we can still define $\Gamma_1\Gamma_2\Gamma_3 = IEI_n$, this is an anti-symmetric matrix with complex eigenvectors hence cannot be used to define chirality for Majorana (real) fermions.

To find the most general continuous unitary symmetry, notice that only II and IE commute with the first two Pauli matrices in $\Gamma_{1,2,3}$. Hence the symmetry transformation needs to be in the form

$$\chi \to (II \otimes g_1 - IE \otimes g_2)\,\chi.$$

Here g_1 and g_2 are orthogonal matrices (which preserve the realness of the Majorana fermion operator and their anti-commutation relation). The condition of g_1 and g_2 being orthogonal matrices is equivalent to requiring $g_1 + ig_2 \in U(n)$[1]. Thus, the unitary continuous symmetry is $U(n)$. In addition, there is time-reversal symmetry. To summarize, Eq.(B.7) is invariant under

$$\begin{aligned}
&U(n) \text{ symmetry :}\\
&U(n) : \chi \to (II \otimes g_1 - IE \otimes g_2)\,\chi \quad \text{where } g := g_1 + ig_2 \in U(n)\\
&\text{Time reversal symmetry :}\\
&T : \chi \to (EZ \otimes I_n)\,\chi
\end{aligned} \tag{B.8}$$

B.2 The mass manifolds, homotopy groups and symmetry transformations

In this appendix we derive the mass manifolds in table 4.2, and the transformation of $Q^{\mathbb{C}}$ and $Q^{\mathbb{R}}$ under the emergent symmetries in table 4.4 for $d = 2, 3$ (the $d = 1$ case has been discussed in section 4.3 and 4.6). In addition, we discuss the relevant homotopy groups of the mass manifolds. For sufficiently large flavor number n, it turns out that the π_{D+1}, relevant to the existence of WZW term, are always equal to $\mathbb{Z}$. On the other hand,

[1] As an algebraic relation, IE plays the role of i here because $(IE)^2 = -I_4$

π_{D-1}, relevant to the existence of non-trivial soliton, are $\mathbb{Z}$ or $\mathbb{Z}_2$ depending on whether the class is complex or real. [2].

B.2.1 Complex class in $(2+1)$-D

In $(2+1)$-D, complex fermion representation, the gamma matrices in Eq.(B.1) are

$$\Gamma_1 = ZI_n, \quad \Gamma_2 = XI_n. \tag{B.9}$$

The most general hermitian mass matrix M satisfying

$$\{M, \Gamma_i\} = 0 \text{ and } M^2 = I_{2n}$$

is of the form

$$M = Y \otimes H := Y \otimes Q^{\mathbb{C}} \tag{B.10}$$

where $Q^{\mathbb{C}} = H$ is an $n \times n$ hermitian matrix satisfying $H^2 = I_n$. This last condition requires the eigenvalues of H to be ± 1. Assuming l of the eigenvalues are $+1$ and $n-l$ are -1, we have

$$Q^{\mathbb{C}} = W \cdot diag(\underbrace{+1, ..., +1}_{l}, \underbrace{-1, ..., -1}_{n-l}) \cdot W^\dagger.$$

Different $Q^{\mathbb{C}}$ are characterized by the unitary matrix (whose columns are eigenvectors) $W \in U(n)$. However, not all W will yield distinct $Q^{\mathbb{C}}$. Under the transformation

$$W \to W \cdot \begin{pmatrix} \tilde{W}_1 & 0 \\ 0 & \tilde{W}_2 \end{pmatrix} \quad \text{where} \quad \tilde{W}_1 \in U(l) \quad \text{and} \quad \tilde{W}_2 \in U(n-l),$$

$Q^{\mathbb{C}}$ is unchanged. Thus the mass manifold $\mathcal{M}$ is the union of the quotient spaces

$$\mathcal{M} = \bigcup_{l=0}^{n} \frac{U(n)}{U(l) \times U(n-l)}.$$

These quotient spaces are called "complex Grassmannians". Note that $\mathcal{M}$ contains $n+1$ disconnected components.

Under the action of the emergent symmetries in Eq.(B.2), the order parameter $Q^{\mathbb{C}}$ transforms as

$$\begin{aligned} Q^{\mathbb{C}} &\xrightarrow{U(n)} g^\dagger \cdot Q^{\mathbb{C}} \cdot g \\ Q^{\mathbb{C}} &\xrightarrow{C} \left(Q^{\mathbb{C}}\right)^T \\ Q^{\mathbb{C}} &\xrightarrow{T} - \left(Q^{\mathbb{C}}\right)^* \end{aligned}$$

[2] Although we shall not further discuss it in this paper, the $\mathbb{Z}$ or $\mathbb{Z}_2$ soliton classifications are originated from K-theory[7] and the Bott periodicity [57, 58]. Therefore this statement holds true in even higher dimensions.

Among them, the time reversal transformation changes the signs of all eigenvalues and thus exchanges l and $n-l$. Therefore only when

$$Q^{\mathbb{C}} \in \frac{U(n)}{U(n/2)\times U(n/2)} \text{ for } n \in \mathit{even}$$

does the time reversal transformed $Q^{\mathbb{C}}$ stay in the same component of the mass manifold. Only in this manifold, fluctuating $Q^{\mathbb{C}}$ can restore the *full* emergent symmetries.

Using the long exact sequence of the homotopy group corresponding to the fibration,

$$0 \to U(\frac{n}{2}) \times U(\frac{n}{2}) \to U(n) \to \frac{U(n)}{U(\frac{n}{2})\times U(\frac{n}{2})} \to 0,$$

we can deduce the homotopy groups of the complex Grassmannian from the homotopy groups of $U(n)$ (see, e.g., [59]). In table B.1 we list the results of the second, third, and fourth homotopy groups. They are relevant for determining the existence of solitons, θ-term, and WZW term. These results are used in appendix B.4.

B.2.2 Real class in $(2+1)$-D

The massless fermion Hamiltonian is given by Eq.(B.3), where the gamma matrices are given by

$$\Gamma_1 = ZI_n \quad \Gamma_2 = XI_n \tag{B.11}$$

The most general purely imaginary *antisymmetric* mass matrix (requirement due to hermiticity and Majorana condition) M satisfying

$$\{M, \Gamma_i\} = 0 \text{ and } M^2 = I_{2n}$$

is of the form

$$M = Y \otimes S := Y \otimes Q^{\mathbb{R}} \tag{B.12}$$

n (even)	Mass manifold	π_2 (soliton)	π_3 (θ term)	π_4 (WZW)
≥ 4	$\frac{U(n)}{U(n/2)\times U(n/2)}$	$\boxed{\mathbb{Z}}$	$\boxed{0}$	$\boxed{\mathbb{Z}}$
2	$\frac{U(2)}{U(1)\times U(1)} = S^2$	$\boxed{\mathbb{Z}}$	$\mathbb{Z}$	$\mathbb{Z}_2$

Table B.1: The homotopy groups of the complex Grassmannian $\frac{U(n)}{U(n/2)\times U(n/2)}$. We box the homotopy group when it is stabilized, i.e., no longer changes with increasing n.

where $Q^{\mathbb{R}} = S$ is an $n \times n$ real symmetric matrix satisfying $S^2 = I_n$. This last condition requires the eigenvalues of $Q^{\mathbb{R}}$ to be ± 1. Assuming l of the eigenvalues are $+1$ and the rest are -1, we have

$$Q^{\mathbb{R}} = W \cdot diag(\underbrace{+1, ..., +1}_{l}, \underbrace{-1, ..., -1}_{n-l}) \cdot W^{\dagger}.$$

Hence different $Q^{\mathbb{R}}$ are characterized by the orthogonal matrix $W \in O(n)$. However, not all W yield distinct $Q^{\mathbb{R}}$. Under the transformation

$$W \to W \cdot \begin{pmatrix} \tilde{W}_1 & 0 \\ 0 & \tilde{W}_2 \end{pmatrix} \quad \text{where} \quad \tilde{W}_1 \in O(l) \quad \text{and} \quad \tilde{W}_2 \in O(n-l),$$

$Q^{\mathbb{R}}$ is unchanged. Thus the mass manifold is the union of quotient spaces called "real Grassmannians"

$$\mathcal{M} = \bigcup_{l=0}^{n} \frac{O(n)}{O(l) \times O(n-l)}.$$

Here, $\mathcal{M}$ contains $n+1$ disconnected components.

Under the action of the emergent symmetries in Eq.(B.4), the order parameter $Q^{\mathbb{R}}$ transforms as

$$\begin{aligned} Q^{\mathbb{R}} &\xrightarrow{O(N)} g^T \cdot Q^{\mathbb{R}} \cdot g \\ Q^{\mathbb{R}} &\xrightarrow{T} -Q^{\mathbb{R}}. \end{aligned}$$

Among them, the time reversal transformation changes the signs of all eigenvalues and thus exchanges l and $n-l$. Therefore only when

$$Q^{\mathbb{R}} \in \frac{O(n)}{O(n/2) \times O(n/2)} \text{ for } n \in even$$

does the time reversal transformed $Q^{\mathbb{R}}$ stay in the same component of the mass manifold. Only in the mass manifold, fluctuating $Q^{\mathbb{R}}$ can restore the *full* emergent symmetries.

Using the long exact sequence of the homotopy group associated with the fibration,

$$0 \to O(\frac{n}{2}) \times O(\frac{n}{2}) \to O(n) \to \frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})} \to 0,$$

we can deduce the homotopy groups of the real Grassmannian from the homotopy groups of $O(n)$ (see e.g., [59]). We list the results of the second, third, and fourth homotopy groups in table B.2. They are relevant for determining the existence of solitons, θ-term, and WZW term. These results are used in appendix B.4.

n (even)	Mass manifold	π_2 (soliton)	π_3 (θ term)	π_4 (WZW)
≥ 10	$\frac{O(n)}{O(n/2)\times O(n/2)}$	$\boxed{\mathbb{Z}_2}$	$\boxed{0}$	$\boxed{\mathbb{Z}}$
2	S^1	0	0	0
4	$\frac{S^2\times S^2}{\mathbb{Z}_2}$	$\mathbb{Z}^2$	$\mathbb{Z}^2$	$\mathbb{Z}_2^2$
6		$\boxed{\mathbb{Z}_2}$	0	$\mathbb{Z}$
8		$\boxed{\mathbb{Z}_2}$	0	$\mathbb{Z}^3$

Table B.2: The homotopy groups of the real Grassmannian $\frac{O(n)}{O(n/2)\times O(n/2)}$. We box the homotopy group when it is stabilized , i.e., no longer changes with increasing n.

B.2.3 Complex class in $(3+1)$-D

The massless fermion Hamiltonian is given by Eq.(B.5), where the gamma matrices are given by

$$\Gamma_1 = ZII_n,\ \Gamma_2 = XII_n,\ \Gamma_3 = YZI_n \tag{B.13}$$

The most general hermitian mass matrix M satisfying

$$\{M, \Gamma_i\} = 0$$

is of the form

$$M = YX \otimes H_1 + YY \otimes H_2 \tag{B.14}$$

Here $H_{1,2}$ are $n \times n$ hermitian matrices. It's easy to check that the extra condition on the mass matrix

$$M^2 = I_{4n}$$

is equivalent to requiring

$$Q^{\mathbb{C}} := H_1 + iH_2 \in U(n)$$

Thus, the mass manifold is $U(n)$. Here the mass manifold contains only a single component.

Under the action of the emergent symmetries in Eq.(B.6), the order parameter $Q^{\mathbb{C}}$ transforms as

$$\begin{aligned} Q^{\mathbb{C}} &\xrightarrow{U(n)\times U(n)} g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+ \\ Q^{\mathbb{C}} &\xrightarrow{C} \left(Q^{\mathbb{C}}\right)^T \\ Q^{\mathbb{C}} &\xrightarrow{T} \left(Q^{\mathbb{C}}\right)^* \end{aligned}$$

Fluctuating $Q^{\mathbb{C}}$ within $U(n)$ can restores the *full* emergent symmetries.

We list the results of the second, third, and fourth homotopy groups in table B.3. They are relevant for determining the existence of solitons, θ-term, and WZW term. These results are used in appendix B.4.

B.2.4 Real class in $(3+1)$-D

The massless fermion Hamiltonian is given by Eq.(B.7), where the gamma matrices are given by

$$\Gamma_1 = ZII_n,\ \Gamma_2 = XII_n,\ \Gamma_3 = YYI_n \tag{B.15}$$

The most general antisymmetric (to ensure hermiticity) mass matrix M satisfying

$$\{M, \Gamma_i\} = 0$$

is of the form

$$M = YX \otimes S_1 + YZ \otimes S_2 \tag{B.16}$$

Here $S_{1,2}$ are $n \times n$ real symmetric matrices. It's easy to check that the extra condition on the mass matrix

$$M^2 = I_{4n}$$

is equivalent to requiring

$$Q^{\mathbb{R}} := S_1 + iS_2 \in \text{symmetric } U(n).$$

According to the "Autonne decomposition" (e.g., corollary 2.6.6 of [25]), any symmetric unitary matrix can be decomposed into

$$Q^{\mathbb{R}} = U \cdot U^T$$

where U is a general $n \times n$ unitary matrix. However, not all U will yield different $Q^{\mathbb{R}}$. The transformation

$$U \to U \cdot O, \text{ where } O \in O(n)$$

n	Mass manifold	π_3 (soliton)	π_4 (θ term)	π_5 (WZW)
≥ 3	$U(n)$	$\boxed{\mathbb{Z}}$	$\boxed{0}$	$\boxed{\mathbb{Z}}$
1	S^1	0	0	0
2	$\frac{S^3 \times S^1}{\mathbb{Z}_2}$	$\boxed{\mathbb{Z}}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$

Table B.3: The homotopy groups of $U(n)$. We box the homotopy group when it is stabilized, i.e., no longer changes with increasing n.

leaves $Q^{\mathbb{R}}$ unchanged. Thus the mass manifold is

$$\mathcal{M}_m = \frac{U(n)}{O(n)}.$$

This mass manifold is called the "real Lagrangian Grassmannian", which contains a single component.

Under the action of the emergent symmetries in Eq.(B.8), the order parameter $Q^{\mathbb{R}}$ transforms as

$$Q^{\mathbb{R}} \xrightarrow{U(n)} g^T \cdot Q^{\mathbb{R}} \cdot g$$
$$Q^{\mathbb{R}} \xrightarrow{T} \left(Q^{\mathbb{R}}\right)^*$$

Fluctuating $Q^{\mathbb{R}}$ in $U(n)/O(n)$ can restore the *full* emergent symmetries.

Using the long exact sequence of the homotopy group associated with the fibration,

$$0 \to O(n) \to U(n) \to \frac{U(n)}{O(n)} \to 0,$$

we can deduce the homotopy groups of the real Lagrangian Grassmannian from the homotopy group of $U(n)$ and $O(n)$ (see e.g., [59]). We list the results of the second, third, and fourth homotopy groups in table B.4. They are relevant for determining the existence of solitons, θ-term, and WZW term. These results are used in appendix B.4.

B.3 The anomalies of the fermion theories

In this section, we shall use the heuristic method introduced in subsection 4.4.2 to determine the anomalies associated with the emergent symmetries of the massless free fermion

n	Mass manifold	π_3 (soliton)	π_4 (θ term)	π_5 (WZW)
≥ 6	$\frac{U(n)}{O(n)}$	$\boxed{\mathbb{Z}_2}$	$\boxed{0}$	$\boxed{\mathbb{Z}}$
1	S^1	0	0	0
2	$\frac{S^1 \times S^2}{\mathbb{Z}_2}$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$
3		$\boxed{\mathbb{Z}_2}$	0	$\mathbb{Z} \times \mathbb{Z}_2$
4		$\boxed{\mathbb{Z}_2}$	$\mathbb{Z}$	$\mathbb{Z} \times \mathbb{Z}_2^2$
5		$\boxed{\mathbb{Z}_2}$	$\boxed{0}$	$\mathbb{Z} \times \mathbb{Z}_2$

Table B.4: The homotopy groups of the real Lagrangian Grassmannian $\frac{U(n)}{O(n)}$. We box the homotopy group when it is stabilized, i.e., no longer changes with increasing n.

theory in (1+1), (2+1) and (3+1)-D. For each massless fermion theory, we shall determine 1) the largest subgroup of the continuous symmetry that is anomaly free, 2) whether the discrete symmetries are anomalous after imposing a regularization mass that is invariant under the anomaly-free part of the continuous symmetry. Readers are referred to table 4.1 for the emergent symmetries of the massless free fermion theories; table 4.2 for the general form of mass terms ($Q^{\mathbb{C},\mathbb{R}}$) , and the topological space (mass manifold) they reside in; table 4.4 for the transformations of $Q^{\mathbb{C},\mathbb{R}}$ under the emergent symmetries.

Complex class in $(1+1)$-D

As discussed in section4.4.2, whether a symmetry group is anomalous depends on whether there exists a regularization mass that is invariant under its action. For continuous symmetries, the existence of an invariant regularization mass guarantees the possibility of gauging such symmetries.

As shown in table 4.1 the continuous part of the emergent symmetries form the $U_+(n) \times U_-(n)$ group, and a general mass term has the following form

$$M = X \otimes H_1 + Y \otimes H_2 \quad \text{where} \quad H_1 + iH_2 := Q^{\mathbb{C}} \in U(n)$$

Under the action of $U_+(n) \times U_-(n)$ these mass terms transform according to

$$Q^{\mathbb{C}} \to g_-^{\dagger} \cdot Q^{\mathbb{C}} \cdot g_+, \quad \text{where} \quad (g_+, g_-) \in U_+(n) \times U_-(n).$$

Since there is no (regularization) mass invariant under the action of the entire $U_+(n) \times U_-(n)$, it follows that $U_+(n) \times U_-(n)$ is anomalous.

The largest anomaly free subgroup is the diagonal $U(n)$, i.e., $g_+ = g_- = g \in U(n)$. In this case we can choose $Q^{\mathbb{C}} = I_n$ (i.e., $H_1 = I_n$ and $H_2 = 0$), such that it is invariant under the diagonal $U(n)$. One can thus use

$$M_{\text{reg}} = X \otimes I_n$$

as the regularization mass.

Note that this mass term is invariant under the time-reversal and charge-conjugation symmetries

$$Q^{\mathbb{C}} \xrightarrow{T} (Q^{\mathbb{C}})^T = I_n$$
$$Q^{\mathbb{C}} \xrightarrow{C} (Q^{\mathbb{C}})^* = I_n.$$

Consequently there is no anomaly for these discrete symmetries after imposing the diagonal $U(n)$-invariant regularization mass.

B.3.1 Real class in $(1+1)$-D

As shown in table 4.1 the continuous part of the emergent symmetries form the $O_+(n) \times O_-(n)$ group, and a general mass term has the following form

$$M = Y \otimes S + X \otimes (iA) \quad \text{where} \quad S + A := Q^{\mathbb{R}} \in O(n).$$

Under the action of $O_+(n) \times O_-(n)$ these mass terms transform according to

$$Q^{\mathbb{R}} \to g_-^T \cdot Q^{\mathbb{R}} \cdot g_+, \quad \text{where} \quad (g_+, g_-) \in O_+(n) \times O_-(n).$$

Since there is no (regularization) mass invariant under the action of the entire $O_+(n) \times O_-(n)$, it follows that $O_+(n) \times O_-(n)$ is anomalous.

The largest anomaly free subgroup is the diagonal $O(n)$, i.e., $g_+ = g_- = g \in O(n)$. In this case $Q^{\mathbb{R}} = I_n$ (i.e., $S = I_n$ and $A = 0$) is invariant under the diagonal $O(n)$. One can thus use

$$M_{\text{reg}} = Y \otimes I_n$$

as the regularization mass.

Since this mass term is invariant under the time-reversal, i.e.,

$$Q^{\mathbb{R}} \xrightarrow{T} (Q^{\mathbb{R}})^T = I_n,$$

there is no anomaly for time-reversal symmetry after imposing the diagonal $O(n)$-invariant regularization mass.

Complex class in $(2+1)$-D

As shown in table 4.1 the continuous part of the emergent symmetries form the $U(n)$ group, and a general mass term has the following form

$$M = Y \otimes Q^{\mathbb{C}} \quad \text{where} \quad Q^{\mathbb{C}} \in \bigcup_l \frac{U(n)}{U(l) \times U(n-l)}$$

Under the action of $U(n)$ these mass terms transform according to

$$Q^{\mathbb{C}} \to g^\dagger \cdot Q^{\mathbb{C}} \cdot g, \quad \text{where} \quad g \in U(n).$$

Because $Q^{\mathbb{C}} = \pm I_n$ is invariant under the action of $U(n)$ one can choose

$$M_{\text{reg}} = \pm Y \otimes I_n$$

as the regularization mass. Hence the entire $U(n)$ is anomaly free.

It is easy to see that $Q^{\mathbb{C}} = \pm I_n$ is the only $U(n)$ preserving mass term. Although this mass term is invariant the charge-conjugation

$$Q^{\mathbb{C}} \xrightarrow{C} (Q^{\mathbb{C}})^T = \pm I_n,$$

it is odd under the time-reversal symmetry

$$Q^{\mathbb{C}} \xrightarrow{T} -(Q^{\mathbb{C}})^* = \mp I_n.$$

Therefore the time-reversal symmetry is anomalous after imposing the $U(n)$-invariant regularization mass.

B.3.2 Real class in $(2+1)$-D

As shown in table 4.1 the continuous part of the emergent symmetries form the $O(n)$ group, and a general mass term has the following form

$$M = Y \otimes Q^{\mathbb{R}} \quad \text{where} \quad Q^{\mathbb{R}} \in \bigcup_l \frac{O(n)}{O(l) \times O(n-l)}.$$

Under the action of $O(n)$ these mass terms transform according to

$$Q^{\mathbb{R}} \to g^T \cdot Q^{\mathbb{R}} \cdot g, \quad \text{where} \quad g \in O(n).$$

Because $Q^{\mathbb{R}} = \pm I_n$ is invariant under the action of $U(n)$ one can choose

$$M_{\text{reg}} = \pm Y \otimes I_n$$

as the regularization mass. Hence the entire $O(n)$ is anomaly free.

It is easy to see that $Q^{\mathbb{R}} = \pm I_n$ is the only $O(n)$ preserving mass term. However, this mass term is odd under the time-reversal symmetry

$$Q^{\mathbb{R}} \xrightarrow{T} -Q^{\mathbb{R}} = \mp I_n.$$

Therefore the time-reversal symmetry is anomalous after imposing the $O(n)$-invariant regularization mass.

Complex class in $(3+1)$-D

As shown in table 4.1 the continuous part of the emergent symmetries form the $U_+(n) \times U_-(n)$ group, and a general mass term has the following form

$$M = M_{\text{reg}} = YX \otimes H_1 + YY \otimes H_2 \ \text{ where } \ Q^{\mathbb{C}} := H_1 + iH_2 \in U(n)$$

Under the action of $U_+(n) \times U_-(n)$ these mass terms transform according to

$$Q^{\mathbb{C}} \to g_-^{\dagger} \cdot Q^{\mathbb{C}} \cdot g_+, \quad \text{where} \ \ (g_+, g_-) \in U_+(n) \times U_-(n).$$

Since there is no (regularization) mass invariant under the action of the entire $U_+(n) \times U_-(n)$, it follows that $U_+(n) \times U_-(n)$ is anomalous.

The largest anomaly free subgroup is the diagonal $U(n)$, i.e., $g_+ = g_- = g \in U(n)$. In this case $Q^{\mathbb{C}} = I_n$ (i.e., $H_1 = I_n$ and $H_2 = 0$) is invariant under the diagonal $U(n)$. One can thus use

$$M_{\text{reg}} = YX \otimes I_n$$

as the regularization mass.

Note that this mass term is invariant under the time-reversal and charge-conjugation symmetries,

$$Q^{\mathbb{C}} \xrightarrow{T} (Q^{\mathbb{C}})^* = I_n$$
$$Q^{\mathbb{C}} \xrightarrow{C} (Q^{\mathbb{C}})^T = I_n$$

Consequently there is no anomaly for these discrete symmetries after imposing the diagonal $U(n)$-invariant regularization mass.

Real class in $(3+1)$-D

As shown in table 4.1, the continuous part of the emergent symmetries form the $U(n)$ group, and a general mass term has the following form

$$M = YX \otimes S_1 + YZ \otimes S_2 \ \text{ where } \ Q^{\mathbb{R}} := S_1 + iS_2 \in U(n)/O(n).$$

Under the action of $U(n)$, these mass terms transform according to

$$Q^{\mathbb{R}} \to g^T \cdot Q^{\mathbb{R}} \cdot g, \quad \text{where} \ \ g \in U(n).$$

Since there is no (regularization) mass invariant under the action of the entire $U(n)$, it follows that $U(n)$ is anomalous.

The largest anomaly-free subgroup is $O(n)$, i.e., $g \in O(n)$. In this case $Q^{\mathbb{R}} = I_n$ (i.e., $S = I_n$ and $S_2 = 0$) is invariant under the diagonal $O(n)$. One can thus use

$$M_{\text{reg}} = YX \otimes I_n$$

as the regularization mass.

Note that this mass term is invariant under the time-reversal transformation,

$$Q^{\mathbb{R}} \xrightarrow{T} (Q^{\mathbb{R}})^* = I_n.$$

Consequently there is no anomaly for time-reversal symmetry after imposing the $O(n)$-invariant regularization mass.

B.4 Fermion integration

In this section, we derive the nonlinear sigma models summarized in section 4.8 and 4.9 by integrating out the gapped fermions.

B.4.1 Integrating out real versus complex fermions

For fermions in the real classes, we face integration of the following form

$$\begin{aligned} Z[Q^{\mathbb{R}}(x)] &= e^{-W[Q^{\mathbb{R}}(x)]} = \int D\chi(x) e^{-S[\chi(x),Q^{\mathbb{R}}(x)]} \quad \text{where} \\ S[\chi, Q^{\mathbb{R}}(x)] &= \int d^D x\, \chi^T \left\{ \partial_0 + \hat{H}[Q^{\mathbb{R}}(x)] \right\} \chi. \end{aligned} \tag{B.17}$$

A convenient trick for doing such integration is to perform the corresponding complex fermion integration and divide the resulting effective action by two.

Too see this, consider two copies of Majorana fermion χ_1 and χ_2 coupled to the *same* $Q^{\mathbb{R}}(x)$. After fermion integration, the result should be the square of that in Eq.(B.17), namely,

$$\begin{aligned} \int D\chi_1\, D\chi_2\, e^{-\left\{S[\chi_1(x),Q^{\mathbb{R}}(x)]+S[\chi_2(x),Q^{\mathbb{R}}(x)]\right\}} &= \left\{Z[Q^{\mathbb{R}}(x)]\right\}^2 = e^{-2W[Q^{\mathbb{R}}(x)]} \\ &:= e^{-\tilde{W}[Q^{\mathbb{R}}(x)]}. \end{aligned}$$

On the other hand, we can combine $\chi_{1,2}$ into a complex fermion field

$$\psi = \chi_1 + i\chi_2,$$

so that the sum of the real fermion actions can be written as a complex fermion action,

$$\begin{aligned}&\chi_1^T\left[\partial_0+\hat{H}(Q^{\mathbb{R}})\right]\chi_1+\chi_2^T\left[\partial_0+\hat{H}(Q^{\mathbb{R}})\right]\chi_2\\&=\psi^\dagger\left[\partial_0+\hat{H}(Q^{\mathbb{R}})\right]\psi.\end{aligned}$$

Note that the cross terms cancel out, due to the anti-commutation relation between χ_1 and χ_2, and the fact that

$$\left[\partial_0+\hat{H}(Q^{\mathbb{R}})\right]^T=-\left[\partial_0+\hat{H}(Q^{\mathbb{R}})\right].$$

Consequently if $\tilde{W}[Q^{\mathbb{R}}(x)]$ is the effective action due to the complex fermion integration, we have

$$W[Q^{\mathbb{R}}(x)]=\frac{1}{2}\tilde{W}[Q^{\mathbb{R}}(x)]. \tag{B.18}$$

Due to Eq.(B.18), we shall focus on the complex fermion integration in the following.

B.4.2 Integrating out complex fermions

To make the action explicitly Lorentz invariant, we rewrite the fermion-boson action as

$$\begin{aligned}S&=\int d\tau\,d\boldsymbol{x}\;\psi^\dagger\Big[\partial_0-i\sum_{i=1}^{d}\Gamma_i\partial_i+m\hat{M}(\tau,\boldsymbol{x})\Big]\psi\\&=\int d\tau\,d\boldsymbol{x}\;\psi^\dagger(-i\gamma_0)\Big[(i\gamma_0)\partial_0-i(i\gamma_0)\sum_{i=1}^{d}\Gamma_i\partial_i+m(i\gamma_0)\hat{M}(\tau,\boldsymbol{x})\Big]\psi,\end{aligned} \tag{B.19}$$

where γ^0 is a $2n\times 2n$ hermitian matrix which anti-commutes with $\{\Gamma_i\}$ and satisfying $(\gamma^0)^2=1$. In general we choose γ^0 to be identity matrix among the flavor degrees of freedom. We will write down γ^0 explicitly for each dimension later on. Here we also extract out the parameter m, which controls the size of the fermion gap. As discussed in section 4.3, we will focus on the $\hat{M}$ belonging to the manifold manifold, i.e., satisfying $\hat{M}^2=1$. Now define

$$\begin{aligned}&\gamma^\mu:=(\gamma^0,-i\gamma^0\;\Gamma_i)\text{ where }i=1,...,d\\&\bar{\psi}:=\psi^\dagger(-i\gamma^0)\\&\boldsymbol{\beta}(\tau,\boldsymbol{x}):=\gamma^0\hat{M}(\tau,\boldsymbol{x})\end{aligned} \tag{B.20}$$

so that Eq.(B.19) turns into

$$S=\int d^Dx\;\bar{\psi}\left[i\gamma^\mu\partial_\mu+im\,\boldsymbol{\beta}(x)\right]\psi:=\int d^Dx\;\bar{\psi}\;\hat{\mathcal{D}}\;\psi$$
$$\text{where }\;\hat{\mathcal{D}}:=i\partial\!\!\!/+i\,m\,\boldsymbol{\beta}(x).$$

Using the anti-commutation relations between $\{\Gamma_i\}$ and γ^0, the γ^μ satisfies the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu}.$$

It's also easy to check that $\boldsymbol{\beta}(x)$, being a function of $Q^{\mathbb{C}}(x)$, is a matrix-valued smooth function of space-time, satisfying

$$\boldsymbol{\beta}(x)^\dagger \cdot \boldsymbol{\beta}(x) = 1.$$

Note that $\boldsymbol{\beta}(x)$ is in general not hermitian.

Fermion integration generates the effective action

$$W = -\ln\det[\hat{\mathcal{D}}] = -\mathrm{Tr}\ln[\hat{\mathcal{D}}].$$

The variation of the effective action W induced by a small variation in $\delta\boldsymbol{\beta}$ (triggered by a small variation in $Q^{\mathbb{C}}$ subject to the constraint $\boldsymbol{\beta}(x)^\dagger \cdot \boldsymbol{\beta}(x) = I$) is given by

$$\begin{aligned}
\delta W =& -\mathrm{Tr}\left[\left(\delta\hat{\mathcal{D}}\right)\hat{\mathcal{D}}^{-1}\right] \\
=& -\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\,\hat{\mathcal{D}}^{-1}\right] \\
=& -\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left(\hat{\mathcal{D}}^\dagger\hat{\mathcal{D}}\right)^{-1}\hat{\mathcal{D}}^\dagger\right] \\
=& -\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left[G_0^{-1} - m((\partial\!\!\!/\boldsymbol{\beta}))\right]^{-1}\hat{\mathcal{D}}^\dagger\right] \\
=& -\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left[G_0^{-1}\left(I - mG_0((\partial\!\!\!/\boldsymbol{\beta}))\right)\right]^{-1}\hat{\mathcal{D}}^\dagger\right] \\
=& -\mathrm{Tr}\left\{im\,\delta\boldsymbol{\beta}\left[\sum_{l=0}^{\infty}[mG_0\,((\partial\!\!\!/\boldsymbol{\beta}))\,]^l\right]G_0\left(i\partial\!\!\!/ - im\boldsymbol{\beta}^\dagger\right)\right\}
\end{aligned}$$

Here the double parentheses in $((\partial\!\!\!/\boldsymbol{\beta}))$ means that the derivative acts only on $\boldsymbol{\beta}$ and nothing afterward, and

$$G_0 := (-\partial^2 + m^2)^{-1}.$$

One can thus express δW in powers of $((\partial\!\!\!/\boldsymbol{\beta}))$. In the following, we shall retain terms where the number of space-time derivatives is less or equal to D. Hence by dimension counting, each of these terms is either relevant or marginal. The expansion is called the gradient expansion in the literature [24]. There are two types of terms having $\leq D$ derivatives, namely,

$$-\mathrm{Tr}\left\{im\,\delta\boldsymbol{\beta}\left[\sum_{l=0}^{D-1}[mG_0\,((\partial\!\!\!/\boldsymbol{\beta}))\,]^l\right]G_0\left(i\partial\!\!\!/\right)\right\} \tag{B.21}$$

and

$$-\mathrm{Tr}\left\{im\,\delta\boldsymbol{\beta}\left[\sum_{l=0}^{D}[mG_0\,((\partial\!\!\!/\boldsymbol{\beta}))\,]^l\right]G_0\left(-im\boldsymbol{\beta}^\dagger\right)\right\} \tag{B.22}$$

It turns out that among all non-vanishing parts of Eq.B.21 and B.22 there is a unique pure imaginary term – the WZW term. The rest are real. In $D = 1 + 1$ and $2 + 1$ the only real term is the stiffness term. In $D = 3 + 1$ there are several extra real terms in addition to the stiffness term. However, all of these extra terms contain four space-time derivatives. Hence they are irrelevant compared with the stiffness term. Therefore the non-linear sigma model with the WZW term contains the most relevant real and imaginary terms after the fermion integration. To avoid sidetracking, we shall leave these details in subsection B.4.4.

Throughout this appendix, we shall adopt the following convention. Tr denotes the trace over both the space-time and the matrices in $\boldsymbol{\beta}$, $\delta\boldsymbol{\beta}$, and γ^μs. tr′ denotes the trace over the matrices in $\boldsymbol{\beta}$, $\delta\boldsymbol{\beta}$, γ^μ. tr_γ denotes the trace over only the γ matrices. tr denotes the trace over the $n \times n$ matrices in $\boldsymbol{\beta}, \delta\boldsymbol{\beta}$. According to the above convention

$$\mathrm{tr}' = \mathrm{tr}_\gamma \times \mathrm{tr}.$$

Moreover, we shall adopt the following short hand

$$\int \frac{d^D k}{(2\pi)^D} := \int_k$$

The stiffness term

The first non-vanishing such term is the stiffness term,

$$\delta W_{\text{stiff}} = \quad -\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\not\partial\boldsymbol{\beta}))\right)G_0\left(i\not\partial\right)\right]. \tag{B.23}$$

Fourier transforming Eq.(B.23), we obtain

$$\begin{aligned}
\delta W_{\text{stiff}} &= -m^2 \int_p \int_q \mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\frac{1}{(p+q)^2+m^2}\not q\boldsymbol{\beta}_q\frac{1}{p^2+m^2}\not p\right] \\
&\approx 2m^2 \int_p \frac{1}{(p^2+m^2)^3}\int_q (q\cdot p)\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}_q^\dagger \not q\not p\right] \\
&= 2m^2 \int_p \frac{p_\mu p_\nu}{(p^2+m^2)^3}\int_q (q_\mu q_\lambda)\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}_q^\dagger\gamma^\nu\gamma^\lambda\right]
\end{aligned} \tag{B.24}$$

As usual

$$\not p := \gamma^\mu p_\mu.$$

In passing to the second line in Eq.(B.24) we have expanded the expression

$$\frac{1}{(p+q)^2+m^2} = \frac{1}{p^2+m^2}\sum_{n=0}^{\infty}\left(-\frac{2p\cdot q+q^2}{p^2+m^2}\right)^n \tag{B.25}$$

and keep the lowest order non-vanishing term. Because

$$\int_p \frac{1}{(p^2+m^2)^3} p_\mu p_\nu = \frac{1}{D}\int_p \frac{p^2}{(p^2+m^2)^3}\delta_{\mu\nu},$$

Eq.(B.24) turns into

$$\begin{aligned}\delta W_{\text{stiff}} &\approx \frac{2m^2}{D}\int_p \frac{p^2}{(p^2+m^2)^3}\int_q (q_\mu q_\lambda)\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}_q^\dagger\gamma^\mu\gamma^\lambda\right] \\ &= \frac{2m^2}{D}\int_p \frac{p^2}{(p^2+m^2)^3}\int_q q^2\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}_q^\dagger\right]\end{aligned} \tag{B.26}$$

In passing to the second line of Eq.(B.26) we have used the fact that

$$q_\mu q_\lambda \gamma^\mu\gamma^\lambda = q^2 I.$$

The p-integration in Eq.(B.26) converges for $(1+1)$-D and $(2+1)$-D, but diverges for $(3+1)$-D. We shall use dimensional regularization ($D = 4-\epsilon$ with $\epsilon \to 0^+$), which leads to

$$\begin{aligned}\delta W_{\text{stiff}} \approx& 2m^2\frac{1}{D}\int_p \frac{p^2}{(p^2+m^2)^3}\int_q q^2\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}_q^\dagger\right] \\ =&\left[\frac{\Gamma(2-\frac{D}{2})}{2(4\pi)^{D/2}}m^{D-2}\right]\int_{\mathcal{M}} d^Dx\,\mathrm{tr}'\left[\partial_\mu(\delta\boldsymbol{\beta})\partial^\mu\boldsymbol{\beta}^\dagger\right]\end{aligned} \tag{B.27}$$

Here $\Gamma(l)$ is the gamma function. For $(3+1)$-D, the dimension regularization is given by

$$\Gamma(2-\frac{D}{2}) = \Gamma(\frac{\epsilon}{2}) \approx \frac{2}{\epsilon} - \gamma + O(\epsilon)$$

where γ is the Euler-Mascheroni constant. Thus, the term whose variation with respect to $\delta\boldsymbol{\beta}$ yields Eq.(B.27) is

$$W_{\text{stiff}}[\boldsymbol{\beta}] = \frac{1}{2\lambda_D^{D-2}}\int_{\mathcal{M}} d^Dx\,\mathrm{tr}'\left[\partial_\mu\boldsymbol{\beta}\partial^\mu\,\boldsymbol{\beta}^\dagger\right] \tag{B.28}$$

where λ has the dimension of length. In the limit that the short-distance cutoff is zero,

$$\frac{1}{\lambda_D^{D-2}} = \left[\frac{\Gamma(2-\frac{D}{2})}{2(4\pi)^{D/2}}m^{D-2}\right]. \tag{B.29}$$

The WZW (topological) term

The second type of non-vanishing term in the gradient expansion is topological in nature, namely, the Wess-Zumino-Witten term

$$\begin{aligned}
\delta W_{\text{WZW}} &= -\text{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\not\partial\boldsymbol{\beta}))\right)^D G_0\left(-im\boldsymbol{\beta}^\dagger\right)\right] \\
&\approx -m^{D+2}\left[\int_p \frac{1}{(p^2+m^2)^{D+1}}\right]\int_{\mathcal{M}} d^Dx\,\text{tr}'\left[\prod_{a=1}^{D}(\gamma^{\mu_a}\partial_{\mu_a}\boldsymbol{\beta})\,\boldsymbol{\beta}^\dagger\,\delta\boldsymbol{\beta}\right] \\
&= -\left[\frac{1}{(4\pi)^{D/2}}\frac{\Gamma(\frac{D}{2}+1)}{\Gamma(D+1)}\right]\int_{\mathcal{M}} d^Dx\,\text{tr}'\left[\prod_{a=1}^{D}(\gamma^{\mu_a}\partial_{\mu_a}\boldsymbol{\beta})\,\boldsymbol{\beta}^\dagger\,\delta\boldsymbol{\beta}\right]
\end{aligned} \tag{B.30}$$

Eq.(B.30) is the difference in the Berry phase between the order parameter configurations $\boldsymbol{\beta}(x)$ and $\boldsymbol{\beta}(x)+\delta\boldsymbol{\beta}(x)$. To determine the Berry phase for a specific $\boldsymbol{\beta}(x)$, we integrate Eq.(B.30) from a reference configuration $\boldsymbol{\beta}(x)$ = constant matrix. The existence of a continuous retraction leading from $\boldsymbol{\beta}(x)$ to the reference configuration relies on

$$\pi_D(\text{mass manifold}) = 0.$$

It turns out this is exactly the condition when the WZW term exists (see later). Under the condition that such continuous retraction exists, we can find a continuous family of configurations $\tilde{\boldsymbol{\beta}}(x,u)$ so that

$$\begin{aligned}
&\tilde{\boldsymbol{\beta}}(x,u=1) = \boldsymbol{\beta}(x) \\
&\tilde{\boldsymbol{\beta}}(x,u=0) = \text{constant matrix.}
\end{aligned} \tag{B.31}$$

We can integrate Eq.(B.30) to yield

$$W_{WZW}[\boldsymbol{\beta}] = -\left[\frac{1}{(4\pi)^{D/2}}\frac{\Gamma(\frac{D}{2}+1)}{\Gamma(D+1)}\right]\int\limits_{\mathcal{B}} du\,d^Dx\,\text{tr}'\left[\prod_{a=1}^{D}(\gamma^{\mu_a}\partial_{\mu_a}\tilde{\boldsymbol{\beta}})\,\tilde{\boldsymbol{\beta}}^\dagger\,\partial_u\tilde{\boldsymbol{\beta}}\right] \tag{B.32}$$

As in the main text, $\mathcal{B}$ is the extension of space-time manifold $\mathcal{M}$, so that

$$\partial\mathcal{B} = \mathcal{M}.$$

In summary, when the fermion flavor number, n, is sufficiently large so that the WZW term is stabilized, the non-linear sigma model action is

$$\begin{aligned}
W[\boldsymbol{\beta}] =& \frac{1}{2\lambda_D^{D-2}}\int_{\mathcal{M}} d^Dx\,\text{tr}'\left[\partial_\mu\boldsymbol{\beta}\partial^\mu\boldsymbol{\beta}^\dagger\right] \\
&-\left[\frac{1}{(4\pi)^{D/2}}\frac{\Gamma(\frac{D}{2}+1)}{\Gamma(D+1)}\right]\int\limits_{\mathcal{B}} du\,d^Dx\,\text{tr}'\left[\prod_{a=1}^{D}(\gamma^{\mu_a}\partial_{\mu_a}\tilde{\boldsymbol{\beta}})\,\tilde{\boldsymbol{\beta}}^\dagger\,\partial_u\tilde{\boldsymbol{\beta}}\right] \\
\text{where } \frac{1}{\lambda_D^{D-2}} =& \left[\frac{\Gamma(2-\frac{D}{2})}{2(4\pi)^{D/2}}m^{D-2}\right].
\end{aligned} \tag{B.33}$$

In the following, we shall apply this result to $(1+1)$-D, $(2+1)$-D, and $(3+1)$-D [3].

B.4.3 The fermion integration results for sufficiently large n so that the WZW term is stabilized

In this subsection, we shall focus on the results of fermion integration when n is sufficiently large so that

$$\pi_{D+1}(\text{mass manifold}) = \mathbb{Z}.$$

The case of small n, before the WZW term is stabilized, will be discussed in appendix B.8.

Complex class in $(1+1)$-D

The fermion action for the complex class in $(1+1)$-D is given by Eq.(4.22)

$$\begin{aligned}
S &= \int d^2x\ \psi^\dagger \left[\partial_0 - i(ZI_n)\partial_1 + m\left(X\otimes H_1 + Y\otimes H_2\right)\right]\psi \\
&= \int d^2x\ \psi^\dagger(-iXI_n)\left[i(XI_n)\partial_0 + i(-YI_n)\partial_1 + im\left(I\otimes H_1 + iZ\otimes H_2\right)\right]\psi \\
&:= \int d^2x\ \bar{\psi}\left[i\partial\!\!\!/ + i\,m\,\boldsymbol{\beta}\right]\psi
\end{aligned}$$

where

$$\begin{aligned}
&\bar{\psi} = \psi^\dagger(-iXI_n) \\
&\gamma^0 = XI_n,\ \gamma^1 = -YI_n,\ \gamma^5 = ZI_n \\
&\boldsymbol{\beta} = I\otimes H_1 + i\gamma^5 I\otimes H_2.
\end{aligned} \tag{B.34}$$

Plugging the above results into equation Eq.(B.28) and Eq.(B.29) the stiffness term is given by

$$\begin{aligned}
W_{\text{stiff}}[Q^{\mathbb{C}}] =& \frac{1}{8\pi}\int d^2x\,\mathrm{tr}\left[\partial_\mu(H_1+iH_2)\partial^\mu(H_1-iH_2)\right] \\
=& \frac{1}{8\pi}\int d^2x\,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right]
\end{aligned}$$

where $Q^{\mathbb{C}} := H_1 + iH_2 \in U(n)$. Substitute Eq.(B.34) into Eq.(B.32) we obtain WZW term as

[3] It turns out that it always contains a level-1 WZW term in even higher dimensions, though we shall not discuss them in the present paper.

$$\begin{aligned}
W_{WZW}[Q^{\mathbb{C}}] =& -\frac{1}{8\pi}\int_{\mathcal{B}} du\, d^2x\, \mathrm{tr}'\left[(\gamma^{\mu_1}\partial_{\mu_1}\tilde{\boldsymbol{\beta}})(\gamma^{\mu_2}\partial_{\mu_2}\tilde{\boldsymbol{\beta}})\,\tilde{\boldsymbol{\beta}}^{\dagger}\,\partial_u\tilde{\boldsymbol{\beta}}\right] \\
=& -\frac{1}{8\pi}\int_{\mathcal{B}} du\, d^2x\, \mathrm{tr}'\Big[(\gamma^{\mu_1}\gamma^{\mu_2})\,\Big(I\otimes\tilde{H}_1 - i\gamma^5 I\otimes\tilde{H}_2\Big)\,\partial_u\Big(I\otimes\tilde{H}_1 + i\gamma^5 I\otimes\tilde{H}_2\Big) \\
&\times\partial_{\mu_1}\Big(I\otimes\tilde{H}_1 - i\gamma^5 I\otimes\tilde{H}_2\Big)\,\partial_{\mu_2}\Big(I\otimes\tilde{H}_1 + i\gamma^5 I\otimes\tilde{H}_2\Big)\Big] \\
=& -\frac{1}{8\pi}\int_{\mathcal{B}} du\, d^2x\, \mathrm{tr}'\Big[\left(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^5\right)\,\Big(I\otimes\tilde{H}_1 - iI\otimes\tilde{H}_2\Big)\,\partial_u\Big(I\otimes\tilde{H}_1 + iI\otimes\tilde{H}_2\Big) \\
&\times\partial_{\mu_1}\Big(I\otimes\tilde{H}_1 - iI\otimes\tilde{H}_2\Big)\,\partial_{\mu_2}\Big(I\otimes\tilde{H}_1 + iI\otimes\tilde{H}_2\Big)\Big] \\
=& -\frac{1}{8\pi}\int_{\mathcal{B}} du\, d^2x\, \left(-2i\epsilon^{\mu_1\mu_2}\right)\,\mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}\dagger}\,\partial_u\tilde{Q}^{\mathbb{C}}\,\partial_{\mu_1}\tilde{Q}^{\mathbb{C}\dagger}\,\partial_{\mu_2}\tilde{Q}^{\mathbb{C}}\Big] \\
=& -\frac{i}{4\pi}\int_{\mathcal{B}} du\, d^2x\, \epsilon^{\mu_1\mu_2}\,\mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_u\tilde{Q}^{\mathbb{C}}\Big)\,\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_1}\tilde{Q}^{\mathbb{C}}\Big)\,\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_2}\tilde{Q}^{\mathbb{C}}\Big)\Big] \\
=& -\frac{i}{4\pi}\times\frac{1}{3}\int_{\mathcal{B}} du\, d^2x\, \epsilon^{\tilde{\mu}_1\tilde{\mu}_2\tilde{\mu}_3}\,\mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_1}\tilde{Q}^{\mathbb{C}}\Big)\,\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_2}\tilde{Q}^{\mathbb{C}}\Big)\,\Big(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_3}\tilde{Q}^{\mathbb{C}}\Big)\Big] \\
=& -\frac{2\pi i}{24\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\Big)^3\Big]
\end{aligned}$$

where

$$\tilde{Q}^{\mathbb{C}} := \tilde{H}_1 + i\tilde{H}_2 \in U(n)$$

is the extension of $Q^{\mathbb{C}} = H_1 + iH_2$ into $\mathcal{B}$, and $\tilde{\mu}_a$ is extended space-time manifold index (i.e., they include u). This extension in Eq.(B.31) is possible because

$$\pi_2(U(N)) = 0.$$

The passing from the 2nd to the 3rd line is due to the fact that when $\tilde{H}_2 \to -\tilde{H}_2$ the whole expression changes sign, hence only the terms with an odd number of $\tilde{H}_2$ survive. The final non-linear sigma model action, namely,

$$W[Q^{\mathbb{C}}] = \frac{1}{8\pi}\int d^2x\,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right] - \frac{2\pi i}{24\pi^2}\int_{\mathcal{B}}\mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\Big)^3\Big] \tag{B.35}$$

is the $U(n)_{k=1}$ WZW theory.

Real class in $(1+1)$-D

The Majorana fermion action for the real class in (1 + 1)-D is given by Eq.(4.26)

$$S = \int d^2x \; \chi^T \left[\partial_0 + i(ZI_n)\partial_1 + m\left(X \otimes (iA) + Y \otimes S\right)\right] \chi,$$

where S and A are symmetric and anti-symmetric matrices, respectively. Upon the complexification described in subsection B.4.1, the form of the action becomes exactly the same as the complex class action in the preceding section, except that

$$H_1 \to iA \quad H_2 \to S.$$

Following the discussion in subsection B.4.1, we can substitute

$$Q^{\mathbb{C}} = H_1 + iH_2 \to i(A + S) := iQ^{\mathbb{R}}$$

into Eq.(B.35) and divide the result by 2 to obtain the following non-linear sigma model action

$$W[Q^{\mathbb{R}}] = \frac{1}{16\pi}\int d^2x \,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}} \partial^\mu (Q^{\mathbb{R}})^T\right] - \frac{2\pi i}{48\pi^2}\int_{\mathcal{B}} \mathrm{tr}\Big\{ \left[(\tilde{Q}^{\mathbb{R}})^T d\tilde{Q}^{\mathbb{R}}\right]^3 \Big\} \tag{B.36}$$

This is the action of the $O(n)_{k=1}$ WZW theory.

Complex class in $(2+1)$-D

The fermion action for complex class in (2 + 1)-D can be constructed from Eq.(B.9) and Eq.(B.10),

$$\begin{aligned} S &= \int d^3x \; \psi^\dagger \left[\partial_0 - i(ZI_n)\partial_1 - i(XI_n)\partial_2 + m\, Y \otimes Q^{\mathbb{C}}\right] \psi \\ &= \int d^3x \; \psi^\dagger (-iYI_n) \left[i(YI_n)\partial_0 + i(XI_n)\partial_1 + i(-ZI_n)\partial_2 + im\, I \otimes Q^{\mathbb{C}}\right] \psi \\ &:= \int d^3x \; \bar{\psi} \left[i\partial\!\!\!/ + im\boldsymbol{\beta}\right] \psi \end{aligned}$$

where

$$\begin{aligned} &\bar{\psi} = \psi^\dagger(-iYI_n) \\ &\gamma^0 = YI_n, \; \gamma^1 = XI_n, \; \gamma^2 = -ZI_n \\ &\boldsymbol{\beta} = I \otimes Q^{\mathbb{C}}. \end{aligned} \tag{B.37}$$

Here $Q^{\mathbb{C}}(x)$ is an $n \times n$ hermitian-matrix-value function satisfying $\left(Q^{\mathbb{C}}\right)^2 = I_n$, forming the mass manifold $\bigcup_{l=0}^{n} \frac{U(n)}{U(l)\times U(n-l)}$ (see appendix B.2). As discussed earlier, the $l = n/2$

component is special because the full emergent symmetries of the fermion theory can be restored upon order parameter fluctuation. Hence as far as bosonization is concerned we will focus on $l = n/2$. However, the following derivation works for other values of l too as long as both l and $n - l$ are sufficiently large for the WZW term to be stabilized.

Substitute Eq.(B.37) into Eq.(B.28) and Eq.(B.29), we obtain the following stiffness term

$$\begin{aligned} W_{\text{stiff}}[Q^{\mathbb{C}}] =& \frac{1}{4\lambda_3} \int d^3x \, \text{tr}' \left[\partial_\mu \boldsymbol{\beta} \partial^\mu \boldsymbol{\beta}^\dagger \right] \\ =& \frac{1}{2\lambda_3} \int d^3x \, \text{tr} \left[\partial_\mu Q^{\mathbb{C}} \partial^\mu Q^{\mathbb{C}} \right], \end{aligned}$$

where λ_3 has the dimension of length and in the limit where the short-distance cutoff is zero,

$$\lambda_3 = \frac{8\pi}{m}. \tag{B.38}$$

Substitution of Eq.(B.37) into equation (B.32) yields the WZW term

$$\begin{aligned} W_{\text{WZW}}[Q^{\mathbb{C}}] =& - \left[\frac{1}{(4\pi)^{3/2}} \frac{\Gamma(\frac{5}{2})}{\Gamma(4)} \right] \int_{\mathcal{B}} du \, d^3x \, \text{tr}' \left[\prod_{a=1}^{D} (\gamma^{\mu_a} \partial_{\mu_a} \tilde{\boldsymbol{\beta}}) \, \tilde{\boldsymbol{\beta}}^\dagger \, \partial_u \tilde{\boldsymbol{\beta}} \right] \\ =& - \frac{i}{32\pi} \int_{\mathcal{B}} du \, d^3x \, \epsilon^{\mu_1\mu_2\mu_3} \, \text{tr} \left[\tilde{Q}^{\mathbb{C}} \, \partial_u \tilde{Q}^{\mathbb{C}} \, \partial_{\mu_1} \tilde{Q}^{\mathbb{C}} \, \partial_{\mu_2} \tilde{Q}^{\mathbb{C}} \, \partial_{\mu_3} \tilde{Q}^{\mathbb{C}} \right] \\ =& - \frac{i}{128\pi} \int_{\mathcal{B}} du \, d^3x \, \epsilon^{\tilde{\mu}_1\tilde{\mu}_2\tilde{\mu}_3\tilde{\mu}_4} \, \text{tr} \left[\tilde{Q}^{\mathbb{C}} \, \partial_{\tilde{\mu}_1} \tilde{Q}^{\mathbb{C}} \, \partial_{\tilde{\mu}_2} \tilde{Q}^{\mathbb{C}} \, \partial_{\tilde{\mu}_3} \tilde{Q}^{\mathbb{C}} \, \partial_{\tilde{\mu}_4} \tilde{Q}^{\mathbb{C}} \right] \\ =& - \frac{2\pi i}{256\pi^2} \int_{\mathcal{B}} \text{tr} \left[\tilde{Q}^{\mathbb{C}} \left(d\tilde{Q}^{\mathbb{C}} \right)^4 \right]. \end{aligned}$$

Here $\tilde{Q}^{\mathbb{C}}$ is the extension of $Q^{\mathbb{C}}$ into $\mathcal{B}$, and $\tilde{\mu}_a$ extended space-time index. The extension in Eq.(B.31) is possible because

$$\pi_3 \left(\frac{U(n)}{U(n/2) \times U(n/2)} \right) = 0.$$

The existence of WZW is indicated by

$$\pi_4 \left(\frac{U(n)}{U(n/2) \times U(n/2)} \right) = \mathbb{Z},$$

with the topological invariant

$$\frac{1}{256\pi^2}\int_{S^4} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right] \in \mathbb{Z}.$$

Comparing with the result of fermion integration, the WZW term is $2\pi i$ times the above topological invariant, implying the level, k, is 1. In summary, the non-linear sigma model action is

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_3}\int d^3x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}}\right] - \frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right]. \tag{B.39}$$

Real class in $(2+1)$-D

The fermion action for real class in $(2+1)$-D can be constructed from Eq.(B.11) and Eq.(B.12),

$$S = \int d^3x\; \chi^T\left[\partial_0 + i(ZI_n)\partial_1 + i(XI_n)\partial_2 + mYQ^{\mathbb{R}}\right]\chi$$

Note that the form of this action is the same as that in the preceding section, except that the fermions are Majorana and $Q^{\mathbb{R}}$ is real symmetric instead of hermitian. According to the discussion in subsection B.4.1, we can replace

$$Q^{\mathbb{C}} \to Q^{\mathbb{R}} \in \frac{O(n)}{O(n/2)\times O(n/2)}$$

in Eq.(B.39) and divide the result by 2. The resulting non-linear sigma model action is

$$W[Q^{\mathbb{R}}] = \frac{1}{4\lambda_3}\int d^3x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}}\partial^\mu Q^{\mathbb{R}}\right] - \frac{2\pi i}{512\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{R}}\,(d\tilde{Q}^{\mathbb{R}})^4\right]. \tag{B.40}$$

Here λ_3 has the dimension of length and in the limit where the short-distance cutoff is zero λ_3 is given by Eq.(B.38). Moreover, $\tilde{Q}^{\mathbb{R}}$ is the extension of $Q^{\mathbb{R}}$ into $\mathcal{B}$. The extension in Eq.(B.40) is possible because

$$\pi_3\left(\frac{O(n)}{O(n/2)\times O(n/2)}\right) = 0.$$

The existence of the WZW term is indicated by

$$\pi_4\left(\frac{O(n)}{O(n/2)\times O(n/2)}\right) = \mathbb{Z},$$

with the topological invariant given by

$$\frac{1}{512\pi^2}\int_{S^4} \mathrm{tr}\left[\tilde{Q}^{\mathbb{R}}\left(d\tilde{Q}^{\mathbb{R}}\right)^4\right] \in \mathbb{Z}.$$

Comparing the WZW term with the topological invariant we conclude Eq.(B.40) is the action for the $\frac{O(n)}{O(n/2)\times O(n/2)}$ non-linear sigma model with $k=1$ WZW term.

Complex class in $(3+1)$-D

The fermion action for complex class in $(3+1)$-D can be constructed from Eq.(B.13) and Eq.(B.14),

$$\begin{aligned}
S =& \int d^4x\ \psi^\dagger\left[\partial_0 - i(ZII_n)\partial_1 - i(XII_n)\partial_2 - i(YZI_n)\partial_3 + m\left(YX\otimes H_1 + YY\otimes H_2\right)\right]\psi \\
=& \int d^4x\ \psi^\dagger(-iYXI_n)\Big[i(YXI_n)\partial_0 + i(XXI_n)\partial_1 + i(-ZXI_n)\partial_2 + i(-IYI_n)\partial_3 \\
&+ im\left(II\otimes H_1 + i\,IZ\otimes H_2\right)\Big]\psi \\
:=& \int d^4x\ \bar{\psi}\left[i\partial\!\!\!/ + im\boldsymbol{\beta}\right]\psi
\end{aligned}$$

where

$$\begin{aligned}
&\bar{\psi} = \psi^\dagger(-iYXI_n)\\
&\gamma^0 = YXI_n,\ \gamma^1 = XXI_n\ ,\gamma^2 = -ZXI_n,\ \gamma^3 = -IYI_n,\ \gamma^5 = IZI_n\\
&\boldsymbol{\beta} = II\otimes H_1 + i\gamma^5 II\otimes H_2.
\end{aligned} \tag{B.41}$$

Substitute Eq.(B.41) into Eq.(B.28) and Eq.(B.29), the stiffness term read

$$\begin{aligned}
W_{\text{stiff}}[Q^{\mathbb{C}}] =& \frac{1}{8\lambda_4^2}\int d^4x\, \mathrm{tr}'\left[\partial_\mu\boldsymbol{\beta}\partial^\mu\boldsymbol{\beta}^\dagger\right]\\
=& \frac{1}{2\lambda_4^2}\int d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right]
\end{aligned}$$

where

$$Q^{\mathbb{C}} = H_1 + iH_2 \in U(n).$$

The parameter λ_4 has the dimension of length and in the limit where the short-distance cutoff is zero,

$$\frac{1}{\lambda_4^2} = \left[\frac{\Gamma(0^+)m^2}{8\pi^2}\right]. \tag{B.42}$$

In the case where the short distance cutoff is finite the coefficient $\Gamma(0^+)$ should be replaced by a cutoff dependent parameter. Substitution Eq.(B.41) into Eq.(B.32) yields the WZW term

$$
\begin{aligned}
W_{\rm WZW}[Q^{\mathbb{C}}] =& -\left[\frac{1}{(4\pi)^2}\frac{\Gamma(3)}{\Gamma(5)}\right]\int\limits_{\mathcal{B}} du\, d^4x\, \mathrm{tr}'\left[(\gamma^{\mu_1}\partial_{\mu_1}\tilde{\boldsymbol{\beta}})(\gamma^{\mu_2}\partial_{\mu_2}\tilde{\boldsymbol{\beta}})\,(\gamma^{\mu_3}\partial_{\mu_3}\tilde{\boldsymbol{\beta}})\,(\gamma^{\mu_4}\partial_{\mu_4}\tilde{\boldsymbol{\beta}})\,\tilde{\boldsymbol{\beta}}^\dagger\,\partial_u\tilde{\boldsymbol{\beta}}\right] \\
=& -\frac{1}{192\pi^2}\int\limits_{\mathcal{B}} du\, d^4x\, \mathrm{tr}'\Big[\,(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_4})\,\left(I\otimes\tilde{H}_1 - i\gamma^5 I\otimes\tilde{H}_2\right)\,\partial_u\left(I\otimes\tilde{H}_1 + i\gamma^5 I\otimes\tilde{H}_2\right) \\
&\times\partial_{\mu_1}\left(I\otimes\tilde{H}_1 - i\gamma^5 I\otimes\tilde{H}_2\right)\,\partial_{\mu_2}\left(I\otimes\tilde{H}_1 + i\gamma^5 I\otimes\tilde{H}_2\right)\Big] \\
&\times\partial_{\mu_3}\left(I\otimes\tilde{H}_1 - i\gamma^5 I\otimes\tilde{H}_2\right)\,\partial_{\mu_4}\left(I\otimes\tilde{H}_1 + i\gamma^5 I\otimes\tilde{H}_2\right)\Big] \\
=& -\frac{1}{192\pi^2}\int\limits_{\mathcal{B}} du\, d^4x\, \mathrm{tr}'\Big[\,(\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3}\gamma^{\mu_4}\gamma^5)\,\left(I\otimes\tilde{H}_1 - iI\otimes\tilde{H}_2\right)\,\partial_u\left(I\otimes\tilde{H}_1 + iI\otimes\tilde{H}_2\right) \\
&\times\partial_{\mu_1}\left(I\otimes\tilde{H}_1 - iI\otimes\tilde{H}_2\right)\,\partial_{\mu_2}\left(I\otimes\tilde{H}_1 + iI\otimes\tilde{H}_2\right)\Big] \\
&\times\partial_{\mu_3}\left(I\otimes\tilde{H}_1 - iI\otimes\tilde{H}_2\right)\,\partial_{\mu_4}\left(I\otimes\tilde{H}_1 + iI\otimes\tilde{H}_2\right)\Big] \\
=& -\frac{1}{192\pi^2}\int\limits_{\mathcal{B}} du\, d^4x\,\left(4\epsilon^{\mu_1\mu_2\mu_3\mu_4}\right)\,\mathrm{tr}\left[\tilde{Q}^{\mathbb{C}\dagger}\,\partial_u\tilde{Q}^{\mathbb{C}}\,\partial_{\mu_1}\tilde{Q}^{\mathbb{C}\dagger}\,\partial_{\mu_2}\tilde{Q}^{\mathbb{C}}\,\partial_{\mu_3}\tilde{Q}^{\mathbb{C}\dagger}\,\partial_{\mu_4}\tilde{Q}^{\mathbb{C}}\right] \\
=& -\frac{1}{48\pi^2}\int\limits_{\mathcal{B}} du\, d^4x\,\epsilon^{\mu_1\mu_2\mu_3\mu_4}\,\mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_u\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_1}\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_2}\tilde{Q}^{\mathbb{C}}\right) \\
&\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_3}\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\mu_4}\tilde{Q}^{\mathbb{C}}\right)\Big] \\
=& -\frac{1}{240\pi^2}\int\limits_{\mathcal{B}} du\, d^4x\,\epsilon^{\tilde{\mu}_1\tilde{\mu}_2\tilde{\mu}_3\tilde{\mu}_4\tilde{\mu}_5}\,\mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_1}\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}_3^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_2}\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}_3^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_3}\tilde{Q}^{\mathbb{C}}\right) \\
&\left(\tilde{Q}_3^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_4}\tilde{Q}^{\mathbb{C}}\right)\left(\tilde{Q}_3^{\mathbb{C}\dagger}\partial_{\tilde{\mu}_5}\tilde{Q}^{\mathbb{C}}\right)\Big] \\
=& -\frac{2\pi}{480\pi^3}\int\limits_{\mathcal{B}}\mathrm{tr}\left[\left(\tilde{Q}^{\mathbb{C}\dagger}d\tilde{Q}^{\mathbb{C}}\right)^5\right]
\end{aligned}
$$

where $\tilde{Q}^{\mathbb{C}}$ is the extension of $Q^{\mathbb{C}}$ into $\mathcal{B}$, and $\tilde{\mu}_a$ is the coordinate index of the extended space-time manifold. The extension in Eq.(B.31) is possible because

$$\pi_4(U(n)) = 0.$$

In passing from the 2nd to the 3rd line is due to the fact that when $\tilde{H}_2 \to -\tilde{H}_2$, the entire expression changes sign, hence only terms with an odd number of $\tilde{H}_2$ survive. The existence of the WZW term is indicated by

$$\pi_5(U(n)) = \mathbb{Z},$$

with the topological invariant given by

$$\frac{i}{480\pi^3}\int_{S^5} \mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\Big)^5\Big] \in \mathbb{Z}.$$

Comparing the WZW term with the topological invariant we conclude the WZW term is at level $k = 1$. In summary, the non-linear sigma model action is given by

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_4^2}\int d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}} \partial^\mu Q^{\mathbb{C}\dagger}\right] - \frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\Big)^5\Big]. \tag{B.43}$$

Real class in $(3+1)$-D

The fermion action for complex class in $(3+1)$-D can be constructed from Eq.(B.15) and Eq.(B.16),

$$S = \int d^4x\ \chi^T \left[\partial_0 - i(XII_n)\partial_1 - i(ZII_n)\partial_2 - i(YYI_n)\partial_3 + m\left(YX \otimes S_1 + YZ \otimes S_2\right)\right]\chi$$

This action has the same form as that in the preceding section, except that the following differences. (i) The fermions are Majorana, (ii) an unitary change of the matrix basis, namely, rotation by $\pi/2$ generated by IXI_n , and (iii) $H_1 \to S_1$ and $H_2 \to S_2$. According to the discussion in subsection B.4.1, we can use the result in the preceding section by substituting $Q^{\mathbb{C}} \to Q^{\mathbb{R}} = S_1 + iS_2$ into Eq.(B.43) and divide the final effective action by 2. The resulting non-linear sigma model action is given by

$$W[Q^{\mathbb{R}}] = \frac{1}{4\lambda_4^2}\int d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{R}} \partial^\mu Q^{\mathbb{R}\dagger}\right] - \frac{2\pi}{960\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{R}\dagger} d\tilde{Q}^{\mathbb{R}}\Big)^5\Big]. \tag{B.44}$$

The existence of the WZW iterm is indicated by

$$\pi_5(U(n)/O(n)) = \mathbb{Z},$$

with the topological invariant given by

$$\frac{i}{960\pi^3}\int_{S^5} \mathrm{tr}\Big[\Big(\tilde{Q}^{\mathbb{R}\dagger} d\tilde{Q}^{\mathbb{R}}\Big)^5\Big] \in \mathbb{Z}.$$

Again, we conclude that the WZW term is at level 1.

B.4.4 The less relevant real terms originate from Eq.(B.21) and Eq.(B.22)

In this subsection, we provide the details which show that in $(1+1)$-D and $(2+1)$-D, the non-vanishing terms in Eq.(B.21) and Eq.(B.22) having $\leq D$ space-time derivatives are the stiffness and WZW terms. In $(3+1)$-D, there are extra real terms. In the following, we shall present a detailed analysis of these potential extra terms.

$(1+1)$**-D**

Given the fact that

$$\boldsymbol{\beta} = I \otimes H_1 + i\gamma^5 I \otimes H_2,$$

the $l=0$ term in Eq.(B.21) read

$$-\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta} G_0 \gamma^\mu (i\partial_\mu)\right].$$

Since $\boldsymbol{\beta}$ only contains I or γ^5, it follows that the this term vanishes when we trace over the gamma matrices because both

$$\mathrm{tr}_\gamma\left[\gamma^\mu\right] = 0, \quad \mathrm{tr}_\gamma[\gamma^\mu\gamma^5] = 0.$$

Similar argument applies to the $l=1$ term in Eq.(B.22).

This leaves the $l=0$ term in Eq.(B.22) as the only term requiring further attention, namely,

$$\begin{aligned}
&-\mathrm{Tr}\left[im\delta\boldsymbol{\beta} G_0(-i\,m\,\boldsymbol{\beta})\right] \\
&= -m^2 \int_{p,q} \frac{1}{p^2+m^2} \mathrm{tr}' \left[\boldsymbol{\beta}^\dagger_{-q} \delta \boldsymbol{\beta}_q\right] \\
&= -m^2 \left(\int_p \frac{1}{p^2+m^2}\right) \int d^2x\, \mathrm{tr}' \left[\boldsymbol{\beta}^\dagger \delta \boldsymbol{\beta}\right] \\
&= -2m^2 \left(\int_p \frac{1}{p^2+m^2}\right) \int d^2x\, \delta \left\{\mathrm{tr}\left[H_1^2 + H_2^2\right]\right\} \\
&= 0.
\end{aligned}$$

In passing to the last line we used the constraint that

$$Q^{\mathbb{C}} = H_1 + iH_2 \in U(n) \Rightarrow H_1^2 + H_2^2 = I_n.$$

Hence the only non-vanishing terms are the stiffness and WZW terms in subsection B.4.2 and B.4.2.

$(2+1)$**-D**

Given the fact that

$$\boldsymbol{\beta} = I \otimes Q^{\mathbb{C}},$$

both the $l=0$ term in Eq.(B.21) and the $l=1$ term in Eq.(B.22) vanishes under tr_γ because

$$\mathrm{tr}_\gamma[\gamma^\mu] = 0.$$

The $l=2$ term in Eq.(B.21) gives

$$\begin{aligned}
&-\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\not\partial\boldsymbol{\beta}))\right)^2 G_0\left(i\not\partial\right)\right] \\
&= m^3 \int_{p,q_1,q_2} \frac{1}{p^2+m^2}\frac{1}{(p+q_1)^2+m^2}\frac{1}{(p+q_1+q_2)^2+m^2}\mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q_1-q_2}(i\not q_2\boldsymbol{\beta}_{q_2})(i\not q_1\boldsymbol{\beta}_{q_1})(i\not p)\right] \\
&\approx m^3 \int_{p,q_1,q_2} \frac{(-2p\cdot(2q_1+q_2))}{(p^2+m^2)^4}\left(2i\epsilon^{\mu\nu\rho}\right) q_2^\mu q_1^\nu p^\rho\,\mathrm{tr}\left[\delta Q^{\mathbb{C}}_{-q_1-q_2} Q^{\mathbb{C}}_{q_2} Q^{\mathbb{C}}_{q_1}\right] \\
&= \frac{-4im^3}{3}\int_{p,q_1,q_2} \frac{p^2}{(p^2+m^2)^4}\epsilon^{\mu\nu\rho} q_2^\mu q_1^\nu (2q_1+q_2)^\rho\,\mathrm{tr}\left[\delta Q^{\mathbb{C}}_{-q_1-q_2} Q^{\mathbb{C}}_{q_2} Q^{\mathbb{C}}_{q_1}\right] \\
&= 0
\end{aligned}$$

In passing to the third line we have traced over the γ matrices, and in passing to the last line we have used the fact that $q_2^\mu q_1^\nu (2q_1+q_2)^\rho$ is symmetric with respect to (ν,ρ) or (μ,ρ), while $\epsilon^{\mu\nu\rho}$ is totally anti-symmetric.

The $l=0$ term in Eq.(B.22) gives

$$\begin{aligned}
&-\mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}G_0\left(-im\boldsymbol{\beta}^\dagger\right)\right] \\
&= -m^2\int_p \frac{1}{p^2+m^2}\int d^3x\,\mathrm{tr}'\left[\delta\boldsymbol{\beta}\boldsymbol{\beta}^\dagger\right] \\
&= -2m^2\int_p \frac{1}{p^2+m^2}\int d^3x\,\mathrm{tr}\left[\delta Q^{\mathbb{C}}\, Q^{\mathbb{C}}\right] \\
&= 0
\end{aligned}$$

In passing to the last line we noted that

$$\left(Q^{\mathbb{C}}\right)^2 = I_n \;\Rightarrow \delta Q^{\mathbb{C}}\, Q^{\mathbb{C}} = -Q^{\mathbb{C}}\delta Q^{\mathbb{C}} \;\Rightarrow \mathrm{tr}\left[\delta Q^{\mathbb{C}}\, Q^{\mathbb{C}}\right] = -\mathrm{tr}\left[Q^{\mathbb{C}}\delta Q^{\mathbb{C}}\right]$$

Upon using the cyclic property of trace we conclude

$$\mathrm{tr}\left[\delta Q^{\mathbb{C}}\, Q^{\mathbb{C}}\right] = 0.$$

The $l=2$ term in Eq.(B.22) gives

$$
\begin{aligned}
&-\operatorname{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\partial\!\!\!/\boldsymbol{\beta}))\right)^2 G_0\left(-im\boldsymbol{\beta}^\dagger\right)\right]\\
&=-m^4\int_{p,q_1,q_2}\frac{1}{p^2+m^2}\frac{1}{(p+q_1)^2+m^2}\frac{1}{(p+q_1+q_2)^2+m^2}\mathrm{tr}'\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}i\not{q}_2\boldsymbol{\beta}_{q_2}i\not{q}_1\boldsymbol{\beta}_{q_1}\right]\\
&\approx -m^4\int_p\frac{1}{(p^2+m^2)^3}\int_{q_1,q_2}\mathrm{tr}'\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}i\not{q}_2\boldsymbol{\beta}_{q_2}i\not{q}_1\boldsymbol{\beta}_{q_1}\right]\\
&\approx -2m^4\int_p\frac{1}{(p^2+m^2)^3}\int d^3x\,\mathrm{tr}\left[Q^{\mathbb{C}}\,\delta Q^{\mathbb{C}}\,\partial_\mu Q^{\mathbb{C}}\,\partial_\mu Q^{\mathbb{C}}\right]\\
&=0
\end{aligned}
$$

In passing from the second to the third line we have used the fact at most three q_i are allowed (otherwise the term becomes irrelevant). Therefore at most we can expand the $\frac{1}{(p+q_1)^2+m^2}\frac{1}{(p+q_1+q_2)^2+m^2}$ to first order in $q_{1,2}$. However, such expansion inevitably comes with a p, and will vanish upon p integration. Thus we can only keep the 0th order term $\frac{1}{p^2+m^2}\frac{1}{(p^2+m^2)^2}$. In passing to the last line we have used $\delta Q^{\mathbb{C}}\,Q^{\mathbb{C}}=-Q^{\mathbb{C}}\delta Q^{\mathbb{C}}$ three times to move $Q^{\mathbb{C}}$ to the end, and use the cyclic property to move it back to the front. In this way, we have proven that the quantity is the minus of itself, hence it is zero.

To summarize, including all (the most and less relevant) terms, the non-linear sigma model is given by

$$
W[Q^{\mathbb{C}}]=\frac{1}{2\lambda_3}\int d^3x\,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}}\right]-\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}}\mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right].
$$

$(3+1)$-D

Given

$$
\boldsymbol{\beta}=I\otimes H_1+i\gamma^5 I\otimes H_2,
$$

the $l=0,2$ terms of Eq.(B.21) and the $l=1,3$ terms of Eq.(B.22) vanishes upon tr_γ. This is because they contain either one or three γ from $\not{p}$ or $\not{q}_i$. Because $\boldsymbol{\beta}$ contributes either I or γ^5. These terms vanish due to the fact that

$$
\mathrm{tr}_\gamma[\gamma^\mu]=\mathrm{tr}_\gamma[\gamma^\mu\gamma^5]=\mathrm{tr}_\gamma[\gamma^\mu\gamma^\nu\gamma^\rho]=\mathrm{tr}_\gamma[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^5]=0.
$$

The $l = 0$ term of Eq.(B.22) gives

$$\begin{aligned}
&- \mathrm{Tr}\left[im\delta\boldsymbol{\beta} G_0(-i\, m\, \boldsymbol{\beta})\right] \\
&= -m^2 \int_{p,q} \frac{1}{p^2+m^2} \mathrm{tr}' \left[\boldsymbol{\beta}^\dagger_{-q} \delta\boldsymbol{\beta}_q\right] \\
&= -m^2 \left(\int_p \frac{1}{p^2+m^2}\right) \int d^4x\, \mathrm{tr}' \left[\boldsymbol{\beta}^\dagger \delta\boldsymbol{\beta}\right] \\
&= -4m^2 \left(\int_p \frac{1}{p^2+m^2}\right) \int d^4x\, \delta \left\{\mathrm{tr}\left[H_1^2 + H_2^2\right]\right\} \\
&= 0.
\end{aligned}$$

In passing to the last line we use the same reasoning as the corresponding term in (1+1)-D.

While maintaining two space-time derivatives, the $l = 1$ term of Eq.(B.21) gives rise to the variation of the stiffness term $\delta W_{\text{stiffness}}$ in subsection B.4.2. Here we retain up to 4 space-time derivatives,

$$\begin{aligned}
&- \mathrm{Tr}\left[im\delta\boldsymbol{\beta}\left(mG_0((\partial\!\!\!/\boldsymbol{\beta}))\right) G_0(i\partial\!\!\!/)\right] \\
&= -m^2 \int_{p,q} \frac{1}{p^2+m^2} \frac{1}{(p+q)^2+m^2} \mathrm{tr}' \left[\delta\boldsymbol{\beta}_{-q}(\gamma^\mu q_\mu)\boldsymbol{\beta}_q(\gamma^\nu p_\nu)\right] \\
&= -m^2 \int_{p,q} \frac{1}{p^2+m^2} \frac{1}{(p+q)^2+m^2} q_\mu p^\nu \mathrm{tr}' \left[\delta\boldsymbol{\beta}_{-q}\gamma^\mu\gamma^\nu\boldsymbol{\beta}^\dagger_q\right] \\
&= \delta W_{\text{stiffness}} - m^2 \int_{p,q} \frac{1}{(p^2+m^2)^2} \left(\frac{4q^2(p\cdot q)}{(p^2+m^2)^2} - \frac{8(p\cdot q)^3}{(p^2+m^2)^3}\right)(q\cdot p)\mathrm{tr}' \left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}^\dagger_q\right] \\
&= \delta W_{\text{stiffness}} - m^2 \int_{p,q} \left(\frac{p^2q^4}{(p^2+m^2)^4} - \frac{p^4q^4}{(p^2+m^2)^5}\right) \mathrm{tr}' \left[\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}^\dagger_q\right] \\
&= \delta W_{\text{stiffness}} - \frac{1}{192\pi^2} \int d^4x\, \mathrm{tr}' \left[\partial^2(\delta\boldsymbol{\beta})\partial^2\boldsymbol{\beta}^\dagger\right] \\
&= \delta W_{\text{stiffness}} - \frac{1}{96\pi^2} \int d^4x\, \mathrm{tr} \left[\partial^2(\delta Q^{\mathbb{C}})\partial^2 Q^{\mathbb{C}\dagger} + \partial^2(\delta Q^{\mathbb{C}\dagger})\partial^2 Q^{\mathbb{C}}\right] \\
&= \delta W_{\text{stiffness}} - \frac{1}{96\pi^2} \int d^4x\, \delta\Big(\mathrm{tr}\left[\partial^2 Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger}\right]\Big)
\end{aligned} \tag{B.45}$$

In passing from the 2nd to the 3rd line, we use the property $\boldsymbol{\beta}\gamma^\nu = \gamma^\nu\boldsymbol{\beta}^\dagger$ for $\mu = 0, 1, 2, 3$. From the 3rd to the 4th line we have used the fact that the trace is only non-zero if the γ^μ and γ^ν are the same. From the 4th to the 5th line, we used the fact that rotational invariance allows the following replacement in the integrand of the p integral

$$p^\mu p^\nu p^\rho p^\sigma \to \frac{1}{D(D+2)} \left(\delta_{\mu\nu}\delta_{\rho\sigma} + \delta_{\mu\rho}\delta_{\nu\sigma} + \delta_{\mu\sigma}\delta_{\nu\rho}\right).$$

(The factor $\frac{1}{D(D+2)}$ can be fixed by taking trace on both sides.) In the 5th line, only the terms in $\delta\boldsymbol{\beta}_{-q}\boldsymbol{\beta}^\dagger_q$ having an even number of γ^5 are non-zero. Moreover, since γ^5 is always accompanied by H_2, we can replace γ^5 with the identity matrix as long as we symmetrize the end result with respect to H_2. After the replacement, $\boldsymbol{\beta}$ becomes $II \otimes Q^{\mathbb{C}}$, the identity matrix can then be trace out, and the symmetrization amounts to sum over the terms with $Q^{\mathbb{C}} = H_1 + iH_2$ replaced by $Q^{\mathbb{C}\dagger} = H_1 - iH_2$. We will use this last trick several times in the following.

The $l = 3$ term of Eq.(B.21) gives

$$
\begin{aligned}
&- \mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\not{\partial}\boldsymbol{\beta}))\right)^3 G_0\left(i\not{\partial}\right)\right] \\
&= m^4 \int_{p,q_1,q_2,q_3} \frac{1}{p^2+m^2}\frac{1}{(p+q_1)^2+m^2}\frac{1}{(p+q_1+q_2)^2+m^2}\frac{1}{(p+q_1+q_2+q_3)^2+m^2} \\
&\times \mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q_1-q_2-q_3}(i\not{q}_3\boldsymbol{\beta}_{q_3})(i\not{q}_2\boldsymbol{\beta}_{q_2})(i\not{q}_1\boldsymbol{\beta}_{q_1})(i\not{p})\right] \\
&\approx m^4 \int_{p,q_1,q_2,q_3} \frac{-2p\cdot(3q_1+2q_2+q_3)}{(p^2+m^2)^5} q_3^\mu q_2^\nu q_1^\rho p^\sigma \mathrm{tr}'\left[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \delta\boldsymbol{\beta}_{-q_1-q_2-q_3}\boldsymbol{\beta}^\dagger_{q_3}\boldsymbol{\beta}_{q_2}\boldsymbol{\beta}^\dagger_{q_1}\right] \\
&= -\frac{m^4}{2}\int_p \frac{p^2}{(p^2+m^2)^5}\int_{q_1,q_2,q_3} q_3^\mu q_2^\nu q_1^\rho (3q_1+2q_2+q_3)^\sigma \left(\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\rho}\delta_{\nu\sigma} + \delta_{\mu\sigma}\delta_{\nu\rho}\right) \\
&\times \mathrm{tr}'\left[\delta\boldsymbol{\beta}_{-q_1-q_2-q_3}\boldsymbol{\beta}^\dagger_{q_3}\boldsymbol{\beta}_{q_2}\boldsymbol{\beta}^\dagger_{q_1}\right]_{\text{even terms in} H_2} \\
&= -\frac{1}{96\pi^2}\int_{q_1,q_2,q_3}\left(3q_1^2(q_2\cdot q_3) - 2q_2^2(q_1\cdot q_3) + q_3^2(q_1\cdot q_2) + 4(q_1\cdot q_2)(q_2\cdot q_3)\right) \\
&\quad \times \mathrm{tr}'\left[\delta Q^{\mathbb{C}}_{-q_1-q_2-q_3} Q^{\mathbb{C}\dagger} Q^{\mathbb{C}}_{q_2} Q^{\mathbb{C}\dagger}_{q_1}\right]_{\text{even } H_2} \\
&= -\frac{1}{96\pi^2}\int d^4x\,\frac{1}{2}\mathrm{tr}\begin{bmatrix}
3\,\delta Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger} + 3\,\delta Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}} \\
-2\,\delta Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger} - 2\,\delta Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \\
+\,\delta Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger} + \,\delta Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \\
+4\,\delta Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu\partial_\nu Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}\dagger} + 4\,\delta Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\mu\partial_\nu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}
\end{bmatrix} \\
&= \frac{1}{96\pi^2}\int d^4x\,\mathrm{tr}\left[Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\begin{pmatrix}
-3\,\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}} \\
+2\,\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \\
-\partial^2 Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \\
-4\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu\partial_\nu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\
+6\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\
-2\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} \\
+2\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}
\end{pmatrix}\right]
\end{aligned}
\tag{B.46}
$$

From the 3rd to the 4th line, we take the terms with even number of γ^5 (thus even number of H_2) from $\boldsymbol{\beta}$ and use the identity $\mathrm{tr}_\gamma\left[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\right] = 4(\delta^{\mu\nu}\delta^{\rho\sigma} - \delta^{\mu\rho}\delta^{\nu\sigma} + \delta^{\mu\sigma}\delta^{\nu\rho})$. The terms with odd number of γ^5 vanish because $\mathrm{tr}_\gamma\left[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^5\right] = 4\epsilon^{\mu\nu\rho\sigma}$ is totally anti-symmetric, while $q_3^\nu q_2^\nu g_1^\rho(3q_1+2q_2+q_3)^\sigma$ is symmetric with respect to either (μ,σ), (ν,σ), or (ρ,σ). From the 4th line to the 5th line, we used the same trick as in Eq.(B.45). From the 6th to the 7th line, $\delta Q^{\mathbb{C}\dagger} = -Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}}Q^{\mathbb{C}\dagger}$ is used repeatedly until δ or ∂ act only on $Q^{\mathbb{C}}$.

The $l = 2$ term of Eq.(B.22) gives

$$
\begin{aligned}
&- \mathrm{Tr}\left[im\,\delta\boldsymbol{\beta}\left(mG_0((\partial\!\!\!/\boldsymbol{\beta}))\right)^2 G_0\left(-im\boldsymbol{\beta}^\dagger\right)\right] \\
&= -m^4 \int_{p,q_1,q_2,q_3} \frac{1}{(p+q_1+q_2)^2+m^2}\frac{1}{(p+q_1)^2+m^2}\frac{1}{p^2+m^2} \\
&\times \mathrm{tr}'\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}(i\not{q}_2\boldsymbol{\beta}_{q_2})(i\not{q}_1\boldsymbol{\beta}_{q_1})\right] \\
&\approx m^4 \int_{p,q_1,q_2,q_3} \frac{q_2^\nu q_1^\mu}{(p^2+m^2)^3}\Big[1 - \frac{q_1^2+(q_1+q_2)^2}{p^2+m^2} + 4\frac{(p\cdot q_1)^2+(p\cdot(q_1+q_2))^2+(p\cdot q_1)(p\cdot(q_1+q_2))}{(p^2+m^2)^2}\Big] \\
&\times \mathrm{tr}'\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}(\gamma_\nu\gamma_\mu\boldsymbol{\beta}^\dagger_{q_2}\boldsymbol{\beta}_{q_1})\right] \\
&= m^4 \int_{p,q_1,q_2,q_3} \frac{(q_1\cdot q_2)}{(p^2+m^2)^3}\Big[1 - \frac{q_1^2+(q_1+q_2)^2}{p^2+m^2} + \frac{p^2\left(q_1^2+(q_1+q_2)^2+q_1\cdot(q_1+q_2)\right)}{(p^2+m^2)^2}\Big] \\
&\times \mathrm{tr}\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}\boldsymbol{\beta}^\dagger_{q_2}\boldsymbol{\beta}_{q_1}\right] \\
&= m^4 \int_{q_1,q_2,q_3} (q_1\cdot q_2)\Big[\frac{1}{32\pi^2 m^2} - \frac{1}{192\pi^2 m^4}\left(q_1^2+q_2^2+q_1\cdot q_2\right)\Big]\mathrm{tr}'\left[\boldsymbol{\beta}^\dagger_{-q_1-q_2-q_3}\delta\boldsymbol{\beta}_{q_3}\boldsymbol{\beta}^\dagger_{q_2}\boldsymbol{\beta}_{q_1}\right] \\
&= \int d^4x\Big[-\frac{m^2}{32\pi^2}\mathrm{tr}'\left[\boldsymbol{\beta}^\dagger\delta\boldsymbol{\beta}\partial_\mu\boldsymbol{\beta}^\dagger\partial_\mu\boldsymbol{\beta}\right] - \frac{1}{192\pi^2}\mathrm{tr}'\left[\boldsymbol{\beta}^\dagger\delta\boldsymbol{\beta}\left(\partial_\mu\boldsymbol{\beta}^\dagger\partial_\mu\partial^2\boldsymbol{\beta}+\partial_\mu\partial^2\boldsymbol{\beta}^\dagger\partial_\mu\boldsymbol{\beta}+\partial_\mu\partial_\nu\boldsymbol{\beta}^\dagger\partial_\mu\partial_\nu\boldsymbol{\beta}\right)\right]\Big] \\
&= \int d^4x\Big[-\frac{m^2}{16\pi^2}\mathrm{tr}\Big[Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} + Q^{\mathbb{C}}\delta Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\Big] \\
&- \frac{1}{96\pi^2}\mathrm{tr}\left[\begin{array}{c} Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}}\left(\partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu\partial^2 Q^{\mathbb{C}}+\partial_\mu\partial^2 Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}+\partial_\mu\partial_\nu Q^{\mathbb{C}\dagger}\partial_\mu\partial_\nu Q^{\mathbb{C}}\right) \\ +Q^{\mathbb{C}}\delta Q^{\mathbb{C}\dagger}\left(\partial_\mu Q^{\mathbb{C}}\partial_\mu\partial^2 Q^{\mathbb{C}\dagger}+\partial_\mu\partial^2 Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}+\partial_\mu\partial_\nu Q^{\mathbb{C}}\partial_\mu\partial_\nu Q^{\mathbb{C}\dagger}\right)\end{array}\right]\Big]
\end{aligned}
$$

In passing from the 6th to the last line we have used the symmetrization trick in arriving at Eq.(B.45). Using $\delta Q^{\mathbb{C}\dagger} = -Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}$, the first term in the last line gives zero. The second term can be evaluated using the same formula repeatedly. After some straightforward expansion, most terms cancel out and we are left with

$$
-\frac{1}{96\pi^2}\int d^4x\,\mathrm{tr}\Big[Q^{\mathbb{C}\dagger}\delta Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}\left(\partial^2 Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} - \partial_\mu Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}} Q^{\mathbb{C}\dagger}\partial^2 Q^{\mathbb{C}}\right)\Big] \tag{B.47}
$$

Summing over Eq.(B.45), Eq.(B.46), and Eq.(B.47), we obtain

$$
\delta W_{\text{stiffness}} + \delta\left(\frac{1}{92\pi^2}\int d^4x\,\mathrm{tr}\left[\begin{array}{c}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\frac{1}{2}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\partial^2 Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger}\end{array}\right]\right). \tag{B.48}
$$

All these terms are real. At low energy and long wavelength they are dominated by the stiffness term. In appendix B.5 we shall refer to the stiffness term plus these extra terms as the "generalized stiffness" term.

To summarize, including all "generalized stiffness" terms, the non-linear sigma model is given by

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_4^2}\int_{\mathcal{M}} d^4x \,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right] - \frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^5\Big]$$
$$+\frac{1}{92\pi^2}\int_{\mathcal{M}} d^4x\,\mathrm{tr}\begin{bmatrix} \partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\frac{1}{2}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\partial^2 Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger}\end{bmatrix} \tag{B.49}$$

B.5 Emergent symmetries of the nonlinear sigma models

In this appendix, we shall generalize the discussions in section 4.11 to $(2+1)$-D and $(3+1)$-D, namely, showing the nonlinear sigma models respect the full emergent symmetries of the massless free fermion theories (see table 4.4, or appendix B.1).

As we explained in appendix B.4.3, the nonlinear sigma models in real classes can be derived from the complex classes by restricting $Q^{\mathbb{R}}$ to the appropriate sub-mass manifold of $Q^{\mathbb{C}}$. Similarly, for each space-time dimension the emergent symmetry group of the real class is a subgroup of the complex class (see table 4.4). Hence, once we have matched the symmetries (between the nonlinear sigma models and fermion theories) for the complex class, it is straightforward to do the same for the real class. All we need to do is to restrict the order parameters to the appropriate sub-mass manifold and the symmetries to the appropriate subgroup. Therefore we shall focus on the complex classes in the following.

Complex class in $(2+1)$-D

The nonlinear sigma model is given by Eq.(B.39), namely,

$$W[Q^{\mathbb{C}}] = \frac{1}{2\lambda_3}\int d^3x\,\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}}\right] - \frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right].$$

(i) *Global $U(n)$*

Using the cyclic invariance of trace, the action in Eq.(B.39) clearly respects the $U(n)$ symmetry

$$Q^{\mathbb{C}} \to g^\dagger \cdot Q^{\mathbb{C}} \cdot g.$$

(ii) *Charge conjugation*

$Q^{\mathbb{C}}$ transforms under the charge conjugation as

$$Q^{\mathbb{C}} \xrightarrow{C} \left(Q^{\mathbb{C}}\right)^T .$$

Under such transformation the stiffness term becomes

$$\begin{aligned}&\frac{1}{2\lambda_3}\int d^3x \,\mathrm{tr}\left[\partial_\mu \left(Q^{\mathbb{C}}\right)^T \partial^\mu \left(Q^{\mathbb{C}}\right)^T\right]\\ &= \frac{1}{2\lambda_3}\int d^3x \,\mathrm{tr}\left[\partial^\mu Q^{\mathbb{C}} \partial_\mu Q^{\mathbb{C}}\right]\end{aligned}$$

Hence is invariant. In passing to the last line we have used the fact that the trace of a transposed matrix is the same as that of the original.

Under charge conjugation the WZW term transforms as

$$\begin{aligned}&-\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right] \xrightarrow{C} -\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[(\tilde{Q}^{\mathbb{C}})^T\left(d(\tilde{Q}^{\mathbb{C}})^T\right)^4\right]\\ &= -\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\left(d\tilde{Q}^{\mathbb{C}}\right)^4 \tilde{Q}^{\mathbb{C}}\right]\\ &= -\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right].\end{aligned}$$

In passing to the second line we have used the transposing invariance of the trace, and the fact the reordering caused by transposing results in an even number of exchanges between the differential 1-forms, hence there is no sign change. The cyclic property of trace is used for the last equality. Therefore the WZW term is charge conjugation invariant.

(iii) *Time reversal*

Under time-reversal $Q^{\mathbb{C}}$ transforms as

$$Q^{\mathbb{C}} \xrightarrow{T} -(Q^{\mathbb{C}})^* = -(Q^{\mathbb{C}})^T.$$

(Here we have used the fact that $Q^{\mathbb{C}}$ is hermitian). This results in the following transformation of the stiffness term

$$\begin{aligned}&\frac{1}{2\lambda_3}\int d^3x \;\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}} \partial^\mu Q^{\mathbb{C}}\right]\\ &\xrightarrow{T} \left(\frac{1}{2\lambda_3}\int d^3x \;\mathrm{tr}\left[\partial_\mu(-Q^{\mathbb{C}*})\partial^\mu(-Q^{\mathbb{C}*})\right]\right)^*\\ &=\frac{1}{2\lambda_3}\int d^3x \;\mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}} \partial^\mu Q^{\mathbb{C}}\right]\end{aligned}$$

In passing to the second line we have used the fact that in Euclidean space-time the Boltzmann weight needs to be complex conjugated under anti-unitary transformation. Therefore, the stiffness term is time reversal invariant.

The WZW term transforms as follows under time reversal

$$\begin{aligned}
&-\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right] \xrightarrow{T} \left(-\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[(-\tilde{Q}^{\mathbb{C}})^*\left(d(-\tilde{Q}^{\mathbb{C}})^*\right)^4\right]\right)^* \\
&= -\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{C}}\left(d\tilde{Q}^{\mathbb{C}}\right)^4\right],
\end{aligned}$$

where the five negative signs associated with transposing are canceled out by the negative sign arising from complex conjugation of i. Thus the WZW term is time reversal invariant.

In summary, the nonlinear sigma model respects the full emergent symmetries of the massless fermion theory (see table 4.4).

Complex class in $(3+1)$-D

The nonlinear sigma model in Eq.(B.49) is given by

$$\begin{aligned}
W[Q^{\mathbb{C}}] =& \frac{1}{2\lambda_4^2}\int_{\mathcal{M}} d^4x\, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}}\partial^\mu Q^{\mathbb{C}\dagger}\right] - \frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^5\Big] \\
&+\frac{1}{92\pi^2}\int_{\mathcal{M}} d^4x\, \mathrm{tr}\begin{bmatrix} \partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\frac{1}{2}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\partial^2 Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger} \end{bmatrix}
\end{aligned}$$

(i) *Global $U(n)\times U(n)$*

Eq.(B.49) is clearly invariant under the $U_+(n)\times U_-(n)$ transformations

$$Q^{\mathbb{C}} \to g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+.$$

This is because in Eq.(B.49) $Q^{\mathbb{C}}$ and $Q^{\mathbb{C}\dagger}$ appears sequentially.

(ii) *Charge conjugation*

Under charge-conjugation $Q^{\mathbb{C}}$ transforms as

$$Q^{\mathbb{C}} \xrightarrow{C} (Q^{\mathbb{C}})^T.$$

Under such transformation the "generalized stiffness" terms transforms as

$$\begin{aligned}
&\frac{1}{2\lambda_4^2}\int d^4x \, \mathrm{tr}\left[\partial_\mu \left(Q^{\mathbb{C}}\right)^T \partial^\mu \left(Q^{\mathbb{C}\dagger}\right)^T\right] \\
&+\frac{1}{92\pi^2}\int d^4x \, \mathrm{tr}\begin{bmatrix} \partial_\mu \left(Q^{\mathbb{C}\dagger}\right)^T \partial_\mu \left(Q^{\mathbb{C}}\right)^T \partial_\nu \left(Q^{\mathbb{C}\dagger}\right)^T \partial_\nu \left(Q^{\mathbb{C}}\right)^T \\ -\frac{1}{2}\partial_\mu \left(Q^{\mathbb{C}\dagger}\right)^T \partial_\nu \left(Q^{\mathbb{C}}\right)^T \partial_\mu \left(Q^{\mathbb{C}\dagger}\right)^T \partial_\nu \left(Q^{\mathbb{C}}\right)^T \\ -\partial^2 \left(Q^{\mathbb{C}}\right)^T \partial^2 \left(Q^{\mathbb{C}\dagger}\right)^T \end{bmatrix} \\
&= \frac{1}{2\lambda_4^2}\int d^4x \, \mathrm{tr}\left[\partial_\mu Q^{\mathbb{C}} \partial^\mu Q^{\mathbb{C}\dagger}\right] + \frac{1}{92\pi^2}\int d^4x \, \mathrm{tr}\begin{bmatrix} \partial_\mu Q^{\mathbb{C}\dagger}\partial_\mu Q^{\mathbb{C}}\partial_\nu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\frac{1}{2}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}}\partial_\mu Q^{\mathbb{C}\dagger}\partial_\nu Q^{\mathbb{C}} \\ -\partial^2 Q^{\mathbb{C}}\partial^2 Q^{\mathbb{C}\dagger} \end{bmatrix}
\end{aligned} \tag{B.50}$$

In arriving at the final line we have used the transposing invariance of the trace. Therefore the "generalized stiffness" terms are charge conjugation invariant.

Under charge conjugation, the WZW term transforms as

$$\begin{aligned}
&-\frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^5\Big] \\
\xrightarrow{C} &-\frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left((\tilde{Q}^{\mathbb{C}})^* d(\tilde{Q}^{\mathbb{C}})^T\right)^5\Big] \\
= &-\frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(d\tilde{Q}^{\mathbb{C}}\,\tilde{Q}^{\mathbb{C}\dagger}\right)^5\Big] \\
= &-\frac{2\pi}{480\pi^3}\int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}}\right)^5\Big]
\end{aligned}$$

In passing to the third line we have used the transposing invariance of the trace. Note that there is no extra sign because the number of exchanges between 1-forms is even (10 times). In arriving at the last line, the last $\tilde{Q}^{\mathbb{C}\dagger}$ is moved to the front by the cyclic invariance of the trace. Thus the WZW term is charge conjugation invariant.

(iii) *Time reversal*

Under time-reversal $Q^{\mathbb{C}}$ transforms as

$$Q^{\mathbb{C}} \xrightarrow{T} (Q^{\mathbb{C}})^*.$$

$$W\left[Q^{\mathbb{C}}\right] \xrightarrow{T} \left(W\left[(Q^{\mathbb{C}})^*\right]\right)^* = W\left[Q^{\mathbb{C}}\right].$$

This is because all the coefficients (including those in front of the generalized stiffness terms and the WZW term) in the nonlinear sigma model are real, the complex conjugation of the Boltzmann weight cancels out with complex conjugation in $Q^{\mathbb{C}*}$.

To summarize, the nonlinear sigma model is invariant under the full emergent symmetries of the massless fermion theory (see table 4.4).

B.6 Anomalies of the nonlinear sigma models

To reveal the ‘t Hooft anomalies of the non-linear sigma model we first need to gauge it. In this section, we shall extend the discussions in section 4.12.1 to gauge the continuous symmetries of nonlinear sigma models in $(1+1)$-D, $(2+1)$-D, and $(3+1)$-D. We shall adopt Witten's trial-and-error method [26].

We have discussed at the beginning of appendix B.5 that the mass manifold and emergent symmetries of the non-linear sigma model of real classes are the submanifold and sub-group of the corresponding sigma model of complex classes. Consequently, once one knows how to gauge the nonlinear sigma models in the complex classes, one simply needs to restrict the order parameters ($Q^{\mathbb{R}}$) to the submanifold, and the gauge group to the subgroup, to derive the gauged non-linear sigma models of real classes.

B.6.1 The (‘t Hooft) anomalies associated with continuous symmetries

Complex class in $(1+1)$-D

The discussion for gauging the nonlinear sigma model of complex class in $(1+1)$-D was already in section 4.12.1. We will not repeat the argument but just quote the result here:

$$\begin{aligned} W[Q^{\mathbb{C}}, A_+, A_-] = &-\frac{1}{8\pi}\int_{\mathcal{M}} d^2x\,\mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}\left(\partial_\mu Q^{\mathbb{C}} - iQ^{\mathbb{C}}A_{+,\mu} + iA_{-,\mu}Q^{\mathbb{C}}\right)\right)^2\right] \\ &- \frac{i}{12\pi}\int_{\mathcal{B}} \mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}}\right)^3\right] - \frac{1}{4\pi}\int_{\mathcal{M}} \mathrm{tr}\Big\{A_+\left(Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}}\right) \\ &+ A_-\left(dQ^{\mathbb{C}}Q^{\mathbb{C}\dagger}\right) + iA_+Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}}\Big\}. \end{aligned} \tag{B.51}$$

Under infinitesimal $U_+(n)\times U_-(n)$ gauge transformation,

$$\begin{aligned} &Q^{\mathbb{C}} \to e^{-i\epsilon_-}Q^{\mathbb{C}}e^{i\epsilon_+} \\ &A_\pm \to A_\pm + d\epsilon_\pm + i[A_\pm, \epsilon_\pm], \end{aligned}$$

Eq.(B.51) acquires an addition piece

$$\delta W = -\frac{i}{4\pi}\int_{\mathcal{M}} \mathrm{tr}\left[A_+d\epsilon_+ - A_-d\epsilon_-\right]. \tag{B.52}$$

Thus Eq.(B.51) is not gauge invariant, revealing the ‘t Hooft anomaly associated with $U_+(n) \times U_-(n)$. However, when one only gauges the diagonal $U(n)$, i.e., $A_+ = A_- := A$ and $\epsilon_+ = \epsilon_- = \epsilon$, the non gauge invariant terms in Eq.(B.52) cancels out. Hence Eq.(B.51) is anomaly free with respect to the diagonal $U(n)$. This agrees with the free fermion anomaly.

Real class in $(1+1)$-D

The gauged nonlinear sigma model for real class in $(1+1)$-D can be derived from the complex class by 1) restricting the order parameter $Q^{\mathbb{C}} \in U(n)$ to the subspace $Q^{\mathbb{R}} \in O(n)$, 2) restricting the gauge group from $U_+(n) \times U_-(n)$ to $O_+(n) \times O_-(n)$, and 3) divide the nonlinear sigma model by a factor of two (see B.4.1). The result is

$$\begin{aligned} W[Q^{\mathbb{R}}, A_+, A_-] = & -\frac{1}{16\pi} \int_{\mathcal{M}} d^2x \, \mathrm{tr} \left[\left((Q^{\mathbb{R}})^T \left(\partial_\mu Q^{\mathbb{R}} - iQ^{\mathbb{R}} A_{+,\mu} + iA_{-,\mu} Q^{\mathbb{R}} \right) \right)^2 \right] \\ & + \frac{2\pi i}{48\pi^2} \int_{\mathcal{B}} \mathrm{tr} \Big[\Big((\tilde{Q}^{\mathbb{R}})^T d\tilde{Q}^{\mathbb{R}} \Big)^3 \Big] + \frac{1}{8\pi} \int_{\mathcal{M}} \mathrm{tr} \Big\{ A_+ \left(dQ^{\mathbb{R}} (Q^{\mathbb{R}})^T \right) \\ & + A_- \left((Q^{\mathbb{R}})^T dQ^{\mathbb{R}} \right) - iA_+ (Q^{\mathbb{R}})^T A_- Q^{\mathbb{R}} \Big\}. \end{aligned} \tag{B.53}$$

Here $A_\pm$ are the gauge fields associated with $O_+(n) \times O_-(n)$. Under the $O_+(n) \times O_-(n)$ gauge transformation,

$$\begin{aligned} & Q^{\mathbb{R}} \to e^{-i\epsilon_-} Q^{\mathbb{R}} e^{i\epsilon_+} \\ & A_\pm \to A_\pm + d\epsilon_\pm + i[A_\pm, \epsilon_\pm], \end{aligned}$$

(here ϵ_+ and ϵ_- are imaginary anti-symmetric matrices) Eq.(B.53) acquires an addition piece

$$\delta W = -\frac{i}{8\pi} \int_{\mathcal{M}} \mathrm{tr} \left[A_+ d\epsilon_+ - A_- d\epsilon_- \right],$$

manifesting the ‘t Hooft anomaly associated with $O_+(n) \times O_-(n)$. Again, when only the diagonal $O(n)$ is gauged, Eq.(B.53) is anomaly-free, consistent with the free fermion prediction.

Complex class in $(2+1)$-D

In the following, we carry out Witten’s method [26] for the non-linear sigma model. The emergent continuous symmetry is $U(n)$ and under gauge transformation $Q^{\mathbb{C}}$ and A change according to

$$\begin{aligned} & Q^{\mathbb{C}} \to Q^{\mathbb{C}} + \delta Q^{\mathbb{C}} \text{ where } \delta Q^{\mathbb{C}} = i[Q^{\mathbb{C}}, \epsilon] \\ & A \to A + \delta A \text{ where } \delta A = d\epsilon + i[A, \epsilon]. \end{aligned} \tag{B.54}$$

The gauge field enters stiffness term in Eq.(B.39) via the minimal coupling,

$$W_{\rm stiff}[Q^{\mathbb{C}},A] = \frac{1}{2\lambda_3}\int\limits_{\mathcal{M}} d^3x\,{\rm tr}\Big[\Big(\partial_\mu Q^{\mathbb{C}} + i[A_\mu, Q^{\mathbb{C}}]\Big)^2\Big]$$

which is gauge invariant.

Following Witten's trial-and-error method, we now determine how gauge field enters through the WZW term. Under Eq.(B.54) the WZW term acquires an addition piece

$$\begin{aligned}
&\delta\left(\int\limits_{\mathcal{B}} {\rm tr}\left[Q^{\mathbb{C}}(dQ^{\mathbb{C}})^4\right]\right)\\
=&\int\limits_{\mathcal{B}} {\rm tr}\left[\delta Q^{\mathbb{C}}(dQ^{\mathbb{C}})^4 + 4Qd(\delta Q^{\mathbb{C}})(dQ^{\mathbb{C}})^3\right]\\
=&\int\limits_{\mathcal{B}} {\rm tr}\left[5\delta Q^{\mathbb{C}}(dQ^{\mathbb{C}})^4 + d\left(4Q^{\mathbb{C}}\delta Q^{\mathbb{C}}(dQ^{\mathbb{C}})^3\right)\right]\\
=&5\int\limits_{\mathcal{B}} {\rm tr}\left[i\left(Q^{\mathbb{C}}\epsilon-\epsilon Q^{\mathbb{C}}\right)(dQ^{\mathbb{C}})^4\right] + \int\limits_{\mathcal{M}} {\rm tr}\left[4Q^{\mathbb{C}} i\left(Q^{\mathbb{C}}\epsilon-\epsilon Q^{\mathbb{C}}\right)(dQ^{\mathbb{C}})^3\right]\\
=&0+8i\int\limits_{\mathcal{M}} {\rm tr}\left[\epsilon(dQ^{\mathbb{C}})^3\right]\\
=&-8i\int\limits_{\mathcal{M}} {\rm tr}\left[d\epsilon\, Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2\right]
\end{aligned} \tag{B.55}$$

In passing to the second line we used the constraint $Q^{\mathbb{C}}dQ^{\mathbb{C}} = -dQ^{\mathbb{C}}\,Q^{\mathbb{C}}$. An integration by part is done from the 2nd to the 3rd line. The 1st term in the 4th line vanishes because we can repeatedly use $Q^{\mathbb{C}}dQ^{\mathbb{C}} = -dQ^{\mathbb{C}}\,Q^{\mathbb{C}}$ and the cyclic invariance of the trace to show

$${\rm tr}\left[\epsilon Q^{\mathbb{C}}(dQ^{\mathbb{C}})^4\right] = {\rm tr}\left[Q^{\mathbb{C}}\epsilon(dQ^{\mathbb{C}})^4\right].$$

To cancel out the gauge dependent part of Eq.(B.55), we add an additional term

$$\text{Added term}\;\; 8i\int\limits_{\mathcal{M}} {\rm tr}\left[A\, Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2\right]. \tag{B.56}$$

Under the gauge transformation Eq.(B.54) this additional term transforms into

$$\begin{aligned}
&\delta\left(8i\int_{\mathcal{M}}\mathrm{tr}\left[AQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2\right]\right)\\
=&8i\int_{\mathcal{M}}\mathrm{tr}\left[\delta AQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2+\delta Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2A+d(\delta Q^{\mathbb{C}})\left(dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}+A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\right)\right]\\
=&8i\int_{\mathcal{M}}\mathrm{tr}\left[\delta A\,Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2+\delta Q^{\mathbb{C}}\begin{pmatrix}(dQ^{\mathbb{C}})^2A+dQ^{\mathbb{C}}\,dA\,Q^{\mathbb{C}}-dQ^{\mathbb{C}}\,A\,dQ^{\mathbb{C}}\\-dA\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}+A\,dQ^{\mathbb{C}}\,dQ^{\mathbb{C}}\end{pmatrix}\right]\\
=&8i\int_{\mathcal{M}}\mathrm{tr}\left[d\epsilon\,Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2+i[A,\epsilon]Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2+i[Q^{\mathbb{C}},\epsilon]\begin{pmatrix}(dQ^{\mathbb{C}})^2A+dQ^{\mathbb{C}}\,dA\,Q^{\mathbb{C}}-dQ^{\mathbb{C}}\,A\,dQ^{\mathbb{C}}\\-dA\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}+A\,dQ^{\mathbb{C}}\,dQ^{\mathbb{C}}\end{pmatrix}\right]\\
=&8i\int_{\mathcal{M}}\mathrm{tr}\left[d\epsilon\,Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2\right]-8\int_{\mathcal{M}}\mathrm{tr}\left[\epsilon\,d\begin{pmatrix}Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}+Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\\Q^{\mathbb{C}}\,dA+dA\,Q^{\mathbb{C}}\end{pmatrix}\right]\\
=&8i\int_{\mathcal{M}}\mathrm{tr}\left[d\epsilon\,Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2\right]+8\int_{\mathcal{M}}\mathrm{tr}\left[d\epsilon\begin{pmatrix}Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}+Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\\Q^{\mathbb{C}}\,dA+dA\,Q^{\mathbb{C}}\end{pmatrix}\right]
\end{aligned}\tag{B.57}$$

The first term of the final result cancels the gauge dependent term of Eq.(B.55) by design. We continue to add the additional term

$$\text{Added term}\quad-8\int_{\mathcal{M}}\mathrm{tr}\left[(AQ^{\mathbb{C}})^2dQ^{\mathbb{C}}\right]\tag{B.58}$$

in an attempt to cancel the term

$$8\int_{\mathcal{M}}\mathrm{tr}\left[d\epsilon\,\left(Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}+Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\right)\right]\tag{B.59}$$

in Eq.(B.57). Under the gauge transformation (Eq.(B.54)) the added term transforms as

$$\begin{aligned}
&\delta\left(-8\int_{\mathcal{M}}\mathrm{tr}\left[(AQ^{\mathbb{C}})^2dQ^{\mathbb{C}}\right]\right)\\
=&-8\int_{\mathcal{M}}\mathrm{tr}\begin{bmatrix}\delta A\left(Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}+Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\\+\delta Q^{\mathbb{C}}\begin{pmatrix}A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A+dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,A\\-dA\,Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}+A\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\\+A\,Q^{\mathbb{C}}\,dA\,Q^{\mathbb{C}}-A\,Q^{\mathbb{C}}\,A\,dQ^{\mathbb{C}}\end{pmatrix}\end{bmatrix}\\
=&-8\int_{\mathcal{M}}\mathrm{tr}\begin{bmatrix}d\epsilon\left(Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}+Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\\+i\,\epsilon\,d\left(-A\,Q^{\mathbb{C}}\,A+Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\end{bmatrix}\\
=&\int_{\mathcal{M}}\mathrm{tr}\begin{bmatrix}-8d\epsilon\left(Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,dQ^{\mathbb{C}}+Q^{\mathbb{C}}\,dQ^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\\+8i\,d\epsilon\,\left(-A\,Q^{\mathbb{C}}\,A+Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\end{bmatrix}
\end{aligned}\tag{B.60}$$

The top line in the final result achieves canceling out Eq.(B.59). Now we focus on canceling out the terms in the bottom line of Eq.(B.60). The term

$$\int_{\mathcal{M}}\mathrm{tr}\left[+8i\,d\epsilon\,\left(Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\,A\,Q^{\mathbb{C}}\right)\right]\tag{B.61}$$

can be canceled out by adding the extra term

$$\text{Added term} \quad -\frac{8i}{3}\int_{\mathcal{M}} \mathrm{tr}\left[(A\,Q^{\mathbb{C}})^3\right]. \tag{B.62}$$

Under gauge transformation the added term transforms as

$$\begin{aligned}
&\delta\left(-\frac{8i}{3}\int_{\mathcal{M}} \mathrm{tr}\left[(A\,Q^{\mathbb{C}})^3\right]\right)\\
=&-8i\int_{\mathcal{M}} \mathrm{tr}\left[\delta(A\,Q^{\mathbb{C}})(A\,Q^{\mathbb{C}})^2\right]\\
=&-8i\int_{\mathcal{M}} \mathrm{tr}\left[\left(d\epsilon\,Q^{\mathbb{C}}+i[A,\epsilon]Q^{\mathbb{C}}+iA[Q^{\mathbb{C}},\epsilon]\right)(A\,Q^{\mathbb{C}})^2\right]\\
=&-8i\int_{\mathcal{M}} \mathrm{tr}\left[d\epsilon\,Q^{\mathbb{C}}(A\,Q^{\mathbb{C}})^2\right]
\end{aligned} \tag{B.63}$$

which indeed cancels Eq.(B.61). The remaining term

$$\int_{\mathcal{M}} \mathrm{tr}\left[+8i\,d\epsilon\,\left(-A\,Q^{\mathbb{C}}\,A\right)\right] \tag{B.64}$$

in Eq.(B.60) can be partially canceled by adding the extra term

$$\text{Added term} \quad 8i\int_{\mathcal{M}} \mathrm{tr}\left[A^3 Q^{\mathbb{C}}\right], \tag{B.65}$$

which transforms as

$$\begin{aligned}
&\delta\left(8i\int_{\mathcal{M}} \mathrm{tr}\left[A^3 Q^{\mathbb{C}}\right]\right)\\
=&8i\int_{\mathcal{M}} \mathrm{tr}\left[(d\epsilon+i[A,\epsilon])\left(A^2\,Q^{\mathbb{C}}+A\,Q^{\mathbb{C}}\,A+Q^{\mathbb{C}}\,A^2\right)+i[Q^{\mathbb{C}},\epsilon]\left(A^3\right)\right]\\
=&8i\int_{\mathcal{M}} \mathrm{tr}\left[d\epsilon\left(A^2\,Q^{\mathbb{C}}+A\,Q^{\mathbb{C}}\,A+Q^{\mathbb{C}}\,A^2\right)\right]
\end{aligned} \tag{B.66}$$

under the gauge transformation. The second term in Eq.(B.66) cancel Eq.(B.64).

At this point, under the gauge transformation, the sum of the original WZW term and the added terms Eq.(B.56),Eq.(B.58),Eq.(B.62), Eq.(B.65) acquires the extra piece

$$\begin{aligned}
&\delta\left(\int_{\mathcal{B}} \mathrm{tr}\left[Q^{\mathbb{C}}(dQ^{\mathbb{C}})^4\right]+8\int_{\mathcal{M}} \mathrm{tr}\begin{bmatrix} i\,A\,Q^{\mathbb{C}}(dQ^{\mathbb{C}})^2-(AQ^{\mathbb{C}})^2dQ^{\mathbb{C}}\\ -\frac{i}{3}(A\,Q^{\mathbb{C}})^3+i\,A^3Q^{\mathbb{C}}\end{bmatrix}\right)\\
=&8\int_{\mathcal{M}} \mathrm{tr}\left[d\epsilon\left(Q^{\mathbb{C}}F+FQ^{\mathbb{C}}\right)\right]
\end{aligned} \tag{B.67}$$

where $F := dA + iA^2$. This last non-gauge invariant term, Eq.(B.67), can also be canceled out by adding

$$\text{Added term} \quad -8\int_{\mathcal{M}} \mathrm{tr}\left[A\,Q^{\mathbb{C}}\,F + A\,F\,Q^{\mathbb{C}}\right]. \tag{B.68}$$

Indeed, under the gauge transformation the added term transforms as

$$\begin{aligned}
&\delta\left(-8\int_{\mathcal{M}} \mathrm{tr}\left[A\,Q^{\mathbb{C}}\,F + A\,F\,Q^{\mathbb{C}}\right]\right)\\
&= -8\int_{\mathcal{M}} \mathrm{tr}\begin{bmatrix} (d\epsilon + i[A,\epsilon])(Q^{\mathbb{C}}\,F + F\,Q^{\mathbb{C}}) \\ +i\,[F,\epsilon](A\,Q^{\mathbb{C}} + Q^{\mathbb{C}}\,A) \\ +i\,[Q^{\mathbb{C}},\epsilon](F\,A + A\,F) \end{bmatrix}\\
&= -8\int_{\mathcal{M}} \mathrm{tr}\left[d\epsilon\,\left(Q^{\mathbb{C}}\,F + F\,Q^{\mathbb{C}}\right)\right]
\end{aligned}$$

which cancels Eq.(B.67). Thus the entire $U(n)$ symmetry can be gauged without anomaly. This is consistent with the free fermion prediction.

In summary, the $U(n)$ gauged nonlinear sigma model in $(2+1)$-D is

$$\begin{aligned}
W[Q^{\mathbb{C}}, A] = \frac{1}{2\lambda_3}\int_{\mathcal{M}} d^3x\, \mathrm{tr}\Big[\Big(\partial_\mu Q^{\mathbb{C}} + i[A_\mu, Q^{\mathbb{C}}]\Big)^2\Big] - \frac{2\pi i}{256\pi^2}\Big\{\int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{C}}\,\Big(d\tilde{Q}^{\mathbb{C}}\Big)^4\Big]\\
+8\int_{\mathcal{M}} \mathrm{tr}\Big[iAQ^{\mathbb{C}}(dQ^{\mathbb{C}})^2 - (AQ^{\mathbb{C}})^2 dQ^{\mathbb{C}} - \frac{i}{3}(AQ^{\mathbb{C}})^3 + iA^3Q^{\mathbb{C}} - AQ^{\mathbb{C}}F - AFQ^{\mathbb{C}}\Big]\Big\}.
\end{aligned} \tag{B.69}$$

Real class in $(2+1)$**-D**

The gauged nonlinear sigma model can be derived from the results in preceding subsection by 1) restricting the order parameter $Q^{\mathbb{C}} \in \frac{U(n)}{U(n/2)\times U(n/2)}$ to the sub-manifold $Q^{\mathbb{R}} \in \frac{O(n)}{O(n/2)\times O(n/2)}$, 2) restricting the gauge group from $U(n)$ to $O(n)$, and 3) divide the effective action by a factor of two (see B.4.1). The resulting gauged nonlinear sigma model action is

$$\begin{aligned}
W[Q^{\mathbb{R}}, A] = \frac{1}{4\lambda_3}\int_{\mathcal{M}} d^3x\, \mathrm{tr}\Big[\Big(\partial_\mu Q^{\mathbb{R}} + i[A_\mu, Q^{\mathbb{R}}]\Big)^2\Big] - \frac{2\pi i}{512\pi^2}\Big\{\int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{R}}\,\Big(d\tilde{Q}^{\mathbb{R}}\Big)^4\Big]\\
+8\int_{\mathcal{M}} \mathrm{tr}\Big[iAQ^{\mathbb{R}}(dQ^{\mathbb{R}})^2 - (AQ^{\mathbb{R}})^2 dQ^{\mathbb{R}} - \frac{i}{3}(AQ^{\mathbb{R}})^3 + iA^3Q^{\mathbb{R}} - AQ^{\mathbb{R}}F - AFQ^{\mathbb{R}}\Big]\Big\}
\end{aligned} \tag{B.70}$$

Here A is the gauge connection for the $O(n)$ gauge group. Again the entire $O(n)$ symmetry is anomaly free, agreeing with the free fermion prediction.

Complex class in $(3+1)$-D

The emergent symmetry is $U_+(n)\times U_-(n)$. The gauged WZW term was written down by Witten [26] with a minor correction in Ref.[60]. To simplify the notation, we will define

$$\alpha_1 := dQ^{\mathbb{C}}\, Q^{\mathbb{C}\dagger}, \quad \alpha_2 := Q^{\mathbb{C}\dagger} dQ^{\mathbb{C}}.$$

The derivation is rather long, so we shall not repeat it here. The result is [26]

$$\begin{aligned}
&W[Q^{\mathbb{C}}, A_+, A_-] \\
&= -\frac{1}{2\lambda_4^2}\int_{\mathcal{M}} d^4x \,\mathrm{tr}\left[\left(Q^{\mathbb{C}\dagger}\left(\partial_\mu Q^{\mathbb{C}} - iQ^{\mathbb{C}}A_{+,\mu} + iA_{-,\mu}Q^{\mathbb{C}}\right)\right)^2\right] \\
&- \frac{2\pi}{480\pi^3}\Bigg\{\int_{\mathcal{B}} \mathrm{tr}\left[(Q^{\mathbb{C}\dagger}dQ^{\mathbb{C}})^5\right] \\
&+ 5\int_{\mathcal{M}} \mathrm{tr}\left[\begin{array}{c}
-i\left(A_+\alpha_2^3 + A_-\alpha_1^3\right) - ((dA_+A_+ + A_+dA_+)\alpha_2 + (dA_-A_- + A_-dA_-)\alpha_1) \\
+dA_-dQ^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger} - dA_+d(Q^{\mathbb{C}\dagger})A_-Q^{\mathbb{C}} + A_+Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}}\alpha_2^2 - A_-Q^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger}\alpha_1^2 \\
+\frac{1}{2}\left((A_-\alpha_1)^2 - (A_+\alpha_2)^2\right) - i\left(A_-^3\alpha_1 + A_+^3\alpha_2\right) \\
+i\left((dA_+A_+ + A_+dA_+)\,Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}} - (dA_-A_- + A_-dA_-)\,Q^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger}\right) \\
-i\left(A_-Q^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger}A_-\alpha_1 + A_+Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}}A_+\alpha_2\right) \\
+\left(A_+^3Q^{\mathbb{C}\dagger}A_-Q^{\mathbb{C}} - A_-^3Q^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger}\right) + \frac{1}{2}(Q^{\mathbb{C}}A_+Q^{\mathbb{C}\dagger}A_-)^2
\end{array}\right]\Bigg\}
\end{aligned} \tag{B.71}$$

Under the infinitesimal gauge transformation, the action transforms as

$$\delta W = \frac{2\pi i}{48\pi^3}\int_{\mathcal{M}} \mathrm{tr}\left[\epsilon_+\Big((dA_+)^2 - \frac{i}{2}d(A_+^3)\Big) - \epsilon_-\Big((dA_-)^2 - \frac{i}{2}d(A_-^3)\Big)\right]$$

The situation is similar to the $(1+1)$-D case: there is an anomaly if we gauge $U_+(n)$ and $U_-(n)$ independently. However, there is no anomaly if we only gauge the diagonal part of $U(n)$. This is consistent with the free fermion prediction.

Real class in $(3+1)$-D

The gauged nonlinear sigma model for the real class in $(3+1)$-D can be derived from the results of the preceding subsection by 1) restricting the order parameter $Q^{\mathbb{C}} \in U(n)$ to the submanifold $Q^{\mathbb{R}} \in \frac{U(n)}{O(n)}$ (the space of symmetric unitary matrix), 2) restricting the gauge group from $U_+(n)\times U_-(n)$ which transforms $Q^{\mathbb{C}}$ according to

$$Q^{\mathbb{C}} \to g_-^\dagger \cdot Q^{\mathbb{C}} \cdot g_+,$$

to the sub-group $U(n)$ (the global symmetry group in the real class is $U(n)$), which transforms $Q^{\mathbb{R}}$ according to

$$Q^{\mathbb{R}} \xrightarrow{u\in U(n)} u^T \cdot Q^{\mathbb{R}} \cdot u,$$

and 3) divide the action by a factor of two (see B.4.1). The resulting gauged nonlinear sigma model is

$$\begin{aligned}
&W[Q^{\mathbb{R}}, A] \\
&= -\frac{1}{4\lambda_4^2}\int\limits_{\mathcal{M}} d^4x \,\mathrm{tr}\left[\left(Q^{\mathbb{R}\dagger}\left(\partial_\mu Q^{\mathbb{R}} - iQ^{\mathbb{R}}A_\mu + i(-A_\mu^T)Q^{\mathbb{R}}\right)\right)^2\right] \\
&- \frac{2\pi}{960\pi^3}\Bigg\{\int_{\mathcal{B}} \mathrm{tr}\left[(Q^{\mathbb{R}\dagger}dQ^{\mathbb{R}})^5\right] \\
&+ 5\int_{\mathcal{M}} \mathrm{tr}\begin{bmatrix}
-i\left(A\alpha_2^3 + (-A^T)\alpha_1^3\right) - \left((dAA + AdA)\alpha_2 + (d(-A^T)(-A^T) + (-A^T)d(-A^T))\alpha_1\right) \\
+d(-A^T)dQ^{\mathbb{R}}AQ^{\mathbb{R}\dagger} - dAd(Q^{\mathbb{R}\dagger})(-A^T)Q^{\mathbb{R}} + AQ^{\mathbb{R}\dagger}(-A^T)Q^{\mathbb{R}}\alpha_2^2 - (-A^T)Q^{\mathbb{R}}AQ^{\mathbb{R}\dagger}\alpha_1^2 \\
+\frac{1}{2}\left(((-A^T)\alpha_1)^2 - (A\alpha_2)^2\right) - i\left((-A^T)^3\alpha_1 + A^3\alpha_2\right) \\
+i\left((dAA + AdA)\,Q^{\mathbb{R}\dagger}(-A^T)Q^{\mathbb{R}} - \left(d(-A^T)(-A^T) + (-A^T)d(-A^T)\right)Q^{\mathbb{R}}AQ^{\mathbb{R}\dagger}\right) \\
-i\left((-A^T)Q^{\mathbb{R}}AQ^{\mathbb{R}\dagger}(-A^T)\alpha_1 + AQ^{\mathbb{R}\dagger}(-A^T)Q^{\mathbb{R}}A\beta\right) \\
+\left(A^3Q^{\mathbb{R}\dagger}(-A^T)Q^{\mathbb{R}} - (-A^T)^3Q^{\mathbb{R}}AQ^{\mathbb{R}\dagger}\right) + \frac{1}{2}(Q^{\mathbb{R}}AQ^{\mathbb{R}\dagger}(-A^T))^2
\end{bmatrix}\Bigg\}.
\end{aligned} \tag{B.72}$$

Here we have used the definition

$$\alpha_1 := dQ^{\mathbb{R}}\,Q^{\mathbb{R}\dagger}, \quad \alpha_2 := Q^{\mathbb{R}\dagger}dQ^{\mathbb{R}}.$$

Under the infinitesimal gauge transformation,

$$\begin{aligned}
&Q^{\mathbb{R}} \to e^{i\epsilon^T}Q^{\mathbb{R}}e^{i\epsilon} \\
&A \to A + d\epsilon + i[A, \epsilon],
\end{aligned}$$

the gauged nonlinear sigma model acquires an addition piece

$$\delta W = \frac{2\pi i}{96\pi^3}\int\limits_{\mathcal{M}} \mathrm{tr}\Big[d\epsilon\left(AdA - \frac{i}{2}A^3\right) + d\epsilon^T\Big((-A^T)d(-A^T) - \frac{i}{2}(-A^T)^3\Big)\Big] \tag{B.73}$$

manifesting the ‘t Hooft anomaly of associated with $U(n)$.

However, if we only gauge the $O(n)$ subgroup of $U(n)$

$$\epsilon^T = -\epsilon, \quad A^T = -A.$$

Under such condition the two terms in Eq.(B.73) cancel. Thus the $O(n)$ subgroup anomaly free. This agrees with the free fermion anomaly.

B.6.2 Anomalies with respect to the discrete groups

After gauging the anomaly-free part of the continuous group, it is straightforward to determine how the resulting action transform under discrete symmetries. The necessary input is the transformation of the gauge field and the $Q^{\mathbb{C},\mathbb{R}}$. Here we simply state the results. In (1+1)-D and (3+1)-D there is no anomaly with respect to discrete symmetries after gauging the anomaly-free part of the continuous symmetries. In $(2+1)$-D, gauging the continuous symmetry breaks the time-reversal symmetry as discussed in subsection 4.12.2.

B.7 Soliton's statistics

As discussed in subsection 4.13.1 of the main text, in $(2+1)$-D and $(3+1)$-D the mass manifolds for $Q^{\mathbb{C},\mathbb{R}}$ support solitons for sufficiently large n (number of flavors). In this appendix, we follow Ref.[27] to determine the statistics of soliton. This is achieved by computing the Berry phase, arising from the WZW term, of an adiabatic self-rotating soliton.

Here is our strategy. (1) We write down the $Q^{\mathbb{C},\mathbb{R}}$ configuration corresponding to a static unit soliton. (2) Based on the result of (1), we write down the $Q^{\mathbb{C},\mathbb{R}}$ configuration corresponding to an adiabatic self-2π-rotating soliton. (3) We plug the $Q^{\mathbb{C},\mathbb{R}}$ configuration constructed in (2) into the WZW term to compute the Berry phase.

Because the space-time manifold S^D is incompatible with the $Q^{\mathbb{C},\mathbb{R}}$ configuration of a single soliton[4], in this section we shall follow Ref.[27] and use

$$\mathcal{M} = S^{D-1} \times S^1$$

as the space-time manifold. Here S^{D-1} is in the spatial manifold and S^1 is the loop in time. The extended manifold needed to define the WZW term is [27]

$$\mathcal{B} = S^{D-1} \times D^2,$$

where D^2 is a two-dimensional disk with the boundary $\partial D^2 = S^1$ being the time loop.

B.7.1 Complex class in $(2+1)$-D

[4]On S^D, the infinite future corresponds to a single point. It follows that $Q^{\mathbb{C},\mathbb{R}}$ is a constant matrix at infinite future. This is incompatible with the single soliton configuration.

The mass manifold is

$$\frac{U(n)}{U(n/2)\times U(n/2)}.$$

For $n \geq 4$ *both* homotopy groups π_2 (relevant to the existence of soliton) and π_4 (relevant to the existence of the WZW term) are stabilized (see table B.1). This is the situation we shall focus on in the following.

To write down a static soliton configuration, let us begin with $n = 2$. This is because as far as π_2 (relevant to the existence of soliton) is concerned, it stabilizes at $n = 2$, for which the mass manifold is

$$\frac{U(2)}{U(1)\times U(1)} = S^2,$$

and $Q^{\mathbb{C}}$ is a 2×2 hermitian matrix. Here a unit soliton is a degree 1 map from the spatial manifold S^2 to the mass manifold S^2. An example of such map is

$$Q^{\mathbb{C}}_{\rm sol}(\theta,\phi) = \vec{n}\cdot\vec{\sigma} \quad \text{where} \quad \vec{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{B.74}$$

where θ and ϕ are the usual coordinates on S^2. This can be verified by computing the topological invariant associated with the soliton quantum number

$$I_2 = \frac{i}{16\pi}\int \text{tr}\left[Q^{\mathbb{C}}_{\rm sol}(dQ^{\mathbb{C}}_{\rm sol})^2\right] = 1.$$

For $n \geq 4$ we can write down a static unit soliton configuration as the *direct sum* of the 2×2 $Q^{\mathbb{C}}_{\rm sol}(\theta,\phi)$ in Eq.(B.74) with a number of Pauli matrices Z, i.e.,

$$Q^{\mathbb{C}}_{\rm sol}(\theta,\phi) = \vec{n}(\theta,\phi)\cdot\vec{\sigma}\oplus Z\oplus Z... = \begin{pmatrix} n_3 & n_1 - in_2 & 0 & 0\\ n_1 + in_2 & -n_3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\oplus Z... \tag{B.75}$$

To construct the configuration of a 2π-self-rotating soliton (around, e.g., the n_x axis) we introduce the following space-time dependent $Q^{\mathbb{C}}$, namely,

$$Q^{\mathbb{C}}(\theta,\phi,\tau) = R^T(\tau)\cdot Q^{\mathbb{C}}_{\rm sol}(\theta,\phi)\cdot R(\tau), \quad \text{where}$$
$$R(\tau) = \left[\begin{pmatrix} e^{+i\frac{\tau}{2}} & 0 & 0 & 0\\ 0 & e^{-i\frac{\tau}{2}} & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\oplus I...\right] \tag{B.76}$$

Here τ ranges from 0 to 2π along the time loop, and $Q^{\mathbb{C}}_{\rm sol}(\theta,\phi)$ is given by Eq.(B.75).

As discussed earlier, in order to calculate the Berry phase arising from the WZW term, we needs to extend the space-time manifold from $\mathcal{M} = S^2\times S^1$ to $\mathcal{B} = S^2\times D^2$. However, this extension is complicated by the fact that the orthogonal matrices $R(\tau)$ in

Eq.(B.76) is not single-valued as τ runs through the time loop (note, however, $Q^{\mathbb{C}}(\theta,\phi,\tau)$ is single valued). To overcome this difficulty, we use the algebraic fact observed by Witten [27] that to reproduce the same $Q^{\mathbb{C}}(\theta,\phi,\tau)$, one can replace the $R(\tau)$ in Eq.(B.76) by the following single valued matrix

$$R(\tau) = \left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\tau} & 0 & 0 \\ 0 & 0 & e^{+i\tau} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \oplus I...\right].$$

After such replacement, one can extend it to $S^2 \times D^2$ by writing

$$\tilde{Q}^{\mathbb{C}}(\theta,\phi,\tau,u) = \tilde{R}^T(\tau,u)\cdot Q^{\mathbb{C}}_{\text{sol}}(\theta,\phi)\cdot \tilde{R}(\tau,u), \quad \text{where}$$
$$\tilde{R}(\tau,u) = \left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \sin u\, e^{-i\tau} & \cos u & 0 \\ 0 & -\cos u & \sin u\, e^{+i\tau} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \oplus I...\right], \tag{B.77}$$

where $u \in [0,\pi]$.

It's straightforward, though slightly tedious, to plug Eq.(B.77) in the WZW term to obtain

$$W_{\text{WZW}}[\tilde{Q}^{\mathbb{C}}] = -\frac{2\pi i}{256\pi^2}\int_{\mathcal{B}} \text{tr}\left[\tilde{Q}^{\mathbb{C}}\,(d\tilde{Q}^{\mathbb{C}})^4\right] = i\pi.$$

Therefore we conclude that the Berry's phase due to the self-rotation is -1, implying the unit soliton is a fermion.

B.7.2 Real class in $(2+1)$-D

The relevant mass manifold is $\frac{O(n)}{O(\frac{n}{2})\times O(\frac{n}{2})}$. From table B.2, both π_2 and π_4 are stabilized for $n \geq 10$. In the following we shall restrict ourselves to such situation.

Unlike the case of complex class, the stabilized π_2 is

$$\pi_2\left(\frac{O(n)}{O(\frac{n}{2})\times O(\frac{n}{2})}\right) = Z_2,$$

rather than Z. As a consequence, unlike the soliton in the preceding section, there is no integral form of topological invariant we can use to test whether a proposed $Q^{\mathbb{R}}_{\text{sol}}$ configuration indeed corresponds to the non-trivial element of Z_2. The purpose of the following subsection is to establish such a testing method.

How to test whether a proposed Z_2 soliton is trivial or not

Let's consider the "fibration"

$$F \xrightarrow{i} E \xrightarrow{p} B. \tag{B.78}$$

Here F stands for "fiber space", E stands for "total space", and B stands for "base space". "Fibration" means that locally (i.e., in a small neighborhood of the base space B), the total space is the Cartesian product of the base space and the fiber space. In Eq.(B.78) i and p stand for the inclusion and projection maps, respectively. They satisfies the property that *image* of i is the *kernel* of p. It is a non-trivial theorem that the fibration in Eq.(B.78) induces the following long exact sequence of mappings between homotopy groups (see, e.g., [59])

$$...\pi_n(F) \xrightarrow{i_*} \pi_n(E) \xrightarrow{p_*} \pi_n(B) \rightarrow \pi_{n-1}(F) \xrightarrow{i_*} \pi_{n-1}(E) \xrightarrow{p_*} \pi_{n-1}(B)... \tag{B.79}$$

Here i_*, p_* stand for the map between mapping classes induced by the inclusion and projection, respectively. Eq.(B.79) has the property that for two consecutive mappings between homotopy groups, the *image* of the preceding map is equal to the *kernel* of the subsequent map.

In our case

$$F = O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right), \quad E = O(n), \quad B = \frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}.$$

The inclusion and projection maps in Eq.(B.78) are defined by

$$\begin{aligned} &(O_1, O_2) \xrightarrow{i} O := \begin{pmatrix} O_1 & 0 \\ 0 & O_2 \end{pmatrix} \text{ where } O_{1,2} \in O\left(\frac{n}{2}\right) \text{ and } O \in O(n) \\ &O \xrightarrow{p} S := O \cdot diag(\underbrace{+1, ..., +1}_{n/2}, \underbrace{-1, ..., -1}_{n/2}) \cdot O^T, \text{ where } S \in \frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}. \end{aligned} \tag{B.80}$$

Our goal is to decide whether a given

$$S^2 \xrightarrow{f_2} \frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})} \tag{B.81}$$

is topologically trivial or not. To answer that we consider the following sub-sequence of Eq.(B.79)

$$\pi_2(O(n)) \xrightarrow{p_*} \pi_2\left(\frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}\right) \xrightarrow{\beta_*} \pi_1\left(O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)\right) \xrightarrow{i_*} \pi_1(O(n)).$$

where it is known that

$$
\begin{aligned}
&\pi_2(O(n)) = 0\\
&\pi_2\left(\frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}\right) = \mathbb{Z}_2\\
&\pi_1\left(O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)\right) = \mathbb{Z}_2 \times \mathbb{Z}_2\\
&\pi_1(O(n)) = \mathbb{Z}_2.
\end{aligned}
$$

The map

$$\pi_1\left(O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)\right) \xrightarrow{i_*} \pi_1(O(n))$$

sends

$$\mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2 \;\text{ via }\; (s_1, s_2) \to (s_1 + s_2 \mod 2).$$

Hence the kernel of this map is $(0,0)$ and $(1,1)$. According to Eq.(B.79) these should be the image of the map

$$\pi_2\left(\frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}\right) \xrightarrow{\beta_*} \pi_1\left(O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)\right),$$

or equivalently,

$$\mathbb{Z}_2 \xrightarrow{\beta_*} \mathbb{Z}_2 \times \mathbb{Z}_2. \tag{B.82}$$

The requirement that the image of the map in Eq.(B.82) be $(0,0), (1,1)$, implies that

$$s \xrightarrow{\beta_*} (s, s). \tag{B.83}$$

Therefore the soliton configuration, which is an representative of the $s = 1$ element of $\pi_2\left(\frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}\right)$, is mapped to a configuration representative of the $(1,1)$ element of $\pi_1\left(O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)\right)$ under β. Hence if we can tell whether a representative map of $\pi_1(O(n/2))$ is trivial or not, we can deduce whether the configuration in Eq.(B.81) is topologically non-trivial by applying β to it.

But this requires us to know how to construct the β map. To achieve that we consider the following commutative diagram (a diagram is commutative if different paths leading from the same initial space to the final space are the same map. The fact that the following diagram is commutative is by construction.)

$$
\begin{array}{ccccc}
S^1 & \xrightarrow{\delta_1} & D^2 & \xrightarrow{\gamma_2} & S^2\\
\downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle \lambda} & & \downarrow{\scriptstyle f_2}\\
O(\frac{n}{2}) \times O(\frac{n}{2}) & \xrightarrow{i} & O(n) & \xrightarrow{p} & \frac{O(n)}{O(\frac{n}{2}) \times O(\frac{n}{2})}
\end{array}
$$

Here δ_1 is the inclusion map which maps S^1 to the boundary of the 2-dimensional disk D^2; γ_2 is the map that compactifies the boundary of D^2 to single point; f_2 is the map in Eq.(B.81) and f_1 is the map obtained by applying β to f_2, i.e., $\beta[f_2] = f_1$. By knowing whether f_1 is a non-trivial map representing $(1,1)$ in Eq.(B.83) we can deduce whether f_2 is a non-trivial soliton configuration. In the commutative diagram λ is the homotopy lift of $f_2 \circ \gamma_2$. The fact that such a lift exists is because D^2 is homeomorphic to a two dimensional cube, i.e., a square, hence by the homotopy lifting property λ exists.

Because the image of $\gamma_2 \circ \delta_1$ is a point, so does the image of $f_2 \circ \gamma_2 \circ \delta_1 = p \circ \lambda \circ \delta_1$. This implies $(\lambda \circ \delta_1)[S^1]$ is in the kernel of the map $O(n) \xrightarrow{p} \frac{O(n)}{O(\frac{n}{2})\times O(\frac{n}{2})}$. Since $(\lambda \circ \delta_1)[S^1]$ is projected to a point in the base space, it must be contained entirely in a single fiber $O\left(\frac{n}{2}\right) \times O\left(\frac{n}{2}\right)$. Therefore the sought-after f_1 is given by

$$f_1 = \lambda \circ \delta_1.$$

The above arguments allows us take the map f_2 as input and produce the map f_1 as output, i.e., we have constructed β.

In the following we apply the construction discussed above to the following proposed soliton configuration[5]

$$f_2 : (\theta, \phi) \to Q^{\mathbb{R}}_{\text{sol}} = (n_1 XI + n_2 EE + n_3 ZI) \oplus Z \oplus Z... \tag{B.84}$$

where

$$\vec{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

Because $Q^{\mathbb{R}}_{\text{sol}}$ is a real symmetric matrix, it can be diagonalized by orthogonal transformation

$$Q^{\mathbb{R}}_{\text{sol}} = W \cdot \left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \oplus Z \oplus Z... \right] \cdot W^T$$

where

$$W(\theta,\phi) = \begin{pmatrix} \cos\frac{\theta}{2}\cos\phi & -\cos\frac{\theta}{2}\sin\phi & -\sin\frac{\theta}{2} & 0 \\ \cos\frac{\theta}{2}\sin\phi & \cos\frac{\theta}{2}\cos\phi & 0 & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & 0 & \cos\frac{\theta}{2}\cos\phi & \cos\frac{\theta}{2}\sin\phi \\ 0 & \sin\frac{\theta}{2} & -\cos\frac{\theta}{2}\sin\phi & \cos\frac{\theta}{2}\cos\phi \end{pmatrix} \oplus I...$$

Naively, one might think $W(\theta,\phi)$ is a mapping between S^2 and $O(n)$. However, this is not true. To see it, let's inspect $W(0,\phi)$ and $W(\pi,\phi)$.

$$W(0,\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix} \oplus I...,$$

[5] It is illuminating to compare $Q^{\mathbb{R}}_{\text{sol}}$ with $Q^{\mathbb{C}}_{\text{sol}}$ in Eq.(B.75), namely, $Q^{\mathbb{R}}_{\text{sol}} = \text{Re}[Q^{\mathbb{C}}_{sol}] \otimes I + \text{Im}[Q^{\mathbb{C}}_{sol}] \otimes E$.

$$W(\pi,\phi) = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \oplus I...$$

The fact that $W(0,\phi)$ depends on ϕ and $W(\pi,\phi)$ does not implies that we should view W as a mapping between D^2 and $O(n)$, where $\theta = 0$ corresponds to the boundary while $\theta = \pi$ corresponds to the center of D^2 (i.e., the radius of D^2 is $\pi - \theta$). In fact, W is the map λ in the commutative diagram, namely,

$$\lambda = W.$$

Hence the map f_1 is given by

$$\begin{aligned} f_1(\phi) = W(0,\phi) &= \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix} \oplus I... \\ &= \left\{ \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \oplus 1... \right\} \oplus \left\{ \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \oplus 1... \right\}. \end{aligned} \tag{B.85}$$

To summarize, given the map f_2 in Eq.(B.84), we have obtained the map f_1 in the commutative diagram via Eq.(B.85).

Now we are ready to determine whether Eq.(B.84) is a topological non-trivial soliton configuration. It is known that the following map from S^1 to $O(n/2)$

$$\tilde{f}_1(\phi) = \left\{ \begin{pmatrix} \cos\phi & \mp\sin\phi \\ \pm\sin\phi & \cos\phi \end{pmatrix} \oplus 1... \right\}$$

is a representative of the generator of $\pi_1(O(n/2)) = \mathbb{Z}_2$. Thus the mapping class of f_1 in Eq.(B.85) is the $(1,1)$ element of $\pi_1(O(n/2) \times O(n/2)) = \mathbb{Z}_2 \times \mathbb{Z}_2$. It follows that f_2 in Eq.(B.84) is a representative of the generator of $\pi_2(\frac{O(n)}{O(n/2)\times O(n/2)}) = \mathbb{Z}_2$, i.e., it is a soliton configuration.

The Berry phase of a self-rotating $\mathbb{Z}_2$ soliton

To calculate the Berry's phase due to a 2π self-rotation of the soliton in Eq.(B.84), we rotate the soliton configuration to produce $Q^{\mathbb{R}}(\theta,\phi,\tau)$ in the same way in the appendix B.7.1. After all dust settles, we end up with

$$Q^{\mathbb{R}}(\theta,\phi,\tau) = \mathrm{Re}[Q^{\mathbb{C}}(\theta,\phi,\tau)] \otimes I + \mathrm{Im}[Q^{\mathbb{C}}(\theta,\phi,\tau)] \otimes E =$$

$$\begin{pmatrix} \cos\theta & 0 & \sin\theta\cos(\phi+\tau) & -\sin\theta\sin(\phi+\tau) \\ 0 & \cos\theta & \sin\theta\sin(\phi+\tau) & \sin\theta\cos(\phi+\tau) \\ \sin\theta\cos(\phi+\tau) & \sin\theta\sin(\phi+\tau) & -\cos\theta & 0 \\ -\sin\theta\sin(\phi+\tau) & \sin\theta\cos(\phi+\tau) & 0 & -\cos\theta \end{pmatrix} \oplus Z \oplus Z...$$

Similarly, we can extend the space-time to one extra dimension by defining

$$\tilde{Q}^{\mathbb{R}}(\theta,\phi,\tau,u) := \mathrm{Re}[\tilde{Q}^{\mathbb{C}}(\theta,\phi,\tau,u)] \otimes I + \mathrm{Im}[\tilde{Q}^{\mathbb{C}}(\theta,\phi,\tau,u)] \otimes E.$$

(It's easy to check it suffices all the properties we want for extension). Plugging the extended $\tilde{Q}_2^{\mathbb{R}}$ into the WZW term, we find

$$W_{WZW}[Q^{\mathbb{R}}] = 2\pi i \left(-\frac{1}{512\pi^2} \int_{\mathcal{B}} \mathrm{tr}\left[\tilde{Q}^{\mathbb{R}}\,(d\tilde{Q}^{\mathbb{R}})^4\right] \right) = i\pi.$$

Hence the soliton is again a fermion.

B.7.3 Complex class in $(3+1)$-D

The mass manifold in this situation is $U(n)$. π_3 (relevant to the existence of soliton) and the π_5 (relevant to the existence of the WZW term) are *both* stabilized for $n \geq 3$. In the following, we shall restrict ourselves to such conditions.

The fact a unit soliton in this mass manifold is a fermion has already been discussed in [27]. We briefly repeat the argument here for completeness. To construct a static soliton we start from $n=2$ (as far as π_3 is concerned, it stabilizes for $n \geq 2$ with $\pi_3(U(n)) = \mathbb{Z}$.) Thus the $n=2$ unit soliton is just the degree one map of $S^3 \to SU(2) \sim S^3$ [6]. We can choose the unit soliton configuration to be

$$Q^{\mathbb{C}}_{\mathrm{sol}}(\vec{\Omega}) = \begin{pmatrix} \Omega_0 + i\Omega_3 & i(\Omega_1 - i\Omega_2) \\ i(\Omega_1 + i\Omega_2) & \Omega_0 - i\Omega_3 \end{pmatrix}$$

where

$$\Omega_0^2 + \Omega_1^2 + \Omega_2^2 + \Omega_3^2 = 1$$

are the coordinate on S^3. For $n \geq 3$ one can write the unit soliton as

$$Q^{\mathbb{C}}_{\mathrm{sol}}(\vec{\Omega}) = \begin{pmatrix} \Omega_0 + i\Omega_3 & i(\Omega_1 - i\Omega_2) \\ i(\Omega_1 + i\Omega_2) & \Omega_0 - i\Omega_3 \end{pmatrix} \oplus 1 \oplus 1 ...$$

Next, we rotate the unit soliton in, say, the Ω_1-Ω_2 plane by 2π. The time-dependent soliton configuration can be written as

$$\begin{aligned} Q^{\mathbb{C}}(\vec{\Omega},\tau) &= \left[\begin{pmatrix} e^{-i\frac{\tau}{2}} & 0 & 0 \\ 0 & e^{+i\frac{\tau}{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus 1... \right] \cdot Q^{\mathbb{C}}_{\mathrm{sol}}(\vec{\Omega}) \cdot \left[\begin{pmatrix} e^{+i\frac{\tau}{2}} & 0 & 0 \\ 0 & e^{-i\frac{\tau}{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus 1.. \right] \\ &= \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{+i\tau} & 0 \\ 0 & 0 & e^{-i\tau} \end{pmatrix} \oplus 1.. \right] \cdot Q^{\mathbb{C}}_{\mathrm{sol}}(\vec{\Omega}) \cdot \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i\tau} & 0 \\ 0 & 0 & e^{+i\tau} \end{pmatrix} \oplus 1.. \right] \end{aligned}$$

[6]Note that $\pi_3(U(n)) = \mathbb{Z}$ originates from the $SU(n)$ part of $U(n)$. Among other things, it means that we can limit ourselves to the $SU(n)$ WZW term for the Berry phase calculation.

where $\tau \in [0, 2\pi] = S^1$ is the time parameter. One can extend the configuration to $\mathcal{B} = S^3 \times D^2$, where D^2 is the two-dimensional disk with radius $u \in [0, \pi]$, by

$$\tilde{Q}^{\mathbb{C}}(\vec{\Omega}, \tau, u) = \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin u\, e^{+i\tau} & -\cos u \\ 0 & \cos u & \sin u\, e^{-i\tau} \end{pmatrix} \oplus 1...\right] \cdot Q^{\mathbb{C}}_{sol}(\vec{\Omega}) \cdot \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin u\, e^{-i\tau} & \cos u \\ 0 & \cos u & \sin u\, e^{+i\tau} \end{pmatrix} \oplus 1..\right] \tag{B.86}$$

Plugging Eq.(B.86) into the WZW term gives

$$W_{WZW}[\tilde{Q}^{\mathbb{C}}] = 2\pi i \left(\frac{i}{480\pi^3} \int_{\mathcal{B}} \mathrm{tr}\Big[\left(\tilde{Q}^{\mathbb{C}\dagger} d\tilde{Q}^{\mathbb{C}} \right)^5 \Big] \right) = i\pi.$$

Hence the unit soliton is a fermion.

B.7.4 Real class in $(3+1)$-D

The mass manifold is $U(n)/O(n)$. Here π_3 (relevant to the existence of soliton) and π_5 (relevant to the existence of the WZW term) are both stabilized for $n \geq 6$. To write down the degree one soliton in $U(n)/O(n)$, let's first look at the fibration

$$O(n) \xrightarrow{i} U(n) \xrightarrow{p} U(n)/O(n). \tag{B.87}$$

Here the projection p is defined by

$$u \xrightarrow{p} u_S = u^T \cdot u, \text{ where } u \in U(n),\ u_S \in U(n)/O(n). \tag{B.88}$$

After the homotopy groups are stabilized, the long exact sequence associated with Eq.(B.87) is given by

$$\begin{array}{ccccccccccc} ...\pi_4(U(n)/O(n)) & \to & \pi_3(O(n)) & \xrightarrow{i_*} & \pi_3(U(n)) & \xrightarrow{p_*} & \pi_3(U(n)/O(n)) & \to & \pi_2(O(n))... \\ ... \quad 0 & \to & \mathbb{Z} & \xrightarrow{i_*} & \mathbb{Z} & \xrightarrow{p_*} & \mathbb{Z}_2 & \to & 0 \quad ... \end{array}$$

This implies that we can construct the unit soliton in $U(n)/O(n)$ by taking a unit soliton in $U(n)$, namely $Q^{\mathbb{C}}_{\text{sol}}$ in appendix B.7.3, and perform the projection map in Eq.(B.88), namely,

$$Q^{\mathbb{R}}_{\text{sol}} = \left(Q^{\mathbb{C}}_{\text{sol}}\right)^T \cdot Q^{\mathbb{C}}_{\text{sol}}.$$

The time dependent $Q^{\mathbb{R}}(\vec{\Omega}, \tau)$ can be constructed by the similar projection of $Q^{\mathbb{C}}(\vec{\Omega}, \tau)$, i.e.,

$$Q^{\mathbb{R}}(\vec{\Omega}, \tau) = \left(Q^{\mathbb{C}}(\vec{\Omega}, \tau)\right)^T \cdot Q^{\mathbb{C}}(\vec{\Omega}, \tau).$$

The extended $\tilde{Q}^{\mathbb{R}}$ can also be constructed by the same projection

$$\tilde{Q}^{\mathbb{R}}(\vec{\Omega},\tau,u) = \left(\tilde{Q}^{\mathbb{C}}(\vec{\Omega},\tau,u)\right)^{T} \cdot \tilde{Q}^{\mathbb{C}}(\vec{\Omega},\tau,u).$$

The result $\tilde{Q}^{\mathbb{R}}(\vec{\Omega},\tau,u)$ can be substituted into the WZW term to obtain

$$W_{WZW}[Q^{\mathbb{R}}] = 2\pi i \left(\frac{i}{960\pi^3} \int_{\mathcal{B}} \mathrm{tr}\left[\left(\tilde{Q}^{\mathbb{R}\dagger} d\tilde{Q}^{\mathbb{R}} \right)^5 \right] \right) = i\pi$$

Therefore the unit soliton is again a fermion.

B.8 Bosonization for small flavor number

In this appendix we discuss the bosonization in cases when n, the number of flavors, is less than the value necessary to stabilize π_{D+1}(mass manifold), or the WZW term.

In some cases, although the homotopy group π_{D+1}(mass manifold) is not yet stabilized, it already contains $\mathbb{Z}$ as a subgroup. For instance, for real class in $(3+1)$-D, $\pi_5(U(3)/O(3)) = \mathbb{Z}\times\mathbb{Z}_2$. The nonlinear sigma model derived from fermion integration in appendix B.4 contains the level-1 WZW term, which is 2π times the topological invariant of the $\mathbb{Z}$ part of π_5. This is also true for $n = 6, 8$ of the non-charge-conserved cases in $(2+1)$-D. In these cases the story is unchanged.

In other cases π_{D+1}(mass manifold) is a finite abelian group, e.g., $\mathbb{Z}_2$, or even 0. This requires a case-by-case study. Here, instead of attempting at studying all possible cases, we shall focus on the case that is relevant to the applications in section 5 of the main text, namely, the case of $n = 2$ complex class in $(2+1)$-D (which is relevant to the discussions in subsection 5.3.2 of the main text).

B.8.1 Complex class in $2+1$ D with $n = 2$

The mass manifold is

$$\frac{U(2)}{U(1)\times U(1)} = S^2$$

and

$$\pi_4(S^2) = 0, \quad \text{but} \quad \pi_3(S^2) = \mathbb{Z}.$$

The generator of $\pi_3(S^2)$ is called Hopf map [61]. The question at hand is whether this signifies the presence of a topological term in the nonlinear sigma model. A similar situation occurs for, e.g., the non-linear sigma model describing the anti-ferromagnetic spin chains in (1 + 1)-D. There, the mass manifold is S^2 and $\pi_3(S^2) = 0$ but $\pi_2(S^2) = \mathbb{Z}$. In the nonlinear sigma model, there is a topological term associated with the π_2 in the non-linear sigma model, the θ term, which is responsible for the difference between the integer and half-integer spin chains [62, 63, 44].

To answer the question posed above, the derivation in appendix B.4 is not applicable. This is because the Hopf term (or the θ term) is invariant under arbitrary infinitesimal deformation of $Q^{\mathbb{C}}$.

B.8.2 Mass manifold enlargement

One way to proceed is to enlarge the mass manifold (or the target space of the order parameter). The idea [64] is as follows. If two order parameter configurations cannot be deformed into each other, as in the case where configurations correspond to different elements of π_D (in this case π_3), we can enlarge the mass manifold so that after the enlargement, one configuration can be continuously deformed to the other. One can then compute the Berry phase difference caused by infinitesimal order parameter variation using the method explained in B.4, and integrate the result. However, it is important to note that enlarging the mass manifold requires adding extra fermion flavors. It is important to make sure that the initial and the final order parameters couple to the added fermion flavors in a trivial way (i.e., in the added flavor space, the order parameters are the same constant) so that the Berry phase difference is originated from the original fermions. Finally, one accounts for the Berry phase by picking the coefficient in front of the π_D (here π_3) topological invariant.

Using this technique, Abanov [54] enlarged $\frac{U(2)}{U(1)\times U(1)} = S^2$ to

$$\frac{U(l+1)}{U(l)\times U(1)} = \mathbb{CP}^l$$

and showed that the nonlinear sigma model from the $n = 2$ fermion integration contains a $\theta = \pi$ Hopf term. In the following we will choose an alternative enlargement, namely,

$$\frac{U(2)}{U(1)\times U(1)} = S^2 \to \frac{U(4)}{U(2)\times U(2)}.$$

We will show that the result is consistent with that of Abanov. Because of the Hopf term, the unit soliton has fermion statistics [55].

B.8.3 The derivation of the Hopf term

The Hopf map is a map from S^3 with coordinate $(\Omega_0, \Omega_1, \Omega_2, \Omega_3)$ where $\sum_{i=0}^{3} \Omega_i^3 = 1$ to S^2 with coordinate (n_1, n_2, n_3) where $\sum_{i=1}^{3} n_i^2 = 1$. More explicitly,

$$\vec{\Omega} \xrightarrow{\text{Hopf}} \vec{n} = z^\dagger \sigma^a z \text{ where } z := \begin{pmatrix} \Omega_0 + i\Omega_1 \\ \Omega_2 + i\Omega_3 \end{pmatrix}. \tag{B.89}$$

The $Q^{\mathbb{C}}$ of the non-linear sigma model is given by

$$Q^{\mathbb{C}}(\vec{\Omega}) = \sum_{a=1}^{3} n^a(\vec{\Omega}) \cdot \sigma^a = 2\, z(\vec{\Omega})\, z(\vec{\Omega})^\dagger - I_2, \tag{B.90}$$

where z is given by Eq.(B.89). In writing down the 2nd equality we have used the identity

$$\sum_{a=1}^{3} \sigma_{ij}^a \sigma_{kl}^a = 2\delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl}.$$

In Eq.(B.90) the 2×2 matrix $Q^{\mathbb{C}}$ has eigenvalues ± 1, and z is the eigenvector associated with eigenvalue $+1$.

In the following, we enlarge the order parameter so that $Q^{\mathbb{C}}$ can be deformed to σ_z. To do so, we add two additional fermion flavors and enlarge the mass manifold to

$$\frac{U(4)}{U(2) \times U(2)}.$$

In the enlarged space the order parameter is given by

$$\begin{aligned} Q'^{\mathbb{C}}(\vec{\Omega}) &= Q^{\mathbb{C}}(\vec{\Omega}) \oplus (-Z) = \begin{pmatrix} Q^{\mathbb{C}}(\vec{\Omega}) & 0 \\ 0 & -Z \end{pmatrix} \\ &= \begin{pmatrix} 2\, z(\vec{\Omega})\, z(\vec{\Omega})^\dagger - I_2 & 0 \\ 0 & -Z \end{pmatrix} \end{aligned} \tag{B.91}$$

where $z(\vec{\Omega})$ is given by Eq.(B.89). Here the fermions associated with extra flavors couple to the mass term $Y \otimes (-Z)$ (see table 4.2). Although the $Q^{\mathbb{C}}$ given by Eq.(B.90) cannot be deformed to a constant configuration in the space $\frac{U(2)}{U(1)\times U(1)} = S^2$ (i.e., within the first 2×2 block) because

$$\pi_3(S^2) = \mathbb{Z},$$

it is possible to deform the 4×4 $Q'^{\mathbb{C}}$ to a constant matrix. This is because

$$\pi_3\left(\frac{U(4)}{U(2) \times U(2)}\right) = 0.$$

Now we explicitly construct such a deformation. First we rewrite Eq.(B.91) as

$$Q'^{\mathbb{C}}(\vec{\Omega}) = \begin{pmatrix} 2\, z'(\vec{\Omega})\, z'(\vec{\Omega})^\dagger - I_3 & 0 \\ 0 & +1 \end{pmatrix}$$
$$\text{where } z'(\vec{\Omega}) := \begin{pmatrix} \Omega_0 + i\Omega_1 \\ \Omega_2 + i\Omega_3 \\ 0 \end{pmatrix}.$$

We then write down a continuous deformation, as a function of u, as follows

$$\tilde{Q}'^{\mathbb{C}}(\vec{\Omega}, u) = \begin{pmatrix} 2\, \tilde{z}'(\vec{\Omega}, u)\, \tilde{z}'^\dagger(\vec{\Omega}, u) - I_3 & 0 \\ 0 & +1 \end{pmatrix}$$
$$\tilde{z}'(\vec{\Omega}, u) := \begin{cases} \Big(-\cos u\, (\Omega_0 + i\Omega_1)\,, -\cos u\, (\Omega_2 + i\Omega_3)\,, \sin u \Big)^T & \text{for } u \in [\frac{\pi}{2}, \pi] \\ \Big(\cos u, 0, \sin u \Big)^T & \text{for } u \in [0, \frac{\pi}{2}). \end{cases} \tag{B.92}$$

This extends the configuration from the space-time $\mathcal{M} = S^3$ at $u = \pi$ to a four dimensional disk $\mathcal{B} = D^4$ with $u \in [0, \pi]$. Here $\partial\mathcal{B} = \mathcal{M}$ and with u as the radial direction of D^4. For $u = \pi$ Eq.(B.92) reduces to Eq.(B.91), while for $u = 0$

$$\tilde{Q}'^{\mathbb{C}}(\vec{\Omega}, u = 0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} = Z \oplus (-Z).$$

Therefore at $u = 0$ and $u = \pi$ the fermions associated with the added flavors couples to exactly the same mass term $Y \otimes (-Z)$ according to table 4.2.

Eq.(B.92) constitutes an extension we need to define the WZW term (which is stabilized at $n = 4$). Now we can plug Eq.(B.92) into the WZW term of the $\frac{U(4)}{U(2)\times U(2)}$ non-linear sigma model. When all dust settles we obtain

$$W_{\text{WZW}}[\tilde{Q}'^{\mathbb{C}}] = -\frac{2\pi i}{256\pi^2} \int_{\mathcal{B}} \text{tr} \left[\tilde{Q}'^{\mathbb{C}} (d\tilde{Q}'^{\mathbb{C}})^4 \right] = i\pi.$$

This result agrees with that of Ref.[54] and suggests the existence of a $\theta = \pi$ Hopf term.

B.8.4 Gauging small n non-linear sigma models

In appendix B.6, we have shown how to gauge the nonlinear sigma models. Recall that the non-trivial gauge coupling terms all originate from the WZW term. For small

n, the WZW term does not exist. One might think we need to re-derive the gauging procedure. Fortunately, we can use the mass manifold enlargement idea discussed in the preceding subsection to derive the coupling between $Q^{\mathbb{C},\mathbb{R}}$ and the gauge field. Without going into details we (1) add additional fermion flavors until the WZW term is stabilized. (2) Proceed as usual to gauge the continuous symmetries. (3) Restrict $Q^{\mathbb{C},\mathbb{R}}$ to the proper sub-mass manifold and the gauge group to the proper subgroup (so that the gauge field does not couple to the added fermion flavors). Following this procedure, we gauged the small n nonlinear sigma model following the same try-and-error method.

As an example, we shall write down the charge-U(1) gauged nonlinear sigma model for $n = 2$ in the $(2+1)$-D complex class. As shown in appendix B.8.3, the bosonized theory the S^2 nonlinear sigma model with the $\theta = \pi$ Hopf term. Plugging $Q^{\mathbb{C}} = n^a\sigma^a$ into the gauge coupling part in Eq.(B.69), we arrive at

$$W[\boldsymbol{\beta}, A] = \frac{1}{\lambda_3}\int\limits_{\mathcal{M}} d^3x\,(\partial_\mu \hat{n})^2 + i\pi H(\hat{n}) + \int\limits_{\mathcal{M}} d^3x\left[iA_\mu\left(\frac{1}{8\pi}\epsilon^{abc}\epsilon^{\mu\nu\rho}n^a\partial_\nu n^b\partial_\rho n^c\right)\right],$$

where the last term makes the S^2 solitons carry $U(1)$ charge. For $Q^{\mathbb{C}}$ in the $l = 0$ and $l = 2$ component of te mass manifold, we have a constant configuration $Q^{\mathbb{C}} = \pm I \in \frac{U(2)}{U(2)\times U(0)}$. Plugging it into Eq.(B.69), we get

$$\pm\frac{i}{4\pi}\int\limits_{\mathcal{M}} d^3x\,\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho, \tag{B.93}$$

which gives $\sigma_{xy} = \pm 1$.

B.9 Massless fermions as the boundary of bulk topological insulators/superconductors

The idea behind our bosonization is to fluctuate the bosonic order parameters ($Q^{\mathbb{C}}$ or $Q^{\mathbb{R}}$) to restore the full emergent symmetries of the massless fermion theory. These order parameters are chosen so that when they are static, any $Q^{\mathbb{C},\mathbb{R}}(\boldsymbol{x})$ configuration will fully gap out the fermions. As shown in appendix B.2, a static $Q^{\mathbb{C},\mathbb{R}}(\boldsymbol{x})$ configuration breaks at least some of the emergent symmetries. Conversely, if the full emergent symmetries are unbroken the fermion spectrum should remain gapless. Putting it succinctly, the emergent symmetries protect the gapless fermions.

The above situation reminds us of the boundary gapless modes of SPTs. Therefore, it is natural to suspect that each of the gapless fermion theories can be realized at the boundary of certain emergent-symmetry-protected SPT. In this appendix, we show that this is indeed the case. Moreover, we shall construct the bulk SPT explicitly.

B.9.1 The $\mathbb{Z}$ classification

As discussed in Ref.[7], the classification of free fermionic SPTs can be determined by checking how many copies of the boundary theory can be "stacked" together before a symmetry allowed mass term emerges. For example, a $\mathbb{Z}_N$ classification implies, after stacking N copies of the massless fermion theory, a mass term can be found without breaking any of the protecting symmetry (here the emergent symmetries). In the following, we show that the emergent-symmetry-protected SPT has $\mathbb{Z}$ classification.

For the sake of generality, we shall use the the Majorana fermion representation, even for complex classes. N copies of the boundary theory is described by the gamma matrices and the matrices that execute symmetry transformations,

$$\begin{aligned}\Gamma_i^{(N)} =& \Gamma_i \otimes I_N, \quad i = 1, ..., d \\ T^{(N)} =& t \otimes I_N \\ U^{(N)} =& u \otimes I_N\end{aligned}$$

Here t and u are orthogonal matrices obeying $\{t, \Gamma_i\} = [u, \Gamma_i] = 0$. The symbol t and u stand for anti-unitary and unitary, respectively. It is important to note that t and u represent the complete set of anti-unitary and unitary transformation matrices, from the product of which all symmetry matrices can be constructed. In addition, Γ_i, t, u stand for the gamma and symmetry matrices for one copy of the massless fermion theory.

Existence of a mass term for the stacked massless fermion theory, implies that there exist a matrix $M^{(N)}$ that anti-commutes with all of the gamma matrices. The general form of $M^{(N)}$ is

$$M^{(N)} = m_s \otimes A_N + m_a \otimes S_N$$

where m_s, S_N and m_a, A_N are symmetric and anti-symmetric matrices, respectively. Since $\Gamma_i^{(N)} = \Gamma_i \otimes I_N,\ T^{(N)} = t \otimes I_N,\ U^{(N)} = u \otimes I_N$ it follows that

$$\{m_{s,a}, \Gamma_i\} = \{m_{s,a}, t\} = [m_{s,a}, u] = 0$$

If such a $m_a \neq 0$ exists, we can use it as the mass term for the original massless fermion theory. This contradicts the statement that under the protection of emergent symmetry there is no mass term. Thus $m_a = 0$ and $M^{(N)}$ reduces to

$$M^{(N)} = m_s \otimes A_N. \tag{B.94}$$

On the other hand, m_s can then be used to construct an anti-unitary symmetry. By our assumption, such anti-unitary symmetry matrix m_s must be the product t's and u's. Thus the matrix

$$T^{(N)\prime} = m_s \otimes I_N$$

is an anti-unitary symmetry matrix of the stacked fermion theory. However such $T^{(N)\prime}$ commutes with Eq.(B.94) which is a contradiction. (Recall that in Majorana fermion representation, a mass matrix must anti-commute with all anti-unitary symmetry matrices.) Therefore $M^{(N)}$ can not exist for any N. Consequently, the classification of the massless fermion theory must be $\mathbb{Z}$.

B.9.2 Construction of the bulk SPT

To construct the bulk SPT, we follow the "holographic construction" in Ref.[8]. In the following, we just summarize the results.

For a massless fermion theory, with gamma matrices $\{\Gamma_i | \, i = 1, 2, ..., d\}$, anti-unitary symmetry t, and unitary symmetries $\{u\}$, we can construct the bulk matrices,

$$\begin{aligned}\Gamma_i^{(\mathrm{bk})} =& \begin{cases}\Gamma_i \otimes Z & \text{for } i = 1, ..., d \\ I_{\mathrm{dim}(\Gamma_\mathrm{i})} \otimes X & \text{for } i = d+1\end{cases} \\ T^{(\mathrm{bk})} =& t \otimes Z \\ U^{(\mathrm{bk})} =& u \otimes I\end{aligned}$$

Here the label (bk) is for distinguishing the bulk from the boundary matrices. In [8], it's shown that as long as the boundary massless fermion is irreducible, and t,u prohibit any mass term, then there is single allowed bulk mass term which respects all the symmetries[7]. Such mass term is given by

$$M^{(\mathrm{bk})} = I_{\mathrm{dim}(\Gamma_\mathrm{i})} \otimes Y.$$

The above mass term can be used to regularize and gap out the fermion in the bulk. In Wilson's regularization, the SPT (single-particle) Hamiltonian in momentum space is given by

$$h^{(\mathrm{bk})}(\boldsymbol{k}) = \sum_{i=1}^{d+1} \sin k_i \, \Gamma_i^{(\mathrm{bk})} + \left(d + 1 + m_B - \sum_{i=1}^{d+1} \cos k_i\right) M^{(\mathrm{bk})}$$

When $m_B < 0$, and when the lattice is cut open in the $(d+1)$th direction (actually the gapless boundary modes exist when the cut is along any direction), the boundary low energy theory is that of the original massless fermions.

B.10 The decoupling of the charge-SU(2) gauge field from the low energy non-linear sigma model after confinement

[7]Here the irreducibility means the gamma and the symmetry matrices cannot be simultaneously block-diagonalized non-trivially. The fact that this is true for our case is because the inclusion of the full emergent symmetries. (Proof omitted.)

In this appendix, we show that the charge-SU(2) gauge field is not coupled to Eq.(5.17). To recap, the charge-SU(2) singlet $Q^{\mathbb{R}}$ is given by

$$Q^{\mathbb{R}} = n_i N_i \quad \text{where}$$
$$N_i = (YXY, IYY, YZY, IIX, IIZ)\,.$$

Following appendix B.6 after gauging the charge-SU(2) symmetry, the $\frac{O(8)}{O(4)\times O(4)}$ nonlinear sigma model with $k = 1$ WZW term becomes

$$\begin{aligned} W[Q^{\mathbb{R}}, a] =& \frac{1}{4\lambda_3}\int_{\mathcal{M}} d^3x \, \mathrm{tr}\Big[\Big(\partial_\mu Q^{\mathbb{R}} + i[a_\mu, Q^{\mathbb{R}}]\Big)^2\Big] - \frac{2\pi i}{512\pi^2}\Big\{\int_{\mathcal{B}} \mathrm{tr}\Big[\tilde{Q}^{\mathbb{R}}\,\Big(d\tilde{Q}^{\mathbb{R}}\Big)^4\Big] \\ &+8\int_{\mathcal{M}} \mathrm{tr}\Big[iaQ^{\mathbb{R}}(dQ^{\mathbb{R}})^2 - (aQ^{\mathbb{R}})^2 dQ^{\mathbb{R}} - \frac{i}{3}(aQ^{\mathbb{R}})^3 + ia^3 Q^{\mathbb{R}} - aQ^{\mathbb{R}} f - afQ^{\mathbb{R}}\Big]\Big\}. \end{aligned} \tag{B.95}$$

Since all N_i commute with the charge-SU(2) group, it follows that $Q^{\mathbb{R}} = n_i N_i$ commutes with charge-SU(2) gauge field a. Hence the gauge coupling term in the stiffness term vanishes.

To show this is also true for the gauged WZW term part, we shall take the

$$\mathrm{tr}\left[aQ^{\mathbb{R}}(dQ^{\mathbb{R}})^2\right]$$

term in Eq.(B.95) as an example. Plugging in $Q^{\mathbb{R}} = n_i N_i$, we obtain

$$\mathrm{tr}\left[aQ^{\mathbb{R}}(dQ^{\mathbb{R}})^2\right] = \sum_{i,j,k} \mathrm{tr}[a\, N_i N_j N_k] n_i\, dn_j\, dn_k.$$

In the following we shall prove that each term in the sum vanishes, i.e.,

$$\mathrm{tr}[a\, N_i\, N_j\, N_k] = 0 \qquad \forall (i,j,k).$$

To achieve that we insert a $N_l^2 = 1$ where $l \neq i,j,k$ into the trace and leave it invariant, i.e.,

$$\mathrm{tr}[aN_iN_jN_k] = \mathrm{tr}[N_l^2 aN_iN_jN_k].$$

Due to the commutivity between N_l and a and the anti-commutivity between N_l and each of the $N_{i,j,k}$, we can move one N_l all the way to the right end and use the cyclic property of trace to put it back to the front

$$\mathrm{tr}[N_l^2 aN_iN_jN_k] = -\mathrm{tr}[N_l aN_iN_jN_kN_l] = -\mathrm{tr}[N_lN_l aN_iN_jN_k] = -\mathrm{tr}[aN_iN_jN_k].$$

Thus

$$\mathrm{tr}[aN_iN_jN_k] = -\mathrm{tr}[aN_iN_jN_k] \Rightarrow \mathrm{tr}[aN_iN_jN_k] = 0.$$

This proof can be applied to all gauge coupling terms in Eq.(B.95) because there is an odd number of $Q^{\mathbb{R}}$s for every term that couples to the gauge field. Therefore the charge-SU(2) gauge field is not coupled to $Q^{\mathbb{R}} = n_i N_i$.

B.11 The WZW term in the $(3+1)$-D real class non-linear sigma model

In this section, we will show that upon the decomposition in Eq.(5.32) of subsection 5.2.3, namely,

$$Q^{\mathbb{R}}(x) = e^{i\theta(x)}\mathcal{G}_S(x),$$

the contribution of the WZW term is solely from the $\mathcal{G}_S(x)$ part, i.e., namely

$$\mathrm{tr}\left[\left(Q^{\mathbb{R}\dagger}dQ^{\mathbb{R}}\right)^5\right] = \mathrm{tr}\left[\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^5\right]$$

First, note that one can at most choose $d\theta$ once in the expansion of $\mathrm{tr}\left[\left(Q^{\mathbb{R}\dagger}dQ^{\mathbb{R}}\right)^5\right] = \mathrm{tr}\left[(\mathcal{G}_S^\dagger d\mathcal{G}_S + id\theta)^5\right]$, otherwise the differential form vanishes because $(d\theta)^2 = 0$. The only term that can possibly survive other than $\mathrm{tr}\left[(\mathcal{G}_S^\dagger d\mathcal{G}_S)^5\right]$ is then of the form

$$\mathrm{tr}\left[d\theta\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^4\right] = d\theta\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^a\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^b\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^c\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^d \mathrm{tr}\left[t^a t^b t^c t^d\right]$$

Here $\{t^a\}$ are the complete basis for the generators of $SU(n)$ in the fundamental representation (note that $\mathcal{G}_S$ are the symmetric special unitary matrices, which are special kind of unitary matrices). In the following, we will show that for every term from the trace $\mathrm{tr}\left[t^a t^b t^c t^d\right]$, it is at least symmetric with respect to two of the indices in a, b, c, d. If so, because the scalar valued one forms $\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^a$ anti-commute with each others $\mathrm{tr}\left[d\theta\left(\mathcal{G}_S^\dagger d\mathcal{G}_S\right)^4\right]$ vanishes.

We shall choose the conventions

$$\mathrm{tr}\left[t^a t^a\right] = \frac{1}{2}\delta_{ab} \tag{B.96}$$

$$[t^a, t^b] = i f_{abc} t^c \tag{B.97}$$

where f_{abc} is the structure constant for $SU(n)$. Here the Einstein summation convention is used. f_{abc} is real and totally anti-symmetric. We shall also define

$$d_{abc} = 2\,\mathrm{tr}\left[\{t^a, t^b\}t^c\right] \tag{B.98}$$

It can be shown simply that due to the cyclic property of trace and the hermiticity of t^a, d_{abc} is real and totally symmetric with respect to a, b, c.

As a pre-step, we would calculate $t^a t^b$. Because the identity matrix I_n together with $\{t^a\}$ form a complete basis for all $n \times n$ complex matrices, we can decompose $t^a t^b$ in terms of them. The coefficients can be calculated making use of Eq.(B.96), Eq.(B.97), and Eq.(B.98),

$$\begin{aligned}
t^a t^b =& \frac{1}{n}\mathrm{tr}\left[t^a t^b\right] I_n + \mathrm{tr}\left[\left(\{t^a, t^b\} + [t^a, t^b]\right) t^c\right] t^c \\
=& \frac{1}{2}\left[\frac{1}{n}\delta_{ab} I_n + (d_{abc} + i f_{abc}) t^c\right].
\end{aligned} \tag{B.99}$$

The equation above implies

$$\{t^a, t^b\} = \frac{1}{n}\delta_{ab} I_n + d_{abc} t^c \tag{B.100}$$

For later usage, we will derive another formula for the product of two f_{abc}s. By direct expansion, one can prove the following identity

$$[t^a, [t^b, t^c]] = \{\{t^a, t^b\}, t^c\} - \{\{t^a, t^c\}, t^b\}.$$

Using of Eq.(B.97) and Eq.(B.100) twice in the equation above, we get

$$f_{abe} f_{cde} = \frac{2}{n}\left(\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}\right) + \left(d_{ace} d_{bde} - d_{ade} d_{bce}\right) \tag{B.101}$$

Now we can calculate $\mathrm{tr}\left[t^a t^b t^c t^d\right]$ by applying Eq.(B.99) twice and carrying out the trace. After some algebra and the help of Eq.(B.101), we get

$$\begin{aligned}
\mathrm{tr}\left[t^a t^b t^c t^d\right] =& \frac{1}{4}\mathrm{tr}\left[\left(\frac{1}{n}\delta_{ab} I_n + (d_{abc} + i f_{abe}) t^e\right)\left(\frac{1}{n}\delta_{cd} I_n + (d_{cdf} + i f_{cdf}) t^f\right)\right] \\
=& \frac{1}{4}\left[\begin{matrix} \frac{1}{n}(\delta_{ab}\delta_{cd} - \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}) \\ + \frac{1}{2}(d_{abe} d_{cde} - d_{ace} d_{bde} + d_{ade} d_{bce}) \\ + \frac{i}{2}(f_{abe} d_{cde} + f_{cde} d_{abe}) \end{matrix}\right]
\end{aligned}$$

By the symmetry properties of δ_{ab} and d_{abc}, every term is least symmetric with respect to two indices. For example, $f_{abe} d_{cde}$ is symmetric with respect to c, d. This concludes our proof.