

Berry phase in quantum field theory: Diabolical points and boundary phenomenaPo-Shen Hsin,¹ Anton Kapustin,¹ and Ryan Thorngren²¹Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA²Center of Mathematical Sciences and Applications, Harvard University, Cambridge, Massachusetts 02138, USA

(Received 6 July 2020; accepted 11 November 2020; published 9 December 2020)

We study aspects of the Berry phase in gapped many-body quantum systems by means of effective field theory. Once the parameters are promoted to space-time-dependent background fields, such adiabatic phases are described by Wess-Zumino-Witten (WZW) and similar terms. In the presence of symmetries, there are also quantized invariants capturing generalized Thouless pumps. Consideration of these terms provides constraints on the phase diagram of many-body systems, implying the existence of gapless points in the phase diagram, which are stable for topological reasons. We describe such diabolical points, realized by free fermions and gauge theories in various dimensions, which act as sources of “higher Berry curvature” and are protected by the quantization of the corresponding WZW terms or Thouless pump terms. These are analogous to Weyl nodes in a semimetal band structure. We argue that in the presence of a boundary, there are boundary diabolical points—parameter values where the boundary gap closes—which occupy arcs ending at the bulk diabolical points. Thus the boundary has an “anomaly in the space of couplings” in the sense of [C. Cordova, D. S. Freed, H. T. Lam, and N. Seiberg, *SciPost Phys.* **8**, 001 (2020) and **8**, 002 (2020)]. Consideration of the topological effective action for the parameters also provides some new checks on conjectured infrared dualities and deconfined quantum criticality in 2+1d.

DOI: [10.1103/PhysRevB.102.245113](https://doi.org/10.1103/PhysRevB.102.245113)**I. INTRODUCTION**

In this paper, we explore topological invariants associated with the global structure of phase diagrams of quantum systems at zero temperature and their physical consequences. An important source of topology in phase diagrams is the presence of so-called “diabolical points” [1] (or more general “diabolical loci”)—isolated parameter values where the ground state differs from the surrounding phase, for example by being degenerate or gapless.

A familiar example is a stationary spin-1/2 in a magnetic field. At zero magnetic field, the energy spectrum is doubly degenerate, while at any nonzero magnetic field there is a unique ground state. One has to tune three real parameters, namely the components of the magnetic field, to encounter this degeneracy. A classic result of von Neumann and Wigner [2] states that this is a generic situation: in quantum systems with a finite number of degrees of freedom and without symmetries¹ energy level degeneracy generically occurs in codimension 3.

Another classic observation, due to M. V. Berry [3], is that such a “diabolical locus” \mathcal{D} can be detected by studying the ground-state wave function on the complement of \mathcal{D} . Consider the case when the parameter space is three-dimensional and \mathcal{D} is a single point. Removing \mathcal{D} from its small neighborhood

we end up with a space which is homotopy equivalent to a two-dimensional sphere. One can show that because of a nonintegrable Berry phase a properly normalized ground-state wave function depending continuously on parameters can not be defined globally on this 2-sphere. The best one can do is to let the ground-state be a normalized continuous section of a nontrivial line bundle over the 2-sphere [4]. A manifestation of this is that the curvature of the Berry connection has a monopole singularity at \mathcal{D} .

In many-body quantum systems, degeneracy of the ground state generically occurs in codimension 1 rather than 3. These are loci where the system undergoes a first-order phase transition. The von Neumann-Wigner result does not apply to such situations, and the degeneracy locus does not leave a topological imprint. Nevertheless, many-body analogues of diabolical points do exist. These are gapless points or gapless loci in the phase diagram which are entirely surrounded by the trivial phase (although they can also occur inside other phases). Such gapless points are not simply endpoints of first-order transitions, they are robust for topological reasons. The simplest example is a free Dirac fermion in spatial dimension d with charge conservation. Depending on whether d is even or odd, the gapless point in this system occurs in codimension 1 or 2. Usually its stability is explained using anomalies (parity anomaly for even d and axial anomaly for odd d). However, there also exist examples of isolated gapless points which cannot be explained using anomalies. We propose a framework which treats all these examples on the same footing.

A natural generalization of the Berry phase to $d > 0$ is the Wess-Zumino-Witten (WZW) term in the effective action

¹When a unitary symmetry is present, energy levels which transform in different representations of the symmetry group generically cross in codimension 1, since there is no level repulsion in this case.

for the parameters. A WZW term in spatial dimension d is described by a closed $(d+2)$ -form on the parameter space which may have monopolelike singularities along gapless loci. Such singularities occur in codimension $d+3$, therefore one might conclude that diabolical points in d dimensions occur in codimension $d+3$. In fact, the situation is more complex, and the WZW form describes just a particular class of topological terms in the effective action. In general, there are several types of diabolical points in a given dimension, and their classification also depends on whether the system is bosonic or fermionic. Imposing symmetry also affects the classification of diabolical points. This was pointed out already by von Neumann and Wigner who showed that in the presence of time-reversal symmetry diabolical points in 0+1d systems occur in codimension 2 [2]. In any case, using the effective action approach one can classify possible diabolical loci and prove their stability.

Another point of view on this problem, which is very powerful, comes from considering the infinite-dimensional space of all field theories, i.e., UV-complete theories defined at some UV scale. This space is decomposed into strata defined by the attractive basins of the RG flow, consisting of field theories which flow to a particular RG fixed point in the IR. Roughly speaking, the number of relevant operators at the fixed point gives the codimension of the stratum.² A generic phase diagram with k parameters is a k -dimensional slice through this infinite dimensional space, transverse to all the strata, meaning a stratum with k relevant operators at the fixed point occurs at isolated points in this k -parameter phase diagram. Protected diabolical loci correspond to the situation where one stratum pokes a hole in another, leaving the latter with some nontrivial topology. This topology, in turn allows for the universal definition of topological terms in the effective action, associated with a certain cohomology of this stratum. Thus the higher Berry phase and protected diabolical loci points of view are equivalent.

An important novelty of working in higher dimensions is the possibility of studying spatial boundaries. We will describe a boundary-bulk correspondence for these “higher” Berry phases, which is reminiscent of the Fermi arcs in Weyl semimetals [6]. Indeed, we observe that relative to some choice of boundary condition, the higher Berry phase implies the existence of a diabolical locus in parameter space where the boundary gap closes, whose boundary in parameter space is in the bulk diabolical locus. In the context of lattice models, this observation was recently made in Refs. [7,8]. We also make contact with the “anomalies in the space of coupling constants” studied recently by Refs. [9,10].

Many of the physical phenomena associated with higher Berry phase are well-known and can be exhibited by free fermion systems. See Ref. [11] for a general approach in this case. A particularly useful collection of such models was defined by Abanov and Wiegmann [12]. As we review in Appendix C 1, they give rise to S^{d+2} level 1 WZW terms and S^d generalized Thouless pumps. We will discuss several

of these theories in some detail, including their boundary phenomena, as well as a free fermion model in 3+1d “missing” from Ref. [12], which is associated with the $p+ip$ superconductor/gravitational Chern-Simons term.

While free fermion theories provide many examples of diabolical points, we stress that our arguments prove their stability on the nonperturbative level. We also discuss some interacting models (gauge theories). The focus of this paper is the long-distance behavior which can be adequately treated using field theory. However, the same phenomena occur in lattice models. Some examples of such models have been constructed in Refs. [7,8], but there are many other examples, some of them quite classical [13].

The paper is organized as follows. In Sec. II, we discuss free fermion diabolical points and their corresponding families in 1+1d. There is a codimension-2 diabolical point protected by the Thouless pump and a codimension-4 diabolical point protected by a higher Berry number. We discuss the physics of interfaces and boundaries for both families.

In Sec. III, we extend the discussion to free fermion diabolical points and their families in 2+1d. There are again two types, corresponding to a family with a charged skyrmion, analogous to the Thouless pump in 1+1d, and a family with a higher Berry number, respectively.

In Sec. IV, we discuss free fermion diabolical points and their families in 3+1d. There is a new codimension-2 diabolical point in this dimension that generates an axion coupling and corresponds to a Thouless pump of a gravitational Chern-Simons term.

In Sec. V, we discuss interacting examples of diabolical points and their families in gauge theories. This includes 2+1d Chern-Simons matter theories realizing the generalized Thouless pump. The nontrivial families are consistent with the recently studied web of 2+1d IR dualities, and we interpret the deconfined quantum critical point between the Néel state and the valence bond solid (VBS) as a protected diabolical point. We also discuss applications to scalar QED in 3+1d.

In Sec. VI, we provide a general proof of the bulk-boundary correspondence for the higher Berry phase and Thouless pumps. The result describes an index theorem whereby there will be topologically protected boundary diabolical points, the sum of whose higher Berry numbers equals the bulk higher Berry number.

In Sec. VII, we discuss some mathematical formalism for describing families of gapped phases, including a classification based on cobordism theory and TQFT.

II. FREE FERMIONS IN 1+1d

A. The Thouless pump and a codimension-2 diabolical point

1. A topologically nontrivial family

As a motivating example, consider a single complex fermion in 1+1d with a Minkowski-signature Lagrangian

$$\mathcal{L}_{\text{Dirac}} = -\bar{\psi}i\partial\psi - i\bar{\psi}(M_1 + iM_2\gamma^0)\psi. \quad (2.1)$$

Here $\partial = \gamma^\mu\partial_\mu$, γ^0 and γ^1 are 2d Dirac matrices satisfying $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, $g^{\mu\nu} = \text{diag}(-1, 1)$ is the Minkowski metric,

²This is complicated by theories with marginally relevant or dangerously irrelevant operators, which demonstrate that this stratification is quite singular [5].

and $\gamma^{01} = \gamma^0\gamma^1$ is the chirality operator.³ Some unusual-looking signs and factors of i in (2.1) and similar formulas below are due to our use of “mostly plus” convention for the Minkowski metric. Our conventions are chosen to facilitate comparison with Ref. [12].

Hermiticity of the Hamiltonian requires M_1 and M_2 to be real, and it is convenient to introduce a complex mass $M = M_1 + iM_2$. We will also denote $\alpha = \arg M$. This Lagrangian has a U(1) symmetry under which ψ has charge 1 and $\bar{\psi} = \psi^\dagger\gamma^0$ has charge -1 . The axial U(1) transformation $\psi \mapsto e^{i\alpha\gamma^{01}}\psi$ is not a symmetry except for $M = 0$, since it has the effect of replacing M with $Me^{2i\alpha}$. For $M \neq 0$, this model describes massive fermionic particles and antiparticles of mass $|M|$, while for $M = 0$ it is gapless. Its properties for spatially varying M have been studied by Goldstone and Wilczek [14]. It is also a special case of the “A” series of Abanov and Wiegmann [12].

In this model, the massless point is an isolated gapless point of codimension 2. One may ask if it is possible to gap it out by adding interactions or couplings to other massive degrees of freedom, without breaking U(1) or bringing in new gapless points from infinity. The answer is no, and one way to see it is to couple the theory for $|M| > 0$ to a background U(1) gauge field and study the effective action after we integrate out the fermion. We will find a topological term whose presence is inconsistent with a completely gapped phase diagram.

From the axial anomaly, one finds that this effective action contains a term

$$\frac{1}{2\pi} \int \alpha \epsilon^{\mu\nu} F_{\mu\nu} d^2x = \frac{1}{2\pi} \int \alpha F, \tag{2.2}$$

where $F = dA = \frac{1}{2}F_{\mu\nu}dx^\mu dx^\nu$ is the gauge field strength. This term does not seem to have much effect if we regard A as a background gauge field and if M is a constant. But as noticed by Goldstone and Wilczek [14], the situation changes if we allow M to be a slowly varying function of coordinates. The above action is then the leading term in the derivative expansion in M and gives rise to a topological contribution to the U(1) current:

$$J_{\text{top}}^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu} \partial_\nu \alpha. \tag{2.3}$$

This U(1) current represents the response of the fermionic vacuum to an adiabatic variation of M . If M depends only on x^1 , it gives a U(1) charge density to the vacuum. The net charge of a configuration where $\alpha = \arg M$ approaches α_\pm as $x \rightarrow \pm\infty$ and slowly varies in between is $(\alpha_+ - \alpha_-)/2\pi + n$, where n is the (signed) number of times M crosses the branch cut for the logarithm. In general the charge is fractional. If $\alpha_+ = \alpha_-$, this charge is just n and can be interpreted as the collective charge of n winding solitons in α [14]. If M depends only on x^0 , there is a uniform vacuum current. In particular, if M is a periodic function of time and winds around the origin N times per period, the net charge which flows through a section of a system during this time is $-N$. This phenomenon is known as the Thouless charge pump [15].

The presence of this topological term implies that we cannot completely remove the gapless point at the origin in a U(1)-invariant way by adding UV degrees of freedom and continuously changing their couplings to the field ψ (although we can deform it into a more general diabolical locus, see Appendix B). This is because for a gapped nondegenerate family on \mathbb{R}^2 , we would need to extend the form $\epsilon^{\mu\nu} \partial_\nu \alpha$, required to express the topological current, to maps M valued anywhere in the plane, which is impossible.

To formalize this argument, we interpret the loop of models defined by $|M| = m > 0$ as a noncontractible loop in the space $\mathfrak{M}_1^{U(1)}$ of 1+1d field theories with a unique gapped ground state (i.e., infrared-trivial theories) and U(1) symmetry.

Consider a general such theory defined over a parameter space P . The low energy limit defines a map from our parameter space $P \rightarrow \mathfrak{M}_1^{U(1)}$. Slowly varying parameters in space-time thus define a map $\phi : X \rightarrow \mathfrak{M}_1^{U(1)}$. We can define an effective action for ϕ . Expanding in derivatives of ϕ , the leading term has the form

$$S_{\text{top}} = \int_X \epsilon^{\mu\nu} A_\mu \partial_\nu \phi^i \tau_{1,i}(\phi) d^2x + \dots = \int_X A \wedge \phi^* \tau_1 + \dots, \tag{2.4}$$

where $\tau_1 = \tau_{1,i}(\phi) d\phi^i$ is a 1-form on $\mathfrak{M}_1^{U(1)}$ (that is, it is universal) and dots denote terms which are independent of both A and the metric. Invariance with respect to infinitesimal gauge transformations $A \mapsto A + df$ requires the 1-form τ_1 to be closed. Invariance with respect to large gauge transformations requires it to have integral periods. That is, for any loop $\phi : S^1 \rightarrow \mathfrak{M}_1^{U(1)}$ one must have $\oint \phi^* \tau_1 \in \mathbb{Z}$. This integer is associated with the Thouless pump of the loop. Indeed, the topological current corresponding to the above action is obtained by varying with respect to A_μ :

$$J_{\text{top}}^\mu = \epsilon^{\mu\nu} \tau_{1,i}(\phi) \partial_\nu \phi^i. \tag{2.5}$$

This current leads to the same physical consequences as the current (2.3). Namely, static configurations of the field ϕ acquire U(1) charge, while a time-dependent ϕ produces a spatial current. Given a loop $\phi : S^1 \rightarrow \mathfrak{M}_1^{U(1)}$ with a base point $\phi_0 \in \mathfrak{M}_1^{U(1)}$, one can attach to it a solitonic configuration on \mathbb{R} which approaches ϕ_0 at $x = \pm\infty$. The charge of this soliton is equal to $\oint \tau_{1,i} d\phi^i$ and is integral. We can also identify S^1 with the time coordinate thus making ϕ a periodic function of time. The net charge transported per period is $\Delta Q = -\oint \tau_{1,i} d\phi^i$.

If a given loop in $\mathfrak{M}_1^{U(1)}$ is contractible, then the integral of τ_1 over this loop vanishes. Mathematically, this follows from the Stokes theorem. The physics of this is also clear: if a slowly varying field configuration on \mathbb{R} can be deformed to the constant one without changing the asymptotics at $x = \pm\infty$ and without creating large gradients, its charge must be zero.⁴ In the case of the complex fermion in 1+1d, it follows from Eq. (2.3) that $\tau_1 = \frac{1}{2\pi} d\alpha$. Therefore the integral of τ_1 over the loop in the M -plane circling the origin is 1, and thus this loop is not contractible. Any nonzero multiple of this

³For definiteness, in terms of the usual Pauli spin matrices, one can choose $\gamma^0 = i\sigma^y$, $\gamma^1 = \sigma^x$, and $\gamma^{01} = \sigma^z$.

⁴If the gradients become large, there may be level crossing, and the vacuum charge may jump discontinuously.

loop is also not contractible in $\mathfrak{M}_1^{U(1)}$. Thus the fundamental group of $\mathfrak{M}_1^{U(1)}$ contains a copy of \mathbb{Z} .⁵ A more precise way to phrase it is this. Let $\mathfrak{M}_1^{\text{Dirac}}$ denote the subspace of $\mathfrak{M}_1^{U(1)}$ corresponding to the family of theories (2.1) with $M \neq 0$. Obviously, $\mathfrak{M}_1^{\text{Dirac}}$ is the M plane with the origin removed and thus $\pi_1(\mathfrak{M}_1^{\text{Dirac}}) = \mathbb{Z}$. The above argument shows that the map $\pi_1(\mathfrak{M}_1^{\text{Dirac}}) \rightarrow \pi_1(\mathfrak{M}_1^{U(1)})$ arising from the embedding $\mathfrak{M}_1^{\text{Dirac}} \rightarrow \mathfrak{M}_1^{U(1)}$ is injective. This implies that the isolated gapless point at $M = 0$ cannot be gotten rid by arbitrary deformations which preserve $U(1)$ symmetry.

Although the diabolical point cannot be completely removed, the nature of the diabolical locus can be modified by deforming the theory for small $|M|$. For example, one can deform the free fermion theory into a more general Luttinger liquid. Depending on how this is done, one can “resolve” the diabolical point into a first-order phase transition line terminating at two critical Ising points, or into an island of a gapless Berezinsky-Kosterlitz-Thouless phase. This is described in more detail in Appendix B.

2. Interfaces

A physical consequence of noncontractibility can be seen when one studies smooth interfaces between different models in the family. For the Dirac fermion (2.1), we will study interfaces where the mass parameter $M = me^{i\alpha}$, $m = |M|$, varies from m at $x = -\infty$ to some other value $me^{i\alpha_0}$ at $x = \infty$. In particular, we will be interested in one-parameter families of interfaces parametrized by $\alpha_0 \in \mathbb{R}/2\pi\mathbb{Z}$. We will show that for at least one of these interfaces, there must be a localized zero mode.

The simplest option is to take m constant and let $\alpha(x)$ be a step function, $\alpha(x) = \alpha_0\theta(x)$. This gives a family of interfaces which is periodic as a function of α_0 but is not smooth in x space. While this is not quite what we need, the advantage of this family of interfaces is that it is soluble [16] and one can find the states localized on the interface exactly. One lets $\psi = \psi(0)e^{-ax}$ for $x > 0$ and $\psi = \psi(0)e^{ax}$ for $x < 0$, with some $a > 0$. Then $\psi(0)$ satisfies

$$\begin{pmatrix} -ia & me^{i\alpha_0} \\ me^{-i\alpha_0} & ia \end{pmatrix} \psi(0) = E\psi(0), \\ \begin{pmatrix} ia & m \\ m & -ia \end{pmatrix} \psi(0) = E\psi(0). \quad (2.6)$$

This gives

$$a = m \sin \frac{\alpha_0}{2}, \quad E = m \cos \frac{\alpha_0}{2}. \quad (2.7)$$

For $\alpha_0 \in (0, 2\pi)$, there is a single mode localized on the interface. Its energy becomes zero at $\alpha_0 = \pi$. The mode becomes nonnormalizable and merges with the continuum for $\alpha_0 = 0$ (or $\alpha_0 = 2\pi$).

⁵According to the classification we discuss in Sec. VII, $\pi_1(\mathfrak{M}_1^{U(1)}) = \mathbb{Z}$. Thus τ_1 is a complete invariant describing the family: two families with the same $\oint \phi^* \tau_1$ may be deformed into each other without closing the gap.

Another soluble case is an interface where $\alpha(x)$ is continuous, varies linearly with x in an interval $0 < x < L$, and is equal to its asymptotic values 0 and α_0 for $x \leq 0$ and $x \geq L$, respectively. This interface was analyzed in Ref. [17]. One finds that for small and positive α_0 there is a single mode on the interface whose energy is just below m and decreases monotonically as α_0 increases. The energy becomes zero for α_0 of order mL . At certain threshold values of α_0 additional interface modes appear. This family of interfaces is not periodic in α_0 , i.e., the interface at $\alpha_0 = 2\pi$ is not identical to the interface at $\alpha_0 = 0$.

To get a family of interfaces which is smooth and periodic in α_0 , one needs to make both $|M|$ and $\alpha = \arg M$ dependent on x . For example, one can take $M(x)$ to take values in a straight-line segment connecting m and $me^{i\alpha_0}$. This family of interfaces does not appear to be soluble. Nevertheless, its qualitative behavior is easy to understand. For $\alpha_0 \ll 2\pi$ this family of interfaces coincides with that studied in Ref. [17], up to terms which are quadratic in α_0 . Thus there is a single normalizable mode on the interface whose energy is slightly below m and approaches m in the limit $\alpha_0 \rightarrow 0$.

As one increases α_0 , additional interface modes may appear. Every time such a mode crosses zero energy from above (respectively from below), the ground-state charge for the effective quantum mechanics describing the interface jumps by $+1$ (respectively -1). It is easy to show that the algebraic number of such crossings (i.e., the number of crossings from above minus the number of crossings from below) as α_0 increases from 0 to 2π is either 1 or -1 . Indeed, charge conjugation symmetry tells us that each crossing at $\alpha_0 = b$ is accompanied by a crossing at $\alpha_0 = 2\pi - b$, and their contributions are opposite.

The only unpaired crossing may occur at $\alpha_0 = \pi$. This situation corresponds to $M(x)$ being real for all x and varying from m to $-m$. It was shown by Jackiw and Rebbi [18] that there is exactly one normalizable zero mode in this case. Thus the algebraic number of crossings is ± 1 . One can verify that it is 1 by taking the limit $L \rightarrow 0$. Then we get an infinitely thin interface studied by MacKenzie and Wilczek [16], and we know that in that case the zero energy level is crossed from above.

To understand the fate of general interfaces, even those which do not come simply as a spatially varying mass profile, we can appeal to the topology of $\mathfrak{M}_1^{U(1)}$. An infrared-trivial interface can be thought of as a path in $\mathfrak{M}_1^{U(1)}$. Let us fix the theory at $x = -\infty$ (the start of the path) to be the Dirac fermion with $M = m > 0$, and we define the space $\mathcal{P}_m \mathfrak{M}_1^{U(1)}$ of paths in $\mathfrak{M}_1^{U(1)}$ beginning at this point. There is a map $r : \mathcal{P}_m \mathfrak{M}_1^{U(1)} \rightarrow \mathfrak{M}_1^{U(1)}$ which is given by the theory at large x . A periodic, smooth 1-parameter family of infrared-trivial interfaces between the Dirac fermion with $M = m > 0$ and the Dirac fermion with $M = me^{i\alpha_0}$ at $x = \infty$ is the same thing as a map $f : S^1_{\alpha_0} \rightarrow \mathcal{P}_m \mathfrak{M}_1^{U(1)}$, such that $r \circ f : S^1_{\alpha_0} \rightarrow \mathfrak{M}_1^{U(1)}$ is our Dirac family with a spatially constant mass $M = me^{i\alpha_0}$. However, this is impossible, because that family is noncontractible in $\mathfrak{M}_1^{U(1)}$, but $\mathcal{P}_m \mathfrak{M}_1^{U(1)}$ is a contractible space: all paths can be continuously retracted to the constant path at m . This is a contradiction, so for any periodic smooth family of

interfaces with these end points, there must be at least one interface with a degenerate ground state.

3. Boundary-bulk correspondence

We may regard such behavior as an anomaly in the space of couplings for the interface, in the sense of Ref. [9]. Imagine integrating out all the bulk modes, leaving only the modes bound to the interface. This gives an effective 0+1d field theory (that is, quantum mechanics) with a U(1) symmetry. For a generic value of α_0 the ground state of this 0+1d field theory is unique and has a well-defined charge. The only exceptions are points in the parameter space where the energy of some modes becomes zero. Every time the energy of a mode crosses zero from above (respectively from below), the ground-state charge increases by 1 (respectively decreases by 1). In view of the above discussion, as one increases α_0 from 0 to 2π the net change of the ground state charge is 1.

Such a situation would be impossible if there were a UV-complete effective quantum mechanics for the interface modes for all α_0 . Indeed, we would expect a well-defined effective action away from the gapless points. This effective action is

$$S_{\text{int}} = \int q(\alpha_0)A, \tag{2.8}$$

where $q : S^1 \rightarrow \mathbb{Z}$ is a locally constant continuous function (the vacuum charge) and A is a background U(1) gauge field restricted to the interface. The function q can change only at gapless points. Since q is single-valued, the sum of jumps of q across all gapless points is obviously zero. This means that the situation described in the previous paragraph cannot occur in a family of UV-complete 0+1d theories parameterized by S^1 . However, it can and does occur in the presence of a 1+1d bulk. The physical reason for this is that at certain points in the parameter space the interface modes become non-normalizable and merge with the bulk excitations. The function q can change discontinuously at these “delocalization” points as well as at gapless points. The sum of jumps over both types of special points vanishes.

One can quantify the anomaly of the boundary 0+1d field theory as the sum of jumps of q at the gapless points, or equivalently as minus the sum of jumps of q over “delocalization” points. The latter viewpoint makes apparent the relation of the anomaly with the Thouless pump invariant of the bulk system: The net charge which flows through the system as one cycles in the parameter space should be equal to the net charge of the interface modes which became delocalized and escaped to infinity. This gives us a “boundary-bulk correspondence”: The sum of jumps of $q(\alpha_0)$ over all gapless points is equal to the Thouless pump invariant of the bulk. This may also be reasoned from the topological terms (2.8) and (2.4). We will revisit this correspondence more generally in Sec. VI.

4. Reducing U(1) to fermion parity and torsion

It is clear from the above discussions that the nontriviality of this family rests on the conservation of U(1) charge. We can consider reducing the symmetry so that only the \mathbb{Z}_2

fermion parity is conserved.⁶ We expect that the Thouless pump invariants will all be reduced from \mathbb{Z} to \mathbb{Z}_2 . Formally, we regard our family (2.1) for some $|M| = m > 0$ as a loop in the space of arbitrary (fermionic) infrared-trivial 1+1d field theories \mathfrak{M}_1 . There is a map $\mathfrak{M}_1^{U(1)} \rightarrow \mathfrak{M}_1$ given by forgetting the U(1) symmetry. We will show the induced map on $\pi_1(\mathfrak{M}_1^{U(1)}) = \mathbb{Z} \rightarrow \pi_1(\mathfrak{M}_1)$ is reduction modulo 2.

We do this in two steps. First, let us argue that the family (2.1) remains nontrivial after breaking U(1) to \mathbb{Z}_2 . This is simply because by the spin-charge relation, our soliton which carries one unit of U(1) charge is fermionic. On the other hand, the constant “loop” (i.e., with constant M) corresponds to a state with an even fermion parity. This means we will continue to have a bulk-boundary correspondence, where the ground state of the boundary changes fermion parity after absorbing a soliton.

Next, we must show that a loop with winding number 2 is contractible in \mathfrak{M}_1 . One way to do it, from the perspective of the diabolical point, is to study two copies of (2.1), which we consider as four real fermions. One can find a U(1)-breaking mass term which completely gaps the theory. We will study this system in detail in Sec. II B.

Another way to do it, from the perspective of the gapped region, is to study a different parametrization of (2.1) over the circle $|M| = m > 0$, where our parameter winds around the origin twice. It does not make sense to extend such a parametrization to the disk $|M| \leq m$. Indeed, the diabolical point at the origin determines the winding number, so any extension to a disk must have a different diabolical locus inside. Doubling the winding number like this may look like a trivial redefinition of the system, but we will show that after breaking U(1) we can extend this family to a disk without closing the gap, i.e., without any diabolical locus.

To do so, let us regard the complex fermion as a pair of real (Majorana) fermions and add yet another massive Majorana fermion field. We can consider a more general family of free field theories describing three Majorana fermions:

$$\mathcal{L} = -\psi^a i \partial \psi^a - i M_{ab} \psi_+^a \psi_-^b. \tag{2.9}$$

Here ψ_+^a and ψ_-^a are right-handed and left-handed components of ψ^a . Hermiticity requires the mass matrix M_{ab} to be real. If we regard ψ_+^a and ψ_-^a as triplets of $O(3)_R$ and $O(3)_L$, then the mass matrix transforms as $M \mapsto O_R^T M O_L$. One can use these transformations to make M diagonal, with real positive entries. These entries determine the physical masses of excitations. The original one-parameter family of theories (2.1) is recovered when one of the physical masses becomes infinite. The rules of the game allow us to add a third Majorana fermion with a very large physical mass and then deform all the physical masses to be the same. Then the mass matrix has the general form $M = mO$, where $O \in O(3)$ and $m > 0$. The space of orthogonal matrices has two components, each of which is topologically equivalent to $\mathbb{R}P^3$. The original family is homotopic to a loop in one of these components. Since

⁶This system satisfies a spin charge relation, so this \mathbb{Z}_2 is a subgroup of U(1). Formally, this means that A is not a U(1) gauge field but a Spin^c structure.

$\pi_1(\mathbb{R}P^3) = \mathbb{Z}_2$, any loop becomes contractible when iterated twice. Thus our family over S^1 which winds twice may be extended to a family over a bounding disk, just using nonzero mass terms—that is, without closing the gap. This shows that loops with an even winding number become contractible when regarded as loops in \mathfrak{M}_1 .⁷

This illustrates a general feature of topological families with torsion invariants [as opposed to ones like the U(1) Thouless invariant, which is \mathbb{Z} valued]. One may say that if the family looks nontrivial for some parameter $\theta \in [0, 2\pi)$, but trivial if that family is defined for a $2\pi n$ periodic parameter, what is the physical difference? There is no physical difference of course, but to obtain the strongest conclusions from our framework one should use the “minimal” parametrization which is relevant to the problem. This is similar to how one may see SPT phases in a spin-1 system like the Haldane chain by studying SO(3) symmetry, which is natural if there are no half-integer spin degrees of freedom, but this SPT phase looks trivial if we consider the symmetry as SU(2). For families presented by an action principle, and especially for families defined near a diabolical locus, there is always a minimal choice of periodicity.

We comment that the usual explanation of the robustness of the gapless point at $M = 0$ involves a mixed 't Hooft anomaly between the vector and axial U(1) symmetry [or, if the U(1) symmetry is broken, an 't Hooft anomaly between fermion parity and axial U(1)]. However, this assumes that the mass term is the only term which breaks axial U(1). This assumption is not very natural, since axial U(1) cannot be a microscopic on-site symmetry for any M if we assume that at some UV scale the theory is described by a lattice Hamiltonian. Our argument does not rely on the axial U(1) symmetry.

While our discussion is phrased in the continuum, when the model (2.1) is discretized, it is equivalent to the Rice-Mele model [13], which describes noninteracting fermionic particles in a sliding potential. Such systems were realized in cold atoms experiments and the Thouless pumping was observed [19,20]. A similar interacting lattice model was studied in Ref. [21] and the emergence of gapless edge modes as one varies the parameter was numerically observed. Reference [21] also proposes that 1d spin chains which exhibit Thouless pumping can be realized using hydrogenated graphene.

B. The WZW term as a higher Berry phase and a codimension-4 diabolical point

1. A topologically nontrivial family

Now we will discuss a family which is topological but which does not rely on any symmetry. In particular, there is no way to understand its nontriviality by any kind of pumping argument. However, many of the same features we explored for the Thouless pump in Sec. II A will appear here. In many ways, this family is a 1+1d analog of the classic example of the Berry phase of a spin-1/2 in a magnetic field.

The family we are interested in is defined by a pair of complex fermions ψ^1, ψ^2 , packaged as a doublet Ψ , with the Lagrangian

$$\mathcal{L}_{\text{AW}} = -\bar{\Psi}i\partial\Psi - i\bar{\Psi}(M^0 + i\gamma^{01}M^k\sigma^k)\Psi, \quad (2.10)$$

where σ^k , $k = 1, 2$, and 3 are Pauli matrices acting in the flavor space of Ψ . This model belongs to the Abanov-Wiegmann “B” series (C2). There are four real parameters $M = (M^0, M^1, M^2, M^3)$. Any symmetries are accidental, and our conclusions will be robust to breaking them, but it is convenient to analyze this simple problem.

As before, this model is trivially gapped away from the massless point at the origin (it is a diabolical point). We study the response to slowly varying parameters $\phi : X \rightarrow S_m^3$ on a sphere

$$S_m^3 = \{M \in \mathbb{R}^4 \mid |M| = m > 0\}.$$

Abanov and Wiegmann [12] studied this problem and found that in the effective action for ϕ , there is a Wess-Zumino-Witten term of level 1, which has the form

$$S_{\text{top}} = \frac{1}{2} \int_X \epsilon^{\mu\nu} \omega_{ij}^{(2)}(\phi) \partial_\mu \phi^i \partial_\nu \phi^j d^2x = \int_X \phi^* \omega^{(2)}, \quad (2.11)$$

where $\omega^{(2)} = \frac{1}{2} \omega_{ij}^{(2)}(\phi) d\phi^i \wedge d\phi^j$ is locally a 2-form on S_m^3 satisfying $d\omega^{(2)} = 2\pi \text{vol}_{S_m^3}$ where $\text{vol}_{S_m^3}$ is the homogeneous volume form on S_m^3 normalized to have volume 1.

More precisely, $\omega^{(2)}$ is a U(1) 2-form gauge field on S_m^3 with Dixmier-Douady-Chern number 1, in other words with 2π units of magnetic flux through S_m^3 . This means that from chart to chart, $\omega^{(2)}$ transforms by 1-form gauge transformations $\omega^{(2)} \mapsto \omega^{(2)} + d\lambda$, where λ is a U(1) gauge field. Note (2.11) is invariant under such transformations. A 2-form gauge background with Dixmier-Douady-Chern number n on S^3 can be constructed by an analog of the “clutching construction” [22], where we define the 2-form on two hemispheres of S^3 and glue them by a gauge transformation along the equatorial S^2 by some λ which has Chern number n on the equator. This number may also be computed by integrating the curvature $\Omega^{(3)} = d\omega^{(2)}$ over S_m^3 . We find it is

$$\Omega^{(3)} = \frac{1}{6\pi |M|^4} \epsilon^{ABCD} M^A dM^B dM^C dM^D. \quad (2.12)$$

See Appendix A for a direct calculation. As a matter of fact, the model (2.10) has Spin(4) symmetry with respect to which Ψ_+ and Ψ_- transform as left-handed and right-handed spinors and M^A transforms as a vector. This symmetry together with the equation $d\Omega^{(3)} = 0$ fix the Berry curvature $\Omega^{(3)}$ up to a numerical factor.

The connection $\omega^{(2)}$ plays the role of the Berry connection for this system, in that it describes a quantized adiabatic response for this system (although it is second order in derivatives), so in this context we will refer to the Dixmier-Douady-Chern class of $\omega^{(2)}$, or equivalently the WZW level, as the higher Berry number. In fact, it is directly related to the 0+1d Berry phase by dimensional reduction (see Appendix C). The coefficient of this term is quantized to be an integer by gauge invariance. We will give a direct argument for this in Sec. VI by relating it to the winding number of the phase of a certain family of partition functions.

⁷According to the classification in Sec. VII, $\pi_1(\mathfrak{M}_1) = \mathbb{Z}_2$, so this pump invariant is again complete.

As with the Thouless pump invariant, the presence of this quantized topological term protects the diabolical point inside S_m^3 from all perturbations. That is, there is no way to extend $\omega_m^{(2)}$ to a 4-ball filling S_m^3 . We can therefore consider this family as representing a $\mathbb{Z} \subset \pi_3(\mathfrak{M}_1)$.⁸

Because of the large symmetry of the model (2.10), the stability of the gapless point at $M = 0$ can be explained using 't Hooft anomalies. The point $M = 0$ has global Spin(4) symmetry, and this symmetry cannot be gauged (it has an 't Hooft anomaly). This means that a gapless point at $M = 0$ is robust with respect to Spin(4)-invariant perturbations.

The argument based on the higher Berry phase continues to apply even if we break this Spin(4) symmetry. However, we cannot guarantee that the diabolical point does not turn into a different diabolical locus. For instance, if we add a chiral mass [not depending on the parameters (v, n_i)] to (2.10), then such a term will dominate near the origin, and gap the theory. However, there will still be some gapless points nearby. This situation is more generic in a four-parameter phase diagram than the single gapless point of (2.10). The presence of gapless points is still protected by the higher Berry number, no matter what perturbations we add. For a bosonic system also realizing this higher Berry phase, see Appendix B.

2. Interfaces

Let us examine the consequences of the higher Berry number for interfaces. Just as for the Thouless pump invariant, the higher Berry number implies that fixing a basepoint $M_* \in S_m^3$, any family of interfaces parametrized by $M \in S_m^3$, which connect the theory at M_* for $x \rightarrow -\infty$ with the theory at M for $x \rightarrow \infty$, must include at least one interface with a degenerate ground state. This follows from the same path space argument we used in Sec. II A 2.

One simple family to analyze in detail is given by taking interfaces which interpolate along a straight line from M to M_* . Then, the interface corresponding to $M = -M_*$ passes through the origin and corresponds to an interface hosting zero modes, while all other interfaces are trivially gapped.⁹ Without loss of generality we can take $M_* = (m, 0, 0, 0)$ and then this special interface corresponds to the inverting mass profile studied by Jackiw and Rebbi [18]. It has two zero modes localized to the place where the interface crosses the origin, therefore the ground state has quadruple degeneracy.

Near this interface in the family, that is for $M = -M_* + \delta M$, we can see how the zero modes at the interface become gapped by studying perturbation theory. The effective Hamiltonian for the interface modes is

$$H_{\text{eff}} = a_\alpha^\dagger \vec{\sigma}_\beta^\alpha a^\beta \cdot \vec{\mu}, \tag{2.13}$$

where a_α , $\alpha = 1, 2$ are the annihilation operators for the interface modes, $\vec{\sigma}$ is the 3-vector of Pauli matrices, and $\vec{\mu}$ is defined by $\delta M = (0, \mu_1, \mu_2, \mu_3)$. This Hamiltonian has a nondegenerate ground state for all $\vec{\mu} \neq 0$ with energy $-2|\vec{\mu}|$.

Moreover, from this 0+1d perspective, we find that the degenerate interface can be considered as a diabolical point!

This diabolical point, being codimension 3, should have some ordinary Berry curvature which protects it. To compute it, we note that both the Fock vacuum $|0\rangle$ and the fully occupied state $a_1^\dagger a_2^\dagger |0\rangle$ are eigenvectors of H_{eff} with energy 0 for all $\vec{\mu}$ and thus do not contribute to the Berry curvature. One may restrict H_{eff} to the two-dimensional subspace spanned by $a_\alpha^\dagger |0\rangle$. In this subspace, H_{eff} reduces the Hamiltonian of a spin-1/2 in a magnetic field 2μ . This is a paradigmatic example of a diabolical point, and the integral of the Berry curvature over the 2-sphere $|\vec{\mu}| = 1$ is 2π .

So we have seen that there is a single degenerate interface for $M = -M_* \in S_m^3$. Near this point, we can study an effective 0+1d problem and we find that this special interface represents a 0+1d diabolical point protected by the ordinary Berry curvature. However, this situation is not consistent from a 0+1d point of view, in that the small sphere which surrounds $M = -M_*$, which we have just argued has ordinary Berry number 1, is the boundary of a disk which closes off on the other side S^3 , which would imply it has ordinary Berry number 0. In this sense, we consider the theory of the interface to have an anomaly: its peculiar phase diagram is only realizable because it exists as an interface. Following the terminology of Ref. [9], the interface theory exhibits ‘‘anomaly in the space of couplings.’’

We will see in Sec. VI that this is a general phenomenon that connects Berry phases and diabolical points in all dimensions. Moreover the number of diabolical points for the interface, weighted by their Berry numbers, is determined by the bulk higher Berry number, namely the integral of $\Omega^{(3)}$ over S^3 .

The reason one gets quadruple degeneracy rather than the double degeneracy of von Neumann and Wigner is that our fermions are free. It is not possible to halve the number of fermions by imposing a Majorana condition on them. If we do this, the mass term in (2.10) will vanish. However, one can halve the degeneracy by adding an interaction of the form $\epsilon(\vec{\mu})(a^\dagger a - 1)^2$, where $\epsilon(0) > 0$. It does not affect the subspace spanned by $a_\alpha^\dagger |0\rangle$, but gives positive energy both to the Fock vacuum and the fully occupied state.

3. No further diabolical points

Let us consider a system of N real fermions in 1+1d with the Lagrangian (2.9). The mass matrix M is real, and the theory is massive if $\det M$ is nonzero. Thus one can identify the parameter space of massive free theories as $GL(N, \mathbb{R})$. It has two connected components labeled by the sign of the determinant of M and denoted $GL_+(N, \mathbb{R})$ and $GL_-(N, \mathbb{R})$. They correspond to two different short-range entangled (SRE) phases of fermions in 1+1d, the nontrivial one being the Kitaev phase with Majorana zero modes at its boundaries. Each of the components is homotopy equivalent to $SO(N)$. For $N > 2$ this space has fundamental group \mathbb{Z}_2 , which corresponds to our \mathbb{Z}_2 fermion parity Thouless pump in Sec. II A 4. For $N > 4$, $\pi_3(SO(N)) = \mathbb{Z}$ [24], which corresponds to the codimension-3 diabolical point we have just discussed (see Appendix A).

⁸By the classification of Sec. VII, this is an isomorphism.

⁹This can be seen perturbatively near the soluble interfaces $M = \pm M_*$. For the others, this can be argued using the results of Appendix A of Ref. [23].

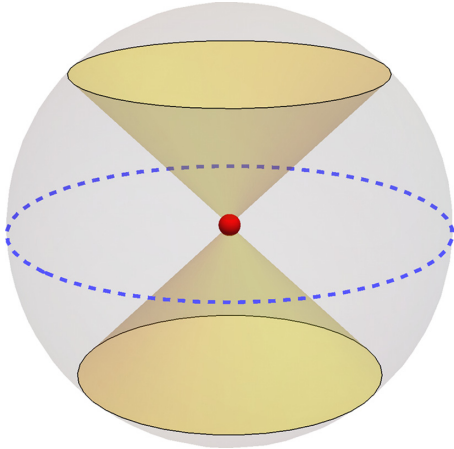


FIG. 1. The nearby phase diagram of the model (3.1), with our two massless Dirac fermions at the origin (red dot), drawn with the vertical axis representing $\cos \nu$, and the horizontal axes schematically representing the three coordinates n_i . The dashed blue circle (actually an S^2) is the family at $\nu = \pi/2$ which realizes the 2+1d Thouless pump. The yellow cones are gapless points where one eigenvalue of the mass matrix changes sign. They separate the Thouless region from the contractible “caps,” where the system realizes the integer quantum Hall effect with Chern-Simons level (i.e., Chern number) $k = \pm 1$, depending on the sign of $\cos \nu$.

There are further stable homotopy groups of $SO(N)$ in arbitrarily high degrees (exhibiting Bott periodicity), which would appear to imply there are free fermion diabolical points of arbitrarily high codimension. However, these points cannot be stable from the point of view of the effective action, since the WZW term (2.11) already has the most number of derivatives of the parameter background ϕ which can be fit in a topological term. This is also reflected in the classification of Sec. VII. Thus we expect all these free fermion families, and indeed any infrared-trivial family in d space dimensions on $S^{m \geq d+3}$ can be extended to an infrared-trivial family on a bounding ball B^{m+1} , so there is no associated protected diabolical point. See Appendix D for more details.

III. FREE FERMIONS IN 2+1d

A. Skyrmion charge and a codimension-3 diabolical point

1. A topologically nontrivial family

Now we consider two 2+1d Dirac fermions $\psi_{1,2}$ with $U(1)$ symmetry $\psi_j \mapsto e^{i\alpha} \psi_j$. We write the doublet $\Psi = (\psi_1, \psi_2)$. The most general free Lagrangian which preserves $U(1)$ symmetry is

$$\mathcal{L} = -\bar{\Psi} i \partial \Psi - im \bar{\Psi} (\cos \nu + n_i \sigma^i \sin \nu) \Psi \quad (3.1)$$

where $\nu \in (0, \pi)$, $n^i n^i = 1$ and σ^i are Pauli matrices acting in flavor space. ν and n^i together parametrize S^3 with radius m in \mathbb{R}^4 . Without loss of generality we can take $m \geq 0$ by $n \rightarrow -n$, $\nu \rightarrow \pi - \nu$. The parameters $\sin \nu n^i$ for fixed ν can be thought of as living on an S^2 of radius $\sin \nu$. The model preserves $SU(2)$ symmetry at $\nu = 0, \pi$ where this S^2 shrinks to zero size. We draw the phase diagram in Fig. 1.

Consider first $\nu = \frac{\pi}{2}$. This family is the 2+1d sibling of (2.1), the next member of the Abanov-Wiegmann A

series (C21). For $m > 0$, the theory is gapped with a unique ground state, while for $m = 0$ the theory is gapless. One can ask if it is possible to remove the gapless point by a $U(1)$ -preserving deformation without bringing new gapless point from infinity. One may even allow deformations which involve additional massive degrees of freedom. We will now argue that this is not possible. If we promote n to be position dependent and turn on the background gauge field A for the $U(1)$ symmetry, then it can be shown [12] that the effective action contains¹⁰ a theta-term for n with $\theta = \pi$ (i.e., the Hopf term) together with the following A -dependent term:

$$S_{\text{top}} = \frac{1}{8\pi} \int \epsilon^{\mu\nu\rho} A_\mu \epsilon^{ijk} n_i \partial_\nu n_j \partial_\rho n_k d^3x. \quad (3.2)$$

A derivation of the A -dependent term is given in Appendix A. The coefficients of these two terms are related by the spin-charge relation [25].

Equation (3.2) is a special case of the following effective action for a family of gapped $U(1)$ -invariant models with scalar parameters ϕ^i :

$$S_{\text{top}} = \frac{1}{2} \int \epsilon^{\mu\nu\rho} A_\mu \partial_\nu \phi^i \partial_\rho \phi^j \tau_{2,ij}(\phi) d^2x = \int A \wedge \phi^* \tau_2, \quad (3.3)$$

where $\tau_2 = \frac{1}{2} \tau_{2,ij}(\phi) d\phi^i \wedge d\phi^j$ is a 2-form on the parameter space. Compare (2.4). Such terms in the effective action were recently considered in the context of lattice models [8]. As described in Ref. [8], they describe a 2+1d analog of the Thouless pump. Gauge-invariance requires τ_2 to be closed and have integral periods. Equivalently, one can say that closedness and integrality is required by the conservation of the topological current and the integrality of the charge of skyrmions (topologically nontrivial configurations of the fields ϕ^i which approach a constant at infinity).

In the case of the model (3.1) at $\nu = \pi/2$, the form τ_2 is proportional to the volume form of S^2 , and the topological current takes the form

$$J_{\text{top}}^\mu = \frac{1}{8\pi} \epsilon^{\mu\nu\rho} \epsilon^{ijk} n_i \partial_\nu n_j \partial_\rho n_k. \quad (3.4)$$

The basic skyrmion is obtained by identifying \mathbb{R}^2 with a point at infinity added with S^2 and letting n to be the identity map $S^2 \rightarrow S^2$. The charge of such a skyrmion is 1.¹¹ Since the form τ_2 is not exact, the 2-parameter family obtained by setting $\nu = \pi/2$ is not contractible in the space $\mathfrak{M}_2^{U(1)}$ of all infrared-trivial $U(1)$ -invariant field theories in 2+1d.

The stability of the gapless point at $m = 0$ can also be explained by the mixed 't Hooft anomaly between the time-reversal symmetry and the $SU(2)$ flavor symmetry which acts on Ψ_1, Ψ_2 as a doublet. Time-reversal symmetry is present for all n if ν is fixed to be $\pi/2$, while $SU(2)$ symmetry is restored

¹⁰In fact, one can use $SU(2)$ rotation $U(x)$ to fix n to be a constant while introducing $SU(2)$ background gauge field $A = iU^{-1}dU$. Then the improperly quantized Chern-Simons term for the background $SU(2)$ gauge field reproduces the $\theta = \pi$ term.

¹¹In some settings, this is referred to as quantum Hall ferromagnetism. See Ref. [26] and references within.

at $m = 0$. Therefore 't Hooft anomaly matching requires gapless modes at $m = 0$. However, this explanation only shows stability with respect to deformations which preserve both symmetries. For example, this kind of argument breaks down if ν is slightly different from $\pi/2$. Also, the mixed 't Hooft anomaly between time-reversal symmetry and $SU(2)$ has order 2, i.e., it becomes trivial if we take an even number of copies of the system (3.1) with $\nu = \pi/2$. The argument based on the effective action (3.3) shows that neither the number of copies, nor time-reversal nor flavor symmetry are essential for the stability of the gapless point at $m = 0$.

2. Interfaces

This noncontractibility has immediate consequences for smooth interfaces between different member of the family. We can fix $n = (0, 0, 1)$ as the basepoint and consider smooth interfaces between the basepoint model and other members of the $\nu = \pi/2$ family obtained by varying mn^i along a straight line segment. By a general argument explained in Sec. VI, if this family of interfaces is continuous, at least one of the interfaces must be gapless. For the model (3.1), a gapless interface occurs for $n = (0, 0, -1)$. This is explained in detail in Appendix C. In brief, one finds that the eigenvalue equation for modes with zero momentum along the interface reduces to the 1d eigenvalue equation for the Jackiw-Rebbi model [i.e., the model (2.1) with $M_2 = 0$ and M_1 varying between m and $-m$]. Thus there is exactly one zero mode in this case, and the interface carries a massless Dirac fermion.

Now consider perturbing n from $(0, 0, -1)$ to a nearby point $(w_1, w_2, -1 + O(w_{1,2}^2))$. Using perturbation theory in w_1, w_2 , it is easy to check that this gaps the Dirac fermions. The low-energy effective action for the interface mode is then described by Eq. (2.1) with $M_1 = w_1, M_2 = w_2$. After integrating out the mode on the interface, we end up with an effective action (2.4), where the 1-form τ_1 on the w plane is given by

$$\tau_1 = \frac{1}{2\pi} d \arg w, \tag{3.5}$$

where $w = w_1 + iw_2$. This form is singular at $w = 0$ which is the location of the interface diabolical point. Thus a codimension-3 diabolical point for the bulk theory leads to a codimension-2 diabolical point for the interface, which is again protected by a Thouless pump.

There is also a point $n = (0, 0, 1)$ where there are no interface modes at all. Note that the problem studied here preserves a $U(1)$ subgroup of $SU(2)$ flavor symmetry, and that $n = (0, 0, 1)$ and $n = (0, 0, -1)$ are the only two points on S^2 fixed by this symmetry. Thus, if we assume that the gapless modes on the interface occurs only at isolated points in the parameter space (which is a generic situation), then $n = (0, 0, -1)$ is the only such point.

Note that if we had a UV-complete theory describing the interface for all values of n , then the sum of periods of τ_1 over all gapless points would necessarily be zero. Here it is equal to 1, thanks to the existence of a point where the interface mode becomes non-normalizable and escapes to the bulk. In the language of Ref. [9], the interface theory has an ‘‘anomaly in the space of couplings.’’

3. A more general family

Now let us move away from $\nu = \frac{\pi}{2}$. Since S^3 does not have noncontractible 2-cycles, we conclude there must be a phase transition as we move away from $\nu = \frac{\pi}{2}$ while keeping the sphere radius m fixed. Indeed, since the determinant of the mass matrix is $m^2 \cos(2\nu)$, there are two cones of gapless field theories at $\nu = \frac{\pi}{4}$ and $\nu = \frac{3\pi}{4}$. They meet at the codimension-4 diabolical point $m = 0$. See Fig. 1.

Let us briefly comment on what protects the codimension-1 gapless cones at $\nu = \frac{\pi}{4}, \nu = \frac{3\pi}{4}$. These are also diabolical points, in a sense, although they are more familiar as a continuous transition between different integer quantum Hall states. To see this, note the masses of the two Dirac fermions are given by the eigenvalues of the mass matrix, which are

$$\sqrt{2}m \cos\left(\nu \pm \frac{\pi}{4}\right). \tag{3.6}$$

Thus, due to the parity anomaly, whenever the mass of one of the Dirac fermion crosses from negative to positive the effective action contains additional term $\frac{1}{4\pi} AdA + 2CS_{\text{grav}}$ where $\int_{M_3} CS_{\text{grav}} = \pi \int_{M_4} \widehat{A}(R)$ with $M_3 = \partial M_4$ is the gravitational Chern-Simons term.

If we study an interface which varies from $\nu < \pi/4$ to $\pi/4 < \nu < 3\pi/4$, these Chern-Simons terms lead to gapless chiral modes near where ν crosses $\pi/4$. This can be thought of as another manifestation of bulk-boundary correspondence.

In light of this, the Abanov-Wiegmann theory does not describe a generic codimension-3 diabolical locus. Suppose we begin with the three-parameter phase diagram of (3.1) with $\nu = \pi/2$, just parametrized by the n_i , and then tune ν slightly away from $\pi/2$. Because the diabolical point at the origin is protected by the Thouless invariant, it cannot disappear completely, but it is unstable in the view of the above. For small n_i , the $SU(2)$ -invariant mass term dominates, and gaps the central point. We find that it is replaced by an island of an integer quantum Hall state with a single gapless Dirac fermion along its two-dimensional boundary. For a three-parameter phase diagram, these diabolical islands are the generic situation, provided $U(1)$ symmetry is maintained.

B. The Wess-Zumino-Witten term and a codimension-5 diabolical point

Even if we do not assume any symmetry, a possible topological term in the effective action for scalar fields ϕ^i is the WZW term

$$S_{\text{top}} = \int_X \phi^* \omega^{(3)}, \tag{3.7}$$

where $\omega^{(3)}$ is a $U(1)$ 3-form gauge field. Compare (2.11). Locally it is a 3-form, but it undergoes a 2-form gauge transformation $\omega^{(3)} \mapsto \omega^{(3)} + dB$ when one changes charts, under which (3.7) is invariant. Here B is a 2-form gauge field, similar to $\omega^{(2)}$ in Sec. II B. Therefore the 4-form $\Omega^{(4)} = d\omega^{(3)}$ is well-defined and closed, but not necessarily exact. Moreover, for $\exp(iS_{\text{top}})$ to be well-defined, periods of G must satisfy certain quantization conditions. For fermionic systems with a general parameter space these conditions are rather nontrivial, see Ref. [7] for a brief discussion. However in the case when the parameter space is S^4 the same argument as for the $1+1d$

WZW term shows that $\int_{S^4} \Omega^{(4)}$ must be an integral multiple of 2π . If the cohomology class of $\Omega^{(4)}$ is nontrivial, the corresponding family of field theories is noncontractible in \mathfrak{M}_3 , the space of infrared-trivial 2+1d field theories.

As a simple example, consider a theory of four complex fermions $\Psi = (\psi_1, \psi_2, \psi_3, \psi_4)$. It is convenient to regard them as a spinor representation of a flavor Spin(5) symmetry, where Ψ transforms as a spinor, even though this symmetry will not be important in our analysis. Following Ref. [12] we consider the Lagrangian

$$\mathcal{L} = -\bar{\Psi}i\partial\Psi - im\bar{\Psi}(\cos\nu\Gamma^5 + n_i\Gamma^i \sin\nu)\Psi, \quad (3.8)$$

where Γ^i are the gamma matrices for the Spin(5) flavor indices, and $n^2 = 1$, $i = 1, 2, 3, 4$. Under the Spin(5) flavor action on Ψ the mass parameters $(m \cos\nu, m \sin\nu n^i)$ transform as a vector. Without loss of generality we take $m \geq 0$ by $\nu \rightarrow \pi - \nu, n \rightarrow -n$. There is also a U(1) symmetry under which $\Psi \mapsto e^{i\nu}\Psi$.

For $m > 0$, the theory is gapped with a unique ground state. The point $m = 0$ is a gapless diabolical point of codimension 5. The parameter space with this point removed is homotopy equivalent to S^4 . This diabolical point is protected by the higher Berry phase on S^4 and thus is stable under arbitrary deformations, including those which break Spin(5) and/or U(1). Indeed, if we promote the parameters $\phi = (\nu, n_i)$ to be position dependent, the effective action of the fermion contains [12] (see also Appendix A) the WZW term (3.7) with

$$\Omega^{(4)} = \frac{1}{8\pi} \epsilon^{ijkl} \sin^3\nu dv n_i dn_j dn_k dn_l. \quad (3.9)$$

This is the volume form of S^4 normalized so that $\int_{S^4} \Omega^{(4)} = 2\pi$. Since this integral is quantized, the family (3.8) is not contractible in the space of infrared-trivial 2+1d field theories.

As in Sec. III A, if we add a constant diagonal mass term to (3.8), this will open a gap near $M = 0$ into a topological phase (equivalent to eight copies of $p + ip$ superconductors). At the boundary of this region two of the eigenvalues of the mass matrix change sign, yielding an S^4 's worth of two massless Dirac fermions, which is protected by the higher Berry number of the surrounding region.

A discussion of interfaces in this model and other models in the ‘‘B’’ series of Abanov and Wiegmann can be found in Appendix C.

IV. FREE FERMIONS IN 3+1d

A. Skyrmion charge and a codimension-4 diabolical point

Consider a model of free fermions in 3+1d with a Lagrangian

$$\mathcal{L} = -\bar{\Psi}i\partial\Psi - i\bar{\Psi}(M^0 + i\gamma^{0123}M^i\sigma^i)\Psi, \quad (4.1)$$

where $\Psi = (\psi^1, \psi^2)$ is a doublet of Dirac fermions, $\gamma^{0123} = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the chirality operator, σ^i for $i = 1, 2, 3$ are Pauli matrices, and $M = (M^0, M^1, M^2, M^3)$ is the 4-vector of mass parameters. Uppercase Roman indices take values from 0 to 3. This is a 3+1d analog of the model (2.1). Its topological properties have been studied by Goldstone and Wilczek [14]. It is also a special case of the ‘‘A’’ series of Abanov and

Wiegmann [12]. The model is gapless for $M = 0$ and gapped otherwise.

We claim that $M = 0$ is a codimension-4 diabolical point protected by U(1) symmetry. To see this, we need to examine topological terms in the effective action which depend on the background U(1) gauge field A as well as scalar parameters ϕ^i representing the slowly varying mass terms M . The term which is analogous to (2.4) and (3.3) is linear in A and has the form

$$\begin{aligned} S_{\text{top}} &= \frac{1}{6} \int_X \epsilon^{\mu\nu\rho\sigma} A_\mu \partial_\nu \phi^i \partial_\rho \phi^j \partial_\sigma \phi^k \tau_{3,ijk}(\phi) d^4x \\ &= \int_X A \wedge \phi^* \tau_3, \end{aligned} \quad (4.2)$$

where τ_3 is a 3-form on the parameter space. Compare (2.4) and (3.3). Gauge invariance requires τ_3 to be closed and to have integral periods. If the restriction of τ_3 to some 3-parameter family is not exact, the corresponding family cannot be contracted to a point in the space $\mathfrak{M}_3^{U(1)}$ of infrared-trivial 3+1d field theories with a U(1) symmetry. Intuitively, this is because the term (4.2) gives rise to a topological current

$$J_{\text{top}}^\mu = \frac{1}{6} \epsilon^{\mu\nu\rho\sigma} \partial_\nu \phi^i \partial_\rho \phi^j \partial_\sigma \phi^k \tau_{3,ijk}(\phi), \quad (4.3)$$

which gives charge to skyrmions [topologically nontrivial configurations of the parameter fields $\phi^i(x)$]. Continuously deforming the model within the class of infrared-trivial field theories cannot change the charge of skyrmions. Note that charge quantization for arbitrary spatial slices is equivalent to the integrality of periods of τ_3 .

In the case of interest to us, the parameters ϕ^i can be identified with M^A . The gapped locus is defined by $M \neq 0$ and is homotopically equivalent to S^3 . This space has $H^3(S^3, \mathbb{Z}) = \mathbb{Z}$, so there is a possibility for a term of the form (4.2). In fact, it was shown in [14] that upon integrating out fermions the model (4.1) generates the effective action (4.2) with the 3-form τ_3 given by

$$\tau_3 = \frac{1}{12\pi^2 |M|^4} \epsilon^{ABCD} M^A dM^B dM^C dM^D. \quad (4.4)$$

This implies that the family (4.1) describes a codimension-4 diabolical point.

The consequences of this for interfaces are similar to the ones in the 1+1d and 2+1d cases. Namely, if we restrict to the S^3 given by $|M| = 1$ and study interfaces from a fixed basepoint p_0 to other points $p \in S^3$, then any family of interfaces which depends continuously on p will have at least one gapless interface. If there is a single such point $p_* \in S^3$, we can consider it as a 2+1d diabolical point for the interface theory, and the effective action for the nearby gapped interfaces will contain a topological term of the form (3.3), where the 2-form τ_2 integrates to 1 on any S^2 surrounding p_* . That is, the field theory on the interface has an ‘‘anomaly in the space of couplings’’ in the sense of Ref. [9]. For a certain nice family of interfaces, the effective 2+1d theory on the interface can be described by the Lagrangian (3.1). This is explained in more detail in Appendix C.

B. WZW terms and a codimension-6 diabolical point

If we do not assume any symmetry, then the most obvious topological term in the effective action is the WZW term

$$S_{\text{top}} = \int_X \phi^* \omega^{(4)}, \tag{4.5}$$

where $\omega^{(4)}$ is a 4-form gauge field on the parameter space, analogous to (2.11), (3.7), and the quantum mechanical Berry phase. That is, locally $\omega^{(4)}$ is a 4-form, but when one goes from chart to chart $\omega^{(4)}$ transforms by a 3-form gauge transformation $\omega^{(4)} \mapsto \omega^{(4)} + dC$, where C is a 3-form gauge field. Equation (4.5) is invariant under such gauge transformations. The gauge curvature $\Omega^{(5)} = d\omega^{(4)}$ is a closed but not necessarily exact 5-form on the parameter space (the WZW 5-form).

As in lower dimensions, if $\Omega^{(5)}$ has a nonzero integral over a family parametrized by S^5 , then that family cannot be contracted to a point within the space \mathcal{M}_3 of infrared-trivial field theories in 3+1 dimensions. In particular, we expect that there are diabolical points in codimension 6 which are stable without imposing any symmetry. An example of such a model has been constructed by Abanov and Wiegmann [12]. Consider four Dirac fermions $\Psi = (\psi^1, \dots, \psi^4)$ transforming in the fundamental representation of $\text{USp}(4)$ [or equivalently, as a spinor of $\text{Spin}(5)$]. Following Ref. [12] we consider the Lagrangian

$$\mathcal{L} = -\bar{\Psi}i\partial\Psi - i\bar{\Psi}(M^0 + i\gamma^{0123}M^i\Gamma^i)\Psi, \tag{4.6}$$

where Γ^i , $i = 1, \dots, 5$, are $\text{Spin}(5)$ Dirac matrices acting in the flavor indices and $M = (M^0, M^1, \dots, M^5)$ is a 6-vector of mass parameters. The model has a gapless point at $M = 0$. Reference [12] shows that upon integrating out the fermions one gets the effective action (4.5) where the 5-form $\Omega^{(5)}$ on $\mathbb{R}^6 \setminus \{0\}$ is given by

$$\Omega = \frac{1}{60\pi^2|M|^6} \epsilon^{ABCDEF} M^A dM^B dM^C dM^D dM^E dM^F. \tag{4.7}$$

The integral of this 5-form over the S^5 is 2π . This implies that $\Omega^{(5)}$ is not exact and therefore the family (4.6) is not contractible in the space of all 3+1d infrared-trivial field theories.

As usual, noncontractibility implies that a family of interfaces which depends continuously on M will have at least one gapless point. As explained in Appendix C, in the case of model (4.6) such a gapless point occurs when one considers an interface between $M = (m, 0, 0, 0, 0, 0)$ and $M = (-m, 0, 0, 0, 0, 0)$. In the neighborhood of this special interface there are almost-gapless interface modes described by the Lagrangian (3.8). This is an example of the boundary-bulk correspondence.

C. Axion couplings

Until now, all diabolical loci not protected by any symmetry arose from WZW terms in the effective action for the parameter fields. Such diabolical loci occur in codimension $d + 3$, where d is the spatial dimension, and are direct descendants of the von Neumann-Wigner points. In 3+1d, a new phenomenon occurs: One encounters diabolical loci in codimension 2, which are related to coupling to background

geometry. Topological invariance requires these terms to have the form

$$S_{\text{top}} = \frac{1}{2\pi} \int_X \text{CS}_{\text{grav}} \wedge \phi^* \rho, \tag{4.8}$$

where ρ is a 1-form on the parameter space with quantized periods. One can always assume that the parameter space is S^1 , in which case $\rho = Nd\alpha$, where α is a 2π -periodic coordinate on S^1 and N is a number. Then the above coupling can be written as an axion coupling

$$S_{\text{top}} = -\frac{N}{384\pi^2} \int_X \alpha \text{Tr}R^2. \tag{4.9}$$

In order for $\exp(iS_{\text{top}})$ to be well-defined, N has to be quantized. The precise quantization condition depends on whether X is allowed to be an arbitrary oriented 4-manifold (the bosonic case) or a spin manifold (the fermionic case). In the former case, N must be a multiple of 16, while in the latter case it must be integral. Whenever N is nonzero, the family parameterized by α is noncontractible, signifying the presence of a diabolical locus in codimension 2. This diabolical locus is stable with respect to arbitrary modifications of the theory in the UV.

Examples of models where gravitational axion couplings arise are well-known. One can take a single Majorana fermion with mass M and set $\alpha = \arg M$. Gravitational anomaly for chiral $\text{U}(1)$ symmetry then leads to the coupling (4.9) with $N = 1$.

The physical consequences of the gravitational axion coupling are also well-known. If α winds k times around the origin in the 12 coordinate plane in space-time, then at the origin of the coordinate plane there must be chiral gapless modes propagating in the x^3 direction with the chiral central charge $c_R - c_L = Nk/2$. If one considers smooth interfaces between all models with $|M| = m > 0$ and the basepoint model $M = m$, then there will be a gapless interface for at least one value of $\alpha = \arg M$. This is so even if one modifies the theory in the UV while preserving the gap for all α . In the case of the massive 3+1d Majorana fermion, the gapless interface carries a massless 2+1d Majorana fermion.

Usually the stability of the gapless point at $M = 0$ is explained in terms of the 't Hooft anomaly for the chiral $\text{U}(1)$ symmetry. This symmetry is restored at $M = 0$ and thus 't Hooft anomaly matching requires the presence of massless modes. This explanation is problematic from the point of view of lattice regularization, since a chiral $\text{U}(1)$ symmetry cannot be realized on a lattice in a completely local way. Our explanation avoids using any symmetries and relies instead on the topology of the parameter space of infrared-trivial field theories.

If one considers gapped models with a global $\text{U}(1)$ symmetry, an axion coupling to a background $\text{U}(1)$ gauge field is also allowed. Such a coupling signifies a diabolical locus in codimension 2 as well. The usual explanation of stability involves a mixed 't Hooft anomaly between two different $\text{U}(1)$ symmetries one of which is broken only by a mass term.

V. INTERACTING GAUGE THEORIES

A. Gauge theories in 2+1d

1. Operation ST and the generalized Thouless pump invariant

Let us describe a general technique for producing new diabolical loci in 2+1d protected by a U(1) Thouless pump. Suppose we begin with such a family. If we turn on nontrivial background gauge field A for the U(1) symmetry and promote the parameters to be position-dependent $\phi : X \rightarrow \mathcal{M}$ on space-time X , then the effective action can be expressed as (3.3). The U(1) current is

$$j^\mu = \frac{1}{2} \epsilon^{\mu\nu\rho} \partial_\nu \phi^i \partial_\rho \phi^j \tau_{2,ij}(\phi). \quad (5.1)$$

Now, let us gauge the U(1) symmetry and study the effective action of the resulting theory. We first add a Chern-Simons counterterm

$$\int_X \left(\frac{1}{4\pi} AdA - \frac{1}{2\pi} AdB \right), \quad (5.2)$$

where B is a background U(1) gauge field, and then promote A to be a dynamical gauge field a . We may also add a Maxwell term for a , but the infrared behavior is determined by the level one Chern-Simons term for a . It makes a massive, and moreover makes the corresponding low-energy theory invertible (see for instance Appendix B of Ref. [27]). Thus we obtain a new system that is invertible away from the diabolical points and has a global U(1) symmetry with a current

$$j'^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu a_\rho \quad (5.3)$$

which couples to the background gauge field B . This is the ST operation of Ref. [28] which originally appeared in Ref. [29].

The new system is described at low energies by a topological U(1) gauge theory with an action

$$S_{\text{eff}} = \int_X \left(\frac{1}{4\pi} ada - \frac{1}{2\pi} adB + a \phi^* \tau_2 \right). \quad (5.4)$$

By shifting $a \mapsto a + B$ in (5.4), we find the effective action

$$S'_{\text{eff}} = \int_X \left(\frac{1}{4\pi} ada + a \phi^* \tau_2 \right) + \int_X \left(-\frac{1}{4\pi} BdB + B \phi^* \tau_2 \right). \quad (5.5)$$

The first term describes a family of infrared-trivial 2+1d field theories which does not couple to the background U(1) gauge field B .¹² Therefore the Thouless pump invariant is determined by the second term and is the same as for the original family (3.3), but diabolical point is modified: it is mapped to its ST transform.

¹²The partition function of this invertible theory is a Hopf-like term

$$e^{iH[\phi]} \equiv \int Dae^{i \int_X \left(\frac{1}{4\pi} ada + a \phi^* \tau_2 \right)}, \quad (5.6)$$

which can be formally expressed as (omitting the gravitational Chern-Simons term $-\int_X 2CS_{\text{grav}}$) $H[\phi] = \pi \int_Y \phi^* (\tau_2)^2$ for spin 4-manifold Y that bounds the space-time. It does not depend on the choice of Y . An example of such a Hopf term is discussed in Ref. [30].

2. $U(1)_1$ with two fermions

Let us apply the above procedure to the theory of two complex 2+1d fermions with SU(2)-covariant vector mass discussed in section III. In this case j in (5.1) is the current for the skyrmion number. The U(1) symmetry transforms the two fermions by the same phase.

The ST operation gives an interacting U(1)₁ Chern-Simons-matter theory with two fermions of charge one. This theory has a magnetic U(1) symmetry whose current is the dual of the gauge field strength. From the previous discussion, away from the $m = 0$ point the theory has a Thouless pump invariant for this U(1) symmetry.

The first term in the effective action (5.5) gives an additional $\theta = \pi$ Hopf term [30], which cancels the Hopf term from two complex fermions discussed in section III. This implies that the skyrmion that carries a unit charge of the topological current j is a boson [30]. Since the equation of motion for a in (5.5) identifies the current (5.1) with (5.3) up to a contact term, the skyrmions correspond to the monopoles in the gauge theory. Indeed, the monopole operators in the U(1) gauge theory with charge one fermions are bosons [31].¹³ The critical point at $m = 0$ is protected by the nonzero skyrmion charge as in the original family of two complex fermions.

Similar to the discussion in Sec. III, we can take $m > 0$ and consider a family of interfaces with fixed base point $n = (0, 0, 1)$ obtained from the family in Sec. III by gauging the U(1) symmetry. From the “boundary-bulk” correspondence this family must contain an gapless interface. An analysis along the lines of Appendix C shows that this gapless interface hosts a massless periodic scalar which is the bosonization of a massless Dirac fermion.

3. $U(N)$ gauge theory with two scalars, deconfined quantum criticality, and boson/fermion duality

In the examples we discussed so far (and the examples discussed in Ref. [12]) the Berry phase or Thouless pump invariant are derived from integrating out massive fermions. Here we present a new class of examples where the Berry phase is obtained from a massive scalar through the Higgs mechanism.

Consider U(1) gauge theory with two complex scalars ϕ of charge one and quartic potential that respects the SU(2) flavor symmetry,¹⁴ aka the $\mathbb{C}P^1$ or abelian Higgs model. We can add the SU(2) vector mass term

$$V(\phi) = m^2 n_i \phi^\dagger \sigma^i \phi + \lambda (\phi^\dagger \phi)^2, \quad (5.7)$$

where we take $m^2 \rightarrow \infty$, $\lambda > 0$ and σ^i are the Pauli matrices for SU(2) isospin and $\sum n_i^2 = 1$.

For large m^2 , due to the Higgs potential the U(1) gauge field obtains a mass square m^2/λ . If we promote the mass parameters n_i to be position dependent, then to leading order

¹³Another way to see this is that the theory obeys spin/charge relation with respect to the dynamical U(1) gauge field, and thus gauge invariant local operators are bosons.

¹⁴The more precise global symmetry is discussed in Ref. [32].

in the $1/m$ expansion there is current

$$j = i(\phi^\dagger d\phi - \text{H.c.}) = -\frac{n_1 dn_2 - n_2 dn_1}{2(1 - n_3)(\lambda/m^2)} + \dots, \quad (5.8)$$

where \dots are suppressed by $1/m^2$. The equation of motion for the $U(1)$ gauge field a to leading order gives

$$a = \frac{n_1 dn_2 - n_2 dn_1}{2(1 - n_3)} + \dots, \quad (5.9)$$

where \dots are suppressed by $1/m^2$. Thus, to leading order,¹⁵

$$\begin{aligned} da &= \frac{dn_1 dn_2}{2(1 - n_3)} + \left(\frac{dn_1 dn_2}{2(1 - n_3)} + \frac{dn_3(n_1 dn_2 - n_2 dn_1)}{2(1 - n_3)^2} \right) \\ &= -\frac{1}{2} \epsilon^{ijk} n_i dn_j dn_k, \end{aligned} \quad (5.11)$$

The relation (5.11) implies that the monopole of the microscopic gauge theory corresponds to the skyrmion configuration of the field n_i .

If we turn on background gauge field A for the magnetic $U(1)$ symmetry with current $j^\mu = -\frac{1}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu a_\rho$, at low energy the theory contains the following effective action:

$$\int A_\mu j^\mu = \frac{1}{8\pi} \epsilon^{\mu\nu\rho} \epsilon^{ijk} \int A_\mu n_i \partial_\nu n_j \partial_\rho n_k. \quad (5.12)$$

Thus this model provides an example of family of interacting bosonic theories that has a nontrivial 2+1d Thouless pump invariant.

For $m^2 = 0$, the model is proposed to describe the deconfined quantum critical point between the Néel phase and the valence bond solid (VBS) [33]. This is a proposed continuous phase transition in a $S = 1/2$ antiferromagnet from a Néel phase, where spin rotation $SO(3)$ is spontaneously broken, to a VBS phase, where site-rotation symmetry is broken. On a square lattice, this group is \mathbb{Z}_4 , on a honeycomb lattice, it is \mathbb{Z}_3 , and so on. The $SO(3)$ spin-rotation is proposed to act as the flavor symmetry of the two scalars, while rotations act as a subgroup of the $U(1)$ monopole symmetry.

The single symmetric relevant operator which tunes the transition is the $SO(3)$ -invariant mass term $M^2 \phi^\dagger \phi$ for the scalars. For $M^2 \gg 0$, the scalars may be integrated out and we are left with a free $U(1)$ gauge theory. If there are no monopoles in the action, this leads to a Coulomb phase where the $U(1)$ monopole symmetry is spontaneously broken. However, in the setting of Neel-VBS, only a finite subgroup, e.g., \mathbb{Z}_4 or \mathbb{Z}_3 , is a microscopic symmetry, so generically there will be charge 4 or 3 (respectively) monopoles in the action, which are irrelevant at the critical point but destabilize the Coulomb phase, leading to a gapped phase where the rotation symmetry is spontaneously broken—this is the VBS phase.

For $M^2 \ll 0$, the scalars condense, and the $U(1)$ gauge field is Higgsed, with a vacuum moduli space $\mathbb{C}\mathbb{P}^1 = S^2$, which means the $SO(3)$ symmetry is spontaneously broken

and the system is in the Néel phase. This direct order-to-order transition may be explained for square, rectangular, or triangular lattices, which have a \mathbb{Z}_2 site-rotation symmetry, by the 't Hooft anomaly of $SO(3) \times U(1)$ [32,34,35]. This anomaly argument is a version of the Lieb-Schultz-Mattis (LSM) theorem. Whether there is actually a continuous phase transition remains controversial. See for instance [36–44]. For the honeycomb lattice, which has only \mathbb{Z}_3 site-rotation symmetry, there is no 't Hooft anomaly in the field theory and no LSM theorem for the lattice, so there is the possibility for an intermediate trivial gapped phase. Possible ground states for this phase were constructed in Refs. [45–47], although it is not clear how to reach them from the $\mathbb{C}\mathbb{P}^1$ model.

Applying the Thouless pump invariant to this study we learn some new conclusions, including for the honeycomb lattice. The $SO(3)$ -vector scalar mass terms $m^2 n_i = N_i$ correspond to Néel polarizing fields, so we will refer to the gapped state obtained by giving these terms a large coefficient as the polarized state. As we have argued, in the three-parameter phase diagram of the N_i , regardless of other local perturbations such as the precise value of M^2 or the strength of symmetry-allowed monopole operators, on a large S^2 there will be a Thouless pump for the unbroken monopole symmetry, in other words for the site-rotation symmetry \mathbb{Z}_p , $p = 2, 3, 4$, or 6 .

The Thouless pump implies that the phase diagram must have some diabolical locus for small \vec{N} . For instance, if we begin in the Néel phase, where $SO(3)$ is spontaneously broken, then for any nonzero \vec{N} the system is polarized, so the gapless Néel phase appears as just a single gapless point at $\vec{N} = 0$, described in the IR by an S^2 sigma model. For $M^2 \gg 0$, on the other hand, we will open up a VBS phase at $\vec{N} = 0$. There will be a phase transition at some nonzero \vec{N} into the polarized state.

As we vary M^2 between these two regimes, there must always be a phase transition somewhere in the slice of the phase diagram corresponding to a fixed M . See Fig. 2. This is true even on the honeycomb lattice, where one might go through an intermediate trivial phase for some M and zero polarizing field. In that case, by the spherical symmetry of the phase diagram, this trivial phase will still be separated from the polarized phase at large \vec{N} by a phase transition at some positive $|N|^2$. Near this phase transition, both nearby phases look trivial. This situation is fine-tuned in the sense that the N_i are the only $SO(3)$ -breaking perturbations. If we add other such perturbations, we can spoil the spherical symmetry of the phase diagram in Fig. 2, resulting in a generic (although still nontrivial) diabolical locus.

If we include in the action a level k Chern-Simons term for the dynamical $U(1)$ gauge field, then (5.11) implies that the Chern-Simons term produces a Hopf term with coefficient k for the parameters n_i [30]. As discussed in Ref. [30], such term results in the skyrmion being a fermion or boson when k is odd or even, in agreement with the spin of the monopole in the microscopic gauge theory. The computation can be generalized straightforwardly to $U(N)$ gauge theory coupled to two complex scalars.

Note that the 2+1d Thouless pump term coincides with that of two complex fermions in (3.4). For $k = 1$, the Hopf term coincides with the $\theta = \pi$ term for n_i in the system of

¹⁵A useful identity is

$$n_3(n_1 dn_2 dn_3 + n_2 dn_3 dn_1 + n_3 dn_1 dn_2) = dn_1 dn_2, \quad (5.10)$$

which can be derived using $n_1^2 + n_2^2 + n_3^2 = 1$ and $n_3 dn_3 = -n_1 dn_1 - n_2 dn_2$.

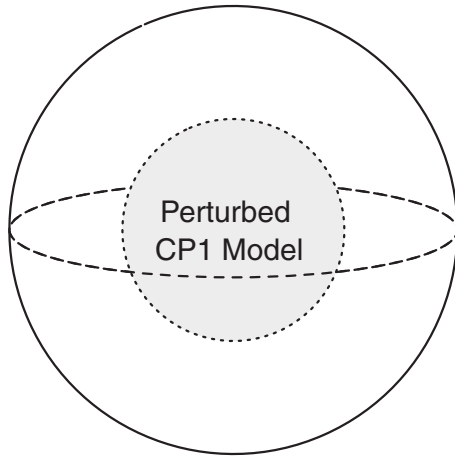


FIG. 2. The three parameter phase diagram of the $\mathbb{C}\mathbb{P}^1$ model with potential (5.7), at some value of M^2 and other perturbations not necessarily at the critical value, and where we have explicitly broken the $\text{SO}(3)$ symmetry by the three mass parameters $N_i = m^2 n_i$, which correspond to Néel polarizing fields in the Néel-VBS transition. The fate of the theory at the origin is unknown except in the large M^2 regions, but there is a Thouless pump for the $\text{U}(1)$ monopole charge on a sphere of large radius which protects a diabolical locus inside for all M^2 . This is true even after breaking $\text{U}(1)$ to a cyclic subgroup of odd order, for which there is no 't Hooft anomaly or LSM constraint. Note that the phase diagram is spherically symmetric so long as the N_i are the only $\text{SO}(3)$ -breaking perturbations.

two complex fermion. Thus the above computation provides a new consistency check for the duality between two complex fermions and $\text{U}(N)_1$ coupled to two complex Wilson-Fisher scalars [48,49],¹⁶

$$2 \text{ free Dirac fermions} \longleftrightarrow \text{U}(N)_1 + 2\phi. \quad (5.13)$$

On the other hand, the theory with a more general Chern-Simons level k may flow to interacting fixed points. In those models, there is a kind of Thouless pump, but the surrounding phase is a nontrivial TQFT, and the skyrmion binds a nontrivial anyon [25].

It is also suspected that there is an emergent $\text{SO}(5)$ symmetry in this model, where the two monopole operators $M_1 \pm M_{-1}$ and the n_i form an $\text{SO}(5)$ vector [43].¹⁷ If we consider perturbations by these operators, the theory describes a diabolical locus in a five-parameter phase diagram, protected by a WZW term of level 1 [51–53]. The protection of this diabolical locus holds without assuming any symmetry at all! The WZW term can be derived using from the Thouless pump above, and vice versa, using equivariant cohomology techniques for computing anomalies of WZW theories [54].

¹⁶The duality holds for $N \geq 2$ since otherwise the mass deformation in the scalar theory contains extra Goldstone mode [49]. We remark that the effective action discussed here cannot be removed by a well-defined local counterterm of the background field since it is not well-defined at $m^2 = 0$. Thus it must agree across the duality.

¹⁷There are also results from conformal bootstrap against a stable critical point with $\text{SO}(5)$ symmetry in $(2+1)d$ [50].

The above computation can be generalized to other dimensions, with the 0-form symmetry for the Thouless pump invariant replaced by higher-form symmetry. We will give an example in Sec. VB 1.

B. Gauge theories in 3+1d

1. $\text{U}(1)$ gauge theory with two scalars

Consider a $\text{U}(1)$ gauge theory in 3+1d coupled to two complex scalars of charge one with an $\text{SU}(2)$ -invariant potential (5.7). The theory has a magnetic $\text{U}(1)$ 1-form symmetry with a current $j^{\mu\nu} = \frac{1}{2\pi} \epsilon^{\mu\nu\rho\sigma} \partial_\rho a_\sigma$ where a is the gauge field [55]. Then by repeating the computation in Sec. VA 3 we find that in the presence of a background 2-form gauge field $B^{(2)}$ for the $\text{U}(1)$ 1-form symmetry, the effective action for the family of theories away from the point $m^2 = 0$ has the following 3+1d Thouless pump invariant

$$S_{\text{top}} = \frac{1}{8\pi} \int \epsilon^{\mu\nu\rho\sigma} \epsilon^{ijkl} B_{\mu\nu}^{(2)} n_i \partial_\rho n_j \partial_\sigma n_k d^4x. \quad (5.14)$$

The significance of this topological terms is as follows. For a fixed m^2 , the parameter space of the family of theories we consider is S^2 . Therefore it makes sense to consider skyrmionic strings: Static topologically nontrivial configurations of the parameter fields n^i which depend on two out of three coordinates on the space \mathbb{R}^3 and approach a constant at infinity. Such strings are classified by a skyrmion 1-form charge taking values in $\pi_2(S^2) = \mathbb{Z}$. The term (5.14) makes skyrmionic strings charged with respect to the 1-form $\text{U}(1)$ symmetry, so that their $\text{U}(1)$ -charge is identified with the skyrmion charge. This protects the gapless locus at $m^2 = 0$ provided 1-form $\text{U}(1)$ symmetry remains unbroken.

The stability of the $m^2 = 0$ point is also protected by an order 2 mixed 't Hooft anomaly between the magnetic 1-form symmetry and the $\text{PSU}(2) = \text{SO}(3)$ flavor symmetry.¹⁸ However, the previous argument using the Thouless pump invariant does not rely on $\text{SO}(3)$ symmetry and thus is still valid for perturbations that break $\text{SO}(3)$. Moreover, such an anomaly becomes trivial when we consider two copies of the system. The argument based on the effective action (5.14) shows that neither the number of copies nor the $\text{SO}(3)$ flavor symmetry are essential for the stability of the phase transition at $m^2 = 0$.¹⁹

¹⁸To see this, we can turn on background $\text{SO}(3)$ gauge field with nontrivial $w_2^{\text{SO}(3)}$ which is the obstruction to lifting the bundle to an $\text{SU}(2)$ bundle, then the Dirac quantization of the $\text{U}(1)$ gauge field a is modified to be

$$\oint \frac{da}{2\pi} = \frac{1}{2} \oint w_2^{\text{SO}(3)} \text{ mod } \mathbb{Z}. \quad (5.15)$$

Thus the coupling $\int B^{(2)} da / (2\pi)$ to background $B^{(2)}$ for the magnetic 1-form symmetry is no-longer 3+1d, and we find the theory lives on the boundary of the SPT phase

$$\pi \int_{5d} \frac{dB^{(2)}}{2\pi} \cup w_2^{\text{SO}(3)}. \quad (5.16)$$

¹⁹The same comment applies to the 2+1d theory discussed in Sec. VA 3 whose discrete 't Hooft anomalies were studied in [32].

VI. THE BULK-BOUNDARY CORRESPONDENCE FOR SPHERES

We will show that given any boundary condition of a d -dimensional system with a higher Berry number on a parameter space with the topology of S^{d+2} , there is at least one parameter value where the boundary gap closes, indicating either gapless edge modes or an edge degeneracy. In the case that the parameter space is $\mathbb{R}^{d+2} \setminus \{0\}$, repeating the argument on successive spheres we find an arc connecting the origin to infinity, along which the boundary gap closes. Furthermore, if we consider varying parameters near the boundary, along a small S^{d+1} which links the singularity, we find the boundary theory has a higher Berry number on the linking S^{d+1} . This by itself is not anomalous, but the fact that this S^{d+1} lies in a larger parameter space where it is homotopically trivial requires the existence of the bulk. Compare Secs. II A 2.

To see there must be a boundary singularity, we will show the partition function vanishes at a certain point by studying the winding number of its phase. Indeed, far from the boundary, but still over a wide-enough region so that we do not cause the bulk gap to close, we can cause the phase of the partition function to wind a number equal to the higher Berry number. We do this by choosing spatially varying parameters which draw a “hoop” in parameter space, such that along a 1-parameter family of such hoops beginning and ending at a chosen configuration of constant parameter, we envelop the target space once. See Fig. 3.

So long as the partition function does not vanish, this winding number is a quantized integer. The nonvanishing of the partition function is ensured by the uniform gap assumption. On the other hand, since we are on a space-time with boundary, we can move the hoop region to the edge, absorbing it and unwinding the one-parameter family so constructed. At the end of this, the winding number must be zero. Therefore the partition function must vanish somewhere during this process, so the assumption of the uniform gap must be violated by the boundary.

We can perform this unwinding more systematically by using the boundary to retract each hoop. See Fig. 4. This defines a two-parameter family $\phi(u, v)$ of backgrounds. Let u denote the parameter of the original family (which may be considered a circular parameter) and v the parameter of the unwinding. We have argued there is some u_*, v_* where the partition function vanishes, hence there is some parameter value p_* in the image of $\phi(u_*, v_*)$ where the boundary is either degenerate or gapless. Typically there could be several such points.

In a small neighborhood of such a point, we may integrate out the bulk and consider p_* as a diabolical point for the effective boundary theory. This implies that such points are isolated in the absence of extra symmetries. Furthermore, they are protected by a higher Berry number of the boundary, which we can compute as follows.

The special point p_* is isolated so there is some $\epsilon > 0$ such that there are no other boundary diabolical points within the parameter values in the image of the strip $v_* - \epsilon < v < v_* + \epsilon, u \in S^1$. We form a 1-parameter family of backgrounds by composing the 1-parameter families $\phi(u, v_* + \epsilon/2)$ and $\phi(2\pi - u, v_* - \epsilon/2)$ (these are composable because all

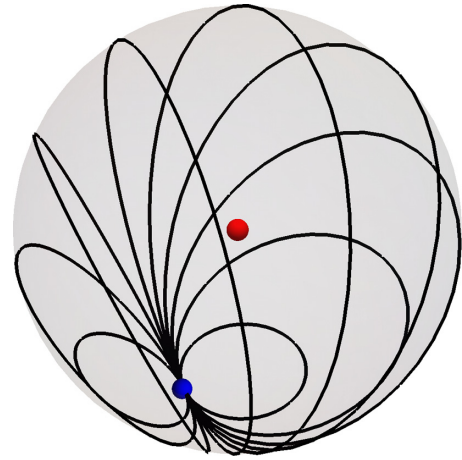


FIG. 3. A 1-parameter family of backgrounds on space-time \mathbb{R}^{d+1} (or a ball B^{d+1}) useful for measuring the higher Berry number around a parameter space S^{d+2} and for arguing the bulk-boundary correspondence. Here we have drawn the case $d = 0$, with each hoop depicting the image of the parameter map ϕ . We fix the parameter values at space-time infinity (or sufficiently near the boundary) at the blue point, where all hoops meet, and varying the parameters in a compact but large region so that over the 1-parameter family, the parameter values wind around the sphere once. Note that this family begins and ends at a background where the parameters are constant, at the blue point. Thus the partition function in this process will return to itself. Assuming a uniform energy gap over the sphere implies that this partition function never vanishes, hence the winding number of the phase of the partition function over this process is a quantized topological invariant, which we identify with the higher Berry number by the effective action (2.11) and its higher-dimensional generalizations. The red point is a hypothetical diabolical point which sources the Berry curvature.

families begin and end at the blue point in Fig. 4). This family is constructed so that the parameter values on the boundary wind once around the boundary diabolical point p_* . Further, because we are in the strip away from any other diabolical points, the partition function is nonvanishing in this family. Finally, the winding number jumps between the families $\phi(u, v_* + \epsilon/2)$ and $\phi(u, v_* - \epsilon/2)$, and we have composed them with different orientations in u , so the one so constructed has nonzero winding number. This winding number protects the boundary diabolical point by our usual arguments, and thus is very analogous to a bulk higher Berry number.

There is an important subtlety, however, which distinguishes this situation from cases where there is no bulk. Indeed, if we sum up the winding numbers defined above over all the boundary diabolical points, we get something nonzero—in fact one can readily see we get the bulk higher Berry number! This “index theorem” is an aspect of anomaly in-flow for the higher Berry phase. Indeed, without the bulk, for a compact parameter space it is easy to see the total index of diabolical points (summing their higher Berry numbers) must be zero.

To put it another way, our one-parameter family we defined above to encircle the boundary diabolical point p_* is homotopic without changing the boundary values to one where the bulk parameter values lie only in the special strip. On

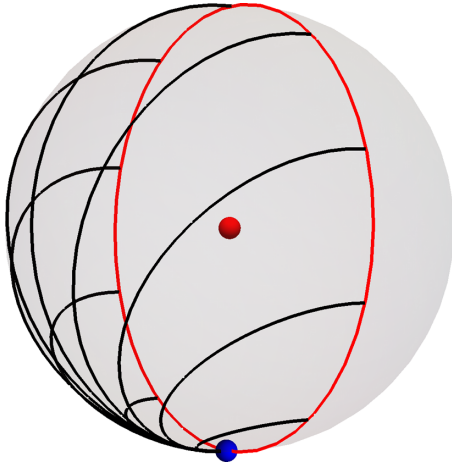


FIG. 4. A two-parameter family of backgrounds where the boundary is used to unwind the one-parameter family in Fig. 3, depicted at the point where the latter family is half-way “absorbed” into the boundary. We have drawn the image of the parameter values at the boundary as a red circle. Because this family of families interpolates between a trivial family and one for which the phase of the partition function winds, the partition function must vanish for some value of the two parameters, which implies the gap must close at the boundary at that value. This occurs precisely where the red circle crosses a boundary diabolical point. We may define a boundary Berry number for such points. We find that the sum of these boundary Berry numbers equals the bulk Berry number.

the other hand we could’ve defined a one-parameter family with the same boundary values but where the bulk parameter values lie *outside* the strip, such that the two one-parameter families are complementary, in that we can glue them together along the boundary to form something homotopic to our bulk 1-parameter family in Fig. 3. Thus, while these one-parameter families have precisely the same boundary parameters, their winding numbers differ by the bulk higher Berry number.

This confirms a point of view in Ref. [9], which is that theories with anomalies in the space of coupling constants are typically presented with a noncompact parameter space, and while the ground state is trivially gapped at infinity, there is some winding that prevents one from compactifying the parameter space. Here we see that we expect such compactification is possible iff one introduces a bulk. We comment on the applications of these arguments to interfaces in Appendix C.

This discussion can be extended to systems with a global symmetry group G . We study families parameterized by S^{d+1-k} , $k \geq 0$, which are characterized by a generalized Thouless pump for which some k -space-time-dimensional G -SPT Σ_k is carried by the skyrmion. After compactifying along a k -dimensional test manifold with background gauge field for this SPT, we can reduce the detection of this SPT on the skyrmion by measuring the phase (or winding number thereof) of the sphere partition function. Introducing a boundary and using it to unwind the skyrmion as in Fig. 4, we again find there is a boundary diabolical point where this partition function vanishes. If $d+1-k=1$, at this point the boundary looks like a phase transition between G -SPTs. If $d+1-k > 1$, the boundary diabolical point is associated

with a Thouless pump over an S^{d-k} . As before, the sum of the G -SPT classes associated with each of these diabolical points must be Σ_k .

VII. CLASSIFICATION OF DIABOLICAL LOCI

In this section, we discuss a conjectural classification of higher Berry numbers and diabolical loci. Consider the space \mathfrak{M}_d of infrared-trivial systems, that is, those with a unique, gapped ground state and no topological degrees of freedom. The connected components of this space [elements of $\pi_0(\mathfrak{M}_d)$] are the short-range entangled (SRE) phases of matter. It is by now well accepted, although it remains unproven, that in the case of bosons (respectively fermions), $\pi_0(\mathfrak{M}_d)$ can be expressed in terms of cobordism groups of oriented (respectively spin) closed manifolds [56,57].

In fact, there is a much stronger conjecture, motivated by TQFT, which identifies the entire homotopy type of \mathfrak{M}_d with that of a space in a certain cobordism spectrum [58]. The fact that the homotopy type forms a loop-spectrum is very powerful, and allows us to immediately derive the classification of SRE phases with global symmetries (i.e., SPT phases) [59–64] as well as classify *families* of SRE phases [65], which is what we are interested in.

For instance, a basic consequence of the existence of a loop-spectrum for SRE phases is the isomorphism

$$\pi_k(\mathfrak{M}_d) = \pi_0(\mathfrak{M}_{d-k}), \quad (7.1)$$

which says that a family of infrared-trivial systems parametrized by S^k can be characterized by a generalized Thouless pump which pumps a $d-k$ -dimensional SRE phase to the boundary.

We see that if d and k are increased together, $d-k$ does not change, so for each $m = d-k$ there is a *series* of families associated to an m -dimensional SRE phase. The Abanov-Wiegmann “A” series in (C21) and (C20), for instance, is associated to the 0+1d fermionic SRE “phase” characterized by the fact that the ground state has fermion number 1 (or 1 mod 2 if only the fermion parity is conserved.) This phase is the generator of $\pi_0(\mathfrak{M}_1) = \Omega_{\text{spin}}^1 = \mathbb{Z}$ in the U(1) conserving case or $\pi_0(\mathfrak{M}_1) = \Omega_{\text{spin}}^1 = \mathbb{Z}_2$ in the case with just fermion parity. The theory studied in Sec. IV C with the axion coupling corresponds to $m=2$, with $\pi_0(\mathfrak{M}_2) = \mathbb{Z}$ generated by the $p+ip$ superconductor.

The “B” series in (C1), (C2) corresponds $m=-1$. Since (-1) -dimensional systems do not make sense, one cannot associate this series to an SRE phase. However, one can set $k=1$ and associate it to a noncontractible loop of SRE “systems” in 0+0 dimensions. A “system” in zero dimensional space-time is characterized by its partition function, which is a complex number. For an SRE “system,” this number is nonzero. The noncontractible family in question is the one where the phase of the partition function winds once around the origin of the complex plane. Although this series has $m=-1$, geometrically it corresponds to the generator of the bordism group of spin 0-manifolds, which is \mathbb{Z} . This degree shift is a consequence of Anderson duality applied to the bordism spectrum [58]. With time reversal symmetry, the partition function is real, so there is no integer

Berry number, but for $m = 0$ there are two SRE “phases,” the trivial one and the nontrivial one. The nontrivial one gives rise to a series of codimension- $d + 2$ diabolical loci of time-reversal-invariant systems in spatial dimension d . The one for $d = 0$ was observed already by von Neumann and Wigner [2] and corresponds to the holonomy of the Berry connection being -1 . Geometrically this series arises from the generator of the bordism group of unoriented points, which is \mathbb{Z}_2 . It does not get shifted in degree by Anderson duality.

These considerations are the interacting analog of the work of Teo and Kane [11] on topologically protected defects in gapped systems of free fermions in d spatial dimensions. They argued that the classification of such defects depends only on $\delta = d - k$, where $k + 1$ is the codimension of the defect. A defect of codimension $k + 1$ in a system of free fermions can be created by making the system’s parameters depend on the coordinates of S^k which surrounds the defect. Therefore a topologically protected defect corresponds to a nontrivial element of $\pi_k(\mathcal{K}_d)$, where \mathcal{K}_d is the space of gapped systems of free fermions in d dimensions. Since the spaces \mathcal{K}_d form a loop spectrum too (the K-theory spectrum [66–68]), one finds again that $\pi_k(\mathcal{K}_d) = \pi_0(\mathcal{K}_{d-k})$. Thus the classification of topologically protected defects depends only on $d - k$ and can be obtained from the periodic table of topological insulators and superconductors [66,67] by a shift $d \mapsto d - k$. In particular, Bott periodicity in K theory implies that the classification of defects is periodic in k with period 2 or 8 depending on the symmetry class.²⁰

If one views $k + 1$ as the codimension of a spatial defect, then k cannot be larger than d . But if one views k as the codimension of a diabolical point in the phase diagram, then k can be arbitrarily large. Bott periodicity thus implies that in a d -dimensional system of free fermions the codimension of a diabolical point can be arbitrarily high. In contrast, in the interacting case there are no diabolical points of codimension greater than $d + 3$. Interactions destabilize all diabolical points of higher codimension. An argument for this is sketched in Appendix D.

For a general parameter space P , the classification implies that families of d -space-dimensional infrared-trivial systems parametrized by P are classified by a certain cobordism group for closed $d + 1$ -manifolds equipped with a map to P , which we

$$\left\{ \begin{array}{l} \text{families of infrared-trivial systems} \\ \text{of bosons or fermions parameterized by } P \end{array} \right\} \simeq \Omega_{\text{SO or Spin}}^{d+1}(P). \tag{7.2}$$

There is a “supercohomology” approximation to this classification which was described in [71]. In particular, if P supports fermion parity Thouless pumps, it can affect the quantization of the WZW term to be fractional relative to the quantization on a sphere. If there is a global symmetry G , we study instead the cobordism group of $P \times BG$, with appropriate twists [57].

²⁰Some of these K-theory invariants, such as the higher Chern numbers, have been interpreted in terms of curvatures of higher form connections on the Brillouin zone in Refs. [69,70].

The group on the right hand side is interpreted as cobordism-invariant effective actions for a $d + 1$ -dimensional space-time with a map to P , *i.e.* a spatiotemporally varying background.

The correspondence between families of theories and theories equipped with varying parameters, or parameters promoted to fields, is quite general, and has been used in many contexts in physics. There is mathematical evidence for this correspondence. For instance, it has been observed in Ref. [65] that the Baez-Dolan-Lurie cobordism hypothesis [72,73] implies an equivalence between families of TQFTs parametrized by a space P and TQFTs for manifolds equipped with a map to P . In the general case, looking at the space of systems \mathfrak{M}_d^T whose IR limit is a $d + 1$ -dimensional TQFT \mathcal{T} , one expects for $0 < k < d + 2$, $\pi_k(\mathfrak{M}_d^T)$ is the group of invertible $d - k$ -dimensional topological defects in \mathcal{T} , $\pi_{d+2}(\mathfrak{M}_d^T) = \mathbb{Z}$, and higher homotopy groups vanish. This is relatively well-understood for $d \leq 2$ [74,75], and an understanding is slowly emerging in higher dimensions [76].

Finally, there is the question of how these various spaces are embedded into the space \mathbf{M}_d of all low-energy theories in d space dimensions. This space may have several components, labeled by anomalies for instance. Our spaces \mathfrak{M}_d are embedded in the component \mathbf{M}_d^0 of “anomaly free” low-energy theories. One might expect this space is homotopic to the space of local lattice Hamiltonian models, a convex space, hence that \mathbf{M}_d^0 is contractible. This would be satisfying, since it implies that every nontrivial homotopy group of \mathfrak{M}_d (or any other stratum) corresponds to some diabolical point on a deeper stratum inside.

Note added. At the same time as this paper, another paper on a closely related subject appeared on the arXiv [77]. That paper studied a phenomenon they dubbed “vacuum crossing,” wherein traversing a noncontractible loop in a region of parameter space with multiple vacua (e.g., in a spontaneous symmetry breaking phase), the system undergoes an adiabatic transformation from one ground state to another. If M is a region of parameter space where the system has n degenerate ground states, then this adiabatic transformation is encoded in a permutation representation $\pi_1(M) \rightarrow S_n$. If these n degenerate ground states form a representation of some spontaneously broken symmetry G , then this permutation must be valued in the commutant of G . If they are associated with anyons, such as when considering ground states of a 2+1d TQFT on a torus, then this permutation representation must respect the braiding action.

One point of view which encompasses both vacuum crossing and the invariants discussed in this paper is to say that for any subset M of parameters where the system flows to the same IR fixed point \mathcal{T} , the *homotopy type* of M acts upon \mathcal{T} as a (higher) symmetry, so that for $0 < k < d + 2$, any element of $\pi_k(M)$ is associated with a symmetry defect of \mathcal{T} (which may be spontaneously broken), while $\pi_{d+2}(M)$ is associated with higher Berry phases. If M is homotopy equivalent to S^1 , this is the permutation representation mentioned above. If M is homotopy equivalent to S^{d+2} , where d is the space dimension, then this is the higher Berry phase. Taking $M = \mathfrak{M}^T$ the space of all theories which flow to \mathcal{T} , there is a universal such action. It would be very interesting to try to understand this in detail. .

ACKNOWLEDGMENTS

We thank Dominic Else, Tobias Holder, Raquel Queiroz, Nathan Seiberg, Ruben Verresen, and Adar Sharon for discussions. A.K. is grateful to Lev Spodyneiko for a collaboration on a closely related project. The work is supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award No. DE-SC0011632, and by the Simons Foundation through the Simons Investigator Award.

APPENDIX A: TOPOLOGICAL ACTIONS FOR FREE FERMIONS IN d DIMENSIONS

1. Wess-Zumino-Witten terms

Consider N free massive fermions in $d + 1$ space-time dimension. The mass term in the Lagrangian has the form $m\bar{\psi}^a P_{ab}\psi^b$, where we parameterized the mass matrix $\mathbf{m} = m\mathbf{P}$ by an overall scale $m > 0$ and a dimensionless matrix \mathbf{P} with entries P_{ab} .

In odd space-time dimensions, Lorenz invariance forces the mass matrix to act trivially on the spinor indices. Unitarity then requires \mathbf{P} to be hermitian. The physical masses of fermions are m times the absolute values of the eigenvalues of this matrix. Thus the model describes massive fermions if and only if \mathbf{P} nondegenerate. If we have N Dirac fermions, we can diagonalize \mathbf{P} with a $U(N)$ transformation, and all eigenvalues are nonzero. We can deform them all to ± 1 while keeping them nonzero. Hence we can assume that $\mathbf{P}^2 = 1$.

In even space-time dimensions we can write $\mathbf{P} = \mathbf{P}_1 + i\gamma^{0\dots d}\mathbf{P}_2$ with hermitian $\mathbf{P}_1, \mathbf{P}_2$ which act trivially on the spinor indices, and with $\gamma^{0\dots d}$ the chirality operator. In the cases of interest to us, these two matrices commute (in fact, \mathbf{P}_1 is always a scalar matrix). For N flavors of fermions a $U(N)$ rotation $\psi \mapsto U\psi$ can diagonalize both \mathbf{P}_1 and \mathbf{P}_2 . The physical masses of fermions squared are m^2 times the eigenvalues of $\mathbf{P}_1^2 + \mathbf{P}_2^2$. If we deform all masses to be m , the matrix \mathbf{P} satisfies $\mathbf{P}^\dagger\mathbf{P} = 1$. The conclusion is that for all d we can assume $\mathbf{P}^\dagger\mathbf{P} = 1$.

The variation of the effective action with respect to the parameter \mathbf{m} is

$$\delta S_{\text{eff}} = \text{Tr} \delta \mathbf{m} (D^\dagger D)^{-1} D^\dagger, \quad (\text{A1})$$

where $D = i\partial + i\mathbf{m}$, and $D^\dagger D = -\partial^2 + \mathbf{m}^\dagger \mathbf{m} - \partial \mathbf{m}$. The topological term in the effective action is

$$\begin{aligned} \delta S_{\text{eff}} &= -i \text{Tr} m^{d+3} \delta \mathbf{P} (-\partial^2 + m^2)^{-1} \cdot ((-\partial^2 + m^2)^{-1} \partial \mathbf{P})^{d+1} \mathbf{P}^\dagger \\ &= -i \theta_0 \int \text{Tr} (\delta \mathbf{P} (\partial \mathbf{P})^{d+1} \mathbf{P}^\dagger), \end{aligned} \quad (\text{A2})$$

where in the last line we used $\mathbf{P}^\dagger\mathbf{P} = 1$, and

$$\begin{aligned} \theta_0 &= \int \frac{d^{d+1}p}{(2\pi)^{d+1}} \frac{m^{d+3}}{(p^2 + m^2)^{d+2}} = i \frac{\pi^{(d+1)/2}}{(2\pi)^{d+1}} \frac{\Gamma(3/2 + d/2)}{\Gamma(d+2)} \\ &= \frac{i}{4^{d+1} \pi^{d/2} \Gamma\left(\frac{d+2}{2}\right)}. \end{aligned} \quad (\text{A3})$$

We can perform the trace over spinor indices using the identities

$$\begin{aligned} \text{odd } d : \quad & \text{Tr} (\gamma^{i_1} \gamma^{i_2} \dots \gamma^{i_{d+1}} \gamma^{0\dots d}) \\ &= i(-2i)^{[(d+1)/2]} \epsilon_{i_1, i_2, \dots, i_{d+1}}, \\ \text{even } d : \quad & \text{Tr} (\gamma^{i_1} \gamma^{i_2} \dots \gamma^{i_{d+1}}) = i(-2i)^{[(d+1)/2]} \epsilon_{i_1, i_2, \dots, i_{d+1}}, \end{aligned} \quad (\text{A4})$$

which can be derived from cyclic property of trace and anticommutation relation of the gamma matrices. In odd space-time dimension (even d) the result is the following Wess-Zumino-Witten term:

$$\theta_0 \frac{(-2i)^{d/2}}{d+2} \int \text{Tr} (\mathbf{P} (d\mathbf{P})^{d+2}). \quad (\text{A5})$$

In even dimensions, one can perform a similar computation with $\mathbf{P} = \mathbf{P}_1 + i\gamma^{0\dots d}\mathbf{P}_2$. We will not give a closed form expression here.

2. Topological terms with a background gauge field

One can perform a similar computation with an additional background gauge field A for an internal Lie group symmetry G . We will assume the mass matrix is invariant under G symmetry. The expansion of $(D^\dagger D)^{-1}$ is modified to be

$$(D^\dagger D)^{-1} = \sum_s (-\partial^2 + m^2)^{-s-1} \sum_r (-i\partial A)^r (m\partial \mathbf{P})^{s-r} + \dots, \quad (\text{A6})$$

where the last term should be understood as a sum over different orderings of ∂A and $\partial \mathbf{P}$. The terms contributing to a $(d+1)$ -form in the effective action correspond to $2r + s - r = d+1$, namely $s = d+1 - r$.

The variation of the effective action thus contains

$$\sum_r (-i)^{r+1} \theta_{r,d} \text{Tr} \delta \mathbf{P} ((\partial A)^r (\partial \mathbf{P})^{d+1-2r}) \mathbf{P}^\dagger, \quad (\text{A7})$$

where

$$\begin{aligned} \theta_{r,d} &= \int \frac{d^{d+1}k}{(2\pi)^{d+1}} \frac{m^{d-2r+3}}{(k^2 + m^2)^{d-r+2}} \\ &= i \frac{\pi^{(d+1)/2}}{(2\pi)^{d+1}} \frac{\Gamma(3/2 + d/2 - r)}{\Gamma(d+2 - r)}. \end{aligned} \quad (\text{A8})$$

a. Odd dimension

In odd space-time dimensions (even d) we can use the trace identity (A4) to obtain the following topological terms in the partition function:

$$\exp \left(i \int_{M_{d+2}} \text{Tr} e^{F/(2\pi)} \chi(\mathbf{P}) \right), \quad (\text{A9})$$

where $\partial M_{d+2} = M_{d+1}$ and

$$\chi(\mathbf{P}) \equiv \sum_s \frac{\sqrt{\pi} i^{s+d/2}}{(2\pi)^{s-d/2}} \frac{\Gamma(s - d/2 + 1/2)}{(2s - d)!} \text{Tr} ((d\mathbf{P})^{2s-d} \mathbf{P}). \quad (\text{A10})$$

b. Even dimension

In even space-time dimensions (odd d), we express the mass matrix as $\mathbf{P} = \mathbf{P}_1 + i\gamma^{0\dots d}\mathbf{P}_2$, where $\mathbf{P}_1, \mathbf{P}_2$ act only on

the flavor indices. We can evaluate the trace in the variation of the effective action by applying the identity (A4) that requires a single overall $\gamma^{0,\dots,d}$. Consider the rightmost $\partial\mathbf{P}$ in (A7) that contributes $\gamma^{0,\dots,d}$

$$\text{Tr}(\mathbf{P}^\dagger \delta\mathbf{P} \dots (i\partial\gamma^{d+2}\mathbf{P}_2)f_k) \quad (\text{A11})$$

where f_k is a product of powers of (∂A) and $(\partial\mathbf{P}_1)$ that has a total of k gamma matrices. Moving $\gamma^{0,\dots,d}$ to the right end introduces $(-1)^k$. We repeat this procedure for the next one to the left contributing $\gamma^{0,\dots,d}$ to the effective action until

$$\begin{aligned} \delta\chi(\mathbf{P}) \equiv & \sum_s \frac{(s-(d+1)/2)!}{(2\pi)^{s-(d+1)/2}} i^{s-\frac{d+1}{2}} \left\{ \sum_{j=0}^{[s-(d+1)/2]} \frac{1}{(2j)!(2s-d-2j-1)!} \text{Tr}(\mathbf{P}_1^\dagger \delta\mathbf{P}_2 - \mathbf{P}_2^\dagger \delta\mathbf{P}_1)(d\mathbf{P}_1)^{2s-d-1-2j}(d\mathbf{P}_2)^{2j} \right. \\ & \left. + \sum_{j=0}^{[s-d/2-1]} \frac{1}{(2j+1)!(2s-d-2j-2)!} \text{Tr}(\mathbf{P}_2^\dagger \delta\mathbf{P}_2 + \mathbf{P}_1^\dagger \delta\mathbf{P}_1)(d\mathbf{P}_1)^{2s-d-2j-2}(d\mathbf{P}_2)^{2j+1} \right\}. \quad (\text{A13}) \end{aligned}$$

For instance, in even space-time dimensions $(d+1)$ the effective action of a single massive fermion with $P_1 = \cos\alpha$, $P_2 = \sin\alpha$ contains the following Wess-Zumino term:

$$i \frac{1}{(2\pi)^{(d+1)/2}((d+1)/2)!} \int_{M_{d+2}} d\alpha \text{Tr} F^{(d+1)/2}. \quad (\text{A14})$$

This is exactly as expected from the Atiyah-Patodi-Singer index theorem [78] that relates the η -invariant of the Dirac operator with $\text{Tr} e^{F/2\pi}$.

APPENDIX B: LUTTINGER LIQUID DEFORMATIONS OF THE 1D THOULESS PUMP

Both the Thouless pump and the diabolical point considered in Sec. II A admit a simple description in terms of a compact boson. Let θ , ϕ be 2π -periodic dual fields describing a 1+1d compact boson. In our convention at radius R , the vertex operator $\exp(in\theta + iw\phi)$ has dimensions

$$(h, \bar{h}) = \left(\frac{1}{2}(n/R + Rw/2)^2, \frac{1}{2}(n/R - Rw/2)^2 \right).$$

The free fermion theory is equivalent to $R=2$, with charge conservation corresponding to the U(1) shift symmetry of θ .

There are two U(1)-invariant and Lorenz-invariant relevant operators, namely $\cos\phi$ and $\sin\phi$. These correspond to the two mass parameters of the Dirac fermion. The axial U(1) symmetry acts as the shift symmetry of ϕ . If we turn on some linear combination of these mass terms, ϕ becomes fixed in the ground state due to a periodic potential. While all these ground states look trivial on their own, there is nevertheless a winding as we vary the argument of the mass around the origin, witnessed by the fixed value of ϕ winding around. Because $\partial_x\phi$ is the current for the θ -shift symmetry, our U(1), we immediately derive the topological current of Sec. II A.

Suppose we deform the family of theories by keeping $R=2$ for sufficiently large $|M|$, but making it larger near $M=0$. If at some $|M|$, the radius R exceeds $2\sqrt{2}$, both $\cos\phi$ and

all $\gamma^{0,\dots,d}$ are moved to the right. If the number of $\gamma^{0,\dots,d}$ is $(2j+1)$, there is additional factor $(-1)^j$ from bringing $\gamma^{0,\dots,d}$ through $\partial\mathbf{P}_2$. Applying the trace identity (A4) erases the gamma matrices and f_k becomes a k -form. Then we move $d\mathbf{P}_2$ across the k -form to the right end, which introduces another sign that compensates $(-1)^k$. This gives

$$\exp\left(i \int_{M_{d+2}} \text{Tr} e^{F/(2\pi)} \chi(\mathbf{P})\right), \quad (\text{A12})$$

where the variation of $\int \chi(\mathbf{P})$ gives the boundary terms

$\sin\phi$ become irrelevant and the model flows to a Berezinskii-Kosterlitz-Thouless (BKT) phase with an algebraic decay of correlators. Thus the diabolical point gets resolved into an island of the BKT phase. On the boundary of this island the model undergoes a phase transition from the gapless BKT phase to the trivial gapped phase.

Alternatively, one can make R smaller than 2. As we tune R past the self-dual radius $R=\sqrt{2}$, two new operators become relevant: $\cos 2\phi$ and $\sin 2\phi$. In a generic 2-parameter phase diagram, this will cause the gapless Luttinger liquid point at zero mass to become destabilized. To see what happens, without loss of generality we may add a uniform perturbation $\cos 2\phi$ to the theory at each point in the 2-parameter phase diagram. Along the $\cos\phi$ axis of the phase diagram, there is an accidental charge conjugation symmetry

$$C : \begin{cases} \phi \mapsto -\phi \\ \theta \mapsto -\theta \end{cases}.$$

Near the origin, where $\cos 2\phi$ dominates $\pm\cos\phi$, C is spontaneously broken, leading to a first order line. By studying the competition between these two potentials, we find this first order line ends at two Ising critical points at $\cos 2\phi \pm 4\cos\phi$. This is depicted in Fig. 5.

If we are willing to break the U(1) symmetry, then we can add $\cos 2\theta$ ($\cos\theta$ violates fermion parity and is not allowed). This creates a similar phase diagram as in the figure, with a first order line along the C -invariant axis ending in two Ising critical points.

If one studies this theory as describing a system of bosons, meaning we can also break the \mathbb{Z}_2 shift symmetry of ϕ , then we can consider the four-parameter phase diagram of $\cos\phi$, $\sin\phi$, $\cos\theta$, and $\sin\theta$. These operators form an SO(4) vector at the SU(2)₁ radius $R=\sqrt{2}$. This implies that in that phase diagram the gapless point is isolated at the origin. The SU(2)₁ WZW term means that the higher Berry number is 1 for the 3-sphere which links the origin. See also Ref. [79].



FIG. 5. A first-order line ending at two Ising critical points, protected by a U(1) symmetry with a Thouless pump invariant encircling the diabolical locus (the interval and its endpoints). Without the U(1) symmetry, the two points can come together and annihilate, leaving a phase diagram which is trivially gapped everywhere. With U(1), however, they can only merge into a gapless Luttinger liquid point. There appears to be no way to rid the phase diagram of gapless points by U(1)-symmetric perturbations. The same conclusions hold even after breaking U(1) down to just the fermion parity \mathbb{Z}_2 subgroup.

APPENDIX C: DIMENSIONAL REDUCTION FOR INTERFACES

1. Interfaces for the Abanov-Wiegmann series

The Abanov-Wiegmann “B” series model in an odd space dimension d is defined as a free theory of $2^{(d+1)/2}$ flavors of $2^{(d+1)/2}$ -component complex fermions. The action has a $\text{Spin}(d+2)$ flavor symmetry, and the fermions transform in its spinor representation. If we denote the generators of the $(d+2)$ -dimensional Clifford algebra Γ^a , where $a = 1, \dots, d+2$, the Lagrangian has the form

$$\mathcal{L}_d^{\text{odd}} = -i\bar{\psi}_j \gamma^\mu \partial_\mu \psi_j - iM^0 \bar{\psi}_j \psi_j - M^a \bar{\psi}_j \gamma^{0, \dots, d} \Gamma_{jk}^a \psi_k, \quad (\text{C1})$$

where $\gamma^{0, \dots, d}$ is the chirality operator. Similarly, in an even space dimension d , we use $2^{d/2+1}$ flavors of $2^{d/2}$ -component complex fermions transforming in the spinor representation of $\text{Spin}(d+3)$, with Clifford algebra generators denoted Γ^a , $a = 1, \dots, d+3$ and the Lagrangian

$$\mathcal{L}_d^{\text{even}} = -i\bar{\psi}_j \gamma^\mu \partial_\mu \psi_j - iM^a \bar{\psi}_j \Gamma_{jk}^a \psi_k. \quad (\text{C2})$$

This parametrization is related to the one in (3.8) by $n_i \sin \nu = M^i$, $i = 1, \dots, d+2$, $m \cos \nu = M^{d+3}$.

Each of these models has a parameter space \mathbb{R}^{d+3} spanned by the M 's, with a diabolical point at the origin and massive elsewhere, giving rise to a WZW term of level 1 for the unit sphere $S^{d+2} \subset \mathbb{R}^{d+3}$.

We will show this by studying interfaces in the mass parameter space, which we realize as paths. First we consider the even dimensional case $d = 2n$, $n > 0$, with 2^{n+1} complex

fermions transforming under the flavor group $\text{Spin}(d+3)$ with the action (C1). We can choose a basis where one of the $\text{Spin}(d+3)$ generators, which we take to be Γ^{d+3} , is of the diagonal form

$$\begin{bmatrix} \mathbb{1}_{2^n \times 2^n} & 0 \\ 0 & -\mathbb{1}_{2^n \times 2^n} \end{bmatrix} = \sigma^z \otimes \mathbb{1}_{2^n \times 2^n}, \quad (\text{C3})$$

while $\Gamma^{0, \dots, d}$ is of the form

$$\begin{bmatrix} 0 & \mathbb{1}_{2^n \times 2^n} \\ \mathbb{1}_{2^n \times 2^n} & 0 \end{bmatrix} = \sigma^x \otimes \mathbb{1}_{2^n \times 2^n}, \quad (\text{C4})$$

where $\sigma^x, \sigma^y, \sigma^z$ are the usual Pauli spin matrices. In this basis, $M^{d+3} \bar{\psi} \Gamma^{d+3} \psi$ looks like a diagonal mass term. We consider a Jackiw-Rebbi-like domain wall setup where M^{d+3} varies along the coordinate x_d from -1 for sufficiently negative x_d to 1 for sufficiently positive x_d . As is well-known, each of the 2^{n+1} fermions will contribute a massless chiral mode localized on the wall. With Γ^{d+3} as above, we find normalizability constrains the zero modes to satisfy

$$\gamma^d \psi_j = \psi_j \quad 1 \leq j \leq 2^n \quad (\text{C5})$$

$$\gamma^d \psi_j = -\psi_j \quad 2^n + 1 \leq j \leq 2^{n+1}. \quad (\text{C6})$$

The resulting low-energy theory thus consists of 2^n complex *nonchiral* fermions localized on the wall we describe as follows. If we choose a basis for the spinor labels where γ^d has the diagonal form (C3), we have

$$\psi_j = \begin{bmatrix} \psi_j^+ \\ \psi_j^- \end{bmatrix}, \quad (\text{C7})$$

where ψ_j^\pm is a 2^n -component spinor. We define the 2^n spinors on the wall (which have the same number of components as the ψ but there are half as many of them) by

$$\chi_j = \begin{bmatrix} \psi_j^+ \\ \psi_{2^n+j}^- \end{bmatrix}, \quad \chi_j^+ = \psi_j^+ \quad \chi_j^- = \psi_{2^n+j}^- \quad 1 \leq j \leq 2^n. \quad (\text{C8})$$

We observe that Γ^{d+3} acts the same as γ^d on these spinors. Further, the interaction in (C1)

$$M^a \bar{\psi}_j \Gamma_{jk}^a \psi_k \quad (\text{C9})$$

ouples ψ_j^\pm 's of both opposite γ^d and Γ^{d+3} . It restricts to an interaction among the χ_j . To express this interaction, we first observe that in our chosen basis, with Γ^{d+2} of the form (C4),

$$M^{d+2} \bar{\psi}_j \Gamma_{jk}^{0, \dots, d} \psi_k = M^{d+2} ((\chi_j^+)^{\dagger} \chi_k^- + (\chi_j^-)^{\dagger} \chi_k^+) \quad (\text{C10})$$

is the diagonal mass term. The rest of the generators of $(d+3)$ -dimensional Clifford algebra can be written in a basis compatible with (C3) and (C4) by letting

$$\Gamma^a = \sigma^y \otimes \tilde{\Gamma}^a, \quad (\text{C11})$$

where $\tilde{\Gamma}^a$, $a = 1, \dots, d+1$ generate a $(d+1)$ -dimensional Clifford algebra. We find that the terms associated to M^a over this range of a thus take the form

$$iM^a (\chi_j^+)^{\dagger} \tilde{\Gamma}_{jk}^a \chi_k^- - iM^a (\chi_j^-)^{\dagger} \tilde{\Gamma}_{jk}^a \chi_k^+ = iM^a \bar{\chi}_j \gamma^d \tilde{\Gamma}_{jk}^a \chi_k. \quad (\text{C12})$$

Summarizing, we have found the effective Lagrangian

$$-i\bar{\chi}_j \gamma^\mu \partial_\mu \chi_j - M^{d+2} \bar{\chi}_j \chi_j - iM^a \bar{\chi}_j \gamma^d \tilde{\Gamma}_{jk}^a \chi_k, \quad a = 1, \dots, d+1 \quad (\text{C13})$$

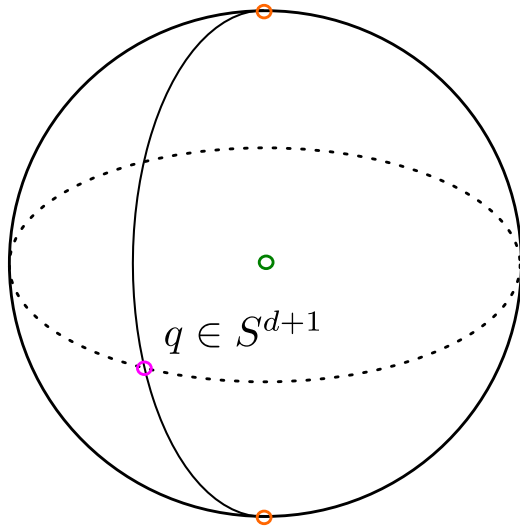


FIG. 6. The spatially varying parameters in S^{d+2} for the domain wall configurations trace out a great semicircle (black) connecting the north and south poles (orange) and are parametrized by a point $q \in S^{d+1}$ (pink), where the semicircle intersects the equatorial $(d+1)$ -sphere (dotted). The winding number of a family of such profiles over S^{d+2} equals the winding number of the intersection points q over S^{d+1} , giving an equality between the Berry number of the d -dimensional theory and its domain wall.

among 2^n complex fermions along the wall. We recognize this is the same as $\mathcal{L}_{d-1}^{\text{odd}}$ in (C1) after identifying $\gamma^{0 \dots (d-1)} = \gamma^d$.

In this case, there is a simple relationship between the Berry number of the d -dimensional theory parametrized by S^{d+2} and the $d-1$ -dimensional theory on the domain wall, parametrized by S^{d+1} . Indeed, the theory on the wall with given parameter values M^a , $a = 1, \dots, d+2$ is defined from the d -dimensional theory with a spatially varying parameters M^a such that along the x_d coordinate, the system parameters draw a great semicircle from the south pole $M^{d+3} = -1$ to the north pole $M^{d+3} = 1$. If the system parameters of $d-1$ -dimensional theory wind once over the equatorial S^{d+1} , we see that the semicircular arcs also wind once over the whole S^{d+2} . Thus the Berry numbers of the two theories are equal. See Fig. 6.

Now we consider the odd dimensional case $d = 2n - 1$, with 2^n -many 2^n -component complex fermions transforming in the flavor group $\text{Spin}(d+2)$ with generators Γ^a , $a = 1, \dots, d+2$. We again consider the Jackiw-Rebbi problem with a spatially varying mass $m(x_d)$ which goes from -1 to 1 over a region near $x_d = 0$. We find 2^n zero modes, one for each fermion, satisfying

$$\gamma^d \psi_j = \psi_j. \quad (\text{C14})$$

We choose a basis for the spinor space where γ^d takes the form $\sigma^z \otimes \mathbb{1}$ of (C3), so that

$$\psi_j = \begin{bmatrix} \psi_j^+ \\ \psi_j^- \end{bmatrix} \quad (\text{C15})$$

and the constraint $\gamma^d \psi_j = \psi_j$ sets ψ_j^- to zero. We define the 2^n domain wall fermions by

$$\chi_j = \psi_j^+. \quad (\text{C16})$$

Note these have half as many components as the ψ_j 's.

Recall that the chirality operator is

$$\gamma^{0 \dots d} = i^{(d-1)/2} \gamma^0 \dots \gamma^d. \quad (\text{C17})$$

When this acts on χ_j it acts as

$$i^{(d-1)/2} \gamma^0 \dots \gamma^{d-1}, \quad (\text{C18})$$

which is identically 1 on this subspace. Thus our interaction simplifies to

$$M^a \bar{\chi}_j \Gamma_{jk}^a \chi_k. \quad (\text{C19})$$

We thus obtain the theory $\mathcal{L}_{d-1}^{\text{even}}$ of (C2) on the wall. See Ref. [80] for a related discussion in the Hamiltonian language.

By the same argument as above, the Berry number of \mathcal{L}_{2n-1} equals that of \mathcal{L}_{2n} . Combining with the other inductive lemma, all of the Berry numbers of the Abanov-Wiegmann family are equal. Since the $D = 1$ case has Berry number 1 (the usual Berry number), all of the Berry numbers are 1.

Another important family is the Abanov-Wiegmann ‘‘A’’ family, which realizes generalized Thouless pumps. In odd space dimensions d , the theories consist of $2^{(d-1)/2}$ -many $2^{(d+1)/2}$ complex fermions ψ_j transforming under the flavor group $\text{Spin}(d)$ with generators Γ^a and the Lagrangian

$$\mathcal{L}_d^{\text{odd}} = i \bar{\psi}_j \gamma^\mu \partial_\mu \psi_j + M^0 \bar{\psi}_j \psi_j + i M^a \bar{\psi}_j \gamma^{0 \dots d} \Gamma_{jk}^a \psi_k, \quad (\text{C20})$$

where $\gamma^{0 \dots d}$ is the chirality operator. Meanwhile, in even space dimensions, the theories consist of $2^{d/2}$ -many $2^{d/2}$ -component complex fermions transforming under $\text{Spin}(d+1)$ with generators Γ^a and the Lagrangian

$$\mathcal{L}_d^{\text{even}} = i \bar{\psi}_j \gamma^\mu \partial_\mu \psi_j + M^a \bar{\psi}_j \Gamma_{jk}^a \psi_k. \quad (\text{C21})$$

One can repeat the derivations above to show that this series also dimensionally reduces along interfaces. We can use this to reduce all the way down to quantum mechanics, where we find a 1-component complex fermion, which depending on the sign of the chemical potential has either a neutral unoccupied or a charged occupied ground state. The generalized Thouless pump follows for the whole series from this simple observation.

2. Anomalous interfaces

In our construction above, the interface was always trivially gapped, and the parameter space of the interface was reduced to a smaller-dimensional sphere S^{d+1} . This is unlike our prescription for studying boundary anomalies in Sec. VI, however they can be related. In particular, we will see that the boundary diabolical point is hiding ‘‘inside’’ the S^{d+1} we defined above.

Indeed, we can define an extension of the interface parameter space from S^{d+1} to the interior ball B^{d+2} by adding a ‘‘radial’’ parameter which interpolates our great circular path with a straight line path. One of these paths passes through the massless point of the d -dimensional theory. In the free fermion theory, it is easy to show that the are gapless modes along the interface for any such path, generalizing the arguments of Jackiw-Rebbi. Even if we add interactions, however, because there is a higher Berry number on the boundary of this B^{d+1} , there is guaranteed to be a diabolical locus inside.

To directly relate the interfaces to the conclusions of Sec. VI however, we have to allow one of the end points to vary over the bulk parameter space, while the other is fixed at some value p_0 , which represents the parameter value “outside of the sample.” In this case, our choice of boundary condition for each bulk parameter p amounts to a choice of path $f_p(u)$ from p to p_0 . This defines a boundary condition for each $p \in S^{d+2}$. Because the sphere is not contractible, at least one of these paths $f_{p_*}(u)$ must pass through the massless point in the origin for some u (or more generally through some diabolical locus analogous to the one defined above). The point p_* thus defines a boundary diabolical point, as required by Sec. VI.

APPENDIX D: DIABOLICAL POINTS OF HIGH CODIMENSION

We will argue by induction that a system in d space dimensions has no protected diabolical points of codimension $m > d + 3$. The argument is rigorous for quantum mechanical systems but only a sketch for higher dimensions.

For quantum mechanical systems, this can be argued as follows. First, by adding to the Hamiltonian, a term proportional to the fermion parity, we can modify our family so that all ground states, including degenerate ones, have the same charge. By working only in this sector, we reduce the problem to eliminating a diabolical point of codimension >3 in an ordinary quantum mechanical system without any special symmetries.

Let us choose local coordinates on the parameter space so that the diabolical locus is at the origin. We can continuously deform our family away from the diabolical locus

(and without creating any other diabolical points) so that at some distance $R > 0$ from the origin the Hamiltonians are all (minus) projectors onto their unique ground state. The space of such projectors is a complex projective space $\mathbb{C}\mathbb{P}^{l-1}$, where l is the dimension of the Hilbert space. For any $m > 2$, by adding spectator degrees of freedom, we can make l big enough so that $\pi_m(\mathbb{C}\mathbb{P}^{l-1}) = 0$.²¹ Once we do so, it is possible to extend our family of projectors to the interior of the ball of radius R . We can then linearly interpolate with our original family to rid ourselves of the diabolical points inside this ball, all the while not modifying the Hamiltonian on the sphere, so in particular no diabolical points leave the ball. We are left with a completely nondegenerate phase diagram.

Now suppose we have shown the result for dimensions smaller than d , and we have an m -parameter phase diagram of d dimensional systems, $m > d + 3$, and a diabolical locus in some compact region. We choose some basepoint $p_0 \in S_R^{m-1}$ on a sphere of radius R large enough to enclose the diabolical locus. We study interfaces from p_0 to other points $p \in S_R^{m-1}$ by choosing paths in our parameter space. This will introduce diabolical loci of codimension $m - 1$ for the interface. By the inductive hypothesis, these interfaces may be smoothed

²¹Recall $\mathbb{C}\mathbb{P}^\infty$ has homotopy type $K(\mathbb{Z}, 2)$. That is, all its homotopy groups except π_2 are zero, while π_2 is isomorphic to \mathbb{Z} . In fact, by these same arguments we find \mathfrak{M}_0 is homotopic to $\mathbb{C}\mathbb{P}^\infty$.

so that there are no diabolical points amount the interface theories. Thinking of these interfaces as paths in \mathfrak{M}_d , we learn that our family on S_R^{m-1} can be extended to a ball B_R^m . Linearly interpolating our original family with this one inside the ball we rid ourselves of any diabolical points while keeping the family unchanged for radii greater than R .

-
- [1] M. V. Berry and M. D. Wilkinson, Proc. R. Soc. London, Ser. A-Math. Phys. Sci. **392**, 15 (1984).
- [2] J. von Neumann and E. P. Wigner, Phys. Z. **30**, 467 (1929).
- [3] M. V. Berry, Proc. Roy. Soc. London Ser. A **392**, 45 (1984).
- [4] B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).
- [5] S. Gukov, J. High Energy Phys. **01** (2016) 020.
- [6] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, Phys. Rev. B **83**, 205101 (2011).
- [7] A. Kapustin and L. Spodyneiko, Phys. Rev. B **101**, 235130 (2020).
- [8] A. Kapustin and L. Spodyneiko, arXiv:2003.09519.
- [9] C. Córdova, D. S. Freed, H. T. Lam, and N. Seiberg, SciPost Phys. **8**, 001 (2020).
- [10] C. Cordova, D. S. Freed, H. T. Lam, and N. Seiberg, SciPost Phys. **8**, 002 (2020).
- [11] J. C. Y. Teo and C. L. Kane, Phys. Rev. B **82**, 115120 (2010).
- [12] A. G. Abanov and P. B. Wiegmann, Nucl. Phys. B **570**, 685 (2000).
- [13] M. J. Rice and E. J. Mele, Phys. Rev. Lett. **49**, 1455 (1982).
- [14] J. Goldstone and F. Wilczek, Phys. Rev. Lett. **47**, 986 (1981).
- [15] D. J. Thouless, Phys. Rev. B **27**, 6083 (1983).
- [16] R. MacKenzie and F. Wilczek, Phys. Rev. D **30**, 2194 (1984).
- [17] W. Keil and R. Kobes, Phys. Rev. D **32**, 2231 (1985).
- [18] R. Jackiw and C. Rebbi, Phys. Rev. D **13**, 3398 (1976).
- [19] M. Lohse, C. Schweizer, O. Zilberberg, M. Aidelsburger, and I. Bloch, Nat. Phys. **12**, 350 (2016).
- [20] S. Nakajima, T. Tomita, S. Taie, T. Ichinose, H. Ozawa, L. Wang, M. Troyer, and Y. Takahashi, Nat. Phys. **12**, 296 (2016).
- [21] J. L. Lado and O. Zilberberg, Phys. Rev. Res. **1**, 033009 (2019).
- [22] M. Atiyah, *K-theory* (CRC Press, 2018).
- [23] R. Verresen (unpublished).
- [24] M. Mimura, J. Math. Kyoto Univ. **6**, 131 (1967).
- [25] D. S. Freed, Z. Komargodski, and N. Seiberg, Commun. Math. Phys. **362**, 167 (2018).
- [26] A. F. Young, C. R. Dean, L. Wang, H. Ren, P. Cadden-Zimansky, K. Watanabe, T. Taniguchi, J. Hone, K. L. Shepard, and P. Kim, Nat. Phys. **8**, 550 (2012).
- [27] N. Seiberg, T. Senthil, C. Wang, and E. Witten, Ann. Phys. **374**, 395 (2016).
- [28] E. Witten, *SL(2,Z) Action on Three-Dimensional Conformal Field Theories with Abelian Symmetry* (World Scientific, 2005), pp. 1173–1200.
- [29] A. Kapustin and M. J. Strassler, J. High Energy Phys. **04** (1999) 021.
- [30] F. Wilczek and A. Zee, Phys. Rev. Lett. **51**, 2250 (1983).

- [31] V. Borokhov, A. Kapustin, and X.-k. Wu, *J. High Energy Phys.* **11** (2002) 049.
- [32] F. Benini, P.-S. Hsin, and N. Seiberg, *J. High Energy Phys.* **04** (2017) 135.
- [33] T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, and M. P. Fisher, *Deconfined Quantum Critical Points* (World Scientific, 2010), pp. 333–343.
- [34] Z. Komargodski, A. Sharon, R. Thorngren, and X. Zhou, *SciPost Phys.* **6** (2019).
- [35] M. A. Metlitski and R. Thorngren, *Phys. Rev. B* **98**, 085140 (2018).
- [36] A. W. Sandvik, *Phys. Rev. Lett.* **98**, 227202 (2007).
- [37] H. Shao, W. Guo, and A. W. Sandvik, *Science* **352**, 213 (2016).
- [38] J. Lou, A. W. Sandvik, and N. Kawashima, *Phys. Rev. B* **80**, 180414(R) (2009).
- [39] A. W. Sandvik, *Phys. Rev. Lett.* **104**, 177201 (2010).
- [40] A. Nahum, J. T. Chalker, P. Serna, M. Ortuño, and A. M. Somoza, *Phys. Rev. Lett.* **107**, 110601 (2011).
- [41] S. Pujari, K. Damle, and F. Alet, *Phys. Rev. Lett.* **111**, 087203 (2013).
- [42] S. Pujari, F. Alet, and K. Damle, *Phys. Rev. B* **91**, 104411 (2015).
- [43] A. Nahum, P. Serna, J. T. Chalker, M. Ortuño, and A. M. Somoza, *Phys. Rev. Lett.* **115**, 267203 (2015).
- [44] R. K. Kaul and A. W. Sandvik, *Phys. Rev. Lett.* **108**, 137201 (2012).
- [45] P. Kim, H. Lee, S. Jiang, B. Ware, C.-M. Jian, M. Zaletel, J. H. Han, and Y. Ran, *Phys. Rev. B* **94**, 064432 (2016).
- [46] C.-M. Jian and M. Zaletel, *Phys. Rev. B* **93**, 035114 (2016).
- [47] H. C. Po, H. Watanabe, C.-M. Jian, and M. P. Zaletel, *Phys. Rev. Lett.* **119**, 127202 (2017).
- [48] O. Aharony, *J. High Energy Phys.* **02** (2016) 093.
- [49] P.-S. Hsin and N. Seiberg, *J. High Energy Phys.* **09** (2016) 095.
- [50] Y. Nakayama and T. Ohtsuki, *Phys. Rev. Lett.* **117**, 131601 (2016).
- [51] A. Tanaka and X. Hu, *Phys. Rev. Lett.* **95**, 036402 (2005).
- [52] T. Senthil and M. P. A. Fisher, *Phys. Rev. B* **74**, 064405 (2006).
- [53] C. Wang, A. Nahum, M. A. Metlitski, C. Xu, and T. Senthil, *Phys. Rev. X* **7**, 031051 (2017).
- [54] M. F. Lapa, C.-M. Jian, P. Ye, and T. L. Hughes, *Phys. Rev. B* **95**, 035149 (2017).
- [55] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, *J. High Energy Phys.* **02** (2015) 172.
- [56] A. Kapustin (unpublished).
- [57] A. Kapustin, R. Thorngren, A. Turzillo, and Z. Wang, *J. High Energy Phys.* **12** (2015) 052.
- [58] D. S. Freed and M. J. Hopkins (unpublished).
- [59] A. Kitaev, talk at “Between Topology and Quantum Field Theory: A conference in celebration of Dan Freed’s 60th birthday,” 2019.
- [60] D. Gaiotto and A. Kapustin, *Int. J. Mod. Phys. A* **31**, 1645044 (2016).
- [61] A. Kapustin and R. Thorngren, *J. High Energy Phys.* **10** (2017) 080.
- [62] C. Z. Xiong, *J. Phys. A: Math. Theor.* **51**, 445001 (2018).
- [63] D. Gaiotto and T. Johnson-Freyd, *J. High Energy Phys.* **05** (2019) 007.
- [64] R. Thorngren (unpublished).
- [65] R. Thorngren and D. V. Else, *Phys. Rev. X* **8**, 011040 (2018).
- [66] A. Kitaev, in *Advances in Theoretical Physics: Landau Memorial Conference*, edited by V. Lebedev and M. Feigel’Man, AIP Conf. Proc. No. 1134 (AIP, New York, 2009), pp. 22–30.
- [67] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, in *Advances in Theoretical Physics: Landau Memorial Conference*, edited by V. Lebedev and M. Feigel’Man, AIP Conf. Proc. No. 1134 (AIP, New York, 2009), pp. 10–21.
- [68] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, *New J. Phys.* **12**, 065010 (2010).
- [69] G. Palumbo and N. Goldman, *Phys. Rev. Lett.* **121**, 170401 (2018).
- [70] G. Palumbo and N. Goldman, *Phys. Rev. B* **99**, 045154 (2019).
- [71] D. S. Freed, *J. Diff. Geom.* **80**, 45 (2008).
- [72] J. C. Baez and J. Dolan, *J. Math. Phys.* **36**, 6073 (1995).
- [73] J. Lurie, *Curr. Dev. Math.* **2008**, 129 (2008).
- [74] P. Etingof, D. Nikshych, and V. Ostrik, *Quantum Topology* **1**, 209 (2010).
- [75] J. Fuchs, C. Schweigert, and A. Valentino, *Commun. Math. Phys.* **321**, 543 (2013).
- [76] Theo Johnson-Freyd (unpublished).
- [77] A. Sharon (unpublished).
- [78] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Math. Proc. Cambridge Philos. Soc.* **77**, 43 (1975).
- [79] F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 1029 (1988).
- [80] R. Queiroz, E. Khalaf, and A. Stern, *Phys. Rev. Lett.* **117**, 206405 (2016).