

Fermion path integrals and topological phases

Edward Witten

School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA

(published 13 July 2016)

Symmetry-protected topological (SPT) phases of matter have been interpreted in terms of anomalies, and it has been expected that a similar picture should hold for SPT phases with fermions. Here a description is given in detail of what this picture means for phases of quantum matter that can be understood via band theory and free fermions. The main examples considered are time-reversal invariant topological insulators and superconductors in two or three space dimensions. Along the way, the precise meaning of the statement that in the bulk of a 3D topological insulator, the electromagnetic θ angle is equal to π , is clarified.

DOI: [10.1103/RevModPhys.88.035001](https://doi.org/10.1103/RevModPhys.88.035001)

CONTENTS

I. Introduction	1	1. Spinors on orientable manifolds	32
A. Goal of this paper	1	2. Fermions on an unorientable manifold	34
B. Some generalities about fermions	2	3. Dirac operators	34
C. Organization and general remarks	4	Appendix B: Examples in Low Dimension	35
II. Pseudoreal Fermions	5	1. Two dimensions	35
A. Topological insulator in $d = 3$	5	2. Three dimensions	36
1. Basics	5	3. Four dimensions	37
2. Global anomaly and spectral flow	6	4. Five dimensions	37
3. T anomaly	7	Appendix C: The η Invariant in Four Dimensions	38
4. Canceling the anomaly from the bulk	8	References	39
5. Physical meaning of $\theta = \pi$	9		
6. Why is the partition function equal to $(-1)^{\mathfrak{S}}$?	10		
7. The index and cobordism invariance	11		
8. The APS index theorem	12		
9. Reflection positivity	14		
10. Running the story in reverse	15		
B. Topological superconductor in $d = 3$	15		
1. Fermions and Pfaffians	15		
2. Vanishing of the spectral flow	16		
3. The sign of the path integral	17		
C. Overview of the remaining cases	18		
III. Real Fermions	19		
A. Symmetry of the spectrum	19		
B. An interlude on the mod 2 index	20		
C. Simple examples of the global anomaly	21		
D. Bulk sTQFT associated with the T-invariant topological superconductor	22		
E. Extending the sTQFT to a manifold with boundary	23		
F. The orientable case and chiral symmetry	23		
G. $(2 + 1)$ -dimensional topological insulator	24		
IV. Complex Fermions	25		
A. Overview	25		
B. Anomaly of the boundary fermions	26		
C. Bulk partition function	28		
D. Anomaly cancellation and the Dai-Freed theorem	29		
E. The case of ν bands	30		
F. Special behavior for $\nu = 8$	30		
G. The $d = 3$ topological insulator on an unorientable spacetime	31		
V. The Majorana Chain	31		
Acknowledgments	32		
Appendix A: Spinors in Riemannian Geometry	32		

I. INTRODUCTION

A. Goal of this paper

Symmetry-protected topological (SPT) bosonic phases of matter can be characterized by anomalies (Chen *et al.*, 2013). The meaning of this statement is as follows.

A d -dimensional SPT phase is gapped in bulk and has no topological order in the usual sense: on a compact d -dimensional spatial manifold M without boundary, it has a unique ground state. But such a theory is not quite trivial. On a compact spacetime manifold X of dimension $D = d + 1$, in the large volume limit, in general in the presence of a suitable background gauge field, and after removing local, nonuniversal terms, the partition function of such a theory is a complex number of modulus 1, $Z_X = e^{i\Phi}$. Here $e^{i\Phi}$ is the partition function of a topological quantum field theory \mathfrak{Z} that is “invertible,” its inverse being the theory with complex conjugate partition function $Z_X^{-1} = e^{-i\Phi}$.

In this situation, typically $e^{i\Phi}$ cannot be defined as a topological invariant in a satisfactory way—consistent with all expected symmetries and with physical principles such as unitarity—if the spatial manifold M (and therefore the spacetime X) has a boundary. To define the theory \mathfrak{Z} in this situation, the boundary of M (and X) must be endowed with additional degrees of freedom of some kind. There are several possibilities, but the case of interest in this paper is that there are gapless degrees of freedom along the boundary of M and X , preserving the symmetries of theory \mathfrak{Z} .

We write \mathfrak{B} for the theory describing the gapless degrees of freedom on the boundary of X . This theory has an “anomaly”

of some kind and does not possess all of the desired physical properties, since if \mathfrak{B} were an entirely satisfactory and symmetry-preserving theory by itself, then including it on the boundary $Y = \partial X$ of X would not resolve whatever difficulty there is in defining theory \mathfrak{Z} on X in the first place. The anomaly in defining theory \mathfrak{B} on ∂X somehow cancels the difficulty in defining theory \mathfrak{Z} on X . Thus, this situation exemplifies, usually in a more abstract way, the idea of anomaly inflow from the bulk to the boundary (Callan and Harvey, 1985). For a general mathematical treatment of this situation, see D. S. Freed (2014). The prototype of this situation in condensed matter physics is the integer quantum Hall effect on a two-manifold M with boundary. On M , or rather on the three-manifold $X = \mathbb{R} \times M$, where \mathbb{R} parametrizes the time, there is a Chern-Simons coupling of the electromagnetic vector potential A :

$$I = k\text{CS}(A), \quad \text{CS}(A) = \frac{e^2}{4\pi} \int_Y d^3x \epsilon^{ijk} A_i \partial_j A_k, \quad k \in \mathbb{Z}. \quad (1.1)$$

This coupling is not gauge invariant on a manifold with boundary, and the associated anomaly is canceled by the coupling of A to chiral edge modes that propagate on the boundary. (If these modes all have unit charge, then their multiplicity is k .) Theory \mathfrak{B} describes those chiral edge modes, and its anomaly is the two (spacetime) dimensional version of the Adler-Bell-Jackiw anomaly for chiral fermions (Adler, 1969; Bell and Jackiw, 1969).

Some of the most interesting SPT phases studied in recent years, and observed experimentally in several important cases, are the topological insulators and superconductors that can be constructed out of free fermions via band theory. For introductions and reviews, see Hasan and Kane (2010), Hasan and Moore (2011), and Qi and Zhang (2011). The crucial symmetries in these examples are time-reversal symmetry T and the $U(1)$ symmetry associated with conservation of electric charge. The purpose of this paper is to understand these free fermion topological phases of matter from the standpoint of anomalies and anomaly inflow. Several results in this direction were obtained previously by Ryu, Moore, and Ludwig, 2010, where it was conjectured, correctly, as will become clear, that a fuller picture would result from consideration of global anomalies. In analyzing this problem, we make use of index theory, applied to fermion path integrals. This is not surprising, since the free fermion phases can be classified via K theory (Kitaev, 2009) [or by classes of matrices, which are related (Ryu *et al.*, 2010)], and index theory was invented (Atiyah and Singer, 1963, 1968) as the analysis associated with K theory.

We should say at the outset that a few of our considerations are somewhat fanciful from the point of view of condensed matter physics. We work in a relativistic framework in which one can consider a theory on an arbitrary D -manifold X , with spatial boundary Y . As in most studies of anomalies, it is technically convenient to take X to have Euclidean signature, and we do so. The relativistic framework is available because the boundary fermions of the standard topological phases are governed by Dirac-like equations with an emergent relativistic

symmetry, but relying on it is unnatural from the point of view of condensed matter physics. Perhaps it is possible to interpret our results in a Hamiltonian framework more natural for condensed matter physics, but we will not try to do this. Experience seems to show that gapped condensed matter systems are frequently governed by fully relativistic theories, but this is not fully understood.

Also, because of their emergent relativistic symmetry, the boundary fermions we study inevitably possess an emergent symmetry that in $D = 4$ is usually called CPT (where C is charge conjugation and P is parity or spatial inversion). However, this formulation is not valid for odd D (even spatial dimension d), since in that case P is contained in the connected part of the spatial rotation group and should be replaced by an operation that reverses the orientation of space. To use a language that is equivalent to the usual CPT statement for even D but is uniformly valid for all D , we refer instead to CRT symmetry, where R is a reflection of one spatial coordinate. CRT symmetry is always valid in a relativistic theory of any dimension, so, although not very natural in condensed matter physics, it is an emergent symmetry of the usual gapless fermions on ∂X for all of the usual free fermion phases. This will be used at some points, notably in Sec. III.G to describe the topological quantum field theory associated with a topological insulator in 2 space dimensions.

Actually we have been a little imprecise in this Introduction in speaking of topological quantum field theory (TQFT). Theories studied in this paper contain fermions, so they require a spin structure on spacetime (or a pin structure in the unorientable case). Hence the structure of interest is analogous to a TQFT except that spacetime is required to have a chosen spin (or pin) structure. We use the name sTQFT (spin topological quantum field theory) to refer to a theory that is like a TQFT except that it includes fermions. For an early study of such theories in the context of $D = 3$ Chern-Simons theory, see Dijkgraaf and Witten (1990). The “s” in sTQFT can refer to either a spin structure or a pin structure, depending on context. The interested reader can consult the Appendixes for explanations of spin and pin structures and related matters.

B. Some generalities about fermions

Now we explain a few generalities about fermions and anomalies that will provide useful background. First, anomalies always come from massless fermions only; massive fermions never contribute to anomalies. A technical explanation is that whatever symmetries a massive fermion ψ may possess are also possessed by a Pauli-Villars regulator field (a field of opposite statistics, obeying the same Dirac equation as ψ but with a much larger mass). So massive fermions can always be regularized in a way that preserves their symmetries, and they never contribute anomalies.

Hence, as one would expect intuitively, in analyzing fermion anomalies, one can ignore gapped degrees of freedom. In fact, we can ignore gappable degrees of freedom—fermions that *could* have a symmetry-preserving bare mass, even if they do not—since again a symmetry-preserving regulator is possible for such fermions. A corollary is that in any relativistic theory of fermions ψ in any dimension, there

is no problem in defining the absolute value $|Z_\psi|$ of the fermion path integral Z_ψ , consistent with all symmetries and physical principles. This is because a conjugate set of fermions $\tilde{\psi}$, transforming under any symmetries as the complex conjugate of ψ , and with an action that is the complex conjugate of the action of ψ , would have partition function \bar{Z}_ψ . The combined system of fermions $\psi \oplus \tilde{\psi}$ is always gappable. This combined system would have a partition function $Z_\psi \bar{Z}_\psi = |Z_\psi|^2$. Thus there is no anomaly in $|Z_\psi|^2$ or equivalently in $|Z_\psi|$.

Implicit in the previous paragraph is that in Euclidean signature, and in contrast to Lorentz signature, fermions do not necessarily transform in a real or self-conjugate representation of the relevant symmetry group (so ψ and $\tilde{\psi}$ may transform differently). For example, the edge modes in the integer quantum Hall effect are described by a fermion field ψ in spacetime dimension 2 that has definite chirality, meaning in Euclidean signature that ψ transforms with spin 1/2 (or spin $-1/2$) under the rotation group $SO(2)$. The chiral asymmetry of the edge modes means that there is no corresponding field of the opposite chirality, transforming as the complex conjugate of ψ . It is only in Lorentz signature that fermions, like boson fields, are real. For instance, continuing with the example of the chiral edge states, the spin 1/2 representation of $SO(2)$, in which a rotation by an angle φ acts as $e^{i\varphi/2}$, becomes after continuation to Lorentz signature a one-dimensional representation of $SO(1,1)$, in which a Lorentz transformation

$$\begin{pmatrix} \cosh w & \sinh w \\ \sinh w & \cosh w \end{pmatrix}$$

acts by $e^{w/2}$, which is real.

In general, relativistic fermions fit in three broad classes, according to how they transform under the appropriate symmetry group K in Euclidean signature. In K , we include the rotation group $SO(D)$, possible discrete symmetries such as ¹T or R, and gauge and/or global symmetries. Fermions may transform in (1) a pseudoreal representation, (2) a real representation, or (3) a complex representation. Of course, mixtures of these cases are also possible; for example, a theory may have some fermions transforming in a representation of one type and some in a representation of the other type. (Also, a reducible representation can be of more than one type; see Sec. III.F.) We encounter all three types in this paper.

An irreducible representation R of the group K is said to be pseudoreal if it admits an invariant, antisymmetric bilinear form ω . In this case, fermions transforming in the representation R , which we denote as ψ^α , $\alpha = 1, \dots, \dim R$, can

¹In Euclidean signature, all symmetries act on the fermions in a linear fashion. The antilinear nature of T arises in analytic continuation back to Lorentz signature. The CRT theorem means that one Euclidean symmetry can be continued to two distinct symmetries in Lorentz signature; for example, a Euclidean signature symmetry CR continues to either the antilinear symmetry T or the linear symmetry CR in Lorentz signature, depending on whether the reflection R acts on the spacetime coordinate that is being analytically continued.

have a K -invariant bare mass $\omega_{\alpha\beta}\psi^\alpha\psi^\beta$. (This is consistent with Fermi statistics, since ω is antisymmetric.) Hence, fermions transforming in a pseudoreal representation will never contribute to anomalies, and strictly speaking case (1) is not relevant to our considerations. However, there is a variant of this case that we might call case (1'). This is the case of a T-conserving theory with fermions that transform in a representation R that is pseudoreal if one omits T in the definition of K , but not if one includes T. Since the fermions are pseudoreal if T is ignored, such a theory can always be quantized in a consistent way, but there may be an anomaly in T symmetry. This anomaly comes from a problem involving the sign of the fermion path integral Z_ψ , as explained in Sec. II. Basic examples are the boundary fermions of a 3D topological insulator or superconductor, formulated on an orientable manifold only. We begin with this example, because the topological invariants involved are relatively familiar. Also, our results in the case of the 3D topological insulator may be particularly interesting, as we get a more precise description of the sense in which (Qi, Hughes, and Zhang, 2008) a 3D topological insulator is characterized by an electromagnetic θ angle equal to π .

Concerning case (2), an irreducible representation R of K is said to be real if it admits an invariant, symmetric bilinear form h . For a fermion ψ transforming in such a representation, a bare mass is forbidden by Fermi statistics, since $h_{\alpha\beta}\psi^\alpha\psi^\beta = 0$ because of Fermi statistics. However, if we double the fermion spectrum, adding a second multiplet $\tilde{\psi}$ also transforming in the representation of R , then a K -invariant bare mass $h_{\alpha\beta}\psi^\alpha\tilde{\psi}^\beta$ becomes possible. Doubling the spectrum replaces the fermion path integral Z_ψ with its square Z_ψ^2 , so Z_ψ^2 is anomaly free and for real fermions, an anomaly can affect only the sign of Z_ψ . The basic examples that we consider are edge states of a 2D topological superconductor or insulator.

Finally, concerning case (3), R is said to be complex if it admits no invariant bilinear form. In this case, the fermion path integral Z_ψ is complex in general. Unlike the other cases, fermions that transform in a complex representation can have perturbative anomalies—the standard Adler-Bell-Jackiw anomaly, and its analogs in other dimension. In this context, a “perturbative” anomaly is one that can be seen when fermions are quantized in a weak gauge or gravitational field, as opposed to “global” or “nonperturbative” anomalies that are invisible in weak background fields and involve global considerations. For example, the edge modes in the integer quantum Hall effect transform in a complex representation, and they do have a perturbative anomaly [which compensates for the lack of gauge invariance of the Chern-Simons coupling (1.1) on a manifold with boundary]. However, in this paper, we consider only cases in which perturbative anomalies are absent, leaving only the more subtle global anomalies. Our basic example of complex fermions is edge modes of a 3D topological superconductor or insulator, now on a possibly unorientable manifold. In particular, in the case of the topological superconductor, it turns out that the anomaly involves a 16th root of unity. [A specific computation for this system exhibiting an anomaly involving an 8th root of unity was performed by Hsieh, Cho, and Ryu (2015); the sense in which the anomaly is of order 16 rather than 8 is

rather subtle, as explained in Sec. IV.F.] The fact that the anomaly involves a 16th root of unity means that, although at the free fermion level, topological superconductors are classified by a \mathbb{Z} -valued invariant, anomalies see only the reduction of this invariant mod 16. In fact, it is known (Fidkowski, Chen, and Vishwanath, 2013; Metlitski *et al.*, 2014; Wang and Senthil, 2014; Kitaev, 2015) that with interactions included the classification of 3D topological superconductors is reduced from \mathbb{Z} to \mathbb{Z}_{16} , so anomalies precisely detect the 16 different classes. For a 3D topological insulator, the anomaly even on an unorientable manifold involves only a sign and the anomaly detects precisely the expected \mathbb{Z}_2 -valued invariant.

Anomalies for the case in which the representation R in Euclidean signature is real were originally studied by Witten (1982, 1985a). Such anomalies can be understood topologically in a fairly direct way, in terms of the mod 2 index of the Dirac operator. The case of pseudoreal fermions can be understood somewhat similarly, using the ordinary Dirac index rather than the mod 2 index, and was originally studied by Niemi and Semenoff (1983), Redlich (1984), and Alvarez-Gaumé, Della Pietra, and Moore (1985). Finally, global anomalies for complex fermions were analyzed by Witten (1985b) and were expressed in terms of the Atiyah, Patodi, and Singer (1975) (APS) η invariant. This relation was later refined in a result that we call the Dai-Freed theorem (Dai and Freed, 1994). The Dai-Freed theorem is useful in a variety of problems involving complex fermions, such as the world sheet path integral of the heterotic string (Witten, 2000). We make use of it in Sec. IV.D.

Part of the motivation for this work was the suggestion (Kapustin *et al.*, 2014) that the partition function of a 3D topological superconductor on a four-manifold would be the exponential of an η invariant. We will see that this proposal follows from a standard characterization of the phase transition between a trivial and topological superconductor. We will show by direct computation of partition functions that, as suggested by previous work,² sTQFTs associated with free fermion states of matter have cobordism-invariant partition functions. The proposal by Kapustin *et al.* (2014) that sTQFTs associated with free fermion states of matter have cobordism-invariant partition functions follows from direct calculation of partition functions for the bulk gapped fermions. The relation of these bulk calculations to the anomalies of boundary fermions follows from the Dai-Freed theorem in general, and from more elementary considerations when the gapless boundary fermions are real or pseudoreal.

²As briefly explained in the discussion of Eq. (5.4) in D. Freed (2014), in general $U(1)$ -valued cobordism invariants are partition functions of unitary TQFTs or sTQFTs. Kapustin *et al.* (2014) showed that in the examples we consider in this paper, the possible cobordism invariants are in one-to-one correspondence with free fermion phases of matter. (In higher dimensions, there are cobordism invariants that do not have an obvious connection to free fermion phases, so they may be associated with interacting phases of matter.) In Secs. II.A.7 and II.A.9, we give some very partial indications of how cobordism invariance is related to physical principles such as unitarity and the behavior under cutting and gluing.

The literature on anomalies is too vast to be properly summarized here. However, several additional papers in which various problems are treated in a similar spirit to the approach here are Bunke (2009), in which considerations of Witten (2000) concerning the heterotic string are formulated rigorously; Diaconescu, Moore, and Witten (2002) with (see Sec. 2) a precise definition of the path integral of M-theory on an 11-manifold³; Freed and Moore (2006), in which a more general version of that problem is treated precisely; Mikhaylov and Witten (2014), in which (see Sec. 5.3) an analog of the topological insulator appears in a problem involving D -branes in string theory; and Monnier (2014), on the anomaly field theory associated with chiral fermions. As this paper neared completion, I became aware of related work involving the η invariant in the context of SPT phases (Metlitski, 2015).

C. Organization and general remarks

The contents of this paper are summarized as follows. In Sec. II, we discuss pseudoreal fermions or more precisely fermions that are pseudoreal if one ignores T symmetry. In Sec. III, we discuss real fermions, and in Sec. IV, we discuss complex fermions. In Sec. V, we briefly consider another example with real fermions: the Majorana chain (Fidkowski and Kitaev, 2009) in $1+1$ dimensions. The overall idea is to start with simple examples and gradually introduce more complicated ones. We give a sort of overview of the different cases in Sec. II.C.

We tried to make this paper relatively self-contained, but some details are contained in the Appendixes. For basic definitions about spinors in Riemannian geometry, see Appendix A. For details in low dimension, see Appendix B. The η invariant in four (and two) dimensions is analyzed in Appendix C.

In this paper, with the sole exception of Sec. V, we only consider time-reversal symmetries with $T^2 = (-1)^F$ [here $(-1)^F$ is the operator that counts fermions mod 2]. This also means that the square of a spatial reflection R [or CR in the presence of a $U(1)$ symmetry] is $+1$, and that when we work on an unorientable manifold, we use a Pin^+ structure, not a Pin^- structure. For an explanation of these concepts, see Appendix A. In Sec. V, all statements are reversed: $T^2 = +1$, $R^2 = (-1)^F$, and the spacetime is endowed with a Pin^- structure.

We frequently introduce a Pauli-Villars regulator field as an aid in defining precisely a continuum theory of fermions. The sign of the regulator mass is then an important variable. In condensed matter physics, there would be no regulator but a physical cutoff. For example, in band theory, gapless relativistic fermions arise at a particular point (or points) in a band, and the topology of the rest of the band plays a role analogous to the sign of the regulator mass in the Pauli-Villars approach.

We usually phrase our arguments in terms of partition functions. There is much more to a physical theory than its

³Because of the mod 8 periodicity of real K theory, this 11-dimensional problem actually has much in common with the case of three-dimensional boundary fermions treated in Sec. II. But the details are different and the anomaly cancels with no need for inflow from 12 dimensions.

partition function, but in the case of free fermions, if there is a satisfactory definition of the partition function, there is always a satisfactory definition of correlation functions. So focusing on the partition function is a convenient shorthand for determining if a theory is well defined. Occasionally the partition function vanishes and one must consider the path integral measure instead.

II. PSEUDOREAL FERMIONS

A. Topological insulator in $d=3$

1. Basics

Our first example will be a topological insulator in $d=3$, studied in this section on an oriented $(3+1)$ -dimensional spacetime X . X has a spatial boundary, which is a $(2+1)$ -dimensional orientable manifold Y . The boundary mode of the topological insulator is a $(2+1)$ -dimensional massless two-component Dirac fermion ψ , with action

$$I = \int_Y d^3x \bar{\psi} i \not{D} \psi. \quad (2.1)$$

Here $\not{D} = \sum_{\mu=0}^2 \gamma^\mu D_\mu$ is the usual Dirac operator. ψ couples to the vector potential A of electromagnetism, just like the electron; indeed, in band theory, ψ originates as a mode of the electron.

It is possible to endow ψ with a Lorentz-invariant mass term. The Dirac equation then becomes⁴

$$(\not{D} - m)\psi = 0, \quad (2.2)$$

with real m . This describes a particle of mass $|m|$ and spin $\frac{1}{2} \text{sign}(m)$. Time reversal and reflection symmetry reverse the sign of the particle spin, so the perturbation to $m \neq 0$ is T and R violating. Concretely, a T transformation⁵

$$\mathsf{T}\psi(t, x_1, x_2) = \gamma_0 \psi(-t, x_1, x_2) \quad (2.3)$$

is easily seen to reverse the sign of m in the Dirac equation. If, however, we have two identical Dirac fermions ψ_1 and ψ_2 , both transforming under T as in Eq. (2.3), then one can add a T -invariant⁶ mass term to the Dirac equation:

⁴In local Lorentz coordinates, our Dirac matrices are real matrices obeying $\{\gamma_a, \gamma_b\} = 2\eta_{ab}$ where $\eta = \text{diag}(-1, 1, 1)$. These conventions are convenient for T -invariant systems with $\mathsf{T}^2 = (-1)^F$, and make it straightforward to compare Majorana and Dirac fermions.

⁵For a Majorana fermion, we have to consider two possible signs in this equation: $\mathsf{T}\psi(-t, x_1, x_2) = \pm \gamma_0 \psi(-t, x_1, x_2)$. For a Dirac fermion with $\mathsf{U}(1)$ symmetry, we can eliminate this sign by transforming, if necessary, $\psi \rightarrow i\psi$.

⁶The mass term in Eq. (2.4) is R violating, assuming that a spatial reflection R is supposed to act the same way on both ψ_1 and ψ_2 by $\mathsf{R}\psi_i(t, x_1, x_2) = \pm \gamma_1 \psi_i(t, -x_1, x_2)$. (The mass term conserves a different R symmetry.) The only general symmetry in relativistic QFT is CRT , which means that T symmetry is equivalent to CR symmetry. Given a T symmetry, there is a canonical CR symmetry, but there is no canonical R symmetry.

$$\left(\not{D} - \begin{pmatrix} 0 & im \\ -im & 0 \end{pmatrix} \right) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = 0. \quad (2.4)$$

Diagonalizing the mass term, one finds that T exchanges two modes of equal mass (and charge) and opposite spin. Since a single Dirac fermion cannot acquire a T -conserving mass, it follows that in dimension $3+1$, if a T -invariant material has a single massless Dirac fermion on its boundary, then this state of affairs is protected by T invariance. We have written this paragraph in Lorentz signature, because this makes the action of T most transparent, but in our analysis of the path integral for ψ and its anomaly, we take Y to have Euclidean signature.

The path integral Z_ψ of the ψ field is formally the determinant of the operator $\mathcal{D} = i\not{D}$:

$$Z_\psi = \det \mathcal{D}. \quad (2.5)$$

The operator \mathcal{D} is Hermitian, so its eigenvalues are real:

$$\mathcal{D}\psi_i = \lambda_i \psi_i, \quad \lambda_i \in \mathbb{R}. \quad (2.6)$$

Formally, the determinant is the product of eigenvalues:

$$\det \mathcal{D} = \prod_i \lambda_i. \quad (2.7)$$

Formally, this is real, since all the factors are real. This reality has a simple physical meaning: it reflects the T invariance of the classical field theory that we are trying to quantize. In general, in any unitary QFT, on an orientable spacetime X of Euclidean signature, the partition function is complex conjugated if one reverses the orientation of X . In a T -invariant theory, this reversal of orientation is a symmetry, so the partition function is always real. All this is true only for orientable X ; for unorientable X , T invariance does not require the partition function to be real. See Sec. IV for examples.

For two massless Dirac fermions, the path integral is

$$Z_\psi^2 = \prod_i \lambda_i^2. \quad (2.8)$$

Every factor is not just real but positive, so formally one certainly expects Z_ψ^2 to be positive, and indeed the path integral for two massless Dirac fermions can be naturally defined to be positive. Being positive, it is equal to $|Z_\psi|^2$, and thus is completely anomaly free, since as explained in Sec. I.B, the absolute value $|Z_\psi|$ of a relativistic fermion path integral can always be defined in a completely anomaly-free way, preserving all symmetries. In particular, the path integral for two massless Dirac fermions on Y can be defined in a T -invariant fashion. This should come as no surprise, since we have seen that a T -invariant mass and therefore a T -invariant Pauli-Villars regulator are possible. [The T -invariant Pauli-Villars regularization is based on the T -invariant massive Dirac equation (2.4), so it uses a regulator of positive mass for one linear combination of the two Dirac fermions and of negative mass for the other.]

With only one Dirac fermion, however, we have a problem, because the formal expression

$$Z_\psi = \prod_i \lambda_i \quad (2.9)$$

is naturally real but not naturally positive. Its sign is roughly speaking the number of λ_i that are negative, mod 2, but the number of negative λ_i is infinite, and there is no natural way to decide if this infinite number is even or odd.

2. Global anomaly and spectral flow

In fact, if we try to define Z_ψ to be positive and gauge invariant, we run into a contradiction. We could pick an arbitrary metric and gauge field, say $g = g_0$ and $A = A_0$, and define Z_ψ to be, say, positive at $(A, g) = (A_0, g_0)$. Then letting A and g vary, we could follow the sign of Z_ψ continuously, saying that this sign changes whenever an eigenvalue of \mathcal{D} passes through 0. However, this procedure leads to a conflict with gauge invariance. Let ϕ be a gauge transformation or the combination of a gauge transformation and a diffeomorphism. Let (A_0^ϕ, g_0^ϕ) be whatever A_0 and g_0 transform into under ϕ . It is always possible to continuously interpolate from (A_0, g_0) to (A_0^ϕ, g_0^ϕ) . One introduces a real parameter s , $0 \leq s \leq 1$, and then one sets

$$A_s = (1-s)A_0 + sA_0^\phi, \quad g_s = (1-s)g_0 + sg_0^\phi. \quad (2.10)$$

These formulas have been chosen so that (A_s, g_s) coincides with (A_0, g_0) at $s = 0$ and with (A_0^ϕ, g_0^ϕ) at $s = 1$. Before accepting this interpolation from (A_0, g_0) to (A_0^ϕ, g_0^ϕ) , we should make sure that A_s and g_s are well defined for all s . A metric g_s is supposed to be positive definite. There is no problem because g_s for $0 < s < 1$ is a linear combination with positive coefficients of the positive metrics g_0 and g_0^ϕ . For A_s , there is actually no analogous condition to be checked.

Now we evolve (A_s, g_s) continuously from $s = 0$ to $s = 1$, and we check how many times Z_ψ changes sign. Gauge invariance implies that Z_ψ should have the same sign at $s = 1$ as at $s = 0$, since (A_s, g_s) at $s = 1$ is gauge equivalent to what it is at $s = 0$. However, in general there is a problem. Gauge invariance implies that the spectrum of the Dirac operator at $s = 1$ is the same as it is at $s = 0$, but between $s = 0$ and $s = 1$, there can be a net ‘‘spectral flow’’ of the Dirac eigenvalues, as shown in Fig. 1. Such a spectral flow is possible only because \mathcal{D} has infinitely many positive and negative eigenvalues.

In the case of the boundary fermions of the 3D topological insulator, there definitely is such an inconsistency in trying to define the sign of Z_ψ , assuming that one expects it to be real. There actually is a topological formula for the net number Δ of eigenvalues flowing through $\lambda = 0$, namely,

$$\Delta = \mathfrak{S}, \quad (2.11)$$

where as described momentarily, \mathfrak{S} is a certain Dirac index in $D = 4$. The corresponding formula for the sign change of Z_ψ between $s = 0$ and $s = 1$ is

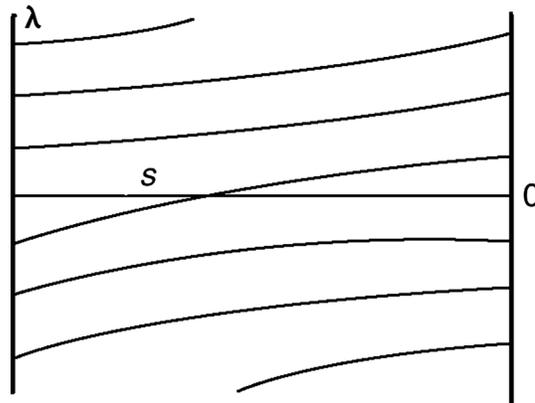


FIG. 1. Spectral flow for a Dirac operator. The vertical axis parametrizes an eigenvalue λ and the horizontal axis parametrizes a parameter s on which the eigenvalues depend. In the case shown, the spectrum is the same at $s = 1$ as at $s = 0$, but there is a net upward flow of one eigenvalue through $\lambda = 0$ between $s = 0$ and $s = 1$. This leads to a sign change of the fermion path integral Z_ψ .

$$Z_\psi \rightarrow Z_\psi(-1)^{\mathfrak{S}}. \quad (2.12)$$

\mathfrak{S} will be a Dirac index on a certain four-manifold X , known as the mapping torus. This manifold will be a calculational tool to study anomalies in the fermion theory on Y . In our discussion of the topological insulator, a four-manifold will generically be called X , whether it is a tool to study anomalies or is the world volume of a physical system. This reason is that the same mathematical considerations will enter in either case. This is not a coincidence but represents the importance of the boundary anomalies in the physics of the topological insulator.

To define X , let I be the unit interval $0 \leq s \leq 1$. We glue together the ends of $I \times Y$ via the diffeomorphism ϕ to build a four-manifold X without boundary.⁷ A manifold constructed in this way is called a mapping torus. Using also the gauge transformation part of ϕ , we glue together the gauge bundles over the ends of $I \times Y$ to make a $U(1)$ gauge bundle $\mathcal{L} \rightarrow X$. Then \mathfrak{S} is the index of the 4D Dirac operator on⁸ X , coupled to \mathcal{L} . (See Sec. II.A.6 for the definition of the Dirac index.) This relation between spectral flow in three dimensions and the Dirac index in four dimensions, originally obtained by Atiyah, Patodi, and Singer (1975), has applications in QCD and has been described in the physics literature (Callan, Dashen, and Gross, 1978; Kiskis, 1978).

What has just been described, in which one looks for a possible inconsistency in the sign or phase of Z_ψ as a function

⁷One can slightly modify the s dependence of the metric so that the metric on X is smooth at the end points $s = 0$ and 1 .

⁸Instead of working on the compact mapping torus X , we could set $u = 1/(1-s) - 1/s = (2s-1)/s(1-s)$, so $0 < s < 1$ corresponds to $-\infty < u < \infty$, and work on the noncompact four-manifold $X' \cong \mathbb{R} \times Y$, with metric $du^2 + g_{s(u)}$, and the corresponding u -dependent gauge field $A_{u(s)}$. The index of the Dirac operator on X' , in the space of square-integrable wave functions, is the same as the index on X .

of a parameter s , is the framework of the 1980s (Witten, 1982, 1985a, 1985b) for studying global anomalies. At one time it was assumed to be the whole story. Nowadays it is clear that this is not the case, and in this paper we need a more complete treatment. The reason that we have begun with the traditional framework is not that we want to delve into history but that this is the most elementary way to introduce the important topological invariants—notably the index \mathfrak{F} in the present example and its analogs in examples that we introduce later.

What is missing in the traditional framework can be described as follows. The absence of an anomaly in the sign of Z_ψ means that on a particular Y , Z_ψ is well defined as a function of (A, g) , up to an overall sign that depends on Y but not on (A, g) . [In the above reasoning, this sign was introduced as the sign of Z_ψ at some arbitrary starting point (A_0, g_0) .] But we need a way to determine this overall sign, and we should certainly not expect to get a satisfactory theory if we define the sign of Z_ψ independently for each Y . A definite physical theory produces definite answers for these signs. Physically, there must be a sensible behavior under various cutting and pasting operations in which various three-manifolds Y_i are cut in pieces and glued together in different ways.

The prototype of such a cutting relation arises in ordinary quantum mechanics, with Hamiltonian H , if a time interval of length t is decomposed in successive intervals of length t_1 and t_2 , where $t = t_1 + t_2$. One has $\exp(-iHt) = \exp(-iHt_2)\exp(-iHt_1)$, leading to a well-known result for transition amplitudes

$$\langle f | \exp(-iHt) | i \rangle = \sum_k \langle f | \exp(-iHt_2) | k \rangle \langle k | \exp(-iHt_1) | i \rangle, \quad (2.13)$$

with a sum over intermediate states. This has an analog in quantum field theory in any dimension, where one cuts a manifold into pieces and claims that the path integral on the whole manifold is the product of path integrals on individual pieces (with a sum over physical states where gluing occurs). The quantum field theory story is much richer than the quantum mechanical case because in higher dimensions there are many more ways to cut a manifold into pieces. Compatibility with cutting and pasting is the essence of the locality of quantum field theory.

In cases relevant to topological states of matter—a good example being the 3D topological superconductor—even if there is no anomaly in the traditional sense of an inconsistency in defining the path integral on a specific Y , there can be an anomaly in the more subtle sense that there is no satisfactory way to define overall signs or phases of the path integral on different Y 's. One needs to take these more subtle anomalies into account as part of the following paradigm: “Anomalies in d dimensions \leftrightarrow SPT phases in $d + 1$ dimensions.”

Taking these anomalies into account means giving an absolute definition of the sign of the fermion path integral Z_ψ for each Y and each (A, g) . We will do this next in the case of the topological insulator, following Alvarez-Gaumé, Della Pietra, and Moore (1985).

3. T anomaly

The fact that there is a problem in defining the sign of Z_ψ for a $(2 + 1)$ -dimensional Dirac fermion ψ does not mean that this theory is inconsistent. It only means that the theory cannot be quantized in a T-invariant (or R-invariant) way. It is possible to define Z_ψ consistently if we do not try to make it real.

After all, ψ could have a gauge-invariant bare mass, which violates T symmetry but otherwise is perfectly physically acceptable. The possibility of a T-violating mass means that, at the cost of losing T symmetry, we can regularize this theory by adding a Pauli-Villars regulator field χ , which one can think of as a bosonic field that obeys a massive Dirac equation $(i\mathcal{D} + i\mu)\chi = 0$, for very large μ . In Euclidean signature, the regularized path integral is

$$Z_{\psi,\text{reg}} = \prod_k \frac{\lambda_k}{\lambda_k + i\mu}. \quad (2.14)$$

Actually, what we have written, although good enough for our purposes, is only an approximation to the Pauli-Villars procedure. In general one introduces a variety of massive Bose and Fermi fields, with different large masses μ_a , $a = 1, \dots, t$, chosen so that the regularized path integral converges for fixed μ_a . Then one takes the limit $\mu_a \rightarrow \infty$, adding local counterterms $W(\mu_a, A, g)$ to the action so that the limit of $Z_{\psi,\text{reg}} \exp[-W(\mu_a, A, g)]$ exists. The limit is the renormalized fermion path integral. We do not need to follow this procedure in detail, because we are interested only in the phase of Z_ψ , and Eq. (2.14) is good enough to motivate the correct formula for this phase. (The counterterms that are needed in three spacetime dimensions are all real and do not affect the discussion of the phase.)

Going back to Eq. (2.14), we see that for large $\mu > 0$, each eigenvalue λ_k contributes a phase i^{-1} or i to Z_ψ , depending on the sign of λ_k . So formally

$$Z_\psi = |Z_\psi| \exp\left(-\frac{i\pi}{2} \sum_k \text{sign}(\lambda_k)\right). \quad (2.15)$$

Thus

$$Z_\psi = |Z_\psi| \exp(-i\pi\eta/2), \quad (2.16)$$

where η [the APS η invariant (Atiyah, Patodi, and Singer, 1975)] is a regularized version of the difference between the number of positive and negative eigenvalues⁹ of \mathcal{D} .

⁹In the present derivation, the λ_k are the eigenvalues of the Dirac operator \mathcal{D} acting on the positively charged Dirac fermion ψ , without including charge conjugate eigenmodes of \mathcal{D} acting on the negatively charged field $\bar{\psi}$. We can also define a similar invariant η in which one sums over modes of both kind. By charge conjugation symmetry, the relation between the two is just $\eta = 2\eta$. Later, when we consider Majorana fermions [with no U(1) symmetry and so no distinction between modes of positive or negative charge], we will want to express all formulas in terms of η .

As usual, the precise regularization does not matter. The original APS definition was¹⁰

$$\eta = \lim_{s \rightarrow 0} \sum_k \text{sign}(\lambda_k) |\lambda_k|^{-s}. \quad (2.17)$$

An equivalent definition would be

$$\eta = \lim_{\varepsilon \rightarrow 0^+} \sum_k \text{sign}(\lambda_k) \exp(-\varepsilon \lambda_k^2). \quad (2.18)$$

For our present application, we do not need to know what is meant by $\text{sign}(\lambda)$ if $\lambda = 0$, because if one of the λ_k is 0, then $|Z_\psi| = 0$ and it does not matter what value we assign to η . However, the formulas of index theory work smoothly if in the definition of η , we set

$$\text{sign}(\lambda) = \begin{cases} 1 & \text{if } \lambda \geq 0, \\ -1 & \text{if } \lambda < 0. \end{cases} \quad (2.19)$$

It would work just as well to take $\text{sign}(\lambda) = -1$ for $\lambda = 0$; the important thing is to treat all zero modes in the same way.

The formula

$$Z_\psi = |Z_\psi| \exp(-i\pi\eta/2), \quad (2.20)$$

together with any standard procedure to define $|Z_\psi|$, gives a satisfactory definition of Z_ψ for all Y and all A, g , with all desirable properties except **T** and **R** symmetry. **T** and **R** symmetry do not hold, since they require Z_ψ to be real. **T** and **R** symmetry have been violated by the choice of sign of the regulator mass. An equally good regularization with the opposite sign would have given the opposite phase to Z_ψ , so we really have two equally good definitions with

$$Z_\psi = |Z_\psi| \exp(\mp i\pi\eta/2). \quad (2.21)$$

Either of these formulas gives the partition function, in a background field (A, g) , of a gapless free fermion QFT that is perfectly unitary and Poincaré invariant—and even conformally invariant—and otherwise physically sensible, but is not **T** or **R** invariant.

To conclude this section, we explain some aspects of the statement that Eq. (2.21) for the path integral is physically sensible. The fermion path integral Z_ψ is supposed to change sign when a fermion eigenvalue passes through 0. Equation (2.20) does have this property. While the first factor $|Z_\psi|$ of course does not change in sign, the second factor $\exp(\mp i\pi\eta/2)$ does change sign, because η jumps by ± 2 when an eigenvalue passes through 0.

We also want to verify that our definition of the partition function is consistent with unitarity. In Euclidean signature, unitarity corresponds to reflection positivity. Let Y_1 and Y_2 be

two identical three-manifolds, endowed with the same gauge fields and spin structures, and let M be their common boundary. Build a three-manifold Y by gluing Y_1 and Y_2 along M (after reversing the orientation of Y_2 so that the orientations match). This construction is illustrated in Fig. 2. Z_ψ should be real and non-negative in this situation for the following reason. The path integral on Y_1 constructs a state $|\Phi\rangle$, a “ket,” in the Hilbert space of quantum states on M . The path integral on Y_2 constructs the corresponding “bra” $\langle\Phi|$ in the same Hilbert space. The full path integral Z_ψ on Y is the inner product $\langle\Phi|\Phi\rangle$. In a unitary quantum field theory, this of course must be real and non-negative.

To show that $Z_\psi \geq 0$, we need to show that $\exp(\mp i\pi\eta/2)$ is equal to 1 in this situation, as long as there are no zero modes (that is, as long as $|Z_\psi| \neq 0$). This is true because the reflection that exchanges Y_1 and Y_2 , and reverses the orientation of Y , anticommutes with the Dirac operator $\mathcal{D} = i\mathcal{D} = i\sum_{\mu=1}^3 \gamma^\mu D_\mu$. For example, if $Y = \mathbb{R}^3$ with a flat metric, a reflection

$$\mathbf{R}\psi(x_1, x_2, x_3) = \gamma^1 \psi(-x_1, x_2, x_3) \quad (2.22)$$

anticommutes with \mathcal{D} . This is actually a universal result for any orientation-reversing symmetry on a manifold of odd dimension. So reflection symmetry in Fig. 2 implies that the eigenvalues of the Dirac operator are invariant under $\lambda \leftrightarrow -\lambda$, and this in turn implies (in the absence of zero modes) that $\eta = 0$ and $\exp(\mp i\pi\eta/2) = 1$.

The fact that a symmetry reversing the orientation of spacetime anticommutes with the Dirac operator is the reason that a mass term added to the Dirac equation violates such symmetries. If **R** commuted with $i\mathcal{D}$, then a massive Dirac equation, which in Euclidean signature would be $(i\mathcal{D} + im)\psi = 0$, would be **R** invariant.

4. Canceling the anomaly from the bulk

The anomaly in **T** invariance just discussed for the boundary fermions of a 3D topological insulator is related to the fact (Qi, Hughes, and Zhang, 2008) that in bulk such a material has an electromagnetic θ angle of π .

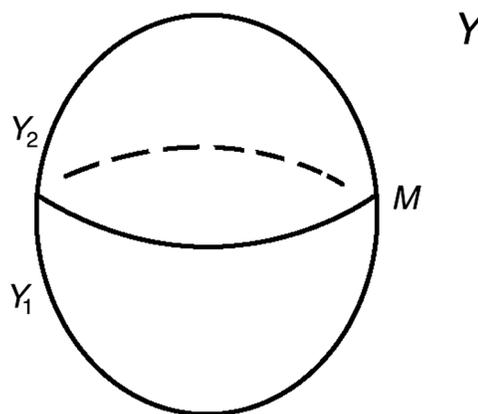


FIG. 2. A mirror-symmetric construction of a three-manifold Y , by gluing together identical pieces Y_1 and Y_2 along their common boundary M , with reversed orientation for Y_2 .

¹⁰Some analytic continuation is required in this definition. One first defines the sum on the right-hand side of Eq. (2.17) for large $\text{Re } s$, where it converges. Then one analytically continues to $s = 0$. Note that the alternative definition in Eq. (2.18) does not require such analytic continuation.

The instanton number of a U(1) gauge field in four spacetime dimensions is defined as

$$P = \frac{e^2}{32\pi^2} \int_X d^4x \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta}, \quad (2.23)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength. If X is a compact manifold with boundary, then P is always an integer. A typical example with $P \neq 0$ is the case that $X = S^2 \times S^2$ with one unit of flux (that is, $\int_{S^2} F = 2\pi/e$) on each factor. This example has $P = 1$.

The electromagnetic θ angle is defined by saying that the P -dependent part of the effective action is

$$I_\theta = \theta P. \quad (2.24)$$

If P can be assumed to be integer valued, then physics will be invariant under $\theta \rightarrow \theta + 2\pi$, since in quantum mechanics, we care about only the value of the action mod $2\pi\mathbb{Z}$.

P is odd under reflection or time reversal, so θ is likewise odd. In a T- or R-invariant theory, θ must equal 0 or π (T and R map $\theta = \pi$ to $\theta = -\pi$, which is equivalent to $\theta = \pi \bmod{2\pi}$). It was shown by Qi, Hughes, and Zhang (2008) that, assuming that $\theta = 0$ in vacuum, $\theta = \pi$ inside a 3D topological insulator.

However, a real topological insulator has a boundary and its world volume X does not fill all of spacetime. In this case, P (defined as an integral over X and not over all of spacetime) is generically not an integer. So it is not T invariant to merely include in the path integral a factor $\exp(\pm i\pi P)$. On an orientable spacetime, T invariance requires that the argument of the path integral should be real, and $\exp(\pm i\pi P)$ certainly does not have this property when P is not \mathbb{Z} valued. Acting with T reverses the sign in the exponent $\exp(\pm i\pi P)$.

More concretely, if we simply include in the functional integral in the presence of a topological insulator a factor $\exp(\pm i\pi P)$, then by a well-known argument this produces on the surface of the topological insulator a Hall conductivity with $\nu = \pm 1/2$ (as usual, ν is the Hall conductivity in units of $e^2/2\pi\hbar$). This is certainly not T conserving.

In the context of a topological insulator, the assertion that $\theta = \pi$ is T conserving really means that it is T conserving in bulk, and that we may be able to maintain T invariance along the boundary if we find the right boundary state. A trivial gapped boundary state is not suitable.

However, the standard boundary state with massless Dirac fermions combines with the bulk system of $\theta = \pi$ to maintain T symmetry. We recall that the partition function of the boundary fermions is

$$Z_\psi = |Z_\psi| \exp(\mp i\pi\eta/2), \quad (2.25)$$

where the sign depends on the choice of regulator. It turns out that if one subtracts from P a gravitational correction that for the moment we denote¹¹ as $\hat{A}(R)$, then the bulk contribution to the path integral with $\theta = \pi$ combines with the partition function of the boundary fermions to give a

T-conserving result. In fact, according to Atiyah, Patodi, and Singer (1975),

$$\exp(\mp i\pi\eta/2) \exp\{\pm i\pi[P - \hat{A}(R)]\} = (-1)^{\mathfrak{S}}, \quad (2.26)$$

where \mathfrak{S} is an integer. As we explain in Sec. II.A.8, \mathfrak{S} is the index of the Dirac operator computed with APS boundary conditions. Hence the complete path integral measure after integrating out the boundary fermions is

$$|Z_\psi| \exp(\mp i\pi\eta/2) \exp\{\pm i\pi[P - \hat{A}(R)]\} = |Z_\psi| (-1)^{\mathfrak{S}}. \quad (2.27)$$

[This formula was essentially found by Mikhaylov and Witten (2014), Sec. 5.3, in a related context involving D -branes in string theory.] This is real and so T conserving.

In the rest of this section, we discuss the physical interpretation of the formula (2.27), and then review the APS formula (2.26).

5. Physical meaning of $\theta = \pi$

To have $\theta = \pi$ in a topological insulator should mean, in some sense, that the path integral measure changes sign when a U(1) instanton moves from outside to inside a topological insulator. To make this concrete, we set up a thought experiment that makes some sense in condensed matter physics in which it is possible to move an instanton in this way.

We take space to be not \mathbb{R}^3 but $\mathbb{R} \times S^2$, with two dimensions compactified to a two-sphere S^2 . (We could just as well replace S^2 with T^2 , a two-torus corresponding to periodic boundary conditions in two spatial directions.) We consider a topological insulator supported on $\mathbb{R}_+ \times S^2 \subset \mathbb{R} \times S^2$ (Fig. 3) where $\mathbb{R}_+ \subset \mathbb{R}$ is a half line $y \geq 0$. Including the time direction, which we parametrize by \mathbb{R}' (another copy of \mathbb{R}), the full spacetime is $\hat{X} = \mathbb{R}' \times \mathbb{R} \times S^2$ and the world volume of the topological insulator is $X = \mathbb{R}' \times \mathbb{R}_+ \times S^2$.

We consider a gauge field that has one unit of magnetic flux on S^2 , and also one unit of electric flux on $\mathbb{R}' \times \mathbb{R}$:

$$\int_{\mathbb{R}' \times \mathbb{R}} dt dx F_{01} = \frac{2\pi}{e}. \quad (2.28)$$

Such a gauge field has the value of 1 of

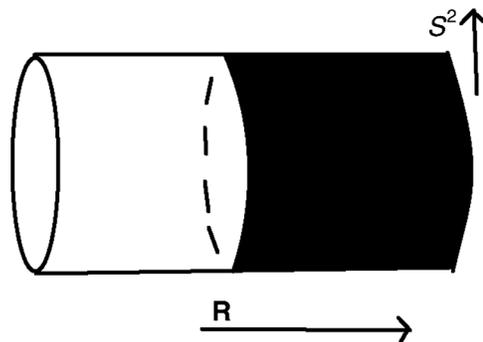


FIG. 3. A topological insulator supported on the right half of $\mathbb{R} \times S^2$ (solid black region).

¹¹We write $\hat{A}(R)$ as an abbreviation for the more usual $\int_X \hat{A}(R)$.

$$\hat{P} = \frac{e^2}{32\pi^2} \int_{\hat{X}} d^4x \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta}. \quad (2.29)$$

Note that this is an integral over the whole spacetime $\hat{X} = \mathbb{R}' \times \mathbb{R} \times S^2$. By contrast, we define P [Eq. (2.23)] as a similar integral over only the world volume $\mathbb{R}' \times \mathbb{R}_+ \times S^2$ of the topological insulator. So \hat{P} is a topological invariant but P is not. If the unit of electric flux in Eq. (2.28) is localized far to the left in Fig. 3, then $\hat{P} = 1$ but $P = 0$. If it is localized far to the right, then $\hat{P} = P = 1$.

To consider in this context the formula

$$|Z_\psi|(-1)^{\mathfrak{F}} \quad (2.30)$$

for the path integral measure, we compactify the time direction and go to Euclidean signature, since the formula was derived in that case. For example, to compute a thermal partition function $\text{Tr} \exp(-\beta H)$, with H the Hamiltonian of the system, we take the time direction to be a circle S of circumference β . As long as the electric flux is deep inside or outside the topological insulator, but at any rate far from its boundary, the boundary fermions do not “see” this flux. Generically (for a generic value of $\oint_S A$, to be precise) the boundary fermions have no zero mode, so $|Z_\psi| > 0$. On the other hand, $\mathfrak{F} = 0$ if the instanton is outside the topological insulator (that is if $P = 0$) and $\mathfrak{F} = 1$ if the instanton is inside the topological insulator (that is if $P = 1$). The path integral measure is always real and so \mathbb{T} conserving, but it passes smoothly from positive to negative values as the instanton is brought inside the topological insulator. The jump in \mathfrak{F} from 0 to 1, and thus the change in sign of the path integral measure, occurs precisely when an eigenvalue of the boundary Dirac operator passes through zero, or in other words when $|Z_\psi| = 0$.

The fact that the path integral is negative when $P = 1$ can be described by saying that the path integral measure of the topological insulator contains a factor $(-1)^P = \exp(i\pi P)$. This gives a precise meaning to the statement that $\theta = \pi$ inside the topological insulator.

Another and perhaps less technical way to give a precise meaning to the statement that $\theta = \pi$ has already been described in the literature (Rosenberg and Franz, 2010). Let us go back to the Lorentz signature picture in which the full spacetime is $\hat{X} = \mathbb{R}' \times \mathbb{R} \times S^2$, the topological insulator lives on $X = \mathbb{R}' \times \mathbb{R}_+ \times S^2$, and the boundary Dirac fermion lives on $\partial X \cong \mathbb{R}' \times \{0\} \times S^2$. To find the quantum states of the boundary fermions, we have to find the eigenmodes of the 2D Dirac operator on S^2 , and then quantize them. In the presence of a unit magnetic flux on S^2 , the 2D Dirac operator has a single charge $-e$ zero mode of one 2D chirality, and a single charge $+e$ zero mode of the opposite 2D chirality. (The existence of these zero modes is guaranteed by the 2D version of the Atiyah-Singer index theorem. The two zero modes are complex conjugate; complex conjugation reverses both the charge and the 2D chirality.) As in Jackiw and Rebbi (1976), quantization of these modes gives a pair of states with electric charge $\pm e/2$. The fact that the boundary of the topological insulator, when pierced by a unit of flux, supports half-integral electric charge can be understood if we

imagine creating this unit of magnetic flux by dragging a magnetic monopole of unit charge into the topological insulator. (We assume that the monopole starts far to the left of the topological insulator in Fig. 3, with its flux in the initial state going off to the left and not entering the topological insulator.) When the monopole enters the material, a unit of flux appears on its surface and accordingly the charge on its surface jumps by a half-integral amount. But the electric charge of the monopole also jumps by a half-integral amount when it enters the topological insulator, because a magnetic monopole in a world with $\theta = \pi$ has half-integral electric charge (Witten, 1979). It has indeed been shown in a lattice model of a topological insulator that when a magnetic monopole enters a topological insulator, half-integral charge appears on the surface of the topological insulator and the monopole itself acquires half-integral charge (Rosenberg and Franz, 2010). This gives a direct interpretation of what it means to say that the topological insulator has $\theta = \pi$.

6. Why is the partition function equal to $(-1)^{\mathfrak{F}}$?

Part of the previous story is that on a compact four-manifold X without boundary, with a background $U(1)$ gauge field A , the topological field theory associated with a topological insulator has a partition function $(-1)^{\mathfrak{F}}$. The interested reader may wonder how this is related to other known descriptions of a topological insulator.

We use the following standard characterization of the phase transition between a topological insulator and an ordinary one. This transition occurs when the mass parameter m of a $D = 4$ Dirac fermion ψ passes through zero and changes sign.

First we recall the definition of the index of the Dirac operator. On an oriented four-manifold, we define the chirality operator

$$\bar{\gamma} = \frac{1}{4!} \epsilon^{ijkl} \gamma_i \gamma_j \gamma_k \gamma_l, \quad (2.31)$$

where ϵ^{ijkl} is the Levi-Civita tensor. We write $\psi = \psi_+ + \psi_-$ where $\psi_\pm = (1/2)(1 \pm \bar{\gamma})\psi$ satisfy $\bar{\gamma}\psi_\pm = \pm\psi_\pm$; we say that ψ_+ and ψ_- have positive or negative chirality. $\bar{\gamma}$ anticommutes with the Dirac operator \mathcal{D} , so \mathcal{D} maps ψ_+ to ψ_- and vice versa. We write n_+ for the dimension of the space of zero modes of \mathcal{D} acting on ψ_+ and n_- for the dimension of the space of zero modes of \mathcal{D} acting on ψ_- . The index of the Dirac operator is $\mathfrak{F} = n_+ - n_-$. It is a basic topological invariant.

To prove topological invariance of \mathfrak{F} , observe that $\bar{\gamma}$ anticommutes with \mathcal{D} , so that if $\mathcal{D}\psi = \lambda\psi$ with $\lambda \neq 0$, then $\mathcal{D}(\bar{\gamma}\psi) = -\lambda\bar{\gamma}\psi$ and $\bar{\gamma}\psi$ must therefore be linearly independent if $\lambda \neq 0$, and hence the chiral projections $\psi_\pm = (1/2)(1 \pm \bar{\gamma})\psi$ are nonzero. Consider the “Hamiltonian” $H = \mathcal{D}^2$. Zero modes of \mathcal{D} are the same as zero modes of H , and more generally eigenstates of H , obeying $H\psi = E\psi$ are linear combinations of solutions of $\mathcal{D}\psi = \pm\sqrt{E}\psi$. If $E \neq 0$, a solution of $\mathcal{D}\psi = \pm\sqrt{E}\psi$ has nonzero chiral projections ψ_\pm with the same E . Hence nonzero energy levels of H occur in pairs with positive and negative chirality. When we vary the gauge field or metric on X , all that can happen to the space of zero energy states is that a pair of states of opposite

chirality can move to or from zero energy. In the process, n_+ and n_- change, but $\mathfrak{F} = n_+ - n_-$ is unchanged.

Let us assume that X and A are such that $\mathfrak{F} \neq 0$ and for definiteness take $\mathfrak{F} > 0$. Generically, for $\mathfrak{F} > 0$, ψ_+ will have \mathfrak{F} zero modes and ψ_- will have none. We now assume that we are in this situation. What about zero modes of the oppositely charged fermion field $\bar{\psi}$? Writing $\bar{\psi}_-$ and $\bar{\psi}_+$ for its components of definite chirality, $\bar{\psi}_-$ will have \mathfrak{F} zero modes and $\bar{\psi}_+$ will have none. (The zero modes of $\bar{\psi}_-$ are complex conjugates of the zero modes of ψ_+ .)

Because of these $2\mathfrak{F}$ zero modes, the fermion path integral vanishes if $m = 0$. However, for $m \neq 0$, the action has a term

$$I = \int d^4x (\cdots - m \bar{\psi}_- \psi_+ + \cdots), \quad (2.32)$$

and accordingly the integrand e^{-I} of the path integral has a factor

$$\exp\left(m \int d^4x \bar{\psi}_- \psi_+\right). \quad (2.33)$$

Expanding this factor, a $\bar{\psi}_-$ zero mode and a ψ_+ zero mode can be lifted at the cost of a factor of m .

Accordingly, for small m , the path integral is proportional to $m^{\mathfrak{F}}$. Therefore, if the path integral is positive definite for, say, $m > 0$, then its sign for $m < 0$ is $(-1)^{\mathfrak{F}}$. (Zero modes beyond the ones we assumed would contribute an even power of m and thus would not affect this sign.) This sign is the universal part of the answer and is the partition function of the topological field theory that one extracts as the long distance limit of the topological insulator.

We treat several similar problems later in a slightly more precise way including an explanation of the regularization; see, for example, Sec. III.D.

7. The index and cobordism invariance

We defined the Dirac index \mathfrak{F} in Sec. II.A.6 by using the fact that, on an oriented four-dimensional spin manifold X without boundary, the equation $D\psi = 0$ for a fermion zero mode splits as separate equations for the chiral projections ψ_+ and ψ_- .

The same is true in any even dimension, and the original Atiyah-Singer index theorem (Atiyah and Singer, 1963, 1968) gives a cohomological formula for \mathfrak{F} . For a U(1) gauge field in $D = 4$, the formula is

$$\mathfrak{F} = \hat{A} - P. \quad (2.34)$$

Here P is the instanton number defined in Eq. (2.23), and

$$\hat{A} = -\frac{1}{48} \int_X \frac{\text{tr} R \wedge R}{(2\pi)^2}, \quad (2.35)$$

with R being the Riemann tensor regarded as a matrix-valued two-form. Equation (2.34) immediately implies¹² that \mathfrak{F} is a cobordism invariant: it vanishes if X is the boundary of an

¹²Here we are reversing the original logic. In the original proof of the index theorem (Atiyah and Singer, 1963), it was shown first that the index is a cobordism invariant, and this knowledge was used to deduce a formula for it.

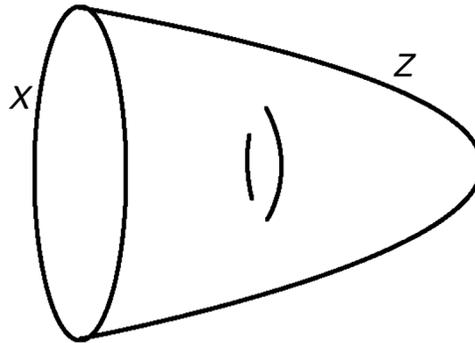


FIG. 4. A topological invariant \mathfrak{F} defined for a manifold X is said to be a cobordism invariant if it vanishes whenever X is the boundary of a manifold Z of one dimension more. If X has some structure [such as an orientation, a spin structure, or a U(1) gauge field] that is required in defining \mathfrak{F} , then this structure is required to extend over Z .

oriented five-dimensional spin manifold Z over which the U(1) gauge field extends (Fig. 4). For in that situation $\hat{A} = P = 0$, so $\mathfrak{F} = 0$. For example, one has

$$\hat{A} = -\frac{1}{48} \int_X \frac{\text{tr} R \wedge R}{(2\pi)^2} = -\frac{1}{48} \int_Z d \frac{\text{tr} R \wedge R}{(2\pi)^2} = 0 \quad (2.36)$$

(here one uses Stokes's theorem to express an integral over X as an integral over Z , and then one uses the Bianchi identity $DR = 0$ to prove that $d \text{tr} R \wedge R = 0$). A similar argument shows that $P = 0$.

Since \mathfrak{F} is a cobordism invariant, so is $(-1)^{\mathfrak{F}}$, which we have seen to be the partition function (on an orientable manifold) of the bulk sTQFT associated with a topological insulator. As we proceed, it will become clear that partition functions of sTQFT's associated with free fermion states of matter are always cobordism invariants. This was conjectured by Kapustin *et al.* (2014).

In general,¹³ starting with such a U(1)-valued cobordism invariant [a homomorphism from an appropriate cobordism group to U(1)], one can always construct an invertible topological field theory with the given invariant as its partition function. The rough idea behind this statement is that cobordism invariance implies relations among partition functions that one would expect on physical grounds in a topological field theory. The simplest example of such a statement is illustrated in Fig. 5. Another example of how cobordism invariance implies a physical property (unitarity) is described at the end of Sec. II.A.9.

It appears to be unknown whether all unitary invertible TQFTs or sTQFTs are associated with cobordism invariants. Since this is unknown, we do not really know if the fact that the usual free fermion phases are associated with cobordism invariants is a consequence just of the fact that they are gapped and have no topological order in bulk, or is more special. At any rate, cobordism invariance is much stronger than topological invariance.

¹³For a brief explanation, see the discussion of Eq. (5.4) in D. Freed (2014).

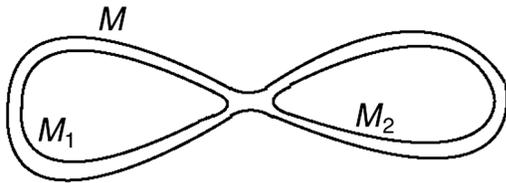


FIG. 5. The simplest illustration of the statement that cobordism invariance leads to relations that are natural in topological field theory. The *connected sum* M of two D manifolds M_1 and M_2 is made by cutting small holes out of M_1 and M_2 and gluing them together along their boundaries. If the space of physical states on a sphere S^{D-1} is one dimensional (as expected in a unitary topological field theory), one can deduce a universal relation between partition functions: $Z_M = gZ_{M_1}Z_{M_2}$, where g is a constant characteristic of the theory. Cobordism invariance implies such a relation with $g = 1$, because M is cobordant to the disjoint union of M_1 and M_2 . This cobordism is sketched in this figure (for the case $D = 1$ with all manifolds being circles). Topological field theories derived from cobordism invariants always satisfy the condition that the space of physical states on any $(D - 1)$ manifold has a dimension of 1, which is why they lead to a relation of the given kind.

8. The APS index theorem

To define the Dirac index \mathfrak{S} in Sec. II.A.6, we used the fact that on an oriented four-manifold, the Dirac equation $\mathcal{D}\psi = 0$ splits as separate equations for the chiral projections ψ_+ and ψ_- of ψ .

The APS index theorem (Atiyah, Patodi, and Singer, 1975) is a generalization of the original Atiyah-Singer index theorem (Atiyah and Singer, 1963, 1968) to a manifold X with boundary. In trying to find such a generalization, one immediately finds that there is no convenient local boundary condition along $Y = \partial X$ that can be used in defining an index. The Dirac operator admits the local boundary condition

$$\gamma \cdot n\psi = \pm\psi, \quad (2.37)$$

where n is the unit normal to Y and $\gamma \cdot n = \gamma^\mu n_\mu$. This boundary condition is the most general one that preserves the local rotation symmetry of the boundary and it is physically sensible in the sense that with this boundary condition the Dirac operator is Hermitian. However, because $\gamma \cdot n$ anticommutes with the chirality operator $\tilde{\gamma}$, this boundary condition mixes the two fermion chiralities and so does not allow one to define a Dirac index.

What can we do to define a Dirac index in this situation? The APS construction starts by replacing a manifold X that has a boundary with a manifold X' without boundary that is not compact but has a complete Riemannian metric. This is done by picking a Riemannian metric on X , and, in particular, on Y , and building X' by gluing to the boundary of X a semi-infinite tube $Y \times \mathbb{R}_+$, where \mathbb{R}_+ is the half line $y \geq 0$ (Fig. 6). The Riemannian metric in the tube is taken to be the obvious one $dy^2 + g_Y$, where g_Y is the chosen metric on Y .

Because X' has no boundary, the Dirac equation on X' does not mix positive and negative chirality spinors. Therefore, one can define spaces H_+ and H_- of *square-integrable* solutions of the Dirac equation on X' of positive and negative chirality.

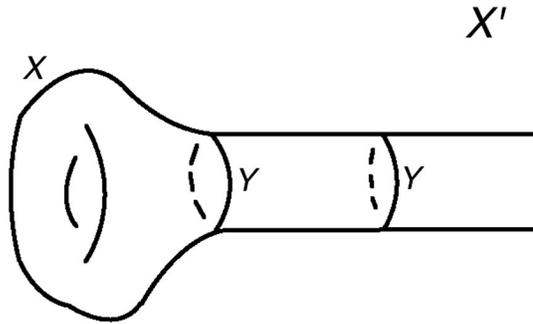


FIG. 6. The complete but not compact manifold X' is built by gluing onto the manifold X , which has a boundary Y , and a semi-infinite tube $Y \times \mathbb{R}_+$.

Their dimensions are integers n_+ and n_- . Assume for the moment that the metric g_Y of Y is sufficiently generic so that the Dirac operator \mathcal{D}_Y of Y has no zero modes. Then the index \mathfrak{S} is defined in the usual way as $\mathfrak{S} = n_+ - n_-$.

The APS index theorem gives a formula for \mathfrak{S} :

$$\mathfrak{S} = \hat{A} - P - \frac{\eta}{2}. \quad (2.38)$$

Here η is the η invariant (2.17) of the Dirac operator \mathcal{D}_Y on Y . Thus $-\eta/2$ appears as a sort of boundary correction to the index theorem.

On a manifold X without boundary, the index of the Dirac operator is an integer and moreover is a topological invariant. When X has a boundary and the index is defined with APS boundary conditions, it is still an integer but it is no longer a topological invariant. The APS index formula makes this clear. As the metric and gauge field on X and Y are varied so that an eigenvalue of \mathcal{D}_Y passes through 0, η jumps by ± 2 , so \mathfrak{S} must jump by ∓ 1 . To understand explicitly how this happens, we look at the Dirac equation in the tube $Y' = Y \times \mathbb{R}_+$. The Dirac operator $\mathcal{D}_{Y'}$ on Y' is defined using gamma matrices γ_y and γ_i , $i = 1, 2$, and 3 with

$$\mathcal{D}_{Y'} = i \left(\gamma^y \frac{\partial}{\partial y} + \sum_{i=1}^3 \gamma^i \frac{D}{Dx^i} \right), \quad (2.39)$$

where x^i , $i = 1, \dots, 3$ are local coordinates on Y . So the equation $\mathcal{D}_{Y'}\psi = 0$ is

$$\left(\frac{\partial}{\partial y} - i \sum_{i=1}^3 i\gamma^y \gamma^i \frac{D}{Dx^i} \right) \psi = 0. \quad (2.40)$$

We want to express this in terms of the Dirac operator on Y , which would be

$$\mathcal{D}_Y = i \sum_{i=1}^3 \hat{\gamma}^i \frac{D}{Dx^i}. \quad (2.41)$$

Here $\hat{\gamma}_i$ are purely three-dimensional gamma matrices obeying $\{\hat{\gamma}_i, \hat{\gamma}_j\} = 2g_{ij}$. In a locally Euclidean frame, we can take them to be $\hat{\gamma}_i = \sigma_i$ (where σ_i are the Pauli matrices) and so to obey

$$\hat{\gamma}_1 \hat{\gamma}_2 \hat{\gamma}_3 = i\epsilon_{123} \quad (2.42)$$

(and cyclic permutations), where ϵ_{ijk} is the Levi-Civita tensor of Y .

In the definition (2.31) of the chirality operator of X , we orient X so that $\epsilon_{yijk} = \epsilon_{ijk}$. The matrices $\gamma'_i = i\gamma_y \gamma_i$ appearing in Eq. (2.40) obey the same Clifford algebra $\{\gamma'_i, \gamma'_j\} = 2g_{ij}$ as $\hat{\gamma}_i$. They also satisfy in a local Euclidean frame $\gamma'_1 \gamma'_2 \gamma'_3 = \bar{\gamma} \cdot i\epsilon_{123}$. From this one can deduce that the matrices $\bar{\gamma} \gamma'_i$ obey the same Clifford algebra and the same equation (2.42) as the purely 3D Dirac matrices $\hat{\gamma}_i$, and so we can set $\bar{\gamma} \gamma'_i = \hat{\gamma}_i$. Then we can rewrite the $D = 4$ Dirac equation (2.40) in the form

$$\left(\frac{\partial}{\partial y} - \bar{\gamma} \mathcal{D}_Y \right) \psi = 0. \quad (2.43)$$

So if ψ_0 is a spinor field on Y satisfying $\mathcal{D}_Y \psi_0 = \lambda \psi_0$, and moreover $\bar{\gamma} \psi_0 = a \psi_0$, $a = \pm 1$, then we can solve the $D = 4$ Dirac equation with

$$\psi = \exp(a\lambda y) \psi_0. \quad (2.44)$$

To get a square-integrable solution of this equation, we need $\lambda < 0$ if $a = 1$ or $\lambda > 0$ if $a = -1$.

Now we can complete the construction by explaining APS boundary conditions on the original manifold with boundary X . Let \mathcal{H} be the space of all spinor fields on Y , and decompose $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$ where \mathcal{H}_+ and \mathcal{H}_- are spanned, respectively, by the eigenstates of \mathcal{D}_Y with $\lambda > 0$ or with $\lambda < 0$. Let ψ_+ and ψ_- be as usual the positive and negative chirality parts of a 4D spinor ψ . Then APS boundary conditions on X say that the restrictions of ψ_{\pm} to $Y = \partial X$ obey

$$\psi_+|_Y \in \mathcal{H}_-, \quad \psi_-|_Y \in \mathcal{H}_+. \quad (2.45)$$

The point is that solutions to the Dirac equation on X that can be extended to square-integrable solutions on X' are precisely those that obey this condition. So we can compute the index \mathfrak{S} either in terms of solutions of the Dirac equation on X that obey this APS boundary condition or in terms of square-integrable solutions on X' .

This should make it clear how \mathfrak{S} can fail to be a topological invariant. When an eigenvalue of \mathcal{D}_Y passes through 0, the spaces \mathcal{H}_{\pm} change discontinuously, the APS boundary conditions change discontinuously, and the index \mathfrak{S} also changes discontinuously.

As long as \mathcal{D}_Y has no zero modes, the spectrum of the Dirac operator on X with APS boundary conditions varies continuously with the metric and gauge fields on X (and, in particular, on Y) and \mathfrak{S} is constant. In the above, we defined \mathfrak{S} , \mathcal{H}_+ , and \mathcal{H}_- only in the generic case that \mathcal{D}_Y has no zero modes. The definitions can be slightly modified so that the APS index formula (2.38) is valid in general, with η defined using Eq. (2.19).

We now make a few more remarks on the APS index theorem. First, in Eq. (2.38), we have written the APS index theorem for $D = 4$. However, this theorem holds in any dimension D . In general, $\hat{A} - P$ has to be replaced by the bulk contribution to the index that is given by the original

Atiyah-Singer index theorem (Atiyah and Singer, 1963, 1968). This bulk contribution can be written in general as $\int_X \hat{A}(R) \text{tr} e^F$, where F is the Yang-Mills field strength or curvature and the trace is taken in the representation that the fermions live in. The boundary contribution is always $-\eta/2$. However, for odd D , the bulk contribution vanishes¹⁴ and the APS index theorem reduces to

$$\mathfrak{S} = -\frac{\eta}{2}. \quad (2.46)$$

Since \mathfrak{S} is always an integer, this shows that for odd D , if a $D - 1$ manifold Y is the boundary of some D -dimensional X over which the relevant structures extend (so that an APS index theorem on X is available), then the η invariant on Y obeys $\eta/2 \in \mathbb{Z}$. In later applications to Majorana fermions, we will want to write this formula in terms of $\mathcal{I} = 2\mathfrak{S}$, $\eta = 2\eta$:

$$\mathcal{I} = -\frac{\eta}{2}. \quad (2.47)$$

Another important consequence of the APS index formula is the relationship between η and the Chern-Simons function

$$\text{CS}(A) = \frac{e^2}{4\pi} \int_Y d^3 x \epsilon^{ijk} A_i \partial_j A_k. \quad (2.48)$$

$\text{CS}(A)$ is only gauge invariant mod $2\pi\mathbb{Z}$ [which is why the coupling k in Eq. (1.1) is an integer in any material that can be described by an effective action for A only], but its variation $\delta \text{CS}(A)$ in a continuous change in A is completely gauge invariant. If the spin manifold Y on which we are trying to define $\text{CS}(A)$ is the boundary of some X (over which A and the spin structure extend), then

$$\text{CS}(A) = 2\pi P \text{ mod } 2\pi\mathbb{Z}, \quad (2.49)$$

where P is the instanton number, integrated over X that appears in the index formula (2.38). This formula depends on the choice of a four-dimensional spin manifold X with boundary Y , along with an extension of A over X , but it is independent of these choices mod $2\pi\mathbb{Z}$. To prove this, suppose that X and X' are spin manifolds with the same boundary Y ; let P and P' be the instanton numbers integrated over X and X' , respectively. Then the difference $P - P'$ is an integer, because it is the instanton number integrated on a closed oriented four-dimensional spin manifold \hat{X} built by gluing X' to X with a reversal of orientation of X' (Fig. 7).

Similarly, one can define a gravitational Chern-Simons term CS_{grav} , also defined mod $2\pi\mathbb{Z}$, that satisfies

$$\text{CS}_{\text{grav}} = 2\pi \frac{\hat{A}}{2}. \quad (2.50)$$

¹⁴The reason for this is that R and F are both two-forms, so $\hat{A}(R) \text{tr} e^F$ is a linear combination of forms of even degree. If D is odd, then $\hat{A}(R) \text{tr} e^F$ has no degree D piece that could be a bulk contribution in the index theorem.

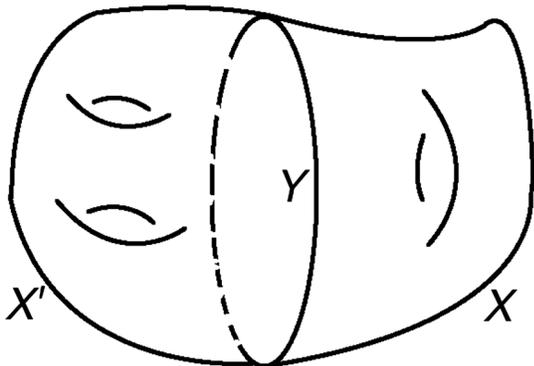


FIG. 7. By gluing together two oriented four-manifolds X and X' that have the same boundary Y (after reversing the orientation of X' so that the orientations are compatible), we build a closed oriented four-manifold \hat{X} . \hat{X} is a spin manifold if X and X' have spin structures that agree along Y .

(It is understood that \hat{A} is integrated over X . The formula is written in terms of $\hat{A}/2$ because \hat{A} is even on a four-dimensional spin manifold without boundary. This fact will be important in Sec. II.B.) Given these definitions, and assuming a suitable X exists, the APS index theorem for X implies

$$\frac{\eta}{2} = \frac{\text{CS}(A)}{2\pi} - 2 \frac{\text{CS}_{\text{grav}}}{2\pi} \text{ mod } \mathbb{Z}. \quad (2.51)$$

The statement is true only mod \mathbb{Z} because it does not take into account the \mathfrak{F} term in the index formula. [Also, $\text{CS}(A)/2\pi$ is only well defined mod \mathbb{Z} so the statement really only makes sense mod \mathbb{Z} .]

In deducing this relation between η and Chern-Simons, we assumed that the manifold Y on which we are studying these objects is that boundary of some X on which we can use the APS index theorem. This is true if Y is a three-dimensional spin manifold with a $U(1)$ gauge field, but in a more general problem to which we might apply the APS index theorem, it will not always be true. Even then, a somewhat weaker statement is true. In general, the difference between $\eta/2$ and the corresponding Chern-Simons invariant is constant (but is not necessarily an integer) as one varies the metric and gauge field on Y , keeping away from jumps in η . To state the relation in generality, we assume that Y has dimension d , and we generically write $\text{CS}(g, A)/2\pi$ for the Chern-Simons invariant that is related to the curvature polynomial¹⁵ $\hat{A} \text{tr} e^F$ that appears in the index theorem in $D = d + 1$ dimensions [generalizing $\hat{A} - P$ for $U(1)$ gauge theory in four dimensions]. Then under any continuous variation of the metric and gauge field (g, A) on Y , modulo even integer jumps in η , we have

$$\frac{\delta\eta}{2} = \frac{\delta\text{CS}(g, A)}{2\pi}. \quad (2.52)$$

This is proved by applying the APS index theorem to the D manifold $X = Y \times [0, 1]$, with one pair of fields (g, A) at

¹⁵A curvature polynomial is just a polynomial in the Riemann curvature R and the gauge theory field strength or curvature F .

one end, and a slightly different pair $(g', A') = (g + \delta g, A + \delta A)$ at the other end. Including the jumps in η , the statement is

$$\frac{\eta}{2} = \frac{\text{CS}(g, A)}{2\pi} + \text{const mod } \mathbb{Z}. \quad (2.53)$$

The constant, which is valued in \mathbb{R}/\mathbb{Z} , is an important topological invariant in some situations. It can appear in global anomalies, because of the relation of global anomalies to η (Sec. IV).

In a context in which one only cares about $\eta/2 \text{ mod } \mathbb{Z}$, and assuming X exists [as it always does for $U(1)$ gauge theory on a three-dimensional spin manifold], Eq. (2.51) can be used to replace η with the more elementary invariants $\text{CS}(A)$ and CS_{grav} . However, the boundary fermions of the 3D topological insulator have a partition function $Z_\psi = |Z_\psi| \exp(-i\pi\eta/2)$, which depends on $(\eta/2) \text{ mod } 2\mathbb{Z}$, not mod \mathbb{Z} . The replacement (2.51) is not adequate for describing those boundary fermions. If we try to make that replacement, we will lose control of the overall sign of Z_ψ , which as we have seen is very important in describing the topological insulator.

9. Reflection positivity

Another question is why the formula $|Z_\psi|(-1)^{\mathfrak{F}}$ for the partition function is consistent with unitarity that is with reflection positivity.

Actually, in the specific case at hand, we can give a trivial answer to this question. We have built $|Z_\psi|(-1)^{\mathfrak{F}}$ as a product of two factors $|Z_\psi| \exp(\mp i\pi\eta/2)$ and $\exp\{\pm i\pi[P - \hat{A}(R)]\}$. We showed the first factor to be reflection positive in the discussion of Fig. 2, and the second factor is also compatible with reflection positivity; otherwise, gauge and gravitational θ angles would in general not be physically sensible. [Concretely, $\exp\{\pm i\theta[P - \hat{A}(R)]\}$ is consistent with reflection positivity for any real θ because P and $\hat{A}(R)$ are odd under reflection and vanish for any reflection-symmetric gauge field and metric.]

However, with a view to other cases, we directly establish the reflection positivity of $|Z_\psi|(-1)^{\mathfrak{F}}$. We want to show that $|Z_\psi|(-1)^{\mathfrak{F}}$ is real and non-negative whenever a spacetime X , with boundary Y , is built by gluing together, along their common boundary M , two mirror image copies of some manifold X_0 , as in Fig. 8. If X has boundary Y , then this boundary is likewise built by gluing together mirror image pieces. In that case, X_0 is really a manifold with corners, as in Fig. 8. We assume that the gauge field, metric, and spin structure on X are invariant under the reflection that exchanges the two copies of X_0 .

Since $|Z_\psi|$ is trivially real and non-negative, to show reflection positivity of $|Z_\psi|(-1)^{\mathfrak{F}}$ it suffices to show that $\mathfrak{F} = 0$ in a reflection-symmetric situation. A reflection reverses the chirality of a $D = 4$ fermion, so in any reflection-symmetric situation, the index \mathfrak{F} computed with APS boundary conditions vanishes. This completes the proof.

Alternatively, we could prove the vanishing of $(-1)^{\mathfrak{F}}$ in the situation of Fig. 8 by using cobordism invariance. This explanation has the advantage that it applies almost

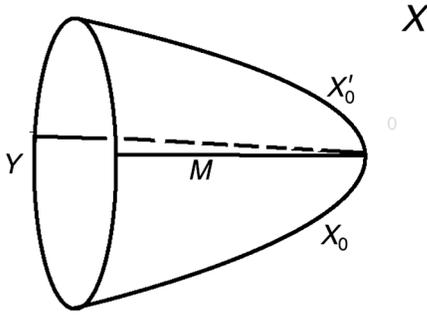


FIG. 8. A spacetime X with boundary Y that is built by gluing together mirror image copies of some manifold X_0 . The copies are denoted X_0 and X_0' . Y is likewise the union of mirror image pieces. X_0 has two types of boundaries: M , along which it is glued to X_0' , and also its intersection with $Y = \partial X$. X_0 has corners along $M \cap Y$.

directly¹⁶ to the other examples that we study in the rest of this paper, since they all involve cobordism-invariant sTQFTs. If X is constructed as in Fig. 8 by “doubling” some manifold X_0 , then X is always the boundary of a manifold Z that can be constructed as follows. Let M be the boundary of X_0 along which the gluing occurs, as in the figure. Let $Z_0 = X_0 \times I$, where I is the unit interval $[0, 1]$, and in X_0 , collapse $M \times I$ to M . This gives a manifold Z with boundary X . The local picture, near M , is that Y looks like $M \times \mathbb{R}_+$ and Z looks like $M \times \mathbb{R}_+^2$, where \mathbb{R}_+ and \mathbb{R}_+^2 are half-spaces in \mathbb{R} or \mathbb{R}^2 . The existence of Z ensures vanishing of a cobordism invariant of X such as $(-1)^{\mathfrak{F}}$.

10. Running the story in reverse

Finally, it may be illuminating to run this story in reverse. We do this by first postulating in bulk a sTQFT, and then determining what sort of symmetry-preserving boundary state it can have.

We consider a $D = 4$ sTQFT, with $U(1)$ symmetry, that is characterized by saying that its partition function on a compact $D = 4$ spin manifold X , in the presence of a background $U(1)$ gauge field A , is $(-1)^{\mathfrak{F}}$, \mathfrak{F} being the index of the Dirac operator. This is the partition function of an sTQFT because the Atiyah-Singer index theorem shows that \mathfrak{F} is a cobordism invariant and in general such a $U(1)$ -valued cobordism invariant is the partition function of a topological field theory.

A topological insulator can have a spatial boundary, and in fact in the real world (as opposed to thought experiments) spatial boundaries are unavoidable. So we would like to extend our sTQFT to make sense on a manifold X with boundary. This is not straightforward. The index \mathfrak{F} of the Dirac operator cannot be defined with local boundary conditions on ∂X , but it can be defined with global APS boundary conditions. But \mathfrak{F} defined this way is not a topological invariant, and moreover as the background data (the metric

¹⁶To do this properly requires some important details about the gluing of the spin or pin structures of X_0 and X_0' that have been described to me by D. Freed. These details have to be used in establishing that the spin or pin structure of X extends over the manifold Z constructed below.

and gauge field on X) are varied, \mathfrak{F} jumps in an unphysical way.

The standard boundary state of the topological insulator has massless Dirac fermions living on ∂X . In the presence of these modes, the jumps in \mathfrak{F} make sense physically. The formula $|Z_\psi|(-1)^{\mathfrak{F}}$ that we have described is a physically sensible formula for the partition function of a combined system consisting of an sTQFT in bulk coupled to Dirac fermions on the boundary. The jumps in \mathfrak{F} compensate for the problem in defining the sign of the fermion partition function.

B. Topological superconductor in $d = 3$

We now turn to the topological superconductor in $3 + 1$ dimensions, formulated for now on an orientable manifold only.

1. Fermions and Pfaffians

The topological superconductor is gapped in bulk (at least for fermionic excitations), but on the boundary it supports a two-component Majorana fermion ψ . The main difference from the analysis of the topological insulator is that ψ does not carry any conserved charge [except $(-1)^F$, the operator that counts fermions mod 2]. In particular, and in contrast to the topological insulator, there is no distinction between fermion modes of positive and negative charge. This necessitates a few changes in the details of our explanations, although the main ideas are the same. In the absence of a conserved additive charge carried by the fermions, we have to use fermion Pfaffians rather than determinants. Moreover, we need a more careful argument to explain a doubling of the spectrum that is rather trivial in the $U(1)$ -invariant case.

A 3D Majorana fermion ψ in Lorentz signature has two real components—they transform in the real two-dimensional representation of $\text{Spin}(1, 2) \cong \text{SL}(2, \mathbb{R})$. After continuing to Euclidean signature, ψ still has two independent components, although they are no longer real fields; they transform in the two-dimensional pseudoreal representation of $\text{Spin}(3) \cong \text{SU}(2)$. [This is simply the spin $1/2$ representation of $\text{SU}(2)$.] We cannot claim that ψ is real, because this would not be compatible with $\text{SU}(2)$ symmetry, but nonetheless the complex conjugate of ψ does not appear in the formalism.¹⁷ When we write the Euclidean version of the action for ψ

$$I_\psi = \int d^3x \bar{\psi} D \psi, \quad (2.54)$$

despite appearances, there is actually no complex conjugation involved in the definition of $\bar{\psi}$. Indeed, for a Majorana fermion continued to Euclidean space, the definition is just $\bar{\psi}_\alpha = \varepsilon_{\alpha\beta} \psi^\beta$, where $\varepsilon_{\alpha\beta}$ is the $\text{Spin}(3) = \text{SU}(2)$ invariant antisymmetric tensor. Thus I_ψ is simply an antisymmetric bilinear expression in the fermion field ψ . It is antisymmetric because of Fermi statistics.

¹⁷This leads to no contradiction because fermion integration is really an algebraic procedure. We define it for an odd variable ψ by saying $\int d\psi \cdot 1 = 0$, $\int d\psi \cdot \psi = 1$, without ever having to claim that ψ is real or to mention its complex conjugate.

It is useful to carefully distinguish between operators and bilinear forms. The usual Hermitian Dirac operator $\mathcal{D} = i\mathcal{D} = i\sum_{i=1}^3 \gamma^i D_i$ is an “operator,” schematically $\psi^\alpha \rightarrow \sum_{\beta} i\mathcal{D}_{\beta}^{\alpha} \psi^{\beta}$. A spinor index of ψ^α can be raised or lowered using the antisymmetric tensor $\varepsilon_{\alpha\beta}$, and if we do this we get the antisymmetric bilinear form $\mathcal{D}_{\gamma\beta} = \varepsilon_{\gamma\alpha} i\mathcal{D}_{\beta}^{\alpha}$.

The path integral of the ψ field is formally the Pfaffian of this antisymmetric form:

$$Z_{\psi} = \text{Pf}(\mathcal{D}). \quad (2.55)$$

Whenever one has a theory of fermions, the quadratic part of the fermion action is always antisymmetric by virtue of Fermi statistics and the corresponding fermion path integral is the Pfaffian of the antisymmetric bilinear form that appears in the action.

When can one speak of a fermion “determinant” rather than a Pfaffian? This is possible if there is a conserved U(1) charge¹⁸ and a splitting of ψ into fields ψ_+ and ψ_- of positive and negative charge, respectively. U(1) symmetry then ensures that in a basis $\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}$ the form \mathcal{D} will be block off diagonal

$$\mathcal{D} = \begin{pmatrix} 0 & \mathcal{D} \\ -\mathcal{D}^{\text{tr}} & 0 \end{pmatrix}. \quad (2.56)$$

Here \mathcal{D} is not subject to any condition of symmetry or antisymmetry, and \mathcal{D}^{tr} is its transpose. In such a situation, one has

$$\text{Pf}(\mathcal{D}) = \det(\mathcal{D}), \quad (2.57)$$

and so the fermion path integral can be written as a determinant.

Going back to the Hermitian Dirac operator \mathcal{D} for the Majorana fermion, its eigenvalues have even multiplicity because of a Euclidean analog of Kramers doubling.¹⁹ If a c -number spinor χ is an eigenvector of the Dirac operator,

$$\mathcal{D}\chi = \lambda\chi, \quad \lambda \in \mathbb{R}, \quad (2.58)$$

then $\tilde{\chi}$ defined by $\tilde{\chi}^\alpha = \varepsilon^{\alpha\beta} \chi_\beta^*$ is an eigenvector with the same eigenvalue

$$\mathcal{D}\tilde{\chi} = \lambda\tilde{\chi}. \quad (2.59)$$

Here we write χ_β^* for the complex conjugate of χ^β . Although there is no field in the Euclidean field theory that corresponds to the complex conjugate of ψ , when we expand ψ in eigenmodes of \mathcal{D} as a step toward performing a functional

integral, it makes sense to consider the behavior of these eigenmodes under complex conjugation. If we define an antilinear operation \mathcal{T} on spinors by

$$(\mathcal{T}\chi)^\alpha = \varepsilon^{\alpha\beta} \chi_\beta^*, \quad (2.60)$$

then $[\mathcal{T}, \mathcal{D}] = 0$ and $\mathcal{T}^2 = -1$. Just as in the study of Kramers doubling in $3 + 1$ dimensions, this implies that the eigenvalues of \mathcal{D} have even multiplicity. The doubling of the spectrum always occurs, for the same reason, for pseudoreal fermions. One simply has to replace $\varepsilon^{\alpha\beta}$ by the invariant antisymmetric bilinear form relevant to a given case.

Let χ and $\mathcal{T}\chi$ be a pair of eigenmodes of \mathcal{D} with the same eigenvalue λ . Because of the definition $\mathcal{D}_{\gamma\beta} = \varepsilon_{\gamma\alpha} \mathcal{D}_{\beta}^{\alpha}$, in a 2×2 block with basis $\begin{pmatrix} \chi \\ \mathcal{T}\chi \end{pmatrix}$, \mathcal{D} takes the form

$$\mathcal{D} = \begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}. \quad (2.61)$$

More generally, consider a basis of eigenmodes of \mathcal{D} of the form $\chi_i, \mathcal{T}\chi_i$ (where $\chi_i, \mathcal{T}\chi_i$ are orthogonal to $\chi_j, \mathcal{T}\chi_j$ for $i \neq j$). In such a basis, \mathcal{D} is block diagonal with 2×2 blocks of the form (2.61). Accordingly, the Pfaffian of \mathcal{D} is formally

$$\text{Pf}(\mathcal{D}) = \prod_i' \lambda_i, \quad (2.62)$$

where the product is over eigenvalues of \mathcal{D} and the symbol \prod' means that in the product, we take only one eigenvalue from each pair.

Let us compare this to the corresponding equation (2.7) for a Dirac fermion. On the left-hand side of Eq. (2.7), we wrote $\det \mathcal{D}$ instead of $\text{Pf}(\mathcal{D})$, reflecting the fact that $\text{Pf}(\mathcal{D})$ can be written as $\det \mathcal{D}$ in the U(1)-invariant case. On the right-hand side of Eq. (2.7), the product over eigenvalues runs only over fermion eigenvalues of positive charge. (In that derivation, \mathcal{D} was defined as the Hermitian Dirac operator acting on a Dirac fermion ψ , meaning a fermion of definite charge, which we may as well consider to be positive charge.) By charge conjugation symmetry of the Dirac equation, the corresponding negatively charged fermion has exactly the same eigenvalues, but they are not considered in Eq. (2.7). The doubling of Dirac eigenvalues that we got in the Majorana case from a relatively subtle argument using the fact that $\mathcal{T}^2 = -1$ is more trivial in the Dirac case: it just means that the eigenvalues are the same for fermions of positive or negative charge. A formula like Eq. (2.62) can be written for any system of pseudoreal fermions; it reduces to something along the lines of Eq. (2.7) when there is a suitable symmetry.

2. Vanishing of the spectral flow

Now let us consider the sign of \mathcal{D} . As in the discussion of the topological insulator, there is potentially a problem in defining the sign of the formal expression (2.62), since there are infinitely many eigenvector pairs with $\lambda < 0$. On a given three-manifold Y , we can start with a particular metric g_0 , define the sign of $\text{Pf}(\mathcal{D})$ as we wish, and then evolve this

¹⁸U(1) can be replaced by some other group as long as there is a splitting of ψ that leads to a structure like that of Eq. (2.56).

¹⁹The mathematical facts that lead to this Euclidean analog of Kramers doubling for c -number modes in three spacetime dimensions are the same ones that lead to conventional Kramers doubling for quantum states in $3 + 1$ dimensions. A definition of \mathcal{T} that is less intrinsic but may look more familiar is given in Appendix B.2; see Eq. (B13).

sign as the metric is varied by counting how many times an eigenvalue pair flows through $\lambda = 0$. To decide if this procedure gives an answer that is invariant under diffeomorphisms not connected to the identity, one follows the procedure involving the mapping torus that was described in Sec. II.A.2. We consider the same spectral flow problem as in Fig. 1, and ask if the number Δ of eigenvalue pairs flowing through $\lambda = 0$ is even or odd.

It turns out that Δ always vanishes. The steps in proving this are as follows. First, the spectral flow arguments (Atiyah, Patodi, and Singer, 1975; Callan, Dashen, and Gross, 1978; Kiskis, 1978) that we mentioned in explaining Eq. (2.11) are also valid for Majorana fermions. The only differences are as follows: (1) when there is no U(1) symmetry, the spectral flow Δ should be defined as the number of eigenvalue pairs flowing through $\lambda = 0$, rather than the number of positive charge eigenvalues flowing through $\lambda = 0$; (2) similarly, when there is no U(1) symmetry, we have to express Δ in terms of a $D = 4$ Dirac index \mathcal{I} computed in the space of all fermion fields, not an index \mathfrak{S} computed in the space of positively charged fermions only. The relation between the two types of index (when they are both defined) is just

$$\mathcal{I} = 2\mathfrak{S}, \quad (2.63)$$

since charge conjugation implies that the $D = 4$ index for negatively charged fermions is the same as that for positively charged ones. As a result the generalization of Eq. (2.11) that is valid without assuming a U(1) symmetry is

$$\Delta = \frac{\mathcal{I}}{2}. \quad (2.64)$$

This formula makes sense because in $D = 4$, \mathcal{I} is always even,²⁰ a fact that will be important in what follows. Similarly the analog of Eq. (2.12) is that the transformation of the fermion path integral under a diffeomorphism ϕ is

$$\text{Pf}(\mathbf{D}) \rightarrow \text{Pf}(\mathbf{D})(-1)^{\mathcal{I}/2}. \quad (2.65)$$

Here \mathcal{I} is the index of the Dirac operator on the mapping torus built from ϕ , taking all fermions into account.

This formula is not limited to the case of a single Majorana fermion coupled to gravity only. It applies (on an orientable three-manifold; we postpone the unorientable case to Sec. IV) to any system of 3D Majorana fermions transforming in a real representation of any compact gauge group. In many cases [such as the topological insulator, where Eq. (2.65) is equivalent to Eq. (2.12)], the formula implies an anomaly in the sign of $\text{Pf}(\mathbf{D})$. However, there is no such anomaly for a single Majorana fermion coupled to gravity only. Even though

²⁰This assertion is proven by an argument similar to the one used to show the doubling of the Dirac spectrum for $d = 3$. The spinor representation of Spin(4) is pseudoreal, and therefore the eigenmodes of a Majorana fermion in $D = 4$ are doubled because of an antilinear symmetry $(\mathcal{I}\chi)^\alpha = \varepsilon^{\alpha\beta}\chi_\beta^*$, $T^2 = -1$, and $[\mathcal{I}, \mathcal{D}] = 0$. The zero modes of positive or negative chirality are both doubled by this argument, and therefore \mathcal{I} is even; see Appendix B.3.

$\mathcal{I}/2$ can be odd in general on a compact $D = 4$ spin manifold without boundary,²¹ it always vanishes on a mapping torus.²²

3. The sign of the path integral

The last statement means that in the traditional sense there is no anomaly in $\text{Pf}(\mathbf{D})$ for a 3D Majorana fermion coupled to gravity only on an oriented three-manifold Y : after picking its sign in an arbitrary fashion at some starting metric g_0 , the sign can be defined in a consistent way for all metrics in a way invariant under orientation-preserving diffeomorphisms. However, it is not physically satisfactory to allow an arbitrary independent sign in the path integral for every Y . As already remarked in Sec. II.A.2, a satisfactory definition of the theory will determine the phase of the path integral for all Y 's, compatibly with physical requirements of cutting and pasting.

We now explain how, even though there is no anomaly in the traditional sense, a physically natural procedure to define the fermion path integral of the $(2 + 1)$ -dimensional Majorana fermion in a consistent way for all Y 's leads to a violation of \mathbb{T} and \mathbb{R} symmetry, just as happens for the topological insulator. Moreover, in the context of this procedure, it will be clear that even though $(-1)^{\mathcal{I}/2} = 1$ for mapping tori, the fact that $(-1)^{\mathcal{I}/2} \neq 1$ for some other (compact, closed, orientable spin) manifolds represents an obstruction to defining the $(2 + 1)$ -dimensional Majorana fermion by itself in a \mathbb{T} -invariant way, without coupling to a bulk topological superconductor. With coupling to the bulk included, we get a well-defined and \mathbb{T} -invariant answer.

Although I do not claim to have a formal proof, one would expect that any other procedure to precisely define the partition function of the $(2 + 1)$ -dimensional Majorana fermion for all Y will lead to the same conclusions. In general, one expects that two reasonable regularizations of the same theory differ by a local counterterm, but there is no possible local counterterm that would change our conclusions.

To define the $(2 + 1)$ -dimensional Majorana fermion, we use the same Pauli-Villars regularization that we used for the Dirac fermion. The analog for the Pfaffian of Eq. (2.14) for the fermion determinant is

$$\text{Pf}(\mathbf{D})_{\text{reg}} = \prod_k' \frac{\lambda_k}{\lambda_k \pm i\mu}, \quad (2.66)$$

²¹For example, $\mathcal{I}/2 = 1$ on a $K3$ surface.

²²One way to prove this is to note that in four dimensions the Dirac index for a Majorana fermion coupled to gravity only is related to the signature σ by $\mathcal{I} = \sigma/8$. But for a mapping torus of any dimension, a relatively elementary topological argument shows that $\sigma = 0$. Let $X \rightarrow S^1$ be a mapping torus, of dimension $4n$ for some n , with fiber Y of dimension $4n - 1$. Consider the subspace $\Gamma \subset H_{2n}(X; \mathbb{Z})$ generated by cycles of the form $W \times p$, with $W \subset Y$ and p a point in S^1 . Modulo torsion, Γ is a null subspace for the intersection pairing on $H_{2n}(X; \mathbb{Z})$. (For any $W, W' \subset Y$, the intersection number of $W \times p$ and $W' \times p$ is 0 because $W' \times p$ is homologous to $W' \times p'$, for some $p' \neq p \in S^1$, and $W' \times p'$ does not intersect $W \times p$.) Poincaré duality can be used to show that Γ is middle dimensional. The existence of a middle-dimensional null subspace of $H_{2n}(X; \mathbb{Z})$ mod torsion shows that the signature of X is 0.

where we allow both signs of the regulator mass. The same steps that led to Eq. (2.21) now give for the Majorana fermion path integral

$$Z_\psi = |Z_\psi| \exp(\mp i\pi\eta/2), \quad (2.67)$$

where we can now define η as a restricted sum

$$\eta = \lim_{s \rightarrow 0} \sum_k' \text{sign}(\lambda_k) |\lambda_k|^{-s} \quad (2.68)$$

in which one only includes one eigenvalue from each pair. However, it is more customary to define an η invariant as a sum over all eigenvalues

$$\eta = \lim_{s \rightarrow 0} \sum_k \text{sign}(\lambda_k) |\lambda_k|^{-s}, \quad (2.69)$$

so that

$$\eta = 2\eta, \quad (2.70)$$

and then we should write the fermion path integral as

$$Z_\psi = |Z_\psi| \exp(\mp i\pi\eta/4). \quad (2.71)$$

This is \mathbb{T} violating because it is not real, but as in Sec. II.A we can restore \mathbb{T} invariance if we assume that the three-manifold Y that supports the massless Majorana fermion is actually the boundary of the world volume X of a topological superconductor. The salient property of the topological superconductor is that in the large volume limit its partition function on a closed (orientable) four-manifold without boundary is $(-1)^{\mathcal{I}/2}$. (This is always ± 1 , since \mathcal{I} is even.) This claim about the partition function of a topological superconductor can be justified by precisely the same reasoning given in Sec. II.A.6 for the topological insulator, starting with the fact that the phase transition between an ordinary and a topological $D = 4$ superconductor occurs when the mass m of a $D = 4$ Majorana fermion passes through 0. What happens to the sign of the path integral when this mass changes sign? To absorb \mathcal{I} fermion zero modes requires $\mathcal{I}/2$ mass insertions, so the partition function is proportional for small m to $m^{\mathcal{I}/2}$. And if it is positive for one sign of m , then for the opposite sign of m it is proportional to $(-1)^{\mathcal{I}/2}$.

The partition function $(-1)^{\mathcal{I}/2}$ can be interpreted in terms of a gravitational θ angle. In general, such an angle gives a factor $\exp(i\theta\mathcal{I}/2)$ in the path integral measure, so the fact that the topological superconductor partition function is $(-1)^{\mathcal{I}/2}$ means that its gravitational θ angle is π . [See Ryu, Moore, and Ludwig (2010) for a previous explanation of this fact.] On a manifold without boundary, the Atiyah-Singer index theorem gives $(-1)^{\mathcal{I}/2} = \exp[\pm i\hat{A}(R)/2]$, but on a manifold with boundary, with \mathcal{I} defined using APS boundary conditions, the APS index theorem gives a boundary correction in this statement:

$$\exp(\mp i\pi\eta/4) \exp[\pm i\pi\hat{A}(R)/2] = (-1)^{\mathcal{I}/2}. \quad (2.72)$$

This formula is the analog of Eq. (2.26).

The combined path integral of the Majorana fermion on $Y = \partial X$ and the topological superconductor in bulk is then

$$|Z_\psi| \exp(\mp i\pi\eta/4) \exp[\pm i\pi\hat{A}(R)/2] = |Z_\psi| (-1)^{\mathcal{I}/2}. \quad (2.73)$$

This is quite analogous to the topological insulator, with the sole difference $\mathfrak{S} \rightarrow \mathcal{I}/2$. Since $|Z_\psi| (-1)^{\mathcal{I}/2}$ is real, \mathbb{T} invariance has been recovered.

Now let us reexamine the partition function of the boundary Majorana fermions on Y . The combined path integral that describes these fermions and the bulk sTQFT on X is supposed to give $|Z_\psi| (-1)^{\mathcal{I}_X/2}$, where we write \mathcal{I}_X (and not just \mathcal{I}) to emphasize the role of X . But does this formula depend on X ? To decide this question, suppose that X and X' are orientable spin manifolds with the same boundary Y . By gluing them together with a reversal of orientation on one of them, we build an orientable spin manifold \hat{X} without boundary (see Fig. 7). The gluing formula for the index with APS boundary conditions gives²³

$$(-1)^{\mathcal{I}_X/2} (-1)^{\mathcal{I}_{X'}/2} = (-1)^{\mathcal{I}_{\hat{X}}/2}. \quad (2.74)$$

Thus in general $(-1)^{\mathcal{I}_X/2}$ and $(-1)^{\mathcal{I}_{X'}/2}$ are different precisely because $(-1)^{\mathcal{I}_{\hat{X}}/2}$ can be nontrivial for a suitable \hat{X} without boundary. Because $(-1)^{\mathcal{I}_X/2}$ and $(-1)^{\mathcal{I}_{X'}/2}$ are different in general, $|Z_\psi| (-1)^{\mathcal{I}_X/2}$ depends on X . This X dependence is the obstruction to defining the $(2+1)$ -dimensional Majorana fermion by itself in a \mathbb{T} -invariant way.

That is the basis for saying that the nontriviality of $(-1)^{\mathcal{I}_X/2}$ for some X , even though it is trivial for mapping tori, represents an anomaly for the $(2+1)$ -dimensional Majorana fermion.

Viewed from the bulk point of view, the significance of the fact that $(-1)^{\mathcal{I}_X/2}$ can be nontrivial may seem more obvious. If $(-1)^{\mathcal{I}_X/2}$ were equal to 1 for all X without boundary, then the corresponding sTQFT would be completely trivial, and we could define it to be still trivial if X has a boundary. Boundary fermions would play no essential role.

In Sec. IV, we reconsider the 3D topological insulator and superconductor on unorientable manifolds. We will see that this does not change much for the topological insulator, but for the topological superconductor it does. What we have said here about the topological superconductor is only a small part of the full story, but treating this part by itself gave us the chance to explain some important points in a simple context.

C. Overview of the remaining cases

Before getting into any details, we give an overview of the rest of this paper. We will discuss the remaining cases of real

²³The theorem says that $\mathcal{I}_{\hat{X}} = \mathcal{I}_X + \mathcal{I}_{X'}$, which clearly implies Eq. (2.74). It follows from the Atiyah-Singer index formula $\mathcal{I}_{\hat{X}} = \int_{\hat{X}} \hat{A}(R) \text{tr} e^F$, together with the APS index theorem, which says that $\mathcal{I}_X = \int_X \hat{A}(R) \text{tr} e^F - \eta/2$, with a similar formula for $\mathcal{I}_{X'}$. When we take the sum $\mathcal{I}_X + \mathcal{I}_{X'}$, the integrals over X and X' add up to the integral over \hat{X} that gives $\mathcal{I}(\hat{X})$, and the boundary terms $-\eta/2$ cancel because η is odd under reversal of orientation.

and complex fermions in a way very similar to the way that we have analyzed pseudoreal fermions.

In each case, there will be a “bulk” invariant that will play the role of the Dirac index \mathfrak{S} or \mathcal{I} in the cases analyzed so far. For real fermions, this invariant will be $(-1)^\zeta$, where ζ is the mod 2 index of a Dirac operator in dimension D . (This concept will be explained in Sec. III.B.) For complex fermions, the analog will be $\exp(-i\pi\eta/2)$, where η is the APS η invariant²⁴ of a Dirac operator in dimension D . In each case, $(-1)^\zeta$ or $\exp(-i\pi\eta/2)$ will be the partition function of a suitable sTQFT in D dimensions.

However, just like $(-1)^{\mathcal{I}/2}$ [or its special case $(-1)^\mathfrak{S}$], the invariants $(-1)^\zeta$ and $\exp(-i\pi\eta/2)$ are not straightforwardly defined on a manifold X with a spatial boundary. When X has a spatial boundary, something has to happen on $Y = \partial X$. Although there are other possibilities, the standard boundary states of topological insulators and superconductors have gapless fermions supported on the boundary. In each case, the fermion partition function Z_ψ is not well defined by itself. But the product $|Z_\psi|(-1)^\zeta$ or $|Z_\psi|\exp(-i\pi\eta/2)$ is well defined and compatible with all physical principles.

Here $|Z_\psi|(-1)^\zeta$ or $|Z_\psi|\exp(-i\pi\eta/2)$ is the partition function of the combined system consisting of gapless fermions on ∂X and an sTQFT in bulk. These combined partition functions cannot be usefully factored as the product of a partition function for the boundary fermions and one for the bulk sTQFT. In each case, the boundary fermions have an anomaly that compensates for the difficulty in defining the bulk sTQFT on a manifold with boundary.

In a certain sense, the invariant $\exp(-i\pi\eta/2)$ is universal and the more elementary invariants $(-1)^{\mathcal{I}/2}$ and $(-1)^\zeta$ can be viewed as special cases to which $\exp(-i\pi\eta/2)$ reduces under favorable circumstances. It logically would have been possible to begin this paper with a general analysis leading to $\exp(-i\pi\eta/2)$ and then to deduce as corollaries the more easily understood formulas involving $(-1)^{\mathcal{I}/2}$ or $(-1)^\zeta$, but this might have made the presentation rather opaque.

III. REAL FERMIONS

In this section, we consider topological states of matter in which the boundary modes, in Euclidean signature, are real fermions. Our basic example will be the $(2+1)$ -dimensional T-invariant topological superconductor. However, we also consider the $(2+1)$ -dimensional topological insulator.

A. Symmetry of the spectrum

In two spacetime dimensions, we need only two gamma matrices. In Euclidean signature, they should obey $\{\gamma_a, \gamma_b\} = 2\delta_{ab}$, and they can be 2×2 real matrices, for example, $\gamma_1 = \sigma_1, \gamma_2 = \sigma_3$.

The operator $\mathcal{D} = \sum_{i=1}^2 \gamma^i D_i$ is therefore a real antisymmetric operator. The corresponding Hermitian Dirac operator

$\mathcal{D} = i\mathcal{D}$ is imaginary and antisymmetric. This is true even on unorientable two-manifold.

Consider the eigenvalue problem $\mathcal{D}\chi = \lambda\chi$ in this situation. The spectrum is invariant under $\lambda \rightarrow -\lambda$, because if $\mathcal{D}\chi = \lambda\chi$, then $\mathcal{D}\chi^* = -\lambda\chi^*$. Note that we have made use of an antilinear operation $\mathcal{T}:\chi \rightarrow \chi^*$ that obeys $\mathcal{T}^2 = 1$. By contrast, for pseudoreal fermions, to prove that the spectrum is doubled for each value of λ , we used in Sec. II.B an antilinear symmetry with $\mathcal{T}^2 = -1$. In the real case, the symmetry $\lambda \rightarrow -\lambda$ gives a pairing of eigenvalues with $\lambda \neq 0$, but there may be unpaired zero modes. (Indeed, these would represent a mod 2 index, which we come to in Sec. III.B.) For the moment we assume that generically there are no zero modes.

The edge modes of a $(2+1)$ -dimensional T-invariant topological superconductor are a $(1+1)$ -dimensional Majorana fermion (coupled to gravity only). There is no U(1) symmetry and the fermion path integral is best understood as the Pfaffian of an antisymmetric bilinear form \mathbf{D} . The definition of \mathbf{D} in terms of the Hermitian Dirac operator \mathcal{D} is slightly different from what it is for pseudoreal fermions, which were discussed in Sec. II.B. Instead of $\mathbf{D}_{\gamma\beta} = \varepsilon_{\gamma\alpha}\mathcal{D}_\beta^\alpha$ as in that case, with $\varepsilon_{\alpha\beta}$ an invariant antisymmetric tensor in the space of fermion fields, we now have $\mathbf{D}_{\gamma\beta} = \delta_{\gamma\alpha}\mathcal{D}_\beta^\alpha$, with $\delta_{\gamma\alpha}$ an invariant symmetric tensor.

The result of this is that \mathbf{D} is the direct sum of 2×2 blocks, with one such block for each eigenvalue pair $\lambda_i, -\lambda_i$ of \mathcal{D} :

$$\mathbf{D} = \bigoplus_i \begin{pmatrix} 0 & \lambda_i \\ -\lambda_i & 0 \end{pmatrix}. \quad (3.1)$$

This contrasts with the pseudoreal case, in which \mathbf{D} has a 2×2 block for each pair of states with the same eigenvalue λ .

An important detail is that if we are given an antisymmetric bilinear form \mathbf{D} , then its “skew eigenvalues” λ_i are uniquely determined only up to sign. An orthogonal transformation by the 2×2 matrix $\text{diag}(1, -1)$ acting in the i th block would reverse the sign of λ_i in Eq. (3.1). So *up to sign* (and modulo the need for regularization), the fermion path integral is

$$\text{Pf}(\mathbf{D}) = \prod_i \lambda_i. \quad (3.2)$$

Broadly speaking, we are in a familiar situation. To decide on the sign of $\text{Pf}(\mathbf{D})$, we need to know, mod 2, how many of the λ_i should be considered negative. In general, such a question does not have a natural answer. However, we observe that as a function of the metric tensor g of a two-manifold Y $\text{Pf}(\mathbf{D})$ should change sign whenever a pair of eigenvalues $\lambda_i, -\lambda_i$ passes through 0. So to get a partial answer, we can pick the sign of $\text{Pf}(\mathbf{D})$ for some chosen metric g_0 on a two-manifold Y , and then determine the sign as a function of g by counting how many times an eigenvalue pair passes through 0 in interpolating from g_0 to g .

To understand the behavior of $\text{Pf}(\mathbf{D})$ under a diffeomorphism ϕ , we have to count mod 2 how many times the sign of $\text{Pf}(\mathbf{D})$ changes when we interpolate, via a family of metrics g_s parametrized by $s \in [0, 1]$, from a starting metric g_0 to its conjugate g_0^ϕ under ϕ (Fig. 9). The difference from the

²⁴Typically $\exp(-i\pi\eta/2)$ is not a topological invariant, but if the original complex fermions have no perturbative anomaly, then the particular invariant $\exp(-i\pi\eta/2)$ that controls their global anomaly is a topological invariant and in fact a cobordism invariant.

mode is a ψ_- zero mode, so the number of zero modes of ψ is trivially even and the mod 2 index inevitably vanishes.

This is not so if we formulate the Majorana fermion on an unorientable two-manifold. For a concrete example, we consider the Klein bottle (KB) constructed as the quotient of the two-torus (3.6) by $z \rightarrow \bar{z} + 1/2$. Equivalently, introducing real coordinates by setting $z = x_1 + ix_2$, the relations are

$$(x_1, x_2) \cong (x_1 + 1, x_2) \cong (x_1, x_2 + 1) \cong (x_1 + 1/2, -x_2). \quad (3.7)$$

We define a pin^+ structure²⁵ on KB by requiring

$$\begin{aligned} \psi(x_1, x_2) &= \psi(x_1 + 1, x_2) = \psi(x_1, x_2 + 1) \\ &= \gamma_2 \psi(x_1 + 1/2, -x_2). \end{aligned} \quad (3.8)$$

One of the two zero modes that ψ would have on T^2 is projected out when we go to KB, so with this pin^+ structure, the mod 2 index of ψ is 1.

For our last example, we consider a two-component Majorana fermion ψ in $D = 3$ spacetime dimensions. There is a Dirac action for such a system, so it is possible to define a mod 2 index. This mod 2 index always vanishes on an orientable three-manifold X . Indeed, on an orientable three-manifold, the eigenvalues of the Dirac operator always have even multiplicity, as shown earlier using the antilinear operation defined in Eq. (2.60). However, on an unorientable three-manifold, the mod 2 index can be nontrivial. A simple example is $X = \text{KB} \times S^1$, where KB and the spin structure on it were defined in Eqs. (3.7) and (3.8), and we take the periodic spin structure on S^1 . The only zero mode of ψ is a mode that is constant in the S^1 direction and whose restriction to KB is the zero mode found in the previous paragraph. So the mod 2 index is 1.

We conclude this introduction to the mod 2 index by stressing that despite its name the mod 2 index is not the mod 2 reduction of an ordinary index or indeed of any integer-valued invariant. In fact, we have given examples in one and three dimensions of a nontrivial mod 2 index, but the ordinary index always vanishes in odd dimensions. The mod 2 version of the Atiyah-Singer index theorem (Atiyah and Singer, 1971) does not give a formula for the mod 2 index,²⁶ analogous to the familiar Atiyah-Singer index formula for the ordinary index. It does imply that the mod 2 index is a cobordism invariant, so, in particular, there is a 3D sTQFT (defined on spin or more precisely pin^+ manifolds) with partition function $(-1)^\xi$. Before discussing what happens when this theory is considered on a manifold with boundary, we present some simple examples of the global anomaly of a 2D Majorana fermion.

C. Simple examples of the global anomaly

We learned in Sec. III.B that in $D = 3$ the mod 2 index of the Dirac operator is always zero on an orientable manifold, but

²⁵This concept, which is the analog of a spin structure on an unorientable manifold, is described in Appendix A.2.

²⁶To be more precise, the formula expresses the mod 2 index in terms of K theory but does not give a formula in a conventional sense as an integral.

that on an unorientable manifold the mod 2 index of a Majorana fermion coupled to gravity only can be nonzero. This should mean that in two spacetime dimensions a Majorana fermion ψ is consistent as long as we formulate it only on an orientable two-manifold and only allow orientation-preserving diffeomorphisms, but becomes anomalous otherwise.

The anomaly if we formulate the Majorana fermion ψ on an unorientable two-manifold or if we allow orientation-reversing symmetries can be seen by elementary examples that do not really require the general formalism. First, let us consider the ψ field on the Klein bottle KB, which we take as an example of an unorientable two-manifold. As we have seen, with a suitable pin structure on KB, the mod 2 index is nonzero. This means that generically the Dirac operator has a single c -number zero mode χ_0 . Other modes are paired in a natural way under the symmetry $\lambda \leftrightarrow -\lambda$ of the spectrum. In such a situation, the measure in the fermion path integral is odd under the symmetry $(-1)^F$ that counts fermions mod 2. Concretely, the fermion partition function on KB vanishes because of the zero mode, but the (unnormalized) one-point function $\langle \int d^2x \sqrt{g} \chi_0 \psi \rangle$ with an operator insertion that removes the zero mode is nonzero. Since this correlation function has an odd number of fermion insertions, it is odd under $(-1)^F$.

Clearly, we can understand this example without thinking about the mapping torus, but how would we understand it in terms of the mapping torus construction? Any diffeomorphism of a manifold Y that can act on fermions at all can act in two different ways, differing by an overall sign. In particular, the trivial diffeomorphism of Y can be taken to act trivially on fermions or to act as multiplication by -1 . If we lift the trivial diffeomorphism of Y to act trivially on fermions, we get the trivial symmetry 1, but if we take it to act as -1 on fermions, we get the symmetry $(-1)^F$. So we think of the trivial symmetry 1 or the nontrivial symmetry $(-1)^F$ as special cases of a diffeomorphism and apply the mapping torus construction.

The mapping torus associated with the trivial diffeomorphism of KB is just the product $\text{KB} \times S^1$. The two lifts of the trivial diffeomorphism to act on fermions correspond to two different pin^+ structures in which the fermions are periodic or antiperiodic around the S^1 direction. (The pin^+ structure on KB is arbitrary and is kept fixed in this discussion.) The trivial symmetry 1 and the symmetry $(-1)^F$ correspond to the antiperiodic and periodic pin^+ structures, respectively. This is analogous to the fact that in statistical mechanics, with Hamiltonian H , to compute a trace $\text{Tr} e^{-\beta H}$ or $\text{Tr} (-1)^F e^{-\beta H}$, we do a path integral on a circle of circumference β with fermions that are, respectively, antiperiodic or periodic around the circle. Now, let us compute the mod 2 index on $\text{KB} \times S^1$ with our two pin^+ structures. If the fermions are antiperiodic on S^1 , corresponding to the trivial symmetry 1, then the lowest eigenvalue of the Dirac operator on $\text{KB} \times S^1$ is π/β (or larger), where again β is the circumference of the circle. There are no zero modes, so the mod 2 index of the mapping cylinder vanishes, implying that the path integral measure on KB is invariant under the trivial symmetry. But if the fermions are periodic on S^1 , corresponding to the symmetry $(-1)^F$, then as seen in Sec. III.B, the mod 2 index on $\text{KB} \times S^1$ is the same as it is on KB. This means that precisely when the mod 2 index on KB is nonzero, the path integral measure on KB is odd under $(-1)^F$.

In these statements, KB could be replaced with any two-manifold Y . Whenever there is a nontrivial mod 2 index on Y , and thus an anomaly in $(-1)^F$, this anomaly shows up in the mod 2 index on the mapping torus $Y \times S^1$.

For another simple example, let us consider the Majorana fermion ψ on the orientable two-manifold T^2 . But now we consider a diffeomorphism ϕ that acts as a reflection of one coordinate, reversing the orientation of T^2 . Taking on T^2 the spin structure that is periodic in both directions, ψ has two zero modes, say χ_1 and χ_2 . Because of these zero modes, the partition function of the ψ field on T^2 vanishes, but the (unnormalized) two-point function

$$\left\langle \int_Y d^2x \sqrt{g} \bar{\chi}_1 \psi \cdot \int_Y d^2x \sqrt{g} \bar{\chi}_2 \psi \right\rangle \quad (3.9)$$

is nonzero. Of the two zero modes, one linear combination is even under ϕ and one is odd, so the matrix element in Eq. (3.9) is odd under ϕ , showing that the path integral measure is odd under ϕ .

Let us again try to understand this global anomaly in the mapping torus construction. We take for T^2 the quotient of the $x_2 - x_3$ plane by $(x_2, x_3) \cong (x_2 + 1, x_3) \cong (x_2, x_3 + 1)$, with the Majorana fermion field ψ assumed to be periodic in both directions. We consider ϕ to act by $(x_2, x_3) \rightarrow (-x_2, x_3)$. To construct the mapping torus associated with ϕ , we introduce a copy of \mathbb{R} parametrized by another variable x_1 and consider the quotient of $\mathbb{R} \times T^2$ by $(x_1, x_2, x_3) \rightarrow (x_1 + 1/2, -x_2, x_3)$. But this is just another way to construct the familiar three-manifold $\text{KB} \times S^1$. We must further say how we want ϕ to act on fermions. We do this by saying that a fermion field ψ on $\text{KB} \times S^1$ is a fermion field on $\mathbb{R} \times T^2$, periodic in the T^2 directions, that obeys $\psi(x_1 + 1/2, -x_2, x_3) = \gamma_2 \psi(x_1, x_2, x_3)$. But this gives the pin^+ structure on $\text{KB} \times S^1$ that was already considered in Sec. III.B. As explained there, the mod 2 index with this pin^+ structure is nonzero, so again the mapping torus construction detects the global anomaly in the reflection symmetry of T^2 .

D. Bulk sTQFT associated with the T-invariant topological superconductor

Now, let us consider the sTQFT whose partition function on a closed three-manifold X with a pin^+ structure is $(-1)^\zeta$. This theory is trivial if X is orientable, but in general not otherwise.

Our first task is to explain why this sTQFT is the bulk state of a T-conserving topological superconductor in $D = 3$ spacetime dimensions. The style of the argument should be familiar from Sec. II.A.6.

We start in Lorentz signature with two 3D fermions ψ_1 and ψ_2 that transform oppositely²⁷ under T:

²⁷In Eq. (2.4), we wrote a massive Dirac equation for a pair of 3D fermions ψ_1 and ψ_2 assumed to transform the same way under T. These, however, were complex Dirac fermions, and the factor of i multiplying the mass term in Eq. (2.4) was necessary for T invariance. For Majorana fermions, the Dirac equation must be real, so this factor of i is not possible; a T-invariant mass term linking two fermions ψ_1 and ψ_2 is possible only if they transform with opposite signs under T.

$$\begin{aligned} \mathbb{T}\psi_1(t, x_1, x_2) &= \gamma_0 \psi_1(-t, x_1, x_2), \\ \mathbb{T}\psi_2(t, x_1, x_2) &= -\gamma_0 \psi_2(-t, x_1, x_2). \end{aligned} \quad (3.10)$$

The opposite signs are chosen so that it is possible to have a T-conserving off-diagonal mass term $m\bar{\psi}_1\psi_2$. (Diagonal mass terms $m'\bar{\psi}_1\psi_1$ or $m''\bar{\psi}_2\psi_2$ are T violating.) The phase transition between a topologically trivial and a topologically nontrivial insulator is achieved when m changes sign. Now, we study as a function of m the partition function of this system on a compact manifold X of Euclidean signature. Suppose that X and the pin^+ structure on X are such that $\zeta \neq 0$. Then setting $m = 0$, both ψ_1 and ψ_2 have an odd number of zero modes (generically 1 each). It takes an odd number of insertions of $m\bar{\psi}_1\psi_2$ to lift these zero modes, and therefore the partition function is proportional to an odd power of m ; thus, it changes sign in passing through $m = 0$ whenever $\zeta \neq 0$. There is no such sign change if $\zeta = 0$. So if the partition function is positive definite for, say, $m > 0$ then its sign is $(-1)^\zeta$ for $m < 0$.

A more precise explanation is as follows. Since ψ_1 and ψ_2 transform oppositely under an orientation-reversing symmetry, they combine together to a single Majorana fermion $\hat{\psi}$ on an oriented three-manifold \hat{X} that arises as the oriented double cover of X . The partition function of the $\psi_1 - \psi_2$ system on X is the same as the $\hat{\psi}$ partition function on \hat{X} . To be precise about this partition function, we add (rather as in Sec. II.A.3) Pauli-Villars regulator fields χ_1 and χ_2 with opposite statistics to ψ_1 and ψ_2 and a T-conserving mass term $\mu\bar{\chi}_1\chi_2$. On \hat{X} , these combine to a single regulator field $\hat{\chi}$. \hat{X} has an orientation-reversing symmetry (the quotient of \hat{X} by this symmetry is the original X). The Dirac operator on \hat{X} anticommutes with such a symmetry [see the discussion of Eq. (2.22)], and hence the nonzero eigenvalues of the Dirac operator on \hat{X} are paired under $\lambda \rightarrow -\lambda$. [There is also a doubling of the spectrum on \hat{X} for each value of λ , because of the antilinear symmetry (2.60), but as in Eq. (2.62) a Majorana fermion partition function is computed by taking only one eigenvalue from each such pair.] With $\hat{\psi}$ having mass m and $\hat{\chi}$ having mass μ , the contribution of a pair $\lambda, -\lambda$ to the fermion path integral is

$$\frac{\lambda + im}{\lambda + i\mu} \frac{-\lambda + im}{-\lambda + i\mu}. \quad (3.11)$$

This is always positive, regardless of the signs of μ and m , and so the contribution of nonzero modes to the regularized path integral is always positive. Now, let us suppose that the massless Dirac operator on X has ι zero modes (with $\iota \cong \zeta \pmod{2}$). Then the field $\hat{\psi}$ on \hat{X} has 2ι zero modes. Since a mass insertion can lift two zero modes of $\hat{\psi}$ or $\hat{\chi}$, every pair of such zero modes gives a factor of m in the $\hat{\psi}$ path integral and a factor of $1/\mu$ in the regulator path integral. As a result the path integral is proportional to $(m/\mu)^\iota$; it is positive definite if m and μ have the same sign, and its sign is $(-1)^\zeta$ otherwise.

The case that m and μ have the same sign corresponds to a topologically trivial T-invariant superconductor. The

topologically nontrivial case is the case that m and μ have opposite signs. This second case is the topologically interesting one and $(-1)^\zeta$ is the partition function of the bulk sTQFT associated with it.

E. Extending the sTQFT to a manifold with boundary

We now extend this sTQFT to the case that X has a spatial boundary Y . For this, we need to extend the definition of the mod 2 index ζ to the case that X has such a boundary. However, there is no symmetry-preserving local boundary condition that will let us define ζ .

Let us look carefully at this crucial point. If X is orientable, then its boundary Y is also orientable, and it is possible to define a local boundary condition on the Dirac equation along the lines of Eq. (2.37). Let \vec{n} be the unit normal vector to Y . Then we can impose the boundary condition

$$\gamma \cdot n\psi|_Y = \pm\psi|_Y, \quad (3.12)$$

with some choice of sign, and these are the most general possible covariant boundary conditions.²⁸ With either choice of sign, this is a physically sensible classical boundary condition,²⁹ and with this boundary condition, we can define the mod 2 index of the Dirac operator. However, this boundary condition does not make sense globally if Y is unorientable. A quick way to see this is as follows. Starting with a two-component fermion field ψ , the boundary condition (3.12) selects one component that can be nonzero along Y , and it is not difficult to see³⁰ that this component has a definite chirality in the two-dimensional sense. But the chirality cannot be defined globally along Y if Y is unorientable, so the boundary condition (3.12) does not make sense globally.

If Y is orientable and we pick an orientation, then the boundary condition (3.12) with some choice of sign makes sense. Defining ζ to be the mod 2 index of the Dirac operator on X with some choice of sign, we can define a 3D sTQFT on a three-manifold with boundary whose partition function is $(-1)^\zeta$. This gives a gapped and topologically unordered but *symmetry-breaking* boundary condition for the 3D sTQFT. It is symmetry breaking because the choice

²⁸A local boundary condition must set to zero one-half of the fermion components along $Y = \partial X$, and Eq. (3.12) is the most general way to do this that is invariant under rotation of the normal plane at a point $p \in Y$. Any other matrix that we might use instead of $\gamma \cdot n$ would violate the local rotation symmetry.

²⁹Quantum mechanically, a $D = 3$ fermion with such a boundary condition will have a perturbative gravitational anomaly along $\partial X = Y$ (Horava and Witten, 1996). That is not important here because we are interested in the mod 2 index of the $D = 3$ fermion, not its partition function.

³⁰In a local Euclidean frame, we can take the three gamma matrices to be $\gamma_a = \sigma_a$, $a = 1, \dots, 3$, obeying $\gamma_1\gamma_2\gamma_3 = i$. Let us assume that the normal direction to the boundary is the 3 direction. The 2D chirality along the boundary is measured by $i\gamma_1\gamma_2 = -\gamma_3$, so the condition (3.12) on the eigenvalue of $\hat{n} = \gamma_3$ selects one 2D chirality.

of sign in Eq. (3.12), which involves a choice of one 2D chirality, breaks reflection or time-reversal symmetry along the boundary.

We can define ζ in a T-invariant way if we use the global APS boundary conditions. Concretely this amounts to replacing X by the noncompact manifold X' of Fig. 6, and defining ζ to be the dimension mod 2 of the space of square-integrable zero modes of the Dirac operator. This gives a T-invariant definition of ζ , but ζ defined this way is not a topological invariant. When the metric on $Y = \partial X$ is varied so that a pair of eigenvalues of the 2D Dirac operator on Y passes through 0, the APS boundary conditions on X jump and ζ also jumps.

At this point $(-1)^\zeta$ jumps in sign, so a bulk theory with partition function $(-1)^\zeta$ is not physically sensible by itself in this situation. However, hopefully it is clear at this point what we should say. The bulk theory with partition function $(-1)^\zeta$ has to be combined with a boundary system consisting of Majorana fermions on the boundary. The partition function of the combined system is

$$|Z_\psi|(-1)^\zeta. \quad (3.13)$$

Here $(-1)^\zeta$ jumps whenever a pair of eigenvalues $\pm\lambda$ of the boundary fermions passes through $\lambda = 0$, that is, whenever the path integral of the boundary fermions would be expected to change sign. So the product is physically sensible.

F. The orientable case and chiral symmetry

Let us now look more closely at the case that Y is orientable. In this case, subject to one restriction that we mention in a moment, we can assume that Y is the boundary of an orientable spin manifold X . This manifold has $\zeta = 0$. Indeed, the argument given in Sec. III.B showing that $\zeta = 0$ if X is orientable remains valid with $\partial X \neq 0$, as long as we use APS boundary conditions.

Since $(-1)^\zeta$ is always 1 in this situation, we conclude that if Y is orientable, we can sensibly define the Majorana fermion partition function on Y to be positive definite. A more direct explanation of this is to observe that if Y is orientable, then the Majorana fermion ψ on Y can be decomposed as a sum of fields ψ_+ and ψ_- of positive and negative chirality; moreover, the action and path integrals of ψ_+ and ψ_- are complex conjugates. So the overall path integral is positive.

As long as Y and X are orientable, the formula (3.13) for the partition function does not depend on X , since $\zeta = 0$ for any X , so it seems that we can forget about X and we get a definition of the massless Majorana fermion theory on Y as a purely two-dimensional theory. However, there is a detail that requires some care. Until this point, we have only assumed that Y is orientable, but we have not actually had to pick an orientation. Momentarily, we will see that to give a complete definition of the Majorana fermion theory on Y requires picking an orientation of Y , and therefore allowing only orientation-preserving diffeomorphisms as symmetries. This should not come as a surprise, because to analyze an

orientation-reversing symmetry of Y the mapping torus construction will tell us to think about the mod 2 index on an unorientable three-manifold X , and this can be nontrivial, as discussed in Sec. III.B.

The reason that some care is needed is that, for given Y , with a given spin structure, a suitable X may not exist. The obstruction to existence of X arises precisely if Y is such that the mod 2 index ζ_Y for a positive chirality fermion ψ_+ on Y is not zero. (By complex conjugation, ζ_Y is also the mod 2 index of a negative chirality fermion ψ_- on Y .) This mod 2 index is a cobordism invariant, and if it is nonzero on Y , then Y with its given spin structure is not the boundary of an orientable spin manifold X .

Suppose that this index is nonzero on Y . Then ψ_+ and ψ_- each have an odd number of zero modes; for simplicity, let us assume that this number is 1, the generic value, and write $\psi_{+,0}$ and $\psi_{-,0}$ for the coefficients of the zero modes in a mode expansion of ψ . The path integral over nonzero fermion modes on Y is naturally positive, as the ψ_- path integral is the complex conjugate of the ψ_+ path integral. Since $\psi_{+,0}$ and $\psi_{-,0}$ are coefficients of zero modes, their path integral vanishes in the absence of operator insertions, so the question we should ask is how to define the path integral measure. But there is no natural way to define the sign of this measure, unless an orientation of Y is chosen: whether we write $d\psi_{+,0}d\psi_{-,0}$ or $d\psi_{-,0}d\psi_{+,0}$ is tantamount to a choice of orientation of Y . We actually considered this situation in Sec. III.C for the case $Y = T^2$ and interpreted the fact that the path integral measure on Y is odd under an orientation-reversing symmetry in terms of the mapping torus construction.

We conclude this discussion with the following remark. Although we have taken the 2D Majorana fermion as our basic example of real fermions, the special behavior for orientable Y that we have just found reflects the fact that this example is somewhat special. The classification of representations of a compact group K as being real, pseudoreal, or complex properly refers to irreducible representations, which are real (they admit an invariant symmetric bilinear form), pseudoreal (they admit an invariant antisymmetric form), or complex (they admit no invariant bilinear form). Reducible representations do not have this classification and a reducible representation may admit both a symmetric bilinear form and an antisymmetric one. That is the case with the $(1+1)$ -dimensional Majorana fermion on an orientable spacetime Y . The spinor representation of $\text{Spin}(2)$ is not irreducible but decomposes as the sum of components of positive or negative chirality. If b is a bosonic spinor of $\text{Spin}(2)$ with chiral components b_+ and b_- , then there exist both an invariant symmetric form $db_+ \otimes db_-$ and an invariant antisymmetric form $db_+ \wedge db_-$. By contrast, if Y is unorientable, then we have to consider the Majorana fermion as a representation of $\text{Pin}^+(2)$. Only the symmetric form and not the antisymmetric one is invariant under the reflections in $\text{Pin}^+(2)$ (which exchange b_+ and b_-), and as a representation of $\text{Pin}^+(2)$, the Majorana fermion is real and not pseudoreal. So that case is a better example of the general paradigm of real fermions.

There are many interesting examples of real fermions in which this issue does not arise (the Euclidean space fermions have a real structure and no pseudoreal one). One such case is the original example in which the mod 2 index was interpreted as a global anomaly for real fermions (Witten, 1982), namely,³¹ an $\text{SU}(2)$ gauge theory in four dimensions with a single multiplet of fermions transforming in the spin $1/2$ representation of $\text{SU}(2)$. A more elementary example, which we discuss in Sec. V, is given by a one-component fermion on the boundary of a $D = 2$ Majorana chain (Fidkowski and Kitaev, 2009).

G. $(2+1)$ -dimensional topological insulator

We now turn to the topological insulator in $2+1$ dimensions.³² This is a system with T symmetry and also with $\text{U}(1)$ symmetry. Crucially, T commutes with the electric charge operator Q , the generator of $\text{U}(1)$. (At the end of this section, we consider the contrary case of a conserved charge that is odd under T .)

The boundary modes of the $(2+1)$ -dimensional topological insulator are a $(1+1)$ -dimensional massless charged Dirac fermion ψ , say of charge 1. Its charge conjugate, of course, has charge -1 . The Dirac fermion ψ can be expanded in terms of two Majorana fermions ψ_1 and ψ_2 as $\psi = (\psi_1 + i\psi_2)/\sqrt{2}$. The $\text{U}(1)$ symmetry of electric charge rotations is

$$\delta\psi_1 = \psi_2, \quad \delta\psi_2 = -\psi_1. \quad (3.14)$$

Since this formula is real and the $(1+1)$ -dimensional Euclidean gamma matrices can also be chosen to be real, we conclude that the boundary fermions of the $(2+1)$ -dimensional topological insulator are real in Euclidean signature. (A remark made in Sec. III.F also applies here, for essentially the same reason: on an orientable two-manifold Y , these fermions can be given a pseudoreal structure as well as a real one.)

Given that the fermions are real, more or less everything stated about the $(2+1)$ -dimensional superconductor has an analog for the topological insulator. On a closed three-manifold X , the sTQFT associated with the topological insulator has a partition function $(-1)^\zeta$, where ζ is the mod 2 index that must be computed in the space of all fermion components [meaning that in computing ζ , we have to include both ψ_1 and ψ_2 or equivalently both charge eigenstates $(\psi_1 \pm i\psi_2)/\sqrt{2}$]. If X has a boundary Y , then $(-1)^\zeta$ cannot be defined as a topological invariant, without breaking the time reversal or reflection symmetry. We can preserve the symmetry in defining $(-1)^\zeta$ if we use APS boundary conditions, but the result is not a topological invariant. It jumps in a way that only makes physical sense if we also include the boundary fermions, and then the partition function of the combined system is as usual $|Z_\psi|(-1)^\zeta$.

³¹The model was pointed out to the author by S. Coleman in Aspen in the summer of 1976. It is difficult to recall what his view was at the time.

³²Some of the following issues were previously treated by Hsieh *et al.* (2014).

But is the invariant $(-1)^\zeta$ nontrivial in this situation? Here we run into an essential novelty. Because the T transformation that is a symmetry of a topological insulator commutes with the electric charge generator Q , it really does correspond in relativistic terms to T and not to CT . In a relativistic theory, there is a CRT theorem, so T symmetry is equivalent to CR symmetry. In Euclidean signature, it is more natural to work with CR (which upon analytic continuation back to Lorentz signature implies both T and CR).

In detail, CR is a symmetry that reverses the orientation of space and also *anticommutes* with the generator Q of electric charge. In terms of gauge theory, the last statement means that the electromagnetic gauge potential A does not transform as a one-form under a spatial reflection, but transforms with an extra minus sign. For example, the reflection $R: (x_1, x_2, x_3) \rightarrow (-x_1, x_2, x_3)$ combined with charge conjugation $C: A \rightarrow -A$ gives the combined CR transformation

$$\begin{aligned} CRA_1(x_1, x_2, x_3) &= A_1(-x_1, x_2, x_3), \\ CRA_i(x_1, x_2, x_3) &= -A_i(-x_1, x_2, x_3), \quad i = 2, 3. \end{aligned} \quad (3.15)$$

Once one understands that an orientation-reversing symmetry should be combined with charge conjugation, it is straightforward to see that the mod 2 index ζ can be nonzero. We start with the two-torus T^2 parametrized by $0 \leq x_1, x_2 \leq 1$ with end points identified. On T^2 , we place a gauge field A with one unit of magnetic flux $\int_{T^2} dx_1 dx_2 F_{12} = 2\pi/e$. Pick a basis ψ_\pm^\pm for the components of a Dirac fermion on T^2 , where the superscript is the $U(1)$ charge ± 1 and the subscript is the 2D chirality. Using standard facts about the 2D Dirac equation in a magnetic field, one finds that ψ_+^+ and ψ_-^- have one zero mode each and ψ_+^- and ψ_-^+ have none. Now consider a reflection $R: (x_1, x_2) \rightarrow (-x_1, x_2)$. One can choose A to be invariant under the combined operation CR (this depends on the fact that $\int_{T^2} F$ is odd under both C and R and so is CR invariant). The effect of CR is to exchange ψ_+^+ with ψ_-^- , so precisely one linear combination of the two zero modes on T^2 is CR invariant. If it were the case that R acted freely on T^2 , we would now consider the quotient of T^2 by the \mathbb{Z}_2 group generated by R to get a 2D example with a nonzero mod 2 index. However, this is not the case. Instead we take a product $T^3 = T^2 \times S^1$, where S^1 is a circle parametrized by $0 \leq x_3 \leq 1$ again with end points identified. We take the gauge field A to be a pullback from T^2 , and we define the orientation-reversing symmetry $R': (x_1, x_2, x_3) \rightarrow (-x_1, x_2, x_3 + 1/2)$. Now R' does act freely on T^3 , and the quotient T^3/\mathbb{Z}_2 , where \mathbb{Z}_2 is generated by R' , is a 3D example with a nontrivial mod 2 index.

It is probably fairly clear at this point that whatever we have said about the $D = 3$ topological superconductor has an analog for the topological insulator.

To conclude, we consider a system with a conserved charge Q that is *odd* rather than even under T . In condensed matter physics, Q could be a component of the electron spin, which for some systems is approximately conserved. A time-reversal symmetry that reverses the only

pertinent conserved charge would be called CT in relativistic language. Then in Euclidean signature, one would have to consider an R symmetry, rather than CR , meaning that the gauge field A transforms under reversal of orientation as an ordinary one-form. In this case, one can show³³ that on any three-manifold X , and for any gauge field, ζ is universally 0. So there is no nontrivial sTQFT with partition function $(-1)^\zeta$. This is consistent with the fact that in $D = 3$ there is according to standard arguments no analog of a topological insulator for a T -invariant system with a $U(1)$ symmetry whose generator Q is assumed to be odd under T . In particular, a Dirac fermion with a conserved charge that is odd under T can be given a symmetry-preserving bare mass.

IV. COMPLEX FERMIONS

A. Overview

Our basic example with complex fermions will be the topological insulator or superconductor in 3 space or 4 spacetime dimensions. We already treated this problem in Sec. II, but there we considered only orientable spacetimes. Here we discuss what happens if the spacetime is permitted to be unorientable. In this case,³⁴ the boundary fermions transform in a complex representation of $\text{Pin}^+(3)$ rather than a pseudoreal representation of $\text{Spin}(3) \cong \text{SU}(2)$. Their partition function, even at a formal level, is now complex valued, not real valued, and we have to consider anomalies in a complex-valued path integral of complex fermions.

We start with the topological superconductor, which proves to be the more interesting case. The discussion proceeds in a more or less familiar way. First, we consider the anomaly of the boundary modes, which are a $d = 3$ Majorana fermion. As usual, this anomaly can be expressed

³³In the case of R rather than CR symmetry, A is a connection on a $U(1)$ bundle $\mathcal{L} \rightarrow X$, where X is a possibly unorientable three-manifold with pin^+ structure. [CR symmetry would mean that the structure group is $O(2)$ rather than $U(1)$.] The first Chern class $c_1(\mathcal{L})$ is Poincaré dual to an embedded circle $U \subset X$. Let \hat{U} be a tubular neighborhood of U in X . Topologically $\hat{U} = U \times D^2$ where D^2 is a two-dimensional disk. Since we are trying to calculate a topological invariant, we can assume that in a suitable gauge, $A = 0$ on the complement of \hat{U} . One can construct an elementary cobordism from X to the disjoint union $X_1 \cup X_2$, where X_1 is obtained from X by cutting out \hat{U} and then gluing back in a copy of \hat{U} with $A = 0$ (so that $A = 0$ everywhere on X_1), and X_2 is obtained by embedding $\hat{U} = U \times D^2$ in $U \times S^2$ (using the obvious embedding $D \subset S^2$, and taking $A = 0$ outside \hat{U}). Let ζ_X , ζ_{X_1} , and ζ_{X_2} be the mod 2 indices on the indicated three-manifolds. By cobordism invariance of the mod 2 index, $\zeta_X = \zeta_{X_1} + \zeta_{X_2}$. On X_1 , $A = 0$ so a charged Dirac fermion on X_1 is equivalent to two Majorana fermions on X_1 coupled to gravity only; they contribute equally to ζ_{X_1} , which therefore trivially vanishes. X_2 is orientable, and ζ_{X_2} vanishes because of the usual antilinear symmetry; thus $\zeta_X = 0$.

³⁴We consider systems with $T^2 = (-1)^F$ so that the appropriate group is $\text{Pin}^+(3)$ and not $\text{Pin}^-(3)$; see Appendix A.2.

as a topological invariant of the mapping torus. For complex fermions in general (Witten, 1985b), the relevant invariant of the mapping torus is $\exp(-i\pi\eta/2)$, where η is an APS η invariant of a Dirac operator in one dimension more—in our example the Dirac operator of a $D = 4$ Majorana fermion. Then, we consider the partition function of the sTQFT associated with the bulk topological superconductor. Using a standard characterization of the phase transition to topological superconductivity in $D = 4$, we show that the partition function of the bulk sTQFT is again $\exp(-i\pi\eta/2)$. Like other bulk invariants we have encountered as sTQFT partition functions, this one cannot be defined with a local, symmetry-preserving boundary condition. We can define it in a symmetry-preserving way using APS boundary conditions, but then it is not a topological invariant and again is not satisfactory by itself. But the product $|Z_\psi| \exp(-i\pi\eta/2)$ is physically sensible as shown by the Dai-Freed theorem (Dai and Freed, 1994). This product should be interpreted as the combined partition function of the boundary fermions and the bulk sTQFT.

Finally, we consider a question that does not have a close analog in cases studied earlier in this paper. This involves a $D = 4$ system whose boundary supports ν identical gapless Majorana fermions. At the free fermion level, ν is an integer-valued invariant, but it has been argued (Fidkowski, Chen, and Vishwanath, 2013; Metlitski *et al.*, 2014; Wang and Senthil, 2014; Kitaev, 2015) that at the interacting level ν is only an invariant mod 16. From the present point of view, this is true because, in $D = 4$, $\exp(-i\pi\eta/2)$ is in general a 16th root of 1, and there is no satisfactory, symmetry-preserving definition of $\exp(-\nu \cdot i\pi\eta/2)$ on a four-manifold with spatial boundary, unless ν is a multiple of 16.

A somewhat similar explanation of the value $\nu = 16$ was proposed by Kapustin *et al.* (2014) based on a suggestion that the relevant sTQFT partition function would be a cobordism invariant. An explicit global anomaly computation (Hsieh, Cho, and Ryu, 2015) showed that the boundary theory with ν identical Majorana fermions is anomalous unless ν is an integer multiple of 8. It is actually rather subtle what happens for $\nu = 8$. We will show that although $\exp(-i\pi\eta/2)$ is a 16th root of unity in general, it is an 8th root of unity in the case of a mapping torus. Thus the $\nu = 8$ theory does not have a traditional anomaly that can be detected via the mapping torus construction. Nevertheless, for $\nu = 8$ there is a more subtle problem, analogous to what we explained in Sec. II.B.3 for a single Majorana fermion considered on an oriented three-manifold only: even though there is no anomaly in the traditional sense, there is also no natural way to specify the sign of the partition function. In Sec. II.B.3, this was because $(-1)^{x/2}$ is in general non-trivial, though it equals 1 for a mapping torus; in the present context, the reason is that the analogous statement holds for the invariant $\exp(-8 \cdot i\pi\eta/2)$.

The $\nu = 8$ problem has one more interesting property that certainly shows its subtlety. It actually is possible to define a gapped symmetry-preserving boundary state without topological order for the $\nu = 8$ theory, that is for the 8th power of the $\nu = 1$ sTQFT, but this boundary state is not compatible with unitarity.

The $d = 3$ topological insulator is also related to the η invariant by the same arguments. But in this case there is no analog of the integer-valued parameter ν . Hence the results are less interesting and going to an unorientable spacetime does not reveal much structure beyond what we explored in Sec. II.

B. Anomaly of the boundary fermions

An unorientable three-manifold Y can always be constructed as a quotient \hat{Y}/\mathbb{Z}_2 , where \hat{Y} is the oriented double cover of Y , and \mathbb{Z}_2 is generated by an orientation-reversing symmetry $\tau: \hat{Y} \rightarrow \hat{Y}$ that obeys $\tau^2 = 1$. Suppose that we are given a τ -invariant spin structure \mathcal{S} on \hat{Y} . This means, in particular, that we have a notion of fermion field $\hat{\psi}$ on \hat{Y} and an action of τ on $\hat{\psi}$, obeying $\tau^2 = 1$.

In this case, we can define a pin^+ structure \mathcal{P} on Y , by saying that a fermion field ψ on Y with this pin^+ structure is a fermion field $\hat{\psi}$ on \hat{Y} that obeys $\tau\hat{\psi} = \hat{\psi}$. Moreover, every pin^+ structure on Y arises in this way, for some choice of τ -invariant spin structure \mathcal{S} on Y .

But we can also define a “complementary” pin^+ structure \mathcal{P}' on Y , by saying that a fermion field ψ' on Y with this alternative pin^+ structure is a fermion field $\hat{\psi}'$ on \hat{Y} that obeys $\tau\hat{\psi}' = -\hat{\psi}'$. Actually, the relationship between \mathcal{P} and \mathcal{P}' is completely reciprocal and neither of them is preferred. When given a geometrical symmetry $\tau: \hat{Y} \rightarrow \hat{Y}$ and a spin structure $\mathcal{S} \rightarrow \hat{Y}$ that is τ invariant up to isomorphism, this does not tell us the sign with which τ should be taken to act on \mathcal{S} . The choice of sign determines what we mean by \mathcal{P} and what we mean by \mathcal{P}' .

On \hat{Y} , we can define the usual Hermitian Dirac operator $\mathcal{D} = i\mathcal{D} = i\sum_k \gamma^k D_k$. However, in odd dimensions, the Dirac operator anticommutes with an orientation-reversing symmetry. This point was illustrated in Eq. (2.22), and, as explained there, is the reason that a mass term added to the Dirac equation in $2 + 1$ dimensions would violate reflection symmetry.

In particular, \mathcal{D} is odd under τ , so if $\tau\psi = \pm\psi$, then $\tau(\mathcal{D}\psi) = \mp \mathcal{D}\psi$. Interpreted on Y , this statement means that there is no self-adjoint Dirac operator acting on a fermion field ψ that is a section of \mathcal{P} , or on a fermion field ψ' that is a section of \mathcal{P}' . Rather, the natural Dirac operator on Y maps \mathcal{P} to \mathcal{P}' and \mathcal{P}' to \mathcal{P} . The operators $\mathcal{D}_{\mathcal{P}}: \mathcal{P} \rightarrow \mathcal{P}'$ and $\mathcal{D}_{\mathcal{P}'}: \mathcal{P}' \rightarrow \mathcal{P}$ are adjoints, so we can combine them to a self-adjoint operator which in the basis $\begin{pmatrix} \mathcal{P} \\ \mathcal{P}' \end{pmatrix}$ is

$$\mathcal{D} = \begin{pmatrix} 0 & \mathcal{D}_{\mathcal{P}'} \\ \mathcal{D}_{\mathcal{P}} & 0 \end{pmatrix}. \quad (4.1)$$

In fact, \mathcal{D} is just the usual self-adjoint Dirac operator of \hat{Y} (for spin structure \mathcal{S}), acting on a spinor field $\hat{\psi}$ that has been reinterpreted as a pair ψ, ψ' on Y .

More fundamentally, the reason that there is no self-adjoint Euclidean signature Dirac operator for ψ or for ψ' is that in Euclidean signature these are complex fermions, not real or pseudoreal ones. As explained in Appendix B.2, the group $\text{Pin}^+(3)$ has two different two-dimensional representations,

which are complex conjugates of each other. The fields ψ and ψ' are associated with these two different representations.

Although we cannot write a self-adjoint Dirac operator for ψ or ψ' only, there is no problem with the classical Dirac action

$$\int d^3x \sqrt{g} \bar{\psi} i \mathcal{D} \psi \quad (4.2)$$

for ψ (or likewise for ψ') only. Recall that, as discussed in Sec. II.B.1, the Dirac action really involves an antisymmetric bilinear form \mathbf{D} rather than an operator \mathcal{D} . Up on \hat{Y} , the relation between the two is $\mathbf{D}_{\gamma\beta} = \varepsilon_{\gamma\alpha} \mathcal{D}_{\beta}^{\alpha}$, where $\varepsilon_{\gamma\beta}$ is the antisymmetric bilinear form on the two-dimensional representation of $\text{Spin}(3) = \text{SU}(2)$. But this form is odd under τ [more generally, it is odd under the disconnected component of $\text{Pin}^+(3)$]. As $\mathcal{D}_{\beta}^{\alpha}$ is also odd under τ , it follows that $\mathbf{D}_{\gamma\beta}$ is even, and thus can be defined only for the τ -even fermion ψ or the τ -odd fermion ψ' .

All of what we have said has an analog that is probably much more familiar on an orientable manifold of, say, dimension 2. In 2 dimensions, we can consider chiral fermions ψ_+ and ψ_- which in Euclidean signature are complex conjugates of each other. The 2D Dirac operator maps ψ_+ to ψ_- and ψ_- to ψ_+ , and to write a self-adjoint Dirac operator one has to combine the two chiralities. But the Dirac action pairs ψ_+ with ψ_+ and ψ_- with ψ_- , so (modulo questions of anomaly cancellation) it makes sense to consider a theory with one chirality only—such as the edge modes of a $(2+1)$ -dimensional chiral topological superconductor.

Returning to the fields ψ and ψ' on a possibly unorientable three-manifold, the Dirac action and path integral for ψ' are the complex conjugates of those for ψ , and the two partition functions Z_{ψ} and $Z_{\psi'}$ are complex conjugates. In fact, concretely ψ and ψ' combine to the fermion $\hat{\psi}$ on \hat{Y} , so formally $Z_{\psi} Z_{\psi'} = Z_{\hat{\psi}}$, and we claim that $Z_{\hat{\psi}} \geq 0$. As usual, we have $Z_{\hat{\psi}} = |Z_{\hat{\psi}}| \exp(-i\pi\eta/4)$. However, as the Dirac operator \mathcal{D} of \hat{Y} is odd under τ , its spectrum is symmetric under $\lambda \leftrightarrow -\lambda$, so nonzero modes on \hat{Y} do not contribute to η . Either there are no zero modes, in which case $\eta = 0$ and $Z_{\hat{\psi}} > 0$, or there are zero modes and $Z_{\hat{\psi}} = 0$.

While the product $Z_{\psi} Z_{\psi'}$ is positive and, in particular, is completely anomaly free, Z_{ψ} and $Z_{\psi'}$ separately are complex and subject to anomalies. They have no perturbative anomalies (in general there are no perturbative anomalies in odd-dimensional spacetimes without boundary), but they have global anomalies. Indeed, the global anomaly of Z_{ψ} was computed recently by Hsieh, Cho, and Ryu (2015) for a particular example. This is analogous to the fact that, in the more familiar case of the chiral fermions ψ_+ and ψ_- in 2 spacetime dimensions, Z_{ψ_+} and Z_{ψ_-} are anomalous (in this case with perturbative as well as global anomalies), but the product $Z_{\psi_+} Z_{\psi_-}$ is non-negative and anomaly free.

A general framework to calculate the global anomaly in a complex fermion path integral was described in Sec. 4 of Witten (1985b). Given a diffeomorphism ϕ of Y (or a combined diffeomorphism and gauge transformation in a more general problem), one constructs as in Sec. II.A.2 a mapping torus X . X is a four-manifold obtained by gluing

together the ends of $Y \times I$, where I is the interval $0 \leq s \leq 1$, via ϕ . We endow X with a metric

$$d\ell^2 = ds^2 + \varepsilon^2 g_s, \quad (4.3)$$

where g_s is an s -dependent family of metrics on Y and ε is a suitable small parameter. In the basis (4.1), one introduces a new gamma matrix

$$\gamma_s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.4)$$

Then one defines a Dirac operator on X :

$$\mathcal{D}_X = i\gamma_s \frac{D}{Ds} + \frac{1}{\varepsilon} \begin{pmatrix} 0 & \mathcal{D}_{\mathcal{P}'} \\ \mathcal{D}_{\mathcal{P}} & 0 \end{pmatrix}. \quad (4.5)$$

Here

$$\mathcal{D}_Y = \begin{pmatrix} 0 & \mathcal{D}_{\mathcal{P}'} \\ \mathcal{D}_{\mathcal{P}} & 0 \end{pmatrix}$$

is the Dirac operator on Y with s -dependent metric g_s , and the explicit factor $1/\varepsilon$ comes because in Eq. (4.3) we use the metric $\varepsilon^2 g_s$. The result of the analysis by Witten (1985b) is to show that the change of the partition function Z_{ψ} under ϕ can be expressed in terms of the η invariant of $D = 4$ Dirac operator \mathcal{D}_X :

$$Z_{\psi} \rightarrow Z_{\psi} \exp(-i\pi\eta/2). \quad (4.6)$$

We do not repeat this calculation in detail and just mention a few points. Because we are trying to compute a topological invariant, we can consider a generic s -dependent family of metrics g_s such that \mathcal{D}_Y has no level crossings as a function of s . By also taking ε sufficiently small, one can reduce to the case that the spectrum of \mathcal{D}_X , and therefore its η invariant, can be computed by an adiabatic approximation.³⁵ In the adiabatic approximation, the contribution of each eigenvalue of \mathcal{D}_Y to $\exp(-i\pi\eta/2)$ comes essentially from the holonomy of its Berry connection (Berry, 1984). [The Berry connection for adiabatic evolution of a given eigenstate is introduced, although not by that name, in the middle of p. 215 of Witten (1985b).] In the adiabatic approximation, the same holonomy gives the contribution of each eigenvalue to the change of phase of Z_{ψ} under ϕ .

A consequence of Eq. (4.6) is that, as in previous examples, the anomaly is a cobordism invariant, meaning that it is trivial if X is the boundary of a five-manifold Z over which the pin^+ structure extends. Here we make use of Eq. (2.47). On a five-dimensional pin^+ manifold, the index \mathcal{I} is always even because of pseudoreality, so Eq. (2.47) implies that if X is a boundary, then η is a multiple of 4 and hence $\exp(-i\pi\eta/2) = 1$.

³⁵If $\lambda_i(s)$ are the eigenvalues of $\varepsilon^{-1} \mathcal{D}_Y$ as a function of s (we assume that these are nondegenerate except for the twofold degeneracy that comes from pseudoreality), the condition for the adiabatic approximation is that $d\lambda_i/ds$ should be sufficiently small compared to the relevant eigenvalue differences. This condition is satisfied for sufficiently small ε , assuming there are no level crossings.

The interested reader who actually consults [Witten \(1985b\)](#) and the later paper by [Dai and Freed \(1994\)](#) in which an important refinement was made will find that it is assumed that Y has even dimension and X has odd dimension. The reason for this was simply that, if Y is orientable, complex fermions on Y exist only if its dimension is even. Applications involving unorientable manifolds were not envisioned so it was natural to assume Y to have even dimension. However, the reasoning in these papers applies equally well to complex fermions in any dimension, even or odd.

In terms of the applications of the results, an important difference comes from the fact that in even dimensions complex fermions generically give rise to perturbative anomalies.³⁶ Moreover, if Y has even dimension, then X has odd dimension, and if it is the boundary of some Z , then Z again has even dimension. If \mathcal{D}_X is a Dirac operator on X and η is its η invariant, then typically $\exp(-i\pi\eta/2)$ is not a cobordism invariant (or even a topological invariant) because there is a curvature term in the APS index theorem on Z . [For Z of dimension 4, this curvature term is the term $\hat{A} - P$ of Eq. (2.38).] However, on an even-dimensional manifold Y , we recover cobordism invariance in the case of complex fermions whose perturbative anomalies cancel.³⁷ This follows from the formula for η given by the APS index theorem. Here of course the global anomaly comes from an η invariant that receives contributions from all fermion multiplets in a given theory; so in using the APS index theorem, we likewise sum over all multiplets. The curvature term in the APS index theorem on Z is precisely the anomaly polynomial of the fermions, so after summing over multiplets, it cancels in a theory that is free of perturbative anomalies. The APS index theorem then shows that in any theory without perturbative anomalies, in even or odd D , the invariant $\exp(-i\pi\eta/2)$ that controls the global anomaly is a cobordism invariant.

The relationship of η to global anomalies is a slight refinement of a result that is much better known: perturbative anomalies in D dimensions are related to Chern-Simons in $D + 1$ dimensions and to a curvature polynomial in $D + 2$ dimensions. In fact the APS index theorem implies [see Eqs. (2.51)–(2.53)] a close relationship between η and Chern-Simons. Usually the distinction between them is important only when one asks rather delicate questions, for example, about global anomalies.

C. Bulk partition function

Here we compute the bulk partition function of the $(3 + 1)$ -dimensional topological superconductor. The strategy should be familiar from Secs. II.A.6 and III.D.

A standard description of the phase transition between an ordinary and a topological superconductor in $3 + 1$

dimensions is that it occurs when the mass m of a $D = 4$ Majorana fermion ψ passes through 0. Even on an unorientable four-manifold X , one can define a Hermitian Dirac operator \mathcal{D} (we recall from Sec. IV.B that this is not true in 3 spacetime dimensions), and its spectrum is doubled by a version of Kramers doubling, reflecting the fact that the Majorana fermion transforms in a pseudoreal representation of $\text{Pin}^+(4)$. For details on these statements, see Appendixes A.3 and B.3. The path integral of the Majorana fermion is a Pfaffian which can be computed by taking one eigenvalue from each pair. We write m for the bare mass of ψ and we regulate the partition function Z_ψ by including a Pauli-Villars field of mass μ obeying the same Dirac equation but with opposite statistics. The regularized partition function is

$$Z_{\psi,\text{reg}} = \prod_k \prime \frac{\lambda_k + im}{\lambda_k + i\mu}, \quad (4.7)$$

where as usual $\prod_k \prime$ is an instruction to take a single eigenvalue from each pair.

There are essentially two cases, depending on whether m and μ have the same sign or opposite signs. If m and μ have the same sign, we get a topologically trivial superconductor. Indeed, in the limit that X is very large, the precise values of m and μ do not matter, only their signs. If m and μ have the same sign, we may as well set $m = \mu$, and then $Z_{\psi,\text{reg}} = 1$ is completely trivial.

The interesting case is that m and μ have opposite signs. In this case, we may as well set $m = -\mu$, so

$$Z_{\psi,\text{reg}} = \prod_k \prime \frac{\lambda_k - i\mu}{\lambda_k + i\mu}. \quad (4.8)$$

A shortcut to analyzing this expression is to factor it

$$Z_{\psi,\text{reg}} = \prod_k \prime \frac{\lambda_k}{\lambda_k + i\mu} \cdot \prod_k \prime \frac{\lambda_k - i\mu}{\lambda_k}. \quad (4.9)$$

The second factor is the inverse complex conjugate of the first, so they contribute the same phase. The first factor was analyzed in Sec. II.A.3, with a result given in Eq. (2.17). In using this result, we have to replace η (computed in the space of positive charge fermions only) with $\eta/2$ (where η is computed using all eigenvalues of the Dirac operator and the factor of $1/2$ reflects the $\prod_k \prime$ symbol). This replacement $\eta \rightarrow \eta/2$ should be familiar from Eqs. (2.68) and (2.69). Thus the two factors in Eq. (4.9) each contribute a factor $\exp(-i\pi\eta/4)$, and the final result for the partition function of the bulk sTQFT of the topological superconductor is $\exp(-i\pi\eta/2)$.

As a check, let us verify that if X is orientable, this answer agrees with the result $(-1)^{I/2}$ that was claimed in Sec. II.B.3. If X is orientable, we can define the chirality operator $\tilde{\gamma}$ [Eq. (2.31)] and the operation $\psi \rightarrow \tilde{\gamma}\psi$ ensures that nonzero eigenvalues of \mathcal{D} are paired under $\lambda \leftrightarrow -\lambda$. (The spectrum also has even multiplicity for each λ because of pseudoreality, so nonzero modes are in quartets $\lambda, \lambda, -\lambda, -\lambda$.) Pairs $\lambda, -\lambda$ do not contribute to η , and for X orientable, η receives contributions only from zero modes. Let n_+ and n_- be the number of zero

³⁶In [Witten \(1985b\)](#) but not in [Dai and Freed \(1994\)](#), it is assumed that perturbative anomalies cancel.

³⁷An example is the standard model of particle physics, in $D = 4$, with gauge group $G = \text{SU}(3) \times \text{SU}(2) \times \text{U}(1)$. In Euclidean signature, the fermions are in a complex but highly reducible representation of $\text{Spin}(4) \times G$. This representation is such that perturbative anomalies cancel.

modes of positive or negative chirality. Each zero mode contributes $+1$ to η [see Eq. (2.19)], so $\eta = n_+ + n_-$ and $\exp(-i\pi\eta/2) = \exp(-i\pi(n_+ + n_-)/2)$. On the other hand, $\mathcal{I} = n_+ - n_-$, so that $(-1)^{\mathcal{I}/2} = (-1)^{(n_+ - n_-)/2}$. These are equal as n_+ and n_- are even. For a more direct explanation, one can return to Eq. (4.8) and see that a quartet of eigenvalues $\lambda, \lambda, -\lambda, -\lambda$ makes no contribution but a pair of zero modes contributes a factor of -1 , just as in $(-1)^{\mathcal{I}/2}$.

Note that in the course of the argument we have shown that if X is orientable, then η is an even integer. But if X is orientable and has an orientation-reversing symmetry, then $n_+ = n_-$ and so η is a multiple of 4.

D. Anomaly cancellation and the Dai-Freed theorem

As usual, there is a problem in defining the partition function $\exp(-i\pi\eta/2)$ of the bulk sTQFT on a four-manifold X with boundary. There does not exist a local, symmetry-preserving, self-adjoint boundary condition for the 4D Dirac operator in Euclidean signature that could be used to give a satisfactory definition of η . We can use nonlocal APS boundary conditions to define η in a symmetry-preserving fashion. But $\exp(-i\pi\eta/2)$ defined this way is not a topological invariant and only makes sense physically when combined with gapless modes on the boundary.

In Sec. IV.B, we explained that the usual boundary fermions of the topological superconductor are complex fermions with a complex path integral when formulated on an unorientable three-manifold Y . Moreover, according to the general global anomaly formula for complex fermion path integrals (Witten, 1985b), the anomaly involves $\exp(-i\pi\eta/2)$ where as in the previous paragraph, η is the η invariant of a $D = 4$ Majorana fermion.

As in other cases that we have studied, what is physically sensible is the product

$$|\text{Pf}(\mathbf{D})| \exp(-i\pi\eta/2), \quad (4.10)$$

where $\text{Pf}(\mathbf{D})$ is the Pfaffian of the boundary fermions and η is the bulk η -invariant computed with APS boundary conditions. Indeed, the assertion that this product is physically sensible is essentially the content of what we will call the Dai-Freed theorem (Dai and Freed, 1994), which is a refinement of the general global anomaly formula.

We do not attempt to describe the proof of the Dai-Freed theorem, but we make a few remarks that may be helpful. First, the theorem is actually phrased in Dai and Freed (1994) in a more abstract way. To explain this, we note that the usual mathematical viewpoint about anomalies is to say that although $\text{Pf}(\mathbf{D})$ cannot be naturally defined as a complex number, or as a complex-valued function of the metric on Y , there is always a unitary complex line bundle over the space of metrics on Y (or metrics plus gauge fields in a more general case) such that $\text{Pf}(\mathbf{D})$ is canonically defined as a section of this line bundle. This line bundle is called the Pfaffian line bundle and we denote it as $Pf(\mathbf{D})$. The Dai-Freed theorem as stated by Dai and Freed (1994) asserts that $\exp(i\pi\eta/2)$ is naturally defined as a section of $Pf(\mathbf{D})$. An equivalent statement is that the product $\text{Pf}(\mathbf{D}) \exp(-i\pi\eta/2)$ (where the two factors are sections of inverse line bundles) is well defined as a complex

number. However, as long as we restrict to the locus on which the boundary Dirac operator has no zero modes, the line bundle $Pf(\mathbf{D})$ has a natural trivialization. Relative to this trivialization, the Pfaffian $\text{Pf}(\mathbf{D})$ is positive and can be replaced by its absolute value, and $\exp(-i\pi\eta/2)$ is likewise a well-defined complex number,³⁸ with the usual definition (2.69).

Finally, we explain in concrete terms the meaning of the statement that the product $|\text{Pf}(\mathbf{D})| \exp(-i\pi\eta/2)$ is physically sensible. For real or pseudoreal fermions, a key point was that $|\text{Pf}(\mathbf{D})|(-1)^\zeta$ or $|\text{Pf}(\mathbf{D})|(-1)^{\mathcal{I}/2}$ behaves sensibly when the boundary Dirac operator develops a zero eigenvalue. This happened because the sign of $(-1)^\zeta$ or $(-1)^{\mathcal{I}/2}$ jumps whenever the path integral of the boundary fermions is expected to change sign. For complex fermions, we are not dealing with sign changes, because $\text{Pf}(\mathbf{D})$ is not naturally real. But it is still essential to understand what happens when zero modes develop.

Near a point at which the boundary Dirac operator develops a pair of zero modes, the associated bilinear form \mathbf{D} , in its canonical form, has a 2×2 block

$$\begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}, \quad (4.11)$$

with small λ . Here λ is a complex number and to make it vanish we must vary two parameters in the metric and/or gauge field on Y . For example, there may be two parameters u_1 and u_2 such that

$$\lambda = u_1 + iu_2. \quad (4.12)$$

Equivalently, setting $u_1 + iu_2 = \varrho e^{i\varphi}$, we have $\lambda = \varrho e^{i\varphi}$. We expect the path integral to be proportional to λ for $\lambda \rightarrow 0$. For the formula $|\text{Pf}(\mathbf{D})| \exp(-i\pi\eta/2)$ for the combined bulk and boundary partition function to be physically sensible near $\lambda = 0$, $\exp(-i\pi\eta/2)$ must not be smooth near $\lambda = 0$. On the contrary, it must be proportional to $e^{i\varphi}$, which is ill defined at $\lambda = 0$. Concretely, this happens because of the way APS boundary conditions depend on φ . APS boundary conditions depend on which modes of the Dirac operator on Y have negative eigenvalues. For very small (positive) ϱ , the Dirac operator on Y has low-lying modes, and which linear combinations of them are eigenmodes with negative eigenvalue depends on φ .

A more comprehensive explanation of the statement that Eq. (4.10) for the fermion path integral is physically sensible involves the following. The variation of this expression in an infinitesimal change in the metric of X depends only on the change in the metric along the boundary of X and is related in the usual way to the expectation value of the fermion stress tensor.

³⁸We do not need to discuss the meaning of $\exp(-i\pi\eta/2)$ when the boundary Dirac operator does have zero modes, since in that case $\text{Pf}(\mathbf{D}) = 0$. As explained by Dai and Freed (1994), this case is subtle because the definition of APS boundary conditions is subtle when the boundary fermions have zero modes.

E. The case of ν bands

At the free fermion level, the topological superconductor is characterized by an integer-valued invariant ν . Let us recall how ν is defined.

On the boundary, we consider n_1 Majorana fermions that transform under T as

$$T\psi(-t, x_1, x_2) = \gamma_0\psi(t, x_1, x_2), \quad (4.13)$$

and n_2 that transform with the opposite sign

$$T\psi(-t, x_1, x_2) = -\gamma_0\psi(t, x_1, x_2). \quad (4.14)$$

A pair of Majorana fermions ψ, ψ' that transform with opposite signs can have a T -conserving bare mass $\bar{\psi}\psi'$. Up to T -invariant perturbations, only the difference $\nu = n_1 - n_2$ is relevant.

In bulk, the basic object, as explained in Sec. IV.C, is a Majorana fermion ψ of mass m that has a regulator of mass μ . The interesting case is that m and μ have opposite signs. In Sec. IV.C, we took $\mu > 0, m < 0$ and arrive at $\exp(-i\pi\eta/2)$ as the sTQFT partition function. With the opposite signs, $\mu < 0$ and $m > 0$, the same derivation would give $\exp(i\pi\eta/2)$. In bulk, ν is defined as the difference between the number of $\mu > 0, m < 0$ pairs and the number of $\mu < 0, m > 0$ pairs. The bulk partition function is $\exp(-\nu i\pi\eta/2)$.

At the free fermion level, ν is an integer-valued invariant, but it is known (Kitaev, 2015) that when T -conserving interactions are taken into account, ν is a topological invariant only mod 16.

It was pointed out by Kapustin *et al.* (2014) that for a four-dimensional Pin^+ manifold, $\exp(-i\pi\eta/2)$ is always (Gilkey, 1985; Stolz, 1988) a 16th root³⁹ of 1. Accordingly, the product of ν copies of the sTQFT whose partition function is $\exp(-i\pi\eta/2)$ is trivial if and only if ν is an integer multiple of 16. This was offered by Kapustin *et al.* (2014) as an explanation of why even at the interacting level ν is well defined mod 16. We have put those considerations on a firmer footing by showing in Sec. IV.C that the partition function of the basic sTQFT associated with the $D = 4$ topological superconductor is indeed $\exp(-i\pi\eta/2)$.

F. Special behavior for $\nu=8$

It was shown by Hsieh, Cho, and Ryu (2015) that, at least in a special case, anomalies associated with the mapping torus detect the value of $\nu \bmod 8$ but not mod 16. We can now explain why this is true. First, $\exp(-i\pi\eta/2)$ is a cobordism invariant of a pin^+ four-manifold X . Moreover, it generates the group of cobordism invariants, and any other $U(1)$ -valued cobordism invariant is a power of this one. In particular, taking w_4 to represent the fourth Stiefel-Whitney class, $(-1)^{w_4}$ is a nontrivial $U(1)$ -valued cobordism invariant so it must be a

power of $\exp(-i\pi\eta/2)$. The power must be 8 since $(-1)^{w_4}$ is valued in $\{\pm 1\}$:

$$(-1)^{w_4} = \exp(-8i\pi\eta/2). \quad (4.15)$$

For example, for $X = \mathbb{R}P^4$, one has $w_4 = 1$, $\exp(-\pi i\eta/2) = \exp(\pm 2i\pi/16)$, with the sign depending on the pin^+ structure.

Thus the $\nu = 8$ case of a topological superconductor has $(-1)^{w_4}$ for its bulk partition function. The Stiefel-Whitney classes are defined without choosing a spin or pin structure, so unlike what happens for other values of ν , the $\nu = 8$ topological superconductor does not depend on the pin^+ structure of the four-manifold X . There is a full-fledged TQFT with partition function $(-1)^{w_4}$.

For a rank 4 real vector bundle, such as the tangent bundle of a four-manifold, w_4 is the same as the mod 2 reduction of the Euler class. So an alternative statement is that the $\nu = 8$ theory has $(-1)^{\chi(X)}$ as its bulk partition function,⁴⁰ where $\chi(X)$ is the Euler characteristic of X .

Now we can explain the result found by Hsieh, Cho, and Ryu (2015) and see that it is general. If X is a mapping torus, meaning that it is total space of a fiber bundle $X \rightarrow S^1$, the fiber being a three-manifold Y , then $\chi(X) = 0$. One way to prove this is to observe that in general the Euler characteristic is multiplicative in fibrations, so that $\chi(X) = \chi(Y)\chi(S^1)$. But this vanishes, since $\chi(S^1) = 0$.

We are then in a similar situation to what we found in Sec. II.B.3 for the same problem on an orientable manifold. If ν is a multiple of 8, this ensures that the theory of ν identical Majorana fermions on a three-manifold Y has no traditional global anomaly—detectable via the invariant $\exp(-\nu i\pi\eta/2)$ of a mapping torus. But to give a natural definition of the sign of the partition function for a system of ν identical $(2+1)$ -dimensional Majorana fermions, with no need for “anomaly inflow” from a four-manifold X , one wants $\exp(-\nu i\pi\eta/2)$ to equal 1 for all four-manifolds, not just mapping tori. For this, we need ν to be a multiple of 16.

The fact that the bulk state for $\nu = 8$ has a partition function $(-1)^{\chi(X)}$ may seem, at first sight, to offer a way to define a symmetry-preserving gapped boundary state without topological order. After all, the Euler characteristic of a manifold with boundary is perfectly well defined and has natural cobordism properties. In the TQFT with partition function $(-1)^{\chi(X)}$, one can define a symmetry-preserving boundary state by saying that the partition function on a four-manifold X with any spatial boundary Y is $(-1)^{\chi(X)}$. This boundary state satisfies cutting and gluing axioms of TQFT, but it is not compatible with reflection positivity or unitarity. For a simple counterexample, take Y to be a 3-sphere and X a 4-ball with boundary Y . Then $(-1)^{\chi(X)} = -1$, but because X can be built by gluing two identical pieces as in Fig. 8 of Sec. II.A.9,

³⁹Moreover, all 16th roots do occur. For $X = \mathbb{R}P^4$, one has $\exp(-i\pi\eta/2) = \exp(\pm 2i\pi/16)$, where the sign depends on the choice of pin^+ structure. For an introduction to these matters, see Appendix C.

⁴⁰We started with w_4 rather than χ because Stiefel-Whitney numbers such as w_4 are always cobordism invariants, which made it easier to explain the relationship to $\exp(-i\pi\eta/2)$. Likewise, the relationship of χ to w_4 is the simplest way to explain that $(-1)^{\chi(X)}$ is a cobordism invariant.

reflection positivity would require that the partition function for X should be positive.

G. The $d=3$ topological insulator on an unorientable spacetime

Now we very briefly consider the $(3+1)$ -dimensional topological insulator on a possibly unorientable spacetime. Everything that we have said about the topological superconductor has a fairly obvious analog, with one exception. In contrast to the integer-invariant ν of the topological superconductor, the topological insulator has only a \mathbb{Z}_2 -valued invariant, even at the free fermion level. Indeed, we have already written in Eq. (2.4) a T -conserving mass term for a pair of $(2+1)$ -dimensional Dirac fermions ψ_1 and ψ_2 that transform the same way under T . T symmetry only protects mod 2 the number of massless charged Dirac fermions.

We therefore expect that it will turn out that $\exp(-i\pi\eta/2)$ is always ± 1 for the topological insulator, even on an unorientable four-manifold. To understand this statement, we first recall that the symmetry that protects the SPT state of a topological insulator is, in relativistic terminology,⁴¹ T and not CT . As discussed in Sec. III.G, this means that in Euclidean signature we should consider the symmetry CR —in other words, a symmetry that reverses the orientation of spacetime also acts as charge conjugation.

Now let X be a possibly unorientable four-manifold and \hat{X} its oriented double cover. The invariant $\exp(-i\pi\eta/2)$ on X that we want to understand is the η invariant of a Dirac operator acting on fermion fields ψ_{\pm}^{\pm} , where the superscript is the $U(1)$ charge and the subscript is the chirality. There is an equivalent definition on \hat{X} in which one considers only fermion fields $\hat{\psi}_{\pm}^{\pm}$ of positive $U(1)$ charge but any chirality: the Dirac spectrum on \hat{X} computed from eigenvalues of the Dirac operator acting on $\hat{\psi}_{\pm}^{\pm}$ only is the same as the Dirac spectrum on X in the full set of fields ψ_{\pm}^{\pm} . Indeed, X is the quotient of \hat{X} by a CR symmetry, so the fields ψ_{+}^{\pm} and ψ_{-}^{\pm} on X combine to $\hat{\psi}_{+}^{\pm}$ on \hat{X} , and similarly ψ_{+}^{\pm} and ψ_{-}^{\pm} on X combine to $\hat{\psi}_{-}^{\pm}$ on \hat{X} . So instead of computing an η invariant by summing over the spectrum of a Dirac operator on X acting on the full set of fields ψ_{\pm}^{\pm} , we can sum over the spectrum of a Dirac operator on \hat{X} acting on $\hat{\psi}_{\pm}^{\pm}$ only.

As \hat{X} is orientable, the Dirac spectrum on \hat{X} is symmetric under $\lambda \leftrightarrow -\lambda$. This follows as usual from considerations of chirality. But a pair of eigenvalues $\lambda, -\lambda$ does not contribute to the η invariant. So in computing η on \hat{X} , we can consider zero modes only. The space of zero modes of the Dirac operator on \hat{X} is even dimensional because of the usual antilinear symmetry of the Dirac equation, so η is an even integer. Consequently, $\exp(-i\pi\eta/2) = \pm 1$, as we aimed to show.

V. THE MAJORANA CHAIN

At a microscopic level, the time-reversal operation that holds in the real world obeys $T^2 = (-1)^F$, where $(-1)^F$ is the operator that counts fermions mod 2. But there are some

⁴¹ C and CT , which exchange electrons with positrons, are never true symmetries in condensed matter physics.

situations in condensed matter physics in which it is natural to consider a time-reversal symmetry that obeys $T^2 = 1$. This usually happens because either spin is unimportant or the symmetry under consideration is really the product RT of time reversal with a reflection of one space coordinate. [In nature $(RT)^2 = 1$, and RT can act as time reversal on a thin layer or long chain that is localized at the fixed point set of R .]

In the context of SPT phases of matter, there is a very interesting example in which a time-reversal symmetry is assumed with $T^2 = 1$. This is the Majorana chain in one space dimension (Fidkowski and Kitaev, 2009).

A boundary of such a system is zero dimensional, that is a point, and we begin there. In this paper, we have built in the condition $T^2 = (-1)^F$ by taking our Lorentz signature gamma matrices to be real matrices obeying $\{\gamma_{\mu}, \gamma_{\nu}\} = 2\eta_{\mu\nu}$, where $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, \dots, 1)$. Then we take T to act by

$$\psi(t, \vec{x}) \rightarrow \pm \gamma_0 \psi(-t, \vec{x}), \quad (5.1)$$

ensuring that $T^2\psi = -\gamma_0^2\psi = -\psi$. Since $\gamma_i^2 = +1$ for $i > 0$, it follows that a spatial reflection squares to $+1$. In Euclidean signature, this leads to a pin^+ structure.

In $0+1$ dimensions, we need only one gamma matrix γ_0 . To get the most economical possible theory, we take $\gamma_0^2 = +1$, rather than -1 as in our discussion so far. Then we can represent γ_0 by the number 1×1 matrix 1 (or -1), and it is possible to consider a 1-component real (Majorana) fermion field ψ with action

$$I = \frac{i}{2} \int dt \psi \frac{d}{dt} \psi. \quad (5.2)$$

We take T to act on such a field by

$$T\psi(t) = \pm \psi(-t), \quad (5.3)$$

with some choice of the sign, so that $T^2 = 1$. A pair of fermions ψ and ψ' that transform with opposite signs can have a T -invariant “mass” term $i\psi\psi'$, and as in Sec. IV.E, the natural invariant is the difference ν between the number of fermions transforming under T with a $+$ sign and the number transforming with a $-$ sign. At the free fermion level, ν is an integer invariant, but (Fidkowski and Kitaev, 2009) at the interacting level, ν is only an invariant mod 8. [This statement means that a $(0+1)$ -dimensional system consisting of eight real fermions all transforming under T with the same sign can be gapped by a suitable T -invariant four-fermion coupling.]

When ν is not a multiple of 8, this system has a subtle (ν -dependent) anomaly in the realization of the symmetries T and $(-1)^F$. We will not repeat this story here, and instead we will ask how the system with ν boundary fermions can naturally appear on the boundary of a T -conserving system in $1+1$ dimensions. The choice $\gamma_0^2 = +1$ that we have made in setting up this problem corresponds to a Pin^- structure, and once we have assumed a Pin^- structure on the $(0+1)$ -dimensional boundary of spacetime, we must also (if we want the Dai-Freed theorem or a similar tool to be available) assume a Pin^- structure on the $1+1$ -dimensional bulk. So we consider fermions with $(1+1)$ -dimensional gamma matrices

that obey $\gamma_0^2 = 1 = -\gamma_1^2$, or more covariantly $\{\gamma_\mu, \gamma_\nu\} = -2\eta_{\mu\nu}$. In Euclidean signature, this becomes $\{\gamma_\mu, \gamma_\nu\} = -2\delta_{\mu\nu}$, where the minus sign corresponds to a Pin^- structure. (2×2 matrices satisfying $\{\gamma_\mu, \gamma_\nu\} = -2\delta_{\mu\nu}$ exist but cannot be chosen to be real. The fact that a bare mass is possible, as discussed next, makes it clear that a Majorana fermion with $T^2 = 1$ is pseudoreal in Euclidean signature.)

In contrast to the case $T^2 = -1$, a 2D Majorana fermion with $T^2 = 1$ can have a T-invariant bare mass. In Lorentz signature, the T-invariant Dirac equation can be written

$$(\gamma^0 \partial_0 + \gamma^1 \partial_1 - m\bar{\gamma})\psi = 0, \quad (5.4)$$

where m is real and $\bar{\gamma} = \gamma^0 \gamma^1$ is the chirality operator. Actually, this equation is T invariant for either sign in $\{\gamma_\mu, \gamma_\nu\} = \pm 2\eta_{\mu\nu}$, but if one writes the equation in Hamiltonian form $i\partial_t \psi = H\psi$ for some linear operator H , one sees that H is only Hermitian for the case of a $-$ sign. So that is the case for which this mass term is physically sensible.

After generalizing to curved spacetime and continuing to Euclidean signature, Eq. (5.4) becomes $(\mathcal{D} - im\bar{\gamma})\psi = 0$. But it is convenient to multiply by $\bar{\gamma}$ and write the equation as

$$(\mathcal{D}' + im)\psi = 0. \quad (5.5)$$

Here $\mathcal{D}' = \bar{\gamma}\mathcal{D}$ is a Hermitian and reflection-symmetric Dirac operator. Because it is reflection symmetric, it makes sense on any two-manifold, possibly unorientable, that is endowed with a Pin^- structure. (See Appendix A.3 for more on this.)

The bulk sTQFT of the Majorana chain is constructed by taking a fermion field ψ of mass m , together with a Pauli-Villars regulator field of mass μ . As in Sec. IV.C, the interesting case is that m and μ have opposite signs. For $\mu > 0$, $m < 0$, the phase of the partition function is $\exp(-i\pi\eta/2)$, by the familiar argument. This is the partition function of the bulk sTQFT associated with the Majorana chain. It is always an 8th root of 1; it equals $\exp(\pm 2i\pi/8)$ for $\mathbb{R}P^2$, where the sign depends on the choice of Pin^- structure (see Appendix C). Because $\exp(-\pi i\eta/2)$ is always an 8th root of 1, the bulk sTQFT of the Majorana chain is trivial for ν multiple of 8. [This was argued by Kapustin *et al.* (2014) based on an assumption of cobordism invariance.]

As usual, if ν is not a multiple of 8, the bulk partition function $\exp(-\nu i\pi\eta/2)$ cannot be defined as a topological invariant in a symmetry-preserving fashion on a possibly unorientable two-manifold with boundary. However, the Dai-Freed theorem applies in this situation and says that the product of the path integral of ν boundary fermions times $\exp(-\nu i\pi\eta/2)$ is well defined. This is the path integral of the combined system consisting of the bulk sTQFT and the boundary fermions.

ACKNOWLEDGMENTS

This research was supported in part by the NSF Grant No. PHY-131-4311. I thank M. F. Atiyah, G. Burdman, D. Freed, A. Kitaev, S. Ryu, and N. Seiberg for comments and discussions.

APPENDIX A: SPINORS IN RIEMANNIAN GEOMETRY

1. Spinors on orientable manifolds

To introduce spinors and the Dirac equation on a Riemannian manifold W of dimension⁴² n , it is useful to first pick a “vielbein” that is an orthonormal basis of tangent vectors e_a , $a = 1, \dots, n$. We denote as e_a^i , $i = 1, \dots, n$, the components of e_a in an arbitrary local coordinate system $x^i = \{x^1, \dots, x^n\}$ on W . The statement that the e_a are orthonormal means that, if g_{ij} is the metric tensor of W , then

$$e_a^i e_b^j g_{ij} = \delta_{ab}. \quad (A1)$$

Equivalently,

$$\delta^{ab} e_a^i e_b^j = g^{ij}. \quad (A2)$$

Indices of e_a^i are raised and lowered with δ_{ab} and g_{ij} and, in particular, e_i^a is the inverse matrix to e_a^i . “Flat space” gamma matrices are $2^{\lfloor n/2 \rfloor} \times 2^{\lfloor n/2 \rfloor}$ matrices γ_a , $a = 1, \dots, n$, obeying

$$\{\gamma_a, \gamma_b\} = 2\delta_{ab}. \quad (A3)$$

The corresponding “curved space” gamma matrices are $\gamma_i = e_i^a \gamma_a$, obeying

$$\{\gamma_i, \gamma_j\} = 2g_{ij}. \quad (A4)$$

While the orientation-preserving rotation group $\text{SO}(n)$ acts on the tangent space to W at a given point $p \in W$, the corresponding group that acts on spinors on W is a double cover of $\text{SO}(n)$ that is known as $\text{Spin}(n)$. Its generators are

$$\Sigma_{ab} = \frac{1}{4}[\gamma_a, \gamma_b]. \quad (A5)$$

The covariant derivative of a spinor field Ψ is defined as

$$\frac{D}{Dx^i} \Psi = \left(\frac{\partial}{\partial x^i} + \omega_i^{ab} \Sigma_{ab} \right) \Psi, \quad (A6)$$

where ω_i^{ab} is the Levi-Civita connection on the tangent bundle of W . We can define a Hermitian Dirac operator

$$\mathcal{D} = i \sum_{i=1}^n \gamma^i D_i. \quad (A7)$$

This construction depended on the choice of a vielbein $\{e_a\}$, which in general can be made only locally. To define a global Dirac operator, one glues together local descriptions. A down-to-earth way to proceed is to cover W with small open sets W_x , $x = 1, \dots, s$, picking a vielbein e_{ax} , $a = 1, \dots, n$, on each W_x . On $W_x \cap W_y$, since e_{ax} and e_{ay} are both orthonormal bases of the tangent bundle of W , they are related by an orthogonal transformation M_{xy} : $e_x = M_{xy} e_y$, or in more detail

⁴²In our applications, W is usually either the spacetime X of dimension D or its boundary $Y = \partial X$ of dimension $d = D - 1$.

$$e_{ax} = \sum_b M_a^b{}_{xy} e_{by}. \quad (\text{A8})$$

(Here $M_a^b{}_{xy}$ are the matrix elements of the $n \times n$ orthogonal matrix M_{xy} .) If W is orientable, as assumed in this section, we can pick the local vielbeins e_x such that M_{xy} is valued in $\text{SO}(n)$ [rather than $\text{O}(n)$]. The transition matrices M_{xy} obey a consistency condition: in triple overlaps $W_x \cap W_y \cap W_z$, we have

$$M_{xy} M_{yz} M_{zx} = 1. \quad (\text{A9})$$

Once a local vierbein is picked in an open set W_x , we can define a spinor field Ψ_x (with components $\Psi_{\alpha x}$, $\alpha = 1, \dots, 2^{\lfloor n/2 \rfloor}$) in this open set. The covariant derivatives and Dirac equation for Ψ_x are as in Eqs. (A1) and (A7). On intersections $W_x \cap W_y$, we compare Ψ_x and Ψ_y via

$$\Psi_x = \hat{M}_{xy} \Psi_y, \quad (\text{A10})$$

where \hat{M}_{xy} is a $2^{\lfloor n/2 \rfloor} \times 2^{\lfloor n/2 \rfloor}$ matrix that is obtained by “lifting” M_{xy} to the spinor representation. Thus \hat{M}_{xy} is an element of $\text{Spin}(n)$, the double cover of $\text{SO}(n)$ associated with spin. Because the sign with which a given element of $\text{SO}(n)$ acts on spinors is not uniquely determined, the lift from M_{xy} to \hat{M}_{xy} is only uniquely determined up to sign. For consistency, the signs must be chosen so that in triple overlaps

$$\hat{M}_{xy} \hat{M}_{yz} \hat{M}_{zx} = 1. \quad (\text{A11})$$

A choice of signs satisfying this condition is called a spin structure on W ; once a spin structure is picked, we can piece together the local descriptions and globally study fermions and the Dirac equation on W . In dimension ≥ 4 , a given W may not admit a spin structure because it may be impossible to satisfy Eq. (A11); in this case, spinors cannot be defined on W (an example is $W = \mathbb{C}\mathbb{P}^2$). This possibility does not play a primary role in the present paper. We simply restrict our attention to spacetimes on which spinors can be defined.

More important for our purposes is the fact that a given W may admit more than one spin structure. If we do find a choice of signs consistent with Eq. (A10), then a different choice $\hat{M}'_{xy} = \hat{M}_{xy} (-1)^{c_{xy}}$, for some \mathbb{Z}_2 -valued function c_{xy} , obeys the same condition if

$$(-1)^{c_{xy}} (-1)^{c_{yz}} (-1)^{c_{zx}} = 1 \quad (\text{A12})$$

for all \mathbf{x} , \mathbf{y} , and \mathbf{z} . We want to impose on c a gauge equivalence $c_{xy} \cong c_{xy} + d_x - d_y$ for any \mathbb{Z}_2 -valued function d_x , since in Eq. (A10) the change in \hat{M}_{xy} resulting from such a change in c_{xy} can be absorbed in redefining the local spinors by $\Psi_x \rightarrow (-1)^{d_x} \Psi_x$. Given the condition (A12) and the gauge invariance just stated, c_{xy} defines an element of the group $H^1(W, \mathbb{Z}_2)$. Any two spin structures on W differ by “twisting” by an element of this group.

For an elementary example of this, take W to be an n torus T^n , with flat metric. Parallel transport of tangent vectors on W is completely trivial: the Levi-Civita connection on tangent vectors is 0. However, in defining spinors on W , we are free to say that a spinor changes sign under parallel transport around a noncontractible loop $\ell \subset T^n$. As there are n independent loops to consider, there are 2^n choices of sign in defining spinors on T^n , and these are the spin structures on T^n . The group $H^1(T^n, \mathbb{Z}_2)$ is isomorphic to \mathbb{Z}_2^n and labels these choices of sign. [For $W = T^n$, but not in general, there is a distinguished spin structure, namely, the “trivial” one in which spinors are parallel transported with no sign changes. As a result, in this example, the set of spin structures can be canonically identified with the group $H^1(W, \mathbb{Z}_2)$.]

A manifold W endowed with a choice of spin structure is called a spin manifold. [Orientability is built into the notion of a spin manifold, since in the starting point we took the transition functions of the tangent bundle to be valued in $\text{SO}(n)$.] If W is a spin manifold, the matrices M_{xy} , understood as $2^{\lfloor n/2 \rfloor} \times 2^{\lfloor n/2 \rfloor}$ matrices in the spin representation of $\text{Spin}(n)$, are transition functions for a vector bundle $\mathcal{S} \rightarrow W$. A section of this bundle is a fermion field on W . We sometimes refer to \mathcal{S} as the spin structure. If \mathcal{S} is one spin structure, then any other spin structure is $\mathcal{S} \otimes \rho$, where $\rho \rightarrow W$ is a real line bundle of order 2 [defined in the above construction by the transition functions $(-1)^{c_{xy}}$].

For even dimension n , the irreducible representation of the Clifford algebra is unique up to isomorphism. This can be proven by writing n gamma matrices in terms of $n/2$ pairs of creation and annihilation operators, which have a unique irreducible representation. But for even n , the irreducible representation of the Clifford algebra is not irreducible as a representation of $\text{Spin}(n)$; it decomposes as the sum of two representations of $\text{Spin}(n)$ of positive or negative chirality. For odd n , matters are different. Up to isomorphism, there is only one spinor representation of $\text{Spin}(n)$, but there are two inequivalent representations of the Clifford algebra. The two representations differ by $\gamma_i \rightarrow -\gamma_i$, which preserves the anticommutation relations of the Clifford algebra and does not affect the $\text{Spin}(n)$ generators $\Sigma_{ij} = (1/4)[\gamma_i, \gamma_j]$, but gives an inequivalent representation of the Clifford algebra. The two representations differ by the sign of the product $\gamma_1 \gamma_2 \cdots \gamma_n$ —which commutes with the Clifford algebra and so is a c number (± 1 or $\pm i$, depending on n) in an irreducible representation of the Clifford algebra. For example, in three dimensions, the two representations of the Clifford algebra correspond in a locally Euclidean frame to $\gamma_i = \pm \sigma_i$, and thus the gamma matrices obey

$$\gamma_i \gamma_j = g_{ij} \pm i \epsilon_{ijk} \gamma^k, \quad (\text{A13})$$

where ϵ_{ijk} is the Levi-Civita tensor. In this formula, $\pm \epsilon_{ijk}$ can be regarded as a choice of orientation of W . Thus the two representations of the Clifford algebra are associated with the two possible orientations of W . The Dirac operator $\mathcal{D} = i \sum_k \gamma^k D_k$ is odd under $\gamma_k \rightarrow -\gamma_k$, and so the sign of the Dirac operator depends on the orientation of

W . The η invariant of the Dirac operator changes sign if one changes the sign of this operator (thereby changing the sign of all of its eigenvalues), and this is why the η invariant is odd under parity, that is under the reversal of orientation.

2. Fermions on an unorientable manifold

Now we consider the unorientable case. On an unorientable n manifold W , the transition matrices M_{xy} of the tangent bundle of W are valued in $O(n)$ rather than $SO(n)$, and the corresponding matrices \hat{M}_{xy} that act on spinors must similarly take values in a double cover of $O(n)$. There are two choices of this double cover,⁴³ depending on whether a spatial reflection \mathbf{R} acting on spinors satisfies $\mathbf{R}^2 = 1$ or $\mathbf{R}^2 = -1$. For most physical applications, the appropriate choice is $\mathbf{R}^2 = 1$. The double cover of $\text{Spin}(n)$ that arises if $\mathbf{R}^2 = 1$ is called $\text{Pin}^+(n)$. [If we choose $\mathbf{R}^2 = -1$, we get a double cover of $\text{Spin}(n)$ that is called $\text{Pin}^-(n)$.]

A reflection symmetry in Euclidean signature can be continued to either a reflection or a time-reversal symmetry in Lorentz signature, depending on whether the reflection acts on the spacetime coordinate that is analytically continued in the change of signature. This process is subtle and, as is well known, \mathbf{T} is antilinear in quantum field theory while \mathbf{R} is linear. Related to this, \mathbf{C} (charge conjugation) enters in the continuation, so \mathbf{R} in Euclidean signature can be continued to \mathbf{CT} in Lorentz signature, and \mathbf{CR} can be continued to \mathbf{T} . If \mathbf{R} (or \mathbf{CR}) squares to $+1$ on fermions, then \mathbf{CT} (or \mathbf{T}) squares to -1 , and vice versa.

To concretely describe the group $\text{Pin}^+(n)$, we work on \mathbb{R}^n and ask how a reflection that reverses the sign of one coordinate acts on spinors. Let w be a unit vector and consider the reflection $x^i \rightarrow x^i - 2w^i w \cdot x$. Setting $\gamma \cdot w = \gamma_i w^i$, the reflection acts on spinors by

$$\psi \rightarrow \gamma \cdot w \psi. \quad (\text{A14})$$

These transformations square to $+1$, and adjoining them to the $\text{Spin}(n)$ transformations generated by $\Sigma_{ab} = (1/4)[\gamma_a, \gamma_b]$ gives $\text{Pin}^+(n)$. One way to describe $\text{Pin}^-(n)$ is simply to take the given reflection to act by $\psi \rightarrow i\gamma \cdot w \psi$, an operation that squares to -1 .

To define a pin^+ structure on W , one lifts the transition functions M_{xy} of the tangent bundle from $O(n)$ to $\text{Pin}^+(n)$ -valued functions \hat{M}_{xy} . In dimension ≥ 2 , there is a possible obstruction to this (for example, $\mathbb{R}\mathbb{P}^2$ admits no pin^+ structure). If W does admit a pin^+ structure, this structure is not necessarily unique; just like spin structures, pin^+

⁴³Here we consider Majorana fermions, that is fermions coupled to gravity only, as is appropriate physically for a topological superconductor. In the presence of the $U(1)$ gauge symmetry of electromagnetism, one can combine a spatial reflection with a gauge transformation and define a symmetry for which \mathbf{R}^2 acts on electrons with an arbitrary phase. Mathematically, this is related to a possible generalization from spin and pin structures to spin_c and pin_c structures. This generalization is useful in analyzing topological states of matter, but is not relevant in the present paper.

structures differ by twisting by an element of $H^1(W, \mathbb{Z}_2)$. A manifold with a choice of pin^+ structure is called a pin^+ manifold.

If W is a pin^+ manifold, the transition functions \hat{M}_{xy} , understood as before as $2^{\lfloor n/2 \rfloor} \times 2^{\lfloor n/2 \rfloor}$ matrices, define a vector bundle $\mathcal{P} \rightarrow W$. We can define a spinor field ψ that is a section of this bundle and the Dirac operator $\mathcal{D} = i\mathcal{D}$ can act on ψ . But some of its properties are different from the orientable case, as seen in the main text and discussed further in Appendix A.3.

If $\mathcal{P} \rightarrow W$ is one pin^+ structure, then just as in the spin case, any other pin^+ structure is $\mathcal{P} \otimes \rho$ for some real line bundle $\rho \rightarrow W$. However, there is a new ingredient when W is unorientable. There is then a canonical real line bundle $\varepsilon \rightarrow W$, namely, the orientation bundle: the holonomy of ε around a closed loop $\ell \subset W$ is -1 if the orientation of W is reversed in going around ℓ , and otherwise it is $+1$. $\mathcal{P}' = \mathcal{P} \otimes \varepsilon$ is a new pin^+ structure on W that we call the pin^+ structure complementary to \mathcal{P} . Concretely, the transition functions \hat{M}'_{xy} of \mathcal{P}' are obtained from the transition functions \hat{M}_{xy} of \mathcal{P} by reversing the sign of \hat{M}_{xy} whenever it is in the orientation-reversing component of $\text{Pin}^+(n)$.

Everything that we have just said for pin^+ has an immediate analog for pin^- except that the obstruction to a pin^- structure begins in dimension 4.

3. Dirac operators

Consider a reflection in \mathbb{R}^n :

$$\mathbf{R}\psi(x_1, x_2, \dots, x_n) = \gamma_1 \psi(-x_1, x_2, \dots, x_n). \quad (\text{A15})$$

It is not difficult to see that such a reflection anticommutes with the Dirac operator

$$\mathcal{D} = i\mathcal{D} = i \sum_{j=1}^n \gamma^j D_j. \quad (\text{A16})$$

More generally, on a spin manifold \hat{W} , any orientation-reversing symmetry anticommutes with \mathcal{D} . Suppose that \hat{W} has an orientation-reversing symmetry τ that acts freely and obeys $\tau^2 = 1$. Then we can pass from \hat{W} to its quotient $W = \hat{W}/\mathbb{Z}_2$. If in addition τ acts as a symmetry of the spin bundle $\mathcal{S} \rightarrow \hat{W}$, still obeying $\tau^2 = 1$, then W is endowed with two pin^+ structures \mathcal{P} and \mathcal{P}' that can be defined as follows. A section ψ of $\mathcal{P} \rightarrow W$ is a section $\hat{\psi}$ of $\mathcal{S} \rightarrow \hat{W}$ that obeys

$$\tau \hat{\psi} = \hat{\psi}, \quad (\text{A17})$$

and a section ψ' of $\mathcal{P}' \rightarrow W$ is a section $\hat{\psi}'$ of $\mathcal{S} \rightarrow \hat{W}$ that obeys

$$\tau \hat{\psi}' = -\hat{\psi}'. \quad (\text{A18})$$

Since the sign with which τ acts on \mathcal{S} depends on an arbitrary choice, the relationship between \mathcal{P} and \mathcal{P}' is completely symmetrical.

Consider any point $p \in \hat{W}$ and a path $\hat{\ell}$ from p to $\tau(p)$. After dividing by τ , $\hat{\ell}$ projects to a closed loop $\ell \subset W$ around which the orientation of W is reversed. The relative minus sign in Eqs. (A17) and (A18) means that the monodromies around ℓ of the pin^+ bundles \mathcal{P} and \mathcal{P}' differ in sign. Thus \mathcal{P} and \mathcal{P}' are complementary pin^+ bundles in the sense of Appendix A.2.

Every pin^+ manifold W arises in this construction, with \hat{W} being the oriented double cover of W .

Now suppose that ψ is a fermion field on W valued in a particular pin^+ structure \mathcal{P} . From what we have said, it is clear that the usual Dirac operator $\mathcal{D} = i\sum_k \gamma^k D_k$ cannot be defined as a self-adjoint operator acting on ψ . On the covering space \hat{W} , \mathcal{D} anticommutes with τ , so it maps sections of \mathcal{P} to sections of \mathcal{P}' and vice versa.

For odd dimension n , there is no way to remedy this situation and there is no self-adjoint Dirac operator acting on a \mathcal{P} -valued field ψ . The upshot of this is that the corresponding path integral Z_ψ can be defined as the Pfaffian of a complex-valued antisymmetric bilinear form, but not as the determinant of a self-adjoint operator. Accordingly it is complex valued, and in general it can be affected by a global anomaly involving a complex number of modulus 1, not just a real number ± 1 . This issue was explored in the main text in Sec. IV.B.

For even n , matters are different and there is always a self-adjoint Dirac operator acting on a fermion field valued in any given pin^+ structure \mathcal{P} . This was assumed in the main text at several points and here we explain the details. First we start on the orientable manifold \hat{W} with spin structure \mathcal{S} . As \hat{W} is orientable, one can define a chirality operator $\bar{\gamma}$ that acts on \mathcal{S} (and leads to a decomposition $\mathcal{S} = \mathcal{S}_+ \oplus \mathcal{S}_-$, where \mathcal{S}_+ and \mathcal{S}_- are the bundles of positive or negative chirality spinors). Assuming that we want $\bar{\gamma}^2 = 1$, the proper definition of $\bar{\gamma}$ depends slightly on whether n is congruent to 0 or 2 mod 4. For example, in two dimensions (and similarly in $4k + 2$ dimensions for any k), one defines

$$\bar{\gamma} = \frac{i}{2!} \epsilon^{ij} \gamma_i \gamma_j. \quad (\text{A19})$$

In four dimensions (and similarly in $4k$ dimensions for any k), one omits the factor of i and defines

$$\bar{\gamma} = \frac{1}{4!} \epsilon^{ijkl} \gamma_i \gamma_j \gamma_k \gamma_l. \quad (\text{A20})$$

On an orientable manifold \hat{W} of even dimension, instead of the usual Dirac operator $\mathcal{D} = i\mathcal{D}$, we can use the alternate Dirac operator

$$\mathcal{D}' = \bar{\gamma}\mathcal{D}. \quad (\text{A21})$$

\mathcal{D}' is self-adjoint, like \mathcal{D} , and the two operators have the same spectrum, since they are conjugate. Indeed, if $U = (1 - i\bar{\gamma})/\sqrt{2}$, then

$$\mathcal{D}' = U\mathcal{D}U^{-1}. \quad (\text{A22})$$

This depends on the fact that $\bar{\gamma}$ commutes with the action of $\text{Spin}(n)$ on the spinor representation, and thus commutes with the $\text{Spin}(n)$ connection that is hidden in the definition of \mathcal{D} .

So on \hat{W} , we could equally well use \mathcal{D} or \mathcal{D}' . However, if we want to descend to the unorientable manifold $W = \hat{W}/\mathbb{Z}_2$, where \mathbb{Z}_2 is generated by the orientation-reversing isometry τ , then it is better to use \mathcal{D}' . Indeed, since $\bar{\gamma}$ and \mathcal{D} are both odd under τ , \mathcal{D}' is even, and it descends to a perfectly good self-adjoint operator acting on the pin^+ bundle \mathcal{P} of W .

In the main text, we sometimes write simply \mathcal{D} rather than \mathcal{D}' for the self-adjoint Dirac operator of an even-dimensional Pin^+ manifold W . A justification for this is that in fact up on \hat{W} what we mean by \mathcal{D} or \mathcal{D}' depends on the choice of representation of the Clifford algebra. We can define new gamma matrices

$$\gamma_i' = U\gamma_i U^{-1}. \quad (\text{A23})$$

Concretely

$$\gamma_i' = -i\bar{\gamma}\gamma_i = \begin{cases} -\epsilon_{ij}\gamma^j & \text{if } n = 2, \\ \frac{i}{3!}\epsilon_{ijkl}\gamma^j\gamma^k\gamma^l & \text{if } n = 4. \end{cases} \quad (\text{A24})$$

In terms of the new gamma matrices, \mathcal{D}' takes the standard form

$$\mathcal{D}' = i\sum_k \gamma'^k D_k. \quad (\text{A25})$$

(The covariant derivatives D_k take the same form in old or new gamma matrices.)

However, if we use the new gamma matrices, then the transformation law (A14) under a reflection [and therefore under any orientation-reversing element of $\text{Pin}^+(n)$] is modified. It becomes

$$\psi \rightarrow i\bar{\gamma}'\gamma' \cdot w\psi. \quad (\text{A26})$$

(Here $\bar{\gamma}'$ is the chirality operator constructed from the new gamma matrices; it coincides with $\bar{\gamma}$.) Actually, this is the most common choice in particle physics in four spacetime dimensions.⁴⁴ The reason that we started with Eq. (A14) instead, apart from the fact that it is fairly standard mathematically, is that it leads for many purposes to a formalism that works uniformly in all dimensions. It is hard to find conventions for fermions that are convenient for all purposes.

APPENDIX B: EXAMPLES IN LOW DIMENSION

We will look more closely at low-dimensional examples that are important in this paper.

1. Two dimensions

To describe spinors on \mathbb{R}^2 , we only need two gamma matrices, and we can pick real 2×2 gamma matrices $\gamma_1 = \sigma_1$, $\gamma_2 = \sigma_3$. The group $\text{Spin}(2)$ is generated by $\Sigma = (1/2)\gamma_1\gamma_2$.

⁴⁴For example, see Eq. (2.33) of Bjorken and Drell (1964), which however is written with an arbitrary phase for Dirac fermions and also is a formula for the transformation of a fermion under parity (a reflection of all three spatial coordinates, not just one).

Abstractly, $\text{Spin}(2)$ is a copy of $\text{SO}(2)$, with the spinors as the fundamental real two-dimensional representation. But, because the eigenvalues of Σ are $\pm i/2$ (not $\pm i$), $\text{Spin}(2)$ is a double cover of the original $\text{SO}(2)$ that acts by rotations on \mathbb{R}^2 .

To get $\text{Pin}^+(2)$, we include the reflections $x^i \rightarrow x^i - 2w^i w \cdot x$, acting by

$$\psi \rightarrow \gamma \cdot w \psi. \quad (\text{B1})$$

This gives a real, two-dimensional representation of $\text{Pin}^+(2)$. The group $\text{Pin}^+(2)$ is abstractly isomorphic to $\text{O}(2)$, although it is a double cover of the usual $\text{O}(2)$ that acts by rotation and reflection of \mathbb{R}^2 .

Naively, one might think that one can define a second real, two-dimensional representation of $\text{Pin}^+(2)$ in which a reflection acts by

$$\psi \rightarrow -\gamma \cdot w \psi. \quad (\text{B2})$$

Shortly we will see that something like this does work in three dimensions, but in two dimensions the two representations are conjugate under $\psi \rightarrow \gamma_1 \gamma_2 \psi$, which represents a π rotation of \mathbb{R}^2 .

Since the spinor representation of $\text{Pin}^+(2)$ is real, there is an antilinear operation

$$\mathcal{T}\psi = \psi^*, \quad \mathcal{T}^2 = 1 \quad (\text{B3})$$

that anticommutes with the Hermitian Dirac operator \mathcal{D}' and ensures that on any two-manifold, orientable or not, the spectrum is symmetric under $\lambda \leftrightarrow -\lambda$.

On an *orientable* two-manifold, we can also define an antilinear operation

$$\mathcal{T}'\psi = \bar{\gamma}\psi^*, \quad (\mathcal{T}')^2 = -1 \quad (\text{B4})$$

that *commutes* with \mathcal{D} (or \mathcal{D}') and ensures that the Dirac eigenvalues all have even multiplicity. Here $\bar{\gamma} = i\gamma_1 \gamma_2 = \sigma_2$ is the chirality operator.

The reason that both structures exist is that, if we restrict to $\text{Spin}(2)$, then the two-dimensional spinor representation is reducible; it decomposes in one-dimensional representations of positive or negative chirality which are complex conjugates to each other. Accordingly, the Dirac operator anticommutes with $\bar{\gamma}$ and we can include it in defining \mathcal{T}' . On an unorientable two-manifold, the definition of \mathcal{T} makes sense, but \mathcal{T}' can no longer be defined, because $\bar{\gamma}$ is not defined globally.

2. Three dimensions

Now we go to three dimensions. For the 2×2 gamma matrices, we can take the Pauli σ matrices

$$\gamma_a = \sigma_a, \quad a = 1, \dots, 3. \quad (\text{B5})$$

The double cover of $\text{SO}(3)$ that acts on spinors is $\text{Spin}(3) = \text{SU}(2)$. Spinors transform in the spin $1/2$ representation of $\text{SU}(2)$. This representation is pseudoreal, since the trivial representation of $\text{SU}(2)$ appears *antisymmetrically* in the tensor product $1/2 \otimes 1/2$. Concretely, the invariant antisymmetric tensor on the spin $1/2$ representation of $\text{SU}(2)$ is given by the 2×2 Levi-Civita symbol $\varepsilon_{\alpha\beta}$, $\alpha,$

$\beta = 1, 2$. If ψ is a fermion field valued in this representation, then the corresponding ‘‘mass term,’’ often written as $\bar{\psi}\psi$, is

$$(\psi, \psi) = \varepsilon_{\alpha\beta} \psi^\alpha \psi^\beta. \quad (\text{B6})$$

The group $\text{Pin}^+(3)$ is obtained by adjoining to $\text{Spin}(3) = \text{SU}(2)$ the reflection matrices $\gamma \cdot w$ for w a unit vector in \mathbb{R}^3 . Such a reflection matrix is a 2×2 unitary matrix of determinant -1 . So $\text{Pin}^+(3)$ is the subgroup of $\text{U}(2)$ consisting of matrices of determinant ± 1 :

$$\text{Pin}^+(3) = \{g \in \text{U}(2) \mid \det g = \pm 1\}. \quad (\text{B7})$$

In contrast to $\text{Spin}(3)$, $\text{Pin}^+(3)$ has two inequivalent two-dimensional irreducible representations R and R' . If ψ and ψ' are valued, respectively, in R and R' , then $g \in \text{Pin}^+(3)$ acts on them by

$$\psi \rightarrow g\psi, \quad \psi' \rightarrow g(\det g)\psi'. \quad (\text{B8})$$

Thus the difference is precisely that a spatial reflection (or any orientation-reversing transformation) acts on ψ and ψ' with opposite signs. The representations R and R' are both complex representations, meaning that they do not admit an invariant bilinear form. Indeed, the mass terms $(\psi, \psi) = \varepsilon_{\alpha\beta} \psi^\alpha \psi^\beta$ and $(\psi', \psi') = \varepsilon_{\alpha\beta} \psi'^\alpha \psi'^\beta$ are both odd under reflection. Under $\psi \rightarrow g\psi$, $\psi' \rightarrow g \det g \psi'$, we have

$$(\psi, \psi) \rightarrow \det g (\psi, \psi), \quad (\psi', \psi') \rightarrow \det g (\psi', \psi'). \quad (\text{B9})$$

But the ‘‘off-diagonal’’ mass term

$$(\psi, \psi') = \varepsilon_{\alpha\beta} \psi^\alpha \psi'^\beta \quad (\text{B10})$$

that links two fermions ψ and ψ' that transform oppositely under \mathbf{R} (or \mathbf{T}) is invariant.

Since there is always an invariant pairing between a representation of a compact group such as $\text{Pin}^+(3)$ and its complex conjugate, the invariance of the pairing (ψ, ψ') means that the representation R is isomorphic to the complex conjugate of R' . Indeed, one can show that the complex conjugates of the matrices g acting on representation R can be conjugated by a unitary transformation to the corresponding matrices $g \det g$ acting on representation R' :

$$\sigma_y g^* \sigma_y^{-1} = g \det g. \quad (\text{B11})$$

To show this, one can consider separately the cases that $g = \gamma \cdot w$ is a reflection, and that $g \in \text{Spin}(3)$, with $\det g = 1$. Equation (B11) is equivalent to $\mathcal{T}g\mathcal{T}^{-1} = g \det g$, where \mathcal{T} is the antilinear operation that we introduce momentarily.

On an orientable three-manifold, the spinor representation of $\text{Spin}(3) \cong \text{SU}(2)$ admits the invariant antisymmetric tensor $\varepsilon_{\alpha\beta}$, as discussed previously. Hence one can define an antilinear operation

$$(\mathcal{T}\psi)^\alpha = \varepsilon^{\alpha\beta} \psi_\beta^*, \quad (\text{B12})$$

where ψ_β^* is the complex conjugate of ψ^β . It obeys $\mathcal{T}^2 = -1$ and commutes with the Dirac operator \mathcal{D} . Accordingly, there

is a version of Kramers doubling: the eigenvalues of \mathcal{D} all have even multiplicity.

We have given a rather intrinsic definition of \mathcal{T} , but some may also want to see a description in local coordinates. After picking a vielbein and “flat space” gamma matrices $\gamma_a = \sigma_a$, and writing \mathcal{K} for complex conjugation, one can define

$$\mathcal{T} = \mathcal{K}\sigma_y, \quad \mathcal{T}^2 = -1. \quad (\text{B13})$$

This operation can be defined using any vielbein, and because it is actually $\text{SU}(2)$ invariant, despite not being written in a way that makes this manifest, the definition of \mathcal{T} does not depend on the choice of vielbein, as long as we consider only vielbeins that determine the same orientation on spacetime (i.e., as long as we allow only orientation-preserving transformations between vielbeins).

A special case of the fact that the spectrum of the Dirac operator has even multiplicity on an orientable three-manifold is that the mod 2 index of this operator vanishes. However, \mathcal{T} anticommutes with the gamma matrices and hence with the reflection elements of $\text{Pin}^+(3)$. As a result, \mathcal{T} cannot be defined as a transformation of spinor fields on an unorientable three-manifold W . Going up to the oriented double cover \hat{W} , with $W = \hat{W}/\mathbb{Z}_2$, \mathcal{T} can be defined but anticommutes with the orientation-reversing generator τ of \mathbb{Z}_2 . So if $\hat{\psi}$ is a spinor field on \hat{W} and $\tau\hat{\psi} = \pm\hat{\psi}$, then $\tau(\mathcal{T}\hat{\psi}) = \mp\mathcal{T}\hat{\psi}$. Recalling the definition of the complementary pin^+ structures \mathcal{P} and \mathcal{P}' , we can interpret this statement on W . It means that although \mathcal{T} cannot be defined to map sections of \mathcal{P} to themselves or sections of \mathcal{P}' to themselves, it makes sense as a map from sections of \mathcal{P} to sections of \mathcal{P}' and vice versa. This means, in particular, that \mathcal{P} and \mathcal{P}' have the same mod 2 index (not necessarily 0, as we know from Sec. III.B). In fact, \mathcal{T} transforms a zero mode of the Dirac operator $\mathcal{D}:\mathcal{P} \rightarrow \mathcal{P}'$ to a zero mode of the adjoint Dirac operator $\mathcal{D}:\mathcal{P}' \rightarrow \mathcal{P}$.

3. Four dimensions

In four dimensions, we need four gamma matrices. We can take them to be 4×4 matrices, but it is not possible for them to be all real. For example, we can take

$$\begin{aligned} \gamma_1 &= \sigma_1 \otimes 1, & \gamma_2 &= \sigma_2 \otimes 1, \\ \gamma_3 &= \sigma_3 \otimes \sigma_1, & \gamma_4 &= \sigma_3 \otimes \sigma_3. \end{aligned} \quad (\text{B14})$$

With this choice, γ_2 is imaginary and the others are real.

The group $\text{Spin}(4)$ is isomorphic to the product of two copies of $\text{SU}(2)$, which we write as $\text{SU}(2)_\ell \times \text{SU}(2)_r$. The action of $\text{Spin}(4)$ commutes with the chirality operator $\bar{\gamma} = \gamma_1\gamma_2\gamma_3\gamma_4$, and accordingly the four-dimensional spinor representation of $\text{Spin}(4)$ splits as the direct sum of two 2-dimensional representations with $\bar{\gamma} = \pm 1$. These are the representations $(1/2, 0)$ and $(0, 1/2)$ of $\text{SU}(2)_\ell \times \text{SU}(2)_r$.

Those representations are both pseudoreal, simply because the two-dimensional representation of $\text{SU}(2)$ is pseudoreal. As a result spinors of $\text{Spin}(4)$ are pseudoreal. If we denote a field valued in the $(1/2, 0)$ representation as $\psi^{\alpha'}$ and one valued in the $(0, 1/2)$ representation as $\psi^{\alpha''}$, then there are invariant antisymmetric tensors $\varepsilon_{\alpha'\beta'}$ and $\varepsilon_{\alpha''\beta''}$ for the $(1/2, 0)$ and

$(0, 1/2)$ representations, respectively. Each of these is uniquely determined by $\text{SU}(2)_\ell$ or $\text{SU}(2)_r$ symmetry up to a constant multiple.

To pass from $\text{Spin}(4)$ to $\text{Pin}^+(4)$, we add reflection symmetries $x^i \rightarrow x^i - 2w^i(w \cdot x)$. A reflection acts as usual by

$$\psi \rightarrow \gamma \cdot w\psi. \quad (\text{B15})$$

Such a reflection exchanges $\text{SU}(2)_\ell$ with $\text{SU}(2)_r$, anticommutes with $\bar{\gamma}$, and exchanges $(1/2, 0)$ with $(0, 1/2)$. So the representation $(1/2, 0) \oplus (0, 1/2)$ of $\text{Spin}(4)$ is irreducible as a representation of $\text{Pin}^+(4)$. There is only one such spinor representation of $\text{Pin}^+(4)$, up to isomorphism, since a sign in Eq. (B15) could be removed by conjugation with $\bar{\gamma}$.

$\text{Spin}(4)$ does not determine a relative normalization between $\varepsilon_{\alpha'\beta'}$ and $\varepsilon_{\alpha''\beta''}$. However, a reflection that exchanges $\text{SU}(2)_\ell$ with $\text{SU}(2)_r$ maps a particular choice of $\varepsilon_{\alpha'\beta'}$ to a choice of $\varepsilon_{\alpha''\beta''}$. With this choice, the sum $\varepsilon_{\alpha'\beta'} \oplus \varepsilon_{\alpha''\beta''}$ is an invariant antisymmetric bilinear form $\varepsilon_{\alpha\beta}$ on the four-dimensional spinor representation of $\text{Pin}^+(4)$. (Indices α and β can now be of either type α', β' or α'', β'' .)

This enables us to define an antilinear operation on spinors on a pin^+ four-manifold W :

$$(\mathcal{T}\psi)^\alpha = \varepsilon^{\alpha\beta}\psi_\beta^*. \quad (\text{B16})$$

Since $\mathcal{T}^2 = -1$, the eigenvalues of the self-adjoint Dirac operator \mathcal{D} on such a manifold have even multiplicity.

One might prefer to see \mathcal{T} defined explicitly using a local vielbein. Using the representation (B14) for the flat space gamma matrices, the definition is

$$\mathcal{T} = \mathcal{K}\gamma_2\bar{\gamma}. \quad (\text{B17})$$

This commutes with $\text{Pin}^+(4)$, so this definition does not really depend on the choice of the vielbein and makes sense globally.

4. Five dimensions

Naively, in this paper, we do not need to know about spinors and the Dirac operator in five dimensions. But in Appendix C, we want to know what the APS index theorem says about the four-dimensional η invariant, and for this one needs to know some facts about five dimensions.

In five dimensions, one can find a 4×4 representation of the Clifford algebra. One uses the four gamma matrices in Eq. (B14) along with $\gamma_5 = \sigma_3 \otimes \sigma_2$. Note that of the five gamma matrices, two are imaginary (γ_2 and γ_5) and the others are real.

Thus the group $\text{Spin}(5)$ has a four-dimensional spinor representation.⁴⁵ The group $\text{Spin}(5)$ is actually isomorphic to $\text{Sp}(4)$, and its four-dimensional spinor representation is simply the fundamental four-dimensional representation of $\text{Sp}(4)$. $\text{Sp}(4)$ is defined as the subgroup of $\text{U}(4)$ that preserves a certain nondegenerate antisymmetric bilinear form, and, in

⁴⁵This representation is pseudoreal in Lorentz signature, so one needs to take two copies of it to make a sensible theory of fermions. This is not really relevant for our limited purposes here.

particular, the four-dimensional representation of $\text{Sp}(4)$ is pseudoreal and spinors of $\text{Spin}(5)$ are pseudoreal.

Concretely, on spinors of $\text{Spin}(5)$, we can define an antilinear operation

$$\mathcal{T} = \mathcal{K}\gamma_2\gamma_5, \quad \mathcal{T}^2 = -1. \quad (\text{B18})$$

Thus on a five-dimensional spin manifold (orientable for the time being), there is a form of Kramers doubling, and all eigenvalues of the Dirac operator have even multiplicity.

So far all this parallels what happens in three dimensions. Now let us go to the unorientable case. Reflections are included in the usual way. The reflection $x^i \rightarrow x^i - 2w^i(w \cdot x)$ acts by

$$\psi \rightarrow \pm(\gamma \cdot w)\psi. \quad (\text{B19})$$

As in three dimensions, the sign is meaningful and gives two distinct four-dimensional representations R and R' of $\text{Pin}^+(5)$. However, there is an important difference. In three dimensions, the representations R and R' are complex representations that are complex conjugates of each other, but in five dimensions, they are each pseudoreal. To see this, we observe that the antilinear operation \mathcal{T} defined in Eq. (B18) commutes with a reflection acting as in Eq. (B19). So \mathcal{T} can be defined separately in R and in R' , and these representations are both pseudoreal. At an elementary level, the difference from three dimensions is that in five dimensions the number of gamma matrices that are imaginary is even.

As in any odd dimension, there is no self-adjoint Dirac operator acting on a field valued in a particular pin^+ structure; rather pin^+ structures come in complementary pairs $\mathcal{P}, \mathcal{P}'$ that are exchanged by the Dirac operator. We can define the *index* of the Dirac operator \mathcal{I} as the number of zero modes of $\mathcal{D}: \mathcal{P} \rightarrow \mathcal{P}'$ minus the number of zero modes of the adjoint operator $\mathcal{D}: \mathcal{P}' \rightarrow \mathcal{P}$. This index is 0 on a five-manifold without boundary, but on a manifold with boundary (and with APS boundary conditions), it can be nonzero, as we learn in Appendix C. However, \mathcal{I} is always even because the antilinear symmetry \mathcal{T} ensures that the number of zero modes of \mathcal{D} acting on either \mathcal{P} or \mathcal{P}' is even.

APPENDIX C: THE η INVARIANT IN FOUR DIMENSIONS

In Secs. IV.E and IV.F, we needed some facts about the η invariant in $D = 4$. Analogous facts in $D = 2$ were invoked in Sec. V. Our goal in this Appendix is to briefly explain these facts. [The arguments we give are similar to the original ones (Gilkey, 1985).]

Our first goal is to show that in $D = 4$, η is always an integer multiple of $1/4$, so that $\exp(-i\pi\eta/2)$ is always a 16th root of 1. The nontrivial case is that X is unorientable; otherwise η is a multiple of 2, as noted in Sec. IV.C. Suppose that X has pin^+ structure \mathcal{P} and complementary pin^+ structure $\mathcal{P}' = \mathcal{P} \otimes \varepsilon$, where ε is a real line bundle, the orientation bundle of X . We can define η invariants $\eta_{\mathcal{P}}, \eta_{\mathcal{P}'}$ for a Majorana fermion valued in either of these pin^+ structures.

The first useful relation is that

$$\eta_{\mathcal{P}} + \eta_{\mathcal{P}'} = 0 \quad \text{mod } 4. \quad (\text{C1})$$

The point is that a Majorana fermion on X coupled to \mathcal{P} together with a Majorana fermion on X coupled to \mathcal{P}' combine together to a Majorana fermion coupled to a spin structure on \hat{X} , the oriented double cover of X . So $\eta_{\mathcal{P}} + \eta_{\mathcal{P}'} = \hat{\eta}$, where $\hat{\eta}$ is the η invariant of the Majorana fermion on \hat{X} . Because \hat{X} is orientable and also has an orientation-reversing symmetry (the quotient by this symmetry being X), $\hat{\eta}$ is a multiple of 4, as noted in Sec. IV.C.

We need one more relation, which is that

$$8\eta_{\mathcal{P}'} = 8\eta_{\mathcal{P}} \quad \text{mod } 4. \quad (\text{C2})$$

Together, these relations imply that $\eta_{\mathcal{P}}$ and $\eta_{\mathcal{P}'}$ are both multiples of $1/4$.

To explain Eq. (C2), it is helpful to generalize the problem. Let $V \rightarrow X$ be any real vector bundle. We consider the Dirac operator \mathcal{D}_V acting on fermions valued in V , that is, acting on sections of $\mathcal{P} \otimes V$. For any V , we can define $\eta_{\mathcal{P} \otimes V}$, the η invariant of the Dirac operator acting on such a field. Here $\eta_{\mathcal{P} \otimes V}$, for any V , is a topological invariant (independent of the metric of X and the connection on V) mod 4. This follows from the APS index theorem, which in odd dimensions takes the simple form (2.47). To apply the theorem without assuming that X is a boundary, we consider the five-manifold $Z = X \times [0, 1]$ with one metric and gauge field (g, A) at one end and some other (g', A') at the other end. The theorem then says that the difference of $-\eta/2$ between the two ends is equal to \mathcal{I} , the index of the Dirac operator on Z with APS boundary conditions. This index is always even, as explained in Sec. B.4, and the difference between the two values of η is a multiple of 4. More directly, what is happening is that as we vary (g, A) on a manifold of even dimensions, η is constant except that it jumps by 2 when an eigenvalue passes through 0. In four dimensions, the eigenvalues have even multiplicity because of a version of Kramers doubling, so the jumps are multiples of 4 and η is a topological invariant mod 4.

We can interpret $8\eta_{\mathcal{P}'}$ as $\eta_{\mathcal{P} \otimes V}$ where $V = \varepsilon^{\otimes 8}$ is the direct sum of eight copies of the real line bundle ε . Similarly $8\eta_{\mathcal{P}} = \eta_{\mathcal{P} \otimes \mathbb{R}^8}$, where \mathbb{R}^8 is a trivial real vector bundle of rank 8. So to establish Eq. (C2), it suffices to prove that $\varepsilon^{\otimes 8}$ is trivial, that is isomorphic to \mathbb{R}^8 .

To show that a real vector bundle V over a four-manifold X is trivial, it suffices to show vanishing of Stiefel-Whitney classes $w_i(V)$, $1 \leq i \leq 4$, and also of a certain class $\lambda \in H^4(X, \mathbb{Z})$ that is subtle to define, but that obeys $2\lambda = p_1(V)$, where p_1 is the first Pontryagin class.

For the first point, the total Stiefel-Whitney class of V is $w(V) = w(\varepsilon^{\otimes 8}) = w(\varepsilon)^8 = [1 + w_1(\varepsilon)]^8 = 1$, where in the last step we use the fact that the relevant binomial coefficients vanish mod 2. So $w_i(V) = 0$ for $i \geq 1$.

To show vanishing of $\lambda(V)$, we first write $V = V_0 \oplus V_0$, with $V_0 = \varepsilon^{\otimes 4}$. So $\lambda(V) = 2\lambda(V_0) = p_1(V_0)$. We are reduced to showing that $p_1(V_0) = 0$. To compute $p_1(V_0)$, we observe that the complexification of V_0 is the direct sum of four copies

of $\varepsilon_{\mathbb{C}} = \varepsilon \otimes_{\mathbb{R}} \mathbb{C}$. In general, if $V_0 \rightarrow X$ is a real vector bundle whose complexification is $\bigoplus_i \mathcal{L}_i$ for some complex line bundles $\mathcal{L}_i \rightarrow X$, then $p_1(V_0) = \sum_i c_1(\mathcal{L}_i)^2$. In the present case, this formula gives $p_1(V_0) = 4c_1(\varepsilon_{\mathbb{C}})^2$, and this vanishes because $c_1(\varepsilon_{\mathbb{C}})$ is 2-torsion. The last statement is a general one that holds for $\varepsilon_{\mathbb{C}} = \varepsilon \otimes_{\mathbb{R}} \mathbb{C}$ for any real line bundle ε over a topological space.

This completes the proof that $\exp(-i\pi\eta/2)$ is a 16th root of 1.

We also want to show that for $X = \mathbb{R}\mathbb{P}^4$, $\exp(-\pi i\eta/2) = \exp(\pm 2i\pi/16)$, with the sign depending on the choice of Pin^+ structure. We start on the five-torus T^5 parametrized by variables x_1, \dots, x_5 of period 1. We give T^5 a spin structure in which spinors are periodic in all directions. The spin bundle S of T^5 has rank 4, since the spinor representation of $\text{SO}(5)$ is four dimensional.

The Dirac operator $\mathcal{D} = i\partial$ on T^5 is self-adjoint, and its index certainly vanishes. Now we replace T^5 by T^5/\mathbb{Z}_2 , where \mathbb{Z}_2 acts by $\mathbf{P}: x_i \rightarrow -x_i$, $i = 1, \dots, 5$. Here T^5/\mathbb{Z}_2 is not a manifold, and we correct for that in a moment. We define \mathcal{H}_+ and \mathcal{H}_- to be the spaces of spinor fields on T^5 that obey, respectively, $\psi(-x) = \psi(x)$ and $\psi(-x) = -\psi(x)$. Because \mathcal{D} anticommutes with \mathbf{P} , it maps \mathcal{H}_+ to \mathcal{H}_- , and vice versa. We define the *index* \mathcal{I} of \mathcal{D} on T^5/\mathbb{Z}_2 to be the number of zero modes of \mathcal{D} acting on \mathcal{H}_+ minus the number of zero modes of \mathcal{D} acting on \mathcal{H}_- . (These operators are adjoints of each other since the full Dirac operator acting on $\mathcal{H}_+ \oplus \mathcal{H}_-$ is self-adjoint.)

It is straightforward to compute \mathcal{I} . The zero modes of \mathcal{D} acting on \mathcal{H}_+ are the four-dimensional space of constant spinors, and there are no zero modes of \mathcal{D} acting on \mathcal{H}_- . So the index is $\mathcal{I} = 4$.

Now T^5/\mathbb{Z}_2 is not a manifold because \mathbf{P} has 32 fixed points: the points with coordinates $x_i \in \mathbb{Z}/2$, $i = 1, \dots, 5$. To get a manifold, we remove from T^5 a small ball around each fixed point to get a manifold T' with boundary. The quotient of T' by \mathbf{P} is then a manifold X with a boundary that consists of 32 copies of $\mathbb{R}\mathbb{P}^4$. X acquires two pin^+ structures $\mathcal{P}, \mathcal{P}'$ whose sections are spinor fields on T' that are even or odd under \mathbf{P} . The Dirac operator \mathcal{D} maps sections of \mathcal{P} to sections of \mathcal{P}' , and vice versa. Hence it has an index \mathcal{I} . A fairly elementary argument⁴⁶ using the conformal invariance of the Dirac

⁴⁶Near one of the fixed points of the action of \mathbf{P} on T^5 , we can replace T^5/\mathbb{Z}_2 by $\mathbb{R}^5/\mathbb{Z}_2$ with Euclidean metric $d\ell^2 = d\vec{x}^2$. A Weyl rescaling converts this to $d\hat{\ell}^2 = d\vec{x}^2 \cdot (1 + 1/|\vec{x}|^2)$. Near $\vec{x} = 0$, this has the tubelike behavior that we described in introducing APS boundary conditions (see Fig. 6). In this case, the tube is a copy of $\mathbb{R}\mathbb{P}^4 \times \mathbb{R}_+$. Because of the conformal invariance of the Dirac equation, a \mathbb{Z}_2 -invariant solution of the Dirac equation on \mathbb{R}^5 that is regular at $\vec{x} = 0$ is equivalent to a solution that is square integrable in the tubelike geometry. So we can equally well compute the index \mathcal{I} by counting regular solutions in the original T^5 geometry or by counting square-integrable solutions after making a Weyl transformation to replace the 32 fixed points by tubes. But the definition of APS boundary conditions ensures that the latter procedure is equivalent to the index computed on X with APS boundary conditions.

operator shows that the Dirac index on X , with APS boundary conditions, has the same value $\mathcal{I} = 4$ found in the previous paragraph on T^5/\mathbb{Z}_2 . In odd dimensions, as in Eq. (2.47) the APS index theorem says that $\eta/2 = -\mathcal{I}$, where \mathcal{I} is the Dirac index on a manifold T' and η is the η invariant of its boundary. In the present case, the boundary consists of 32 copies of $\mathbb{R}\mathbb{P}^4$ (all with the same pin^+ structure, since they are permuted by obvious symmetries). So the formula becomes $16\eta_{\mathbb{R}\mathbb{P}^4} = -\mathcal{I} = -4$, or $\eta_{\mathbb{R}\mathbb{P}^4} = -1/4$.

We leave it to the interested reader to show in a similar way that in two dimensions $\exp(-i\pi\eta/2)$ is an eighth root of 1, and $\eta_{\mathbb{R}\mathbb{P}^2} = \pm 1/2$.

REFERENCES

- Adler, S., 1969, "Axial Vector Vertex In Spinor Electrodynamics," *Phys. Rev.* **177**, 2426.
- Alvarez-Gaumé, L., S. Della Pietra, and G. Moore, 1985, "Anomalies And Odd Dimensions," *Ann. Phys. (N.Y.)* **163**, 288.
- Atiyah, M. F., and I. M. Singer, 1963, "The Index Of Elliptic Operators On Compact Manifolds," *Bull. Am. Math. Soc.* **69**, 422.
- Atiyah, M. F., and I. M. Singer, 1968, "The Index Of Elliptic Operators: I," *Ann. Math.* **87**, 484.
- Atiyah, M. F., and I. M. Singer, 1971, "The Index Of Elliptic Operators: V," *Ann. Math.* **93**, 139.
- Atiyah, M. F., V. K. Patodi, and I. M. Singer, 1975, "Spectral Asymmetry and Riemannian Geometry. I," *Math. Proc. Cambridge Philos. Soc.* **77**, 43.
- Bell, J. S., and R. Jackiw, 1969, "A PCAC Puzzle: $\pi^0 \rightarrow \gamma\gamma$ In The Sigma Model," *Nuovo Cimento A* **60**, 47.
- Berry, M., 1984, "Quantal Phase Factors Accompanying Adiabatic Changes," *Proc. R. Soc. A* **392**, 45.
- Bjorken, J. D., and S. D. Drell, 1964, *Relativistic Quantum Mechanics* (McGraw-Hill, New York).
- Bunke, U., 2009, "String Structures And Trivialisations Of A Pfaffian Line Bundle," *arXiv:0909.0846*.
- Callan, Jr., C. G., R. Dashen, and D. J. Gross, 1978, "Towards A Theory Of The Strong Interactions," *Phys. Rev. D* **17**, 2717.
- Callan, Jr., C. G., and J. A. Harvey, 1985, "Anomalies And Fermion Zero-Modes On Strings And Domain Walls," *Nucl. Phys. B* **250**, 427.
- Chen, X., Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, 2013, "Symmetry Protected Topological Orders And The Group Cohomology Of Their Symmetry Group," *Phys. Rev. B* **87**, 155114.
- Dai, X., and D. S. Freed, 1994, " η -Invariants And Determinant Lines," *J. Math. Phys. (N.Y.)* **35**, 5155.
- Diaconescu, E., G. W. Moore, and E. Witten, 2002, " E_8 Gauge Theory, and a Derivation of K-Theory From M-Theory," *Adv. Theor. Math. Phys.* **6**, 1031.
- Dijkgraaf, R., and E. Witten, 1990, "Topological Gauge Theories And Group Cohomology," *Commun. Math. Phys.* **129**, 393.
- Fidkowski, L., X. Chen, and A. Vishwanath, 2013, "Non-Abelian Topological Order On The Surface Of A 3d Topological Superconductor From An Exactly Solved Model," *Phys. Rev. X* **3**, 041016.
- Fidkowski, L., and A. Kitaev, 2009, "The Effects Of Interactions On The Topological Classification Of Free Fermion Systems," *arXiv:0904.2197*.
- Freed, D., 2014, "Short Range Entanglement And Invertible Field Theories," *arXiv:1406.7278*.
- Freed, D., and G. Moore, 2006, "Setting The Quantum Integrand Of M-Theory," *Commun. Math. Phys.* **263**, 89.

- Freed, D. S., 2014, “Anomalies And Invertible Field Theories,” [arXiv:1404.7224](https://arxiv.org/abs/1404.7224).
- Gilkey, P., 1985, “The η -Invariant Of Even-Dimensional Pin_c Manifolds,” *Adv. Math.* **58**, 243.
- Hasan, M. Z., and C. L. Kane, 2010, “Topological Insulators,” *Rev. Mod. Phys.* **82**, 3045.
- Hasan, M. Z., and J. E. Moore, 2011, “Three-Dimensional Topological Insulators,” *Annu. Rev. Condens. Matter Phys.* **2**, 55.
- Horava, P., and E. Witten, 1996, “Heterotic and Type I String Dynamics From Eleven Dimensions,” *Nucl. Phys. B* **460**, 506.
- Hsieh, C.-T., G. C. Cho, and S. Ryu, 2015, “Global Anomalies On The Surface Of Symmetry-Protected Topological Phases in (3 + 1) Dimensions,” [arXiv:1503.01411](https://arxiv.org/abs/1503.01411).
- Hsieh, C.-T., O. M. Sule, G. Y. Cho, S. Tyu, and R. Leigh, 2014, “Symmetry-Protected Topological Phases, Generalized Laughlin Argument, and Orientifolds,” *Phys. Rev. B* **90**, 165134.
- Jackiw, R., and C. Rebbi, 1976, “Solitons With Fermion Number 1/2,” *Phys. Rev. D* **13**, 3398.
- Kapustin, A., R. Thorngren, A. Turzillo, and Z. Wang, 2014, “Fermionic Symmetry Protected Topological Phases And Cobordisms,” [arXiv:1406.7329](https://arxiv.org/abs/1406.7329).
- Kiskis, J., 1978, “Fermion Zero-Modes And Level Crossing,” *Phys. Rev. D* **18**, 3690.
- Kitaev, A., 2009, “Periodic Table For Topological Insulators And Superconductors,” [arXiv:0901.2686](https://arxiv.org/abs/0901.2686).
- Kitaev, A., 2015, “Homotopy-Theoretic Approach to SPT Phases in Action: \mathbb{Z}_{16} Classification of Three-Dimensional Superconductors,” lecture notes available at <http://www.ipam.ucla.edu/abstract/?tid=12389&pcode=STQ2015>.
- Metlitski, M., 2015, “S-Duality of U(1) Gauge Theory With $\theta = \pi$ on Non-Orientable Manifolds: Applications to Topological Insulators and Superconductors,” [arXiv:1510.05663](https://arxiv.org/abs/1510.05663).
- Metlitski, M. A., L. Fidkowski, X. Chen, and A. Vishwanath, 2014, “Interaction Effects On 3d Topological Superconductors: Surface Topological Order From Vortex Condensation, the 16-Fold Way, And Fermionic Kramers Doublets,” [arXiv:1406.3032](https://arxiv.org/abs/1406.3032).
- Mikhailov, V., and E. Witten, 2014, “Branes And Supergroups,” [arXiv:1410.1175](https://arxiv.org/abs/1410.1175).
- Monnier, S., 2014, “Hamiltonian Anomalies From Extended Field Theories,” [arXiv:1410.7442](https://arxiv.org/abs/1410.7442).
- Niemi, A., and G. Semenoff, 1983, “Anomaly-Induced Fermion Fractionization And Effective Actions In Odd-Dimensional Spacetimes,” *Phys. Rev. Lett.* **51**, 2077.
- Qi, X.-L., T. Hughes, and S.-C. Zhang, 2008, “Topological Field Theory Of Time-Reversal Invariant Insulators,” [arXiv:0802.3537](https://arxiv.org/abs/0802.3537).
- Qi, X.-L., and S.-C. Zhang, 2011, “Topological Insulators and Superconductors,” *Rev. Mod. Phys.* **83**, 1057.
- Redlich, N., 1984, “Parity Violation And Gauge Noninvariance of the Effective Gauge Field Action In Three Dimensions,” *Phys. Rev. D* **29**, 2366.
- Rosenberg, C., and M. Franz, 2010, “Witten Effect In A Crystalline Topological Insulator,” [arXiv:1001.3179](https://arxiv.org/abs/1001.3179).
- Ryu, S., J. E. Moore, and A. W. W. Ludwig, 2010, “Electromagnetic And Gravitational Responses And Anomalies In Topological Insulators And Superconductors,” [arXiv:1010.0936](https://arxiv.org/abs/1010.0936).
- Ryu, S., A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, 2010, “Topological Insulators And Superconductors: Tenfold Way And Dimensional Hierarchy,” *New J. Phys.* **12**, 065010.
- Stolz, S., 1988, “Exotic Structures On 4-Manifolds Detected By Spectral Invariants,” *Inventiones Mathematicae* **94**, 147.
- Wang, C., and T. Senthil, 2014, “Interacting Fermionic Topological Insulators/Superconductors In Three Dimensions,” *Phys. Rev. B* **89**, 195124.
- Witten, E., 1979, “Dyons Of Charge $e\theta/2\pi$,” *Phys. Lett. B* **86**, 283.
- Witten, E., 1982, “AnSU(2) Anomaly,” *Phys. Lett. B* **117**, 324.
- Witten, E., 1985a, “Global Anomalies In String Theory,” in *Symposium On Anomalies, Geometry, Topology*, edited by W. A. Bardeen and A. R. White (World Scientific, Singapore), pp. 61–99, <https://www.sns.ias.edu/witten/symp-global-anomalies-stringtheory-1985>.
- Witten, E., 1985b, “Global Gravitational Anomalies,” *Commun. Math. Phys.* **100**, 197.
- Witten, E., 2000, “World-Sheet Corrections Via D-Instantons,” *J. High Energy Phys.* **02**, 030.