

# Anomaly Inflow and the $\eta$ -Invariant

Edward Witten<sup>1</sup> and Kazuya Yonekura<sup>2</sup>

<sup>1</sup>*School of Natural Sciences, Institute for Advanced Study  
Einstein Drive, Princeton, NJ 08540 USA*

<sup>2</sup>*Department of Physics, Tohoku University, Sendai 980-8578, Japan*

## Abstract

Perturbative fermion anomalies in spacetime dimension  $d$  have a well-known relation to Chern-Simons functions in dimension  $D = d + 1$ . This relationship is manifested in a beautiful way in “anomaly inflow” from the bulk of a system to its boundary. Along with perturbative anomalies, fermions also have global or nonperturbative anomalies, which can be incorporated by using the  $\eta$ -invariant of Atiyah, Patodi, and Singer instead of the Chern-Simons function. Here we give a nonperturbative description of anomaly inflow, involving the  $\eta$ -invariant. This formula has been expected in the past based on the Dai-Freed theorem, but has not been fully justified. It leads to a general description of perturbative and nonperturbative fermion anomalies in  $d$  dimensions in terms of an  $\eta$ -invariant in  $D$  dimensions. This  $\eta$ -invariant is a cobordism invariant whenever perturbative anomalies cancel. (To be submitted to the proceedings of the Shoucheng Zhang Memorial Workshop.)

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>A Precise Formula For Anomaly Inflow</b>	<b>4</b>
2.1	Massive Dirac Fermion With Local Boundary Condition . . . . .	4
2.2	Integrating Out A Fermion . . . . .	9
2.3	The Chiral Fermion Partition Function as a State Overlap . . . . .	13
2.4	The General Case . . . . .	15
2.5	Treatment of Zero-Modes . . . . .	24
<b>3</b>	<b>The Anomaly</b>	<b>30</b>
3.1	Defining the Phase of the Path Integral . . . . .	31
3.2	Topological Field Theory And Cobordism . . . . .	34
3.3	Deformation Classes And Anomalies . . . . .	36
<b>4</b>	<b>Examples In Dimensions <math>d = 1, 2, 3, 4</math></b>	<b>39</b>
4.1	$d = 1$ . . . . .	40
4.2	$d = 3$ . . . . .	45
4.3	$d = 2$ . . . . .	47
4.4	$d = 4$ . . . . .	51

# 1 Introduction

Shoucheng Zhang made many outstanding contributions in condensed matter physics. He started his career in particle physics and was always very interested in the interplay between condensed matter physics and relativistic physics. He made numerous lasting contributions in areas that combine ideas of the two fields.

Fermion anomalies – originally the one-loop triangle anomaly of Adler, Bell, and Jackiw [1, 2] – have been important in particle physics since their discovery, and have proved to be important in condensed matter physics. In particular, they played a role in some of Shoucheng Zhang’s important contributions. The lecture of one of us at the Shoucheng Zhang Memorial Workshop was devoted to a particular question about fermion anomalies [3]. (The question is briefly described at the end of section 4.1.) The conference organizers suggested that written contributions to the proceedings could be on broader themes related to the topics of the lectures. Here we will describe a nonperturbative approach to fermion anomaly inflow.

A rather well-known fact about perturbative fermion anomalies is that the perturbative anomaly in spacetime dimension  $d$  is related to a Chern-Simons function in dimension  $D = d + 1$  [4–6]. This idea has a beautiful manifestation in the idea of anomaly inflow [7], which is important in both particle physics and condensed matter physics. The classic example in condensed matter physics is the integer quantum Hall effect. The worldvolume of a quantum Hall sample has spacetime dimension  $D = 3$  and its boundary has dimension  $d = 2$ . The “edge mode” of a quantum Hall system is – in relativistic language – a massless electrically charged chiral fermion. The coupling of this mode to electromagnetism has an anomaly – analogous to the original triangle anomaly. In the bulk of a quantum Hall system, the effective action for the gauge field  $A$  of electromagnetism includes a multiple of the Chern-Simons coupling  $\int A \wedge dA$ . This coupling is not gauge-invariant on a manifold with boundary, and that failure of gauge invariance cancels the quantum anomaly of the edge mode.

Along with perturbative anomalies, fermions can also have global or nonperturbative anomalies that are not visible in perturbation theory [8]. The relation between perturbative anomalies in  $d$  dimensions and a Chern-Simons function in  $d+1$  dimensions can be generalized to include nonperturbative anomalies [9]. One just has to replace the Chern-Simons function by the  $\eta$ -invariant of Atiyah, Patodi, and Singer (APS) [10]. The  $\eta$ -invariant is equivalent in perturbation theory to a Chern-Simons function, but it contains additional nonperturbative information. The relation between anomalies and the  $\eta$ -invariant motivated a mathematical result that we will call the Dai-Freed theorem [11]. See [12] for an explanation of the Dai-Freed theorem from a physical point of view. The Dai-Freed theorem makes it possible to get more precise results about fermion anomalies than were available otherwise; see for example [13], section 2.2.

Since anomaly inflow is an important aspect of understanding perturbative fermion anomalies, and since fermions do have nonperturbative anomalies, one would like to formulate anomaly inflow in a

way that incorporates nonperturbative anomalies as well as perturbative ones. The Dai-Freed theorem strongly suggests how to do this. The basic idea is to replace the Chern-Simons function with an  $\eta$ -invariant defined with APS boundary conditions. The resulting nonperturbative formula for anomaly inflow (eqn. (2.52) below) was described in [14]. However, this formula has not been fully justified. The main goal of the present article is to do this.

We present our nonperturbative approach to fermion anomaly inflow in section 2. We take a viewpoint that is natural in modern condensed matter physics. An arbitrary, possibly anomalous, fermion system in dimension  $d$  can be regarded as the boundary state of an anomaly-free gapped fermion system in dimension  $D = d + 1$ . Then the anomaly inflow formula is simply obtained, in principle, by integrating out the gapped fermions in the bulk. Since it is already known that integrating out a massive fermion gives a Chern-Simons term perturbatively [15] and an  $\eta$ -invariant nonperturbatively [16], it is fairly clear that this process will give an answer of a familiar kind but with Chern-Simons replaced with  $\eta$ . Less obvious is why the  $\eta$ -invariant should be computed with APS boundary conditions, as suggested by the Dai-Freed theorem. Explaining this is one of the main tasks of section 2. To avoid extraneous details, we begin section 2 assuming that the bulk fermion is a Dirac fermion, but then we extend the discussion to apply to the anomalies of a completely general relativistic fermion system. (In a somewhat similar way, though presented more abstractly, the Dai-Freed theorem was generalized recently to an arbitrary relativistic fermion system in the appendix of [17].)

In section 3, we explain that the nonperturbative formula for anomaly inflow implies a general description of perturbative and nonperturbative fermion anomalies in dimension  $d$  in terms of an exponentiated  $\eta$ -invariant in dimension  $d + 1$ . This generalizes the global anomaly formula obtained in [9] (where  $d$  was assumed to be even and only orientable manifolds were considered), and also gives more precise information about the fermion path integral in anomaly-free examples. In section 4, we use the description via the  $\eta$ -invariant to analyze potential global anomalies in various examples in dimensions  $d = 1, 2, 3, 4$ . One of the examples we consider involves a celebrated contribution by Shoucheng Zhang [18], involving the electromagnetic  $\theta$  angle in the bulk of a  $3+1$ -dimensional topological insulator. Another example we consider is the Standard Model of particle physics.

Though in the present article we consider only fermion anomalies, we should note that in modern condensed matter and particle physics, anomalies in bosonic systems play a role as well. In particular, in condensed matter physics, a symmetry protected topological (SPT) phase of matter is a gapped phase with some global symmetry group  $G$ , such that the phase is in some sense topologically nontrivial but would become trivial if the symmetry is explicitly broken. A version of anomaly inflow is important in understanding SPT phases [19–21], with group cohomology playing the role that for fermion anomalies is played by the  $\eta$ -invariant. In all cases, the anomaly in dimension  $d$  is related to an “invertible topological field theory” in dimension  $d + 1$ . For a general description of this point of view about anomalies, see [22, 23].

## 2 A Precise Formula For Anomaly Inflow

In this section, we study massive fermions on a manifold  $Y$  with boundary  $W$ .  $W$  will have dimension  $d$ , while  $Y$  has dimension  $D = d + 1$ . The goal is to get a precise formula for anomaly inflow from  $Y$  to  $W$ .

### 2.1 Massive Dirac Fermion With Local Boundary Condition

We begin with a massive Dirac fermion  $\Psi$  on a  $D$ -dimensional manifold  $Y$ . A Dirac fermion is just a fermion field  $\Psi$  whose components are all charged under a global or gauge  $U(1)$  symmetry. The conjugate field  $\bar{\Psi}$  carries opposite  $U(1)$  charges. The assumption of a  $U(1)$  symmetry is certainly nongeneric, but we begin with this case in order to explain the main idea of the derivation without potentially distracting details.

The discussion will be quite general;  $D$  may be even or odd,  $Y$  may be orientable or unorientable, and  $\Psi$  may be coupled to gauge fields as well as to the geometry of  $Y$ . We assume only that  $Y$  is endowed with the appropriate structure (such as a spin or pin structure, a  $\text{spin}_c$  structure, etc.) so that a suitable action for  $\Psi$  exists, of the usual form for a Dirac fermion:

$$I = - \int_Y d^D x \sqrt{g} \bar{\Psi} (\not{D} + m) \Psi. \quad (2.1)$$

$\Psi$  is coupled to the Riemannian metric  $g$  of  $Y$  and possibly to a background gauge field  $A$ . The Dirac operator is defined in the usual way,  $\not{D}_Y = \sum_{\mu=1}^D \gamma^\mu D_\mu$ , with  $\{\gamma_\nu, \gamma_\nu\} = 2g_{\mu\nu}$ . On a manifold without boundary, the operator  $\mathcal{D}_Y = i\not{D}_Y$  is self-adjoint.

Now suppose that the manifold  $Y$  has boundary  $W = \partial Y$ , with a metric that looks like a product near the boundary. Thus near the boundary the metric of  $Y$  is  $ds_Y^2 = d\tau^2 + ds_W^2$ , where  $ds_W^2$  is a metric on  $W$ , and  $\tau$  parametrizes the normal direction. We can normalize  $\tau$  to vanish along  $W$  and to be negative away from  $W$ . We assume this product description is valid at least in a range  $(-\epsilon, 0]$ .

We impose on  $\Psi$  the following local boundary condition:

$$\mathbf{L} : (1 - \gamma^\tau) \Psi|_{\tau=0} = 0 \quad (2.2)$$

where  $\gamma^\tau$  is the gamma matrix in the direction  $\tau$ . Before proceeding with the analysis, we will say a word about this boundary condition.

The boundary condition (2.2) is elliptic but not self-adjoint.<sup>1</sup> Ellipticity of the boundary condition

---

<sup>1</sup>Self-adjoint local boundary conditions are also possible, of course, at least in some cases. For recent analysis of a self-adjoint local boundary condition that leads to no localized boundary mode (but may explicitly violate symmetries such as time-reversal or reflection symmetry), see e.g. [24–26].

is necessary and sufficient to ensure that the Dirac operator with this boundary condition has the properties needed to make Euclidean field theory well-defined. For example, on a compact manifold, the Dirac operator with an elliptic boundary condition has a discrete spectrum, and (after possibly removing finitely many zero-modes) it has a Green's function with the usual properties (such as short distance singularities of a standard form). Technically, the condition for ellipticity is as follows. Suppose that  $W = \mathbb{R}^d$  and  $Y = \mathbb{R}^d \times \mathbb{R}_- = \mathbb{R}^{d+1}$ , where  $\mathbb{R}_-$  is a half-line parametrized by  $\tau \leq 0$ . A boundary condition on the Dirac equation at  $\tau = 0$  is elliptic if, after dropping from the equation lower order terms (notably mass terms and couplings to background fields), or equivalently after taking the momentum along  $\mathbb{R}^d$  to be extremely large, the equation and its adjoint have no solutions that satisfy the boundary condition and vanish for  $\tau \rightarrow -\infty$ . The local boundary condition (2.2) has this property, because any solution of the free Dirac equation on  $\mathbb{R}_-^{d+1}$  that has nonzero momentum along the boundary and vanishes for  $\tau \rightarrow -\infty$  has components with  $\gamma^\tau = -1$  as well as components with  $\gamma^\tau = +1$ .

A boundary condition on the Euclidean Dirac operator is self-adjoint if, with this boundary condition,  $i\not{D}_Y$  is self-adjoint on a manifold with boundary. As usual, a self-adjoint operator has real eigenvalues. The boundary condition (2.2) is not self-adjoint; an attempt to prove self-adjointness will fail because of a surface term at  $\tau = 0$ . Accordingly, though the operator  $i\not{D}_Y$  with boundary condition **L** has a discrete spectrum (if  $Y$  is compact), its eigenvalues are not real. In fact, a self-adjoint boundary condition could not lead to the potentially anomalous boundary physics that we are interested in. In Euclidean signature, self-adjointness of a local boundary condition is not a physical requirement. The actual condition that a local boundary condition should satisfy in order to be physically sensible is the following. Suppose that one analytically continues  $W$  to Lorentz signature, with time coordinate  $x^0$ . One wants the boundary condition to be such that the Hamiltonian that propagates a state in the  $x^0$  direction is self-adjoint. One can verify that this is true for the boundary condition **L**, using the fact that  $\{\gamma^0, \gamma^\tau\} = 0$ .

The action (2.1) can be written near the boundary as

$$I = - \int_Y d^D x \sqrt{g} \bar{\Psi} \gamma^\tau \left( \frac{\partial}{\partial \tau} + \mathcal{D}_W + \gamma^\tau m \right) \Psi. \quad (2.3)$$

where<sup>2</sup>

$$\mathcal{D}_W = \sum_{\mu \neq \tau} \gamma^\tau \gamma^\mu D_\mu \quad (2.4)$$

is a self-adjoint Dirac operator on the boundary  $W$ . Notice that  $\gamma^\tau$  and  $\mathcal{D}_W$  anticommute with each other. If  $d$  is even, the operator  $\gamma^\tau$  measures what is usually called the “chirality” of a fermion that propagates along  $W$ . We will call  $\gamma^\tau$  a chirality operator in any case, though for odd  $d$  this is not standard terminology.

---

<sup>2</sup>  $\mathcal{D}_W$  differs from the usual Dirac operator  $i \sum_{\mu=1}^d \gamma^\mu D_\mu$  of  $W$  only by the use of a different set of gamma matrices ( $-i\gamma^\tau \gamma^\mu$  instead of  $\gamma^\mu$ ) that obey the same algebra and lead to the same rotation generators  $\frac{1}{4}[\gamma_\mu, \gamma_\nu]$ . So the two operators are equivalent.

Our discussion so far makes sense for either sign of  $m$ , but in what follows, given our choice of boundary condition, the interesting case is  $m < 0$ . For  $m < 0$ , the Dirac equation  $(\not{D} + m)\Psi = 0$  has a mode localized near the boundary and given by

$$\Psi = \chi \exp(|m|\tau), \quad (1 - \gamma^\tau)\chi = 0, \quad \mathcal{D}_W \chi = 0, \quad (2.5)$$

where  $\chi$  is a fermion field on  $W$ . Since it vanishes exponentially for  $\tau \ll 0$ , this mode is localized along  $W$ . Since it satisfies  $\mathcal{D}_W \chi = 0$ , it propagates along  $W$  as a massless fermion. Finally, as it obeys  $\gamma^\tau \chi = \chi$ , it is a chiral fermion along  $W$ . For  $m > 0$ , there is no such boundary-localized mode.

To quantize the fermion field  $\Psi$ , some sort of regulator is needed. A useful choice for our purposes is a simple Pauli-Villars regulator, defined by adding a very massive field of opposite statistics, satisfying the same Dirac equation (possibly with a different mass parameter) and the same boundary condition. We take the Pauli-Villars regulator field to have a positive mass parameter, since we do not want the regulator field to have a low energy mode propagating along the boundary, which would be quite unphysical.

Now we want to compute the partition function of the above massive fermion  $\Psi$  on the manifold  $Y$  with the boundary condition  $\mathbf{L}$ . For this purpose, it is useful to think of  $\tau$  as a Euclidean time coordinate, and the boundary  $W = \partial Y$  as a time-slice. In this point of view, the path integral over  $Y$  gives a physical state vector which we denote as  $|Y\rangle$ . This is a state vector in the Hilbert space  $\mathcal{H}_W$  of  $W$ . The boundary condition  $\mathbf{L}$  also corresponds to a state vector  $|\mathbf{L}\rangle \in \mathcal{H}_W$ , which we will later discuss explicitly. Then, the partition function on  $Y$  with the boundary condition  $\mathbf{L}$  is just given by

$$Z(Y, \mathbf{L}) = \langle \mathbf{L} | Y \rangle. \quad (2.6)$$

The important point is now as follows. We take the mass  $|m|$  to be very large so that its Compton wavelength  $1/|m|$  is much shorter than the typical scale of the manifolds  $Y$  and  $W$ ; equivalently, we take the length scale of  $Y$  and  $W$  to be much larger than  $1/|m|$ . In this limit, there is a large mass gap in the Hilbert space  $\mathcal{H}_W$ . Moreover, the path integral on a cylindrical region  $(-\epsilon, 0] \times W$  gives the Euclidean time evolution  $e^{-\epsilon H}$  where  $H$  is the Hamiltonian. If the mass gap is large so that  $\epsilon|m| \gg 1$ , the factor  $e^{-\epsilon H}$  plays the role of the projection operator to the ground state  $|\Omega\rangle$ :

$$e^{-\epsilon H} \simeq |\Omega\rangle\langle\Omega| \quad (\epsilon|m| \gg 1). \quad (2.7)$$

This is valid up to errors which are exponentially suppressed by  $e^{-\epsilon|m|}$ , which we neglect. In this limit,  $|Y\rangle$  is proportional to the ground state,

$$|Y\rangle \propto |\Omega\rangle, \quad (2.8)$$

where we may assume  $\langle\Omega|\Omega\rangle = 1$ . So we can rewrite the partition function as

$$Z(Y, \mathbf{L}) = \langle \mathbf{L} | \Omega \rangle \langle \Omega | Y \rangle. \quad (2.9)$$

In this way we can split the bulk contribution  $\langle\Omega|Y\rangle$  and the boundary contribution  $\langle\mathbf{L}|\Omega\rangle$ .

The ground state  $|\Omega\rangle$  has a phase ambiguity, and hence the splitting into  $\langle\Omega|Y\rangle$  and  $\langle L|\Omega\rangle$  also has this ambiguity. In the context of the Dai-Freed theorem [11], a physical interpretation of this ambiguity is as follows [12]. The space of ground states is a one-dimensional subspace of the Hilbert space  $\mathcal{H}_W$ . We denote this one-dimensional subspace as  $\mathcal{L}_W \subset \mathcal{H}_W$ . Now suppose that we make an adiabatic change of background fields such as, e.g., the metric  $g$  of  $W$  or an external gauge field  $A$ . The ground state changes adiabatically, and evolves with a Berry phase. Let  $\mathcal{W}$  be the parameter space of the background fields. The existence of such a Berry phase means that we may get a nontrivial holonomy when going around a loop in  $\mathcal{W}$ . The one-dimensional spaces  $\mathcal{L}_W$  at each point of  $\mathcal{W}$  combine into a rank one complex vector bundle (a complex line bundle) over  $\mathcal{W}$  with nontrivial holonomies determined by the Berry phases. Then instead of  $\langle\Omega|Y\rangle$  and  $\langle L|\Omega\rangle$ , which do not have well-defined phases, we can define the following:

$$|\Omega\rangle\langle\Omega|Y\rangle \in \mathcal{L}_W, \quad \langle L|\Omega\rangle\langle\Omega| \in \mathcal{L}_W^{-1}. \quad (2.10)$$

These formulas are well-defined, since they do not depend on the phase of the state  $|\Omega\rangle$ . But instead of a number  $\langle\Omega|Y\rangle$  we get a vector  $|\Omega\rangle\langle\Omega|Y\rangle \in \mathcal{L}_W$ , and likewise instead of a number  $\langle L|\Omega\rangle$  we get a vector  $\langle L|\Omega\rangle\langle\Omega| \in \mathcal{L}_W^{-1}$ . All this is related to the following fact [27]: the partition function of the chiral fermion  $\chi$  on the boundary  $W$  is naturally understood as a section of a line bundle  $\mathcal{L}_W^{-1}$  (called the determinant line bundle) rather than a complex number. Also, the factor  $\langle\Omega|Y\rangle$  is an exponentiated  $\eta$ -invariant that can be regarded as an element of  $\mathcal{L}_W$  [11]. See [12] for more details on these points, from a physical perspective.

Somewhat surprisingly, however, as long as we are in a situation in which the Dirac operator on  $W$  generically has no zero-mode,<sup>3</sup> it is possible to avoid introducing the line bundle  $\mathcal{L}_W$ . One can avoid the line bundle  $\mathcal{L}_W$  by using instead of  $|\Omega\rangle$  a state which has no phase ambiguity. As long as the boundary Dirac operator  $\mathcal{D}_W$  has no zero modes, we can consider the global Atiyah-Patodi-Singer (APS) boundary condition [10], which we denote as **APS**. This boundary condition will be described explicitly in section 2.2. The path integral on  $Y$  with boundary condition **APS** corresponds to a state vector  $|\text{APS}\rangle \in \mathcal{H}_W$ , which we will also describe explicitly in section 2.2. Because  $|Y\rangle$  is a multiple of  $|\Omega\rangle$ , we can rewrite (2.9) in the following longer but useful form:

$$Z(Y, L) = \frac{\langle L|\Omega\rangle\langle\Omega|\text{APS}\rangle}{|\langle\text{APS}|\Omega\rangle|^2} \cdot \langle\text{APS}|Y\rangle. \quad (2.11)$$

In this formula, the inner product  $\langle\text{APS}|Y\rangle$  is the partition function of the massive fermion  $\Psi$  on  $Y$  with the APS boundary condition. In section 2.2, we show that at long distances (that is, if  $Y$  and  $W$  are large compared to  $1/|m|$ ), we have, modulo nonuniversal factors that do not affect the analysis of anomalies,<sup>4</sup>

$$\langle\text{APS}|Y\rangle = \exp(-i\pi\eta_D), \quad (2.12)$$

---

<sup>3</sup>The case that  $\mathcal{D}_W$  does generically have a zero-mode will be analyzed in section 2.5. As explained there, this case is somewhat more exotic since, for example, it does not arise if  $W$  is connected.

<sup>4</sup>These are factors that can be removed by adding to the action the integral of a gauge invariant function of the background fields  $g, A$ . For example, the ground state energy per unit volume of the  $\Psi$  field is a nonuniversal effect that can be removed by adding to the action a multiple of the volume of  $Y$ .



where  $\eta_D$  is the APS  $\eta$ -invariant (defined with APS boundary conditions along  $W = \partial Y$ ) for the Dirac fermion  $\Psi$ ;  $\eta_D$  will be introduced in section 2.2. The remaining factor  $\langle L|\Omega\rangle\langle\Omega|\text{APS}\rangle/|\langle\text{APS}|\Omega\rangle|^2$  in  $Z(Y, L)$  will be analyzed in section 2.3, and we will see that (modulo nonuniversal factors) it equals  $|\text{Det } \mathcal{D}_W^+|$ , the absolute value of the path integral on  $W$  for the massless chiral fermion  $\chi$ . (The superscript  $+$  in  $\mathcal{D}_W^+$  means that here  $\mathcal{D}_W$  is taken to act only on the field  $\chi$  of positive chirality.) It is essential that an absolute value comes in here. In general, the path integral of the boundary mode  $\chi$  might be anomalous, so the corresponding determinant, without taking an absolute value, may not be well-defined. However, an anomaly always affects only the phase of the path integral, so the absolute value  $|\text{Det } \mathcal{D}_W^+|$  is always anomaly-free.

Combining these claims about the various factors in  $Z(Y, L)$ , we get the following formula for the partition function on  $Y$  with boundary condition  $L$ :

$$Z(Y, L) = |\text{Det } \mathcal{D}_W^+| \exp(-i\pi\eta_D). \quad (2.13)$$

We claim that (for fermions charged under a  $U(1)$  symmetry) this is the general statement of anomaly inflow, including all perturbative and nonperturbative anomalies.

If one is only interested in perturbative anomalies, then eqn. (2.13) can be replaced by a slightly simpler and probably more familiar version:

$$Z(Y, L) \stackrel{?}{=} \text{Det } \mathcal{D}_W^+ \exp(-iCS(g, A)). \quad (2.14)$$

Here  $CS(g, A)$  is a Chern-Simons interaction that depends on the background fields  $g, A$  on  $Y$ . In perturbation theory, this formula is a satisfactory expression of the idea of anomaly inflow [7]. The idea behind this formula is that the path integral of the boundary fermion is the determinant  $\text{Det } \mathcal{D}_W^+$ , and integrating out the massive fermion  $\Psi$  on  $Y$  induces the Chern-Simons coupling  $CS(g, A)$ . When one proves the gauge-invariance of  $CS(g, A)$ , one encounters a surface term, as a result of which  $CS(g, A)$  is not gauge-invariant on a manifold with boundary. This failure of gauge invariance of  $CS(g, A)$  on a manifold with boundary precisely matches the perturbative anomaly of a chiral fermion on the boundary [4–6], as a result of which the formula (2.14) is satisfactory as far as perturbative anomalies are concerned.

In general, however, in a topologically nontrivial situation, the result of integrating out a massive Dirac fermion, even on a manifold without boundary, is not  $\exp(-iCS(g, A))$  but  $\exp(-i\pi\eta_D)$  [14, 16]. This fact will be reviewed in section 2.2. The two expressions are equivalent in perturbation theory around Euclidean space, that is in perturbation theory around  $Y = \mathbb{R}^D$ ,  $g_{\mu\nu} = \delta_{\mu\nu}$ , and  $A = 0$ . But in general, the description by  $\exp(-iCS(g, A))$  misses many topological subtleties that are important in an understanding of global or nonperturbative anomalies and anomaly inflow and that are properly described by  $\exp(-i\pi\eta_D)$ . This is true for all  $D$ , but perhaps it is most obvious if  $D$  is even. If  $D$  is even, there are no conventional Chern-Simons couplings, and in perturbation theory around Euclidean space, one sees nothing. But  $\eta_D$  can still be nontrivial and describes anomaly inflow.

Bearing in mind that  $\exp(-iCS)$  should be replaced by  $\exp(-i\pi\eta_D)$  if we want a nonperturbative description, let us ask what might be the correct general formula for anomaly inflow, replacing (2.14).

Part of the Dai-Freed theorem [11], which has been elucidated from a physical point of view in [12], is an abstract definition of  $\exp(-i\pi\eta_D)$  (with APS boundary conditions along  $W = \partial Y$ ) as a section of the line bundle  $\mathcal{L}_W$ . In that framework, the obvious candidate for improving on (2.14) is  $Z(Y, \mathbb{L}) = \text{Det}\mathcal{D}_W^+ \cdot \exp(-i\pi\eta_D)$ . Neither factor is gauge-invariant separately, but the Dai-Freed theorem shows that the product is well-defined and completely free of any perturbative or nonperturbative anomaly.

This formulation is correct but unnecessarily abstract. As long as the operator  $\mathcal{D}_W$  generically has no zero-mode, it is possible to simply define  $\eta_D$  as a number; this is the route that we will follow.<sup>5</sup> Then the formula takes the form of eqn. (2.13). In this formulation, the two factors are separately gauge-invariant, but they are not separately physically sensible. Indeed, because of the absolute value,  $|\text{Det}\mathcal{D}_W^+|$  does not vary smoothly near a value of background fields  $g, A$  at which  $\mathcal{D}_W$  has a zero-mode (and therefore  $\text{Det}\mathcal{D}_W^+ = 0$ ), while  $\exp(-i\pi\eta_D)$  does not vary smoothly near such a point because the definition of  $\eta_D$  breaks down. The Dai-Freed theorem says that the product in (2.13) is always smoothly varying.

Thus eqn. (2.13) is the natural generalization of the more familiar but imprecise formula (2.14). But to our knowledge a full justification of eqn. (2.13) has not previously been given. We will address this issue in sections 2.3 and 2.4, after first explaining the basic facts about the  $\eta$ -invariant.

## 2.2 Integrating Out A Fermion

Though our main interest in the present paper is in integrating out a massive fermion, it is also of interest to know what happens if one integrates out a massless fermion. It turns out to be useful to consider this case first.

We consider a Dirac fermion  $\Psi$  on a  $D$ -manifold  $Y$ , initially without boundary.  $\Psi$  has a self-adjoint Dirac operator  $\mathcal{D}_Y = i\not{D}_Y$ . It has real eigenvalues that we denote  $\lambda_k$ .

Formally, the path integral of  $\Psi$  is  $\text{Det}\mathcal{D}_Y = \prod_k \lambda_k$ . This of course needs to be regulated. We do so with the aid of a Pauli-Villars regulator of mass  $M$ . A regulated version of the determinant is<sup>6</sup>

$$\text{Det}\mathcal{D}_Y = \prod_k \frac{\lambda_k}{\lambda_k + iM}. \quad (2.15)$$

---

<sup>5</sup>Fundamentally, it is not necessary to introduce the line bundle  $\mathcal{L}_W$  because as long as  $\mathcal{D}_W$  has no zero-mode,  $\text{Det}\mathcal{D}_W^+$  is nonzero and trivializes  $\mathcal{L}_W$ . Relative to this trivialization,  $\text{Det}\mathcal{D}_W^+$  is positive (so it is replaced by its absolute value  $|\text{Det}\mathcal{D}_W^+|$ ) and  $\exp(-i\pi\eta_D)$  becomes a number. To understand what follows, it is not necessary to think in such abstract terms.

<sup>6</sup>Actually, to completely regularize the determinant, one needs several Pauli-Villars fields with different masses and statistics. The details do not really modify the following derivation. The same remark applies whenever we consider expressions that require regularization.

Now let us determine the phase or argument of  $\text{Det } \mathcal{D}_Y$ . This is ill-defined if any of the  $\lambda_k$  vanish, so we assume they are nonvanishing. Assuming  $M > 0$  and  $M \gg |\lambda_k|$ , we see that the argument of  $\lambda_k/(\lambda_k + iM)$  is approximately  $-\frac{i\pi}{2}\text{sign}(\lambda_k)$ . So the argument of  $\text{Det } \mathcal{D}_Y$  is a regularized version of  $-\frac{i\pi}{2}\sum_k \text{sign}(\lambda_k)$ .

The Atiyah-Patodi-Singer  $\eta$ -invariant is a regularized version of  $\sum_k \text{sign}(\lambda_k)$ . The precise regularization does not matter. We can take, for example,<sup>7</sup>

$$\eta_D = \lim_{\epsilon \rightarrow 0^+} \sum_k \exp(-\epsilon|\lambda_k|)\text{sign}(\lambda_k). \quad (2.16)$$

The Pauli-Villars regularization in (2.15) gives a different regularization of  $\sum_k \text{sign}(\lambda_k)$  (it gives a regularization since the argument of  $\lambda_k/(\lambda_k + iM)$  vanishes for  $|\lambda_k| \rightarrow \infty$ ). The two regularizations are equivalent in the limit  $M \rightarrow \infty$  or  $\epsilon \rightarrow 0^+$ .

The phase of the path integral for a massless fermion (with a positive regulator mass) is therefore  $\exp(-i\pi\eta_D/2)$  [16]. This is a refinement of a well-known perturbative calculation, expanding around Euclidean space with a flat metric and  $A = 0$ , in which one finds a Chern-Simons coupling [15]. The nonperturbative statement with  $\eta_D$  can be related to the perturbative statement based on a Chern-Simons coupling by using the APS index theorem [10] for a manifold with boundary. Recall first that if  $X$  is a  $D+1$ -manifold without boundary, then the original Atiyah-Singer index theorem [28] gives a formula for the index  $\mathcal{I}$  of a Dirac operator on  $X$  as the integral over  $X$  of a certain differential form<sup>8</sup>  $\Phi_{d+2}$  on  $X$ :  $\mathcal{I} = \int_X \Phi_{d+2}$ . If instead  $X$  has a nonempty boundary  $Y$ , the APS index theorem gives a more general formula

$$\mathcal{I} = \int_X \Phi_{d+2} - \frac{\eta_D}{2}, \quad (2.17)$$

where  $\eta_D$  is the  $\eta$ -invariant of a suitable fermion field on  $Y$ . (Roughly, a fermion field  $\Psi$  on  $X$  reduces along  $Y = \partial X$  to two copies of a fermion field on  $Y$ , and  $\eta_D$  is here the  $\eta$ -invariant of one such copy. For more detail, see section 3.2.) In perturbation theory around Euclidean space, one can assume that a suitable  $X$  exists with  $\partial X = Y$  (if  $Y = \mathbb{R}^D$ ,  $X$  can be a half-space in  $\mathbb{R}^{D+1}$ ). Moreover in perturbation theory, one does not see the integer  $\mathcal{I}$ . Under these conditions, eqn. (2.17) reduces to  $\eta_D/2 = \int_X \Phi_{d+2}$ , so one can think of  $\eta_D/2$  as a boundary term related to the characteristic class  $\Phi_{d+2}$ , or in other words as a Chern-Simons interaction in a generalized sense. In a topologically nontrivial situation,  $X$  may not exist and if it does the integer  $\mathcal{I}$  cannot necessarily be neglected. So one has to describe the phase of the path integral for a massless fermion as  $\exp(-i\pi\eta_D/2)$ , not as the exponential of a Chern-Simons coupling.

Now let us consider integrating out a massive fermion, still on a  $D$ -manifold  $Y$  without boundary. In this case, the path integral is  $\text{Det } (\mathcal{D}_Y + im)$ . This formally equals  $\prod_k (\lambda_k + im)$ . A regularized version

---

<sup>7</sup>The subscript  $D$  in  $\eta_D$  is for Dirac; for a Dirac fermion  $\Psi$ , we define  $\eta_D$  by summing over modes of  $\Psi$ , ignoring the modes of the conjugate field  $\bar{\Psi}$ . We will shortly introduce a related invariant  $\eta$  that is defined by summing over all fermion modes regardless of charge.

<sup>8</sup>Concretely, for the case that  $\Psi$  transforms in a representation  $V$  of some gauge group,  $\Phi_{d+2} = \hat{A}(R) \text{Tr}_V \exp(F/2\pi)$ , where  $R$  is the Riemann tensor,  $\hat{A}(R)$  is a certain polynomial in  $R$ , and  $F$  is the gauge field strength.

is  $\prod_k (\lambda_k + im)/(\lambda_k + iM)$ . This can be factored as a ratio of two factors that we have already studied:

$$\prod_k \frac{\lambda_k}{\lambda_k + iM} \prod_{k'} \frac{\lambda_{k'} + im}{\lambda_{k'}}. \quad (2.18)$$

The resulting phase is trivial if  $m$  and  $M$  have the same sign. Indeed, in the previous derivation, only the sign of  $M$  was relevant, so if  $m$  and  $M$  have the same sign, we can set  $m = M$ . But then the product in eqn. (2.18) is simply 1. The interesting case is therefore the “topologically nontrivial” case in which the physical fermion and the regulator have mass parameters with opposite signs. For example (as in section 2.1), we may take  $m < 0$  and  $M > 0$ . Then eqn. (2.18) is the product of a factor we have already studied divided by the complex conjugate of the same factor. The argument of the path integral is hence  $\exp(-i\pi\eta_D)$ .

This derivation has been so general that it also applies if  $Y$  has a boundary with a self-adjoint boundary condition, the role of self-adjointness being to ensure that the Dirac eigenvalues are real. In particular, an example of a self-adjoint boundary condition is the global boundary condition of Atiyah, Patodi, and Singer [10]. This accounts for the claim that  $\langle \text{APS} | Y \rangle = \exp(-i\pi\eta_D)$  (eqn. (2.12)).

But what is the APS boundary condition? It is not a local boundary condition but a rather subtle global one. The original APS definition (in the notation of the present paper) was that  $\Psi$ , restricted to  $W = \partial Y$ , is required to be a linear combination of eigenfunctions of  $\mathcal{D}_W$  with positive eigenvalue.<sup>9</sup> For our purposes, however, the most useful way to describe the APS boundary condition is to say that  $|\text{APS}\rangle$  is the state  $|\Omega_{m=0}\rangle$  that would be the fermion ground state if one sets  $m = 0$ . This can be regarded as a state in the  $m \neq 0$  Hilbert space (though it is certainly not the ground state of the  $m \neq 0$  Hamiltonian); this is made explicit in section 2.3. Because of our assumption that the operator  $\mathcal{D}_W$  has no zero-mode, the  $m = 0$  ground state  $|\Omega_{m=0}\rangle$  is uniquely determined up to phase.<sup>10</sup> After Pauli-Villars regularization, there is no phase ambiguity in the state  $|\text{APS}\rangle$ , because in defining  $|\text{APS}\rangle$ , we use the same massless ground state  $|\Omega_{m=0}\rangle$ , with the same phase, for the physical fermion and for the regulator. The choice of phase cancels between the physical fermion field and the regulator.

We conclude with three remarks that are important in generalizations.

First, what happens if we replace the local boundary condition  $\mathbf{L}$  of equation (2.2) with the opposite boundary condition  $\mathbf{L}'$  defined by  $(1 + \gamma^\tau)\Psi|_W = 0$ ? In this case, we get no boundary-localized mode if we set  $m < 0$ , but for  $m > 0$ , we get a boundary-localized mode with the opposite chirality to that in

---

<sup>9</sup>This boundary condition does not have a sensible continuation to the case that  $W$  has Lorentz signature. The properties of the APS boundary condition such as nonlocality and the absence of a continuation to Lorentz signature motivated some study of the APS theorem in a physical context [29]. See also [30] for a new treatment of the APS index theorem. For an early use of the APS boundary condition in the physics literature in a different context, see [31]. Notice that the local boundary condition  $\mathbf{L}$  is completely physically sensible, and we just introduce the APS boundary condition as a way to compute the partition function (2.6).

<sup>10</sup>If  $\mathcal{D}_W$  has zero-modes, then quantizing these modes gives a space of degenerate ground states of the  $m = 0$  theory and the simple definition of the APS state and boundary condition breaks down. As a result, a more elaborate definition of APS boundary conditions is needed in that situation. This case is treated in section 2.5.

eqn. (2.5). In quantizing this theory, we have to take  $M < 0$ , since we do not want the regulator field to have a massless mode propagating on the boundary. Reversing the sign of both  $m$  and  $M$  relative to the above derivation has the effect of complex-conjugating the product in eqn. (2.18), so now the phase is  $\exp(i\pi\eta_D)$ . In other words, anomaly inflow for a boundary mode of one chirality requires a phase  $\exp(-i\pi\eta_D)$ , and anomaly inflow for the opposite chirality requires a complex conjugate phase  $\exp(i\pi\eta_D)$ . The fact that the sign of the inflow depends on the chirality of the boundary mode is a standard result, usually deduced perturbatively in terms of Chern-Simons couplings rather than an  $\eta$ -invariant.

Second, in this derivation, we have assumed that  $\Psi$  carries a  $U(1)$  gauge or global symmetry. A “Dirac fermion” is usually understood as a fermion field such that  $\Psi$  has positive  $U(1)$  charge and its conjugate  $\bar{\Psi}$  (which appeared in the original action (2.1) though we have not had to discuss it subsequently) has negative  $U(1)$  charge. The Dirac operator  $\mathcal{D}_Y$  could be taken to act on fermions of either positive or negative charge, that is, on either  $\Psi$  or  $\bar{\Psi}$ . We defined  $\eta_D$  by summing over eigenvalues of  $\mathcal{D}_Y$  acting on  $\Psi$  only. Obviously, we could have defined a similar invariant  $\eta$  by summing over all eigenvalues of  $\mathcal{D}_Y$ , acting on either  $\Psi$  or  $\bar{\Psi}$ . The relation between  $\eta$  and  $\eta_D$  is simply  $\eta = 2\eta_D$ , since complex conjugation can be used to show that the eigenvalues of  $\mathcal{D}_Y$  acting on  $\Psi$  or on  $\bar{\Psi}$  are the same.<sup>11</sup> In the absence of a  $U(1)$  symmetry, there is no distinction between  $\Psi$  and  $\bar{\Psi}$ , so the only natural definition involves a sum over all eigenvalues of  $\mathcal{D}_Y$ . In other words, to express our results in a way that generalizes naturally to an arbitrary fermion system, we should use  $\eta$  rather than  $\eta_D$ . In terms of  $\eta$ , the effect of integrating out a massive fermion (with  $m < 0$ ,  $M > 0$ ) is a factor  $\exp(-i\pi\eta/2)$ . This is the formula that we will use when we drop the assumption of  $U(1)$  symmetry, starting in section 2.4.

Finally, in the above derivation, we assumed that  $\mathcal{D}_Y$  has no zero-modes, so that all  $\lambda_k$  are nonzero. We did this because we started with a massless fermion; the path integral of a massless fermion vanishes when there is a zero-mode, and does not have a well-defined phase. However, a zero-mode of  $\mathcal{D}_Y$  causes no problem in defining the phase that is induced by integrating out a *massive* fermion. Taking  $m = -M$ , each mode contributes to the regularized fermion path integral a factor  $(\lambda - iM)/(\lambda + iM)$ , which is  $-1$  if  $\lambda = 0$ . So in including zero-modes, we want each such mode to give a factor of  $-1$ . For this, we define

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

and then we define  $\eta_D$  (and  $\eta = 2\eta_D$ ) precisely as in (2.16). Then the phase induced by integrating out a fermion with  $m < 0$ ,  $M > 0$  is  $\exp(-i\pi\eta_D)$  or  $\exp(-i\pi\eta/2)$ , whether zero-modes are present or not. We should warn the reader, however, that this definition of  $\eta$  or  $\eta_D$  including the zero-modes is not completely standard. In the original APS paper [10] (and many subsequent mathematical references), such a sum with the zero-modes included is called  $2\xi$  rather than  $\eta$ ;  $\eta$  is defined precisely as in eqn. (2.16), but summing over nonzero modes only.

---

<sup>11</sup>We could have made the whole derivation thinking of  $\mathcal{D}_Y$  as an operator acting on  $\bar{\Psi}$  rather than  $\Psi$ , leading to an  $\eta$ -invariant defined in terms of the eigenvalues of  $\mathcal{D}_Y$  acting on  $\bar{\Psi}$ . So  $\eta$ -invariants defined using eigenvalues of  $\mathcal{D}_Y$  acting on  $\Psi$  or on  $\bar{\Psi}$  must be equal.

## 2.3 The Chiral Fermion Partition Function as a State Overlap

To complete the proof of the formula (2.13) for the partition function, we must calculate the state overlaps  $\langle \mathbb{L} | \Omega \rangle$  and  $\langle \text{APS} | \Omega \rangle$  that appear in eqn. (2.11). We will see that  $\langle \mathbb{L} | \Omega \rangle$  is essentially the partition function of the boundary chiral fermion  $\chi$ .<sup>12</sup>

They are computed by a straightforward quantization of the fermion  $\Psi$  on the space  $W$ . We make a Wick rotation  $\tau \rightarrow it$ , and the Hamiltonian is easily read off from the action (2.1) as

$$H = \int_W d^d x \sqrt{g} \Psi^\dagger (\mathcal{D}_W + \gamma^\tau m) \Psi, \quad (2.20)$$

where we have used the usual relation  $\bar{\Psi} = \Psi^\dagger \gamma^\tau = \Psi^\dagger (i\gamma^t)$  (for Dirac fermions).

The operators  $\gamma^\tau$  and  $\mathcal{D}_W$  anticommute, so we can make a mode expansion of the following kind. All modes of the  $\Psi$  field on  $W$  can be expressed as linear combinations of pairs of modes  $(\psi_{+,a}, \psi_{-,a})$  which satisfy

$$\gamma^\tau \psi_{\pm,a} = \pm \psi_{\pm,a}, \quad \mathcal{D}_W \psi_{+,a} = \lambda_a \psi_{-,a}, \quad \mathcal{D}_W \psi_{-,a} = \lambda_a \psi_{+,a}. \quad (2.21)$$

Here  $\lambda_a$  are the positive square roots of the eigenvalues of  $(\mathcal{D}_W)^2$ . Since we assume that  $\mathcal{D}_W$  has no zero-mode, all modes have  $\lambda_a > 0$  and are paired in the above fashion. We expand

$$\Psi = \sum_a (A_{+,a} \psi_{+,a} + A_{-,a} \psi_{-,a}). \quad (2.22)$$

The Hamiltonian is given by

$$H = \sum_a (A_{+,a}^\dagger, A_{-,a}^\dagger) \begin{pmatrix} m & \lambda_a \\ \lambda_a & -m \end{pmatrix} \begin{pmatrix} A_{+,a} \\ A_{-,a} \end{pmatrix}. \quad (2.23)$$

The operators  $A_{\pm,a}$  are annihilation operators, and  $A_{\pm,a}^\dagger$  are the corresponding creation operators.

We can consider each value of  $a$  separately, so we omit  $\sum_a$  (or  $\prod_a$ , depending on the context) until the end of this section. After defining

$$(\cos(2\theta_a), \sin(2\theta_a)) = \frac{(m, \lambda_a)}{\sqrt{m^2 + \lambda_a^2}}, \quad (2.24)$$

the Hamiltonian is easily diagonalized as

$$H = \sqrt{\lambda_a^2 + m^2} (B_{1,a}^\dagger B_{1,a} - B_{2,a}^\dagger B_{2,a}), \quad (2.25)$$

---

<sup>12</sup> A similar phenomenon that a state vector in  $d+1$ -dimensions is related to a partition function in  $d$ -dimensions was also observed in a different system [32]. There, the  $d$ -dimensional theory is a free  $U(1)$  Maxwell theory with  $d=4$  and the  $D=d+1$ -dimensional theory is a gauge theory which flows to a topological field theory in the low energy limit. Their result might be interpreted as anomaly inflow of the global anomaly of the Maxwell theory under S-duality [33–35].



where

$$B_{1,a} = \cos \theta_a A_{+,a} + \sin \theta_a A_{-,a}, \quad B_{2,a} = -\sin \theta_a A_{+,a} + \cos \theta_a A_{-,a}. \quad (2.26)$$

The ground state  $|\Omega\rangle$  is specified by the conditions  $B_{1,a}|\Omega\rangle = 0$  and  $B_{2,a}^\dagger|\Omega\rangle = 0$ , or  $0 = \langle\Omega|B_{1,a}^\dagger = \langle\Omega|B_{2,a}$ .

The local boundary condition  $\mathbf{L}$  given in (2.2) requires the condition that  $A_{-,a} = 0$  at the boundary, while  $A_{+,a}$  is free. Then the corresponding state  $|\mathbf{L}\rangle$  is specified by the condition that  $\langle\mathbf{L}|A_{-,a} = 0$  and  $\langle\mathbf{L}|A_{+,a}^\dagger = 0$ . (See [12] for more details about the relation between boundary conditions and state vectors.) The APS boundary state  $|\text{APS}\rangle$  is the ground state with  $m$  set to 0, which corresponds to  $\theta_a = \pi/4$  for all  $a$  and hence  $\cos \theta_a = \sin \theta_a$ . So  $|\text{APS}\rangle$  is characterized by  $\langle\text{APS}|(A_{-,a} - A_{+,a}) = 0$  and  $\langle\text{APS}|(A_{-,a} + A_{+,a})^\dagger = 0$ .

These conditions leave us free to multiply the states  $\langle\mathbf{L}|$  and  $\langle\text{APS}|$  by arbitrary nonzero complex numbers. We could constrain them to have unit norm, but this leaves an undetermined phase. However, the ambiguities do not matter. This is because we will eventually take the ratio between the negative mass theory  $m < 0$  (for the physical fermion) and the positive mass theory  $m > 0$  (for the Pauli-Villars regulator). The ambiguities cancel out in the ratio, because the boundary conditions  $\langle\mathbf{L}|$  and  $\langle\text{APS}|$  do not depend on the mass parameter. The ground state  $|\Omega\rangle$  has a phase ambiguity, which is related to the appearance of the line bundle  $\mathcal{L}_W$ , as discussed in section 2.1. This ambiguity is physically unavoidable because of Berry phases when background fields are varied. But this ambiguity cancels out in the product  $\langle\mathbf{L}|\Omega\rangle\langle\Omega|\text{APS}\rangle$ . Thus overall normalization factors will always cancel out and we can use any explicit state vectors satisfying the above conditions.

Let  $|E\rangle$  be a state vector with

$$A_{+,a}|E\rangle = A_{-,a}|E\rangle = 0. \quad (2.27)$$

Then we can realize the relevant state vectors as

$$|\Omega\rangle = B_{2,a}^\dagger|E\rangle, \quad \langle\mathbf{L}| = \langle E|A_{-,a}, \quad \langle\text{APS}| = \langle E|(A_{-,a} - A_{+,a}). \quad (2.28)$$

From these expressions, we easily get

$$\langle\text{APS}|\Omega\rangle = \cos \theta_a + \sin \theta_a \rightarrow 1 \quad (\lambda_a \ll |m|), \quad (2.29)$$

and

$$\langle\mathbf{L}|\Omega\rangle = \cos \theta_a \rightarrow \begin{cases} 1 & (m > 0, \quad \lambda_a \ll |m|) \\ \lambda_a/(2|m|) & (m < 0, \quad \lambda_a \ll |m|). \end{cases} \quad (2.30)$$

Essentially,  $\cos \theta_a$  for  $m < 0$  is the eigenvalue  $\lambda_a$  normalized by  $2|m|$  as long as  $\lambda_a \ll |m|$ . But  $|m|$  plays the role of a regulator. In the limit  $\lambda_a/|m| \rightarrow \infty$ , we have  $\cos \theta_a \rightarrow 1/\sqrt{2}$ , independent of  $a$  or  $m$ . Upon taking the ratio between the theories with  $m < 0$  and  $m > 0$ , the factors of  $\cos \theta_a$  associated to eigenvalues with  $|\lambda_a| \gg m$  cancel out, and hence the ultraviolet is regularized. Therefore, after taking the ratio, we finally get

$$\frac{\langle\mathbf{L}|\Omega\rangle\langle\Omega|\text{APS}\rangle}{|\langle\text{APS}|\Omega\rangle|^2} = \prod_a \left( \frac{\lambda_a}{2|m|} \right)_{\text{reg}} = |\text{Det}(\mathcal{D}_W^+)| \quad (2.31)$$

where the subscript “reg” means that the product is regularized by  $\cos\theta_a$  in the way just mentioned. Note that the determinant of the nonchiral Dirac operator  $\mathcal{D}_W$  would have a factor of  $\lambda_a^2$  for each pair of modes, while in eqn. (2.31), there is just one factor of  $\lambda_a$  for each pair. That is why the right hand side of eqn. (2.31) is  $|\text{Det } \mathcal{D}_W^+|$  (where as before  $\mathcal{D}_W^+$  is the chiral Dirac operator on  $W$ ), not  $\text{Det } \mathcal{D}_W$ . Note that the  $\lambda_a$  are all positive, and that  $|\text{Det } \mathcal{D}_W^+|$  is the same as  $|\text{Det } \mathcal{D}_W|^{1/2}$ .

Combining this result with eqn. (2.11) and with what we learned in section 2.2, it follows that the total partition function of the bulk massive fermion  $\Psi$  with the boundary condition  $\mathbf{L}$  is given by

$$Z(Y, \mathbf{L}) = |\text{Det}(\mathcal{D}_W^+)| \exp(-\pi i \eta_D). \quad (2.32)$$

as we claimed in eqn. (2.13), and as suggested by the Dai-Freed theorem [14, 36].

## 2.4 The General Case

We have so far described anomaly inflow for a fermion all of whose components are charged under a  $U(1)$  symmetry. Here we want to generalize the construction to describe anomaly inflow for an arbitrary set of relativistic boundary fermions. For this, we will start with an arbitrary relativistic fermion field  $\chi$  on the  $d$ -manifold  $W$ , and then find a massive fermion system on  $Y$  that is related to  $\chi$  by anomaly inflow. This problem has also been studied recently in the appendix of [17], in the context of generalizing the Dai-Freed theorem to an arbitrary system of relativistic fermions.

We will ultimately work in Euclidean signature, but the necessary conditions to define a physically sensible theory are most naturally stated in Lorentz signature. The field  $\chi$  at each point in  $W$  will take values in a vector space  $S$  that will be a representation of a group that includes Lorentz symmetries possibly together with some gauge or global symmetries. In the Lorentz group, we possibly include disconnected components, related to time-reversal and reflection symmetry; also the symmetry group that acts on the fermions is really a  $\mathbb{Z}_2$  central extension of a product of the Lorentz group with a group of gauge or global symmetries. The central  $\mathbb{Z}_2$  subgroup is generated by the operator  $(-1)^F$  that distinguishes bosons and fermions. This  $\mathbb{Z}_2$  central extension may be a spin group, a pin group, or a refinement such as  $\text{spin}_c$ . We treat all cases uniformly.<sup>13</sup>

We will denote the components of  $\chi$  generically as  $\chi^a$ , where  $a$  runs over a basis of  $S$ . The only assumption that we make about  $\chi$  is that it has a first order action of the general form

$$\frac{i}{2} \int d^d x \sqrt{g} \sum_{a,b} \chi^a \sigma_{ab}^\mu D_\mu \chi^b, \quad (2.33)$$

with some matrices  $\sigma_{ab}^\mu$ , which obey constraints that will be described in a moment. We do not make any assumption about  $S$  beyond whatever is needed to ensure the existence of a physically sensible

---

<sup>13</sup> The symmetry group of a relativistic fermion system in any dimension  $d \geq 2$  is as just described. For  $d = 1$ , relativity loses its force and there are more general possibilities. We will run into this in section 4.1.



action of this form. We have not written a mass term in eqn. (2.33) and the more interesting case (since anomalies are possible) is that the symmetries do not allow a mass term. Possible mass terms and possible additional couplings, such as couplings to scalar fields, do not modify the following analysis in any essential way.

Because  $\chi$  satisfies fermi statistics, and in view of the possibility of integration by parts, only the symmetric part of  $\sigma^\mu$  really contributes to the action, so we can assume that  $\sigma_{ab}^\mu = \sigma_{ba}^\mu$ . In Lorentz signature, the  $\chi^a$  are real (after quantization, they become hermitian operators) and the matrices  $\sigma^\mu$  are real, so as to make the action and the Hamiltonian real.<sup>14</sup> Also, in Lorentz signature, one wants  $\sigma^0$  to be a positive-definite matrix for unitarity (positivity of  $\sigma^0$  ensures that if  $\beta = \sum_a r_a \chi^a$  with real coefficients  $r_a$ , then after quantization,  $\{\beta, \beta\} > 0$ , consistent with positivity of the Hilbert space inner product). Here  $\sigma^0$  is the time component of  $\sigma^\mu$  in any local Lorentz frame. (In some particular local Lorentz frame, one can make a linear transformation of the  $\chi^a$  to set  $\sigma^0 = 1$ , but this statement would then only hold in that particular frame.)

The Dirac equation is

$$\sigma_{ab}^\mu D_\mu \chi^b = 0. \quad (2.34)$$

The left hand side of this equation is not valued in the same vector space  $S$  in which  $\chi$  takes values, but rather in the dual space  $\tilde{S}$ . The duality is clear from the fact that the action (which is a pairing of  $\chi^a$  with the left hand side of eqn. (2.34)) exists. The  $\sigma^\mu$  are not gamma matrices, since they map the vector space  $S$  not to itself, but rather to a dual space  $\tilde{S}$ . Fields valued in  $S$  and  $\tilde{S}$  will be represented by spinor fields  $\chi^a$  or  $\tilde{\chi}_a$  with “index up” or “index down.”

To get a dispersion relation of the appropriate form for relativistic fermions, the matrices  $\sigma^\mu$  should satisfy the following. Let  $g_{\mu\nu}$  be the spacetime metric (in a local Lorentz frame), and define  $\bar{\sigma}_\mu = g_{\mu\mu} \sigma_\mu^{-1}$ , with no summation over  $\mu$ . (Thus in Euclidean signature with  $g_{\mu\nu} = \delta_{\mu\nu}$ ,  $\bar{\sigma}_\mu$  is just the inverse of  $\sigma_\mu$ .) Since  $\sigma_\mu$  maps  $S$  to  $\tilde{S}$ , its inverse matrix  $\sigma_\mu^{-1}$  maps  $\tilde{S}$  to  $S$ , and therefore  $\bar{\sigma}_\mu$  does the same. Then we want

$$\bar{\sigma}_\mu \sigma_\nu + \bar{\sigma}_\nu \sigma_\mu = 2g_{\mu\nu}. \quad (2.35)$$

This equation makes sense: the  $\sigma_\mu$  map  $S$  to  $\tilde{S}$  and the  $\bar{\sigma}_\mu$  map in the opposite direction, so the product maps  $S$  to itself and can be a  $c$ -number. From eqn. (2.35) we can deduce that  $\chi$  obeys the expected dispersion relation, since eqn. (2.34) implies (in flat space) that

$$0 = (\bar{\sigma}^\nu \partial_\nu)(\sigma^\mu \partial_\mu) \chi = g^{\mu\nu} \partial_\mu \partial_\nu \chi. \quad (2.36)$$

Therefore we require (2.35) as a condition on  $\sigma^\mu$ . Note that the Clifford algebra of eqn. (2.35) implies Lorentz invariance with generators  $\sigma_{\mu\nu} = \frac{1}{4}(\bar{\sigma}_\mu \sigma_\nu - \bar{\sigma}_\nu \sigma_\mu)$ . These generators map  $\chi$  to  $\chi$ , so the original theory of  $\chi$  only was in fact Lorentz-invariant.

---

<sup>14</sup>The real and imaginary parts of a complex field are real, so there is no loss of generality in assuming the  $\chi^a$  to be real. If there are gauge and global symmetries, then their generators acting on  $\chi$  are real matrices. Although  $\sigma^\mu$  and the gauge and global symmetry generators are all real, in general one cannot write  $S = S' \otimes S''$  with  $S'$  and  $S''$  being real vector spaces such that the  $\sigma^\mu$  are bilinear forms on  $S'$  and the gauge and global symmetry generators act only on  $S''$ . The existence or not of such a decomposition will play no role the present paper.

As an example of this, in two dimensions with Euclidean signature,  $\sigma_1$  and  $\sigma_2$  can be the  $1 \times 1$  matrices 1 and  $i$ . Then  $\bar{\sigma}_1$  and  $\bar{\sigma}_2$  are 1 and  $-i$ . The fermion  $\chi$  has a single component, with definite chirality. In Lorentz signature,  $\chi$  is called a Majorana-Weyl fermion. In this example, the conjugate fermion  $\tilde{\chi}$  that is introduced in a moment has opposite chirality.

We now introduce a second spinor field  $\tilde{\chi}_a$  that transforms in the dual representation  $\tilde{S}$ . (In a particular example, the original representation might be self-dual so it may happen that  $\chi$  and  $\tilde{\chi}$  actually transform the same way. We proceed the same way whether this is so or not. Note that even if  $\chi$  and  $\tilde{\chi}$  transform the same way, a mass term for  $\chi$  might be forbidden by fermi statistics, so anomalies may be possible.) For  $\tilde{\chi}$  we can always write an action

$$-\frac{i}{2} \int d^d x \sqrt{g} \sum_{a,b} \tilde{\chi}_a \bar{\sigma}^{\mu ab} D_\mu \tilde{\chi}_b, \quad (2.37)$$

since the matrices  $\bar{\sigma}^\mu$  map  $\tilde{S}$  back to  $S$ . Here we have put a minus sign in the action relative to eqn. (2.33), because  $\bar{\sigma}^0 = -(\sigma^0)^{-1}$  is negative-definite in Lorentz signature. We write  $\mathcal{D}_W^+ = -\sigma^\mu D_\mu$  and  $\mathcal{D}_W^- = \bar{\sigma}^\mu D_\mu$ .

Now we can combine  $\chi$  and  $\tilde{\chi}$  to a fermi field  $\Psi = \begin{pmatrix} \chi \\ \tilde{\chi} \end{pmatrix}$ . Acting on  $\Psi$ , we define the matrices

$$\gamma_\mu = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix}. \quad (2.38)$$

These are gamma matrices  $\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}$  by virtue of eqn. (2.35).

The action for  $\Psi$  is the sum of the actions for  $\chi$  and for  $\tilde{\chi}$  plus a possible mass term:

$$-im \int d^d x \sqrt{g} (\tilde{\chi}, \chi) \quad (2.39)$$

where  $(\tilde{\chi}, \chi) = \tilde{\chi}_a \chi^a$ . We can introduce an antisymmetric bilinear form  $\langle \ , \ \rangle$  on the space  $S \oplus \tilde{S}$ , invariant under all symmetries, by

$$\langle \Psi_1, \Psi_2 \rangle = (\tilde{\chi}_1)_a (\chi_2)^a - (\chi_1)^a (\tilde{\chi}_2)_a. \quad (2.40)$$

One can verify that

$$\langle \gamma^\mu \Psi_1, \Psi_2 \rangle = -\langle \Psi_1, \gamma^\mu \Psi_2 \rangle. \quad (2.41)$$

The action can be written in the manifestly Lorentz-invariant form.

$$-\frac{i}{2} \int d^d x \sqrt{g} \langle \Psi, (\gamma^\mu D_\mu + m) \Psi \rangle. \quad (2.42)$$

The identify (2.41), together with integration by parts and the antisymmetry of  $\langle \ , \ \rangle$ , can be used to verify that the equation of motion derived from this action is the expected  $(\not{D} + m)\Psi = 0$ .

Crucially, it is also possible to add another dimension. One adds a new coordinate  $\tau$  and a new gamma matrix

$$\gamma^\tau = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.43)$$

that obeys  $(\gamma^\tau)^2 = 1$  and anticommutes with the previous gamma matrices. One can verify the analog of (2.41):

$$\langle \gamma^\tau \Psi_1, \Psi_2 \rangle = -\langle \Psi_1, \gamma^\tau \Psi_2 \rangle. \quad (2.44)$$

With a new gamma matrix, one can extend  $d$ -dimensional Lorentz symmetry to  $D = (d+1)$ -dimensional Lorentz symmetry, the new generators being  $\frac{1}{4}[\gamma^\tau, \gamma^\mu]$ . The form  $\langle \cdot, \cdot \rangle$  possesses  $D$ -dimensional Lorentz symmetry.<sup>15</sup> Then there can be a new term in the action

$$-i \int d^d x d\tau \sqrt{g} (\tilde{\chi}, D_\tau \chi) = -\frac{i}{2} \int d^d x d\tau \sqrt{g} \langle \Psi, \gamma^\tau D_\tau \Psi \rangle. \quad (2.45)$$

Combining (2.42) (or rather its integral over  $\tau$ ) and (2.45), we construct the action of the massive fermion  $\Psi$  in  $D = d + 1$  dimensions

$$-\frac{i}{2} \int d^{d+1} x \sqrt{g} \langle \Psi, (\not{D} + m) \Psi \rangle. \quad (2.46)$$

From this action, one can derive the expected  $D$ -dimensional Dirac equation  $(\not{D} + m)\Psi = 0$  (all facts used to prove this statement in dimension  $d$  also hold in dimension  $D = d + 1$ ). The action and the Dirac equation have manifest  $D$ -dimensional Lorentz invariance. Therefore, this action can be defined on any  $D$ -manifold  $Y$  that is endowed with all the structures<sup>16</sup> (such as orientations, spin structures, gauge bundles) that were needed to define the original fermion field  $\chi$  on  $W$ .

With the  $d$ -manifold  $W$  still understood to have Lorentz signature, the original  $d$  gamma matrices (2.38) are all real, and of course  $\gamma^\tau$  as defined in eqn. (2.43) is also real. So all gamma matrices are real, and therefore the Lorentz generators are also real.  $\chi$  was real to begin with, and  $\tilde{\chi}$ , transforming as the dual to  $\chi$ , is also real. So there is a physically sensible theory of a real  $\Psi$  field in  $D$  dimensions with Lorentz signature. This is important in order to make the formula we will get for anomaly inflow physically meaningful.

Thus, on a completely general  $D$ -manifold with Lorentz signature,  $\Psi$  can be considered real. What happens when we go to Euclidean signature? Then  $\Psi$  is generally no longer real, but in fact it becomes

---

<sup>15</sup>To check invariance under the additional generators, the identity we need is  $\langle \gamma^\tau \gamma^\mu \Psi_1, \Psi_2 \rangle + \langle \Psi_1, \gamma^\tau \gamma^\mu \Psi_2 \rangle = 0$ , which follows immediately from (2.41) and (2.44). Note also that if the original  $d$ -dimensional theory has a time-reversal or reflection symmetry, the  $D$ -dimensional theory has the same symmetry. This follows from Lorentz invariance in  $D$  dimensions together with the discrete symmetry in  $d$  dimensions.

<sup>16</sup>By a spin structure or gauge bundle on  $Y$ , we always mean a spin structure or gauge bundle on  $Y$  that extend the spin structure and gauge bundle of  $W$ . Along  $\partial Y$ , a specific isomorphism is chosen between the spin structure and gauge bundle of  $Y$  and those of  $W$ .

pseudoreal.<sup>17</sup> We rotate to Euclidean signature by a Wick rotation of  $W$ , say  $x^0 \rightarrow -ix_E^0$ , where  $x^0$  parametrizes a time direction in  $W$ . The corresponding transformation of gamma matrices is  $\gamma^0 = -i\gamma_E^0$ . In Lorentz signature, all gamma matrices were real, so in Euclidean signature, they are all real except for one, namely  $\gamma_E^0$ . Let  $*$  be complex conjugation. Evidently,  $*$  anticommutes with  $\gamma_E^0$  and commutes with other gamma matrices. Accordingly,

$$C = *(-i\gamma_E^0) = *\gamma^0 \quad (2.47)$$

anticommutes with all gamma matrices.  $C$  therefore commutes with the  $D$ -dimensional rotation generators  $\frac{1}{4}[\gamma_\mu, \gamma_\nu]$ , and of course it commutes with all gauge and global symmetry generators (which are real and commute with the gamma matrices). So  $C$  commutes with all symmetries.  $C$  is antilinear and satisfies

$$C^2 = -1. \quad (2.48)$$

The existence of an antilinear operator  $C$  that commutes with all symmetries and obeys  $C^2 = -1$  means that  $\Psi$  is in a pseudoreal representation of the symmetry group. (This statement follows more abstractly from the existence of the antisymmetric bilinear form  $\langle \cdot, \cdot \rangle$  that preserves all symmetries. But that reasoning would not give an explicit construction of  $C$ .) Since  $C$  is antilinear and anticommutes with gamma matrices, it commutes with the self-adjoint Dirac operator  $\mathcal{D}_Y = i\sum_{\mu=1}^D \gamma^\mu D_\mu$ . Moreover, the fact that  $C$  is invariant under all symmetries means that the construction can be made on an arbitrary  $D$ -manifold  $Y$  that admits all the relevant structures, though we started by singling out a particular “time” direction.<sup>18</sup>

Using the antilinear operator  $C$  that was just introduced, we can define a hermitian form on the space of  $\Psi$  fields,

$$(\Psi_1, \Psi_2) = \langle C\Psi_1, \Psi_2 \rangle, \quad (\Psi_1, \Psi_2)_Y = \int_Y d^D x \sqrt{g} (\Psi_1, \Psi_2). \quad (2.49)$$

Calling  $(\cdot, \cdot)$  a hermitian form means that it is linear in the second variable and antilinear in the first. This hermitian form is clearly invariant under all symmetries, since it was constructed from the invariant ingredients  $\langle \cdot, \cdot \rangle$  and  $C$ . Explicitly if  $\Psi = \begin{pmatrix} \chi \\ \tilde{\chi} \end{pmatrix}$ , then  $C\Psi = \begin{pmatrix} \bar{\sigma}^0 \tilde{\chi}^* \\ \sigma^0 \chi^* \end{pmatrix}$ , where  $*$  is complex conjugation. So

$$(\Psi, \Psi) = \sigma_{ab}^0 \chi^{a*} \chi^b - \bar{\sigma}^{0ab} \tilde{\chi}_a^* \tilde{\chi}_b. \quad (2.50)$$

Since  $\sigma^0$  is positive-definite and  $\bar{\sigma}^0$  is negative-definite, we see that  $(\Psi, \Psi)$  is positive-definite. Moreover, the identities that have been described imply that, assuming  $Y$  has no boundary,  $\mathcal{D}_Y = i\gamma^\mu D_\mu$  is

---

<sup>17</sup>In some cases,  $\Psi$  can also be given a real structure in Euclidean signature. This happens if  $\Psi$  has  $2k$  components, and the full symmetry group including rotations and gauge and global symmetries is a subgroup of  $U(k)$ . This group is a subgroup of  $Sp(2k)$  (corresponding to the pseudoreal structure of  $\Psi$  described in the text) and of  $O(2k)$  (corresponding to an additional real structure). When  $\Psi$  does have a real structure, this can be used to simplify the analysis of anomalies. However, we will focus on the pseudoreal structure that  $\Psi$  carries universally.

<sup>18</sup>Concretely, to make this construction in curved spacetime, one uses the vierbein formalism that is anyway necessary to define fermions in curved spacetime, and one uses the above formulas in a locally Euclidean frame. Because  $C$  commutes with the rotation generators in any locally Euclidean frame, the definitions of  $C$  in different locally Euclidean frames are compatible.

self-adjoint with respect to the hermitian form  $(\cdot, \cdot)_Y$ :

$$(\mathcal{D}_Y \Psi_1, \Psi_2)_Y = (\Psi_1, \mathcal{D}_Y \Psi_2)_Y. \quad (2.51)$$

Though  $\Psi$  is not real on a general  $D$ -manifold with Euclidean signature, it can be considered real on a  $D$ -manifold that is presented with a given factorization as  $W \times \mathbb{R}$ , where  $W$  has Euclidean signature (and we only care about the behavior of  $\Psi$  under symmetries that preserve this factorization). One way to explain this statement is to observe that whether the  $\mathbb{R}$  direction is Lorentzian or Euclidean does not affect the behavior under symmetries of  $W$ , and if the  $\mathbb{R}$  direction is Lorentzian, we are on a Lorentz signature manifold, so  $\Psi$  can be considered real. Since the point is important, we will give another explanation. The vector spaces  $S$  and  $\tilde{S}$  in which  $\chi$  and  $\tilde{\chi}$  take values were dual to each other when  $W$  has Lorentz signature, and this duality persists after continuation to Euclidean signature. But in Euclidean signature, the symmetry group – a double cover of the group of rotations and possible gauge or global symmetries – is compact. The dual of a representation of a compact Lie group is isomorphic to the complex conjugate representation. This means that when  $W$  has Euclidean signature, as far as symmetries of  $W$  are concerned,  $S$  and  $\tilde{S}$  can be regarded as complex conjugate vector spaces, while in Lorentz signature they were each real and were not related to each other by complex conjugation. Hence  $\chi$  and  $\tilde{\chi}$  transform as complex conjugates and  $\Psi$  can be considered real and takes values in a real vector space  $\mathbf{S}$ . Informally  $\mathbf{S} = S \oplus \tilde{S}$  (the precise statement is that  $\mathbf{S}$  is a real vector space whose complexification has such a decomposition).

We stress that there are two senses in which  $\Psi$  might be considered real. In the starting point, in Lorentz signature,  $\Psi$  is naturally real. This is important for showing that our study of anomaly inflow applies to a physically sensible quantum field theory. In Lorentz signature, for  $\Psi$  to be real just means that  $\chi$  and  $\tilde{\chi}$  are both real. But in our calculations, we will be in Euclidean signature, and then it is useful that a different real structure can be defined in the case of a  $D$ -manifold that locally is a product  $\mathbb{R} \times W$ . In this second real structure,  $\tilde{\chi}$  is proportional to the complex conjugate of  $\chi$ . The relationship between them is described more precisely presently.

The Dirac operator  $\mathcal{D}_W = \sum_{\mu=1}^d \gamma^\tau \gamma^\mu D_\mu$  (see footnote 2 for its relation to the usual Dirac operator  $i \sum_{\mu=1}^d \gamma^\mu D_\mu$ ) can be written  $\mathcal{D}_W = \begin{pmatrix} 0 & \mathcal{D}_W^- \\ \mathcal{D}_W^+ & 0 \end{pmatrix}$ , where  $\mathcal{D}_W^+ = -\sigma^\mu D_\mu$ ,  $\mathcal{D}_W^- = \bar{\sigma}^\mu D_\mu$  are the Dirac operators for  $\chi$ ,  $\tilde{\chi}$  that we had in the beginning before going to  $D$  dimensions.

Now consider the theory of the  $\Psi$  field on a  $D$ -manifold  $Y$  with boundary  $W$ , with the local boundary condition  $\mathbf{L}$  defined by  $\tilde{\chi}|_W = 0$ . For  $m < 0$ , and treating  $Y$  as a product near the boundary,  $\Psi$  has a boundary-localized mode given by the ansatz of eqn. (2.5). In particular,  $\tilde{\chi}$  vanishes identically (in the approximation that  $Y = W \times \mathbb{R}_-$ ), and  $\chi$  vanishes exponentially fast away from  $W$ . The effective theory for this boundary-localized mode is the original purely  $d$ -dimensional theory (2.33) for  $\chi$  only. This means that anomaly inflow from  $Y$ , generated by integrating out the massive field  $\Psi$ , will cancel any anomaly of the original  $d$ -dimensional theory of  $\chi$ .

To describe this anomaly inflow precisely, we would like to generalize to this situation the formula (2.13) for  $Z(Y, \mathbf{L})$ . It is not difficult to guess the generalization. First of all, integrating out the massive field  $\Psi$  will generate a factor of  $\exp(-i\pi\eta/2)$ . (When there is no  $U(1)$  symmetry, we have to use  $\eta$ , defined by summing over all eigenvalues of  $\mathcal{D}_Y$ , rather than  $\eta_D = \eta/2$ , defined by summing only over eigenvalues of positive charge. This was explained at the end of section 2.2.) Also, by fermi statistics, the kinetic operator  $\mathcal{D}_W^+ = -\sigma^\mu D_\mu$  of  $\chi$  can be viewed as an antisymmetric matrix. The path integral of  $\chi$  is its Pfaffian,  $\text{Pf}(\mathcal{D}_W^+)$ . When there is a  $U(1)$  symmetry carried by all the fermions, this Pfaffian can be viewed as the determinant of a kinetic operator that acts on a smaller set of fields (those of positive  $U(1)$  charge), but in the absence of a  $U(1)$  symmetry, the path integral is best understood as a Pfaffian. So the natural analog of eqn. (2.13) is

$$Z(Y, \mathbf{L}) = |\text{Pf}(\mathcal{D}_W^+)| \exp(-i\pi\eta/2), \quad (2.52)$$

and we will aim to justify this formula by adapting the derivation of section 2.3.

The first step is straightforward.  $Z(Y, \mathbf{L})$  can be expressed in terms of inner products by the same logic as before, leading to eqn. (2.11). Here  $\langle \text{APS} | Y \rangle$  is known from the arguments of section 2.2, and we will slightly modify the arguments of section 2.3 to compute the inner products  $\langle \mathbf{L} | \Omega \rangle$  and  $\langle \text{APS} | \Omega \rangle$ .

In general, when one quantizes fermions, to get a Hilbert space with a positive-definite inner product, the fermions that are being quantized carry a real structure. (If complex fermion fields are present, one can take their real and imaginary parts.) Real fermion fields become hermitian operators after quantization. And on the space of real fermion fields, there is a positive-definite inner product that appears in the canonical anticommutation relations. So in order to quantize the  $\Psi$  field on  $W$ , we want to describe explicitly the real structure that is appropriate if  $Y = \mathbb{R} \times W$ , and the natural positive-definite inner product in this real structure.

From our definitions, it follows immediately that if  $W$  has Lorentz signature with time coordinate  $x^0$ , and  $\gamma^0$  is the corresponding gamma matrix (defined in eqn. (2.38)), then  $-\langle \Psi_1, \gamma^0 \Psi_2 \rangle = \sigma_{ab}^0 \chi_1^a \chi_2^b - \bar{\sigma}^{0ab} (\tilde{\chi}_1)_a (\tilde{\chi}_2)_b$ . Since  $\sigma^0$  is positive-definite and  $\bar{\sigma}^0$  is negative-definite, it follows that the inner product  $(\Psi_1, \Psi_2)_0 = -\langle \Psi_1, \gamma^0 \Psi_2 \rangle$  is positive-definite, as long as  $\chi$  and  $\tilde{\chi}$  are real. We actually want to take  $W$  to be Euclidean and make a Wick rotation  $\tau = it$  of the coordinate orthogonal to  $W$ . Lorentz invariance of the pairing  $\langle \Psi_1, \Psi_2 \rangle$  means, of course, that any statement with  $x^0$  viewed as a time coordinate and  $\tau$  as a space coordinate has an analog if the roles are reversed. Thus, setting  $\gamma^t = -i\gamma^\tau$ , and now assuming  $W$  to be Euclidean, we define the inner product

$$(\Psi_1, \Psi_2)_t = -\langle \Psi_1, \gamma^t \Psi_2 \rangle = \langle \Psi_1, i\gamma^\tau \Psi_2 \rangle, \quad (2.53)$$

where  $\gamma^\tau$  was defined in eqn. (2.43). This will be positive-definite if we place on  $\Psi$  the appropriate Wick-rotated reality condition. What reality condition will do the job? The positivity of  $(\ , \ )_t$  is ensured if we impose on  $\Psi$  a reality condition such that

$$(\Psi_1, \Psi_2)_t = (\Psi_1, \Psi_2), \quad (2.54)$$

where the right-hand-side is the positive definite hermitian inner product introduced in eqn. (2.49). Notice that the left-hand-side is linear in  $\Psi_1$ , while the right-hand-side is antilinear in  $\Psi_1$ , so this equation imposes a certain reality condition on  $\Psi_1$ . Explicitly, since  $(\Psi_1, \Psi_2)_t = \langle -i\gamma^\tau \Psi_1, \Psi_2 \rangle$  and  $(\Psi_1, \Psi_2) = \langle C\Psi_1, \Psi_2 \rangle$ , the reality condition is given by

$$i\gamma^\tau C\Psi = \Psi. \quad (2.55)$$

This is a consistent reality condition, since  $C$  and  $\gamma^\tau$  anticommute and hence  $(i\gamma^\tau C)^2 = -C^2 = 1$ . More explicitly, in terms of  $\chi$  and  $\tilde{\chi}$ , eqn. (2.55) becomes  $\tilde{\chi}_a = -i\sigma_{ab}^0(\chi^b)^*$  and  $\chi^a = i\bar{\sigma}^{0ab}(\tilde{\chi}_b)^*$ , where  $*$  is complex conjugation.<sup>19</sup> This is consistent with the claim that when  $W$  has Euclidean signature,  $\tilde{\chi}$  and  $\chi$  transform as complex conjugates.

We define  $\bar{\chi}$  as

$$\bar{\chi}_a = \sigma_{ab}^0(\chi^b)^*. \quad (2.56)$$

Then the reality condition is

$$\tilde{\chi} = -i\bar{\chi}. \quad (2.57)$$

We impose this condition in canonical quantization of the system when  $W$  has Euclidean signature and  $t = -i\tau$  is the time coordinate. In particular, the product  $\bar{\chi}\chi = \bar{\chi}_a\chi^a = (\chi^b)^*\sigma_{ab}^0\chi^b$  is positive-definite and invariant under all symmetries on  $W$ . The matrix  $\sigma^0$  plays the role of a hermitian metric on  $S$ , invariant under all symmetries.

Using  $\tilde{\chi} = -i\bar{\chi}$ , the action after the Wick rotation  $\tau \rightarrow it$  is

$$I = \int d^d x dt \sqrt{g} \left[ i\bar{\chi}\partial_t\chi - m\bar{\chi}\chi + \frac{i}{2}\chi\sigma^\mu D_\mu\chi + \frac{i}{2}\bar{\chi}\bar{\sigma}^\mu D_\mu\bar{\chi} \right], \quad (2.58)$$

where  $\bar{\chi}\partial_t\chi = \bar{\chi}_a\partial_t\chi^a$ ,  $\chi\sigma^\mu D_\mu\chi = \chi^a\sigma_{ab}^\mu D_\mu\chi^b$  and so on. The Hamiltonian is

$$H = \int d^d x \sqrt{g} \left[ m\bar{\chi}\chi + \frac{i}{2}\chi\mathcal{D}_W^+\chi + (\text{h.c.}) \right] \quad (2.59)$$

where  $\mathcal{D}_W^+ = -\sigma^\mu D_\mu$  and (h.c.) is the hermitian conjugate of  $\frac{i}{2}\chi\mathcal{D}_W^+\chi$ .

We work in the subspace in which  $(\mathcal{D}_W^+)^\dagger(\mathcal{D}_W^+)\chi = \lambda_a^2\chi$ . We can assume that  $\lambda_a > 0$ , since until section 2.5, we assume there are no zero-modes. The existence of the action  $\int \chi\mathcal{D}_W^+\chi$ , consistent with

---

<sup>19</sup>This way of writing the reality condition looks noncovariant, but it actually is covariant, since eqn. (2.55) is a manifestly covariant version. As always, formulas such as  $\tilde{\chi}_a = -i\sigma_{ab}^0(\chi^b)^*$  are written in a locally Euclidean frame. With the representation of the gamma matrices that led to the explicit definition (2.47) of  $C$ , some rotation generators of the tangent space of  $W$  are real and some are imaginary; when this is taken into account, the formula  $\tilde{\chi}_a = -i\sigma_{ab}^0(\chi^b)^*$  is covariant under a change of the locally Euclidean frame, as is the definition of  $C$ . A different representation of the gamma matrices (in which  $\gamma_\tau$  is imaginary and the gamma matrices of  $W$  are real) would make the relation between  $\tilde{\chi}$  and  $\chi^*$  look more natural while obscuring the relationship to the starting point, which was a physically sensible theory of the  $\chi$  field in Lorentz signature.



fermi statistics, means that the differential operator  $\mathcal{D}_W^+$  can be regarded as an antisymmetric matrix. Thus we can put it in a block diagonal form with  $2 \times 2$  blocks of the form  $\lambda_a \epsilon_{ij}$ , where  $\epsilon_{ij}$  is the totally antisymmetric  $2 \times 2$  matrix with  $\epsilon_{12} = 1$ . We can normalize the corresponding orthonormal eigenmodes  $\psi_1$  and  $\psi_2$  such that

$$\int d^d x \sqrt{g} \bar{\psi}^i \psi_j = \delta_j^i, \quad \int d^d x \sqrt{g} \psi_i \mathcal{D}_W^+ \psi_j = i \lambda_a \epsilon_{ij}, \quad (2.60)$$

where  $\bar{\psi}_a^i = \sigma_{ab}^0 (\psi_i^b)^*$ . The phase of  $\int d^d x \sqrt{g} \psi_1 \mathcal{D}_W^+ \psi_2$  can be freely chosen by rotating the phases of  $\psi_1$  and  $\psi_2$ , and we have chosen it to be  $i$  to slightly simplify the later equations.

Using these modes, we expand  $\chi$  and  $\bar{\chi}$  as

$$\chi = A_+ \psi_1 + A_-^\dagger \psi_2, \quad \bar{\chi} = A_+^\dagger \bar{\psi}^1 + A_- \bar{\psi}^2. \quad (2.61)$$

The reason for the notation  $A_+$  and  $A_-^\dagger$  for the coefficients of the expansion of  $\chi$  is to make the Hamiltonian below to be of the same form as in the case of a Dirac fermion studied in section 2.3. By canonical quantization, the above action gives the anticommutation relations  $\{A_+, A_+^\dagger\} = 1$ ,  $\{A_-, A_-^\dagger\} = 1$ , with others zero. The Hamiltonian is

$$H = m(A_+^\dagger A_+ - A_-^\dagger A_-) + \lambda_a (A_-^\dagger A_+ + A_+^\dagger A_-) = (A_+^\dagger, A_-^\dagger) \begin{pmatrix} m & \lambda_a \\ \lambda_a & -m \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix}. \quad (2.62)$$

This is the same as we had for a Dirac fermion.

The boundary conditions have basis-independent characterizations as follows. We denote the Hamiltonian with the parameters  $m, \lambda_a$  as  $H(m, \lambda_a)$ . The ground state  $|\Omega\rangle$  is the lowest energy state of the Hamiltonian  $H(m, \lambda_a)$ . The APS boundary condition  $|\text{APS}\rangle$  is the lowest energy state of the Hamiltonian  $H(0, \lambda_a)$  (i.e. the Hamiltonian of the massless theory  $m = 0$ ). The state  $|\text{L}\rangle$  defined by the local boundary condition  $\text{L}$  is the lowest energy state of the Hamiltonian  $H(|m|, 0)$  (i.e. the Hamiltonian of the positive mass theory in the limit  $\lambda_a/|m| \rightarrow 0$ ).

Now the problem reduces to the case of the Dirac fermion studied in the previous section. The Hamiltonian as well as the boundary conditions are completely the same. The only change is that the infinite product over  $a$  now gives a Pfaffian:

$$\frac{\langle \text{L} | \Omega \rangle \langle \Omega | \text{APS} \rangle}{|\langle \text{APS} | \Omega \rangle|^2} = \prod_a \left( \frac{\lambda_a}{2|m|} \right)_{\text{reg}} = |\text{Pf}(\mathcal{D}_W^+)|. \quad (2.63)$$

Indeed, eqn. (2.60) shows that  $|\text{Pf}(\mathcal{D}_W^+)|$  is a regularized version of  $\prod_a \lambda_a$ .

Together with the fact that  $\langle \text{APS} | Y \rangle$  is the path integral of the massive fermion field  $\Psi$  on  $Y$  with APS boundary conditions, and so is equal to  $\exp(-i\pi\eta_Y/2)$ , eqn. (2.63) is what we need to justify eqn. (2.52).



## 2.5 Treatment of Zero-Modes

So far we have assumed that the boundary Dirac operator  $\mathcal{D}_W$  has no zero-modes. We have justified the anomaly inflow formula (2.52) under this assumption. As long as there are no zero-modes for generic background fields  $g, A$ , this formula gives a good characterization of anomaly inflow. (The formula remains valid if the background fields are varied so that zero-modes appear, since the left and right hand sides of eqn. (2.52) both vanish in that case.)

We need a new derivation for the case that for arbitrary background fields,  $\mathcal{D}_W$  has zero-modes. This might happen because of a nonzero index or mod 2 index that implies the existence of zero-modes.<sup>20</sup> Actually, in the anomaly inflow problem, we consider a  $d$ -manifold  $W$  that by definition is the boundary of some  $Y$ . The index and the mod 2 index of the Dirac operator on  $W$  are cobordism invariants, so they vanish.<sup>21</sup> But it may happen that  $W$  is the union of disconnected components  $W_i$ . In that case, although the overall Dirac index and mod 2 index on  $W$  will vanish, on individual components there may be a nonvanishing index or mod 2 index. That is the situation in which generically (and in fact always)  $\mathcal{D}_W$  has zero-modes. Though this situation may at first sight seem rather esoteric, it actually plays a role in understanding relatively simple examples, as we will see in section 4. Therefore, it is important to generalize the anomaly inflow formula (2.52) to cover this situation.

In doing so, we will use the fact that there is always a generic number of zero-modes, the minimum number allowed by any index or mod 2 index theorem. Let  $b_i$  be the generic number of zero-modes of the chiral Dirac operator  $\mathcal{D}_W^+$  on  $W_i$ . Then summing over all components, the generic number of zero-modes of the chiral Dirac operator  $\mathcal{D}_W^+$  on  $W$  is  $b = \sum_i b_i$ . Since (when  $W$  has Euclidean signature), complex conjugation exchanges  $\chi$  and  $\tilde{\chi}$ , the number of zero-modes of the opposite chirality Dirac operator  $\mathcal{D}_W^-$  is also  $b$ . In the anomaly inflow problem, the number  $b$  is always even, because the mod 2 index is a cobordism invariant and  $W$  is assumed to be the boundary of some  $Y$ . So we set  $b = 2\nu$ .

---

<sup>20</sup>The mod 2 index  $\zeta_W$  of the Dirac operator  $\mathcal{D}_W^+$  is the number of zero-modes of  $\chi$  mod 2. As we will explain in section 4.1, fermi statistics imply that it is a topological invariant. When  $\zeta_W \neq 0$ ,  $\mathcal{D}_W^+$  must have a zero-mode. (By complex conjugation,  $\mathcal{D}_W^-$  has the same number of zero-modes as  $\mathcal{D}_W^+$ .) An ordinary Dirac index can likewise imply the existence of zero-modes. For example, in  $d = 4$ , suppose that  $\chi$  is a Majorana fermion coupled to gravity only. Then the number of positive chirality zero-modes of  $\chi$  minus the number of negative-chirality zero-modes of  $\chi$  is the Dirac index  $\mathcal{I}$ ; when it is nonzero, again  $\mathcal{D}_W^+$  must have zero-modes.

<sup>21</sup>Cobordism invariance of the mod 2 index  $\zeta_W$  of  $\mathcal{D}_W^+$  is an easy consequence of the anomaly inflow construction. Because the equation  $(\not{D}_Y + m)\Psi = 0$  on  $Y$  with the local boundary condition  $\mathbf{L}$  that leads to anomaly inflow can be derived from an action consistent with fermi statistics, the number  $\zeta_Y$  of zero-modes of this equation mod 2 is a deformation invariant and in particular independent of  $m$  (see section 4.1 for this argument). Taking  $m \gg 0$ , the equation  $(\not{D}_Y + m)\Psi = 0$  has no approximate solutions on  $Y$  that satisfy the boundary condition, so  $\zeta_Y = 0$ . Taking  $m \ll 0$ , the approximate solutions of the equation are the same as the zero-modes of  $\chi$  on  $W$ , so  $\zeta_W = \zeta_Y$  and hence  $\zeta_W = 0$ . Note that zero-modes of  $\mathcal{D}_W^+$  may not correspond to *exact* zero-modes of  $\not{D}_Y + m$ , but small corrections to the spectrum do not affect the number of zero-modes of an antisymmetric matrix mod 2, and so do not affect the statement that  $\zeta_W = \zeta_Y$ . For the ordinary index  $\mathcal{I}$ , the Atiyah-Singer theorem gives a formula  $\mathcal{I} = \int_W \Phi$  for some characteristic class  $\Phi$ . If  $W = \partial Y$  and the structures needed to define  $\Phi$  extend over  $Y$  then  $\mathcal{I} = \int_W \Phi = \int_Y d\Phi = 0$ , so  $\mathcal{I}$  is likewise invariant under cobordism. This statement can also be proved without knowing the index formula by adapting the proof we explained for the mod 2 index.

Generalizing eqn. (2.52), we will describe a formula that characterizes anomaly inflow as long as  $\mathcal{D}_W^+$  has the generic number  $2\nu$  of zero-modes. When the background fields are varied so that additional zero-modes appear, the formula will remain valid but will become a trivial identity  $0 = 0$ .

We denote orthonormal bases of the spaces of zero-modes of  $\mathcal{D}_W^+$  as  $\psi_i$ , where  $i = 1, \dots, 2\nu$ . By definition, the  $\psi_i$  are modes of the positive chirality field  $\chi$ . The zero modes for the negative chirality field  $\tilde{\chi}$  are  $\bar{\psi}^i = \sigma^0(\psi_i)^*$ . We impose the orthonormality condition

$$\int d^d x \sqrt{g} \bar{\psi}^i \psi_j = \delta_j^i. \quad (2.64)$$

We neglect nonzero modes because their treatment is completely the same as before. We expand  $\Psi = \begin{pmatrix} \chi \\ \tilde{\chi} \end{pmatrix}$  as

$$\chi = \sum_{i=1}^{2\nu} A^i \psi_i, \quad \tilde{\chi} = -i \sum_{i=1}^{2\nu} \bar{A}_i \bar{\psi}^i. \quad (2.65)$$

When the theory is formulated on the Euclidean signature manifold  $W$  or on  $W \times \mathbb{R}$  (where  $\mathbb{R}$  parametrizes the “time” in the Hamiltonian framework), the reality condition on  $\Psi$  is  $\tilde{\chi} = -i\bar{\chi}$  and hence  $\bar{A}_i$  is the complex conjugate of  $A^i$ . Upon quantization, they become hermitian conjugate operators, the nonzero canonical anticommutators being

$$\{A^i, \bar{A}_j\} = \delta_j^i. \quad (2.66)$$

The Hamiltonian is

$$H = \int d^d x \sqrt{g} m \bar{\chi} \chi = m \sum_{i=1}^{2\nu} \bar{A}_i A^i. \quad (2.67)$$

If the mass parameter is positive  $m > 0$ , the ground state is specified by  $A^i|\Omega\rangle = 0$ , while if it is negative  $m < 0$ , the condition is  $\bar{A}_i|\Omega\rangle = 0$ . The local boundary condition  $\mathbf{L}$  is  $\langle \mathbf{L} | \bar{A}_i = 0$ , which is the same as the one for the ground state of the positive mass theory.

When we formulate the theory on a general Euclidean signature manifold  $Y$ ,  $\Psi$  is no longer real<sup>22</sup> and we consider the Dirac operator  $\mathcal{D}_Y$  acting on a complex-valued field  $\Psi$ . If  $Y$  has boundary  $W$ , it would be meaningful to say that  $\Psi$  is real along the boundary, but neither the local boundary condition  $\mathbf{L}$  nor the APS boundary condition, which we describe shortly, imposes such a constraint. We have already noted that  $\mathbf{L}$  constrains  $\bar{A}_i$  but not  $A^i$  along the boundary, and similarly the APS boundary condition will put a constraint on  $\bar{A}_i$  and  $A^i$  that is not consistent with  $\Psi$  being real along the boundary.

---

<sup>22</sup>In fact, the reality condition (2.55) cannot be extended into the bulk of Euclidean  $Y$ , because this condition involves  $(\Psi_1, \Psi_2)_t = \langle \gamma^t \Psi_1, \Psi_2 \rangle$  which is not invariant under  $D$ -dimensional Lorentz transformations in  $Y$ . It is invariant only under  $d$ -dimensional Lorentz transformations in  $W$ .

Quantum mechanically, the constraint will mean that certain non-hermitian linear combinations of  $A^i$  and  $\bar{A}_i$  annihilate the state  $|\text{APS}\rangle$  (similarly to the fact that the non-hermitian operators  $\bar{A}_i$  annihilate the state  $\langle \text{L}|$ ). Note that because of the canonical anticommutation relations (2.66), there is no state that is annihilated by a hermitian linear combination of the  $A^i$  and  $\bar{A}_j$ , since the square of such an operator is strictly positive. So no reasonable boundary condition can make  $\Psi$  real along the boundary at the classical level.

Let us proceed to the discussion of the APS boundary condition. When zero-modes are present, the APS boundary condition is subtle. Remember that in the absence of zero-modes, the state  $|\text{APS}\rangle$  associated with the APS boundary condition is the ground state of the  $\Psi$  field for  $m = 0$ . When the operator  $\mathcal{D}_W$  has zero-modes, the  $m = 0$  theory does not have a unique ground state (even up to an overall scalar multiple) but a nontrivial space of ground states, obtained by quantizing the zero-modes. Hence, to define an APS boundary condition, it is necessary to pick a particular state (or more precisely, a particular one-dimensional subspace) in the space of ground states. A completely arbitrary choice will not do, as we want the APS boundary condition to be such that the path integral of the  $\Psi$  field on  $Y$  with APS boundary conditions will generate the usual phase  $\exp(-i\pi\eta_Y/2)$ . In order for this to happen, the equation  $(\mathcal{D}_Y + im)\Psi = 0$  should be the equation of motion derived from the fermion action, and in addition  $\mathcal{D}_Y$  must be self-adjoint. These were inputs to the derivation in section 2.2 (the first condition is needed so that the fermion path integral is the Pfaffian of  $\mathcal{D}_Y + im$ , and the second makes the eigenvalues of  $\mathcal{D}_Y$  real, as assumed in the derivation). Each of these requirements is nontrivial.

First, we ask whether the equation of motion derived from the usual Dirac action  $\int_Y d^D x \sqrt{g} \langle \Psi, (\not{D} + m)\Psi \rangle$  is the Dirac equation  $(\mathcal{D}_Y + im)\Psi = 0$ , or whether the equation of motion contains additional delta function terms supported on the boundary. To avoid such terms, we need

$$\langle \Psi_1, \mathcal{D}_Y \Psi_2 \rangle_Y = \langle \mathcal{D}_Y \Psi_1, \Psi_2 \rangle_Y \quad (2.68)$$

for  $c$ -number (commuting) fermion fields  $\Psi_1, \Psi_2$ , where

$$\langle \Psi_1, \Psi_2 \rangle_Y = \int d^D x \sqrt{g} \langle \Psi_1, \Psi_2 \rangle. \quad (2.69)$$

To prove eqn. (2.68), we have to integrate by parts, and we encounter a surface term proportional to

$$- \int d^d x \sqrt{g} \langle \Psi_1, \gamma^t \Psi_2 \rangle = \int d^d x \sqrt{g} (\Psi_1, \Psi_2)_t, \quad (2.70)$$

where  $(\ , \ )_t$  was introduced in eqn. (2.53). We want this expression to vanish whenever  $\Psi_1$  and  $\Psi_2$  satisfy the APS boundary condition. Expanding

$$\Psi_1 = \begin{pmatrix} \chi_1 \\ \tilde{\chi}_1 \end{pmatrix} = \sum_{i=1}^{2\nu} \begin{pmatrix} A^i \psi_i \\ -i \bar{A}_i \bar{\psi}^i \end{pmatrix}, \quad \Psi_2 = \begin{pmatrix} \chi_2 \\ \tilde{\chi}_2 \end{pmatrix} = \sum_{i=1}^{2\nu} \begin{pmatrix} B^i \psi_i \\ -i \bar{B}_i \bar{\psi}^i \end{pmatrix}, \quad (2.71)$$

we get

$$\int d^d x \sqrt{g} (\Psi_1, \Psi_2)_t = \sum_{i=1}^{2\nu} (A^i \bar{B}_i + \bar{A}_i B^i), \quad (2.72)$$

where we have used the orthonormality condition (2.64). As noted earlier, along  $W$ ,  $\Psi$  is not constrained to be real, so in this formula,  $A^i$  and  $\bar{A}_j$  (and similarly  $B^i$  and  $\bar{B}_j$ ) should be understood as independent complex variables.

Now let us look at the condition for self-adjointness of  $\mathcal{D}_Y$ . When  $Y$  has no boundary,  $\mathcal{D}_Y$  is hermitian with respect to the hermitian form  $(\ , \ )_Y$  that was introduced in eqn. (2.49):  $(\Psi_1, \mathcal{D}_Y \Psi_2)_Y = (\mathcal{D}_Y \Psi_1, \Psi_2)_Y$ . When  $Y$  has a boundary, integration by parts generates a surface term  $\int d^d x \sqrt{g} \langle C \Psi_1, \gamma^\tau \Psi_2 \rangle$ . From the definition of  $C$  given in eqn. (2.47) along with  $\sigma^0(\psi_i)^* = \bar{\psi}^i$ ,  $\bar{\sigma}^0(\bar{\psi}^i)^* = -\psi_i$ , we get

$$\Psi_1 = \sum_{i=1}^{2\nu} \begin{pmatrix} A^i \psi_i \\ -i \bar{A}_i \bar{\psi}^i \end{pmatrix} \implies C \Psi_1 = \sum_{i=1}^{2\nu} \begin{pmatrix} -i(\bar{A}_i)^* \psi_i \\ (A^i)^* \bar{\psi}^i \end{pmatrix}. \quad (2.73)$$

Therefore, the surface term is given by

$$\int d^d x \sqrt{g} \langle C \Psi_1, \gamma^\tau \Psi_2 \rangle = \sum_{i=1}^{2\nu} ((A^i)^* B^i - (\bar{A}_i)^* \bar{B}_i). \quad (2.74)$$

For the surface terms (2.72) and (2.74) to vanish, we impose a boundary condition of the following sort. Let  $J_{ij} = -J_{ji}$  be a matrix that is antisymmetric and also unitary. Then we impose

$$\bar{A}_i = \sum_j J_{ij} A^j. \quad (2.75)$$

(with the same condition on  $B^i$ ,  $\bar{B}_i$ , since  $\Psi_1$  and  $\Psi_2$  take values in the same space.) The antisymmetry of the matrix  $J$  guarantees that the term (2.72) vanishes and hence the Pfaffian of  $\mathcal{D}_Y$  is well-defined. The unitarity of  $J$  guarantees that the term (2.74) vanishes and hence the eigenvalues are real. For the existence of an antisymmetric unitary  $J$ , it is crucial that the total mod 2 index is zero and hence the index  $i$  takes the values  $i = 1, \dots, 2\nu$ . If the number of zero modes were odd, an antisymmetric matrix necessarily would have a zero eigenvalue and could not be unitary.

There is no unique way to choose the matrix  $J_{ij}$ . That leads to the non-uniqueness of APS boundary conditions in the presence of zero modes. We allow any choice of  $J$ . But we note that it is always possible to choose a basis of zero modes  $\psi_i$  to put  $J$  in the standard form

$$J_{2i-1, 2i} = -J_{2i, 2i-1} = -1, \quad \text{other } J_{ij} = 0. \quad (2.76)$$

Quantum mechanically,  $A^i$  and  $\bar{A}_j$  become operators that satisfy the canonical anticommutation relations (2.66). The condition (2.75) means that the state  $|\text{APS}\rangle$  satisfies

$$\langle \text{APS} | (\bar{A}_i - \sum_j J_{ij} A^j) = 0. \quad (2.77)$$

The operators  $\bar{A}_i - \sum_j J_{ij} A^j$  are a maximal set of anticommuting operators constructed from the zero-modes, so these constraints are consistent and uniquely determine a linear combination of the ground states (up to a scalar multiple). On the other hand, not every linear combination of the ground states satisfies a condition of this form. The above detailed analysis has singled out a preferred class of ground states. There is, however, no way to avoid making a choice of the matrix  $J$ . For nonzero modes, there was no need to make such a choice; one simply says that the state  $|\text{APS}\rangle$  is annihilated by the negative energy modes of  $\Psi$ .

Let us go back to the Hamiltonian framework on  $W$ . Let  $|E\rangle$  be a state vector with  $A_i|E\rangle = 0$ ,  $i = 1, \dots, 2\nu$ . For notational simplicity we focus on a block of four operators  $(A^{2k-1}, A^{2k}, \bar{A}_{2k-1}, \bar{A}_{2k})$  (with some fixed value of  $k$ ) and omit  $\prod_k$  in the following equations. Then we have

$$\langle \mathbf{L} | = \langle E |, \quad \langle \text{APS} | = \langle E | (-\bar{A}_{2k} + A^{2k-1})(\bar{A}_{2k-1} + A^{2k}), \quad (2.78)$$

and

$$|\Omega\rangle = \begin{cases} |E\rangle, & m > 0 \\ \bar{A}_{2k}\bar{A}_{2k-1}|E\rangle & m < 0 \end{cases} \quad (2.79)$$

From this, we get

$$\langle \text{APS} | \Omega \rangle = 1 \quad (2.80)$$

and

$$\langle \mathbf{L} | \Omega \rangle = \begin{cases} 1 & m > 0 \\ 0 & m < 0 \end{cases} \quad (2.81)$$

This result could have been anticipated. For  $m > 0$ , there are no localized chiral fermions on the boundary and hence nothing happens. On the other hand, when  $m < 0$ , we have a localized chiral fermion whose partition function  $\text{Pf}(\mathcal{D}_W^+)$  vanishes because of the zero-modes.

To get something nonzero for  $m < 0$ , we have to insert an operator. We choose some operator  $O$  that is a function of the field  $\chi$  on  $W = \partial Y$ . Then instead of a vacuum path integral, we compute a path integral with an insertion of  $O$ . We write  $\langle O \rangle_{Y, \mathbf{L}}$  for an unnormalized path integral on  $Y$  with this insertion, and with boundary condition  $\mathbf{L}$ . Before going on, let us note that  $\langle O \rangle_{Y, \mathbf{L}}$  is completely well-defined and anomaly-free. It is an observable in the  $D$ -dimensional theory of the massive field  $\Psi$ ; because this field is massive, it admits Pauli-Villars regularization and is free of any anomaly.

We can get an illuminating formula for  $\langle O \rangle_{Y, \mathbf{L}}$  as follows. Reasoning as in the derivation of eqn. (2.9), in the long distance limit we have  $\langle O \rangle_{Y, \mathbf{L}} = \langle \mathbf{L} | O | \Omega \rangle \langle \Omega | Y \rangle$ . Choosing an APS boundary condition and repeating the derivation that led to eqn. (2.11) gives

$$\langle O \rangle_{Y, \mathbf{L}} = \frac{\langle \mathbf{L} | O | \Omega \rangle \langle \Omega | \text{APS} \rangle}{|\langle \text{APS} | \Omega \rangle|^2} \cdot \langle \text{APS} | Y \rangle. \quad (2.82)$$

In other words, the only change from the previous derivation is that  $\langle \mathbf{L} | \Omega \rangle$  is replaced by  $\langle \mathbf{L} | O | \Omega \rangle$ .

It does not matter very much which  $O$  we pick, as long as  $\langle \mathbf{L} | O | \Omega \rangle$  is generically nonzero. A minimal choice is to take  $O = \chi(x_1)\chi(x_2)\cdots\chi(x_{2\nu})$  for points  $x_1, x_2, \dots, x_{2\nu} \in W$ . The zero-mode part of  $\chi$

is  $\chi(x) = \sum_i A^i \psi_i(x)$ . The matrix element  $\langle \mathbf{L} | \chi(x_1) \dots \chi(x_{2\nu}) | \Omega \rangle$  is a product of a zero-mode factor and a factor coming from nonzero modes. The factor that comes from nonzero modes can be analyzed precisely as in sections 2.3 and 2.4 and is  $|\text{Pf}'(\mathcal{D}_W^+)|$ , where  $\text{Pf}'$  is the Pfaffian in the space orthogonal to the zero modes. The zero-mode factor is

$$\langle \mathbf{L} | \chi(x_1) \dots \chi(x_{2\nu}) | \Omega \rangle_0 = \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^{2\nu} \psi_i(x_{\sigma(i)}) \quad (2.83)$$

where the sum is over all permutations  $\sigma$  of  $2\nu$  numbers, and we have restored  $\prod_i$  in the notation. This follows from the formulas given earlier for the states  $|\mathbf{L}\rangle$  and  $|\Omega\rangle$ . Combining the two factors,

$$\langle \mathbf{L} | \chi(x_1) \dots \chi(x_{2\nu}) | \Omega \rangle = |\text{Pf}'(\mathcal{D}_W^+)| \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^{2\nu} \psi_i(x_{\sigma(i)}) . \quad (2.84)$$

The bulk contribution is exactly the same as before, namely

$$\langle \text{APS} | Y \rangle = \exp(-i\pi\eta_Y/2) , \quad (2.85)$$

where  $\eta_Y$  is computed using the chosen APS boundary conditions. In fact, we have defined this boundary condition so that  $\mathcal{D}_Y$  has all the properties that were needed in section 2.2 for the derivation of eqn. (2.85).

Combining these results, we get

$$\begin{aligned} \langle \chi(x_1) \dots \chi(x_{2\nu}) \rangle_{Y,\mathbf{L}} &= \frac{\langle \mathbf{L} | \chi(x_1) \dots \chi(x_{2\nu}) | \Omega \rangle \langle \Omega | \text{APS} \rangle}{|\langle \text{APS} | \Omega \rangle|^2} \langle \text{APS} | Y \rangle \\ &= |\text{Pf}'(\mathcal{D}_W^+)| \exp(-i\pi\eta_Y/2) \sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^{2\nu} \psi_i(x_{\sigma(i)}) . \end{aligned} \quad (2.86)$$

This is the result with a minimal operator insertion that gives a nonzero result. We could also consider some other choice of  $O$ . The only consequence is to modify the factor  $\langle \mathbf{L} | O | \Omega \rangle$ . The anomaly inflow factor  $\exp(-i\pi\eta_Y/2)$  is unchanged. This is important because – in the context of the derivation explained in section 3 – it will imply that the anomaly of the path integral of the original  $\chi$  field on  $W$  does not depend on what operator insertion is made. This is equally true whether  $\nu = 0$  – the case assumed in sections 2.3 and 2.4 – or  $\nu > 0$ , as analyzed here. The only difference is that for  $\nu = 0$ , the minimal choice of  $O$  is  $O = 1$ , so in the previous analysis, we did not introduce  $O$  explicitly.

We conclude with one last remark. One might worry that the above result depends on the choice of a basis of zero modes  $\psi_i$ , because if we transform the basis as  $\psi_i \rightarrow \sum_j \psi_j U_i^j$  with a unitary matrix  $U_j^i$ , the factor  $\sum_{\sigma} \text{sign}(\sigma) \prod_{i=1}^{2\nu} \psi_i(x_{\sigma(i)})$  is multiplied by  $\det U$ . However, if we change the basis in this way, we are also changing the APS boundary condition, since we have fixed  $J$  to the standard form

(2.76). The Dai-Freed theorem [11] states that if we change the APS boundary condition in this way, the exponentiated  $\eta$ -invariant  $\exp(-i\pi\eta_Y/2)$  changes by  $\det U^{-1}$  so that the above correlation function is independent of  $U$ . In this way the final result is independent of any choice of basis vectors or APS boundary condition. Alternatively, since in this derivation we knew at the beginning that  $\langle O \rangle_{Y,L}$  is well-defined, and the APS boundary condition was only introduced as a way to calculate it, one may see the above result as a physical proof of this part of the Dai-Freed theorem [12].

Of course, the result for  $\langle O \rangle_{Y,L}$  will in general depend on the choice of  $Y$  and of its spin structure. Here is a special case – an important one, though unfortunately somewhat technical to describe. In general, if  $W$  has  $s$  connected components  $W_1, W_2, \dots, W_s$ , and we are given a spin structure on  $Y$  whose restriction to the boundary is isomorphic to some given spin structures on  $W_1, W_2, \dots, W_s$ , then there are  $2^s$  such isomorphisms,<sup>23</sup> as a given isomorphism on any given component can be multiplied by  $(-1)^F$ . According to the definition of footnote 16 in section 2.4, the choice of such an isomorphism is part of the definition of a spin structure on  $Y$ . However, an overall  $(-1)^F$  gauge transformation on  $Y$ , changing the sign of all fermions, would “flip” the isomorphism on each of the  $W_i$ . So spin structures on  $Y$  come in groups of  $2^{s-1}$  that differ only by the chosen isomorphisms with the given spin structures on  $W_1, W_2, \dots, W_s$ . Suppose that we “flip” the isomorphism on one of the boundary components, say  $W_1$ . This is gauge-equivalent to changing the APS boundary condition on  $Y$  without changing the bases of zero-modes on  $W$ , so it multiplies  $\exp(-i\pi\eta_Y/2)$  by  $\det U^{-1}$ , with no compensating factor of  $\det U$ . Here  $U$  is the matrix that acts as  $-1$  on zero-modes supported on  $W_1$ , and  $+1$  on zero-modes supported on other components. As a result,  $\langle O \rangle_{Y,L}$  is multiplied by  $(-1)^{\zeta_1}$ , where  $\zeta_1$  is mod 2 index on  $W_1$  (the number of  $\chi$  zero-modes on  $W_1$ , mod 2). If  $(-1)^{\zeta_1}$  can be nontrivial, then the product of an odd number of  $\chi$  fields on  $W_1$  can have an expectation value in the original  $d$ -dimensional theory on  $W_1$ . Since  $\chi$  is odd under  $(-1)^F$ , this represents an anomaly in  $(-1)^F$  symmetry. This particular anomaly is reproduced in the  $D$ -dimensional formalism by the spin structure dependence that was just described.

### 3 The Anomaly

In section 2, we have obtained a general formula describing anomaly inflow for an arbitrary fermion field  $\chi$  on a manifold  $W$ . Implicit in this formula, as we will now explain, is a description of the anomaly of  $\chi$ . The anomaly inflow involved an  $\eta$ -invariant on a manifold  $Y$  with boundary  $W$ , and this will enable us to express the anomaly in  $d$  dimensions in terms of an  $\eta$ -invariant in dimension  $D = d + 1$ .

The fact that an anomaly in  $d$  dimensions is naturally related to some quantity in  $d + 1$  dimensions is familiar for perturbative anomalies. The most familiar version of the statement is that the perturbative anomaly in  $d$  dimensions is related to a Chern-Simons function in  $d + 1$  dimensions [4–6]. Our point here is to explain that nonperturbative or global fermion anomalies can be incorporated in this statement

---

<sup>23</sup>We assume that a specific isomorphism between the gauge bundle of  $\partial Y$  and that of  $W$  has been fixed. In general, if the gauge group has a nontrivial center, then  $\langle O \rangle_{Y,L}$  may depend on the choice of this isomorphism.



by just replacing the Chern-Simons function with  $\eta$ .

The relation between the global fermion anomaly in  $d$  dimensions and an  $\eta$ -invariant in  $d + 1$  dimensions was originally found in [9]. The original derivation involved computing what one might call the holonomy of a Berry connection on the determinant (or Pfaffian) line bundle and expressing this in terms of  $\eta$ . Our derivation here, inspired by the Dai-Freed theorem [11, 12] (which itself was partly inspired by the computation in [9]), is more general as it applies in any dimension, even or odd, and on any manifold, orientable or not, on which a fermion system can be defined. (The Dai-Freed theorem has been extended to this more general context in the appendix to [17].) Also, our derivation gives a more precise answer than was sought in the original work. When there is no anomaly, we determine the actual phase of the fermion path integral on  $W$ . In early work on anomalies, one aimed to show that the overall phase of the path integral was not affected by any inconsistency or anomaly, but one did not aim to get a formula for that phase. To determine the phase is much more precise than just showing that the phase can be defined. For the use of the Dai-Freed theorem and related ideas to define the absolute phase of a path integral in particular cases see [13] section 2.2, [37] section 2, and [38].

### 3.1 Defining the Phase of the Path Integral

Let us start with a fermion field  $\chi$  on a  $d$ -manifold  $W$ , and try to define the corresponding partition function. If  $\chi$  can have a bare mass, the partition function can always be defined using Pauli-Villars regularization. Even if a bare mass is not possible, Pauli-Villars regularization can always be used to define the absolute value of the fermion partition function. But in general, if a bare mass is not possible, there is no simple direct way to define the phase of the fermion partition function, and there may be an anomaly that makes it impossible to get a satisfactory definition of this phase.

However, from section 2, we know how to define the partition function of a modified system if  $W$  is the boundary of some manifold  $Y$  over which all the structures (such as a spin or pin structure and possibly a gauge bundle) needed to define the original fermion field  $\chi$  on  $W$  have been extended. We regard  $\chi$  as a boundary mode of a massive fermion field  $\Psi$  on  $Y$ , whose partition function is

$$Z(Y, \mathbb{L}) = |\text{Pf}(\mathcal{D}_W^+)| \exp(-i\pi\eta_Y/2), \quad (3.1)$$

where we put the subscript  $Y$  on  $\eta$  to indicate that it is computed on the manifold  $Y$ . This is the partition function of a combined system consisting of a massless fermion on  $W$  with anomaly inflow from a massive fermion in bulk.

If the expression (3.1) actually does not depend on the choice of  $Y$ , we can regard it as a definition of the path integral for the  $\chi$  field on  $W$ . So let us investigate the dependence on  $Y$ . (If a suitable  $Y$  does not exist at all, then it is necessary to generalize the procedure. We postpone this issue for the moment.) To investigate the dependence on  $Y$ , let  $Y'$  be another manifold with the same boundary  $\partial Y' = W$ . The ratio between the partition functions  $Z(Y, \mathbb{L})$  and  $Z(Y', \mathbb{L})$  is given by  $\exp(-i\pi(\eta_Y - \eta_{Y'})/2)$ . Let



$\overline{Y}$  be the closed manifold<sup>24</sup> which is constructed by gluing  $Y$  and the orientation reversal of  $Y'$  along their common boundary  $W$ . The gluing theorem for  $\eta$  [11] says that we have

$$\exp(-i\pi(\eta_Y - \eta_{Y'})/2) = \exp(-i\pi\eta_{\overline{Y}}/2). \quad (3.2)$$

The physical interpretation of this gluing theorem is as follows. The partition function of the massive fermion on the closed manifold  $\overline{Y}$  is represented in the path integral formulation as

$$Z(\overline{Y}) = \langle Y' | Y \rangle, \quad (3.3)$$

where  $|Y\rangle, |Y'\rangle$  are physical states on the Hilbert space  $\mathcal{H}_W$  as introduced in section 2.1. As explained in that section, these states are proportional to the ground state  $|Y\rangle \propto |\Omega\rangle$ ,  $|Y'\rangle \propto |\Omega\rangle$ , and hence we can write

$$\langle Y' | Y \rangle = \frac{\langle Y' | \text{APS} \rangle \langle \text{APS} | Y \rangle}{|\langle \text{APS} | \Omega \rangle|^2}. \quad (3.4)$$

By computing as in the section 2.2, we get the gluing formula (3.2). The universal phase in the numerator on the right hand side is  $\exp(-i\pi(\eta_Y - \eta_{Y'})/2)$  (note that reversing the orientation of  $Y'$  reverses the sign of  $\eta_{Y'}$ ) and the denominator is positive. The universal phase on the left hand side is  $\exp(-i\pi\eta_{\overline{Y}}/2)$ .

Therefore, the dependence of the partition function  $Z(Y, \mathbf{L})$  on  $Y$  is characterized by

$$\frac{Z(Y, \mathbf{L})}{Z(Y', \mathbf{L})} = \exp(-i\pi\eta_{\overline{Y}}/2). \quad (3.5)$$

If  $\Upsilon_{\overline{Y}} = \exp(-i\pi\eta_{\overline{Y}}/2)$  is always equal to 1 for any closed manifold  $\overline{Y}$ ,  $Z(Y, \mathbf{L})$  does not depend on the choice of  $Y$ . In that case,  $Z(Y, \mathbf{L})$  can serve as a definition of the path integral  $Z(W)$  for the original fermion system on  $W$ , which therefore is completely anomaly-free. If instead  $\Upsilon_{\overline{Y}}$  is nontrivial, then  $Z(Y, \mathbf{L})$  does depend on  $Y$  and cannot serve as a satisfactory definition of  $Z(W)$ . However, we should ask if there is some other definition that we should use instead of  $Z(Y, \mathbf{L})$ . This will be discussed in section 3.3. The conclusion will be to show that the anomaly can be eliminated by using a better definition if and only if  $\eta_{\overline{Y}}$  can be written as a local integral over  $\overline{Y}$ .

In the above derivation, we have implicitly assumed that the operator  $\mathcal{D}_W$  has no zero-modes, so that the anomaly can be probed by studying the partition function  $Z(Y, \mathbf{L})$ . When that is not the case, to get something nonzero, one has to consider a path integral  $\langle O \rangle_{Y, \mathbf{L}}$  with some operator insertion, as analyzed in section 2.5, and the APS boundary condition is no longer unique. However, the above derivation is still valid in this more general situation. The only properties of the APS boundary condition that are needed in the derivation are that the universal phase of  $\langle \text{APS} | Y \rangle$  is  $\exp(-i\pi\eta_Y/2)$ , and that the matrix element  $\langle \text{APS} | \Omega \rangle$  is nonzero. The APS boundary condition was chosen in section 2.5 to ensure the first property, and the second property also holds in general (eqn. (2.80)).

---

<sup>24</sup>A closed manifold is a compact manifold without boundary. In what follows,  $\overline{Y}$  is always a closed  $D$ -manifold, and  $Y$  is a  $D$ -manifold that might have a boundary.

When  $\exp(-i\pi\eta_{\overline{Y}}/2) = 1$  for any closed manifold  $\overline{Y}$  that carries the appropriate structures, this tells us in a very strong sense that the original fermion system on  $W$  was anomaly-free. Not only is there no anomaly that would prevent us from determining the phase of the path integral of the original fermion system, but we have actually determined this phase in eqn. (3.1). In sections 3.2 and 3.3, we will explore this result more fully. But first we fill a gap in the explanation that we have given so far.

In this derivation, we assumed that  $W$  is the boundary of some manifold  $Y$  over which all relevant structures are extended. What happens if a suitable  $Y$  does not exist? This actually does not mean that the original theory on  $W$  cannot be defined. It means that if it is possible to define this theory, then the definition is not unique. The condition under which it is possible to define the theory is the same as before:  $\exp(-i\pi\eta_{\overline{Y}}/2)$  should equal 1 for any closed manifold  $\overline{Y}$  carrying the appropriate structures (or a slight generalization of this described in section 3.3).

The nonuniqueness arises for the following reason. The very fact that  $W$  is not the boundary of any  $Y$  over which the appropriate structures extend means that there is a nontrivial cobordism group  $\Gamma$  whose elements are  $d$ -manifolds carrying such structures modulo those that are boundaries. The group operation in  $\Gamma$  is disjoint union of manifolds. Any homomorphism  $\varphi$  from  $\Gamma$  to  $U(1)$  gives the partition function of an “invertible” topological field theory [38–41]. This is a purely  $d$ -dimensional theory whose partition function on a manifold  $W$  is  $\varphi(W)$ . When  $\Gamma$  is nontrivial, we cannot expect to uniquely determine the partition function of the original fermion theory on  $W$  by any general arguments, because any definition that is consistent with all general principles of quantum field theory could always be modified by multiplying it by  $\varphi(W)$ . In general, different regularizations of the same theory will give results that differ by such a factor.

We can proceed as follows, as in section 2.6 of [36]. Rather than being abstract, we will assume that  $\Gamma = \mathbb{Z}_k$  for some integer  $k$ . This means that there is some manifold  $W_0$  such that  $W_0$  is not the boundary of any suitable  $Y$ , but the disjoint union of  $k$  copies of  $W_0$  is such a boundary. Writing  $W'$  for this disjoint union, we determine the partition function  $Z(W')$  from the formula (3.1). Since  $W'$  is the disjoint union of  $k$  copies of  $W_0$ , we interpret  $Z(W')$  as  $Z(W_0)^k$ . Now we define  $Z(W_0)$  as  $(Z(W'))^{1/k}$ , with some choice of the  $k^{th}$  root. It is in this choice of a  $k^{th}$  root that the nonuniqueness enters. Now given any  $W$ , since  $W_0$  generates the cobordism group, there is some integer  $r$  such that  $W''$ , defined as the disjoint union of  $W$  with  $r$  copies of  $W_0$ , is the boundary of some  $Y$ . Thus we can use eqn. (3.1) to define  $Z(W'')$ . We interpret this as  $Z(W)Z(W_0)^r$ , since  $W''$  is a disjoint union of  $W$  with  $r$  copies of  $W_0$ , and we define  $Z(W) = Z(W'')/Z(W_0)^r$ .

In this construction, changing the  $k^{th}$  root in the definition of  $Z(W_0)$  will multiply  $Z(W)$  for any  $d$ -manifold  $W$  by  $\varphi(W)$ , for some  $\varphi : \Gamma \rightarrow U(1)$ . Since  $W_0$  was assumed to generate  $\Gamma$ ,  $\varphi$  is completely determined by  $\varphi(W_0)$ .

### 3.2 Topological Field Theory And Cobordism

To understand the global anomaly more deeply, we have to return to the APS index theorem [10], which was briefly introduced in section 2.2.

Suppose that the  $d + 1$ -manifold  $\bar{Y}$  is the boundary of a  $d + 2$ -manifold  $X$ . Suppose also that we are given a fermion field  $\Psi$  on  $\bar{Y}$  with a self-adjoint Dirac operator  $\mathcal{D}_{\bar{Y}} = i \sum_{\mu=1}^{d+1} \gamma^\mu D_\mu$ , and that all structures (such as spin structures and gauge bundles) needed to define  $\mathcal{D}_{\bar{Y}}$  have been extended over  $X$ . The APS index theorem relates  $\eta_{\bar{Y}}$ , the  $\eta$ -invariant of  $\mathcal{D}_{\bar{Y}}$ , to the index  $\mathcal{I}$  of a certain Dirac operator  $\mathcal{D}_X$  on  $X$ .  $\mathcal{D}_X$  is defined using the same doubling procedure that we used in section 2.4 to go from  $W$  to  $\bar{Y}$ . In brief, we introduce a second copy  $\tilde{\Psi}$  of  $\Psi$  and a combined field  $\hat{\Psi} = \begin{pmatrix} \Psi \\ \tilde{\Psi} \end{pmatrix}$ . Acting on  $\hat{\Psi}$ , we define  $d + 2$  gamma matrices rather as in eqns. (2.38) and (2.43):

$$\Gamma^\mu = \begin{pmatrix} 0 & \gamma^\mu \\ \gamma^\mu & 0 \end{pmatrix}, \quad \Gamma^\tau = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (3.6)$$

This enables us to define a Dirac operator  $\mathcal{D}_X = i(\Gamma^\mu D_\mu + \Gamma^\tau D_\tau)$  on  $\bar{Y} \times \mathbb{R}_-$ . By Lorentz invariance of the construction, this operator can be defined on any  $X$  with boundary  $\bar{Y}$  such that the relevant structures on  $\bar{Y}$  have been extended over  $X$ .

There is one very important difference from the previous case. In section 2.4, we started with a  $d$ -dimensional fermion field  $\chi$  that did not necessarily have a self-adjoint Dirac operator. Doubling involved introducing a dual field  $\tilde{\chi}$  that did not necessarily transform the same way as  $\chi$  under the symmetry group. As a result, in general the nonzero blocks in eqn. (2.38) are not equal; they consist of distinct matrices  $\sigma_\mu$  and  $\bar{\sigma}_\mu$ . But in the present discussion, we started with a  $d + 1$ -dimensional fermion field that by hypothesis does have a self-adjoint Dirac operator, and the additional field  $\tilde{\Psi}$  that we introduced in the doubling was just a second copy of  $\Psi$ . Accordingly the nonzero blocks in the definition of  $\Gamma^\mu$  are equal. This gave more freedom in the definition of  $\Gamma^\tau$  than we had in the previous case, and we have taken advantage of that freedom in making a convenient choice of  $\Gamma^\tau$ . The extra freedom also means that in addition to  $\Gamma^\tau$ , we can define a “chirality” operator:

$$\bar{\Gamma} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.7)$$

Because  $\bar{\Gamma}$  anticommutes with all gamma matrices  $\Gamma^\mu$  and  $\Gamma^\tau$ , it anticommutes with the Dirac operator  $\mathcal{D}_X = i(\Gamma^\mu D_\mu + \Gamma^\tau D_\tau)$ . Hence we can define the index  $\mathcal{I}$  of  $\mathcal{D}_X$ : the number of zero-modes of  $\mathcal{D}_X$  with  $\bar{\Gamma} = 1$  minus the number with  $\bar{\Gamma} = -1$ . This did not have an analog when we went from  $d$  dimensions to  $d + 1$  dimensions.

We stress that this index can in general be nonzero in any dimension, even or odd. The Atiyah-Singer index theorem implies that the index of an elliptic operator on an odd-dimensional manifold without

boundary vanishes, but this is not so on an odd dimensional manifold  $X$  with boundary  $\overline{Y}$ . In general, the index of  $\mathcal{D}_X$  can be nonzero in any dimension.

So far, everything makes sense for any  $d + 1$ -dimensional fermion field  $\Psi$  with a self-adjoint Dirac operator  $\mathcal{D}_{\overline{Y}}$ . But we are actually interested in the case that  $\Psi$  was itself defined by the doubling procedure starting from some fermion field  $\chi$  in  $d$  dimensions. In that case, there is an antilinear operator  $\mathbb{C}$ , defined in eqn. (2.47), that anticommutes with all gamma matrices while commuting with all symmetries and with the Dirac operator  $\mathcal{D}_{\overline{Y}}$ . After another round of doubling, the Dirac operator  $\mathcal{D}_X$  on  $X$  has an antilinear symmetry  $\overline{\mathbb{C}}$  that also obeys  $\overline{\mathbb{C}}^2 = -1$ . The definition of  $\Gamma^\tau$  was chosen to make the definition simple:

$$\overline{\mathbb{C}} = \begin{pmatrix} \mathbb{C} & 0 \\ 0 & \mathbb{C} \end{pmatrix}. \quad (3.8)$$

Given that  $\mathbb{C}$  anticommutes with the  $\gamma^\mu$ ,  $\overline{\mathbb{C}}$  anticommutes with  $\Gamma^\mu$  and  $\Gamma^\tau$ , and hence commutes with  $\mathcal{D}_X = i(\Gamma^\mu D_\mu + \Gamma^\tau D_\tau)$ .

The existence of an antilinear transformation  $\overline{\mathbb{C}}$  that commutes with  $\mathcal{D}_X$  and squares to  $-1$  implies that all eigenvalues of  $\mathcal{D}_X$  have even multiplicity, by a kind of Kramers doubling of eigenvalues. In particular, the index  $\mathcal{I}$  of  $\mathcal{D}_X$  is even. We will see shortly why this is important.

Now let us look at the APS index formula for  $\mathcal{I}$ :

$$\mathcal{I} = \int_X \Phi_{d+2} - \frac{\eta_{\overline{Y}}}{2}. \quad (3.9)$$

If  $\overline{Y}$  is empty, this reduces to the usual Atiyah-Singer index formula  $\mathcal{I} = \int_X \Phi_{d+2}$ .  $\Phi_{d+2}$  can be expressed in terms of the gauge field strength and the Riemann tensor, as already remarked in footnote 8 of section 2.2. However, for us, the most useful characterization of  $\Phi_{d+2}$  is that it is the anomaly  $d + 2$ -form of the  $d$ -dimensional fermion  $\chi$ . This follows directly from the APS index formula. The anomaly  $d + 2$ -form  $\Phi_{d+2}$  is by definition the polynomial in the Riemann tensor and the gauge field strength whose associated Chern-Simons  $d + 1$ -form is related by anomaly inflow to the perturbative anomaly of  $\chi$ . But the formula (3.9) shows that modulo the integer  $\mathcal{I}$  (which plays no role in perturbation theory) and provided  $X$  exists (which can be assumed in perturbation theory)  $\eta_{\overline{Y}}/2$  is the Chern-Simons form associated to  $\Phi_{d+2}$ . Moreover, we have learned in section 2 that  $\eta_{\overline{Y}}/2$  is related by anomaly inflow to the perturbative (and nonperturbative) anomaly of  $\chi$ . So  $\Phi_{d+2}$  is the anomaly  $d + 2$ -form.

Therefore,  $\Phi_{d+2}$  vanishes if and only if the original theory of the  $\chi$  field in  $d$  dimensions is free of perturbative anomalies. For example, this is always the case for odd  $d$ ; in odd dimensions, there are no perturbative anomalies, and the Atiyah-Singer index theorem shows that  $\Phi_{d+2} = 0$  for odd  $d$ . For even  $d$ , perturbative anomalies are possible, of course, but many interesting theories – such as the Standard Model of particle physics – are free of them. These are again theories with  $\Phi_{d+2} = 0$ .

When there is no perturbative anomaly, there may still be a global anomaly. As we discussed in section 3.1, the anomaly is governed by  $\exp(-i\pi\eta_{\overline{Y}}/2)$  for closed  $d + 1$ -manifolds  $\overline{Y}$ . This can definitely

be nontrivial even when there is no perturbative anomaly. However, from the APS index formula, we can reach an important conclusion about the function  $\exp(-i\pi\eta_{\overline{Y}}/2)$  when  $\Phi_{d+2} = 0$ . Suppose that  $\overline{Y}$  is the boundary of a  $d+2$ -manifold  $X$  over which the relevant structures extend. The index formula then reduces to  $\eta_{\overline{Y}} = -2\mathcal{I}$ . Since  $\mathcal{I}$  is an even integer,  $\eta_{\overline{Y}}$  is an integer multiple of 4, and therefore in this situation  $\exp(-i\pi\eta_{\overline{Y}}/2) = 1$ . In other words, we have learned that when there is no perturbative anomaly, the function  $\exp(-i\pi\eta_{\overline{Y}}/2)$  that governs the global anomaly is a cobordism invariant: it is trivial on any  $\overline{Y}$  that is the boundary of some  $X$ . It may be a nontrivial cobordism invariant; if  $\overline{Y}$  is not the boundary of any  $X$ , it may happen that  $\exp(-i\pi\eta_{\overline{Y}}/2) \neq 1$ . This is precisely the case that the original theory of the  $\chi$  field in dimension  $d$  has a nontrivial global anomaly.

As a special case of what we have said,  $\exp(-i\pi\eta_{\overline{Y}}/2)$  is a topological invariant, unchanged in any continuous deformations of metrics and gauge fields. Cobordism invariance is much stronger than topological invariance.

The APS index formula and the associated cobordism invariance of the global anomaly when perturbative anomalies vanish was used in [43] (without using the term “cobordism”) for certain applications to string theory. The framework was more restrictive than that of the present paper: in relating anomalies to  $\eta$ -invariants,  $d$  was assumed to be even and all manifolds were assumed orientable. Nowadays, there is a more general framework for understanding the role of cobordism invariance. Any gapped theory with no topological order reduces at long distances (modulo nonuniversal terms that can be removed by local counterterms) to an “invertible” topological quantum field theory whose partition function is a cobordism invariant. This was conjectured in [39], and proved under some axioms of locality and unitarity in [40, 41]. A system of free massive fermions coupled to background fields is an example of a gapped system with no topological order, and our discussion establishes directly the claim about cobordism invariance at long distances for such a system.

### 3.3 Deformation Classes And Anomalies

In section 3.1, assuming that  $\Upsilon_{\overline{Y}} = \exp(-i\pi\eta_{\overline{Y}}/2)$  is trivial for every closed  $D$ -manifold  $\overline{Y}$ , we have found a satisfactory partition function for the purely  $d$ -dimensional theory of the  $\chi$  field,<sup>25</sup> namely  $Z_W^0 = |\text{Pf } \mathcal{D}_W^+| \exp(-i\pi\eta_Y/2)$ , where  $Y$  is any appropriate manifold with boundary  $W$ . We call this  $Z_W^0$  (and not just  $Z_W$ ) because we will consider some more general possibilities in a moment.

What happens if  $\Upsilon_{\overline{Y}}$  is not always trivial? Then the formula  $Z_W^0$  is not a satisfactory definition for the partition function of the  $\chi$  field as a purely  $d$ -dimensional theory, but we should ask if this formula can be improved. Let us try some other definition  $Z_W = |\text{Pf } \mathcal{D}_W^+| \exp(-i\pi\eta_Y/2) \exp(iQ_Y)$ , where to begin with  $Q_Y$  is some unknown function of the background fields  $g, A$  on  $Y$  (we can assume that  $Q$  is

---

<sup>25</sup>Refinements discussed in section 3.1, where  $\mathcal{D}_W$  has zero-modes (so that the path integral measure cannot be characterized by a partition function) or  $W$  is not a boundary (leading to a slightly more elaborate discussion), can be straightforwardly included in the following. We omit them for brevity.

real as the anomaly only affects the phase of the partition function).

In order for  $Z_W$  to be independent of the background fields  $g, A$  on  $Y$  once  $Y$  has been chosen,  $Q_Y$  must be unaffected by any variation of the background fields on  $Y$  that leaves them fixed along  $W$ . (Therefore, when presently we discuss the variation of  $Q_Y$  under a change of  $g, A$  along  $W$ , it will not matter how  $g, A$  are being changed away from  $W$ .) For  $Z_W$  to be independent of the choice of  $Y$ , we need  $\exp(iQ_Y)$  to satisfy conditions similar to those that were used in the previous analysis of  $Z_W^0$ :  $\exp(iQ_Y)$  must satisfy the same gluing law (3.2) as  $\exp(-i\pi\eta_Y/2)$ , and for a closed  $D$ -manifold  $\bar{Y}$ , we need

$$\exp(-i\pi\eta_{\bar{Y}}/2) \exp(iQ_{\bar{Y}}) = 1. \quad (3.10)$$

These conditions could be trivially satisfied with  $\exp(iQ_Y) = \exp(i\pi\eta_Y/2)$ . However, we need more. Since  $Z_W^0$  is the partition function of a physically sensible system consisting of the  $\chi$  field plus massive ( $D$ -dimensional) regulator degrees of freedom, it is manifestly physically sensible (at least in a  $D$ -dimensional sense) and we did not have to discuss what properties make it physically sensible. An important aspect is the following. Consider making a small variation of the background fields  $g, A$  along  $W$ . The variation of the logarithm of the partition function of  $\chi$  should be given by the one-point function of the stress tensor  $T$  or the current operator  $J$  of  $\chi$ . In any theory free of perturbative anomaly, these one-point functions  $\langle T \rangle$  and  $\langle J \rangle$  can be regularized in a way consistent with all physical principles, including conservation of  $T$  and  $J$ . (If there is a global anomaly, it may appear when one tries to integrate  $\delta \log Z / \delta g$  and  $\delta \log Z / \delta A$  to determine  $Z$ .) The regularization is unique up to the possibility of adding to  $T$  and  $J$  a function of the background fields that is local, gauge-invariant, and conserved. The derivation that led to the formula  $Z_W^0$  for the partition of the  $\chi$  field plus massive degrees of freedom makes it manifest that  $\delta \log Z_W^0 / \delta g$  and  $\delta \log Z_W^0 / \delta A$  are related in the expected way to  $\langle T \rangle$  and  $\langle J \rangle$ .

What happens if we replace  $Z_W^0$  with  $Z_W = Z_W^0 \exp(iQ_Y)$ ? Since  $\log Z_W = \log Z_W^0 + iQ_Y$ , this shifts  $\delta \log Z / \delta g$  and  $\delta \log Z / \delta A$  by  $i\delta Q_Y / \delta g$  and  $i\delta Q_Y / \delta A$ . For  $Z_W$  to be a physically sensible candidate formula for a partition function of the  $\chi$  field, it must be possible to interpret  $i\delta Q_Y / \delta g$  and  $i\delta Q_Y / \delta A$  as contributions to  $T$  and  $J$ . So  $\delta Q_Y / \delta g$  and  $\delta Q_Y / \delta A$  must be gauge-invariant, local (and conserved) functions of the background fields along  $W$ .

For this to be the case,  $Q_Y$  must be the integral over  $Y$  of some local operator  $\Phi$ :  $Q_Y = \int_Y \Phi$ .  $\Phi$  must be such that  $Q_Y$  respects all symmetries of the theory, including possible time-reversal or reflection symmetries. Eqn. (3.10) tells us that in the case of a closed manifold  $\bar{Y}$ ,

$$\exp(-i\pi\eta_{\bar{Y}}/2) = \exp\left(-i \int_{\bar{Y}} \Phi\right), \quad (3.11)$$

so  $\exp(-i \int_{\bar{Y}} \Phi)$  is a topological invariant and in fact a cobordism invariant. For a local operator  $\Phi$  to have that property,  $\Phi$  must be a  $D$ -form constructed as a gauge-invariant polynomial in the Riemann tensor  $R$  and the gauge field strength  $F$ , plus a possible exact form  $d\Lambda$  (here  $\Lambda$  is a gauge-invariant  $(D-1)$ -form, locally constructed from  $g, A$ ). Adding  $d\Lambda$  to  $\Phi$  will modify  $Z_W$  by  $Z_W \rightarrow$



$Z_W \exp(i \int_Y d\Lambda) = Z_W \exp(i \int_W \Lambda)$ . This is equivalent to adding to the action of the original theory on  $W$  a  $c$ -number term  $-i \int_W \Lambda$  constructed from the background fields. That does not affect the consistency of the theory, so an exact term in  $\Phi$  is not important. So in short, the case that  $\Phi$  can help in eliminating an anomaly is that  $\Phi$  is a polynomial in  $R$  and  $F$ .  $\Phi$  is supposed to be a  $D$ -form, so this is only possible if  $D$  is even. For such a  $\Phi$ ,  $\int_{\overline{Y}} \Phi$  is a characteristic class.

Whenever  $\Upsilon_{\overline{Y}} = \exp(-i\pi\eta_{\overline{Y}}/2)$  can be expressed as in eqn. (3.11) in terms of a characteristic class  $\int_{\overline{Y}} \Phi$ , we can define a purely  $d$ -dimensional partition function for the field  $\chi$ . Under these assumptions,

$$Z_W = |\text{Pf } \mathcal{D}_W^+| \exp(-i\pi\eta_Y/2) \exp\left(i \int_Y \Phi\right) \quad (3.12)$$

depends only on  $W$  and not on  $Y$ .<sup>26</sup>

When  $Y$  has boundary  $W$ ,  $CS(g, A) = \int_Y \Phi$  is a generalized Chern-Simons function of the fields  $g, A$  on  $W$ , and is independent of how those fields are extended over  $Y$  (and of the choice of  $Y$ ) modulo the values of  $\int_{\overline{Y}} \Phi$  for closed manifolds  $\overline{Y}$ . (This is proved by using the analog of eqn. (3.2) with  $\pi\eta_Y/2$  replaced by  $\int_Y \Phi$ .) If  $\int_{\overline{Y}} \Phi$  takes values in  $2\pi\mathbb{Z}$ , then  $\exp(iCS(g, A))$  depends only on the background fields on  $W$ , and not on anything about  $Y$ . In that case, we say that  $CS(g, A)$  is a properly normalized Chern-Simons action; it makes sense as a purely  $d$ -dimensional coupling. However, eqn. (3.11) tells us that whenever  $\Upsilon_{\overline{Y}}$  is nontrivial to begin with,  $\int_{\overline{Y}} \Phi$  is not valued in  $2\pi\mathbb{Z}$  and  $CS(g, A)$  is not a properly normalized Chern-Simons action. It does not make sense by itself as a  $d$ -dimensional action; rather, it is being used to cancel an anomaly. The classic example of this situation is the “parity anomaly” for odd  $d$  [15], in which  $CS(g, A)$  is  $\frac{1}{2}$  of a properly normalized Chern-Simons coupling.<sup>27</sup> We will discuss that case from the present point of view in section 4.2.

Whenever there is a  $\Phi$  that satisfies (3.11), it is not unique, because we can always shift  $\Phi \rightarrow \Phi + \Phi'$ , where  $\int_{\overline{Y}} \Phi'$  is a characteristic class normalized to take values in  $2\pi\mathbb{Z}$ . This will have the effect of shifting the effective action on  $W$  by a properly quantized, physically sensible Chern-Simons function of the background fields. This shift can be interpreted as the result of using a different regularization of the underlying theory.

Now let us discuss the same subject from the point of view of the massive theory in the  $D$ -dimensional bulk. The massive  $\Psi$  field has a single ground state when quantized on any  $(D - 1)$ -manifold, so at long distances the theory of the  $\Psi$  field becomes an “invertible topological field theory.” The partition function of this invertible topological field theory on a closed manifold  $\overline{Y}$  is the cobordism invariant  $\exp(-i\pi\eta_{\overline{Y}}/2)$ . Let us ask what are the moduli of this theory as an invertible topological field theory. In general, a first order deformation with small parameter  $\varepsilon$  of any quantum field theory multiplies the

---

<sup>26</sup>Actually, there is one last requirement to make the definition (3.12) physically sensible:  $\Phi$  should be odd under reflections, to ensure reflection positivity of the theory. Accordingly, the case that  $\Phi$  is the polynomial in  $R$  related to the Euler characteristic is not satisfactory. Reflection positivity implies that the Euler characteristic should appear in the Euclidean signature effective action with a real coefficient, not an imaginary one.

<sup>27</sup>Somewhat similar are fractional Chern-Simons counterterms that arise in certain situations [42].

partition function by  $\exp(i\varepsilon \int_Y \langle \mathcal{O} \rangle)$  for some local operator  $\mathcal{O}$ . In an invertible topological field theory,  $\mathcal{O}$  is just a function of background fields  $g, A$  (since there are no other local operators), and so  $\langle \mathcal{O} \rangle$  reduces to the classical value of  $\mathcal{O}$  in the given background fields. For  $\exp(i\varepsilon \int_Y \mathcal{O})$  to be a topological invariant (and in fact a cobordism invariant), the exponent must be a characteristic class. In other words, the moduli of an invertible topological field theory precisely correspond to the possibility of multiplying the partition function by  $\exp(i \int_Y \Phi)$ , where as before  $\Phi$  is a gauge-invariant polynomial in the curvature and field strength.

We are free here to take  $\Phi = \sum_i t_i \Phi_i$ , where  $\Phi_i$  are a basis of possible gauge-invariant polynomials and  $t_i$  are arbitrary real numbers. In particular, we can interpolate continuously between any given  $\Phi$  and  $\Phi = 0$ . Therefore, eqn. (3.11) is precisely the condition under which the invertible topological field theory with partition function  $\exp(-i\pi\eta_Y/2)$  can be deformed, through a family of invertible topological field theories, to a trivial theory. In other words, the anomaly of the original theory in  $d$  dimensions is really associated not with the invertible topological field theory associated to  $\exp(-i\pi\eta_Y/2)$  but with the deformation class of this theory. There is no anomaly if and only if this invertible topological field theory is deformable to a trivial one.

This result is natural from the point of view of condensed matter physics. In the discussion of symmetry protected topological (SPT) phases [19–21], we consider two phases to be equivalent if they can be continuously deformed to each other without any phase transition while preserving the relevant symmetries. In particular, since the coefficients with which the polynomials  $\Phi_i$  appear in the effective action can be varied continuously, these coefficients are not universal and are not part of the classification of SPT phases. The SPT theory of the massive field  $\Psi$  is considered nontrivial if the invertible topological field theory associated to it, with partition function  $\exp(-i\pi\eta_Y/2)$ , is not deformable to a trivial theory. As we have seen, this is the case in which the theory of the  $\chi$  field cannot be defined as a purely  $d$ -dimensional theory.

## 4 Examples In Dimensions $d = 1, 2, 3, 4$

In this section, we consider some examples of the use of the  $\eta$ -invariant to analyze global anomalies in dimensions  $d = 1, 2, 3, 4$ .

The cases of odd  $d$  have a different flavor, because for odd  $d$  there are no perturbative anomalies, and all anomalies are global from the beginning. We will consider the odd  $d$  cases first.

To keep our examples simple, we will primarily work on orientable manifolds only; in other words, we will generally not incorporate time-reversal or reflection symmetry. When  $d$  is odd, the global anomaly is related to an  $\eta$ -invariant in an even dimension  $D = d+1$ . On an even-dimensional orientable manifold, the  $\eta$ -invariant reduces to a more familiar invariant – an index or a mod 2 index. Therefore our odd  $d$  examples could be described more directly in terms of the index or the mod 2 index rather than the



$\eta$ -invariant. (See [43] and [14] for such a treatment of the  $d = 1$  and  $d = 3$  examples, respectively.) However, it seems useful to explain that all examples can be deduced from the  $\eta$ -invariant.

In even  $d$ , we consider examples in which perturbative anomalies cancel, and analyze the global anomalies using the  $\eta$ -invariant.

## 4.1 $d = 1$

In  $d = 1$ , we consider a system of  $n$  real fermions with a classical  $O(n)$  symmetry and a simple action

$$I = \int dt \frac{i}{2} \sum_{j=1}^n \chi_j \frac{d}{dt} \chi_j. \quad (4.1)$$

For simplicity, to begin with we take  $n$  even, do not incorporate time-reversal symmetry, and consider background  $SO(n)$  gauge fields only. We briefly explain at the end what happens if one relaxes one of those conditions.

A compact 1-manifold  $W$  will have to be a circle. The circle has two possible spin structures, as  $\chi$  could be either periodic or antiperiodic in going around the circle. We will refer to these spin structures as Ramond (R) or Neveu-Schwarz (NS).

The Hamiltonian derived from the action  $I$  simply vanishes. Though we could add additional terms to the action to get a nonzero Hamiltonian, instead we will turn on a background  $SO(n)$  gauge field on the circle. We write  $U$  for the holonomy of this gauge field. In its canonical form,  $U$  is block diagonal with  $2 \times 2$  blocks of the form

$$\begin{pmatrix} \cos \theta_k & \sin \theta_k \\ -\sin \theta_k & \cos \theta_k \end{pmatrix}, \quad k = 1, \dots, n/2. \quad (4.2)$$

As this formula makes clear, shifting any of the  $\theta_k$  by an integer multiple of  $2\pi$  does not change  $U$  and so is a gauge transformation of the background gauge field.

Upon quantization, the  $\chi_k$  satisfy  $\{\chi_k, \chi_{k'}\} = \delta_{kk'}$  so (up to a factor of  $\sqrt{2}$ ) they are gamma matrices. Hence the group that acts on the quantum Hilbert space  $\mathcal{H}$  is not the classical symmetry group  $SO(n)$  but its double cover  $\text{Spin}(n)$ . The fact that the group that acts quantum mechanically is a double cover of the classical symmetry group is an anomaly of sorts. To see that this can be understood as an anomaly in the conventional sense – an ill-definedness of the path integral – we consider the path integral on a circle in the presence of a background gauge field with holonomy  $U$ .

The path integral with antiperiodic boundary conditions for the fermions (NS spin structure) computes  $\text{Tr}_{\mathcal{H}} U$ . Using the explicit description of  $\mathcal{H}$  as a spinor representation of  $SO(n)$ , we can write a

formula for this trace:

$$\mathrm{Tr}_{\mathcal{H}} U = \prod_{k=1}^{n/2} 2 \cos(\theta_k/2). \quad (4.3)$$

Here we see the anomaly: shifting any one of the  $\theta_k$  by  $2\pi$  is a gauge transformation of the background gauge field, but it changes the sign of the path integral.

The path integral with periodic boundary conditions for the fermions (R spin structure) computes  $\mathrm{Tr}_{\mathcal{H}} (-1)^F U$ . Again we can write an explicit formula:

$$\mathrm{Tr}_{\mathcal{H}} (-1)^F U = \pm \prod_{k=1}^{n/2} 2i \sin(\theta_k/2). \quad (4.4)$$

Again we see the anomaly, but now there is a new ingredient: the overall sign of the path integral depends on an arbitrary choice. One way to explain this fact is the following. The operator  $(-1)^F$  is characterized by the fact that it anticommutes with the elementary fermions, and its square is 1. But these conditions do not determine the overall sign of the operator  $(-1)^F$ , and without more input there is no natural way to fix this sign. We can define  $(-1)^F$  up to sign by the product  $\chi_1 \chi_2 \cdots \chi_n$ , but as the  $\chi_k$  anticommute with each other, there is no way to determine the overall sign of this expression without knowing something about how the  $\chi_k$  should be ordered. We run into the same issue from the point of view of path integrals. Consider the system of  $n$  real fermions  $\chi_k$  with periodic spin structure in the special case  $U = 1$ . Each of the  $\chi_k$  has a zero-mode; let us call these modes  $\chi_k^0$ . The measure for the fermion zero-modes is, up to sign,  $d\chi_1^0 d\chi_2^0 \cdots d\chi_n^0$ . But again, the sign of this measure depends on how we order the  $d\chi_k^0$ .

Now let us look at this anomaly from the perspective of the  $\eta$ -invariant. Applied to a 1-component real fermion in dimension  $d = 1$ , the doubling procedure of section 2.4 produces a 2-component real fermion field  $\Psi$  in dimension  $D = d + 1 = 2$ . So starting with the  $n$  fields  $\chi_1, \dots, \chi_n$  in dimension 1 transforming in the fundamental representation of  $O(n)$ , we get in 2 dimensions an  $n$ -component Majorana fermion field  $\Psi$  in the fundamental representation of  $O(n)$ .

To study the anomaly by our general procedure, we regard the circle  $W$  (at least in the NS case; see below for the R case) as the boundary of a two-manifold  $Y$ , over which the spin structure of  $W$  is extended. Since we do not incorporate time-reversal symmetry as part of the discussion, we can consider the original circle  $W$  to be oriented, and then in the procedure of section 2.4, we consider only oriented  $Y$ . The anomaly is given by  $\exp(-i\pi\eta_{\overline{Y}}/2)$ , where  $\eta_{\overline{Y}}$  is the  $\eta$ -invariant of the self-adjoint Dirac operator  $\mathcal{D}_{\overline{Y}}$  of the  $\Psi$ -field on a closed manifold  $\overline{Y}$ .

On an even-dimensional orientable manifold  $Y$ , there is always a chirality operator  $\overline{\gamma}$  (usually called  $\gamma_5$  in four-dimensional particle physics) that anticommutes with the self-adjoint Dirac operator  $\mathcal{D}_Y$ . Accordingly, the nonzero eigenvalues of  $\mathcal{D}_Y$  are equal and opposite in pairs: if  $\mathcal{D}_Y \Psi = \lambda \Psi$ , then

$\mathcal{D}_Y(\bar{\gamma}\Psi) = -\lambda\bar{\gamma}\Psi$ . From the definition of the  $\eta$ -invariant

$$\eta_Y = \lim_{\epsilon \rightarrow 0} \sum_k \exp(-\epsilon|\lambda_k|) \text{sign}(\lambda_k) \quad (4.5)$$

(where the function  $\text{sign}(x)$  was defined in eqn. (2.19)), we see at once that a pair of eigenvalues  $\lambda, -\lambda$  do not contribute. Therefore, in this situation,  $\eta_Y$  simply equals the number of linearly independent zero-modes of  $\mathcal{D}_Y$ .

A zero-mode of  $\mathcal{D}_Y$  can have either positive or negative chirality. In  $d = 2$ , complex conjugation exchanges the two types of mode. So the number of zero-modes of  $\mathcal{D}_Y$  is 2 times the number of zero-modes of positive chirality.

It is physically sensible in two dimensions to consider a fermion field  $\psi$  of positive chirality only, transforming in a real representation of some symmetry group. For our case, the relevant real representation is the vector representation of  $\text{SO}(n)$  or  $\text{O}(n)$ . Such a field can have an action

$$\int d^2x \sqrt{g} (\psi, \mathcal{D}_Y^+ \psi), \quad (4.6)$$

where  $\mathcal{D}_Y^+$  is the Dirac operator acting on a fermion field of positive chirality. By fermi statistics, we can here think of  $\mathcal{D}_Y^+$  as an antisymmetric matrix. The canonical form of such a matrix is

$$\begin{pmatrix} 0 & a_1 & & & & \\ -a_1 & 0 & & & & \\ & & 0 & a_2 & & \\ & & -a_2 & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ & & & & & & 0 \end{pmatrix}, \quad (4.7)$$

with skew “eigenvalues”  $a_i$  that appear in  $2 \times 2$  blocks, and unpaired zero-modes. The only way that the number of zero-modes can change is that one of the  $a_i$  can become zero or nonzero. When that happens, the number of zero-modes jumps by  $\pm 2$ . So if  $\zeta$  is the number of zero-modes mod 2,  $\zeta$  is invariant in any continuous deformation.  $\zeta$  is called the mod 2 index of the chiral Dirac operator on  $Y$ . The derivation shows that any fermion system with an action consistent with fermi statistics has a mod 2 index (this invariant is not always interesting as for many fermion systems it identically vanishes or is the mod 2 reduction of a more familiar invariant). In our problem of the chiral fermion field  $\psi$  in the vector representation of  $\text{SO}(n)$  or  $\text{O}(n)$ , this means that the number  $\zeta$  of zero-modes of  $\psi$  mod 2 is a topological invariant: it is unchanged if one varies the metric of  $Y$  or the background gauge field.

Since  $\eta_Y$  is twice the number of zero-modes of  $\psi$ , we have  $\eta_Y = 2\zeta \bmod 4$ . Hence  $\exp(-i\pi\eta_Y/2) = (-1)^\zeta$ . So the anomaly reduces to a sign  $\pm 1$ , as we saw more directly in eqn. (4.3). To evaluate the anomaly more explicitly, first note that an  $\text{SO}(n)$  bundle  $E \rightarrow \bar{Y}$ , for any closed two-manifold  $\bar{Y}$ , is

classified topologically by its second Stieffel-Whitney class  $w_2(E)$ . Hence the anomaly can only depend on  $w_2(E)$ . Furthermore, the structure group of  $E$  can be reduced to an  $\text{SO}(2)$  subgroup. (See the end of section 4.4 for this argument.) We have  $\text{SO}(2) \cong \text{U}(1)$ . From the point of view of  $\text{U}(1)$ , the vector representation of  $\text{SO}(n)$  consists of  $n - 2$  neutral fermions and two components of charge  $\pm 1$ . In the field of a  $\text{U}(1)$  gauge field with first Chern class  $k$ , generically one of the charged components has  $|k|$  zero-modes, and the other has none. Since we assume  $n$  even, the  $n - 2$  neutral fermions do not contribute to the mod 2 index. So  $\zeta$  is the mod 2 reduction of  $k$ . But the mod 2 reduction of  $k$  coincides with  $w_2(E)$ . So  $\zeta = w_2(E)$  and the anomaly is  $(-1)^{w_2(E)}$ . In particular, the anomaly does not depend on the spin structure and could have arisen in a purely bosonic theory. Since  $w_2(E)$  is the obstruction to lifting the structure group of  $E$  from  $\text{SO}(n)$  to its double cover  $\text{Spin}(n)$ , the anomaly would disappear if we view the original system (4.1) as a system with  $\text{Spin}(n)$  rather than  $\text{SO}(n)$  symmetry. This is consistent with the observation that we made at the outset: the Hilbert space  $\mathcal{H}$  of this theory furnishes a representation of  $\text{Spin}(n)$ , not  $\text{SO}(n)$ .

Now let us consider the case of Ramond spin structure. A single circle  $W$  with Ramond spin structure is not the boundary of any spin manifold  $Y$ . But two such circles are the boundary of such a  $Y$  (we can take  $Y$  to be a cylinder). The cobordism group in this problem is  $\mathbb{Z}_2$ ; for a generator, we can pick a circle  $W_0$  with Ramond spin structure and with any chosen holonomy  $U_0$  for the background gauge field. Since  $W_0$  is not a boundary, our formalism gives no natural way to compute the sign  $\text{Tr}(-1)^F U_0$ . But once we fix this sign, any other Ramond sector path integral is determined by the procedure explained at the end of section 3.1. Thus the Ramond sector path integral is uniquely determined up to an arbitrary overall sign, as we saw more directly in eqn. (4.4).

Finally, we briefly consider three generalizations that were mentioned at the outset.

(1) For odd  $n$ , consider the path integral of the theory (4.1) on  $W = S^1$  with a spin structure of R type. With such a spin structure, each of  $\chi_1, \dots, \chi_n$  has a zero-mode, so to get a nonzero path integral, we need to insert the product of all these fields. With this insertion, we get  $\langle \text{Tr}(-1)^F \chi_1 \chi_2 \cdots \chi_n \rangle \neq 0$ . But for odd  $n$ , the operator  $\chi_1 \chi_2 \cdots \chi_n$  is odd under  $(-1)^F$ . So we have found an anomaly in  $(-1)^F$ .

Let us try to recover this anomaly from a  $D = 2$  point of view. For odd  $n$ , the derivation showing that the anomaly is governed by  $(-1)^\zeta$  is still valid. The only difference is that  $(-1)^\zeta$  can be nontrivial even if we ignore the  $\text{O}(n)$  symmetry and do not turn on any background  $\text{O}(n)$  gauge field. For an example, take  $Y$  to be a two-torus with fermions periodic in both directions. Then the Dirac operator on  $Y$  for a one-component chiral fermion  $\psi$  has a single zero-mode (the “constant” mode of  $\psi$ ), and so  $(-1)^\zeta = -1$ . Taking  $n$  identical chiral fermions does not change  $(-1)^\zeta$  if  $n$  is odd. So there is an anomaly even if the only symmetry we consider is  $(-1)^F$ . This symmetry is implicit whenever we discuss fermions and spin structures.

This anomaly in  $(-1)^F$  has a variety of applications in different areas of physics; for example, see [44–47]. An application to Type II superstring theory, closely related to the recent paper [48], was actually the subject of the lecture by one of us at the Shoucheng Zhang Memorial Workshop [3].

(2) Another generalization is to allow background gauge fields of  $O(n)$  rather than  $SO(n)$ . In discussing this, for simplicity we take  $n$  even. Let  $W$  be a circle with NS spin structure and with a background  $O(n)$  gauge field with monodromy  $R_1 = \text{diag}(-1, 1, 1, \dots, 1)$ . In this background,  $\chi_1$  has a zero-mode and  $\chi_r$ ,  $r > 1$ , does not. So again we see an anomaly in  $(-1)^F$ , this time a “mixed anomaly” between  $(-1)^F$  and the  $O(n)$  symmetry, since to detect it we had to turn on a background  $O(n)$  gauge field.

Replacing  $O(n)$  by a double cover  $\text{Pin}^+(n)$  or  $\text{Pin}^-(n)$  will not eliminate this anomaly.  $R_1$  can be lifted to  $\text{Pin}^+(n)$  or  $\text{Pin}^-(n)$ , though not uniquely, and the same analysis leads to the same anomaly in  $(-1)^F$  in the presence of a background field.

From a two-dimensional point of view, the anomaly is still governed by  $(-1)^\zeta$ . But in contrast to what we found in discussing  $SO(n)$  for even  $n$ ,  $(-1)^\zeta$  now depends on the spin structure. To see this, take  $Y$  to be a two-torus  $S_A^1 \times S_B^1$ , where the first circle  $S_A^1$  has trivial  $O(n)$  gauge field and the second circle  $S_B^1$  has an  $O(n)$  gauge field with monodromy  $R_1$ . Then  $(-1)^\zeta$  can again be computed just by counting “constant” fermion modes; it equals 1 if  $S_A^1$  has NS spin structure, and equals  $-1$  if  $S_A^1$  has spin structure of R type (regardless of the spin structure of  $S_B^1$ ). Thus there is an anomaly, and, since  $(-1)^\zeta$  depends on the spin structure, the anomaly is not given by a cohomological formula and cannot be reproduced in a purely bosonic theory. That is consistent with the fact that in the boundary description, the anomaly involves  $(-1)^F$ .

One can relate the two ways of seeing the anomaly by trying to compute  $\langle \text{Tr} (-1)^F R_1 \chi_1 \rangle^2$  via a path integral on the cylinder  $Y = S^1 \times I$ , where  $S^1$  is a circle with Ramond spin structure and  $I$  is an interval. We assume that the gauge field is a pullback from  $S^1$  and has monodromy  $R_1$ . Once we specify the spin structure on  $S^1$ , there are two possible spin structures on  $Y$ , for a reason explained at the end of section 2.5. (After picking an isomorphism between the two boundary circles and their spin bundles, one can say that the two spin bundles on  $Y$  differ by the sign of parallel transport of a fermion from one end of the cylinder to the other.) As explained at the end of section 2.5, the sign one gets for  $\langle \text{Tr} (-1)^F R_1 \chi_1 \rangle^2$  depends on the spin structure of  $Y$ , and this dependence reproduces the  $(-1)^F$  anomaly of the original  $d = 1$  theory on a single boundary component.

The theory (4.1) can certainly be quantized, with the Hilbert space providing an irreducible representation of the Clifford algebra  $\{\chi_i, \chi_j\} = \delta_{ij}$ . But after quantization, the “internal symmetry”  $R_1$  anticommutes with the “spacetime symmetry”  $(-1)^F$ , rather than commuting. As a result, the symmetry group after quantization does not fit the general framework introduced at the start of section 2.4, where  $(-1)^F$  is supposed to be central, commuting with the full symmetry group (as noted in footnote 13, there are more general possibilities in  $d = 1$ ). The symmetry group of the quantized  $d = 1$  theory does not extend naturally to  $D = 2$  (in dimension  $D \geq 2$ , the element  $(-1)^F \in \text{Spin}(D)$  is always central in any relativistic fermion theory). Hence it is not immediately obvious how to deduce the symmetry group of the quantized  $d = 1$  theory from the  $D = 2$  anomaly.

(3) To include time-reversal symmetry, we should consider  $W$  to be unoriented and allow unorientable

$Y$ . The obvious time-reversal symmetry of eqn. (4.1) that acts on the fermions by  $\mathsf{T}\chi_j(t)\mathsf{T}^{-1} = \chi_j(-t)$  satisfies  $\mathsf{T}^2 = 1$  at the classical level and corresponds to a  $\text{pin}^-$  structure in 1 dimension. To incorporate this symmetry in the boundary theory, one should allow  $Y$  to be an unorientable manifold endowed with a  $\text{pin}^-$  structure. In this case, as there is no longer a chirality operator  $\bar{\gamma}$  that anticommutes with  $\mathcal{D}_Y$ , it is no longer true that  $\exp(-i\pi\eta_{\bar{Y}}/2) = \pm 1$  for a closed manifold  $\bar{Y}$ , and instead  $\exp(-i\pi\eta_{\bar{Y}}/2)$ , for a single Majorana fermion, can be an arbitrary eighth root of 1. (This can be proved along lines explained in Appendix C of [14] for the analogous problem in four dimensions.) Indeed, in the time-reversal invariant case, the theory (4.1) that we started with does have a mod 8 anomaly from a Hamiltonian point of view [49]; that is, it is anomalous unless  $n$  is a multiple of 8. The interpretation of this in terms of cobordism was originally suggested in [39]. If  $n$  is even, it is possible to define a time-reversal symmetry such that  $\mathsf{T}^2 = (-1)^F$  at the classical level. This leads to a slightly different analysis, in which the description by the  $\eta$ -invariant again reduces to a mod 2 index.

## 4.2 $d = 3$

Our example in  $d = 3$  will actually involve one of the celebrated contributions of Shoucheng Zhang [18].

First consider in  $d = 3$  a massless Dirac fermion  $\chi$  coupled with charge 1 to a  $U(1)$  gauge field. Such a field could have a gauge-invariant bare mass, so there will be no anomaly spoiling the gauge symmetry. However, a bare mass of the  $\chi$  field would explicitly violate the time-reversal and reflection symmetries of the massless theory. So there is a possibility of an anomaly that would spoil those symmetries. Indeed, this model is the original context for the “parity” anomaly [15]: the field  $\chi$  cannot be quantized, purely in three-dimensional terms, in a way that preserves time-reversal and reflection symmetry. We essentially computed the anomaly in section 2.2, when we explained that integrating out the  $\chi$  field on a manifold  $W$ , with a particular regulator mass, generates a phase  $\exp(-i\pi\eta_{D,W}/2)$  (as we are dealing with a Dirac fermion, we use  $\eta_{D,W}$ , the  $\eta$ -invariant on  $W$  for charge 1 modes only, rather than  $\eta_W = 2\eta_{D,W}$ , the  $\eta$ -invariant on  $W$  for all modes of charge  $\pm 1$ ). The partition function of  $\chi$ , including this phase, is

$$|\text{Det } \mathcal{D}_\chi| \exp(-i\pi\eta_{D,W}/2), \quad (4.8)$$

where  $\mathcal{D}_\chi$  is the Dirac operator of the  $\chi$  field. This phase is odd under time-reversal or reflection symmetry and cannot be removed by any counterterm. Reversing the sign of the regulator mass would give the opposite phase  $\exp(i\pi\eta_{D,W}/2)$ , still violating time-reversal and reflection symmetry. The anomaly is a mod 2 anomaly, because a pair of  $\chi$  fields could be quantized with regulator masses of opposite signs, in which case the phases cancel out and all symmetries are preserved.

We are interested here in the anomalous case with a single  $\chi$  field. In the theory of topological insulators, instead of trying to quantize  $\chi$  by itself in purely three-dimensional terms, one views it as a field that propagates on the boundary of a four-manifold  $Y$ , the worldvolume of a topological insulator. The combined system can then be quantized in a way that preserves time-reversal and reflection symmetry. In this interpretation, the  $U(1)$  gauge field is just the usual gauge field of electromagnetism. The result



of Shoucheng Zhang, together with T. Hughes and X. Qi [18], was to show that the electromagnetic  $\theta$ -angle in the bulk of a topological insulator is equal to  $\pi$ . (A striking consequence of this can be explicitly demonstrated in a lattice model of a topological insulator [50]: a magnetic monopole of unit charge immersed in a  $3+1$ -dimensional topological insulator will acquire a half-integral electric charge.) Let us see how the result that  $\theta = \pi$  may be understood from the point of view of the present paper.

The doubling procedure applied to the  $\chi$  field produces a massive charge 1 Dirac fermion  $\Psi$  in dimension  $d + 1 = 4$ . With the couplings, boundary conditions, and regulator described in sections 2.1 and 2.2,  $\Psi$  provides a model of a topological insulator; in particular,  $\chi$  can be interpreted as a boundary localized mode of  $\Psi$ , and the path integral of  $\Psi$  gives a consistent framework that describes  $\chi$  together with massive degrees of freedom in the bulk. According to the general formalism, the universal part of the path integral of  $\Psi$  on a manifold  $Y$  with boundary  $W$  is

$$|\text{Det}\mathcal{D}_W^+| \exp(-i\pi\eta_{D,Y}). \quad (4.9)$$

We want to show that this formula is consistent with time-reversal and reflection symmetry, and moreover we would like to recover the result of Shoucheng Zhang and his colleagues showing that the electromagnetic  $\theta$ -angle of the bulk theory is  $\theta = \pi$ . For simplicity, we work on oriented manifolds only. The self-adjoint Dirac operator  $\mathcal{D}_Y$  anticommutes with a chirality operator  $\overline{\gamma}$ , so just as in section 4.1, its nonzero modes do not contribute to  $\eta_{D,Y}$ ;  $\eta_{D,Y}$  is simply equal to the number of linearly independent (charge 1) zero-modes of  $\mathcal{D}_Y$ . However, unlike the two-dimensional case, in four dimensions, complex conjugation does not exchange zero-modes of positive and negative chirality. On the contrary, if  $n_+$  and  $n_-$  are the numbers of linearly independent zero-modes of  $\mathcal{D}_Y$  (acting on fermions of charge 1) with positive or negative chirality, then the index of  $\mathcal{D}_Y$  is  $\mathcal{I} = n_+ - n_-$ . On the other hand,  $\eta_{D,Y}$  is the total number of zero-modes:  $\eta_{D,Y} = n_+ + n_-$ . We see that  $\eta_{D,Y} \cong \mathcal{I} \pmod{2}$ , and hence  $\exp(-i\pi\eta_{D,Y}) = (-1)^{\mathcal{I}}$ . Thus the combined path integral of the bulk and boundary modes of the topological insulator – or more precisely its universal part – is

$$|\text{Det}\mathcal{D}_W|(-1)^{\mathcal{I}}. \quad (4.10)$$

This is real, and thus manifestly consistent with time-reversal and reflection symmetry.

To understand the result of Hughes, Qi, and Zhang concerning the  $\theta$ -angle, we write  $(-1)^{\mathcal{I}} = \exp(i\pi\mathcal{I})$  and think of  $-i\pi\mathcal{I}$  as a contribution to the effective action. For discussing the bulk effective action, we can temporarily work on a closed manifold  $\overline{Y}$ . In that case the Atiyah-Singer index formula gives  $\mathcal{I} = \int_{\overline{Y}} \Phi$  with

$$\Phi = \widehat{A}(R) + \frac{1}{2} \frac{F \wedge F}{(2\pi)^2}, \quad (4.11)$$

where  $\widehat{A}(R)$  is a certain quadratic polynomial in the Riemann tensor. Because of the electromagnetic contribution to  $\Phi$ , the term  $-i\pi\mathcal{I} = -i\pi \int \Phi$  in the action corresponds to an electromagnetic  $\theta$ -angle  $\theta = \pi$ .

A key fact in this derivation was that on a closed manifold,  $\exp(-i\pi\eta_{\overline{Y},D}) = \exp(i\pi\mathcal{I}) = \exp(i\pi \int_{\overline{Y}} \Phi)$  is the exponential of the integral of a characteristic class. Therefore we are in the situation that was



analyzed in detail in section 3.3. As in eqn. (3.12), it is possible to give a purely three-dimensional formula for the partition function of  $\chi$  by replacing  $\exp(-i\pi\eta_{Y,D})$  with  $\exp(-i\pi\eta_{Y,D})\exp(-i\pi\int_Y\Phi)$ :

$$Z_W = |\text{Det } \mathcal{D}_W^+| \exp(-i\pi\eta_{D,Y}) \exp\left(-i\pi\int_Y\Phi\right) = |\text{Det } \mathcal{D}_W^+| \exp(i\pi\mathcal{I}) \exp\left(-i\pi\int_Y\Phi\right). \quad (4.12)$$

The general formalism tells us that the right hand side of eqn. (4.12) will make sense in purely three-dimensional terms. We can confirm this by using the APS index theorem:

$$\mathcal{I} = \int_Y \Phi - \frac{\eta_{D,W}}{2}. \quad (4.13)$$

So in fact eqn. (4.12) is equivalent to

$$Z_W = |\text{Det } \mathcal{D}_W^+| \exp(-i\pi\eta_{D,W}/2). \quad (4.14)$$

This is the purely three-dimensional, but time-reversal violating, formula that was explained more directly in eqn. (4.8), except that after doubling, we write  $\mathcal{D}_W^+$  for  $\mathcal{D}_\chi$ . Time-reversal violation entered this derivation when we canceled the  $Y$  dependence of  $(-1)^{\mathcal{I}} = \exp(\pm i\pi\mathcal{I})$  with a factor of  $\exp(-i\pi\int_Y\Phi)$ . Time-reversal would map this to a conjugate construction using  $\exp(+i\pi\int_Y\Phi)$ .

So we recover the familiar fact that this system can be quantized as a purely three-dimensional theory if we are willing to give up time-reversal and reflection symmetry. Alternatively, we can consider the  $\chi$  field to propagate on the boundary of a four-manifold, and use the time-reversal invariant partition function (4.10). If we wish to define the  $\chi$  field on unorientable three-manifolds, then time-reversal and reflection symmetry are essential and the purely three-dimensional quantization is not available. In this case, we have to consider the  $\chi$  field as living on the surface of a topological insulator. The appropriate formula for the partition function is eqn. (4.9), in which the  $\eta$ -invariant no longer reduces to  $\mathcal{I}$ .

### 4.3 $d = 2$

Our remaining examples will be cases with nontrivial cancellation of perturbative anomalies.

In  $d = 2$ , we consider first a  $U(1)$  gauge theory with positive chirality Dirac fermions of charges  $n_1, n_2, \dots, n_s$  and negative chirality Dirac fermions of charges  $m_1, m_2, \dots, m_s$ . The perturbative gauge theory anomaly cancels if and only if

$$\sum_{i=1}^s n_i^2 = \sum_{i=1}^s m_i^2. \quad (4.15)$$

By taking equal numbers of positive and negative chirality fermions, we have ensured cancellation of perturbative gravitational anomalies. Therefore, when eqn. (4.15) holds, this theory is free of perturbative anomalies.

A possible global anomaly would be controlled as usual by an  $\eta$ -invariant in dimension  $d+1=3$ . In the present case, let us write  $\eta_{D,r}$  for the  $\eta$ -invariant of a charge  $r$  Dirac fermion on a three-manifold  $Y$ . (We use  $\eta_D$  as we are dealing with Dirac fermions.) The global anomaly of our given theory is controlled by  $\Upsilon = \exp\left(-i\pi\left(\sum_{i=1}^s \eta_{D,n_i} - \sum_{j=1}^s \eta_{D,m_j}\right)\right)$ . (As explained at the end of section 2.2, reversing the sign of the fermion chirality in two dimensions reverses the sign with which the  $\eta$ -invariant appears in the exponent.) However, as is always the case when perturbative anomalies are absent, the factor  $\Upsilon$  that controls the global anomaly is a cobordism invariant. The cobordism of a three-dimensional closed spin manifold  $\bar{Y}$  with a  $U(1)$  gauge field is trivial (any such  $\bar{Y}$  is the boundary of some  $X$  over which the spin structure and  $U(1)$  gauge field extend). So  $\Upsilon$  always equals 1 on a closed manifold, and a theory of this kind that is free of perturbative anomalies is also always free of global or nonperturbative anomalies.<sup>28</sup>

We can get an example in  $d=2$  that does have a nontrivial global anomaly if we replace  $U(1)$  by a  $\mathbb{Z}_k$  subgroup, for some  $k$ . Then there is no perturbative anomaly as long as there are equally many positive and negative chirality fermions (to avoid a gravitational anomaly). In particular, there is no condition analogous to (4.15). For generic choices of the  $\mathbb{Z}_k$  quantum numbers of the fermions, such a theory will have a global anomaly.

We will consider a special case in a moment, but first we explain the motivation to consider this special case. Spacetime supersymmetry in string theory was originally discovered by Gliozzi, Olive, and Scherk (GOS) [51]. Their original insight was that the partition function of 8 chiral fermions in two dimensions (in genus 1, where they computed explicitly) vanishes if it is summed over spin structures; this vanishing was a reflection of spacetime supersymmetry. (8 is the number of light cone oscillator modes in the Ramond-Neveu-Schwarz model.) Though this point was not made explicitly until later, it only makes sense to add together the partition functions with different spin structures if the anomaly does not depend on the spin structure. Later discoveries involved elaborations of the original GOS analysis. One always finds that spacetime supersymmetry in string theory depends on the fact that the anomaly of 8 chiral fermions in two dimensions is independent of the spin structure. It turns out that this is true for a system of  $k$  chiral fermions if and only if  $k$  is divisible by 8.

A closely related problem has been much studied in the context of condensed matter physics. In that context, one studies “symmetry protected topological” (SPT) states [19–21], which are states that are topologically nontrivial when some global symmetry is taken into account but become trivial if that global symmetry is explicitly broken. Such a system can have an anomalous boundary state in one dimension less. A much-studied special case is a fermionic system in spacetime dimension  $D=3$  with

---

<sup>28</sup>One could avoid relying on a knowledge of the cobordism group by making an argument similar to what we will make for the Standard Model in section 4.4. A  $U(1)$  gauge field is classified topologically by its first Chern class. In three dimensions, this is dual to an embedded circle  $C \subset \bar{Y}$ . Using this, and cobordism invariance of  $\Upsilon$ , one can reduce as in section 4.4 to two special cases: (a) the gauge field is trivial, or (b)  $\bar{Y} = S^2 \times S^1$  with a gauge field is a pullback from  $S^2$ . In case (a), trivially  $\Upsilon = 1$  since the  $U(1)$  charges do not matter, and in case (b), arguing as in section 4.4 one shows that  $\Upsilon = 1$ . This argument will also work for a theory with gauge group  $U(1)^n$  for any  $n$ : if such a theory in two dimensions has no perturbative anomaly, it also has no global anomaly.

a global  $\mathbb{Z}_2$  symmetry [52–54]. Those systems have a  $\mathbb{Z}_8$  classification. An example associated to a nonzero element  $k \in \mathbb{Z}_8$  has a boundary state in dimension  $d = 2$  that carries the anomaly that one would find in the GOS calculation if the number of chiral fermions considered were  $k$  (or any integer congruent to  $k \bmod 8$ ) rather than 8.

To put this problem in the context of gauge theory, we consider a  $\mathbb{Z}_2$  gauge theory in two dimensions, with  $k$  positive chirality real fermions that are invariant under  $\mathbb{Z}_2$ , and  $k$  negative chirality real fermions that transform as  $-1$  under the nontrivial element of  $\mathbb{Z}_2$ . This system has no perturbative anomaly, but an explicit genus 1 calculation shows that if  $k$  is not divisible by 8, it has a global anomaly. We will explore from the vantage point of the present paper the absence of anomalies when  $k$  is a multiple of 8. (For a previous analysis, see the discussion of eqn. (24) in [43]. An explicit computation of the  $\eta$ -invariant when  $k$  is not a multiple of 8 to show the  $\mathbb{Z}_8$  classification was done in [55].)

Let  $W$  be a two-manifold with spin structure  $\alpha$  and some background  $\mathbb{Z}_2$  bundle, which we can think of as a real line bundle  $L$  with structure group  $\mathbb{Z}_2 = \{\pm 1\}$ . In the model just introduced, positive chirality fermions (being invariant under  $\mathbb{Z}_2$ ) are coupled only to the spin structure  $\alpha$ . But negative chirality fermions are coupled to the spin structure  $\alpha$  and also to  $L$ . Effectively the negative chirality fermions are coupled to a new spin structure  $\beta = \alpha \otimes L$ . Thus for studying anomalies, we can forget about the  $\mathbb{Z}_2$  gauge field and just say that fermions of positive or negative chirality are coupled to different spin structures  $\alpha$  or  $\beta$ .

The spin cobordism problem for a two-manifold with two spin structures  $\alpha, \beta$  (or even just with one spin structure) is not trivial. This is related to the fact that in string theory, the sign of the GOS projection in the Ramond sector is not uniquely determined and could be reversed. It is also related to the existence of two different Type II superstring theories.<sup>29</sup> We will pass over such issues here and just ask if there is any consistent way to define the theory with 8 positive chirality fermions coupled to spin structure  $\alpha$  and 8 negative chirality fermions coupled to  $\beta$ . As usual, the potential obstruction is a global anomaly that can be measured by an  $\eta$ -invariant.

In detail, let  $\bar{Y}$  be a closed three-manifold with spin structures  $\alpha, \beta$ . The global anomaly is then measured by  $\exp\left(-\frac{\pi i}{2} 8(\eta_{\bar{Y}, \alpha} - \eta_{\bar{Y}, \beta})\right)$  where  $\eta_{\bar{Y}, \alpha}$  and  $\eta_{\bar{Y}, \beta}$  are  $\eta$ -invariants on  $\bar{Y}$  for a Majorana fermion coupled to spin structure  $\alpha$  or  $\beta$ . We note that this is trivial if and only if one always has

$$\exp\left(-\frac{\pi i}{2} 8\eta_{\bar{Y}, \alpha}\right) = \exp\left(-\frac{\pi i}{2} 8\eta_{\bar{Y}, \beta}\right), \quad (4.16)$$

or in other words if and only if the anomaly for 8 positive chirality fermions in two dimensions does not depend on the spin structure. This is how we formulated the question initially.

In this form, it is not immediately obvious how to answer the question. But a more general question is easier to answer. Consider a two-dimensional theory with gauge group  $\text{Spin}(8)$ . This group has

---

<sup>29</sup>The cobordism invariant of a Riemann surface with one spin structure is  $(-1)^\zeta$ , which was already discussed, with references to various applications, at the end of section 4.1. With two spin structures  $\alpha, \beta$ , one has  $(-1)^{\zeta_\alpha}$  and  $(-1)^{\zeta_\beta}$ .

two spinor representations – spinors of positive or negative Spin(8) chirality. Each of these is a real representation of dimension 8. Let us call the two representations  $S_+$  and  $S_-$ . We consider a two-dimensional theory with gauge group Spin(8), with 8 positive chirality real fermions in the representation  $S_+$ , and 8 negative chirality fermions in the representation  $S_-$ . This theory is free of perturbative anomalies, because the representations  $S_+$  and  $S_-$  have the same dimension and quadratic Casimir operator. We will show that the theory is also free of global anomalies.

To answer this question, we consider a closed three-manifold  $\bar{Y}$  with spin structure  $\alpha$  and some Spin(8) bundle. The global anomaly of the Spin(8) theory described in the last paragraph is measured by  $\Upsilon = \exp\left(-\frac{\pi i}{2}(\eta_{\bar{Y},\alpha,S_+} - \eta_{\bar{Y},\alpha,S_-})\right)$ . The notation is hopefully clear;  $\eta_{\bar{Y},\alpha,S_\pm}$  is the  $\eta$ -invariant on  $\bar{Y}$  for a Majorana fermion in the representation  $S_\pm$  coupled to some background Spin(8) gauge field, as well as to the spin structure  $\alpha$ .

Since this theory has no perturbative anomaly,  $\Upsilon$  is a cobordism invariant and in particular it is a topological invariant. But any Spin(8) bundle on a three-manifold is topologically trivial (since  $\pi_i(\text{Spin}(8)) = 0$  for  $i \leq 2$ ). So we can continuously deform to the case that the background Spin(8) gauge field is trivial, in which case trivially  $\Upsilon = 1$ . Thus the Spin(8) theory under consideration has no global anomaly, and for any background Spin(8) gauge field,

$$\exp\left(-\frac{\pi i}{2}\eta_{\bar{Y},\alpha,S_+}\right) = \exp\left(-\frac{\pi i}{2}\eta_{\bar{Y},\alpha,S_-}\right). \quad (4.17)$$

It follows from this, together with a judicious embedding of  $\mathbb{Z}_2$  in Spin(8), that the  $\mathbb{Z}_2$  theory with 8 positive chirality neutral fermions and 8 negative chirality charged fermions is also free of global anomaly. For this, we embed  $\mathbb{Z}_2$  in SO(8) so that the nontrivial element  $x \in \mathbb{Z}_2$  maps to the central element  $-1 \in \text{SO}(8)$ . The element  $-1 \in \text{SO}(8)$  can be lifted to Spin(8) in two ways. We can pick a lift so that  $x$  acts as  $+1$  on  $S_+$  and as  $-1$  on  $S_-$ . (With the other lift, these signs are reversed.) Now we consider the identity (4.17), specialized to the case that the background Spin(8) gauge field actually has structure group  $\mathbb{Z}_2$ , embedded in Spin(8) as just described. With this choice, the Spin(8) identity (4.17) reduces to the identity (4.16), which says that the anomaly of 8 chiral fermions does not depend on the spin structure. Indeed, for a background Spin(8) gauge field that is induced from a  $\mathbb{Z}_2$  bundle  $L$  by embedding  $\mathbb{Z}_2$  in Spin(8) in the way that we have described, the vector bundle over  $\bar{Y}$  corresponding to  $S_+$  is a rank 8 trivial bundle, and the vector bundle over  $\bar{Y}$  corresponding to  $S_-$  is the direct sum of 8 copies of  $L$ . On the left hand side of eqn. (4.17),  $\eta_{\bar{Y},\alpha,S_+}$  reduces in this example to  $8\eta_{\bar{Y},\alpha}$  on the left hand side of eqn. (4.16), while on the right hand side of eqn. (4.17),  $\eta_{\bar{Y},\alpha,S_-}$  similarly reduces to  $8\eta_{\bar{Y},\beta}$  on the right hand side of eqn. (4.16). So eqn. (4.17) does reduce to eqn. (4.16).

The only property of Spin(8) that we used was that it is connected and simply-connected; the only important property of the fermion representation was that it has no perturbative anomaly. So a two-dimensional theory with a connected and simply-connected gauge group and no perturbative anomaly also has no global anomaly.

#### 4.4 $d = 4$

For an example in  $d = 4$  with nontrivial cancellation of perturbative anomalies, we can take the Standard Model of particle physics. Does it have any global anomaly?

The Standard Model can actually be embedded in the  $SU(5)$  grand unified theory [56] which is also free of perturbative anomalies. It turns out that the  $SU(5)$  grand unified model has no global anomaly. Thus, the phase of the fermion path integral of the  $SU(5)$  grand unified theory, coupled to an arbitrary background metric and  $SU(5)$  gauge field, can be defined in a consistent way. Specializing to the case that the structure group of the background gauge field reduces to the gauge group of the Standard Model, it follows that the Standard Model is also free of global anomalies. This was originally shown in [57], Example 3.4. Here we will explain how to establish the result in the framework of the present paper. See also [58] for another discussion.

Since the Standard Model does not have a time-reversal or reflection symmetry, we formulate it only on oriented four-manifolds  $W$ , and in the anomaly inflow problem, we consider only oriented manifolds  $Y$ . Since the Standard Model has fermions whose definition requires a spin structure, both  $W$  and  $Y$  are endowed with spin structures.

We are not going to be able to get a unique answer for the phase of the path integral of the  $SU(5)$  model in an arbitrary background gauge and gravitational field. The reason is that a four-dimensional spin manifold  $W$  with an  $SU(5)$  bundle is not necessarily the boundary of a five-manifold  $Y$  over which the spin structure and  $SU(5)$  bundle of  $W$  can be extended. The relevant cobordism group is  $\mathbb{Z} \times \mathbb{Z}$ . The two integer-valued invariants are  $\sigma(W)/16$  (where  $\sigma$  is the signature) and the  $SU(5)$  instanton number.

For generators of the cobordism group, we can use the following two manifolds  $W_1$  and  $W_2$ . For  $W_1$ , we take a K3 surface, with some chosen metric and orientation, and with the background gauge field being  $A = 0$ . For  $W_2$ , we take a four-sphere  $S^4$  with some chosen metric and orientation and with some chosen gauge field  $A_0$  of instanton number 1.

We have no way to determine the phases of the path integral measure for those two examples, so we make arbitrary choices. One can think of those choices as representing a precise definition of what is meant by the gravitational  $\theta$ -angle and the  $SU(5)$   $\theta$ -angle including quantum effects of the fermions.

Any other  $W$  is cobordant, by some manifold  $Y$  over which the relevant structures extend, to a linear combination  $n_1 W_1 + n_2 W_2$  for some integers  $n_1$  and  $n_2$ . (By  $n_1 W_1$  or  $n_2 W_2$  with  $n_1$  or  $n_2$  negative, one means  $|n_1|$  or  $|n_2|$  copies of  $W_1$  or  $W_2$  with orientation reversed.) So once the gauge and gravitational  $\theta$ -angles have been fixed, the procedure of section 3.1 gives a definition of the path integral measure for any  $W$ . *A priori*, this definition might depend on  $Y$ .

To know that the phase does not depend on  $Y$  so that the  $SU(5)$  grand unified theory has no global anomaly, we need to know that  $\Upsilon_{\overline{Y}} = \exp(-i\pi\eta_{\overline{Y}}/2)$  is trivial for any closed five-manifold  $\overline{Y}$  with  $SU(5)$

gauge field. Here  $\eta_{\overline{Y}}$  is the  $\eta$ -invariant of the operator  $\mathcal{D}_{\overline{Y}}$ , acting on a five-dimensional field  $\Psi$  that is obtained by combining the Standard Model fermions  $\chi$  with dual fields  $\tilde{\chi}$ . We will use cobordism invariance, plus an examination of some special cases, to show that  $\Upsilon$  is always trivial.

We can proceed as follows. First of all, spin cobordism is trivial in five dimensions, so in case the  $SU(5)$  gauge field is trivial,  $\overline{Y}$  is the boundary of a spin manifold and  $\Upsilon_{\overline{Y}} = 1$ . Actually any fermion coupled only to gravity can have a bare mass in  $d = 4$  and hence does not have any pure gravitational anomaly.<sup>30</sup> Second, an  $SU(5)$  bundle  $E$  over a five-manifold  $\overline{Y}$  is completely classified by its second Chern class  $c_2(E)$ . (This follows from the fact that homotopy groups  $\pi_d(SU(5))$  vanish for  $d \leq 4$  except for  $\pi_3(SU(5)) = \mathbb{Z}$ , which is related to the second Chern class.) Moreover, in five dimensions,  $c_2(E)$  is dual to an embedded circle  $C \subset \overline{Y}$ . That means that, topologically, the  $SU(5)$  gauge field describes some integer number  $\nu$  of instantons propagating along the circle  $C$ . In other words, we can deform the  $SU(5)$  gauge field so that it is trivial except very near some circle  $C$ , while in the normal plane to  $C$  the integral that defines the instanton number integrates to  $\nu$ .

In this situation, by an elementary cobordism,<sup>31</sup>  $\overline{Y}$  is cobordant to a disjoint union  $Y_1 + Y_2$ , where  $Y_1$  and  $Y_2$  are as follows.  $Y_1$  is a copy of  $\overline{Y}$ , but with  $A = 0$ .  $Y_2$  is a copy of  $S^4 \times C$  ( $S^4$  being a four-sphere) with instanton number  $\nu$  on  $S^4$ .

We already know that  $Y_1$  has  $\Upsilon_{Y_1} = 1$ . We therefore only have to investigate  $Y_2$ .

There are two possible spin structures on  $Y_2 = S^4 \times C$ , since the spin structure around  $C$  may be of R or NS type. In the NS case (antiperiodic fermions),  $S^4 \times C$  is the boundary of  $S^4 \times D$ , where  $D$  is a two-dimensional disc, and the  $SU(5)$  gauge field on  $S^4 \times C$  extends over  $S^4 \times D$ . Therefore,  $\Upsilon_{Y_2} = 1$  in this case.

In the R case (periodic fermions), there is no obvious six-dimensional spin manifold  $X$  with boundary  $S^4 \times C$  over which the instanton bundle on  $S^4$  extends. Such an  $X$  actually can be constructed,<sup>32</sup> but

---

<sup>30</sup>Since a bare mass is possible, one suspects that there will be a simple direct proof, without knowing about the cobordism group, that, for a  $d = 4$  Majorana fermion  $\chi$  coupled only to gravity, in five dimensions one has  $\Upsilon_{\overline{Y}} = 0$ . Such a proof may be constructed as follows. A rank five Clifford algebra has an irreducible representation of dimension four, so there exists on a five-dimensional spin manifold  $\overline{Y}$  a self-adjoint Dirac operator  $\mathcal{D}_{\overline{Y}}^0 = i\mathcal{D}$  acting on a four-component fermion field. It is not possible for the five gamma matrices to be all real, but it is possible for three of them to be real while two, say  $\gamma_1$  and  $\gamma_2$ , are imaginary. Using this fact in a locally Euclidean frame, one can define an antilinear operator  $C = *\gamma_1\gamma_2$  that commutes with all gamma matrices and anticommutes with  $i\mathcal{D}$ . Therefore, the  $\eta$ -invariant of  $\mathcal{D}_{\overline{Y}}^0$  reduces to the number of its zero-modes. That number is always even, since  $C^2 = -1$ , so  $\eta(\mathcal{D}_{\overline{Y}}^0)$  is an even integer. The doubling procedure applied to the four-dimensional Majorana fermion  $\chi$  produces an operator  $\mathcal{D}_{\overline{Y}}$  that is the direct sum of two copies of  $\mathcal{D}_{\overline{Y}}^0$ . ( $\chi$  has four components, so after doubling we get an eight component fermion in five dimensions, whose Dirac operator is the direct sum of two copies of  $\mathcal{D}_{\overline{Y}}^0$ .) So  $\eta_{\overline{Y}} = 2\eta(\mathcal{D}_{\overline{Y}}^0)$  is a multiple of 4 and  $\Upsilon_{\overline{Y}} = \exp(-i\pi\eta_{\overline{Y}}/2) = 1$ .

<sup>31</sup>Instead of cobordism, one can use a cut and paste argument that is explained near the end of this section.

<sup>32</sup>This is slightly technical. First, by an elementary cobordism or by a cut and paste argument (as described later), one can replace  $S^4$  with another convenient manifold containing an instanton. A useful choice is  $S^2 \times T^2$  where  $T^2$  is a two-torus, which we take with a spin structure periodic in both directions. Then  $Y_2 = S^4 \times S^1$ , with Ramond spin structure on  $S^1$ , is replaced by  $Y'_2 = S^2 \times T^2 \times S^1$ , where now  $T^2 \times S^1$  is a three-torus with spin structure periodic in



without having to know this, we can proceed as follows. We can deform the metric on  $S^4 \times C$  to be a product and we can choose the  $SU(5)$  gauge field on  $S^4 \times C$  to be a pullback from  $S^4$ . This condition means that if  $C$  is parametrized by an angle  $\tau$ , then the gauge field on  $S^4 \times C$  is independent of  $\tau$  and has no component in the  $\tau$  direction. Concretely then, if the metric of  $C$  is chosen to be  $d\tau^2$ , the self-adjoint Dirac operator on  $Y_2$  takes the form

$$\mathcal{D}_{Y_2} = i\gamma^\tau \frac{\partial}{\partial \tau} + \mathcal{D}_{S^4}, \quad (4.18)$$

where  $\mathcal{D}_{S^4}$  is the self-adjoint Dirac operator of  $S^4$ .

The chirality operator  $\bar{\gamma}_{S^4}$  of  $S^4$  (the product of the four gamma matrices of  $S^4$ ) anticommutes with  $\mathcal{D}_{S^4}$  and commutes with  $\gamma^\tau$ . If we combine  $\bar{\gamma}_{S^4}$  with a reflection  $\tau \rightarrow -\tau$ , we get an operator that anticommutes with  $\mathcal{D}_{Y_2}$ . Hence the nonzero eigenvalues of  $\mathcal{D}_{Y_2}$  come in pairs  $\lambda, -\lambda$ . As usual, such pairs do not contribute to the  $\eta$ -invariant. Hence,  $\eta_{\mathcal{D}_{Y_2}}$  is just the number of zero-modes of  $\mathcal{D}_{Y_2}$ .

Since

$$\mathcal{D}_{Y_2}^2 = -\frac{\partial^2}{\partial \tau^2} + \mathcal{D}_{S^4}^2, \quad (4.19)$$

a zero-mode of  $\mathcal{D}_{Y_2}$  is independent of  $\tau$  and is a zero mode of  $\mathcal{D}_{S^4}$ . Conversely, any such mode is a zero-mode of  $\mathcal{D}_{Y_2}$ . So  $\eta_{Y_2}$  is the same as the number of zero-modes of  $\mathcal{D}_{S^4}$ .

Because of the doubling that is involved in going from  $W$  to  $Y$  in this construction, the operator  $\mathcal{D}_{S^4}$  acts on the Standard Model fermions  $\chi$  plus a dual or complex conjugate set of fermions  $\tilde{\chi}$ . Let  $n$  be the number of zero modes of the Dirac operator on  $S^4$  acting on the original Standard Model fermions  $\chi$ . Complex conjugation exchanges zero-modes of  $\chi$  with zero-modes of  $\tilde{\chi}$ , so the number of zero-modes of  $\mathcal{D}_{S^4}$  is  $2n$ . Hence  $\eta_{Y_2} = 2n$ .

As we will explain in a moment,  $n$  is always even. So  $\eta_{Y_2}$  is a multiple of 4, which implies that  $\exp(-i\pi\eta_{Y_2}/2) = 1$ .

To show that  $n$  is always even, let  $n_+$  and  $n_-$  be the number of positive or negative chirality zero-modes of  $SU(5)$  model fermions, coupled to some background  $SU(5)$  gauge field on  $S^4$ . Then  $n = n_+ + n_-$ , and this is congruent mod 2 to the index  $\mathcal{I} = n_+ - n_-$ . The index theorem for  $SU(5)$  gauge theory shows that for the  $SU(5)$  model fermions,  $\mathcal{I}$  is always even (in fact always a multiple of 4), so  $n$  is always even.

It is inevitable that in one way or another we were going to have to check that in the  $SU(5)$  grand

---

all directions.  $T^2 \times S^1$ , with this spin structure, is the boundary of a “half-K3 surface,” that is, a four-manifold  $Q$  that maps to a disc  $D$  with generic fiber an elliptic curve. In particular, the map  $Q \rightarrow D$  has a section  $s : D \rightarrow Q$ . We can use  $S^2 \times Q$  as the six-manifold of boundary  $Y_2'$ . Picking any point  $p \in S^2$ , the world-volume of the instanton in  $S^2 \times Q$  can be taken to be the two-manifold  $p \times s(D)$ . We proceed with the argument in the text because it seems illuminating to know how to directly evaluate the  $\eta$ -invariant, rather than relying on a technical argument about cobordism.



unified model, the number of fermion zero-modes in an instanton field on  $S^4$  is always even. If this number could be odd, it really would represent an anomaly [8, 59].

An alternative to cobordism in reducing to the two special cases of  $Y_1$  and  $Y_2$  is the following. Let  $Y' = S^4 \times S^1$  with trivial gauge field. Given any closed five-manifold  $\bar{Y}$ , to show  $\Upsilon_{\bar{Y}} = 1$ , let  $\hat{Y} = \bar{Y} + Y'$ . Since  $\Upsilon_{Y'} = 1$ , we have  $\Upsilon_{\hat{Y}} = \Upsilon_{\bar{Y}}$  and we want to show that  $\Upsilon_{\hat{Y}} = 1$ . We will reduce to the special cases  $\Upsilon_{Y_1} = \Upsilon_{Y_2} = 1$  by showing that  $\Upsilon_{\hat{Y}} = \Upsilon_{Y_1} \Upsilon_{Y_2}$ . Since cutting  $S^4$  on the equator decomposes it into two copies of a four-dimensional ball  $B_4$ ,  $Y'$  can be cut to make two copies of  $B_4 \times S^1$ . The boundary of  $B_4 \times S^1$  is  $S^3 \times S^1$ . Likewise by cutting along the boundary of a tubular neighborhood of  $C \subset \bar{Y}$ , we can remove from  $\bar{Y}$  a copy of  $B_4 \times S^1$ , leaving behind another manifold  $\tilde{Y}$ , also with boundary  $S^3 \times S^1$ . At this point we have decomposed  $\hat{Y}$  to the disjoint union of  $\tilde{Y}$  and three copies of  $B_4 \times S^1$ . Of these four manifolds, only one copy of  $B_4 \times S^1$  has a nontrivial gauge field. Exchanging that copy of  $B_4 \times S^1$  with one of the others and then gluing the pieces back together, we get  $Y_1 + Y_2$ . In other words, we have described a procedure of cutting and regluing that converts  $\hat{Y} = \bar{Y} + Y'$  into  $Y_1 + Y_2$ . The gluing law (3.2) for the  $\eta$ -invariant implies that  $\Upsilon$  is invariant under cutting and regluing, so we conclude that  $\Upsilon_{\hat{Y}} = \Upsilon_{Y_1} \Upsilon_{Y_2}$ . Since invariance under cobordism can essentially be deduced from invariance under cutting and regluing and vice versa [40, 41], it is inevitable that a cobordism argument can be expressed in terms of cutting and regluing.

What happens if (still assuming there is no time-reversal or reflection symmetry) we replace  $SU(5)$  by another compact gauge group  $G$ ? Since  $\pi_2(G) = 0$  for any  $G$ , we can repeat the above argument verbatim if  $\pi_0(G) = \pi_1(G) = \pi_4(G) = 0$ . For example, the grand unified theories based on  $Spin(10)$  or  $E_6$  satisfy those conditions.<sup>33</sup> We conclude that if  $G$  satisfies those conditions and moreover the fermion representation is free of perturbative anomaly and the number of fermion zero-modes in a  $G$ -bundle on  $S^4$  is always even, then the model is completely anomaly-free. With some additional knowledge of Lie group topology, one can omit the assumption that  $\pi_4(G) = 0$ . First of all, with  $\pi_0(G) = \pi_1(G) = 0$  but without assuming  $\pi_4(G) = 0$ , the classification of  $G$ -bundles on a five-manifold  $\bar{Y}$  is modified only in a very simple way. The only change is that we have to allow for the possibility that a  $G$ -bundle over  $\bar{Y}$  can be modified in a small neighborhood of a point  $p \in \bar{Y}$  by twisting by a nontrivial element of  $\pi_4(G)$ . This means that in the cobordism analysis, we have to allow a third special case  $Y_3$ , namely a five-sphere  $S^5$  with some  $G$ -bundle. To proceed farther, we need some specific facts about Lie group topology. For simple  $G$ , one has  $\pi_4(G) = 0$  except for  $G = Sp(2k)$ , which satisfies  $\pi_4(Sp(2k)) = \mathbb{Z}_2$ . If  $G$  is semi-simple rather than simple, then  $\pi_4(G)$  is a product of copies of  $\mathbb{Z}_2$ , one for each  $Sp(2k)$  factor in  $G$ . A nontrivial  $Sp(2k)$ -bundle over  $S^5$  associated to the nonzero element of  $\pi_4(Sp(2k))$  can be constructed as follows. Think of an instanton as a particle in five dimensions. Consider an instanton propagating in  $S^5$  in such a way that its worldline is an embedded circle  $C$ . To make from this a nontrivial  $Sp(2k)$  bundle, the instanton should undergo a  $2\pi$  rotation as it propagates around  $C$ . Starting with this description of the nontrivial  $G$ -bundles on  $S^5$ , one can use a cobordism argument – or a cut and paste argument, as in the last paragraph – to show that  $\Upsilon_{Y_3}$ , for a  $G$ -bundle on  $S^5$ , is equal to  $\Upsilon_{Y_2}$ , for a corresponding

---

<sup>33</sup>The  $Spin(10)$  theory has a refinement in which we take the symmetry group of the fermions to be  $(Spin(4) \times Spin(10))/\mathbb{Z}_2$  rather than  $Spin(4) \times Spin(10)$ . Absence of a global anomaly in this case was argued in [60] based on cobordism invariance of the anomaly. It is possible to modify the following argument for this case. See also [61].

$G$ -bundle on  $Y_2 = S^4 \times S^1$ . So the condition  $\Upsilon_{Y_3} = 1$  does not add anything new. In fact, the mod 2 cobordism invariant associated to  $\pi_4(\mathrm{Sp}(2k)) = \mathbb{Z}_2$  is the mod 2 index of the Dirac operator with values in the fundamental representation of  $\mathrm{Sp}(2k)$ . This invariant is nonzero for  $Y_2 = S^4 \times S^1$  with Ramond spin structure around  $S^1$  [59], so in effect our analysis for  $Y_2$  already incorporated the role of  $\pi_4(G)$ .

The discussion so far demonstrates the general utility of cobordism invariance and a cut and paste argument. It is possible to reorganize the argument in a perhaps more elementary way as a reduction from any connected and simply-connected  $G$  to the case  $G = \mathrm{SU}(2)$ . In doing so, we may assume that  $G$  is simple; otherwise, we make the following analysis for each simple factor of  $G$ . In this version of the argument, we de-emphasize the role of cobordism invariance and proceed as much as possible using only the fact that when a four-dimensional theory has no perturbative anomaly, the global anomaly is a topological invariant in five dimensions. To make the argument explicit, let us first consider gauge groups  $G = \mathrm{Spin}(n), \mathrm{SU}(n)$ , and  $\mathrm{Sp}(2n)$ . The bundle associated to the fundamental representation of  $G$  is a real, complex, or quaternionic vector bundle of rank  $n$ , respectively. Let  $\alpha$  be 1, 2 or 4 for  $\mathrm{Spin}$ ,  $\mathrm{SU}$  or  $\mathrm{Sp}$ , respectively. The real rank of the vector bundle is  $\alpha n$ . If  $\alpha n > D$ , there is a section of the vector bundle which is nonzero everywhere. This is simply because a sufficiently generic vector field with  $\alpha n$  real components is always nonzero as a function of  $D$  variables  $x^1, \dots, x^D$  if  $\alpha n > D$ . By taking such a nonzero section, the structure group of a vector bundle can be reduced from  $\mathrm{Spin}(n)$ ,  $\mathrm{SU}(n)$ , or  $\mathrm{Sp}(2n)$  to  $\mathrm{Spin}(n-1)$ ,  $\mathrm{SU}(n-1)$ , and  $\mathrm{Sp}(2n-2)$ , respectively. By repeating this process, the structure group in  $D = 5$  dimensions is topologically reduced to  $\mathrm{Spin}(5)$ ,  $\mathrm{SU}(2)$ , and  $\mathrm{Sp}(2)$ . But we have  $\mathrm{Spin}(5) \cong \mathrm{Sp}(4)$  and this can be further reduced to  $\mathrm{Sp}(2) \cong \mathrm{SU}(2)$ , so we can always reduce the structure group from  $G$  to  $\mathrm{SU}(2)$ . Using “obstruction theory” and a knowledge of the homotopy groups  $\pi_i(G)$ ,  $i \leq 4$ , one can show that this is also possible if  $G$  is a connected and simply-connected exceptional Lie group. Therefore, for any  $G$ , the global anomaly can always be captured for background fields valued in a subgroup  $\mathrm{SU}(2) \subset G$ . An irreducible representation of  $\mathrm{SU}(2)$  is either strictly real or pseudoreal. A strictly real representation does not contribute to the anomaly because a mass term is possible in  $d = 4$ . The exponentiated  $\eta$ -invariant of a pseudoreal representation is given by the mod 2 index in  $D = 5$ . This vanishes if the number of zero-modes in an instanton field on  $S^4$  is always even. But to prove that this last statement holds for a general five-manifold (and therefore that the anomaly associated to the mod 2 index in  $D = 5$  is entirely captured by a counting of zero-modes on  $S^4$ ) appears to require cobordism or cut and paste arguments such as we have explained above.

At any rate, the conclusion is that a four-dimensional theory with connected and simply-connected gauge group has no anomalies beyond the familiar ones. If one drops the requirement for the gauge group to be connected, then, as in  $d = 2$ , there definitely are new anomalies.

## Acknowledgments

The work of KY is supported by JSPS KAKENHI Grant-in-Aid (Wakate-B), No.17K14265. Research of EW is supported in part by NSF Grant PHY-1911298.

## References

- [1] S. L. Adler, “Axial Vector Vertex in Spinor Electrodynamics,” *Phys. Rev.* **177**, 2426 (1969). doi:10.1103/PhysRev.177.2426
- [2] J. S. Bell and R. Jackiw, “A PCAC puzzle:  $\pi^0 \rightarrow \gamma\gamma$  in the  $\sigma$  Model,” *Nuovo Cim. A* **60**, 47 (1969). doi:10.1007/BF02823296
- [3] E. Witten, “Nonsupersymmetric D-Branes and the Kitaev Fermion Chain,” lecture at the Shoucheng Zhang Memorial Workshop, available at [https://glam.stanford.edu/sites/g/files/sbiybj10026/f/may3-1-2\\_witten.pdf](https://glam.stanford.edu/sites/g/files/sbiybj10026/f/may3-1-2_witten.pdf).
- [4] R. Jackiw, “Topological Investigations Of Quantized Gauge Theories,” in B. S. DeWitt et. al., eds., *Relativity, Groups, and Topology, II*, Les Houches 1983, reprinted in updated form in S. B. Treiman et. al., eds., *Current Algebra and Anomalies* (World-Scientific, 1985).
- [5] B. Zumino, “Chiral Anomalies In Differential Geometry,” in B. S. DeWitt et. al., eds., *Relativity, Groups, and Topology, II*, Les Houches 1983, reprinted in S. B. Treiman et. al., eds., *Current Algebra and Anomalies* (World-Scientific, 1985).
- [6] R. Stora, *Algebraic Structure And Topological Origin of Anomalies*, in *Progress in Gauge Field Theory* (Plenum, 1984),
- [7] C. G. Callan, Jr., and J. A. Harvey, “Anomalies And Fermion Zero-Modes On Strings And Domain Walls,” *Nucl. Phys.* **B250** (1985) 427-36.
- [8] E. Witten, “An SU(2) Anomaly,” *Phys. Lett.* **117B** (1982) 324-8.
- [9] E. Witten, “Global Gravitational Anomalies,” *Commun. Math. Phys.* **100**, 197 (1985). doi:10.1007/BF01212448
- [10] M. F. Atiyah, V. K. Patodi and I. M. Singer, “Spectral Asymmetry and Riemannian Geometry 1,” *Math. Proc. Cambridge Phil. Soc.* **77**, 43 (1975). doi:10.1017/S0305004100049410
- [11] X. z. Dai and D. S. Freed, “Eta Invariants and Determinant Lines,” *J. Math. Phys.* **35**, 5155 (1994) Erratum: [*J. Math. Phys.* **42**, 2343 (2001)] doi:10.1063/1.530747 [hep-th/9405012].
- [12] K. Yonekura, “Dai-Freed Theorem and Topological Phases of Matter,” *JHEP* **1609**, 022 (2016) doi:10.1007/JHEP09(2016)022 [arXiv:1607.01873 [hep-th]].
- [13] E. Witten, “World-Sheet Corrections Via  $D$ -Instantons,” *JHEP*0002:030 (2000). arXiv:hep-th/9907041.
- [14] E. Witten, “Fermion Path Integrals And Topological Phases,” *Rev. Mod. Phys.* **88**, no. 3, 035001 (2016) doi:10.1103/RevModPhys.88.035001, 10.1103/RevModPhys.88.35001 [arXiv:1508.04715 [cond-mat.mes-hall]].
- [15] N. Redlich, “Gauge Noninvariance and Parity Violation of Three-Dimensional Fermions,” *Phys. Rev. Lett.* **52** (1984) 18.
- [16] L. Alvarez-Gaume, S. Della Pietra and G. W. Moore, “Anomalies and Odd Dimensions,” *Annals Phys.* **163**, 288 (1985). doi:10.1016/0003-4916(85)90383-5

- [17] D. S. Freed and M. J. Hopkins, “M-Theory Anomaly Cancellation,” arXiv:1908.09916 [hep-th].
- [18] Xiao-Liang Qi, T. Hughes, and Shou-Cheng Zhang, “Topological Field Theory Of Time-Reversal Invariant Insulators,” Phys. Rev. **B78** (2008) 195424, arXiv.:0802.3537.
- [19] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, “Symmetry-Protected Topological Orders in Interacting Bosonic Systems,” Science **338** (2012) 1604.
- [20] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, “Symmetry Protected Topological Orders and the Group Cohomology of Their Symmetry Group,” Phys. Rev. **B87** (2013) 155114.
- [21] T. Senthil, “Symmetry Protected Topological Phases of Quantum Matter,” arXiv:1405.4015.
- [22] D. S. Freed, “Anomalies and Invertible Field Theories,” Proc. Symp. Pure Math. **88**, 25 (2014) doi:10.1090/pspum/088/01462 [arXiv:1404.7224 [hep-th]].
- [23] S. Monnier, “A Modern Point Of View On Anomalies,” arXiv:1903.02828.
- [24] M. Kurkov and D. Vassilevich, “Parity Anomaly in Four Dimensions,” Phys. Rev. D **96**, no. 2, 025011 (2017) doi:10.1103/PhysRevD.96.025011 [arXiv:1704.06736 [hep-th]].
- [25] M. Kurkov and D. Vassilevich, “Gravitational Parity Anomaly with and without Boundaries,” JHEP **1803**, 072 (2018) doi:10.1007/JHEP03(2018)072 [arXiv:1801.02049 [hep-th]].
- [26] I. Fialkovsky, M. Kurkov and D. Vassilevich, “Quantum Dirac Fermions in a Half-space and Their Interaction with an Electromagnetic Field,” Phys. Rev. D **100**, no. 4, 045026 (2019) doi:10.1103/PhysRevD.100.045026 [arXiv:1906.06704 [hep-th]].
- [27] M. F. Atiyah and I. M. Singer, “Dirac Operators Coupled to Vector Potentials,” Proc. Nat. Acad. Sci. **81**, 2597 (1984). doi:10.1073/pnas.81.8.2597
- [28] M. F. Atiyah and I. M. Singer, “The Index Of Elliptic Operators, I,” Ann. Math. **87** (1968) 484-530.
- [29] H. Fukaya, T. Onogi and S. Yamaguchi, “Atiyah-Patodi-Singer Index from the Domain-Wall Fermion Dirac Operator,” Phys. Rev. D **96**, no. 12, 125004 (2017) doi:10.1103/PhysRevD.96.125004 [arXiv:1710.03379 [hep-th]].
- [30] A. Dabholkar, D. Jain, and A. Rudra, “APS  $\eta$ -Invariant, Path Integrals, and Mock Modularity,” arXiv:1905.05207.
- [31] M. Hortacsu, K. D. Rothe and B. Schroer, “Zero Energy Eigenstates for the Dirac Boundary Problem,” Nucl. Phys. B **171**, 530 (1980). doi:10.1016/0550-3213(80)90384-3
- [32] D. Belov and G. W. Moore, “Conformal Blocks for AdS(5) Singletons,” hep-th/0412167.
- [33] E. Witten, “On  $S$ -Duality in Abelian Gauge Theory,” Selecta Math. **1**, 383 (1995) doi:10.1007/BF01671570 [hep-th/9505186].
- [34] N. Seiberg, Y. Tachikawa and K. Yonekura, “Anomalies of Duality Groups and Extended Conformal Manifolds,” PTEP **2018**, no. 7, 073B04 (2018) doi:10.1093/ptep/pty069 [arXiv:1803.07366 [hep-th]].
- [35] C. T. Hsieh, Y. Tachikawa and K. Yonekura, “On the Anomaly of the Electromagnetic Duality of the Maxwell Theory,” arXiv:1905.08943 [hep-th].

- [36] E. Witten, “The ‘Parity’ Anomaly On An Unorientable Manifold,” *Phys. Rev. B* **94**, no. 19, 195150 (2016) doi:10.1103/PhysRevB.94.195150 [arXiv:1605.02391 [hep-th]].
- [37] D.-E. Diaconescu, G. W. Moore, and E. Witten, “ $E_8$  Gauge Theory, and a Derivation of K Theory From M Theory,” *Adv. Theor. Math. Phys.* **6** (2003) 1031-1134.
- [38] D. S. Freed and G. W. Moore, “Setting The Quantum Integrand of  $M$ -Theory,” *Commun. Math. Phys.* **263** (2006) 89-132, [hep-th/0409135].
- [39] A. Kapustin, R. Thorngren, A. Turzillo and Z. Wang, “Fermionic Symmetry Protected Topological Phases and Cobordisms,” *JHEP* **1512**, 052 (2015) [*JHEP* **1512**, 052 (2015)] doi:10.1007/JHEP12(2015)052 [arXiv:1406.7329 [cond-mat.str-el]].
- [40] D. S. Freed and M. J. Hopkins, “Reflection Positivity and Invertible Topological Phases,” arXiv:1604.06527 [hep-th].
- [41] K. Yonekura, “On the Cobordism Classification of Symmetry Protected Topological Phases,” *Commun. Math. Phys.* **368**, no. 3, 1121 (2019) doi:10.1007/s00220-019-03439-y [arXiv:1803.10796 [hep-th]].
- [42] C. Closset, T. Dumitrescu, G. Festuccia, Z. Komargodski, and N. Seiberg, “Comments on Chern-Simons Contact Terms In Three Dimensions,” *JHEP* **1209** (2012) 091, arXiv:1206.5218.
- [43] E. Witten, “Global Anomalies In String Theory,” in W. A. Bardeen and A. R. White, eds., *Symposium on Anomalies, Geometry, Topology* (World Scientific Publishing Company, 1986), available at <https://www.sns.ias.edu/sites/default/files/files/global-anomalies-in-stringtheory-1985.pdf>.
- [44] A. Y. Kitaev, “Unpaired Majorana Fermions in Quantum Wires,” *Phys. Usp.* 44 no. 10S (2001) 131-6, arXiv:cond-mat/0010440 [cond-mat.mes-hall].
- [45] A. Kapustin and N. Seiberg, “Coupling a QFT to a TQFT and Duality,” *JHEP* **04** (2014) 001, arXiv:1401.0740 [hep-th].
- [46] R. Dijkgraaf and E. Witten, “Developments in Topological Gravity,” arXiv:1804.03275 [hep-th].
- [47] D. Stanford and E. Witten, “JT Gravity And The Ensembles Of Random Matrix Theory,” arXiv:1907.03363 [hep-th].
- [48] J. Kaidi, J. Parro-Martinez, and Y. Tachikawa, “GSO Projections Via SPT Phases,” arXiv:1908.04805 [hep-th].
- [49] L. Fidkowski and A. Kitaev, “The Effects Of Interactions On The Topological Classification Of Free Fermion Systems,” *Phys. Rev.* **B81** (2010) 134509, arXiv:0904.2197.
- [50] C. Rosenberg and M. Franz, “Witten Effect In A Crystalline Topological Insulator,” arXiv:1001.3179.
- [51] F. Gliozzi, D. Olive, and J. Scherk, “Supersymmetry, Supergravity Theories, and the Dual Spinor Model,” *Nucl. Phys.* **B122** (1977) 253-90.
- [52] X.-L. Qi, “A New Class of  $(2 + 1)$ -Dimensional Topological Superconductors with  $\mathbb{Z}_8$  Topological Classification,” *New Journal of Physics* **15** (2013) 065002 [arXiv:1202.3983].

- [53] H. Yao and S. Ryu, “Interaction Effect on Topological Classification of Superconductors in Two Dimensions,” *Phys. Rev.* **B88** (2013) 064507 [arXiv:1202.5805].
- [54] Z.-C. Gu and M. Levin, “Effect of Interactions on Two-Dimensional Fermionic Symmetry-Protected Topological Phases with  $\mathbb{Z}_2$  Symmetry,” *Phys. Rev.* **B89** (2014) 201113 [arXiv:1304.4569].
- [55] Y. Tachikawa and K. Yonekura, “Why are Fractional Charges of Orientifolds Compatible with Dirac Quantization?” arXiv:1805.02772 [hep-th].
- [56] H. Georgi and S. L. Glashow, “Unity of All Elementary Particle Forces,” *Phys. Rev. Lett.* **32** (1974) 438-41.
- [57] D. S. Freed, “Pions and Generalized Cohomology,” *J. Diff. Geom.* **80** (2008) 45-77, arXiv:hep-th/0607134.
- [58] I. García-Etxebarria and M. Montero, “Dai-Freed Anomalies in Particle Physics,” *JHEP* **1908**, 003 (2019) doi:10.1007/JHEP08(2019)003 [arXiv:1808.00009 [hep-th]].
- [59] J. Wang, X. G. Wen and E. Witten, “A New SU(2) Anomaly,” *J. Math. Phys.* **60**, no. 5, 052301 (2019) doi:10.1063/1.5082852 [arXiv:1810.00844 [hep-th]].
- [60] J. Wang and X. G. Wen, “A Non-Perturbative Definition of the Standard Models,” arXiv:1809.11171 [hep-th].
- [61] Z. Wan and J. Wang, “Higher Anomalies, Higher Symmetries, and Cobordisms I: Classification of Higher-Symmetry-Protected Topological States and Their Boundary Fermionic/Bosonic Anomalies via a Generalized Cobordism Theory,” arXiv:1812.11967 [hep-th].