Physics of the family index theorem

Kazuya Yonekura

Department of Physics, Tohoku University, Sendai 980-8578, Japan

Abstract: This is a lecture note on a physical interpretation of the geometric family index theorem. The main focus is on the structure of partition functions of chiral fermions in d dimensions and massive fermions in $d+1$ dimensions. These are described by Pfaffians of some Dirac operators.

Contents

1 Introduction

Fermions are one of the most basic and important ingredients in physics. The reason is simply that the matter of our world, including ourselves, consists of fermions, such as electrons and quarks. So I hope there is no need to motivate the reader why we have to understand their mathematical structure better. Index theorems are such mathematics.

There are more specific reasons why index theorems related to fermions are important. One of the reasons is that properties of some material, which are called topological material, are direct consequences of some index theorems. Examples include integer quantum Hall systems and topological insulators. The properties of electric currents on the boundaries of these materials are consequences of the family index theorem, which is the main subject of this note.

Index theorems also appear in various places in string theory and play important roles. Fortunately, we can describe both topological material and string theory by using the same mathematics, at least in effective field theory approach. (I do not explain what effective

field theory is.) The purpose of this note is to describe this mathematical structure of fermions.

I am a physicist and I will not try to make the discussions mathematically rigorous which may make this note difficult to understand for mathematicians. But at the same time, I also give somewhat abstract discussions which make it difficult to understand for physicists. I am also not sure what common knowledge to assume for the physics or mathematics side of the audience, so I may use some facts without explanation which may not be familiar to some readers. I apologize for them in advance.

2 Basic setup

2.1 Symmetry group

Throughout the note we will use the language of relativistic quantum field theory. Spacetimes are Riemannian geometry, so they are manifolds with a metric tensor. Unless otherwise stated, we work with a Wick-rotated metric, so a metric is positive definite.

However, more detailed structure is sometimes necessary. For example, fermions usually require spin structure on spacetime manifolds. Also, electromagnetic fields are described by U(1) fiber bundles.

To accommodate a broad class of cases, we follow Freed and Hopkins [1]. We notice that a metric reduces the structure group of the tangent bundle to the orthogonal group $\mathrm{O}(d)$, where d is the dimension of the manifold. We will define H structure on manifolds, which is a generalization of structure such as spin structure and U(1) bundles.

We consider a sequence of compact Lie group groups H_d for $d = 1, 2, \cdots$ with the following properties. First we require that there is a homomorphism for each d,

$$\rho : H_d \to \mathrm{O}(d) \tag{2.1}$$

such that its image contains the connected part of the orthogonal group, $\mathrm{SO}(d) \subset \rho(H_d)$.

We require that spacetime manifolds are equipped with a principal H_d bundle P such that its image under ρ (or more precisely the bundle $P \times_\rho \mathrm{O}(d)$) is isomorphic to the $\mathrm{O}(d)$ frame bundle associated to the tangent bundle. In such a case, we say that the manifold has an H_d structure (or more simply an H structure). We call a manifold equipped with an H structure as an H manifold.

Let us give a few examples. If $H_d = \mathrm{SO}(d)$ and $\rho : \mathrm{SO}(d) \to \mathrm{O}(d)$ is the obvious inclusion map, then an H structure is just an orientation of a manifold. If $H_d = \mathrm{Spin}(d)$, an H structure is a spin structure. We can also take $H_d = [\mathrm{Spin}(d) \times \mathrm{U}(1)]/\mathbb{Z}_2$, where the nontrivial element of $\mathbb{Z}_2$ is the product of the center $-1 \in \mathrm{Spin}(d)$ and the order 2 element $-1 \in \mathrm{U}(1)$. In this case it is called a spin^c structure of a manifold. Nonorientable cases $\rho(H_d) = \mathrm{O}(d)$ are also important. For example, topological superconductors are characterized by the group $H_d = \mathrm{Pin}^+$ or Pin^-, where $\mathrm{Pin}^\pm$ are some double covers of $\mathrm{O}(d)$ such that its connected component is $\mathrm{Spin}(d)$. Topological insulators are characterized by $H_d = \mathrm{Pin}^\pm(d) \ltimes \mathrm{U}(1)$ (or $[\mathrm{Pin}^\pm(d) \ltimes \mathrm{U}(1)]/\mathbb{Z}_2$) We omit details.

Physically, the group H_d is the symmetry group of the physical system. We always have the Wick-rotated version of the Lorenz symmetry $\mathrm{SO}(d)$, but we may have more

symmetries depending on the system. For example, if there is a U(1) symmetry of the electromagnetic field, the symmetry group contains U(1) at the level of Lie algebra. Thus a fermion coupled to a U(1) gauge field has the Lie algebra $\mathfrak{h}_d \supset \mathfrak{so}(d) \times \mathfrak{u}(1)$ where $\mathfrak{h}_d$ is the Lie algebra of H_d. But the global structure of H_d can be complicated, as in the case of a spin^c structure mentioned above.

We also need the following requirement when we consider manifolds with boundaries. We can consider an inclusion

$$\cdots \subset \mathrm{O}(d) \subset \mathrm{O}(d+1) \subset \cdots \tag{2.2}$$

in an obvious way. Then we require that there is a corresponding inclusion

$$\cdots \subset H_d \subset H_{d+1} \subset \cdots \tag{2.3}$$

such that there is a commutative diagram

$$\begin{array}{ccccccc} \cdots & \longrightarrow & H_d & \longrightarrow & H_{d+1} & \longrightarrow & \cdots . \\ & & \downarrow & & \downarrow & & \\ \cdots & \longrightarrow & \mathrm{O}(d) & \longrightarrow & \mathrm{O}(d+1) & \longrightarrow & \cdots \end{array} \tag{2.4}$$

where horizontal lines are the inclusions mentioned above, and vertical lines are $\rho : H_d \to \mathrm{O}(d)$. Moreover, we assume that each square is a pullback diagram, meaning that if we have $h \in H_{d+1}$ and $o \in \mathrm{O}(d)$ such that $\rho(h) = o$, then h is actually an element of H_d.

This requirement is imposed so that if a $(d+1)$-dimensional manifold with a boundary has an H_{d+1} structure, then its boundary has a canonical H_d structure. The reason is as follows. We can take (say) the outward normal vector field to the boundary. It reduces the structure group from $\mathrm{O}(d+1)$ to $\mathrm{O}(d)$. Then by the above assumption about the diagram, the structure group H_{d+1} is reduced to H_d on the boundary. Whether we use the outward or inward normal vector fields depend on the context. (See e.g. Sec. 3.1.) By changing from the outward to inward vectors, we can define the opposite H structure, whose simplest example is the orientation reversal for $H_d = \mathrm{SO}(d)$. See [1, 2] for details.

2.2 Clifford module

To describe fermions, we need Clifford modules in the language of mathematics or gamma matrices in the language of physics. First let us describe it locally.

A vector space S is a Clifford module associated to the vector space $\mathbb{R}^d$ if it has the following properties. For any vector $v \in \mathbb{R}^d$, we have an action of v on S as a linear map. We denote the corresponding linear map by putting a hat on a vector as $\widehat{v} \in \mathrm{End}(S)$. Thus we have a map

$$(\mathbb{R}^d, S) \ni (v, s) \mapsto \widehat{v}s \in S, \tag{2.5}$$

and it is required to be linear in both variables. Moreover, the above action is required to have the property that

$$\widehat{v}\widehat{v} = v^2 1, \tag{2.6}$$

where 1 here is the identity map and v^2 is the squared length of v computed by the standard metric of the vector space $\mathbb{R}^d$.

In physics, it is more common to use an explicit basis. Let

$$(e_a) = (e_1, \cdots, e_d) \tag{2.7}$$

be the standard orthonormal basis of $\mathbb{R}^d$. We denote the linear map $\widehat{e}_a$ corresponding to e_a as γ_a,

$$\gamma_a = \widehat{e}_a. \tag{2.8}$$

A general vector v can be expanded in terms of the basis (e_a) as $v = \sum_a v_a e_a$. Following Einstein, we omit the summation symbol $\sum_a$ when we sum over indices which appear repeatedly in a single term. For example, we write $v = v_a e_a$. In the orthonormal frame, we have $v^2 = v_a v_a$. Therefore, the condition $\widehat{v}\widehat{v} = v^2$ implies that

$$\{\gamma_a, \gamma_b\} := \gamma_a\gamma_b + \gamma_b\gamma_a = 2\delta_{ab} \tag{2.9}$$

where δ_{ab} is Kronecker delta. We call γ_a as gamma matrices.

We want to incorporate the symmetry group H_d in the above story. We often omit the subscript d since it just implies the dimension and we write the group as H.

We require the vector space S to be not only a Clifford module, but also a representation space of the group H under some representation r. Now we have the action of both $v \in \mathbb{R}^d$ and $h \in H$ on S. They are required be compatible in the sense that

$$r(h)\widehat{v}s = \widehat{\rho(h)v}\, r(h)s, \tag{2.10}$$

where $\rho : H_d \to \mathrm{O}(d)$ is the homomorphism which we have discussed before in the definition of H. In other words,

$$r(h)\widehat{v}r(h)^{-1} = \widehat{\rho(h)v}. \tag{2.11}$$

This gives some restriction on the representation r.

Practically, in most cases (except for gravitinos in supersymmetric theories and non-unitary fields such as bc ghost fields in string theory), this condition is achieved as follows. The Lie algebra $\mathfrak{h}$ of the compact Lie group H can be decomposed into simple factors and $\mathfrak{u}(1)$ factors. Moreover, there is a surjective map $\rho : \mathfrak{h} \to \mathfrak{so}(d)$. These facts imply that there is a subalgebra $\mathfrak{so}(d) \subset \mathfrak{h}$ which is mapped isomorphically to $\mathfrak{so}(d)$ under ρ. Corresponding to generators T_{ab} of $\mathfrak{so}(d)$ with $T_{ab} = -T_{ab}$ and

$$[T_{ab}, T_{cd}] = \delta_{bc}T_{ad} - \delta_{ac}T_{bd} - \delta_{bd}T_{ac} + \delta_{ad}T_{bc}, \tag{2.12}$$

it is possible to construct a representation r such that

$$r(T_{ab}) = \frac{1}{4}(\gamma_a\gamma_b - \gamma_b\gamma_a). \tag{2.13}$$

One can check that the right hand side satisfies the above algebra of T_{ab} by using (2.9). This gives a spin 1/2 representation in the physics language. More general representations can be

constructed by taking a tensor product of this representation with another representation on which γ_a acts trivially.

For physical theories, we need two products in S : an invariant antisymmetric bilinear form, and an invariant positive definite hermitian form. A bilinear form is necessary to write down a Lagrangian. A hermitian form follows from unitarity of physical theories as explained in e.g. [3]. In this note we simply assume that both of them exist.

Let us first discuss an invariant antisymmetric bilinear form $\langle\ ,\ \rangle$. It is a bilinear map

$$S \times S \ni (s_1, s_2) \mapsto \langle s_1, s_2 \rangle \in \mathbb{C}. \tag{2.14}$$

with the properties that

$$\langle r(h)s_1, r(h)s_2 \rangle = \langle s_1, s_2 \rangle\,, \quad \langle s_1, s_2 \rangle = -\langle s_2, s_1 \rangle\,, \quad \langle s_1, \widehat{v}s_2 \rangle = -\langle \widehat{v}s_1, s_2 \rangle\,. \tag{2.15}$$

The first condition is the invariance of the bilinear form under the symmetry group H. The second condition is motivated by the fact that fermion path integrals in physics are integration over Grassmann variables ψ, which satisfy $\psi_1\psi_2 = -\psi_2\psi_1$. For Grassmann variables which take values in S, we get $\langle \psi_1, \psi_2 \rangle = \langle \psi_2, \psi_1 \rangle$ and hence $\langle \psi, \psi \rangle$ can be nonzero. This is used for fermion mass terms. The last condition $\langle s_1, \widehat{v}s_2 \rangle = -\langle \widehat{v}s_1, s_2 \rangle$ will be required by a similar reason that the fermion kinetic term is well-defined.

In a special case that S consists of a Clifford module $\widetilde{S}$ and its dual space $\widetilde{S}^*$ as $S = \widetilde{S} \oplus \widetilde{S}^*$, we can always define a bilinear form as follows. We denote the standard pairing between a vector space $\widetilde{S}$ and its dual space $\widetilde{S}^*$ as $\langle\ ,\ \rangle_{\text{dual}}$. The action of $\widehat{v}$ on $\widetilde{S}^*$ is defined by the condition that $\langle \widetilde{s}', \widehat{v}\,\widetilde{s} \rangle_{\text{dual}} = -\langle \widehat{v}\,\widetilde{s}', \widetilde{s} \rangle_{\text{dual}}$ where $\widetilde{s} \in \widetilde{S}$ and $\widetilde{s}' \in \widetilde{S}^*$. For elements $s_1 = (\widetilde{s}_1, \widetilde{s}_1')$ and $s_2 = (\widetilde{s}_2, \widetilde{s}_2')$ of S, we define

$$\langle s_1, s_2 \rangle = \langle \widetilde{s}_1', \widetilde{s}_2 \rangle_{\text{dual}} - \langle \widetilde{s}_2', \widetilde{s}_1 \rangle_{\text{dual}}\,. \tag{2.16}$$

This case is relevant for Dirac fermions. Dirac fermions are possible in any dimension and for any symmetry group H. More generally, without such splitting of S as $S = \widetilde{S} \oplus \widetilde{S}^*$, the corresponding fermions are called Majorana fermions. Majorana fermions are more general than Dirac fermions, since Dirac fermions are the special case of Majorana fermions in which S happens to be of the form $S = \widetilde{S} \oplus \widetilde{S}^*$.

We also assume to have a positive definite hermitian form $\mathsf{h}(s_1, s_2)$. The existence of a blinear form $\langle \cdot, \cdot \rangle$ and a hermitian form $\mathsf{h}(\cdot, \cdot)$ implies that S is pseudoreal. Then there exists an antilinear map $\mathsf{C} : S \to S$ with $\mathsf{C}^2 = -1$. By changing the normalization of $\mathsf{h}(\cdot, \cdot)$ if necessary, we assume

$$\mathsf{h}(s_1, s_2) = \langle \mathsf{C}(s_1), s_2 \rangle\,. \tag{2.17}$$

The requirement of unitarity in physics turns out to require that $\widehat{v}$ for real vectors v are hermitian matrices with respect to the inner product $\mathsf{h}(\cdot, \cdot)$, which means that $\mathsf{h}(\widehat{v}s_1, s_2) = \mathsf{h}(s_1, \widehat{v}s_2)$. Also, the representation r of H is required to be a unitary representation. By using (2.15) and (2.17), these requirements are translated to the properties of C given by

$$\mathsf{C}(\widehat{v}s) = -\widehat{v}\mathsf{C}(s), \qquad \mathsf{C}(r(h)s) = r(h)\mathsf{C}(s). \tag{2.18}$$

Next let us discuss global issues on a manifold with an H structure. Let P be the principal bundle of the H structure, and let S be a Clifford module as above. By definition of an H structure, the tangent bundle TX is isomorphic to

$$TX = P \times_\rho \mathbb{R}^d. \tag{2.19}$$

We define a Clifford module bundle SX as

$$SX = P \times_r S. \tag{2.20}$$

We can define an action of a vector $v \in T_pX$ at a point $p \in X$ to elements $s \in S_pX$ of the fiber S_pX by using (2.5) in the obvious way. This action is well-defined (i.e. does not depend on a choice of local trivialization of P) due to the consistency condition (2.11).

The bilinear form at each fiber S_pX is also defined by (2.14), and it does not depend on local trivialization due to the invariance of it under H. For sections $s_1, s_2 \in \Gamma(SX)$ of SX, we can define the associated antisymmetric bilinear form as

$$\langle s_1, s_2 \rangle_X = \int_X \langle s_1, s_2 \rangle , \tag{2.21}$$

where the integration measure is defined by using the Riemann metric. The antilinear map C is also globally well defined. We can also define the positive definite hermitian form as $\mathsf{h}(s_1, s_2)_X = \langle \mathsf{C}(s_1), s_2 \rangle_X$.

2.3 Dirac operator

To describe a physical system, it is also necessary to introduce a connection (gauge field) on the H bundle P. A connection on P induces a connection on the tangent bundle TX and the Clifford module SX. On the other hand, the Riemann metric on TX gives the Levi-Civita connection. We require that the connection on P is such that its induced connection on TX is the same as the Levi-Civita connection.

One of the consequences is as follows. Let $v \in \Gamma(TX)$ be a section of TX. Then we get a section $\widehat{v} \in \Gamma(\mathrm{End}(SX))$. The covariant derivative D satisfies $D\widehat{v} = \widehat{Dv}$. This is valid under the condition (2.11).

More explicitly, we can regard the gamma matrices γ_a as a section of $TX \otimes \mathrm{End}(SX)$ by combining them as $\gamma = e_a \gamma_a$. Locally γ is an element of $\mathbb{R}^d \otimes \mathrm{End}(S)$. The condition (2.11) implies that γ is an invariant tensor under the action of H. Then globally it gives a section $\gamma \in \Gamma(TX \otimes \mathrm{End}(SX))$. The above condition on the connection is equivalent to the statement that γ is covariantly constant,

$$D\gamma = 0. \tag{2.22}$$

Having introduced a covariant derivative, we can define a Dirac operator which acts on sections of SX.

$$\not{D} = \gamma \cdot D = \gamma_a D_a \tag{2.23}$$

where D_a is the covariant derivative in the direction e_a. In more detail, the meaning of $\gamma \cdot D$ is as follows. For a section $s \in \Gamma(SX)$, the covariant derivative gives a new section $Ds \in \Gamma(T^*X \otimes SX)$. We can combine it with $\gamma \in \Gamma(TX \otimes \mathrm{End}(SX))$ by using the natural pairing between TX and T^*X, and SX and S^*X, to get $\gamma \cdot Ds \in \Gamma(SX)$.

By recalling (2.15), using integration by parts and the property (2.22), we obtain

$$\left\langle s_1, \not{D} s_2 \right\rangle_X = \left\langle \not{D} s_1, s_2 \right\rangle_X = - \left\langle s_2, \not{D} s_1 \right\rangle_X . \tag{2.24}$$

This is valid if the manifold does not have a boundary. (The case of manifolds with boundaries will become important later.) In other words, for a Grassmann field $\psi \in \Gamma(SX)$, the quantity $\left\langle \psi, \not{D}\psi \right\rangle_X$ can be nonzero. This is the desired property for the kinetic term of a fermion field.

The properties (2.18) and the fact that C is antilinear (rather than linear) implies that

$$\mathsf{C}(\mathsf{i}\not{D} s) = \mathsf{i}\not{D}\mathsf{C}(s). \tag{2.25}$$

Then $\mathsf{i}\not{D}$ is self-adjoint with respect to the hermitian form $\mathsf{h}(\cdot,\cdot)_X$ on closed manifolds.

3 Fermions

3.1 Axioms

In the previous section, we have done enough preparation to discuss fermions. Before going to the discussions of fermions, however, let us mention some general axioms.

In the discussions below, all manifolds are assumed to have an H structure with a connection for a given symmetry group H. When we write a manifold such as X or Y, we are also implicitly including the information of a metric and a connection on it.

A quantum field theory in $D = d + 1$ spacetime dimensions is partly characterized by the following axioms. First, to each d-dimensional closed manifold (i.e. compact manifold without boundary) X, there is an associated Hilbert space $\mathcal{H}(X)$.

Next consider a $(d+1)$-dimensional compact manifold Y with boundaries

$$\partial Y = X_{\mathrm{out}} \sqcup X_{\mathrm{in}}, \tag{3.1}$$

where $\sqcup$ means disjoint union. We have to specify H structures on the boundaries. The manifold Y is assumed to have an H_{d+1} structure. By taking the outward normal vector field to the boundary X_{out}, the structure group $\mathrm{O}(d+1)$ of the tangent bundle is reduced to $\mathrm{O}(d)$. Then, the H_{d+1} structure on Y induces a H_d structure on X_{out} as discussed in Sec. 2.1. On the boundary X_{in}, we use the inward normal vector field to X_{in} and by using it, a H_d structure on X_{in} is defined.

To each $(d+1)$-dimensional compact manifold Y with boundaries $\partial Y = X_{\mathrm{out}} \sqcup X_{\mathrm{in}}$ as above (or more precisely to each triple $(Y, X_{\mathrm{in}}, X_{\mathrm{out}})$), we associate a linear map

$$\mathcal{Z}(Y) : \mathcal{H}(X_{\mathrm{in}}) \to \mathcal{H}(X_{\mathrm{out}}). \tag{3.2}$$

Moreover, we require that these maps can be composed in the following way. Suppose we have two manifolds $\partial Y^1 = X^1_{\text{out}} \sqcup X^1_{\text{in}}$ and $\partial Y^2 = X^2_{\text{out}} \sqcup X^2_{\text{in}}$ with $X^1_{\text{out}} = X^2_{\text{in}}$.[1] Then we can glue Y^1 and Y^2 at $X^1_{\text{out}} = X^2_{\text{in}}$. We assume that the manifold after gluing is smooth. For simplicity, we guarantee the smoothness by requiring that the neighborhood of any boundary component is of a product form

$$(-\epsilon, 0] \times X_{\text{out}}, \qquad [0, \epsilon) \times X_{\text{in}} \tag{3.3}$$

where ϵ is a sufficiently small positive number. The H structure and in particular the Riemann metric on Y is also a product of the H structures on $(-\epsilon, 0]$ or $[0, \epsilon)$ and X, with the trivial connection on $(-\epsilon, 0]$ and $[0, \epsilon)$. Then the manifold obtained by gluing Y^1 and Y^2 are smooth. We denote the glued manifold as $Y^2 \cdot Y^1$. Then the linear map $\mathcal{Z}(Y)$ is required to satisfy

$$\mathcal{Z}(Y^2 \cdot Y^1) = \mathcal{Z}(Y^2)\mathcal{Z}(Y^1). \tag{3.4}$$

Notice that both sides are maps $\mathcal{H}(X^1_{\text{in}}) \to \mathcal{H}(X^2_{\text{out}})$.

A special case of manifolds is the empty set $X = \varnothing$ regarded as a d-dimensional manifold. In this case, we set $\mathcal{H}(\varnothing) = \mathbb{C}$. Then, if a $(d+1)$-dimensional manifold Y has no boundary $\partial Y = \varnothing$, the quantity $\mathcal{Z}(Y)$ takes values in $\mathbb{C}$. This is called the partition function on Y. Recall that a symbol Y also contains the information of a metric and a connection on it, which we call as background fields. Thus $\mathcal{Z}(Y)$ is a functional of the background fields.

3.2 Massive fermions in $D = d + 1$ dimensions

Now we discuss massive fermion theories in $D = d+1$ dimensional spacetime. We consider a fermion field Ψ. A theory is specified by writing down the action $\mathcal{S}$,

$$-\mathcal{S} = -\frac{1}{2} \left\langle \Psi, (\not{D} + m)\Psi \right\rangle_Y , \tag{3.5}$$

where m is a real parameter. The Feynman path integral is $\int [D\Psi] \exp(-\mathcal{S})$ for the Grassmann field Ψ.

Among other things, the above action means the following thing on a closed manifold Y (i.e. compact and $\partial Y = \varnothing$). Let s_1, s_2 be sections of SX. Then by (2.15) and (2.24) we have

$$\left\langle s_1, (\not{D} + m)s_2 \right\rangle_Y = - \left\langle s_2, (\not{D} + m)s_1 \right\rangle_Y . \tag{3.6}$$

Thus we can regard $(\not{D} + m)$ (or more precisely $\left\langle \cdot, (\not{D} + m) \cdot \right\rangle$) as an antisymmetric form on $\Gamma(SX)$. Then the partition function $\mathcal{Z}_\Psi(Y)$ is given by

$$\mathcal{Z}_\Psi(Y) \sim \mathrm{Pf}(\not{D} + m) \tag{3.7}$$

[1] More precisely, when we write an equation like $\partial Y = X_{\text{out}} \sqcup X_{\text{in}}$, we actually mean that there is an isomorphism from boundaries of Y to $X_{\text{out}} \sqcup X_{\text{in}}$. We will not try to be precise in this note.

where $\mathrm{Pf}(\not{D}+m)$ formally means the Pfaffian of the antisymmetric form $\langle \cdot, (\not{D}+m)\cdot \rangle$ on the vector space $\Gamma(SX)$.

There are some problems with the above formal Pfaffian. First, the Pfaffian of an antisymmetric matrix A depends on a choice of a basis of the vector space. Suppose we make change from one basis to another by a matrix B. Then the matrix A is changed to $B^T AB$ where B^T is the transpose of B. Under this change, the Pfaffian is changed to $\mathrm{Pf}(B^T AB) = \mathrm{Det}(B)\,\mathrm{Pf}(A)$.

Another problem is that the Pfaffian is taken in the infinite dimensional vector space $\Gamma(SX)$. For example, a similar quantity $\mathrm{Det}(\not{D}+m)$ involves the product over eigenvalues of $\mathsf{i}\not{D}$,

$$\mathrm{Det}(\not{D}+m) \sim \prod(-\mathsf{i}\lambda + m). \tag{3.8}$$

where the product runs over all eigenmodes of $\mathsf{i}\not{D}$. The imaginary unit i here is put since $\mathsf{i}\not{D}$ is a self-adjoint operator so λ is real. The absolute value of this quantity may be regularized by regularizations such as zeta function regularization and heat kernel regularization. But for our purposes, the most significant problem is the phase of the partition function. In the region of large eigenvalues $\lambda \to \pm\infty$, each factor of the product has a phase factor $-\mathsf{i}\,\mathrm{sign}(\lambda)$ where $\mathrm{sign}(\lambda) = \lambda/|\lambda|$. It is not clear how to handle the product of infinitely many such factors. Pf is roughly a square root of Det, so it has a similar problem.

These problems can be avoided by what is called Pauli-Villars regularization in physics. Let us consider a quantity

$$\frac{\mathrm{Pf}(\not{D}+m)}{\mathrm{Pf}(\not{D}+M)}, \tag{3.9}$$

where M is a real parameter. By taking the ratio between two Pfaffians, it is now independent of a choice of a basis. Also, for large eigenvalues λ of $\mathsf{i}\not{D}$, we have

$$\frac{-\mathsf{i}\lambda + m}{-i\lambda + M} \to 1 \quad (|\lambda| \to \infty). \tag{3.10}$$

Thus, although the product is over infinitely many eigenvalues, most of the factors are close to 1. By using this fact, we can make the above ratio of the Pfaffians well-defined. The denominator $\mathrm{Pf}(\not{D}+M)$ is interpreted to come from a Feynman path integral of a field called the Pauli-Villars field. More precisely, we need to do more for a complete regularization (see Appendix A for a sketch of an example), but we neglect that issue in this note.

The Pauli-Villars field is an unphysical field and we want to take a limit $|M| \to \infty$. The choice of the sign of M has an important consequence which we will discuss later, and in this note we use $M > 0$. When we take the limit $M \to \infty$, we need to include what is called a counterterm $\mathcal{S}_{\mathrm{c.t.}}$ which is required for renormalization in physics. A counterterm is a local polynomial of the curvature tensors and their derivatives which is invariant under the symmetry group H. For example, we always have the Riemann curvature tensor R_{abcd}, and if there is no other curvature tensor, a counterterm is a polynomial of the form

$$\mathcal{S}_{\mathrm{c.t.}} = \int \left(c_1 + c_2 R + c_3 R^2 + c_4 R_{ab}R_{ab} + c_5 R_{abcd}R_{abcd} + c_6 (D_a R)(D_a R) + \cdots\right), \tag{3.11}$$

where the integral is over the spacetime manifold with the measure determined by the Riemann metric. If there are other curvature tensors, we also include them in an obvious way. The coefficients $c_1, c_2, \cdots$ are chosen by the following principle. When we take the limit $M \to \infty$, the quantity (3.9) diverges. However, it is expected that the divergence can be cancelled by multiplying $\exp(-\mathcal{S}_{\text{c.t.}})$ to (3.9) if we choose the coefficients $c_1, c_2, \cdots$ appropriately. In other words, we expect to get a finite limit

$$\mathcal{Z}_\Psi(Y) = \lim_{M\to\infty} \left[\frac{\mathrm{Pf}(\not{D} + m)}{\mathrm{Pf}(\not{D} + M)} \exp(-\mathcal{S}_{\text{c.t.}}) \right] \tag{3.12}$$

by appropriately choosing the coefficients of the counterterm. The partition function $\mathcal{Z}(Y)$ is defined by this limit.

The above requirement for the existence of a finite limit does not uniquely specify the coefficients $c_1, c_2, \cdots$. We can always shift them by some values which are independent of M or at least finite in the limit $M \to \infty$. There are some physical restrictions on them (such as unitarity of quantum theory), but it is unavoidable that a certain amount of ambiguities remain. Usually a choice of them is a part of the data of the physical theory which specifies physical parameters, such as the strength of gravitational force. But for the purposes of the present note, it will be convenient to choose them by some different requirement which we will discuss later.

3.3 Chiral fermions in d dimensions

Next we discuss chiral fermions in a d-dimensional manifold X. The reason why we have considered massive fermions in $D = d+1$ dimensions above and we consider chiral fermions in d dimensions rather than D dimensions will become clear later.

To define chiral fermions, it is important to introduce a $\mathbb{Z}_2$-graded Clifford module S. In addition to gamma matrices γ_a $(a = 1, \cdots, d)$ and action of $h \in H$, we introduce an additional operator $\overline{\gamma}$, which we call a chirality operator, acting on S. It is required to satisfy

$$\overline{\gamma}^2 = 1, \qquad \{\overline{\gamma}, \gamma_a\} = 0, \qquad r(h)\overline{\gamma}r(h)^{-1} = \overline{\gamma}. \tag{3.13}$$

The first equation implies that $\overline{\gamma}$ has eigenvalues ± 1 which is used for $\mathbb{Z}_2$ grading. The second equation means that $\overline{\gamma}$ anticommutes with gamma matrices and hence gamma matrices change the eigenvalue of $\overline{\gamma}$. The third equation means that $\overline{\gamma}$ is invariant under H.

We also require that

$$\langle s_1, \overline{\gamma} s_2 \rangle = - \langle \overline{\gamma} s_1, s_2 \rangle . \tag{3.14}$$

This condition is necessary to define chiral fermions below.

By using $\overline{\gamma}$, we decompose S as

$$S = S_+ \oplus S_-, \tag{3.15}$$

where $S_\pm$ are eigenspaces with eigenvalues $\overline{\gamma} = \pm 1$, respectively. Then we can define the corresponding bundles $S_\pm X$. We say that elements of S_+ (or the corresponding bundle) as having the positive chirality and S_- as having the negative chirality.

Let s_1, s_2 be sections of SX. Then we have

$$\left\langle s_1, \not{D}\overline{\gamma} s_2 \right\rangle_X = -\left\langle s_1, \overline{\gamma}\not{D} s_2 \right\rangle_X = +\left\langle \overline{\gamma} s_1, \not{D} s_2 \right\rangle_X \tag{3.16}$$

where we have used (3.13) and (3.14). This means that $\left\langle s_1, \not{D} s_2 \right\rangle_X$ is automatically zero if s_1 and s_2 have different chiralities, but it can be nonzero if they have the same chirality.

Let ψ be a Grassmann field which is a section of $\Gamma(S_+X)$. This ψ is a chiral fermion. We can have a nonzero action

$$-\mathcal{S} = -\frac{1}{2}\left\langle \psi, \not{D}\psi \right\rangle_X . \tag{3.17}$$

However, there is a crucial difference from the previous case. (3.14) implies that $\langle s_1, s_2 \rangle$ is automatically zero if s_1 and s_2 have the same chirality. This means that the mass term of a simple form $m\langle \psi, \psi \rangle_X$ is not possible. In some cases there may be more invariant antisymmetric bilinear forms in S other than $\langle \cdot, \cdot \rangle$ depending on the representation r of H, and mass terms may be possible in those cases. However, for the purposes of general argument, we proceed without mass terms.

There is a big problem if a mass term is not possible. The partition function of ψ on a closed manifold X is supposed to be

$$\mathcal{Z}_\psi(X) \sim \mathrm{Pf}(\not{D}_+) \tag{3.18}$$

where $\not{D}_+$ is the restriction of $\not{D}$ to positive chirality sections $\Gamma(S_+X)$. However, without a mass term, the Pauli-Villars regularization is not possible. The absolute value of this Pfaffian may be defined by e.g. zeta function or heat kernel regularization, but we cannot determine its phase.

The most essential problem is the following. Let us first discuss the case of a finite dimensional vector space V. Let A be an antisymmetric bilinear form on V. If we take an explicit basis (v_α), we can represent A as an antisymmetric matrix $A^{(v)}_{\alpha\beta}$. Then

$$\mathrm{Pf}(A^{(v)}) = \frac{1}{2^n n!}\epsilon^{\alpha_1 \cdots \alpha_{2n}} A_{\alpha_1\alpha_2} \cdots A_{\alpha_{2n-1}\alpha_{2n}}, \tag{3.19}$$

where $\dim V = 2n$, and $\epsilon^{\alpha_1 \cdots \alpha_{2n}}$ is the totally antisymmetric tensor with $\epsilon^{12\cdots 2n} = 1$. If we change the basis as

$$w_\alpha = v_\beta B^\beta{}_\alpha, \tag{3.20}$$

Then A is changed as $A^{(w)}_{\alpha\beta} = A^{(v)}_{\gamma\delta} B^\gamma{}_\alpha B^\delta{}_\beta$. Pfaffian changes as

$$\mathrm{Pf}(A^{(w)}) = \mathrm{Det}(B)\,\mathrm{Pf}(A^{(v)}). \tag{3.21}$$

This transformation implies that $\mathrm{Pf}(A)$ can be most naturally regarded as an element of a one dimensional vector space as follows. Given a vector space V, we denote its top exterior power as $\mathrm{Det}\, V$,

$$\mathrm{Det}\, V = \wedge^{\dim V} V = V \wedge \cdots \wedge V. \tag{3.22}$$

This is a one-dimensional vector space. Let V^* be the dual space to V, and let $(\widetilde{v}^\alpha)$ be the dual basis to (v_α). $\mathrm{Det}\, V^*$ has a basis $\widetilde{v}^1 \wedge \cdots \widetilde{v}^{2n}$. We find that

$$\mathrm{Pf}(A) := \mathrm{Pf}(A^{(v)}) \widetilde{v}^1 \wedge \cdots \wedge \widetilde{v}^{2n} \tag{3.23}$$

is independent of a choice of a basis. Therefore, we can most naturally regard $\mathrm{Pf}(A)$ as an element of $\mathrm{Det}\, V^*$.

Let us return to the problem of chiral fermions. Based on the above discussion, we can most naturally regard $\mathrm{Pf}(\not{D}_+)$ as an element of a one dimensional vector space called the Pfaffian line,

$$\mathcal{L}_X = \mathrm{Det}\, \Gamma(S_+X)^*. \tag{3.24}$$

Therefore, $\mathrm{Pf}(\not{D}_+)$ is not a number, but an element of a vector space and hence there is no natural phase of it unless some canonical trivialization of $\mathrm{Det}\, \Gamma(S_+X)^*$ is given.

The vector space $\Gamma(S_+X)^*$ is infinite dimensional, and hence one may wonder how to define a space like $\mathrm{Det}\, \Gamma(S_+X)$. This can be defined as follows. Let us consider eigenfuncitons $s_\alpha \in \Gamma(S_+X)$ with nonzero eigenvalues of $(\mathrm{i}\not{D})^2$,

$$(\mathrm{i}\not{D})^2 s_\alpha = \lambda_\alpha^2 s_\alpha. \tag{3.25}$$

We have

$$\lambda_\beta^2 \langle s_\alpha, s_\beta \rangle = \left\langle s_\alpha, (\mathrm{i}\not{D})^2 s_\beta \right\rangle = \left\langle (\mathrm{i}\not{D})^2 s_\alpha, s_\beta \right\rangle = \lambda_\alpha^2 \langle s_\alpha, s_\beta \rangle . \tag{3.26}$$

Therefore, we have $\langle s_\alpha, s_\beta \rangle = 0$ unless $\lambda_\alpha^2 = \lambda_\beta^2$. This means that we can decompose $\Gamma(S_+X)$ into subspaces

$$\Gamma(S_+X) = \bigoplus_{\lambda^2} V_{\lambda^2} \tag{3.27}$$

where V_{λ^2} is spanned by eigenmodes s_α with the eigenvalue $\lambda_\alpha^2 = \lambda^2$.

Within each subspace V_{λ^2} with $\lambda^2 \neq 0$, we can take a canonical trivialization of $\mathrm{Det}\, V_{\lambda^2}$ as follows. On this subspace, there is a non-degenerate antisymmetric bilinear form

$$J_{\alpha\beta}^{(\lambda^2)} = \frac{1}{\sqrt{\lambda^2}} \left\langle s_\alpha, \not{D} s_\beta \right\rangle_X . \tag{3.28}$$

The fact that this is non-degenerate for $\lambda^2 \neq 0$ can be shown by taking $s_\alpha = \mathsf{C}(\not{D} s_\beta)$ for arbitrary s_β, where C is the antilinear map introduced around (2.17). We take a basis in which this $J_{\alpha\beta}^{(\lambda^2)}$ takes the standard form

$$(J_{\alpha\beta}^{(\lambda^2)}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{3.29}$$

Such a choice is not unique, but any choice leads to the same trivialization of $\mathrm{Det}\, V_{\lambda^2}$. The reason is as follows. Suppose that we change the basis by a matrix U. To preserve $J^{(\lambda^2)}$, we require $U^T J^{(\lambda^2)} U = J^{(\lambda^2)}$. Taking the Pfaffian of both sides, we get $\mathrm{Det}(U)\,\mathrm{Pf}(J^{(\lambda^2)}) = \mathrm{Pf}(J^{(\lambda^2)})$ or $\mathrm{Det}(U) = 1$. Therefore, the basis of $\mathrm{Det}\, V_{\lambda^2}$ is invariant under this change.

We can use the canonical trivialization of $\mathrm{Det}\, V_{\lambda^2}$ above for all spaces V_{λ^2} with $\lambda^2 \neq 0$. Then we define

$$\mathrm{Det}\,\Gamma(S_+X)^* = \bigotimes_{\lambda^2 \leq \Lambda} \mathrm{Det}\, V_{\lambda^2}. \tag{3.30}$$

where Λ is an arbitrary positive number. Because of the trivialization $\mathrm{Det}\, V_{\lambda^2} \cong \mathbb{C}$, the only important factor is that of $\lambda^2 = 0$. So the value of Λ is irrelevant. However, the above expression may be more illuminating in later discussions. The right hand side of (3.30) is constructed from finite dimensional vector spaces and hence it is well-defined. Notice that if there are no modes with zero eigenvalue $\lambda^2 = 0$, then we have $\mathrm{Det}\,\Gamma(S_+X)^* \cong \mathbb{C}$. In this case, $\mathrm{Pf}(\not{D}_+) \in \mathrm{Det}\,\Gamma(S_+X)^*$ is just given by $\mathrm{Pf}(\not{D}_+) = \prod \mathrm{Pf}(\sqrt{\lambda^2} J^{(\lambda^2)})$. If V_{λ^2} is odd dimensional, we simply set $\mathrm{Pf}(\not{D}_+) = 0$.[2]

We must be concerned with zero modes (i.e. modes with $\lambda^2 = 0$) by the following reason. If there is no particular reason from index theorems of the Dirac operator $\mathsf{i}\not{D}$, there are generically no zero modes for a generic background field. Here, a background field means the data of a manifold X with an H structure and an explicit metric and a connection on it. However, physically we are interested not only in a single background field, but the space of all possible background fields. Let $\widetilde{\mathcal{M}}$ be the moduli space of all background fields, and $\mathcal{M}$ be the space $\widetilde{\mathcal{M}}$ divided by the diffeomorphism and gauge transformation groups. This means that a point on $\mathcal{M}$ is given by X, which includes the information of a metric and a connection on it, up to diffeomorphisms and gauge transformations. (This description is not rigorous but we do not try to be precise.) When we vary $X \in \mathcal{M}$, the Dirac operator $\mathsf{i}\not{D}$ changes. Then at some point in $\mathcal{M}$, we may get zero modes. The space $\mathcal{M}$ is huge and hence we expect the appearance of zero modes at some points of $\mathcal{M}$. Near such points, we cannot trivialize V_{λ^2} for the eigenvalue λ^2 which becomes zero at the point. Then we have to use the expression (3.30) near those points. By varying X, the Pfaffian lines $\mathcal{L}_X$ defined in (3.24) are combined to a line bundle $\mathcal{L}$ on $\mathcal{M}$ whose fiber at $X \in \mathcal{M}$ is $\mathcal{L}_X$. This is called a Pfaffian line bundle. Then $\mathrm{Pf}(\not{D}_+)$ is a section of this Pfaffian line bundle.

Now we can ask whether the Pfaffian line bundle $\mathcal{L}$ can be trivialized or not on $\mathcal{M}$. For example, if the first Chern class $c_1(\mathcal{L})$ is nontrivial, it is impossible to trivialized $\mathcal{L}$. Then $\mathrm{Pf}(\not{D}_+)$ cannot be regarded as taking values in $\mathbb{C}$. Thus the chiral fermion ψ is not a "quantum field theory" in the sense of Sec. 3.1.

In fact, we will determine an explicit differential geometric connection on the Pfaffian line bundle $\mathcal{L}$. This is studied mathematically by Bismut and Freed (see [4] and references therein) which was partly motivated by Witten's global anomaly formula [5]. We will give a physical argument for it [3, 6]. As a corollary, we can derive the curvature from the

[2] $\dim V_{\lambda^2} \bmod 2$ is a $\mathbb{Z}_2$ index of $\not{D}_+$.

connection and hence get an explicit differential form expression for the first Chern class $c_1(\mathcal{L})$. This is called an anomaly polynomial in physics and it contains the information of perturbative anomalies of ψ [7–10]. However, $\mathcal{L}$ contains more topological structure than just the differential form expression for $c_1(\mathcal{L})$. Moreover, it is not sufficient to ask whether there exists a trivialization of $\mathcal{L}$ or not. We must ask whether there is a canonical trivialization which satisfies physical principles. This is studied by Wittten [11–13], based on a theorem by Dai and Freed [14]. We discuss some of these points in the next section based on the understanding of the situation in terms of the physics of massive fermions in $D = d + 1$ dimensions [3, 6]. This understanding is along the lines of the study of topological material.

4 Fermions and the geometric family index theorem

4.1 Massive fermions on a manifold with boundary

First let us make the following simple observation. In the case of chiral fermions, we have the $\mathbb{Z}_2$-graded Clifford algebra generated by $\overline{\gamma}$ and γ_a for $a = 1, \cdots, d$. Their algebra is

$$\{\overline{\gamma}, \overline{\gamma}\} = 1, \quad \{\overline{\gamma}, \gamma_a\} = 0, \quad \{\gamma_a, \gamma_b\} = \delta_{ab}. \tag{4.1}$$

If we simply set $\gamma_0 = \overline{\gamma}$, they are just $(d+1)$-dimensional Clifford algebra $\{\gamma_a, \gamma_b\} = \delta_{ab}$ for $a, b = 0, 1, \cdots, d$. Also notice that under the subgroup $H_d \subset H_{d+1}$, the matrix γ_0 is invariant and hence the identification $\overline{\gamma} = \gamma_0$ is valid including the action of H. We would like to see that there is indeed a relation between chiral fermions in d dimensions and massive fermions in $d + 1$ dimensions.

To see the relation, we consider a $(d+1)$-dimensional manifold Y with a boundary $\partial Y = X$, and study massive fermions on it.

Near the boundary, we assume that Y is of the form $Y \supset (-\epsilon, 0] \times X$ with the standard metric and the trivial connection on $(-\epsilon, 0]$, for some small positive constant ϵ. We denote the standard coordinate of $(-\epsilon, 0]$ as $\tau \in (-\epsilon, 0]$. Then we impose a boundary condition on a massive fermion Ψ given by

$$\mathsf{L}: \ (\gamma_0 - 1)\Psi|_{\partial Y} = 0. \tag{4.2}$$

In more detail, the meaning of this boundary condition is as follows. γ_0 is the gamma matrix in the direction $(-\epsilon, 0]$. Since $\gamma_0^2 = 1$, it has eigenvalues ± 1. The above boundary condition means that the components of Ψ with the negative eigenvalue $\gamma_0 = -1$ are set to zero at the boundary.

Under the above boundary condition L, the Dirac operator $\mathrm{i}\not{D}_Y$ on Y is not self-adjoint. The reason is as follows. Recall that the positive definite hermitian form is given by $\langle \mathsf{C}(s_1), s_2 \rangle$ as discussed in Sec. 2.2. We consider integration by parts as

$$\int_Y \langle \mathsf{C}(s_1), \mathrm{i}\not{D}_Y s_2 \rangle = \int_Y \langle \mathsf{C}(\mathrm{i}\not{D}_Y s_1), s_2 \rangle + \int_X \langle \mathsf{C}(s_1), \mathrm{i}\gamma_0 s_2 \rangle \, . \tag{4.3}$$

The second term is a surface term. Now, suppose that s_2 satisfies the boundary condition $\gamma_0 s_2 = s_2$ on $\partial Y = X$. Then we ask what boundary condition on s_1 makes the surface

term vanish. If we take the same boundary condition for s_1 and s_2, the surface term is not guaranteed to vanish. We can easily demonstrate it by taking $s_1 = s_2$, and use the fact that $\langle \mathsf{C}(s), s\rangle$ is positive definite. The boundary condition which guarantees $\langle \mathsf{C}(s_1), s_2\rangle = 0$ for arbitrary s_2 with $\gamma_0 s_2 = s_2$ is to require that $\gamma_0 s_1 = -s_1$. This is because

$$\langle \mathsf{C}(s_1), \gamma_0 s_2\rangle = \langle \mathsf{C}(\gamma_0 s_1), s_2\rangle\,. \tag{4.4}$$

Thus $\langle \mathsf{C}(s_1), s_2\rangle$ vanishes if s_1 and s_2 have different eigenvalues under γ_0.

We can still consider the eigenvalue problem for the operator

$$\Delta_m = (\not{D}_Y + m)^\dagger (\not{D}_Y + m), \tag{4.5}$$

where $(\not{D}_Y + m)^\dagger$ is the adjoint of $(\not{D}_Y + m)$ with respect to the hermitian form $\langle \mathsf{C}(s_1), s_2\rangle$ for $\gamma_0 s_1 = -s_1$ and $\gamma_0 s_2 = +s_2$. This suggests that for the eigenvalue problem for Δ_m, we should impose the boundary condition such that

$$(1-\gamma_0)s|_{\partial Y} = 0, \qquad (1+\gamma_0)(\not{D}_Y + m)s|_{\partial Y} = 0. \tag{4.6}$$

Then we expect to get well-defined engenfunctions and eigenvalues.

We denote the partition function of the massive fermion on Y with the boundary condition L given in (4.2) as $\mathcal{Z}(Y, \mathsf{L})$;

$$\mathcal{Z}_\Psi(Y) = \lim_{M\to\infty}\left[\frac{\mathrm{Pf}(\not{D}_Y + m)}{\mathrm{Pf}(\not{D}_Y + M)}\exp(-\mathcal{S}_{\text{c.t.}})\right], \tag{4.7}$$

where now $\mathcal{S}_{\text{c.t.}}$ may include an integral of a local polynomial of curvatures not only in Y but also on ∂Y. The absolute value of the Pfaffian is interpreted as

$$|\,\mathrm{Pf}(\not{D}_Y + m)| = (\mathrm{Det}\,\Delta_m)^{1/4}, \tag{4.8}$$

where the exponent $1/4 = (1/2)\cdot(1/2)$ comes from the fact that $\mathrm{Pf} \sim (\mathrm{Det})^{1/2}$ and $(\not{D}_Y + m) \sim \Delta^{1/2}$. The right hand side may be regularized by e.g. zeta function or heat kernel regularization. We define the absolute value of the Pauli-Villars contribution in the same way.

Including the phase, the ratio of the Pfaffians is determined as follows. We have a complete set of eigenfunctions $s_\alpha^{(m)}$ for Δ_m, and another complete set of eigenfunctions $s_\alpha^{(M)}$ for Δ_M. We can define matrices

$$A^{(m)}_{\alpha\beta} = \left\langle s^{(m)}_\alpha, (\not{D}_Y + m)s^{(m)}_\beta\right\rangle, \quad A^{(M)}_{\alpha\beta} = \left\langle s^{(M)}_\alpha, (\not{D}_Y + M)s^{(M)}_\beta\right\rangle. \tag{4.9}$$

They are related by some unitary matrix $U^\alpha{}_\beta$ such that

$$s^{(M)}_\alpha = s^{(m)}_\beta U^\beta{}_\alpha. \tag{4.10}$$

Then we formally get

$$\frac{\mathrm{Pf}(\not{D}_Y + m)}{\mathrm{Pf}(\not{D}_Y + M)} = \frac{\mathrm{Pf}(A^{(m)})}{\mathrm{Pf}(A^{(M)})}\,\mathrm{Det}(U). \tag{4.11}$$

where the property of Pfaffian (3.21) is used. The matrices appearing here are infinite dimensional, but physically we expect that this formal expression can be regularized to get a well-defined value. The intuition is that for engenvalues λ^2 of Δ which are very large, $\lambda^2 \gg M^2, m^2$, we expect that eigenfunctions $s^{(m)}_\alpha$ and $s^{(m)}_\alpha$ coincide $s^{(m)}_\alpha \to s^{(M)}_\alpha$ for $\lambda^2 \to \infty$.

4.2 Chiral fermions as boundary modes of massive fermions

The crucial point for the physics of chiral fermions is as follows. A chiral fermion appears on the boundary of a massive fermion with the boundary condition (4.2) if we take m to be negative. (We will take M to be positive so that there is no chiral fermion from the unphysical Pauli-Villars field.)

For concreteness, let us first consider a manifold

$$Y = (-\infty, 0] \times X. \tag{4.12}$$

A quick way to see the appearance of a chiral fermion is as follows. In this paragraph only, we take X to have a Lorentz signature metric. Then we consider time evolution of the quantum field Ψ, now regarded as a quantum mechanical operator in the Heisenberg picture. The equation of motion is given by

$$0 = (\not{D}_Y + m)\Psi = (\not{D}_X + \overline{\gamma}\partial_\tau + m)\Psi, \tag{4.13}$$

where τ is the coordinate of $(-\infty, 0]$, $\overline{\gamma} = \gamma_0$ is the gamma matrix in this direction, and $\not{D} = \sum_{a=1}^{d} \gamma_a D_a$. This equation has solutions of the form

$$\Psi = \psi \exp(-m\tau), \qquad \overline{\gamma}\psi = \psi, \qquad \not{D}_X \psi = 0, \tag{4.14}$$

where ψ depends only on the coordinates of X. The equations $\overline{\gamma}\psi = \psi$ and $\not{D}_X\psi = 0$ imply that ψ satisfies the equations required for a chiral fermion on X with the chirality operator $\overline{\gamma}$. This class of solutions is exponentially localized near the boundary $\tau = 0$ if m is negative. On the other hand, if m is positive, this is exponentially growing and not normalizable at all. In this sense, these solutions exist only for $m < 0$.

To see more about localized modes, let us study eigenvalues of Δ_m given in (4.5) on the space (4.12) with a Euclidean signature metric on X. Δ_m can be written as

$$\Delta_m = m^2 - \partial_\tau^2 - \not{D}_X^2. \tag{4.15}$$

Also, the boundary condition (4.6) for the eigenvalue problem of Δ_m can be rewritten as

$$(1 - \overline{\gamma})s|_{\partial X} = 0, \qquad (1 + \overline{\gamma})(m + \partial_\tau)s|_{\partial X} = 0 \tag{4.16}$$

where we have used $\gamma_0 = \overline{\gamma}$, $\{\overline{\gamma}, \not{D}_X\} = 0$, and $(1+\overline{\gamma})\overline{\gamma} = (1+\overline{\gamma})$. Therefore, the eigenvalues of Δ_m are just of the form

$$\Delta_m = m^2 + \omega^2 + \lambda_X^2, \tag{4.17}$$

where λ_X^2 represents eigenvalues of $-\not{D}_X^2$, and ω^2 represents eigenvalues of $-\partial_\tau^2$ with the boundary condition (4.16). Eigenmodes of $-\partial_\tau^2$ with eigenvalue ω^2 are explicitly given by

$$s = \left((m - \mathrm{i}\omega)e^{\mathrm{i}\omega\tau} - (m + \mathrm{i}\omega)e^{-\mathrm{i}\omega\tau}\right) s_+ + (e^{\mathrm{i}\omega\tau} - e^{-\mathrm{i}\omega\tau})s_- \tag{4.18}$$

where $s_\pm$ are independent of τ and $\overline{\gamma}s_\pm = \pm s_\pm$. One can check that this satisfies the boundary condition. We require that these modes do not grow at $\tau \to -\infty$. Then ω must

be real, or $\omega = \pm \mathrm{i}m$ for $m < 0$. Solutions for $\omega = \mathrm{i}m$ and $-\mathrm{i}m$ are just equivalent, so we only need to consider one of them. In this case, we get

$$s \propto e^{-m\tau} s_+, \qquad \Delta_m s = \lambda_X^2 s. \tag{4.19}$$

Therefore, we obtain the following conclusion. There are two types of eigenvalues of Δ_m for $m < 0$. One type is of the form $m^2 + \omega^2 + \lambda_X^2$ for real ω. This is always greater than or equal to m^2, and the corresponding modes do not decay at $\tau \to -\infty$. The other type of eigenmodes is the form λ_X^2, and the corresponding modes are localized near the boundary by the exponential factor $e^{-m\tau}$. On the other hand, if $m > 0$, only the type of eigenvalues $m^2 + \omega^2 + \lambda_X^2 \geq m^2$ exists.

We want to obtain a boundary chiral fermion, so we take $m < 0$. On the other hand, we do not want the unphysical Pauli-Villars field to have such a boundary chiral field, so we take $M > 0$. If we use a different boundary condition $(1 + \gamma_0)\Psi|_{\partial Y} = 0$ rather than $(1 - \gamma_0)\Psi|_{\partial Y} = 0$, we can realize chiral fermions with negative chirality $\overline{\gamma}\psi = -\psi$ for $m > 0$ and $M < 0$.

For more general manifolds Y, we have assumed that the neighborhood of the boundary $\partial Y = X$ is of the form $(-\epsilon, 0] \times X$. Although the above analysis is not exactly valid, we expect that there are small eigenvalues of order λ_X^2 if we take $|m|\epsilon \gg 1$. The reason is that the eigenmodes with $\omega = \mathrm{i}m$ found above are exponentially localized near the boundary, so it is exponentially small at $\tau = -\epsilon$ by a factor $e^{-|m|\epsilon}$. By taking $|m| \to \infty$, the approximate eigenvalues λ_X^2 are expected to become exact. Based on this consideration and physical intuition, we might expect the following. If we take the counterterm $-\mathcal{S}_{\text{c.t.}}$ appropriately, the absolute value of the partition function $\mathcal{Z}(Y, \mathsf{L})$ is given by

$$|\mathcal{Z}_\Psi(Y, \mathsf{L})| \xrightarrow{m \to -\infty} |\operatorname{Pf}(\not{D}_X)| = |\mathcal{Z}_\psi(X)| \tag{4.20}$$

where we have defined $\mathcal{Z}_\psi(X) = \operatorname{Pf}(\not{D}_X)$ which takes values in the Pfaffian line $\mathcal{L}_X$. A physical intuition behind this claim is that eigenvalues $m^2 + \omega^2 + \lambda_X^2$ with real ω go to infinity in the large mass limit $|m| \to \infty$, so they are high energy or in other words short distance contributions, and hence they may be cancelled by choosing an appropriate counterterm $\mathcal{S}_{\text{c.t.}}$. This is not exactly true for the phase part of $\mathcal{Z}_\Psi(Y, \mathsf{L})$ and that point will be crucial later, but for the absolute value we have (4.20).

4.3 Pfaffian lines and physical Hilbert spaces

Now we make an important observation. In Sec. 3.1, we discussed some axioms of quantum field theory. Massive fermions are expected to satisfy these axioms. In massive fermions, there is a one-dimensional subspace $\mathcal{H}_0(X) \subset \mathcal{H}(X)$ of the Hilbert space on X which is spanned by the ground state (i.e. the lowest energy state). Moreover, in the large mass limit $|m| \to \infty$, all other states have very high energies and we can neglect these states for low energy (long distance) physics. Therefore, we just think $\mathcal{H}(X)$ to be the space of the ground state in the following, and omit the subscript 0. This approximation becomes exact in the limit $|m| \to \infty$. In particular, $\dim \mathcal{H}(X) = 1$ in this approximation.

One of the axioms of Sec. 3.1 states the following. Suppose we have a manifold Y with a boundary $\partial Y = X$ whose H structure is determined by using the outward normal vector

to ∂Y. Then we have a map $\mathcal{Z}_\Psi(Y) \in \mathrm{Hom}(\mathcal{H}(\varnothing), \mathcal{H}(X)) \cong \mathcal{H}(X)$ where we have used $\mathcal{H}(\varnothing) \cong \mathbb{C}$. Let us denote it as $|Y\rangle$,

$$|Y\rangle = \mathcal{Z}_\Psi(Y) \in \mathcal{H}(X). \tag{4.21}$$

We have introduced a boundary condition (4.2). In the language of the Hilbert space, it is interpreted as an (unbounded) linear map from $\mathcal{H}(X)$ to $\mathbb{C}$, which we denote as $\langle \mathsf{L}|$,

$$\langle \mathsf{L}| : \mathcal{H}(X) \to \mathbb{C}, \tag{4.22}$$

such that the partition function $\mathcal{Z}(Y, \mathsf{L})$ with the boundary condition L is given by

$$\mathcal{Z}(Y, \mathsf{L}) = \langle \mathsf{L}|Y\rangle \in \mathbb{C}. \tag{4.23}$$

We mentioned that we can focus on the ground state, which we denote as $|\Omega\rangle$. This is a basis vector of $\mathcal{H}_0(X)$, and it is unique up to phase since a physical state is determined by a ray in the Hilbert space and we can take the absolute value as $\langle\Omega|\Omega\rangle = 1$. Therefore, we have $|Y\rangle \propto |\Omega\rangle$ in the limit $|m| \to \infty$. So we can write

$$\mathcal{Z}(Y, \mathsf{L}) = \langle \mathsf{L}|\Omega\rangle\langle\Omega|Y\rangle. \tag{4.24}$$

Thus the partition function consists of the inner product of two vectors,

$$|Y\rangle = |\Omega\rangle\langle\Omega|Y\rangle \in \mathcal{H}, \qquad \langle \mathsf{L}|\Omega\rangle\langle\Omega| \in \mathcal{H}^*. \tag{4.25}$$

They are defined to be independent of a choice of the phase of $|\Omega\rangle$.

By definition, we have $|Y\rangle = \mathcal{Z}_\Psi(Y)$. On the other hand, we have seen in the previous subsection that the boundary condition L leads to the chiral fermion ψ which is localized near the boundary. The chiral fermion has $\mathcal{Z}_\psi(X) = \mathrm{Pf}(\not{D}_X)$ which is an element of the Pfaffian line $\mathcal{L}_X$. This is determined completely on the boundary. The quantity $\langle \mathsf{L}|\Omega\rangle\langle\Omega|$ is also determined on the boundary.

Based on the above observations, we claim the following. First, we claim that the Pfaffian line is actually isomorphic to $\mathcal{H}^*(X)$ (regarded as a one-dimensional vector space in the limit $|m| \to \infty$),

$$\mathcal{L}_X \cong \mathcal{H}^*(X). \tag{4.26}$$

Under this isomorphism, we claim

$$\mathcal{Z}_\psi(X) = \langle \mathsf{L}|\Omega\rangle\langle\Omega|. \tag{4.27}$$

Then we have (under an appropriate choice of the counterterm $\mathcal{S}_{\text{c.t.}}$),

$$\mathcal{Z}_\Psi(Y, \mathsf{L}) = \mathcal{Z}_\psi(X)\mathcal{Z}_\Psi(Y), \tag{4.28}$$

where the limit $m \to -\infty$ is implicit. These are the main claims. This is based on the analysis of the previous subsection that ψ appears as a boundary mode of Ψ. See [3] for more details.

Let us check the above idea. We mentioned that the absolute value of $\mathcal{Z}_\Psi(Y, \mathsf{L})$ is proportional to $\Delta_m(Y)^{1/4}$ as in (4.8). It contains a product $\prod(\lambda_X^2)^{1/4} = |\mathrm{Pf}(\not{D}_X)| = |\mathcal{Z}_\psi(X)|$ which comes from the eigenvalues of the localized modes. Thus, both sides of (4.28) contains this product of eigenvalues. Other eigenvalues of Δ_m are all greater than or equal to m^2, and they are included in the part $\mathcal{Z}_\Psi(Y)$. We claim that we can take $|\mathcal{Z}_\Psi(Y)| = 1$ by choosing the counterterm $\mathcal{S}_{\text{c.t.}}$ appropriately in the limit $|m| \to \infty$.

The mathematics corresponding to the above claims is known as the Dai-Freed theorem [14].

4.4 Connection on the Pfaffian line bundle

Using the identification between the Pfaffian line $\mathcal{L}$ and the dual of the Hilbert space $\mathcal{H}$, we can now discuss a connection on the line bundle $\mathcal{L}$ over the moduli space of metrics and connections $\mathcal{M}$ which we have introduced in Sec. 3.3.

The Hilbert spaces form a bundle $\mathcal{H}$ over $\mathcal{M}$ by combining $\mathcal{H}(X)$ for all $X \in \mathcal{M}$ as in the case of the Pfaffian lines.

On $\mathcal{M}$, we also take a universal fiber bundle

$$\begin{array}{ccc} X & \longrightarrow & \mathcal{F} \\ & & \Big\downarrow{\scriptstyle \pi} \\ & & \mathcal{M} \end{array} \tag{4.29}$$

Namely, a fiber at $X \in \mathcal{M}$ is the manifold X itself. Since the notation is now quite confusing, we represent points on $\mathcal{M}$ as p_X. Then, the fiber at p_X is $X = \pi^{-1}(p_X)$ where X is just regarded as a manifold, while p_X contains the information of a metric and connection on X.

We take a metric and connection on the total space $\mathcal{F}$ such that its restriction on each fiber $X = \pi^{-1}(p_X)$ is the metric and connection specified by the point $p_X \in \mathcal{M}$. We do not try to make these discussions rigorous, but just notice that we can model the situation by considering a finite dimensional manifold M and a fiber bundle $X \to F \to M$. Thus the infinite dimensionality and singularities of $\mathcal{M}$ can be avoided in the following discussion.

Once we specify such a metric and connection on $\mathcal{F}$, we can define a canonical connection on the bundle $\mathcal{H} \cong \mathcal{L}^*$ as follows. We consider a path γ on $\mathcal{M}$,

$$\gamma : [0, 1] \to \mathcal{M}. \tag{4.30}$$

The pullback of $\mathcal{F}$ by γ, $\gamma^* F$, is a bundle whose total space is a $(d+1)$-dimensional manifold. Let us set $Y = \gamma^* F$. We have a metric and connection on Y by pullback from $\mathcal{F}$. Y has boundaries $\partial Y = X_{\text{in}} \sqcup X_{\text{out}}$, where $X_{\text{in}} = \pi^{-1}(\gamma(0))$ and $X_{\text{out}} = \pi^{-1}(\gamma(1))$. One of the axioms discussed in Sec. 3.1 says that for such Y, we have an associated map

$$\mathcal{Z}_\Psi(Y) : \mathcal{H}(X_{\text{in}}) \to \mathcal{H}(X_{\text{out}}). \tag{4.31}$$

This is independent of reparametrization of $[0, 1]$, since the metric and connection is determined by the pullback from $\mathcal{F}$. We claim that this defines parallel transport and a

unitary connection on the Hilbert space bundle $\mathcal{H}$ on $\mathcal{M}$, if we choose the counterterm $\mathcal{S}_{\text{c.t.}}$ appropriately. We demonstrate it by explicitly computing its curvature 2-form below.

The holonomy of the connection associated to a loop

$$\gamma : S^1 \to \mathcal{M} \tag{4.32}$$

can be computed as the partition function $\mathcal{Z}_\Psi(Y)$ on the closed manifold $Y = \gamma^* \mathcal{F}$. This holonomy is called a Berry phase of the ground state $|\Omega\rangle$ in physics. Let us determine this holonomy.

More generally, we compute the partition function $\mathcal{Z}_\Psi(Y)$ on any closed manifold Y. It is defined by (4.7) and we also take the limit $m \to -\infty$. We assume that we can freely take the limits in any order so that on closed manifolds we have

$$\mathcal{Z}_\Psi(Y) = \lim_{M\to\infty} \left[\frac{\mathrm{Pf}(\not{D}_Y - M)}{\mathrm{Pf}(\not{D}_Y + M)} \exp(-\mathcal{S}_{\text{c.t.}}) \right], \tag{4.33}$$

where we have set $m = -M$. Each eigenvalue of $\mathsf{i}\not{D}_Y$ appears twice due to the existence of the antilinear map C discussed in Sec. 2.2. The reason is as follows. Let s be a section with eigenvalue $\mathsf{i}\not{D}_Y s = \lambda s$. Then, by using $\mathsf{C}(\mathsf{i}\not{D}_Y s) = \mathsf{i}\not{D}_Y \mathsf{C}(s)$, we see that $\mathsf{C}(s)$ is also an eigenfunction with eigenvalue λ. It is not possible to have $\mathsf{C}(s) \propto s$ because if $\mathsf{C}(s) = \alpha s$ for some $\alpha \in \mathbb{C}$ then $\mathsf{C}^2(s) = \mathsf{C}(\alpha s) = \overline{\alpha}\mathsf{C}(s) = |\alpha|^2 s$, but this contradicts with $\mathsf{C}^2 = -1$. Therefore, s and $\mathsf{C}(s)$ are different eigenfunctions with the same eigenvalue. Then the ratio of the Pfaffians above is simply given by

$$\frac{\mathrm{Pf}(\not{D}_Y - M)}{\mathrm{Pf}(\not{D}_Y + M)} = \prod{}' \left(\frac{-\mathsf{i}\lambda - M}{-\mathsf{i}\lambda + M} \right) \tag{4.34}$$

where the product $\prod'$ is over all pairs $(s, \mathsf{C}(s))$. Namely, we take only one λ from two eigenfunctions $(s, \mathsf{C}(s))$. Notice that the absolute value of each factor $(-\mathsf{i}\lambda - M)/(-\mathsf{i}\lambda + M)$ is 1, so it is a phase. This justifies the claim that $\mathcal{Z}_\Psi(Y)$ can be taken to have absolute value 1 under an appropriate choice of the counterterm $\mathcal{S}_{\text{c.t.}}$ (at least if Y is closed.) We can just take $\mathcal{S}_{\text{c.t.}}$ to be pure imaginary, such as $\mathcal{S}_{\text{c.t.}} = 0$, in the above equation.

We define $s(\lambda)$ as

$$\frac{-\mathsf{i}\lambda - M}{-\mathsf{i}\lambda + M} = \exp(-2\pi \mathsf{i} s(\lambda)), \qquad -\frac{1}{2} < s(\lambda) \le \frac{1}{2}. \tag{4.35}$$

This has a limit

$$s(\lambda) \xrightarrow{M\to\infty} \begin{cases} +\frac{1}{2} & \lambda \ge 0 \\ -\frac{1}{2} & \lambda < 0 \end{cases} \tag{4.36}$$

We define the Atiyah-Patodi-Singer (APS) η invariant [15] of a Dirac operator $\mathsf{i}\not{D}_Y$ as

$$\eta(\mathsf{i}\not{D}_Y) = \lim_{M\to\infty} \sum s(\lambda) \tag{4.37}$$

where the sum is taken over all eigenmodes, i.e., two contributions from two modes $(s, \mathsf{C}(s))$. This infinite sum requires regularization, and we assume some appropriate regularization has been done. Then we get

$$\begin{aligned}&\prod{}'\left(\frac{-\mathsf{i}\lambda - M}{-\mathsf{i}\lambda + M}\right) = \exp\left(-2\pi\mathsf{i}{\sum}'s(\lambda)\right) = \exp\left(-\pi\mathsf{i}\sum s(\lambda)\right)\\ &\to \exp\left(-\pi\mathsf{i}\eta(\mathsf{i}\not{D}_Y)\right).\end{aligned} \tag{4.38}$$

We finally get

$$\mathcal{Z}_\Psi(Y) = \exp\left(-\pi\mathsf{i}\eta(\mathsf{i}\not{D}_Y) - \mathcal{S}_{\text{c.t.}}\right). \tag{4.39}$$

This is valid for general closed manifolds Y.

In particular, if we take $Y = \gamma^*\mathcal{F}$, then we obtain the formula for the holonomy of the unitary connection on $\mathcal{H} \cong \mathcal{L}^*$ around a loop $\gamma : S^1 \to \mathcal{M}$.

We can compute the curvature of the above connection, by using the APS index theorem [15]. (See also [16–18].) First we need to explain the setup for the APS index theorem in our context of manifolds with H structure.

For the APS index theorem, we will consider a $(d+2)$-dimensional manifold, and hence we need to construct a $\mathbb{Z}_2$-graded Clifford module $\widetilde{S}$ associated to the vector space $\mathbb{R}^{d+2}$. $\widetilde{S}$ is also required to be a representation of H_{d+2}. We take it to be

$$\widetilde{S} = S \oplus S, \tag{4.40}$$

where S is the Clifford module for $\mathbb{R}^{d+1}$ that we have been using so far. A chirality (or $\mathbb{Z}_2$-grading) operator $\overline{\Gamma}$ and gamma matrices Γ_a $(a = -1, 0, \cdots, d)$ acting on $\widetilde{S}$ are defined as

$$\overline{\Gamma} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \Gamma_{-1} = \begin{pmatrix} 0 & \mathsf{i} \\ -\mathsf{i} & 0 \end{pmatrix}, \qquad \Gamma_a = \begin{pmatrix} 0 & \gamma_a \\ \gamma_a & 0 \end{pmatrix} \quad (a = 0, \cdots, d). \tag{4.41}$$

These gamma matrices satisfy the Clifford algebra. We also want to have an H_{d+2} representation on $\widetilde{S}$ satisfying the condition (2.11). For example, it is possible to define an H_{d+2} representation if the representation of $\mathfrak{so}(d+1) \subset \mathfrak{h}$ on S has been constructed purely by the Clifford algebra so that the corresponding fermion has spin 1/2. See Appendix B for that construction. For more general spins, it is not always true that $\widetilde{S}$ can be made into a representation of H_{d+2}. But here we simply assume that S is such that $\widetilde{S}$ is a representation of H_{d+2}. By using $\widetilde{S}$, we can define H structure on $(d+2)$-dimensional manifolds.

Now let us state the APS index theorem. We consider a $(d+2)$-dimensional manifold Z with an H structure and a $\mathbb{Z}_2$-graded Clifford module bundle $\widetilde{S}Z$. Z has a boundary ∂Z. On this manifold, we can define an index of the Dirac operator $\mathsf{i}\not{D}_Z$. By the APS index theorem, this index is given as follows. There is a certain polynomial of the curvature tensors which is determined by the representation of H. This polynomial is the one which appears in the usual Atiyah-Singer index theorem in $d+2$ dimensions, and we denote it as

$\mathcal{I}_{d+2}$, which is a $(d+2)$-form. Then the index on a manifold Z with boundary ∂Z is given by

$$\mathrm{index}(\mathrm{i}\not{D}_Z) = \int_Z \mathcal{I}_{d+2} + \eta(\mathrm{i}\not{D}_{\partial Z}). \tag{4.42}$$

Thus, compared to the case of the Atiyah-Singer index theorem, there is a boundary contribution $\eta(\mathrm{i}\not{D}_{\partial Z})$ which is determined completely on the boundary. We remark that we need $\mathcal{I}_{d+2}$ as a differential form, rather than just as a de Rham cohomology element.

The index is even in our case. This can be seen by defining an antilinear operator $\widetilde{\mathsf{C}}$ as

$$\widetilde{\mathsf{C}} = \begin{pmatrix} \mathsf{C} & 0 \\ 0 & \mathsf{C} \end{pmatrix}. \tag{4.43}$$

It satisfies $\widetilde{\mathsf{C}}(\mathrm{i}\not{D}_Z s) = \mathrm{i}\not{D}_Z \widetilde{\mathsf{C}}(s)$ and $\widetilde{\mathsf{C}}(\overline{\Gamma} s) = \overline{\Gamma}\widetilde{\mathsf{C}}(s)$ as one can check from the properties of C given in (2.18) (at least in the case of spin 1/2 constructed in Appendix B). It also satisfies $\widetilde{\mathsf{C}}^2 = -1$. Thus for a zero mode s, we get another zero mode $\widetilde{\mathsf{C}}(s)$ which has the same chirality (i.e. the same eigenvalue under $\overline{\Gamma}$) as s. This is the reason that the index is even,

$$\mathrm{index}(\mathrm{i}\not{D}_Z) \in 2\mathbb{Z}. \tag{4.44}$$

Now, suppose that the boundary of Z is Y. Then by using the above facts, we get

$$\exp\left(-\pi\mathrm{i}\eta(\mathrm{i}\not{D}_Y)\right) = \exp\left(\pi\mathrm{i}\int_Z \mathcal{I}_{d+2}\right). \tag{4.45}$$

Let us consider a map from a 2-dimensional disk D^2 to $\mathcal{M}$,

$$\widetilde{\gamma} : D^2 \to \mathcal{M}, \tag{4.46}$$

such that its restriction to the boundary $S^1 = \partial D^2$ is a loop $\gamma : S^1 \to \mathcal{M}$. Then we define $Z = \widetilde{\gamma}^*\mathcal{F}$ and $Y = \gamma^*\mathcal{F}$. Applying the above formula to this case, we get

$$\exp\left(-\pi\mathrm{i}\eta(\mathrm{i}\not{D}_Y)\right) = \exp\left(\pi\mathrm{i}\int_{\widetilde{\gamma}(D^2)}\int_X \mathcal{I}_{d+2}\right), \tag{4.47}$$

where we have computed the integral over Z by first integrating over fibers X and then on $\widetilde{\gamma}(D^2) \subset \mathcal{M}$.

Recall that a counterterm $\mathcal{S}_{\text{c.t.}}$ is an integral of an invariant local polynomial of curvatures and their derivatives on Y,

$$\mathcal{S}_{\text{c.t.}} = \int_Y \mathcal{L}_{\text{c.t.}}. \tag{4.48}$$

We restrict to the case that $\mathcal{L}_{\text{c.t.}}$ which is relevant to (4.39) is given by a characteristic $(d+1)$-form which takes values in the orientation bundle so that its integral on Y is well-defined. (For $\mathcal{S}_{\text{c.t.}}$ appearing in other formulas, $\mathcal{L}_{\text{c.t.}}$ need not be a characteristic class.)

If there is no such characteristic class, we simply take $\mathcal{S}_{\text{c.t.}} = 0$. By Stokes theorem, the contribution of $\mathcal{L}_{\text{c.t.}}$ vanishes when Y is of the form $Y = \partial Z$. Thus we get

$$\mathcal{Z}_\Psi(\gamma^* S^1) = \exp\left(- \int_{\widetilde{\gamma}(D^2)} \mathsf{F} \right) \tag{4.49}$$

where

$$\mathsf{F} = -\pi \mathsf{i} \int_X \mathcal{I}_{d+2}. \tag{4.50}$$

F is given by an integral of a $(d+2)$-form on a d-dimensional manifold X (i.e. pushforward), so it is a 2-form. Since $\mathcal{Z}_\Psi(\gamma^* S^1)$ is the holonomy on the loop γ, the formula (4.49) implies that F is the curvature tensor of the connection. The first Chern class of $\mathcal{H} \cong \mathcal{L}^*$ at the de Rham cohomology level is given by

$$c_1(\mathcal{L}^*)|_{\text{de Rham}} = \frac{\mathsf{i}\mathsf{F}}{2\pi} = \frac{1}{2} \int_X \mathcal{I}_{d+2}. \tag{4.51}$$

This is the famous perturbative anomaly for chiral fermions as written in the language of the family index theorem $\mathcal{L}$ [7–10].

4.5 Invertible topological phases

As is well-known in the theory of line bundles, a line bundle with a connection can have more structure than the curvature. More detailed information is contained in holonomies, and this was the understanding of global anomalies in the past [5]. Holonomies are computed by $\mathcal{Z}_\Psi(Y) = \exp\left(-\pi \mathsf{i} \eta(\mathsf{i} \not{D}_Y)\right)$ for $Y = \gamma^* S^1$. More recently, it was recognized [11–13] that the quantity $\mathcal{Z}_\Psi(Y)$ is important for not only Y of the form $\gamma^* \mathcal{F}$, but for arbitrary Y. This is the current understanding of global anomalies of fermions, and this understanding works very well in both topological material [11, 19] and string theory [12, 13, 20, 21]. Let us sketch the reason why arbitrary Y is important.

We have explained in Sec. 3.3 that there is a problem in defining the partition function of a chiral fermion $\mathcal{Z}_\psi(X) = \mathrm{Pf}(\not{D}_+)$ because it takes values in Pfaffian lines rather than $\mathbb{C}$. However, the discussions of this section suggests the following definition. The partition function $\mathcal{Z}_\Psi(Y, \mathsf{L})$ of a massive fermion on a manifold with boundary $\partial Y = X$ with the boundary condition L is almost the partition function of the chiral fermion ψ which appears as the boundary mode. Therefore, we try to define the chiral fermion partition function by $\mathcal{Z}_\Psi(Y, \mathsf{L})$ itself.[3] However, the problem is that $\mathcal{Z}_\Psi(Y, \mathsf{L})$ depends not only on X, but on a choice of Y.

Let us take another manifold Y' with the same boundary $\partial Y' = X$. If $\mathcal{Z}_\Psi(Y, \mathsf{L}) = \mathcal{Z}_\Psi(Y', \mathsf{L})$ for any Y, Y', then it is independent of a choice of Y and we can regard it as a

[3] The following discussions assume that X is such that there exists Y with $X = \partial Y$. If there is no such Y, that means that X is a nontrivial element of a certain bordism group. Then the phase of the partition function $\mathcal{Z}_\psi(X)$ depends on "generalized theta angles", which are cobordism invariant contributions to $\mathcal{Z}_\psi(X)$ [1, 2, 19]. Thus we need to choose the phase of $\mathcal{Z}_\psi(X)$ in each bordism class, that corresponds to a choice of generalized theta angles. In general, there is no canonical choice.

definition of the chiral fermion partition function on X. So let us compute the ratio between the two partition functions. By using (4.24) and the claim that $\langle\Omega|Y\rangle$ has absolute value 1, we compute

$$\frac{\mathcal{Z}_\Psi(Y,\mathsf{L})}{\mathcal{Z}_\Psi(Y',\mathsf{L})} = \frac{\langle\mathsf{L}|\Omega\rangle\langle\Omega|Y\rangle}{\langle\mathsf{L}|\Omega\rangle\langle\Omega|Y'\rangle} = \frac{\langle\Omega|Y\rangle}{\langle\Omega|Y'\rangle} = \langle Y|\Omega\rangle\langle\Omega|Y\rangle = \langle Y'|Y\rangle \tag{4.52}$$

where in the last step we have used the fact that the Hilbert space is spanned by $|\Omega\rangle$. By the axioms discussed in Sec. 3.1, the last quantity $\langle Y'|Y\rangle$ is the partition function on a closed manifold Y_{closed} obtained by gluing Y and Y' along the boundary W. For the gluing, we need to change the H structure on Y' so that the boundary X is regarded as an ingoing X_{in}, while the boundary of Y is outgoing X_{out}. See [1, 2] for detailed discussions of how this can be done. Therefore, we finally get

$$\frac{\mathcal{Z}_\Psi(Y,\mathsf{L})}{\mathcal{Z}_\Psi(Y',\mathsf{L})} = \mathcal{Z}_\Psi(Y_{\text{closed}}). \tag{4.53}$$

Therefore, if $\mathcal{Z}_\Psi(Y_{\text{closed}})$ is always equal to 1, then we get $\mathcal{Z}_\Psi(Y',\mathsf{L}) = \mathcal{Z}_\Psi(Y,\mathsf{L})$. In other words, the partition function $\mathcal{Z}_\Psi(Y_{\text{closed}})$ measures the obstruction to defining the chiral fermion partition function. This is the nonperturbative anomaly of the chiral fermion. Notice that $\mathcal{Z}_\Psi(Y)$ depends on a choice of a counterterm $\mathcal{S}_{\text{c.t.}}$, so the nonperturbative anomaly is described by $\exp(-\pi i\eta(i\not{D}_Y))$ up to $\exp(-\mathcal{S}_{\text{c.t.}})$. If there exists a choice of $\mathcal{S}_{\text{c.t.}}$ such that $\mathcal{Z}_\Psi(Y_{\text{closed}}) = 1$ for any Y_{closed}, then the chiral fermion does not have an anomaly. Otherwise it has an anomaly.

The $(d+1)$-dimensional theory with only a single ground state on any closed manifold is called invertible topological phases which was first introduced in [22]. See [23, 24]. Those theories give the low energy effective field theories of the bulk of various interesting physical systems, such as integer quantum Hall systems and topological insulators. Their partition function characterizes the anomalies of the boundary modes, such as chiral fermions, in the above way. We remark that although we focused on the computation of the curvature F of the Pfaffian line bundle in the previous subsection, theories with zero curvature $\mathsf{F} = 0$ is as interesting as the case of nonzero curvature. Topological insulators are examples with $\mathsf{F} = 0$. See [1, 2, 19] for classification of theories with $\mathsf{F} = 0$ in terms of cobordism invariants.

Acknowledgments

I would like to thank Yasuyuki Kawahigashi for invitation to the summer school on mathematical physics at the University of Tokyo, Komaba. KY is partly supported by JSPS KAKENHI Grant-in-Aid (Wakate-B), No.17K14265.

A Remark on Pauli-Villars regularization

We do not try to be rigorous about regularization, but here we want to sketch an example of a regularization by using the Pauli-Villars regularization only. Let us denote $m = M_0$.

Then we consider

$$\frac{\mathrm{Pf}(\not{D}+M_0)\,\mathrm{Pf}(\not{D}+M_2)\cdots\mathrm{Pf}(\not{D}+M_{2N-2})}{\mathrm{Pf}(\not{D}+M_1)\,\mathrm{Pf}(\not{D}+M_3)\cdots\mathrm{Pf}(\not{D}+M_{2N-1})}, \tag{A.1}$$

where the number N needs to be large enough to satisfy the condition which we will discuss below.

Let us roughly estimate when this kind of regularization makes the partition function well-defined. In a region of large eigenvalues $|\lambda|\to\infty$, we expand

$$\begin{aligned}\log\prod_{k=0}^{N-1}\left(\frac{-\mathsf{i}\lambda+M_{2k}}{-\mathsf{i}\lambda+M_{2k+1}}\right)&=\sum_{\ell=0}^{2N-1}(-1)^\ell\log\left(1+\frac{\mathsf{i}M_\ell}{\lambda}\right)\\ &=\sum_{n\geq 1}\frac{(-1)^{n-1}}{n}\frac{\mathsf{i}^n}{\lambda^n}\sum_{\ell=0}^{2N-1}(-1)^\ell M_\ell^n.\end{aligned} \tag{A.2}$$

The divergence from the infinite product is a short distance divergence, so it is enough to work in a flat space to find the leading divergent contribution since any manifold is locally flat. Then we have Fourier modes $e^{\mathsf{i}\,k\cdot x}$ for flat space coordinates $x\in\mathbb{R}^D$ and wave numbers $k\in\mathbb{R}^D$. We may replace the sum over λ with the phase space integral

$$\sum_\lambda\to\int\frac{\mathrm{d}^Dx\mathrm{d}^Dk}{(2\pi)^D} \tag{A.3}$$

Also, the eigenvalues λ behave roughly as $\pm|k|$. The integral over x gives the volume of the manifold. The integral over k is, after replacing $\lambda\to\pm|k|$,

$$\int\frac{\mathrm{d}^Dk}{(2\pi)^D}\frac{(\pm 1)}{|k|^n}. \tag{A.4}$$

This is absolutely convergent in the large $|k|$ region if $n>D$. Thus, for the partition function to be well-defined, we require

$$\sum_{\ell=0}^{2N-1}(-1)^\ell M_\ell^n=0\quad\text{for } n\leq D. \tag{A.5}$$

We choose M_ℓ to satisfy this condition.

B Constructing a Clifford module for $\mathbb{R}^{d+2}$ with a H_{d+2} representation

Suppose that we have a Clifford module S associated to the vector space $\mathbb{R}^{d+1}$ which is also a representation of H_{d+1} such that it satisfies the conditions discussed in Sec. 2. Moreover, we assume that the representation of the subalgebra $\mathfrak{so}(d+1)\subset\mathfrak{h}_{d+1}$ is constructed from the Clifford algebra as follows. Let T_{ab} $(a,b=0,\cdots,d)$ be generators of $\mathfrak{so}(d+1)$ with $T_{ba}=-T_{ab}$. They satisfy the algebra

$$[T_{ab},T_{cd}]=\delta_{bc}T_{ad}-\delta_{ac}T_{bd}-\delta_{bd}T_{ac}+\delta_{ad}T_{bc}. \tag{B.1}$$

We assume that their representation on S is constructed by the gamma matrices γ_a as

$$r(T_{ab}) = \frac{1}{4}(\gamma_a\gamma_b - \gamma_b\gamma_a). \tag{B.2}$$

This is the standard representation for a spin 1/2 fermion.

We define a Clifford module $\widetilde{S}$ for $\mathbb{R}^{d+2}$ as

$$\widetilde{S} = S \oplus S. \tag{B.3}$$

The gamma matrices Γ_a $(a = -1, 0, \cdots, d)$ acting on $\widetilde{S}$ are defined as

$$\Gamma_{-1} = \begin{pmatrix} 0 & \mathsf{i} \\ -\mathsf{i} & 0 \end{pmatrix}, \qquad \Gamma_a = \begin{pmatrix} 0 & \gamma_a \\ \gamma_a & 0 \end{pmatrix} \quad (a = 0, \cdots, d). \tag{B.4}$$

It is easy to check that they satisfy the $(d+2)$-dimensional Clifford algebra

$$\{\Gamma_a, \Gamma_b\} = 2\delta_{ab} \qquad (a, b = -1, 0, \cdots, d). \tag{B.5}$$

Our purpose in this appendix is to show that we can define a representation $\widetilde{r}$ of H_{d+2} acting on $\widetilde{S}$ which satisfies the conditions discussed in Sec. 2.

Because the commutative diagram (2.4) is a pullback diagram (see the explanation below (2.4)), we can always split an element $h \in H_{d+2}$ as $h = h_1 h_2$, where h_1 is in the subgroup $H_{d+1} \subset H_{d+2}$, and h_2 is in the subgroup $\mathrm{Spin}(d+2) \subset H_{d+2}$ generated by the subalgebra $\mathfrak{so}(d+2) \subset \mathfrak{h}_{d+2}$. Then h_2 can be represented by taking the representation of generators as

$$\widetilde{r}(T_{ab}) = \frac{1}{4}(\Gamma_a\Gamma_b - \Gamma_b\Gamma_a) \qquad (a, b = -1, 0, \cdots, d). \tag{B.6}$$

This defines a representation $\widetilde{r}(h_2)$ of the subgroup $\mathrm{Spin}(d+2)$. On the other hand, $h_1 \in H_{d+1}$ can be represented as

$$\widetilde{r}(h_1) = \begin{pmatrix} r(h_1) & 0 \\ 0 & r(h_1) \end{pmatrix}. \tag{B.7}$$

For elements $h \in H_{d+1} \cap \mathrm{Spin}(d+2) = \mathrm{Spin}(d+1)$, the above two definitions coincide due to our assumption (B.2). Thus we define $\widetilde{r}(h) = \widetilde{r}(h_1)\widetilde{r}(h_2)$. We will later check that this gives a homomorphism $r(hh') = r(h)r(h')$ for arbitrary h, h'.

We notice that for $h_1 \in H_{d+1}$, we can use (2.11) and the definitions of Γ_a to get

$$\widetilde{r}(h_1^{-1})\Gamma_a\widetilde{r}(h_1) = \rho(h_1)_{ab}\Gamma_b \tag{B.8}$$

where $\rho : H_{d+2} \to \mathrm{O}(d+2)$ is one of the defining data of H_{d+2} as discussed in Sec. 2, and $\rho(h)_{ab}$ are the matrix elements of $\rho(h)$ as an $\mathrm{O}(d+2)$ matrix with respect to the standard basis of $\mathbb{R}^{d+2}$. The same equation is also valid for $h_2 \in \mathrm{Spin}(d+2)$ by the standard Clifford algebra. Therefore, for general $h \in H_{d+2}$, we get

$$\widetilde{r}(h^{-1})\Gamma_a\widetilde{r}(h) = \rho(h)_{ab}\Gamma_b. \tag{B.9}$$

This is the $(d+2)$-dimensional version of the condition (2.11).

Finally, we want to show that $\widetilde{r}$ defined above satisfies $\widetilde{r}(hh') = \widetilde{r}(h)\widetilde{r}(h')$ for arbitrary h, h'. Two such elements $h = h_1 h_2$ and $h' = h_1' h_2'$ can be multiplied as $hh' = h_1 h_1' (h_1'^{-1} h_2 h_1') h_2'$. We need to study the factor $(h_1'^{-1} h_2 h_1')$.

To simplify the notation, we rename h_1' as $h \in H_{d+1}$. Also, we can replace h_2 by generators T_{ab} of the Lie subalgebra $\mathfrak{so}(d+2) \subset \mathfrak{h}_{d+2}$. The subalgebra generated by $h^{-1} T_{ab} h$ is the same subalgebra $\mathfrak{so}(d+2) \subset \mathfrak{h}_{d+2}$ generated by T_{ab}, since both of their images under $\rho : H_{d+2} \to \mathrm{O}(d+2)$ generate $\mathfrak{so}(d+2)$. More explicitly, by acting ρ, we get

$$\rho(h^{-1} T_{ab} h) = \rho(h^{-1}) \rho(T_{ab}) \rho(h) = \rho(h)_{ac} \rho(h)_{bd} \rho(T_{bc}). \tag{B.10}$$

Therefore, we get

$$h^{-1} T_{ab} h = \rho(h)_{ac} \rho(h)_{bd} T_{bc}. \tag{B.11}$$

Then by acting $\widetilde{r}$ to this equation, we obtain

$$\begin{aligned} &\widetilde{r}(h^{-1} T_{ab} h) = \rho(h)_{ac} \rho(h)_{bd}\, \widetilde{r}(T_{ab}) = \rho(h)_{ac} \rho(h)_{bd}\, \frac{1}{4} (\Gamma_a \Gamma_b - \Gamma_b \Gamma_a) \\ =& \widetilde{r}(h^{-1}) \frac{1}{4} (\Gamma_a \Gamma_b - \Gamma_b \Gamma_a) \widetilde{r}(h) \end{aligned} \tag{B.12}$$

where in the last step we have used (B.9). Therefore, we have obtained

$$\widetilde{r}(h^{-1} T_{ab} h) = \widetilde{r}(h^{-1}) \widetilde{r}(T_{ab}) \widetilde{r}(h). \tag{B.13}$$

This equation implies $\widetilde{r}(h_1^{-1} h_2 h_1) = \widetilde{r}(h_1^{-1}) \widetilde{r}(h_2) \widetilde{r}(h_1)$ for $h_1 \in H_{d+1}$ and $h_2 \in \mathrm{Spin}(d+2)$.

Now we can compute $\widetilde{r}(hh')$ for arbitrary h and h'. We split them as $h = h_1 h_2$ and $h' = h_1' h_2'$ for $h_1, h_1' \in H_{d+1}$ and $h_2, h_2' \in \mathrm{Spin}(d+2)$. Then by using $hh' = h_1 h_1' (h_1'^{-1} h_2 h_1') h_2'$, $(h_1'^{-1} h_2 h_1') \in \mathrm{Spin}(d+2)$, and the definition of $\widetilde{r}$, we get

$$\begin{aligned} &\widetilde{r}(hh') = \widetilde{r}(h_1) \widetilde{r}(h_1') \widetilde{r}(h_1'^{-1} h_2 h_1') \widetilde{r}(h_2') = \widetilde{r}(h_1) \widetilde{r}(h_1') \widetilde{r}(h_1'^{-1}) \widetilde{r}(h_2) \widetilde{r}(h_1') \widetilde{r}(h_2') \\ &= \widetilde{r}(h_1) \widetilde{r}(h_2) \widetilde{r}(h_1') \widetilde{r}(h_2') = \widetilde{r}(h) \widetilde{r}(h'). \end{aligned} \tag{B.14}$$

This completes the construction of a representation $\widetilde{r}$ of H_{d+2} on $\widetilde{S}$ with the desired property (B.9).

We can also introduce a $\mathbb{Z}_2$ grading to the above Clifford module, defined by

$$\overline{\Gamma} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{B.15}$$

It commutes with $\widetilde{r}(H)$ and anticommutes with Γ_a.

References

[1] D. S. Freed and M. J. Hopkins, *Reflection positivity and invertible topological phases*, arXiv:1604.06527.

[2] K. Yonekura, *On the cobordism classification of symmetry protected topological phases*, *Commun. Math. Phys.* **368** (2019), no. 3 1121–1173, [arXiv:1803.10796].

[3] E. Witten and K. Yonekura, *Anomaly Inflow and the η-Invariant*, in *The Shoucheng Zhang Memorial Workshop*, 9, 2019. arXiv:1909.08775.

[4] D. S. Freed, *ON DETERMINANT LINE BUNDLES*, *Conf. Proc.* **C8607214** (1986) 189–238.

[5] E. Witten, *GLOBAL GRAVITATIONAL ANOMALIES*, *Commun. Math. Phys.* **100** (1985) 197.

[6] K. Yonekura, *Dai-Freed theorem and topological phases of matter*, *JHEP* **09** (2016) 022, [arXiv:1607.01873].

[7] L. Alvarez-Gaume and E. Witten, *Gravitational Anomalies*, *Nucl. Phys. B* **234** (1984) 269.

[8] L. Alvarez-Gaume and P. H. Ginsparg, *The Topological Meaning of Nonabelian Anomalies*, *Nucl. Phys. B* **243** (1984) 449–474.

[9] O. Alvarez, I. Singer, and B. Zumino, *Gravitational Anomalies and the Family's Index Theorem*, *Commun. Math. Phys.* **96** (1984) 409.

[10] M. Atiyah and I. Singer, *Dirac Operators Coupled to Vector Potentials*, *Proc. Nat. Acad. Sci.* **81** (1984) 2597–2600.

[11] E. Witten, *Fermion Path Integrals And Topological Phases*, *Rev. Mod. Phys.* **88** (2016), no. 3 035001, [arXiv:1508.04715].

[12] E. Witten, *The "Parity" Anomaly On An Unorientable Manifold*, *Phys. Rev. B* **94** (2016), no. 19 195150, [arXiv:1605.02391].

[13] E. Witten, *World sheet corrections via D instantons*, *JHEP* **02** (2000) 030, [hep-th/9907041].

[14] X.-z. Dai and D. S. Freed, *eta invariants and determinant lines*, *J. Math. Phys.* **35** (1994) 5155–5194, [hep-th/9405012]. [Erratum: J.Math.Phys. 42, 2343–2344 (2001)].

[15] M. Atiyah, V. Patodi, and I. Singer, *Spectral asymmetry and Riemannian Geometry 1*, *Math. Proc. Cambridge Phil. Soc.* **77** (1975) 43.

[16] H. Fukaya, T. Onogi, and S. Yamaguchi, *Atiyah-Patodi-Singer index from the domain-wall fermion Dirac operator*, *Phys. Rev. D* **96** (2017), no. 12 125004, [arXiv:1710.03379].

[17] A. Dabholkar, D. Jain, and A. Rudra, *APS η-invariant, path integrals, and mock modularity*, *JHEP* **11** (2019) 080, [arXiv:1905.05207].

[18] H. Fukaya, M. Furuta, S. Matsuo, T. Onogi, S. Yamaguchi, and M. Yamashita, *The Atiyah-Patodi-Singer index and domain-wall fermion Dirac operators*, arXiv:1910.01987.

[19] A. Kapustin, R. Thorngren, A. Turzillo, and Z. Wang, *Fermionic Symmetry Protected Topological Phases and Cobordisms*, *JHEP* **12** (2015) 052, [arXiv:1406.7329].

[20] Y. Tachikawa and K. Yonekura, *Why are fractional charges of orientifolds compatible with Dirac quantization?*, *SciPost Phys.* **7** (2019), no. 5 058, [arXiv:1805.02772].

[21] C.-T. Hsieh, Y. Tachikawa, and K. Yonekura, *Anomaly inflow and p-form gauge theories*, arXiv:2003.11550.

[22] D. S. Freed and G. W. Moore, *Setting the quantum integrand of M-theory*, *Commun. Math. Phys.* **263** (2006) 89–132, [hep-th/0409135].

[23] D. S. Freed, *Anomalies and Invertible Field Theories*, *Proc. Symp. Pure Math.* **88** (2014) 25–46, [arXiv:1404.7224].

[24] S. Monnier, *A Modern Point of View on Anomalies*, *Fortsch. Phys.* **67** (2019), no. 8-9 1910012, [arXiv:1903.02828].