

THE ATIYAH-SINGER INDEX THEOREM

Dylan Pentland

Contents

1	Overview	1
2	Clifford bundles and Dirac operators	4
2.1	Clifford bundles	4
2.2	Dirac operators	6
2.3	Chern classes on Clifford bundles	8
3	The analytic and topological index	11
3.1	The analytic index	11
3.2	The topological index	13
4	Heat kernels	17
4.1	Calculus of Dirac operators	17
4.2	The supertrace of e^{-tD^2}	21
5	The proof of the index theorem	23
5.1	The McKean-Singer formula	23
5.2	Getzler's asymptotics	27
5.3	Putting it all together	35

1 Overview

There are various classical results in differential geometry which arise as subcases of the Atiyah-Singer index theorem that were discovered much before it, such as the Hirzebruch-Riemann-Roch theorem, the Hirzebruch signature theorem, or even the more classical Gauss-Bonnet theorem. The common thread between these results is that they express the analytic index of some differential operator in terms of cohomology, giving a topological formula. The question of finding a general formula for a larger class of differential operators was posed in 1960 by Gel'fand in [Gel60].

The Atiyah-Singer index theorem was originally proved in [AS63] in 1963 by Atiyah and Singer, where they applied cobordism theory to write down the analytic index in terms of characteristic classes. This was later reformulated in terms of K -theory in [AS68]. The idea is that the analytic index of an elliptic pseudo-differential operator D depends only on the symbol $\sigma(D)$. Further, the symbol naturally produces a class $[\sigma(D)] \in K(\mathrm{TX})$, and from there one defines a topological index on K -theory that is shown to agree with the analytic index. It turns out that the definition of the topological index is quite simple for K -theory. Let $i : \mathrm{TX} \rightarrow \mathrm{TR}^n$ be the map induced by an embedding of X into \mathbf{R}^n , and let $i_! : K(\mathrm{TX}) \rightarrow K(\mathrm{TR}^n)$ denote the wrong-way map (coming from the Thom isomorphism). We then set $\mathrm{t}\text{-ind}(D) := i_![\sigma(D)] \in \mathbf{Z}$, using the isomorphism $K(\mathrm{TR}^n) \simeq \mathbf{Z}$.

The K -theory formulation of the index theorem is then the following.

THEOREM 1.1. Let X be a compact manifold, and D an elliptic pseudo-differential operator on X . Then

$$\mathrm{a}\text{-ind}(D) = \mathrm{t}\text{-ind}(D).$$

What is shown in the proof is that when we interpret the notion of an index as certain assignments of homomorphisms $\mathrm{ind}_X : K(\mathrm{TX}) \rightarrow \mathbf{Z}$ for each compact Riemannian manifold X , then an index is uniquely determined by being the identity $K(*) \rightarrow \mathbf{Z}$ when $X = *$ and behaving as

$$\mathrm{ind}_X \circ i_! = \mathrm{ind}_Y$$

for a proper embedding $Y \rightarrow X$. It then suffices to prove the analytic and topological indices satisfy these properties and therefore must agree.

A downside of this formulation is that $K(\mathrm{TX})$ is a global object and it would be preferable to have an integral formula for the topological index, for example like we have in the Hirzebruch-Riemann-Roch theorem. This is indeed the case, and takes a similar form involving the Todd and Chern classes! It can be proven that

$$\mathrm{t}\text{-ind}(D) = (-1)^n \int_{\mathrm{TX}} \pi^* \mathrm{Td}(\mathrm{TX} \otimes \mathbf{C}) \mathrm{ch}([\sigma(D)]).$$

Here, π is the projection $\mathrm{TX} \rightarrow X$ and $\dim_{\mathbf{R}} X = n$. This is local in the sense that it is now an integral, and hence can actually be computed locally (unlike the K -theory definition).

A natural question to ask is whether or not one can deduce the integral formula directly, without involving global objects like $K(\mathrm{TX})$ in the proof. One way to do this is through the heat kernel approach, which is the method we will be using. We will focus on a more restrictive version of the index theorem for Dirac operators for this method, but it is worth noting that this is not an important restriction since elliptic pseudo-differential

operators look locally like Dirac operators, and the general index theorem can be deduced from the Dirac operator case. The version of the theorem we will show is the following, and definitions of each side will be given in later sections.

THEOREM 1.2 (Atiyah-Singer). Let X be an even-dimensional compact oriented Riemannian manifold, and let \mathcal{E} be a graded Clifford bundle on X with associated Dirac operator D . Then

$$\text{a-ind}(D) = \text{t-ind}(D),$$

where $\text{t-ind}(D)$ is defined as

$$\int_X \widehat{A}(X) \wedge \text{ch}(\mathcal{E}/S)$$

where $\text{ch}(\mathcal{E}/S)$ is the *twisted Chern class* of a Clifford module and \widehat{A} is the \widehat{A} -genus.

This can be deduced from the version of the theorem using K -theory, as previously mentioned. We will give a version of the proof of this theorem which directly proves the (now locally computable) integral formula for the topological index using heat kernels closely following the exposition in [Roe99]. The basic idea of the heat kernel proof is that both the analytic and topological indices have interpretations as supertraces via $\text{str}(e^{-tD^2})$, but in different limits. For the analytic index, this is not hard to show for $t \rightarrow \infty$, and for $t \rightarrow 0$ through some difficult asymptotics we can obtain the topological index. To prove the index theorem, one shows that this supertrace is independent of t and so both limits must be the same.

In more more detail, the first thing to show is that

$$\text{a-ind}(D) = \text{str}(e^{-tD^2})$$

as $t \rightarrow \infty$. Then, the McKean-Singer formula gives for all $t > 0$

$$\text{a-ind}(D) = \text{str}(e^{-tD^2}) = \int_X \text{str}(K_t(x, x)) \text{dvol}(x)$$

where K_t is the *heat kernel*. In particular, the supertrace is independent of t . Being able to directly write down an integral for the analytic index is part of the significance of this approach: by writing the index this way, we can more directly express it as an integral over the heat kernel and the main task is now to relate this integral to the one defining the topological index.

Appealing to independence of t we can analyze the integral $\int_X \text{str}(K_t(x, x)) \text{dvol}(x)$ in the limit as $t \rightarrow 0$ to compare to $\text{t-ind}(D)$. Using an asymptotic expansion of the heat

kernel due to Getzler in the limit as $t \rightarrow 0$, plugging this into the supertrace will yield

$$\int_X \text{str}(K_t(x, x)) \text{dvol}(x) = (4\pi)^{-n/2} \int_X \det^{1/2} \left(\frac{R/2}{\sinh(R/2)} \right)_{[n]} \wedge \exp(-F^{\mathcal{E}/S}).$$

Here, $\omega_{[n]}$ denotes the n -form part of $\omega \in \Omega^\bullet(X)$. One can see that the n -form part of $\det^{1/2} \left(\frac{R/2}{\sinh(R/2)} \right)$ term is just the \widehat{A} -genus applied to X and $\exp(-F^{\mathcal{E}/S})$ is the twisted Chern character. Thus, this is precisely the topological index, and so $\text{a-ind}(D) = \text{t-ind}(D)$.

Notation. Throughout, we will assume that X is a compact oriented even-dimensional Riemannian manifold unless otherwise stated. We will use $\text{a-ind}(D)$ for the analytic index of a differential operator, and $\text{t-ind}(D)$ for the topological index.

We will use \mathcal{E} to refer to a Clifford bundle on X , and D for the associated Dirac operator.

2 Clifford bundles and Dirac operators

2.1 Clifford bundles

In this section, I will discuss some preliminaries on Clifford modules and Dirac operators following [Roe99] and [Fre87].

First, we will recall the classical notion of a Clifford algebra.

DEFINITION 2.1. Let V be a finite-dimensional real vector space and let Q be a quadratic form. The algebra $\text{Cl}(V, Q)$ is defined as

$$\text{T}(V) / \langle v \otimes w + w \otimes v + 2Q(v, w) \rangle,$$

where $\text{T}(V)$ denotes the tensor algebra of V .

There is an involution α on $\text{Cl}(V, Q)$ induced by the map $v \mapsto -v$. This map preserves the quadratic form and hence induces an automorphism α of the Clifford algebra.

As α is an involution, one may decompose $\text{Cl}(V, Q)$ into positive and negative eigenspaces of α . This gives a $\mathbf{Z}/2\mathbf{Z}$ grading

$$\text{Cl}(V, Q) = \text{Cl}_-(V, Q) \oplus \text{Cl}_+(V, Q).$$

The fact that the underlying vector space is canonically $\mathbf{Z}/2\mathbf{Z}$ -graded, or is a super vector space, means this is a superalgebra.

DEFINITION 2.2. A *graded Clifford module* \mathcal{E} is a module over $\text{Cl}(V, Q)$ equipped with a $\mathbf{Z}/2\mathbf{Z}$ -graded $\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-$ such that Clifford multiplication by $v \in V$ interchanges \mathcal{E}_+ and \mathcal{E}_- .

With this being taken as the local definition, we want to generalize to bundles on an even-dimensional compact oriented Riemannian manifold X . Call the metric on X g .

DEFINITION 2.3. Let $\text{Cl}(X)$ denote the natural bundle with fibers $\text{Cl}(T_x X, g_x)$. A Clifford bundle is a vector bundle \mathcal{E} with an action map

$$\text{cl} : \text{Cl}(X) \rightarrow \text{End}(\mathcal{E})$$

equipped with a Hermitian metric h and connection ∇ such that

- The Clifford action of each $v \in T_x X$ is skew-adjoint on \mathcal{E}_x with respect to h . That is,

$$h(\text{cl}(v)(s_1), s_2) + h(s_1, \text{cl}(v)(s_2)) = 0.$$

- The connection $\nabla_{\mathcal{E}}$ is compatible with the Levi-Civita connection ∇^{LC} on X , in the sense that

$$[\nabla_{\mathcal{E}}, \text{cl}(v)] = \text{cl}(\nabla^{\text{LC}}(v))$$

for $v \in C^\infty(\text{Cl}(X))$. Note here ∇^{LC} is really the induced connection on $\text{Cl}(TX)$.

We then want the graded version of this definition, which is now straightforward.

DEFINITION 2.4. A *graded Clifford bundle* is a Clifford bundle \mathcal{E} equipped with a decomposition

$$\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-$$

respecting the metric h and connection $\nabla_{\mathcal{E}}$ such that on each fiber Clifford multiplication exchanges \mathcal{E}_+ and \mathcal{E}_- (i.e. each fiber is a graded Clifford module).

REMARK 2.5. We will frequently use the adjective ‘super’ to indicate that an object comes equipped with a $\mathbf{Z}/2\mathbf{Z}$ -grading compatible with its usual structure. We could then also call a graded Clifford bundle a Clifford superbundle.

These adaptations are largely straightforward, for example a super vector space is a $\mathbf{Z}/2\mathbf{Z}$ -graded vector space $V = V_+ \oplus V_-$. We will often consider the supertrace of an endomorphism f , which on a vector space is $\text{tr}(f|_{V_+}) - \text{tr}(f|_{V_-})$.

2.2 Dirac operators

Clifford bundles on X come equipped with Dirac operators, and if graded come with graded Dirac operators.

DEFINITION 2.6. The Dirac operator D of a Clifford bundle \mathcal{E} is a first-order differential operator on $C^\infty(\mathcal{E})$ defined via the composition

$$D : C^\infty(\mathcal{E}) \xrightarrow{\nabla_{\mathcal{E}}} C^\infty(T^*X \otimes \mathcal{E}) \xrightarrow{\iota_h} C^\infty(TX \otimes \mathcal{E}) \xrightarrow{\text{cl}} C^\infty(\mathcal{E}).$$

where ι_h is given by the canonical identification of T^*X and TX given a Hermitian metric h (coming from \mathcal{E}) and the final map is induced by the Clifford action cl .

In local orthonormal frame e_i , we have

$$Ds = \sum_i \text{cl}(e_i) \nabla_i s.$$

For a graded Clifford bundle, the same method produces a Dirac operator D . It is now graded, in the sense it swaps sections of \mathcal{E}_+ and \mathcal{E}_- . This is simply due to the Clifford multiplication involved in the last part of the definition of D . This tells us that it has the form

$$D = \begin{pmatrix} 0 & D_- \\ D_+ & 0 \end{pmatrix}.$$

The operator D is formally self-adjoint on X as it is compact and oriented, meaning

$$\int_X \langle Ds_1, s_2 \rangle \text{dvol} = \int_X \langle s_1, Ds_2 \rangle \text{dvol}.$$

Note that this holds without assuming X is compact if we just take compactly supported sections. This is because we have the identity

$$\langle Ds_1, s_2 \rangle = \text{div}(V) + \langle s_1, Ds_2 \rangle$$

where $V = \sum_i \langle \text{cl}(e_i) s_1, s_2 \rangle e_i$ in an orthonormal frame e_i . Upon verifying this by expanding in local coordinates, since the integral of div will vanish the claim follows. Due to the form of D , this tells us $D_+^* = D_-$.

Classically, a Dirac operator means it squares to the Laplacian operator on a vector bundle. Theorem 2.7 justifies the name. Recall the symbol of a linear differential operator $D : C^\infty(\mathcal{E}) \rightarrow C^\infty(\mathcal{E})$ is the map

$$\sigma(D) : T^*X \otimes \mathcal{E} \rightarrow \mathcal{E}$$

given locally by $\sigma(D)_x : df_x \mapsto [D, f]_x \in \text{End}(\mathcal{E}_x)$. Here, we take $f \in C^\infty(X)$ and regard multiplication by it as an endomorphism of \mathcal{E} .

THEOREM 2.7. The operator D is a first-order elliptic operator. Moreover, the principal symbol of the operator D^2 is $\sigma(D)_x = -\|df\|^2$ so D^2 is a generalized Laplacian.

Proof. We just need to compute $[D, f]_x$. We write D as

$$C^\infty(\mathcal{E}) \xrightarrow{\nabla_{\mathcal{E}}} C^\infty(T^*X \otimes \mathcal{E}) \xrightarrow{\iota_h} C^\infty(TX \otimes \mathcal{E}) \xrightarrow{\text{cl}} C^\infty(\mathcal{E}),$$

and then note that the Leibniz rule for $\nabla_{\mathcal{E}}$ gives $[\nabla_{\mathcal{E}}, f] = df$ and so $[D, f] = \text{cl}(df)$ (here we used the identification ι_h). This shows the first claim. This also shows $[D^2, f]$ is $\text{cl}(df)^2 = -\|df\|^2$. \square

There is also an extremely useful Weitzenböck identity relating D^2 to a connection Laplacian.

THEOREM 2.8 (Weitzenböck identity). We have

$$D^2 s = \nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}} s + Ks$$

where Ks is the Clifford contraction of curvature and $\nabla_{\mathcal{E}}^*$ is the formal adjoint (with respect to the metric on \mathcal{E}).

Proof. This can be proved by expanding in a local orthonormal frame but becomes somewhat lengthy. See Proposition 2.2.26 in [Nic13] for a detailed derivation. \square

The operator $\nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}}$ appearing here should be thought of as analogous to the connection Laplacian, following the usual definition $\nabla^{LC,*} \nabla^{LC}$ on X for the connection Laplacian.

EXAMPLE 2.9. Let $D = d + d^*$ be the de Rham operator acting on Ω_X^\bullet , with grading operator given by $(-1)^q$ on Ω_X^q (this means the $+1$ eigenspace is \mathcal{E}_+ and the -1 eigenspace is \mathcal{E}_-). Exterior algebras $\Lambda^\bullet V$ naturally carry an action of $\text{Cl}(V, Q)$, which gives us the graded Clifford bundle structure. Giving the obvious connection d and a hermitian metric, we see the Dirac operator is $d + d^*$.

Note that D squares to the Hodge laplacian, which we know by the usual Weitzenböck formula differs from the connection Laplacian by a term involving the curvature. We see both of the previous theorems here: D^2 is a generalized Laplacian, and moreover it differs from the connection Laplacian by a curvature term.

2.3 Chern classes on Clifford bundles

The usual idea of a Chern class is that topological complex vector bundles on X can be regarded as principal $U(n)$ -bundles, and so are classified by $[X, BU(n)]$. A natural transformation

$$[-, BU(n)] \rightarrow H^i(-, \mathbf{Z})$$

amounts to an element of $H^i(BU(n), \mathbf{Z})$ by the Yoneda lemma. Upon computing the cohomology of $BU(n)$, we see it is generated by universal characteristic classes given by the Chern classes $c_i \in H^{2i}(BU(n), \mathbf{Z})$. This construction gives us the topological definition of the classes $c_i(-) \in H^{2i}(X, \mathbf{Z})$ which assemble into a total Chern class $c(-)$. The total Chern class $c(-) := \sum_i c_i(-)$ has the following properties.

- We have $c_0(\mathcal{E}) = 1$ and $c_i = 0$ for $i > \text{rank}(\mathcal{E})$.
- (Naturality) For a continuous map $f : Y \rightarrow X$, $c(f^*\mathcal{E}) = f^*c(\mathcal{E})$.
- (Whitney sum) If $\mathcal{E} \simeq \mathcal{E}' \oplus \mathcal{E}''$, then $c(\mathcal{E})$ is the cup product $c(\mathcal{E}') \smile c(\mathcal{E}'')$.
- (Normalization) If \mathcal{E} is the tautological line bundle on \mathbf{CP}^∞ , $c_1(\mathcal{E}) = e(\mathcal{E}_{\mathbf{R}})$ where e is the Euler class and $\mathcal{E}_{\mathbf{R}}$ is the underlying real bundle.

These properties uniquely characterize the total Chern class, which can be seen by applying the splitting principle. You can also check that $c_1(\mathcal{E})$ agrees with the Euler class for $\mathcal{O}(-1)$ on \mathbf{CP}^1 for the normalization, as this is easier to compute.

Chern classes can also be understood via Chern-Weil theory, which is a more productive perspective here. There, the following definition is made:

DEFINITION 2.10. Let V be a complex vector bundle. We define

$$\text{ch}(V) := [\text{tr}(\exp(-F_{\nabla}/2\pi i))] \in H^\bullet(X, \mathbf{Q}).$$

Here $F_{\nabla} \in \Omega^2(\text{End}(TX))$ is the curvature of any connection on V .

To check this agrees with the purely topological definition, we just verify the previously stated axioms (none of which are too difficult).

The definition of the twisted Chern class of a Clifford module is made in a similar spirit, but cannot be made as directly: both tr and F_{∇} are not appropriate for Clifford bundles, as they do not take into account that we have changed the category from vector bundles to graded Clifford bundles. For example, the notion of trace comes from vector spaces but we want our definition to come from the correct notion for graded Clifford modules.

Let \mathcal{E} be a Clifford bundle with Dirac operator

$$D : C^\infty(\mathcal{E}_+) \rightarrow C^\infty(\mathcal{E}_-).$$

The connection $\nabla_{\mathcal{E}}$ gives rise to a curvature $F_{\nabla_{\mathcal{E}}} \in \Omega^2(\text{End}(\mathcal{E}))$. We have an isomorphism

$$\text{End}(\mathcal{E}) \simeq \text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}),$$

where we are considering everything as super vector bundles. Here, $\text{Cl}(X)$ again refers to the bundle with fiber $\text{Cl}(T_x X, g_x)$ used in the definition of a Clifford bundle.

Thus, we may regard the curvature $F_{\nabla_{\mathcal{E}}}$ as an element of $\Omega^2(\text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}))$. It is also possible to regard the curvature R of the Levi-Civita connection on X as a section of this bundle.

Naturally, $R \in \Omega^2(\mathfrak{so}(TX))$. But there is a natural map

$$\mathfrak{so}(TX) \rightarrow \text{Cl}(X)$$

induced by fiberwise applying the isomorphism between $\mathfrak{so}(V)$ and $\mathfrak{spin}(V)$, which is the degree two component in the natural grading of $\text{Cl}(V, Q) = \text{Cl}^0 \oplus \text{Cl}^1 \oplus \dots \oplus \text{Cl}^n$ where $n = \dim V$. In particular, we may regard R as a section in $\Omega^2(\text{Cl}(X))$, and therefore a section $\text{cl}(R)$ of $\Omega^2(\text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}))$.

In local coordinates, by definition this means

$$\text{cl}(R) = \frac{1}{2} \sum_{k < \ell} g(R(e_i, e_j)e_k, e_\ell) \text{cl}(e^k) \text{cl}(e^\ell).$$

The map $\text{cl} : \text{Cl}(X) \rightarrow \text{End}(\mathcal{E})$ in the definition of a Clifford bundle, but here applied to covectors via the canonical identification through the metric g .

LEMMA 2.11. The difference

$$F^{\mathcal{E}/S} := F_{\nabla_{\mathcal{E}}} - \text{cl}(R) \in \Omega^2(\text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}))$$

actually lies in $\Omega^2(\text{End}_{\text{Cl}(X)}(\mathcal{E}))$.

Proof. See Proposition 2.2.23 in [Nic13]. □

DEFINITION 2.12. The section $F^{\mathcal{E}/S} = F_{\nabla_{\mathcal{E}}} - \text{cl}(R)$ of $\Omega^2(\text{End}_{\text{Cl}(X)}(\mathcal{E}))$ is the *twisting curvature* of \mathcal{E} .

REMARK 2.13. This is related to the curvature K in the Weitzenböck identity. After

expanding in local coordinates, one sees

$$\mathbf{K} = \frac{s(g)}{4} + \text{cl}(F^{\mathcal{E}/S}),$$

where s is the scalar curvature.

Next, we need to define the relative trace. We need a notion of supertrace on $\text{End}_{\text{Cl}(X)}(\mathcal{E})$, and the supertrace on $\text{End}(\mathcal{E})$ is not suitable as it does not use the Clifford bundle structure. However, there is a good way to identify $\text{End}_{\text{Cl}(X)}(\mathcal{E})$ with endomorphisms of a superbundle so that we can apply the usual notion.

DEFINITION 2.14. Assume $\dim_{\mathbf{R}} V$ is even. There exists a graded Clifford module S called the *spinor module* such that induced map

$$\text{Cl}(V) \otimes \mathbf{C} \rightarrow \text{End}_{\mathbf{C}}(S)$$

via the structure map is an isomorphism. This graded Clifford module S is unique up to isomorphism.

The existence and uniqueness claims in this definition are shown in Proposition 2.2.4 of [Nic13].

This gives rise to a *twisting space* $\mathcal{E}/S := \text{Hom}_{\text{Cl}(V) \otimes \mathbf{C}}(S, V)$. Then we have an isomorphism of superalgebras

$$\text{End}_{\text{Cl}(V)}(\mathcal{E}) \simeq \text{End}_{\mathbf{C}}(\mathcal{E}/S).$$

Call the image of F under this isomorphism F/S .

DEFINITION 2.15. Let \mathcal{E} be a graded Clifford module and let $F \in \text{End}_{\text{Cl}(V)}(\mathcal{E})$. We define the relative supertrace $\text{str}^{\mathcal{E}/S}(F) := \text{str}(F/S)$, using the natural superalgebra structure on $\text{End}_{\mathbf{C}}(\mathcal{E}/S)$.

We can globalize this definition to apply to a graded Clifford bundle \mathcal{E} by applying the relative supertrace fiberwise. With the appropriate definitions in place, we can define the twisted Chern class in the same spirit as Chern-Weil theory.

DEFINITION 2.16. Let \mathcal{E} be a graded Clifford module. The twisted Chern class is given by

$$\text{ch}(\mathcal{E}/S) := [\text{str}^{\mathcal{E}/S}(\exp(-F^{\mathcal{E}/S}/2\pi i))] \in H^{\bullet}(X, \mathbf{Q})$$

where $F^{\mathcal{E}/S}$ is the twisting curvature and $\text{str}^{\mathcal{E}/S}$ is the relative supertrace.

3 The analytic and topological index

3.1 The analytic index

With the definitions we have made so far, the definition analytic index is now fairly easy to write down. Beginning with a linear operator $T : V \rightarrow W$ of vector spaces, we can note that

$$\dim \ker T - \dim \operatorname{coker} T$$

is $\dim V - \dim W$ and hence constant, despite neither of these numbers being constant.

The operator D we consider is a map $C^\infty(X) \rightarrow C^\infty(X)$. In particular, we want to use infinite dimensional Hilbert spaces and not vector spaces. Moreover, the natural general class of operator to consider is that of Fredholm operators: these are operators T with $\ker T$ and $\operatorname{coker} T$ finite dimensional.

THEOREM 3.1. Let $\operatorname{Fred}(V)$ be the space of Fredholm operators of an infinite dimensional separable Hilbert space V . Then there is an isomorphism

$$\operatorname{ind} : [X, \operatorname{Fred}(V)] \simeq K^0(X).$$

This result is useful as motivation but not directly relevant to the proof of the index theorem, so we will only give a sketch illustrating the main idea.

Proof sketch from [Fre15]. Let $f : X \rightarrow \operatorname{Fred}(V)$ be a continuous map. To its homotopy class $[f]$, we can associate a virtual vector bundle $\operatorname{ind}(f) = \mathcal{E}$ with

$$\mathcal{E}_x = [\ker f(x)] - [\operatorname{coker} f(x)].$$

This gives a well-defined morphism of monoids

$$[X, \operatorname{Fred}(V)] \rightarrow K^0(X)$$

where $[X, \operatorname{Fred}(V)]$ obtains a monoid structure via composition in $\operatorname{Fred}(V)$. Given that $K^0(X)$ is an abelian group, it suffices to prove that this is an isomorphism of monoids.

The main idea behind surjectivity is that any class in $K^0(X)$ actually has the form $[\mathcal{E}] - N$ for an integer N and vector bundle \mathcal{E} . Since V is infinite dimensional separable, after picking a basis $\{e_i\}_{i \geq 1}$ we have shift operators T_k which send $e_i \mapsto e_{i-k}$, or 0 if $i \leq k$. These have $\dim \ker T_k - \dim \operatorname{coker} T_k = k$ since the cokernel is trivial. This means the image contains \mathbf{Z} , so it suffices to produce $[\mathcal{E}] - N_{\mathcal{E}}$ where $N_{\mathcal{E}}$ is some integer depending only on \mathcal{E} .

Given \mathcal{E} , we now want to construct the corresponding function $f : X \rightarrow \operatorname{Fred}(V)$. Embed \mathcal{E} in a trivial bundle $\tilde{\mathcal{E}}$, and let $p_x \in \operatorname{End}(\tilde{\mathcal{E}})$ be the orthogonal projection with

kernel \mathcal{E}_x . Embed $\tilde{\mathcal{E}} \rightarrow V$ and extend p_x by the identity on \mathcal{E}_x^\perp . The element $x \mapsto p_x \in [X, \text{Fred}(V)]$ gets sent to $[\mathcal{E}] - [\text{coker } p_x]$, where $\text{coker } p_x$ is constant. This proves surjectivity.

Next, we want to show injectivity. Suppose that $\text{ind}(f) = 0$. Then for some finite-dimensional vector space W there we have

$$\varphi : [\ker f(x)] \oplus \underline{W} \simeq [\text{coker } f(x)] \oplus \underline{W}$$

where \underline{W} denotes the trivial vector bundle with fiber W at each point. This is because $\mathcal{E} = [\ker f(x)] - [\text{coker } f(x)]$ has a trivialization.

Using the fiberwise isomorphisms φ_x , we obtain for $x \in X$ isomorphisms

$$f_x + \varphi_x : V \oplus W \rightarrow V \oplus W$$

where f_x is an isomorphism between $(\ker f(x))^\perp$ and $(\text{coker } f(x))^\perp$ induced by f . Then the map $t \mapsto f + t\varphi$ is a homotopy from f to this family $f + \varphi$. The latter is a family of isomorphisms, however, and it turns out that $\text{End}(V)^\times$ is contractible in the norm topology for any Hilbert space V . It follows that f was homotopically trivial to begin with. \square

Take $X = \text{pt}$. Then Theorem 3.1 says $\pi_0 \text{Fred}(V) \simeq \mathbf{Z}$, and the map is now given by

$$\dim \ker T - \dim \text{coker } T$$

which is an integer since for a Fredholm operator both spaces are finite-dimensional. Note that

$$\dim \text{coker } T = \dim \ker T^*,$$

so we can redefine this as $\dim \ker T - \dim \ker T^*$. This defines the index of a Fredholm operator and gives us an interpretation as measuring the connected component it is in. Moreover, on top of calculating connected components, the result also tells us that the homotopy type of $\text{Fred}(V)$ is $\text{BU} \times \mathbf{Z}$ as this is the homotopy type representing the functor K^0 .

In the context of a Dirac operator viewed as a Fredholm operator, noting $(D_+)^* = D_-$ we arrive at the following definition of the analytic index.

DEFINITION 3.2. Let D be a Dirac operator on a graded Clifford bundle. Then

$$D = \begin{pmatrix} 0 & D_- \\ D_+ & 0 \end{pmatrix}.$$

We define

$$\text{a-ind}(D) := \dim \ker D_+ - \dim \ker D_-.$$

Note that this is equivalently the supertrace of the projection $\pi : \mathcal{E} \rightarrow \ker D$, regarding the fibers as super vector spaces.

EXAMPLE 3.3. Continuing the example of the de Rham operator, if $D = d + d^*$ acting on Ω_X^\bullet , we have $\text{a-ind}(D) = \chi(X)$. De Rham's theorem tells us that

$$\sum_i (-1)^i \dim_{\mathbf{R}} H^i(X; \mathbf{R}) = \sum_i (-1)^i \dim_{\mathbf{R}} \Omega_X^i(X)$$

Note that after using de Rham's theorem to translate the sum to singular cohomology, we need to use that X is compact and oriented and even dimensional to apply Poincaré duality. This gives us the alternating sum of ranks of homology groups, which is the actual definition of $\chi(X)$.

An important fact about the analytic index is that it can in fact be calculated using the supertrace of rapidly decreasing functions of D^2 . We will see later it can be calculated as a supertrace of e^{-tD^2} independently of t . In fact, one can develop a calculus for Dirac operators in which $f(D^2)$ makes sense for any rapidly decaying function f (and we again have a similar supertrace interpretation) which can be used to prove the topological invariance of the analytic index.

3.2 The topological index

Next, we move to the definition of the topological index. This involves the construction of the \widehat{A} -genus, and the twisted Chern class of a graded Clifford bundle.

As we have already described Chern classes on Clifford bundles, we only need to define the \widehat{A} -genus which was originally one of the motivations for the discovery of the index theorem. The exposition on genus theory in this section is based on [MS74].

DEFINITION 3.4. A genus is a ring homomorphism

$$\Phi : \text{MSO}_* \rightarrow \mathbf{R}$$

where \mathbf{R} is some ring and MSO_* is the oriented cobordism ring.

For the \widehat{A} -genus, this is a homomorphism

$$\text{MSO}_* \rightarrow \mathbf{Q}.$$

Once MSO_* has been computed, it is possible to give a complete description of all genera to the rational numbers.

THEOREM 3.5 (Thom). We have $\text{MSO}_* \otimes \mathbf{Q} \simeq \mathbf{Q}[\mathbf{CP}^2, \mathbf{CP}^4, \mathbf{CP}^6, \dots]$. Here, we mean the cobordism classes of \mathbf{CP}^{2k} .

This tells us that the oriented cobordism ring is just a polynomial ring over \mathbf{Q} with generators in degree $4k$ rationally. It is now possible to identify what exactly the homomorphisms to \mathbf{Q} are. This also tells you that the data of a genus is just the assignments $\Phi(\mathbf{CP}^{2k})$.

To *compute* the value of the genus on a particular manifold X is a different question. The idea is to give a formula in terms of Pontryagin numbers, based on the sequence $\Phi(\mathbf{CP}^{2k})$. A Pontryagin number can be defined in terms of Chern classes.

DEFINITION 3.6. Let V be a real vector bundle. The Pontryagin class p_k is defined as $c_{2k}(V \otimes \mathbf{C})$ in $H^{4k}(X, \mathbf{Q})$, with total Pontryagin class $p = \sum_k p_k \in H^{4\bullet}(X, \mathbf{Q})$. A Pontryagin number is the output of pairing a cup product of Pontryagin classes p_{k_i} living in the top cohomology group of X with the fundamental class $[X]$.

Given a genus Φ specified by $\Phi(\mathbf{CP}^{2k})$, actually giving the formula for $\Phi(X)$ in terms of Pontryagin classes requires some additional setup. Define

$$\log_{\Phi}(x) = \sum_{k \geq 0} \Phi(\mathbf{CP}^{2k}) \frac{x^{2k+1}}{2k+1}$$

and set the characteristic series to be

$$K_{\Phi}(u) = \frac{u}{\log_{\Phi}^{-1}(u)}.$$

Finally, to give a formula, we define

$$K_{\Phi,n}(p_1, \dots, p_n) = [K_{\Phi}(u_1)K_{\Phi}(u_2) \dots]_{4n}$$

where $[-]_{4n}$ means taking the degree $4n$ part with u_i in degree $2i$. Here, the p_i are the i th symmetric polynomials in the u_j^2 . Hirzebruch's formula is then as follows.

THEOREM 3.7 (Hirzebruch). Given the sequence $\Phi(\mathbf{CP}^{2k})$, we can compute

$$\Phi(X) = \int_X K_{\Phi,n}(p_1, \dots, p_n)$$

where p_n is the n th Pontryagin class and $\dim X = 4n$.

Note that $K_{\Phi,n}(p_1, \dots, p_n)$ lives in the top cohomology class, so we can integrate it. Note that by the result that the oriented cobordism ring has generators in dimension $4k$, in other dimensions manifolds are cobordant to the identity (rationally). Thus, this tells us how to compute a genus via Pontryagin numbers.

DEFINITION 3.8. The \widehat{A} -genus is defined by the characteristic series

$$K_{\widehat{A}}(u) = \frac{\frac{1}{2}u}{\sinh(\frac{1}{2}u)}.$$

This yields the formula $\widehat{A}(X) = 1 - \frac{1}{24}p_1 + \frac{1}{5760}(-4p_2 + 7p_1^2) + \dots \in H^\bullet(X)$.

To get the actual number given by $\widehat{A} : \text{MSO}_* \rightarrow \mathbf{Q}$, you pair with the fundamental class. If X is a smooth manifold, we can also integrate a differential form.

Note that the \widehat{A} -genus arises in the framework we have already set up: when X is a closed spin manifold of even dimension, $\widehat{A}(X) = \text{a-ind}(D)$ where D is the Dirac operator on the spinor bundle \mathcal{S} by the index theorem. In particular, the \widehat{A} -genus is actually an *integer* in this case. But as the following examples show, this is not at all clear from the definition.

EXAMPLE 3.9. Consider $X = \mathbf{CP}^2$. Then we have a total Pontryagin class

$$p(X) = (1 + c^2)^3$$

where $c \in H^2(X, \mathbf{Z})$ is a generator. This means $\int_X p_1$ will give 3, so we get $\widehat{A}(\mathbf{CP}^2) = -1/8$.

As a more nontrivial example, let X be a degree four hypersurface in \mathbf{CP}^3 . This is a spin manifold since it is a K3 surface, defined by an exact sequence

$$0 \longrightarrow \mathcal{O}_{\mathbf{CP}^3}(-4) \longrightarrow \mathcal{O}_{\mathbf{CP}^3} \longrightarrow \mathcal{O}_X \longrightarrow 0.$$

Here, we can compute

$$\frac{-1}{24}p_1(X) = \frac{1}{24}(2c_2(X) - c_1(X)^2).$$

Now $c_1(X) = 0$ since the canonical bundle of this K3 surface is trivial: it is $\mathcal{O}_X(4 - 4) \simeq \mathcal{O}_X$.

Now $c_2(X) = 24$. Using the exact sequence, we see $\chi(\mathcal{O}_X) = 2$. But then Noether's formula for $\chi(\mathcal{O}_X)$ says this is $\frac{c_1(X)^2 + c_2(X)}{12}$, so $c_2(X) = 24$. Putting this all together, the \widehat{A} -genus is 2, an integer! Note that we are implicitly using a comparison between characteristic classes in algebraic geometry and topology here.

For our purposes, we would like a more explicit formula for a differential form corresponding to the \widehat{A} -genus. This is what we will use for $\widehat{A}(X)$ in the definition of $\text{t-ind}(D)$.

THEOREM 3.10. As usual, let X be a compact oriented even-dimensional Riemannian manifold. Equip TX with the Levi-Civita connection ∇^{LC} . Then

$$\widehat{A}(X) = \int_X \det^{1/2} \left(\frac{R/2}{\sinh(R/2)} \right).$$

where R is the curvature tensor regarded as a matrix of 2-forms in $\Omega^2(X, \text{End}(TX))$ so that we can take its determinant.

Proof. Assume $\dim X = 4n$. In terms of curvature, the total Pontryagin class has a formula

$$p = [\det(1 - R/2\pi)] = \prod_{i=1}^{2n} (1 + x_i^2)$$

where x_i are the Chern roots. Similarly, by expanding we obtain

$$\det^{1/2} \left(\frac{R/2}{\sinh(R/2)} \right) = \prod_{i=1}^{2n} \frac{\frac{x_i}{2}}{\sinh(\frac{1}{2}x_i)}$$

in terms of Chern roots.

Now take our definition

$$K_{\widehat{A}}(u) = \frac{\frac{1}{2}\sqrt{u}}{\sinh(\frac{1}{2}\sqrt{u})}.$$

Translating this into Chern roots, we get the exact same formula. Indeed, the form $\widehat{A}(X)$ is given by $K_{\widehat{A},n}(p_1, \dots, p_n)$, and writing p_j as the j th symmetric polynomial in the x_i^2 we see from Theorem 3.7 that

$$K_{\widehat{A},n}(p_1, \dots, p_n) = \left[\prod_{i=1}^{2n} \frac{\frac{1}{2}u_i}{\sinh(\frac{1}{2}u_i)} \right]_{4n} = \left[\prod_{i=1}^{2n} \frac{\frac{1}{2}x_i}{\sinh(\frac{1}{2}x_i)} \right]_{4n}$$

as $x_i^2 = u_i$. This then agrees with the previous formula (terms outside of degree $4n$ do not matter: the integral only cares about the degree $4n$ part of the form).

Note that we are free to assume $\dim X = 4n$, as in dimension $4n + 2$ the same formulas show both are zero (the Pontryagin classes appear in degrees divisible by 4). \square

This will be essential in proving the index theorem, as the heat kernel method can only hope to provide a differential form and not a cohomology class. It also provides a concise formula for the differential form we can hope to prove, unlike the formula for the cohomology class which requires some work to expand.

We can now define the topological index.

DEFINITION 3.11. Let \mathcal{E} be a graded Clifford bundle on an even-dimensional compact oriented Riemannian manifold X and let D be the associated Dirac operator. Then

$$\text{t-ind}(D) := \int_X \widehat{A}(X) \wedge \text{ch}(\mathcal{E}/S).$$

4 Heat kernels

4.1 Calculus of Dirac operators

Knowing D^2 is a generalized Laplacian, the heat equation on a graded Clifford module \mathcal{E} is given by

$$\frac{\partial s}{\partial t} + D^2 s = 0.$$

As with the usual heat equation, given an initial value s_0 which is smooth there is a unique smooth solution s_t for $t \geq 0$. To show uniqueness, note that two different solutions s_t and \tilde{s}_t have a difference which is a solution for $s_0 = 0$ that is nonzero. But one can compute

$$\frac{\partial}{\partial t} \|s_t\|^2 = -\langle D^2 s_t, s_t \rangle - \langle s_t, D^2 s_t \rangle = -2\|D^2 s_t\|^2 \leq 0.$$

It follows that $\|s_t\|^2 \leq \|s_0\|^2$, so if $s_0 = 0$ the only solution is the zero solution. Thus, solutions are unique.

For existence, there is an operator e^{-tD^2} which comes into play. Indeed, we can observe that

$$s_t := e^{-tD^2} s_0$$

gives a solution. This is because

$$\frac{\partial s_t}{\partial t} + D^2 s_t = -D^2 e^{-tD^2} s_0 + D^2 e^{-tD^2} s_0 = 0.$$

To make sense of these statements and develop the theory of the heat equation further, it is useful to develop a general functional calculus for Dirac operators on Clifford bundles \mathcal{E} .

Let $\text{Spec } D$ denote the set of eigenvalues of D . Being formally self-adjoint, eigenvalues of D will be real and it can be shown they form a discrete subset of \mathbf{R} . Moreover, they are countable and yield a decomposition

$$L^2(\mathcal{E}) = \bigoplus_{\lambda} H_{\lambda}$$

into finite dimensional eigenspaces. Recall that earlier we had thought of D as an operator purely on $C^{\infty}(\mathcal{E})$; here we are thinking of it as an *unbounded* operator on $L^2(\mathcal{E})$, so it only needs to be defined on the dense subspace $C^{\infty}(\mathcal{E})$.

Any section $s \in L^2(\mathcal{E})$ then has a Fourier expansion

$$s = \sum_{\lambda \in \text{Spec } D} s_{\lambda}$$

where s_{λ} is the component which belongs to the λ -eigenspace of D .

DEFINITION 4.1. Suppose that f is a bounded function on $\text{Spec } D \subset \mathbf{R}$. Define

$$f(D)s := \sum_{\lambda \in \text{Spec } D} f(\lambda)s_{\lambda}.$$

This now defines what is meant by e^{-tD^2} : we use $\lambda \mapsto e^{-t\lambda^2}$ as our function.

We first need to check for reasonable f that $f(D)$ is actually a well-behaved operator, since it is not immediately clear the output is in $L^2(\mathcal{E})$ or that it maps smooth sections to smooth sections. This requires the following estimate.

THEOREM 4.2 (Elliptic estimate for Dirac operators). Let D be a Dirac operator on a Clifford bundle \mathcal{E} . For any $k > 0$ there is a constant C_k such that for any $s \in C^{\infty}(\mathcal{E})$ we have

$$\|s\|_{k+1} \leq C_k(\|s\|_k + \|Ds\|_k)$$

where $\|\cdot\|_k$ denotes the Sobolev k -norm on $C^{\infty}(X)$. Note that this is non-canonical on $C^{\infty}(X)$ for X a compact Riemannian manifold, but well-defined up to equivalence).

Proof. See [Roe99] Proposition 5.16. □

We can then deduce the following lemma.

LEMMA 4.3. Let f be a bounded function on $\text{Spec } D$. The operator $f(D)$ is well-defined on $L^2(\mathcal{E})$, and maps $C^\infty(\mathcal{E})$ to $C^\infty(\mathcal{E})$.

Proof. Since $f(\lambda)$ is bounded, we have $\|\sum_{\lambda \in \text{Spec } D} f(\lambda)s_\lambda\| \leq N\|s\|$ for some N . The main difficulty is checking that this operator respects smooth sections.

To see this, we note that a section s of \mathcal{E} is smooth if and only if each component s_λ of the Fourier expansion has

$$\|s_\lambda\| = O(|\lambda|^{-N})$$

for each N . That is, the terms of the Fourier expansion are rapidly decreasing. It is clear that a smooth section satisfies this estimate, for the same reasons as usual Fourier analysis.

By Theorem 4.2, this means we have the estimate

$$\|s_\lambda\|_k \leq C_k \lambda^k \|s_\lambda\| \leq C_k \lambda^k \|s\|.$$

for the Sobolev k -norm (note that the 0th Sobolev k -norm is the L^2 norm). Here, we repeatedly apply the elliptic estimate for s_λ to reduce down to 0th Sobolev space, which is isomorphic to $L^2(\mathcal{E})$ by Parseval's theorem. Then we use $\|s_\lambda\| \leq \|s\|$. The rapid decay condition on s_λ implies convergence in each Sobolev space, and so for the resulting sum $\sum_\lambda s_\lambda$ all derivatives exist. It follows that s is smooth if and only if s_λ are rapidly decaying.

Now the claim that $f(D)$ sends smooth sections to smooth sections is clear, since the s_λ will remain rapidly decaying after applying $f(D)$. \square

If we control the growth of f , we can hope to show $f(D)$ produces smooth sections.

DEFINITION 4.4. A function f on $\text{Spec } D$ is called *rapidly decreasing* if $f(\lambda)$ is $O(|\lambda|^{-N})$ for each N .

DEFINITION 4.5. An operator $A : L^2(\mathcal{E}) \rightarrow L^2(\mathcal{E})$ is *smoothing* if we have for a particular section K of $\mathcal{E} \boxtimes \mathcal{E}^*$ over $X \times X$ that

$$As(y) = \int_X K(x, y)s(y)d\text{vol}(x)$$

for all smooth sections s .

The following proposition confirms this expectation that $f(D)$ produces smooth sections and justifies the terminology smoothing. Note that knowing A is smoothing is much

more useful than just knowing it produces smooth sections since it gives us the existence of a kernel K .

PROPOSITION 4.6. If f is rapidly decreasing, $f(D)$ maps $L^2(\mathcal{E})$ into $C^\infty(\mathcal{E})$. Moreover $f(D)$ a smoothing operator.

Proof. The first claim is immediate: if f is rapidly decreasing, then the Fourier components of $f(D)s$ are rapidly decreasing and hence $f(D)s$ is smooth.

Next, we consider the claim that $f(D)$ is smoothing (that is, it is described as $f(D)s(y) = \int_X K_t(x, y)s(y)d\text{vol}(x)$ for a kernel K_t). The main idea is that the projection operator

$$\pi_\lambda : s \mapsto s_\lambda$$

onto the λ -eigenspace is smoothing, which is just due to Fourier analysis.

Moreover we can control the growth of the smoothing kernel K_λ for π_λ , by choosing $\ell(k)$ for each k so that there is a Sobolev k -norm on $X \times X$ where the $\|K_\lambda\|_k \leq C_k \lambda^{\ell(k)}$. This part is somewhat nontrivial.

Let π be a smoothing operator with kernel K . We'll start with proving the claim that $\|K\|_2 \leq C\|\pi\|_{\text{op}}$ for some constant C depending on X where the latter is the operator norm

$$\|\pi\|_{\text{op}} = \max_{f, \|f\|_2=1} \left| \iint_{X \times X} K(x, y)f(y)d\text{vol}(y) \right|$$

of π as an operator $L^2(\mathcal{E}) \rightarrow C^0(\mathcal{E})$.

Let us consider the case of a trivial vector bundle \mathcal{E} where the integral kernel K is just a function. We can observe that the square of the L^2 norm of K is given by

$$\iint_{X \times X} |K(x, y)|^2 d\text{vol}(x)d\text{vol}(y) = \int_X \pi(\overline{K(x, -)})d\text{vol}(x),$$

which is seen by expanding π as an integral of K_λ (we use the conjugate so we get $|K(x, y)|^2$).

To prove $\|K\|_2 \leq C\|\pi\|_{\text{op}}$, we use the formula $\int_X \pi(\overline{K(x, -)})d\text{vol}(x)$ for the square of the L^2 norm of K . Now using the operator norm of π , we have a pointwise bound

$$\int_X \pi(\overline{K(x, -)})d\text{vol}(x) \leq \int_X \|\pi\|_{\text{op}} \|K(x, -)\|_2 d\text{vol}(x).$$

That is, $\|K(x, y)\|_2^2 \leq \int_X \|\pi\|_{\text{op}} \|K(x, -)\|_2 d\text{vol}(x)$. This is bounded above by

$$\left(\int_X \|\pi\|_{\text{op}}^2 \right)^{1/2} \left(\int_X \|K(x, -)\|_2^2 d\text{vol}(x) \right)^{1/2}$$

by Hölder's inequality applied to $\|K(x, -)\|_2$ and the constant $\|\pi\|_{\text{op}}$. But now we see $\int_X \|K(x, -)\|_2^2 \text{dvol}(x)$ is just $\int_{X \times X} |K(x, y)|^2 \text{dvol}(y) \text{dvol}(x) = \|K(x, y)\|_2^2$. Cancelling away a factor of $\|K(x, y)\|_2$, it follows that $\|K(x, y)\|_2 \leq \text{vol}(X)^{1/2} \|\pi\|_{\text{op}}$. By considering local trivializations, the case for general \mathcal{E} follows.

From this, specializing to the projection π_λ we want to show that the Sobolev k -norm of K_λ is bounded by $C_k \lambda^\ell$ for ℓ depending only on k . For $k = 0$, this is the L^2 norm, and what we have just shown is that $\|K_\lambda\|_2 \leq C \|\pi_\lambda\|_{\text{op}}$. The operator norm of an orthogonal projection is ≤ 1 , so the claim for $k = 0$ follows. Inductively, we can apply Theorem 4.2 to the Clifford module $\mathcal{E} \boxtimes \mathcal{E}^*$ on $X \times X$. We obtain

$$\|K_\lambda\|_{k+1} \leq C_k (\|K_\lambda\|_k + \|DK_\lambda\|_k),$$

where D is the induced Dirac operator on $\mathcal{E} \boxtimes \mathcal{E}^*$.

It follows that for rapidly decreasing f that

$$f(D) = \sum_{\lambda \in \text{Spec } D} f(\lambda) \pi_\lambda$$

converges in the space of smoothing kernels on $X \times X$ due to the bound on Sobolev k -norms of the kernels K_λ . In particular, we can write the desired smoothing kernel K as $\sum_\lambda f(\lambda) K_\lambda$ and this converges. \square

Specializing to $f(D) = e^{-tD^2}$, the functional calculus for Dirac operators now gives precise meaning to this for each t . Moreover, due to the exponential decay, we know that applying e^{-tD^2} will output smooth solutions, so it indeed makes sense to use $e^{-tD^2} s_0$ to give a solution to the heat equation on X .

We also now know, again due to the exponential decay, that there is a time-dependent *heat kernel* K_t such that

$$e^{-tD^2} s(y) = \int_X K_t(x, y) s(y) \text{dvol}(x).$$

4.2 The supertrace of e^{-tD^2}

Recall $L^2(\mathcal{E})$ has a decomposition according to $\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-$, and so we define $\text{str}(A) := \text{tr}(\varepsilon A)$ where ε is the grading operator induced on $L^2(\mathcal{E})$ by the decomposition: it is -1 on $L^2(\mathcal{E}_-)$ and 1 on $L^2(\mathcal{E}_+)$.

However, it is not entirely clear that $\text{tr}(\varepsilon A)$ has meaning. For this, we will need some harmonic analysis. If we have smoothing operator A on $L^2(\mathcal{E})$ given by a kernel K then we can write it $A = BC$ where B and C are Hilbert-Schmidt operators¹ represented by

¹Recall this means the Hilbert-Schmidt norm, essentially the matrix norm, is finite.

continuous kernels. This property is called being of trace-class, giving us a reasonable definition

$$\mathrm{tr}(A) := \langle B^*, C \rangle_{\mathrm{HS}} = \mathrm{tr}(BC).$$

where the inner product is the Hilbert-Schmidt inner product. Note that this is independent of the factorization.

The reason this factorization is possible is because $(1 + D^2)^{-N}$ for $N \gg 0$ will be Hilbert-Schmidt with continuous kernel as by repeating the argument of Proposition 4.6 shows we can get just a continuous kernel if $f(\lambda) = O(\lambda^{-N})$ for some large N . Then we can use this and $(1 + D^2)^N A$: the operator A is Hilbert-Schmidt since it is given by a smooth kernel, and $(1 + D^2)^N$ is bounded so the product is Hilbert-Schmidt. We are therefore in the position to make the following definition.

DEFINITION 4.7. Let $A : L^2(\mathcal{E}) \rightarrow L^2(\mathcal{E}) \cap C^\infty(\mathcal{E})$ be a smoothing operator where \mathcal{E} is a graded Clifford bundle. We set

$$\mathrm{str}(A) = \mathrm{tr}(\varepsilon A),$$

where we use that εA is again smoothing and has a decomposition BC into Hilbert-Schmidt operators to define

$$\mathrm{tr}(\varepsilon A) := \mathrm{tr}(BC).$$

We note that the supertrace $\mathrm{str}(e^{-tD^2})$ can be computed by an integral as well, using the following lemma with $A = e^{-tD^2}$ as the smoothing operator (which is now possible due to Proposition 4.6).

LEMMA 4.8. Let A be a smoothing operator given by a kernel K on $L^2(\mathcal{E})$ where \mathcal{E} is a graded Clifford bundle. Then

$$\mathrm{str}(A) = \int_{\mathbf{X}} \mathrm{str}K(x, x) \mathrm{dvol}(x).$$

Proof. Factorize $A = BC$ and B and C are Hilbert-Schmidt with continuous kernels.

We can see

$$K(x, z) = \int_{\mathbf{X}} K_B(x, y) K_C(y, z) \mathrm{dvol}(y)$$

since if we integrate with this formula for $K(x, y)$ we get

$$\int_{\mathbf{X}} K(x, y) s(y) \mathrm{dvol}(x) = BCs(y).$$

Now we can observe that $\|A\|_{\text{HS}}^2 = \iint_{X \times X} |K(x, y)|^2 d\text{vol}(x) d\text{vol}(y)$; the polarization identity tells us that therefore a similar integral formula holds for the Hilbert-Schmidt inner product, which will yield

$$\text{tr}(A) = \iint_{X \times X} K_B(x, y) K_C(y, x) d\text{vol}(x) d\text{vol}(y).$$

But using the previous expression for $K(x, y)$, it is then clear that $\text{tr}(A) = \int_X K(x, x) d\text{vol}(x)$. Generalizing to Clifford bundles, we obtain $\text{tr}(A) = \int_X \text{tr} K(x, x) d\text{vol}(x)$, from which the supertrace version follows by dividing into \mathcal{E}_+ and \mathcal{E}_- . \square

With this discussion out of the way, we can now explain the connection between the heat kernel and the analytic index.

LEMMA 4.9. As $t \rightarrow \infty$, we have $\text{str}(e^{-tD^2}) = \text{a-ind}(D)$.

Proof. By definition, as $t \rightarrow \infty$ the operator e^{-tD^2} tends towards the orthogonal projection onto $\ker D$. The supertrace of this operator is $\dim \ker D_+ - \dim \ker D_-$. \square

The heat kernel then gives us an integral yielding the analytic index in the limit as $t \rightarrow \infty$, giving some hope at relating it to the topological index. We will see in §5.1 that this supertrace is independent of t , and in fact as $t \rightarrow 0$ the integral of the heat kernel will be how we obtain the topological index.

It is therefore important to understand the asymptotic expansion of the heat kernel for small t . We will later refine this asymptotic expansion to be more usable using Getzler's approach.

5 The proof of the index theorem

5.1 The McKean-Singer formula

Recall that we observed in Lemma 4.9 that

$$\text{a-ind}(D) = \text{str}(e^{-tD^2})$$

in the limit as $t \rightarrow \infty$. Taking this argument a bit further, we can actually see this supertrace is independent of t , so there is no need to take a limit as $t \rightarrow \infty$.

THEOREM 5.1 (McKean-Singer, [MJS67]). Let X be a compact oriented Riemannian manifold, and \mathcal{E} a graded Clifford module on X with Dirac operator D . We have for all $t > 0$

$$\text{a-ind}(D) = \text{str}(e^{-tD^2}) = \int_X \text{str}(K_t(x, x)) \text{dvol}(x).$$

We regard $K_t(x, x)$ as an element of $\mathcal{E}_x \otimes \mathcal{E}_x^* \simeq \text{End}(\mathcal{E}_x)$ to take the supertrace.

Proof. We adapt the version of the argument presented in §4.2.3 of [Nic13]. Roe gives a more general argument in Proposition 11.9 of [Roe99] for certain operators of the form $f(D)$ and also outlines the same argument in the special case of e^{-tD^2} .

The idea of the proof comes from the fact that D^2 has a square root D to get isomorphisms between eigenspaces. In particular, by definition we have

$$\text{str}(e^{-tD^2}) = \sum_{\lambda \in \text{Spec } D} (n_\lambda^+ - n_\lambda^-) e^{-t\lambda^2}$$

where n_λ^\pm is the dimension of the \pm component of the eigenspace for λ . Since D is odd, this means D^2 is even and so will commute with the $\mathbf{Z}/2\mathbf{Z}$ grading on $C^\infty(\mathcal{E})$ induced by the one on \mathcal{E} .

Now we use the fact that D gives an isomorphism between the positive and negative components of the λ -eigenspace when $\lambda > 0$, being odd by definition. In particular, $(n_\lambda^+ - n_\lambda^-) = 0$ for $\lambda > 0$ due to these isomorphisms. It follows that only the 0 eigenspace can contribute to the supertrace: this contribution will be $\dim \ker D_+ - \dim \ker D_-$, which is the analytic index of D .

For the final equality $\text{str}(e^{-tD^2}) = \int_X \text{str}(K_t(x, x)) \text{dvol}(x)$, we note that e^{-tD^2} is smoothing and is given by the kernel $K_t(x, x)$ by Proposition 4.6. We can pull the supertrace inside by Lemma 4.8. \square

The definition of D depends on the metric g on X as well as a connection $\nabla_{\mathcal{E}}$ on \mathcal{E} . The previous result tells us these choices do not matter.

COROLLARY 5.2. The analytic index $\text{a-ind}(D)$ is independent of the metric on X and connection on \mathcal{E} .

Proof. The index varies smoothly in these quantities by Theorem 5.1 as $K_t(x, x)$ does, but is an integer. Thus, it is constant. \square

We now see that we can write the analytic index as an integral of $\text{str}(K_t(x, x))$ over X . It remains to relate this integral to $\widehat{A}(X) \wedge \text{ch}(\mathcal{E}/S)$, which will be done using asymptotics of the heat kernel as $t \rightarrow 0^+$.

The obvious first approximation comes from the classic heat kernel on Euclidean space, given by

$$\frac{1}{(4\pi t)^{n/2}} \exp(-|x - y|^2/4t).$$

On our Riemannian manifold X , using the intuition that the heat kernel should approximate the Euclidean one as it is locally isomorphic to Euclidean space, we take

$$h_t(x, y) = \frac{1}{(4\pi t)^{n/2}} \exp(-d(x, y)^2/4t)$$

where $d(x, y)$ comes from the metric g on our manifold X .

The following theorem tells us about the higher order terms of the approximation $h_t(x, y) \approx K_t(x, y)$.

THEOREM 5.3 (Asymptotic expansion). Let X be a compact Riemannian manifold equipped with a Clifford module \mathcal{E} and Dirac operator D . Let $K_t(-, -)$ denote the heat kernel of X . Then there is an asymptotic expansion

$$K_t(x, y) \sim h_t(x, y)(\Theta_0(x, y) + t\Theta_1(x, y) + t^2\Theta_2(x, y) + \dots)$$

where the Θ_j are smooth sections of $\mathcal{E} \boxtimes \mathcal{E}^*$. Moreover, $\Theta_j(x, x)$ can be computed via algebraic expressions involving the metric, connection coefficients, and derivatives of both.

Sketch. We follow the argument for Theorem 7.15 in [Roe99]. Suppose that the heat kernel $K_t(x, y)$ is represented, in local coordinates near y , by

$$(x, t) \mapsto h_t(x)(u_0(x) + tu_1(x) + \dots),$$

where $h_t(x)$ is the Euclidean heat kernel. Let $r^2 = \sum_{ij} g^{ij} x^i x^j$, so that

$$h_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-r^2/4t}.$$

For any section s of $\mathcal{E} \otimes \mathcal{E}_y^*$, we have the formula

$$\frac{1}{h_t} (\partial_t + D^2) h_t s = \partial_t s + D^2 s + \frac{r}{4gt} \frac{\partial g}{\partial r} s + \frac{1}{t} \nabla_{r\partial/\partial r} s \quad (1)$$

To make sense of $\frac{\partial g}{\partial r}$, we actually use $g = \det(g_{ij})$. The derivation of (1) uses a few relations, all of which are not difficult to prove.

- $D^2(fs) = (\Delta f)s - 2\nabla_{\nabla f} s + fD^2s.$

-
- $\nabla h_t = -\frac{h}{2t} r \frac{\partial}{\partial r}$.
 - $\frac{\partial h_t}{\partial t} + \Delta h_t = \frac{r h}{4gt} \frac{\partial g}{\partial r}$, where $g = \det(g_{ij})$ as in (1).

We apply the first item to expand $\frac{1}{h_t}(\partial_t + D^2)(h_t s)$, and the second two to simplify to the right hand side of (1). The second item lets us turn $\nabla_{\nabla h_t}$ into the $\nabla_{r\partial/\partial r}$, and the third item lets us remove the Δf term.

From this, we will want to put $s \sim u_0 + t u_1 + t^2 u_2 + \dots$ where u_i are independent of t , and then try to use (1) to recursively find the u_i for s such that $(\partial_t + D^2)(h_t s) = 0$. We then obtain by looking at the coefficient of t^{i-1} term by term on the right hand side of (1) that

$$i u_i + D^2 u_{i-1} \frac{r}{4g} \frac{\partial g}{\partial r} u_i + \nabla_{r\partial/\partial r} u_i = 0.$$

Now this can give us a recurrence. Introducing an integrating factor of $g^{1/4}$, we obtain

$$\nabla_{\partial/\partial r} r^i g^{1/4} u_i = -r^{i-1} g^{1/4} D^2 u_{i-1}$$

for $i \geq 1$ and that the right hand side is 0 when $i = 0$. This means u_0 is determined by its value at 0 which is fixed as 1 by convention, and that u_i is determined by u_{i-1} up to a constant of order r^{-i} . But smoothness at the origin makes this vanish, so the u_i are uniquely determined. We fixed y when we began, and it is straightforward to assemble this into the desired asymptotic expansion for the Θ_i . By working through what happens in the recurrence, we obtain the last condition about the Θ_i being computed by algebraic expressions in the metric, connection coefficients, and derivatives of both. \square

Using Theorem 5.3, we get

$$\text{str}(e^{-tD^2}) \sim \frac{1}{(4\pi t)^{n/2}} \left(\int_X \text{str}(\Theta_0) \text{dvol} + t \int_X \text{str}(\Theta_1) \text{dvol} + \dots \right).$$

By Theorem 5.1, we obtain a less useful index theorem

$$\text{a-ind}(D) = \frac{1}{(4\pi)^{n/2}} \int_X \text{str}(\Theta_{n/2}) \text{dvol}.$$

This is still not obvious, as the leading part of the asymptotic expansion is $O(t^{-n/2})$ and so it appears the supertrace should diverge as $t \rightarrow 0$. We really need to appeal to the independence of t to ensure this makes sense.

Note that this result also explains why we assume the dimension $n = \dim X$ is even throughout. If it is odd, then the analytic index is *always* zero, so there is no content to the theorem. It is possible to work out small cases of the index theorem by directly working out an expression for $\Theta_{n/2}$, but this gets increasingly difficult as n grows larger.

5.2 Getzler's asymptotics

We previously obtained

$$\text{a-ind}(D) = \frac{1}{(4\pi)^{n/2}} \text{str}(\Theta_{n/2}) \text{dvol}.$$

Having observed that computing $\Theta_{n/2}$ directly to obtain the supertrace is impractical, we will need an alternate method to obtain the index theorem that bypasses writing down a formula for $\Theta_{n/2}$. In the previously established asymptotic formula

$$K_t(x, y) \sim h_t(x, y)(\Theta_0(x, y) + t\Theta_1(x, y) + \dots)$$

we will note $\text{str}(\Theta_{n/2})$ can be understood in terms of a symbol map $\sigma_\bullet : C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*) \rightarrow C^\infty(\mathbf{C}[[\text{TX}]] \otimes \Lambda^\bullet(\text{TX}) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}))$ (which can be applied to the Θ_i in particular). It turns out we can understand the heat symbol

$$h_t(\sigma_0\Theta_0 + t\sigma_2\Theta_1 + \dots)$$

as this satisfies a differential equation we can explicitly solve. We will then relate this back to the supertrace computation we attempted earlier. The material in this section is largely from Chapter 12 of [Roe99], which gives an exposition to the paper [Get86].

We have already seen some of the first ideas leading to Getzler's symbol map. When understanding the twisted Chern class, we exploited the isomorphism of super vector bundles

$$\text{End}(\mathcal{E}) \simeq \text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}).$$

This induces a filtration on $\text{End}(\mathcal{E})$. Namely, we have a filtration on $\text{Cl}(X)$ coming from the filtration on $\text{Cl}(V, Q)$ induced by the grading of $T(V)$ (it is not a grading because we imposed relations). Putting $\text{End}_{\text{Cl}(X)}(\mathcal{E})$ in degree zero, this induces a filtration on $\text{End}(\mathcal{E})$ called the Getzler filtration.

The Clifford algebra $\text{Cl}(V, Q)$ also came equipped with a symbol map

$$\sigma : \text{Cl}(V, Q) \rightarrow \Lambda^\bullet V$$

sending v to its corresponding element in the associated graded algebra, which is $\Lambda^\bullet V$. This globalizes to a symbol map

$$\sigma : \text{End}(\mathcal{E}) \simeq \Lambda^\bullet \text{TX} \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}),$$

which naturally has graded components σ_m coming from the filtrations on both sides.

Our goal is to be able to formulate a version of the symbol σ that can be applied to the $\Theta_i(x, y)$ in the asymptotic expansion of Theorem 5.3. These are smooth sections of $\mathcal{E} \boxtimes \mathcal{E}^*$, but we can use the symbol above on $\text{End}(\mathcal{E})$. We have a Taylor expansion map

$$\Sigma : C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*) \rightarrow C^\infty(\mathbf{C}[[\text{TX}]] \otimes \text{End}(\mathcal{E})).$$

Here, $\mathbf{C}[[\mathrm{TX}]]$ is the bundle whose fibers over each $x \in X$ are the rings of formal power series $\mathbf{C}[[\mathrm{T}_x X]]$. We will describe this map fiberwise. Fixing $y \in X$, given a section $s \in C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*)$ there is a Taylor expansion in x of the function

$$s_y(x) : x \mapsto s(x, y),$$

which we can view as a section of $\mathcal{E} \otimes \mathcal{E}_y^*$. The function $s_y(x)$ has the form $s_y(x) \sim \sum_\alpha s_\alpha x^\alpha$ by using the usual Taylor expansion. The s_α here are sections of $\mathcal{E} \otimes \mathcal{E}_y^*$ parallel along geodesics coming from y . In particular, the s_α are determined by the value at $x = y$ which live in $\mathrm{End}(\mathcal{E}_y)$. Hence, the Taylor expansion at y naturally is viewed as a section of $C^\infty(\mathbf{C}[[\mathrm{T}_y X]] \otimes \mathrm{End}(\mathcal{E}_y))$, and these expansions smoothly assemble into Σ .

Now we can apply the previous symbol map on $\mathrm{End}(\mathcal{E})$ to produce a final symbol map

$$\sigma_\bullet : C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*) \rightarrow C^\infty(\mathbf{C}[[\mathrm{TX}]] \otimes \Lambda^\bullet(\mathrm{TX}) \otimes \mathrm{End}_{\mathrm{Cl}(X)}(\mathcal{E})).$$

There is once again a clear filtration on the target of the symbol map, by putting $\mathrm{End}_{\mathrm{Cl}(X)}(\mathcal{E})$ in degree zero, $\Lambda^\bullet(\mathrm{TX})$ with its natural filtration, and $\mathbf{C}[[\mathrm{TX}]]$ the filtration where a monomial x^α has degree $-|\alpha|$ (we use the metric g on each tangent space $\mathrm{T}_x X$ to give this meaning). We again call this the Getzler filtration.

DEFINITION 5.4. We say $s \in C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*)$ has Getzler order $\leq m$ if $\sigma_\bullet(s)$ lies in the degree $\leq m$ part of the Getzler filtration on $C^\infty(\mathbf{C}[[\mathrm{TX}]] \otimes \Lambda^\bullet(\mathrm{TX}) \otimes \mathrm{End}_{\mathrm{Cl}(X)}(\mathcal{E}))$.

The symbol map σ_\bullet gives graded components which we denote σ_m , and we use σ_m^0 for the constant term in the Taylor series $\sigma_m(s)$.

It is worth remarking that $C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*)$ can be regarded as consisting of smoothing operators (by writing down the corresponding kernels). However, the symbol does not respect the composition of smoothing operators. It does respect the action of the algebra of differential operators $\mathcal{D}(\mathcal{E})$ on $C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*)$.

The key idea relating the symbol back to the supertrace is the following.

LEMMA 5.5. We have

$$\mathrm{str}(\Theta_{n/2}(x, x)) = (-2i)^{n/2} \mathrm{str}^{\mathcal{E}/S}(\sigma_n^0(\Theta_{n/2})).$$

Proof. Recall that $\Theta_{n/2}(x, x)$ at x yields an element of $\mathcal{E}_x \otimes \mathcal{E}_x^* \simeq \mathrm{End}(\mathcal{E}_x)$, hence it defines an endomorphism we can take the supertrace of.

This is a consequence of a more general identity, which tells us that for an endomorphism $c \otimes F$ of a Clifford module \mathcal{E} (using the decomposition $\mathrm{End}(\mathcal{E}) \simeq \mathrm{Cl}(V) \otimes \mathrm{End}_{\mathrm{Cl}(V)}(\mathcal{E})$)

we have

$$\text{str}(c \otimes F) = \text{str}(c)\text{str}^{\mathcal{E}/S}(F).$$

Here, we take the supertrace of c regarded as an endomorphism of the spinor module S . Moreover, via representation theory one can show $\text{str}(c) = (-2i)^{\dim V/2} c_{1,2,\dots,\dim V}$ when we pick a basis e_i of V and write $c = \sum_{\mathcal{E} \subset \{1,2,\dots,\dim V\}} c_{\mathcal{E}} \prod_{e \in \mathcal{E}} e$.

Now this tells us that $\text{str}(\Theta_{n/2}(x, x))$ can be computed in terms of the top degree component of $\Theta_{n/2}$ in the Geztler filtration coming from $\text{End}(\mathcal{E}) \simeq \text{Cl}(X) \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E})$. This is precisely what the symbol σ_n^0 gives us by construction, which proves the claim. \square

Motivated by this, we will aim to explicitly obtain the heat *symbol* $W = h_t(\sigma_0\Theta_0 + t\sigma_2\Theta_1 + \dots + t^{n/2}\sigma_n\Theta_{n/2})$ rather than the full asymptotic expansion of $K_t(x, x)$, as we are only interested in the supertrace. This will allow us to give an explicit formula for $\sigma_n^0\Theta_{n/2}$.

In doing this we will need to find a differential equation W satisfies, but for this we need a symbol for differential operators and not just smoothing operators.

DEFINITION 5.6. Let $\text{Diff}(\mathcal{E})$ be the algebra of differential operators on \mathcal{E} . This is the algebra generated by Clifford multiplications, covariant derivatives, and sections of $\text{End}_{\text{Cl}(X)}(\mathcal{E})$ with the natural relations.

We give $\text{Diff}(\mathcal{E})$ a filtration as follows:

- Elements of $\text{End}_{\text{Cl}}(\mathcal{E})$ have degree zero.
- Elements of the form $\text{cl}(V)$ for V a vector field on X have degree one.
- Covariant derivatives ∇_V for V a vector field on X also have degree one.

If V is a vector space, let $\text{PDiff}(V)$ be the graded algebra of polynomial differential operators on V generated by $x^\alpha \frac{\partial^\beta}{\partial x^\beta}$ in degree $|\beta| - |\alpha|$. This yields a bundle $\text{PDiff}(TX)$.

There is a filtered symbol map

$$\sigma_\bullet : \text{Diff}(\mathcal{E}) \rightarrow C^\infty(\text{PDiff}(TX) \otimes \Lambda^\bullet TX \otimes \text{End}_{\text{Cl}(X)}(\mathcal{E}))$$

where we give the target the filtration given by assigning $\text{End}_{\text{Cl}(X)}(\mathcal{E})$ degree zero, and the other components their natural graded structures.

It is uniquely determined by what it does on generators of $\text{Diff}(\mathcal{E})$:

- For $F \in \text{End}_{\text{Cl}(X)}(\mathcal{E})$, we have $\sigma_0(F) = F$.
- Clifford multiplication $\text{cl}(V)$ by V is sent by σ_1 to exterior multiplication by V .
- Covariant derivatives are sent by σ_1 to $\partial_V + \frac{1}{4}(RV, -)$.

In the last item we regard R as a 2-form with values in $\text{End}(\text{TX})$ so that $v \mapsto (R_x V_x, v)$ is a linear map $\text{T}_x X \rightarrow \Lambda^2 \text{T}_x^* X$. Via the metric, identify the tangent and cotangent spaces so we regard this as a polynomial function on $\text{T}_x X$ with values in $\Lambda^2 \text{T}_x X$. It follows $\frac{1}{4}(RV, -)$ lies in $\text{PDiff}(\text{TX}) \otimes \Lambda^\bullet \text{TX}$.

We omit the proof the following lemma, since what is actually notable is the corollary.

LEMMA 5.7. Let $T \in \text{Diff}(\mathcal{E})$ be one of the generators: an element of $\text{End}_{\text{Cl}(X)}(\mathcal{E})$, Clifford multiplication, or a covariant derivative. Let i be the degree in $\text{Diff}(\mathcal{E})$ it lies in (either 0 or 1). If f is a smoothing operator,

$$\sigma_{i+j}(Tf) = \sigma_i(T)\sigma_j(f)$$

where Tf is again a smoothing operator.

COROLLARY 5.8. The symbol map σ_\bullet is well-defined on $\text{Diff}(\mathcal{E})$, and satisfies

$$\sigma_{i+j}(Tf) = \sigma_i(T)\sigma_j(f)$$

for $T \in \text{Diff}(\mathcal{E})$ and $f \in C^\infty(\mathcal{E} \boxtimes \mathcal{E}^*)$ corresponding to a smoothing operator.

Proof. As given, the symbol $\sigma_i(T)$ is potentially different for distinct ways of writing T in terms of generators of $\text{Diff}(\mathcal{E})$. Let T and T' be two such representations of the same element of $\text{Diff}(\mathcal{E})$. Repeatedly applying Lemma 5.7, we obtain

$$\sigma_{i+j}(Tf) = \sigma_i(T)\sigma_j(f) = \sigma_i(T')\sigma_j(f).$$

They both give the same value for $\sigma_{i+j}(Tf)$ as they are representations of the same element of $\text{Diff}(\mathcal{E})$. As $\sigma_j(f)$ can be an arbitrary formal power series, this implies $\sigma_i(T) = \sigma_i(T')$. Thus, the symbol is well-defined on $\text{Diff}(\mathcal{E})$. This implies that

$$\sigma_{i+j}(Tf) = \sigma_i(T)\sigma_j(f)$$

independently of the representation of T in generators. □

We are now ready to study the heat symbol $W = h_t(\sigma_0\Theta_0 + t\sigma_2\Theta_1 + \dots + t^{n/2}\sigma_n\Theta_{n/2})$, as we can formulate correctly the differential equation it satisfies.

THEOREM 5.9. Assume $\dim_{\mathbf{R}} X = n$, which must be even due to the constraints on

X. As $t \rightarrow 0$, we have an asymptotic expansion

$$K_t(x, y) \sim h_t(x, y) \sum_{i \geq 0} t^i \Theta_i(x, y).$$

The $\Theta_i(x, y)$ are of Getzler order $\leq 2i$. Then the heat symbol

$$W = h_t(\sigma_0 \Theta_0 + t \sigma_2 \Theta_1 + \dots + t^{n/2} \sigma_n \Theta_{n/2})$$

satisfies

$$\partial_t W + \sigma_2(D^2)W = 0$$

and moreover is the *unique* solution of the form $h_t(v_0 + tv_1 + \dots + t^{n/2}v_{n/2})$ where v_i has order $\leq 2i$ and $v_0 = 1$.

Once we have this, getting the asymptotic result we need is a matter of directly solving this differential equation. Due to the uniqueness claim, all that is necessary is to verify a solution works.

Proof. To see this, we just need to derive the asymptotic expansion as in Theorem 5.3 and pay close attention to the construction of Θ_i . We started by writing $K_t(x, y)$ in local coordinates near y as

$$(x, t) \mapsto h_t(x)(u_0(x) + tu_1(x) + \dots).$$

Then, we ultimately derived recurrences

$$\nabla_{\partial/\partial r} r^i g^{1/4} u_i = -r^{i-1} g^{1/4} D^2 u_{i-1}$$

for $i \geq 1$ and that the right hand side is 0 when $i = 0$. Here, $r = \sum_{ij} g^{ij} x^i x^j$. We argued using smoothness that the value $u_0(0) = 1$ uniquely determines the expansion from here.

Now, take Taylor expansions on both sides of this recurrence and use that $\nabla_{\partial/\partial r}$ annihilates synchronous sections (parallel along geodesics through y). Then we can argue by induction that the u_i , as sections of $\mathcal{E} \otimes \mathcal{E}_y^*$, have Getzler order $\leq 2i$ (so the same applies to Θ_i with the actual map σ , not just the fibers at each y). Moreover, we see inductively that

$$\partial_r(r^i \sigma_{2i}(u_i)) = r^{i-1} \sigma_2(D^2) \sigma_{2i-2}(u_{i-1}).$$

Here, we are using both symbol maps: differential operators in $\text{Diff}(\mathcal{E})$ apply on both sides, it is necessary to use Corollary 5.8 to evaluate the symbol.

These are the exact same relations for the asymptotic expansion coefficients of $\partial_t W + \sigma_2(D^2)W = 0$, which determine solutions uniquely given $u_0(0) = 1$. This proves the claim. \square

For solving this explicitly, we need to understand $\sigma_2(D^2)$. This can be done by explicit computation in local coordinates using the formula

$$D^2 = \nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}} + \frac{1}{4}s + \text{cl}(F^{\mathcal{E}/S}).$$

where s is the scalar curvature.

LEMMA 5.10. We have

$$\sigma_2(D^2) = - \sum_i \left(\frac{\partial}{\partial x^i} + \frac{1}{4} \sum_j R_{ij} x^j \right)^2 + F^{\mathcal{E}/S}.$$

The matrix R is skew-symmetric, and F is a 2-form with values in $\text{End}_{\text{Cl}(X)}(\mathcal{E})$.

Proof. Pick a local orthonormal frame e_i with coordinate functions x^i . As remarked, we can apply the identity

$$D^2 = \nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}} + \frac{1}{4}s + \text{cl}(F^{\mathcal{E}/S}).$$

We can see $\sigma_2(s) = 0$ and

$$\sigma_2(\text{cl}(F^{\mathcal{E}/S})) = F^{\mathcal{E}/S},$$

since $\text{cl}(F^{\mathcal{E}/S}) = \sum_{i < j} \text{cl}(e_i) \text{cl}(e_j) F^{\mathcal{E}/S}(e_i, e_j)$. This means applying σ_2 gives just $F^{\mathcal{E}/S} = \sum_{i < j} F^{\mathcal{E}/S}(e_i, e_j) e_i \wedge e_j$.

What remains is to compute $\sigma_2(\nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}})$. We can do this at a point $x \in X$, whereupon picking a local orthonormal frame we have the formula at

$$\nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}} = - \sum_i \nabla_i^2$$

at x . The symbol is then $-\sum_i \sigma_1(\nabla_i)^2$, which we know from what the symbol does to generators of $\text{Diff}(\mathcal{E})$. We obtain

$$\sigma_2(\nabla_{\mathcal{E}}^* \nabla_{\mathcal{E}}) = - \sum_i \left(\partial_i + \frac{1}{4} \sum_j R_{ij} x^j \right)^2.$$

Putting everything together, we get the claim. \square

The idea behind solving $\partial_t W - \sigma_2(D^2)W = 0$ is to try to reduce to the case of a harmonic oscillator.

PROPOSITION 5.11 (Mehler's formula). The differential equation

$$\frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} + a^2 x^2 w = 0$$

for $a \in \mathbf{R}$ has an explicit solution given by

$$w(x, t) = \left(\frac{a}{2\pi \sinh(2at)} \right)^{1/2} \exp\left(\frac{-ax^2 \coth(2at)}{2} \right).$$

This is the kernel for the harmonic oscillator. Now we are ready to explicitly solve $\partial_t W - \sigma_2(D^2)W = 0$ on X . This is the analytic heart of the proof where we first see characteristic classes appear in the form of $\det^{1/2}\left(\frac{tR/2}{\sinh tR/2}\right)$ for the \widehat{A} -genus and $\exp(-tF^{\mathcal{E}/S})$ for the twisted Chern class.

PROPOSITION 5.12. The differential equation

$$\partial_t W - \sigma_2(D^2)W = 0$$

has a solution which is explicitly given by

$$W(x, t) = (4\pi t)^{-n/2} \det^{1/2}\left(\frac{tR/2}{\sinh tR/2}\right) \exp\left(-\frac{1}{4t} \left\langle \frac{tR}{2} \coth \frac{tR}{2} x, x \right\rangle\right) \exp(-tF^{\mathcal{E}/S}).$$

Proof. We will perform a series of reductions to show it suffices to know Mehler's formula to give a solution. First, it suffices to reduce to the case of scalars and solve

$$\frac{\partial w}{\partial t} - \sum_i \left(\frac{\partial}{\partial x^i} + \frac{1}{4} \sum_j R_{ij} x^j \right)^2 w + Fw = 0$$

where R_{ij} is a skew-symmetric *matrix of scalars* and F is a *scalar*. If we show

$$w(x, t) = (4\pi t)^{-n/2} \det^{1/2}\left(\frac{tR/2}{\sinh tR/2}\right) \exp\left(-\frac{1}{4t} \left\langle \frac{tR}{2} \coth \frac{tR}{2} x, x \right\rangle\right) \exp(-tF)$$

is a solution, we are actually done.

The proposed global solution $W(x, t)$ is required to solve $\partial_t W - \sigma_2(D^2)W = 0$. We can expand $\sigma_2(D^2)$ to get it in the previous form but instead with F and R in terms of 2-forms rather than scalars. By taking the solution $w(x, t)$ in the scalar case and viewing everything as formal power series in R and F (which converge, since entries

are 2-forms which are nilpotent), we see it formally follows from the scalar case that $W(x, t)$ is a solution.

Once we have made this reduction, it is clear that it suffices to solve the differential equation explicitly when $F = 0$ by change of variables. Further, since R is a skew-symmetric matrix, by a change of basis we can break it into blocks $\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$. This reduces us to the case where R is a 2 by 2 matrix, which can now be done via Mehler's formula.

Indeed, we note that the differential equation for w is given by

$$\frac{\partial w}{\partial t} + Hw = 0$$

where

$$H = \underbrace{\left[-(\partial_x^2 + \partial_y^2) - 16\theta^2(x^2 + y^2) \right]}_{H_0} + \underbrace{\left[\frac{1}{2}\theta(x\partial_y - y\partial_x) \right]}_{H_1}.$$

The purpose of separating the terms H_0 and H_1 in square brackets is that $H_1 w = 0$ by rotational invariance of the proposed solution for w . Thus, we just need to check $\partial_t w + H_0 w = 0$.

Then we can separate variables into x and y ; individually, in each of these (for example, x) we have the differential equation

$$\frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} - \frac{1}{16}\theta^2 x^2 w = 0$$

which has explicit solution $(4\pi t)^{-1/2} \left(\frac{it\theta/2}{\sinh(it\theta/2)} \right)^{1/2} \exp(-\frac{1}{8}i\theta x^2 \coth(it\theta/2))$ by Mehler's formula. Note that previously we stated it for $a \in \mathbf{R}$, but by analytic continuation of the solution it is fine to take sufficiently small $a \in \mathbf{C}$.

Taking the product of these yields

$$w(x, t) = (4\pi t)^{-1} \left(\frac{it\theta/2}{\sinh(it\theta/2)} \right) \exp \left(-\frac{1}{8}i\theta |x|^2 \coth(it\theta/2) \right)$$

where we now take x as a vector in two coordinates. This is a solution of the desired form when $R = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$ and $F = 0$. The general case then follows. \square

COROLLARY 5.13. The solution $W(x, t)$ as in Proposition 5.12 is equal to

$$h_t(\sigma_0\Theta_0 + t\sigma_2\Theta_1 + \dots + t^{n/2}\sigma_n\Theta_{n/2}).$$

Proof. This follows from the uniqueness in Theorem 5.9 once we verify that the solution $W(x, t)$ of Proposition 5.12 has the form $h_t(v_0 + tv_1 + \dots + t^{n/2}v_{n/2})$ where v_i has order $2i$ and $v_0 = 1$, as this tells us $W(x, t) = h_t(\sigma_0\Theta_0 + t\sigma_2\Theta_1 + \dots + t^{n/2}\sigma_n\Theta_{n/2})$. \square

Having obtained an explicit version of a heat kernel asymptotic, we are now interested in the supertrace of $\Theta_{n/2}(x, x)$. Fortunately, this is easily computable.

Thus, it just a matter of computing $\sum_{i=0}^{n/2} \sigma_{2i}^0(\Theta_i)$ and taking the n -form part.

COROLLARY 5.14. We have

$$\sum_{i=0}^{n/2} \sigma_{2i}^0(\Theta_i) = \det^{1/2} \left(\frac{R/2}{\sinh R/2} \right) \exp(-F^{\mathcal{E}/S}).$$

Proof. Evaluating $W(x, t)$ at $x = 0$ and $t = 1$ to get the σ_{2i}^0 (recall these are constant terms in the Taylor expansion), we get the result. \square

5.3 Putting it all together

With Corollary 5.14 in place we just need to combine all of the main results we've seen so far.

THEOREM 5.15 (Atiyah-Singer). Let X be a compact oriented Riemannian manifold, and let \mathcal{E} be a graded Clifford bundle with Dirac operator D . Then

$$\text{a-ind}(D) = \text{t-ind}(D).$$

Proof. As $t \rightarrow \infty$, we have already seen

$$\text{a-ind}(D) = \lim_{t \rightarrow \infty} \text{str}(e^{-tD^2})$$

in Lemma 4.9. By the McKean-Singer formula, the analytic index can be expressed as

$$\text{a-ind}(D) = \text{str}(e^{-tD^2}) = \int_X K_t(x, x) \text{dvol}(x)$$

independently of t . Thus it can also be computed by finding the limit as $t \rightarrow 0^+$. In §5.1, we noted that this implies

$$\text{a-ind}(D) = \frac{1}{(4\pi)^{n/2}} \int_X \text{str}(\Theta_{n/2}) \text{dvol}.$$

Now note that

$$\text{str}(\Theta_{n/2}) = (-2i)^{n/2} \text{str}^{\mathcal{E}/S} \sigma_n^0(\Theta_{n/2}).$$

The right hand side $(-2i)^{n/2} \text{str}^{\mathcal{E}/S} \sigma_n^0(\Theta_{n/2})$ by Corollary 5.14 yields

$$\text{str}(\Theta_{n/2}) = \left[(-2i)^{n/2} \det^{1/2} \left(\frac{R/2}{\sinh(R/2)} \right) \wedge \text{str}(\exp(-F^{\mathcal{E}/S})) \right]_{[n]}.$$

Here, the subscript denotes the n -form part. This is precisely $(-2i)^{n/2} (2\pi i)^{n/2} = (4\pi)^{n/2}$ times the n -form part of $\widehat{A}(X) \wedge \text{ch}(\mathcal{E}/S)$ since in the definition of the Chern class we divide $F^{\mathcal{E}/S}$ by $2\pi i$. We obtain

$$\text{a-ind}(D) = \int_X \widehat{A}(X) \wedge \text{ch}(\mathcal{E}/S)_{[n]} = \text{t-ind}(D).$$

□

References

- [AS63] Michael Atiyah and Isadore M Singer, *The index of elliptic operators on compact manifolds*, Bulletin of the American Mathematical Society **69** (1963), no. 3, 422–433.
- [AS68] Michael Atiyah and Isadore Singer, *The index of elliptic operators: I*, Annals of mathematics (1968), 484–530.
- [Fre87] Daniel Freed, *Geometry of dirac operators*, 1987.
- [Fre15] ———, *Math 392c: K-theory*, 2015.
- [Gel60] Israel M Gel'fand, *On elliptic equations*, Russian Mathematical Surveys **15** (1960), no. 3, 113.
- [Get86] Ezra Getzler, *A short proof of the local atiyah-singer index theorem*, Topology **25** (1986), no. 1, 111–117.
- [MJS67] Henry P McKean Jr and Isadore M Singer, *Curvature and the eigenvalues of the laplacian*, Journal of Differential Geometry **1** (1967), no. 1-2, 43–69.
- [MS74] John Willard Milnor and James D Stasheff, *Characteristic classes*, no. 76, Princeton university press, 1974.

[Nic13] Liviu I Nicolaescu, *Notes on the atiyah-singer index theorem*, University of Notre Dame (2013).

[Roe99] John Roe, *Elliptic operators, topology, and asymptotic methods*, CRC Press, 1999.