Outline of the Theory and Perspective

de Rham complex and Dolbeault complex

The de Rham complex and the Dolbeault complex are two complexes that appear in the geometry of differentiable manifolds and the geometry of complex manifolds, respectively.

The former, resp. the latter, has its origin in Cartan's discovery of the exterior derivative, resp. Cauchy's discovery of the Cauchy-Riemann equation.[1] In the former case, its connection with geometry was revealed when the de Rham theorem clarified its relation to Betti numbers. In the latter case, it goes back to an older time, when Riemann found the ancestor of the Riemann-Roch theorem, i.e., the inequality involving the genus of the Riemann surface. However, the former was understood as the device deriving invariants of manifolds and the latter was understood as some properties of various holomorphic line bundles (or divisors). Hence they play different roles. This is still the case now.

These two trends met in a clear form under the viewpoint of unification in the theory of Hodge and Kodaira, which followed Weyl's work. Cohomology is associated with a complex and has geometric meaning. (For clear understanding, it was also necessary to establish the notion of sheaves and their Čech cohomology, which is equivalent to the two kinds of complexes mentioned above.) Hodge-Kodaira described their cohomologies in terms of solutions of the equation of Laplace type.

Considering the alternating sum of the dimension of the cohomology of the complex instead of the dimension of the cohomology in each degree, we lose a lot of detailed information, but we acquire simple properties (with respect to exact sequences). This fact is clear

[1] Manifesting the origin may depend on personal opinion. Other descriptions in this part are also like virtual history.

in an algebraic setting. However, the alternating sums for the two complexes have developed in different ways.

For the de Rham complex, its alternating sum is called the Euler characteristic. As its name indicates, the Euler characteristic itself was known a long time ago. However, it was after Poincaré started the subject, which is now called topology, and then considered homology, that the geometric meaning of the Euler characteristic began to be clarified. The Poincaré-Hopf theorem is famous in its relation to the number of zeros of a vector field on a manifold. This theorem states that the Euler characteristic can be expressed as the integration of a characteristic class called the Euler class (in general, the integration of characteristic classes are called characteristic numbers).

For the Dolbeault complex (or the corresponding Čech complex), it was already a historical achievement to construct the higher cohomology groups as the objects for taking the alternating sum. The Riemann-Roch theorem gives the formula for the alternating sum in the case of complex dimension 1. Based on results for low-dimensional cases, such as complex surfaces, due to Kodaira and others, it was Hirzebruch who finally gave the formula for the general case (nonsingular projective algebraic varieties). The end result is given as the combination of various characteristic numbers.

Vaguely speaking, the alternating sum is invariant under quite a large class of deformations of the complex. In order to investigate properties of the alternating sum by use of the freedom of such deformations, it is necessary to make the precise meaning of "deformation" that would be as broad as possible.

From a modern viewpoint, the ideas of "derived categories" (in algebraic geometry) and "stable homotopy" (in topology) provide a framework in which equivalence classes under deformations (with certain freedom of deformations) are described. It was Grothendieck who introduced cohomology (of sheaves) in algebraic geometry suitable for the former notion. In particular, in our current problem on the alternating sum of the dimensions of cohomology groups of the complex, the notion of the K-group, which naturally emerges from these frameworks, is useful. When Grothendieck extended the Riemann-Roch theorem to families of algebraic varieties of arbitrary characteristic, rather than to a single algebraic variety, he used the notion of the K-group. Following such developments, the topological version of the K-group was later introduced by Atiyah as the set of a certain kind of stable homotopy classes.

The freedom that Hirzebruch could use for extending the Riemann-Roch theorem was rather limited as compared to the freedom Grothendieck later enjoyed when he invented a very general approach to algebraic geometry and used it for the above-mentioned extension of the Riemann-Roch theorem. Instead, Hirzebruch used, in an essential step, a deep theorem, which belongs to differential topology. It is the signature theorem, which is also due to Hirzebruch himself.

Hirzebruch's signature theorem

Hirzebruch's signature theorem is purely a theorem in differential topology. Similar to the fact that the Euler characteristic partially reflects the information on the dimension of the de Rham cohomology of the manifold, the signature partially reflects the information on the cup product on the de Rham cohomology. This theorem describes the signature of the manifold explicitly using characteristic numbers.

In order to prove the signature theorem, which is formulated for each (oriented) closed manifold, Hirzebruch did not discuss each individual manifold, but used Thom's cobordism theory, which deals with all closed manifolds at once.

Consider the set of all diffeomorphism classes of closed manifolds, which are not necessarily connected, for all dimensions. To introduce the algebraic structure on this set, define the sum by the disjoint union as spaces and the product by the product as spaces. Take the quotient by the relation defined by the requirement that a manifold, which is the boundary of some compact manifold, is equivalent to 0. This is the cobordism ring. In fact, this ring is a commutative ring (in the sense of graded rings) with the identity being the empty set. We can also define the cobordism ring of closed manifolds, or closed manifolds equipped with any other suitable structure. On the other hand, the totality of characteristic classes has a structure of a commutative ring. The characteristic numbers are obtained by the integration of the characteristic classes of the tangent bundle over the manifold. This integration defining the characteristic numbers yields a bilinear form between the cobordism ring and the ring of characteristic classes. Thom's theory claims that this bilinear form is non-degenerate in the case of the cobordism ring of oriented closed manifolds after tensoring $\mathbb{Q}$.

Hirzebruch noticed that the signature gives a ring homomorphism from the cobordism ring to $\mathbb{Z}$. By Thom's theory, this immediately implies that the signature is expressed in terms of characteristic numbers. Once the explicit expression is predicted, it is easily verified by confirming the prediction for generators of the cobordism ring.

How does the proof of Thom's cobordism theory go? Logically, it consists of the following four steps.

(1) Determine the ring of all characteristic classes (the point is that any vector bundle is realized as a subbundle of a trivial vector bundle. The stable homotopy class of the embedding is uniquely determined. It reduces the problem to the computation of the cohomology of the Grassmann variety (the classifying space)).
(2) Estimate the size of the cobordism ring from below (list all candidates for generators. In practice, candidates are given by projective spaces of all dimensions).
(3) Estimate the size of the cobordism ring from above (the point is that any manifold is realized as a closed subspace of the Euclidean space. The Pontrjagin-Thom construction reduces the problem to the computation of the homotopy of the space called the Thom complex).
(4) Observe that the bilinear form between (1) and (2) is non-degenerate and that the size in (3) is at most the size of (1).

Then we find that (2) and (3) are equal and that the bilinear form between (1) and (3) must be non-degenerate. This is the desired statement.

The first two proofs of the index theorem

Using the de Rham cohomology and the Hodge theory, the signature is expressed as the dimension of Ker minus that of Coker, namely the "index" of a certain elliptic differential operator.

Similarly, the Poincaré-Hopf theorem related to the de Rham complex and the Riemann-Roch theorem related to the Dolbeault complex are interpreted as the formulas of the index of elliptic linear differential opeators, which are obtained by folding the $\mathbb{Z}$-graded complexes to $\mathbb{Z}_2$ ones.

Atiyah-Singer generally proved the formula for the index of elliptic linear differential operators in terms of the combination of characteristic numbers, which includes the above three cases. This is the "index theorem".

The first proof of the index theorem was the extension of Hirzebruch's argument using Thom's cobordism theory.

Then Atiyah-Singer announced the second proof of the index theorem. In this proof, a more direct framework was given using the (topological) K group beforehand of the formula in terms of characteristic numbers.

The second proof contains three points.

Compared with the approaches taken by Grothendieck and Hirzebruch, where they remained in the framework of algebraic geometry and topology as much as possible, Atiyah-Singer's second proof of the index theorem uses the freedom of "pseudo-differential operators", which are closely related to the elliptic differential operators, which are the original objects of the theorem. Then the first point of the proof is to check fundamental properties of elliptic pseudo-differential operators.

Thom's cobordism theory, which Hirzebruch used, requires a meta-global argument, which captures all manifolds and all characteristic classes. From the viewpoint of studying individual manifolds, it may be considered as an indirect tool. To capture all manifolds at once, the argument is based on the fact that any manifold can be embedded in the Euclidean space of sufficiently large dimension. Atiyah-Singer's second proof of the index theorem directly used this fact as a geometric idea. It is only necessary to embed each individual manifold, and one needs not to capture all manifolds at once. This is the second point.

The third point is "Bott's periodicity theorem". It is, in the final stage, used for verifying the index theorem in the case of the Euclidean space, to which the index theorem in the general case is reduced. Atiyah pointed out that Bott's periodicity theorem is easily explained using the notion of the topological K-group and the notion of the "index for families of elliptic linear differential operators". Each of these notions were modelled on the extension of the Riemann-Roch theorem due to Grothendieck. However, the only known proof, which covers the case with (non-commutative) compact group action, is Atiyah's argument, which uses the index.

Other proofs of the Atiyah-Singer index theorem

The second proof and the formulation using the K-group therein are suitable for extending the index theorem to wider situations. For example, the formulation and the proof for the case of the index of families immediately follow.

On the other hand, there are some other approaches for the original index theorem dealing with the numerical index (or characteristic numbers).

- The method of the heat kernel.[2] The local properties, which the index enjoys, are realized by the study of the asymptotic expansion of the trace of the heat kernel. It is also a starting point toward the extension of the index theorem for manifolds with boundary due to Atiyah-Patodi-Singer.
- The proof using stochastic analysis. It is related to the method of the heat kernel. However, in this approach, we, in a sense, consider the stage before (pseudo-) differential operators, and directly study the situation when the heat kernel emerges.
- Explanation using symplectic geometry on the loop space. This consideration uses the Duistermaat-Heckman formula, which is similar to the residue formula for integration. It is an intuitive argument by Witten, but an appropriate mathematical framework has not been given yet.

Both the first and second proofs due to Atiyah-Singer fully used the topological freedom. In other regards, roughness of the mesh in such arguments is just enough for the proof and the minimal understanding of the index theorem. On the other hand, other proofs rely on more detailed geometry. We may say that these arguments do not reduce the problem to simpler cases by using the topological freedom, but may help to understand the index theorem as geometry that arises in discussion of differential operators.

It is fundamental in the index theorem, or the notion of the index, that certain finite-dimensional information is defined by cancellation between infinite-dimensional parts. We may say that each of these other proofs helps in the direct understanding of such a mechanism of cancellation.

Nowadays, when discussing the index theorem, we cannot ignore its relation to physics. The above argument due to Witten is one such example. In this book, we do not describe their relation, but only list some names of related physicists and mathematicians.

---

[2] Cf. Tomoyoshi Yoshida, "Index Theorem for Dirac Operators" (Kyoritsu Shuppan Publisher, 1998, in Japanese).

Alvarez-Gaumé, Witten, Getzler, Quillen, Mathai,
Bismut, Berline, Vergne, etc.

We make only one remark that there is a phenomenon, called "supersymmetry" in physics, which is a mechanism of cancellation. The place where the index theorem is related to physics is in the relation to this mechanism.[3]

Characteristic features of this book

In this book, we will present the second proof of the Atiyah-Singer index theorem as far as we can, avoiding the use of pseudo-differential operators. We also intend to explain some applications of the index theorem.

So far, we have explained the index theorem according to its time evolution. An outline of the theory will be given in Chapter 1, titled "Prelude". Chapter 1 is a "sketch", and some parts use terminology without giving the definition. Thus it is not necessary that the reader read it carefully at first. The real arguments start in Chapter 2. By looking at the corresponding part of Chapter 1 again before reading the real arguments, the outline of the theory may play the role of side remarks or comments.

In this book, we deal only with "operators of Dirac type associated with $\mathbb{Z}_2$-graded Clifford modules" rather than with all elliptic differential operators. The key is the behaviour of operators of Laplace type, which is the square of operators of Dirac type. We present the formulation of the index theorem before introducing the K-group and the characteristic classes.

Our guiding principle is the "localization". This is the phenomenon showing that the index can be captured by looking at parts of the manifold without dealing with the whole manifold. The essential part of this localization arises, prior to the properties of the index, as the naive phenomenon that eigenfunctions for the operator of Laplace type with a potential decay exponentially, whereas the potential function takes large values.

The localization to open submanifolds is formulated as the "excision theorem". The localization to closed submanifolds is studied through the tensor product of the operator of Dirac type on the closed submanifold and a certain operator along its normal direction.

[3]It is not necessarily the association with linear differential equations where the supersymmetry arises.

For the proof of the index theorem, we only use differential operators (we do not use pseudo-differential operators). The index will be defined for operators of Dirac type not only on closed submanifolds, but also on open submanifolds with certain conditions at infinity for the operator. We do not assume orientability of manifolds and do not deal with odd-dimensional manifolds. Recently, such objects have appeared in some applications (for example, the index theorem on the 3-dimensional Euclidean space is necessary for analysis of monopoles).

Operators on open manifolds play an important role not only in such applications, but also in theoretical considerations. The most fundamental role in this book is the "unit", namely the operator of Dirac type on the Euclidean space, which is called the "supersymmetric harmonic oscillator" in physics.

Prerequisites.

In geometry, we assume a basic knowledge of manifolds, differential forms and vector bundles. There are also some places where the reader will need a little knowledge in topology (e.g., covering spaces, homotopy groups, the Poincaré duality, etc.).

In algebra, we assume knowledge of linear algebra, including the exterior product, the tensor product, etc. We also use very basic facts on the representation theory of finite groups and compact groups in § 2. 3 and § 4. 4.

In analysis, we assume knowledge of local properties of partial differential equations on the Euclidean space. However, the place where some prerequisite in analysis in the first volume is required is only in Chapter 4 (the Rellich theorem, *a-priori* estimates, regularity for solutions of elliptic differential operators, etc.). If the reader is interested only in geometric arguments, then (s)he can skip Chapter 4 and proceed to later chapters. The second volume [4] contains the study of non-linear differential equations, and we quote some facts from functional analysis.

[4]Editor's Note: *Index Theorem. 2* is forthcoming.

Terminology and notation.

- $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Z}$: the sets of complex numbers, real numbers, and integers, respectively. We also use some standard notation without mentioning them here.
- $\mathbb{Z}_2$: we abbreviate $\mathbb{Z}/2\mathbb{Z}$ in this way (it does not stand for the 2-adic numbers).
- $\operatorname{Ker} f, \operatorname{Coker} f,\ \operatorname{Im} f$: for a linear mapping f, we denote its kernel, cokernel and image by these symbols.
- id_E: the identity on E.
- Unless otherwise stated, "vector spaces" are finite dimensional.
- Unless otherwise stated, "manifolds" are differentiable manifolds, "vector bundles" are smooth vector bundles, and "functions" and "sections" are smooth.
- For a smooth vector bundle F on a manifold X, we denote by $(F)_x$ the fiber at $x \in X$. We also denote by $\Gamma(F)$ the set of all smooth sections of F.
- ϵ_E: the endomorphism of a $\mathbb{Z}_2$-graded vector space (or vector bundle) $E = E^0 \oplus E^1$, which coincides with id_{E^0} on E^0 and with $-\mathrm{id}_{E^1}$ on E^1.
- Since a "$\mathbb{Z}_2$-graded Clifford module with compact support" is too long, we simply call it a "pair".
- $H^k(X, \theta)$: the de Rham cohomology with coefficients in the local system θ on X.
- We tried to use notation systematically throughout the book; e.g., most of the time E stands for a vector space or Euclidean space T a vector bundle with Euclidean metric, F a vector bundle, which we take the tensor product as coefficients, but we could not avoid coincidences of notation. For example, F_A is the symbol for the curvature and may be confused with the symbol F for a vector bundle. The reader is asked to be careful with such cases.

CHAPTER 1

Prelude

The index is the information extracted from the linear mapping between infinite dimensional vector spaces, which comes from linear differential operators, so that it is invariant under "small perturbations". The index theorem provides the way of computing the index. It is easy to formulate and prove the index theorem in the one-dimensional case, so we start by explaining it.

1. What Is the Index?

1.1. Definition of the Index. Let $f : E^0 \to E^1$ be a linear mapping between finite dimensional vector spaces E^0 and E^1. Then we have

$$\dim \operatorname{Ker} f - \dim \operatorname{Coker} f = \dim E^0 - \dim E^1.$$

We interpret this fact in two ways.

(1) Suppose that we want to compute the difference of the dimension of the vector spaces E^0 and E^1. If a linear mapping f is given, it helps us with the computation as follows. The key point is that f induces an isomorphism from any complementary subspace F of $\operatorname{Ker} f$ to the image $\operatorname{Im} f$. In particular, their dimensions are the same so that they cancel each other when we consider the difference between the dimensions. Hence it is enough to compute the difference between the dimensions of $\operatorname{Ker} f$ and the complementary subspace of $\operatorname{Im} f$.

(2) The dimension of $\operatorname{Ker} f$ and $\operatorname{Coker} f$ depend on the mapping f. But the difference between the dimensions is independent of f. Therefore it gives a quantity, which carries a basic feature, like a "spine" of the mapping f.

These two points are obvious when we deal with finite dimensional vector spaces. The feature, however, becomes different when we deal with infinite dimensional vector spaces.

Let $P : \mathcal{E}^0 \to \mathcal{E}^1$ be a linear mapping between infinite dimensional vector spaces $\mathcal{E}^0$ and $\mathcal{E}^1$. Then the argument (1) above does not make sense, since each term of

$$\dim \mathcal{E}^0 - \dim \mathcal{E}^1$$

is infinity and does not make sense as it is.

How is the argument (2) above? Suppose that $\operatorname{Ker} P$ and $\operatorname{Coker} P$ are finite dimensional, so that the difference

$$\dim \operatorname{Ker} P - \dim \operatorname{Coker} P$$

is defined as an integer. In such a case, P is called **Fredholm**. However, the difference above depends on P.

For example, let us consider the set of all sequences of real numbers indexed by integers $(a_k)_{k\in\mathbb{Z}}$. This set is a real vector space in an obvious way. We define linear mappings P_0 and P_1 from $\mathcal{E}^0$ to $\mathcal{E}^1$, where $\mathcal{E}^0 = \mathcal{E}^1 = \{(a_k)_{k\in\mathbb{Z}} | a_k \in \mathbb{R}\}$.

Let P_0 be the identity mapping. Then $\operatorname{Ker} P_0 = \{0\}$ and $\operatorname{Coker} P_0 = \{0\}$, thus we have

$$\dim \operatorname{Ker} P_0 - \dim \operatorname{Coker} P_0 = 0.$$

On the other hand, we define P_1 as follows. Let (a_k) be a sequence as above. For $k < 0$, we put $b_k = a_k$. For $k \geq 0$, we put $b_k = a_{k+1}$. Then we define

$$P_1((a_k)_{k\in\mathbb{Z}}) = (b_k)_{k\in\mathbb{Z}}.$$

Then we have

$$\operatorname{Ker} P_1 = \mathbb{R}(\cdots, 0, 0, 1, 0, 0, \cdots), \quad \operatorname{Coker} P_1 = \{0\},$$

and, hence,

$$\dim \operatorname{Ker} P_1 - \dim \operatorname{Coker} P_1 = 1.$$

Comparing P_0 and P_1, the difference between them is caused by "infinite dimensionality" in the sense that the infinitely many basis elements are shifted. The difference between the dimension of the kernel and the dimension of the cokernel is invariant when we perturb P_0 in "finite dimensional way". More precisely, it is an exercise in linear algebra to show that, for a Fredholm linear mapping $P : \mathcal{E}^0 \to$

$\mathcal{E}^1$, $P+Q$ is also a Fredholm linear mapping, where $Q : \mathcal{E}^0 \to \mathcal{E}^1$ is a linear mapping with the finite dimensional image, and satisfies

$$\dim \operatorname{Ker}(P+Q) - \dim \operatorname{Coker}(P+Q) = \dim \operatorname{Ker} P - \dim \operatorname{Coker} P.$$

From consideration of the above, neither of two interpretations can work directly in the infinite dimensional case. However, this discussion provides a strategy.

(1) Suppose that we would like to define something like differences between dimensions of infinite dimensional vector spaces $\mathcal{E}^0$ and $\mathcal{E}^1$. When a Fredholm mapping P is given, we can use it and proceed as follows. The key point is that the restriction of P to any complementary subspace $\mathcal{F}$ to $\operatorname{Ker} P$ is an isomorphism onto $\operatorname{Im} P$. We regard that they cancel each other. Then the difference between the dimensions of the remaining parts (complementary subspaces) can be taken as the definition of the difference between the dimensions of the total spaces.

(2) Dimensions of $\operatorname{Ker} P$ and $\operatorname{Coker} P$ depend on a Fredholm mapping P. But the difference between these dimensions hardly changes under "small perturbation" of P. For example, it is invariant under addition of Q, where Q is a linear mapping with finite dimensional image. Hence this quantity must be considered as a basic property of P, something like a "spine".

In this way, we can define something like the difference between the dimensions of $\mathcal{E}^0$ and $\mathcal{E}^1$ for a triple $(\mathcal{E}^0, \mathcal{E}^1, [P])$. Here, P is a Fredholm operator and $[P]$ is an equivalence class, which only captures its "spine".

Conversely, the supplementary datum which is necessary for defining the difference between the dimensions of infinite dimensional vector spaces is provided by a certain equivalence class of P.

What is the definition of an appropriate equivalence relation? We, at least, want P and $P+Q$ to belong to the same equivalence class. If we require only this property, the equivalence class is too narrow.

There is a framework for dealing with such a problem in the most general way, which is provided by functional analysis.

For example, suppose that the norms are equipped to $\mathcal{E}^0$ and $\mathcal{E}^1$ so that they are complete, i.e., they are Banach spaces. We assume that P is continuous (in other words, bounded). Then the closure of the set of linear mappings with finite dimensional image is the set of compact operators, namely, operators which map a bounded set to a relatively compact set. In another respect, the operator norm measures the difference between operators, and we have one way of making precise the meaning of "small perturbation". In this situation, we can prove that the difference between the dimension of the kernel and the cokernel is invariant under continuous deformation through Fredholm operators. Then, in this functional analytical framework, we get the best possible definition for the equivalence relation given by the condition that operators belong to the same connected component of the space of Fredholm operators. Also, Fredholm operators are simply characterized as those which induce isomorphisms between Banach spaces up to compact operators[1].

REMARK 1.1. If the difference between the dimension of the kernel and the dimension of the cokernel is not zero, this implies existence of non-zero elements either in the kernel or the cokernel. In analysis, to show existence, without explicit construction, in an abstract way, some completeness condition is usually required. (Recall the relation between the mean value theorem and the completeness of real numbers.) By requiring the completeness for norm spaces, the property of Fredholmness is preserved under continuous deformation.

As a theory, it is simple and beautiful. It is also useful, in practice, to apply this functional analytical framework, when we deal with concrete examples of Fredholm operators appearing in analysis. Moreover, there is a subject[2] which discusses how much of geometry can be reconstructed by a thorough procedure in such a framework, beyond just a generalization, and tries to make a new sense of geometry.

It is, however, possible to investigate the operators more concretely than general theory in functional analysis, when they arise as differential operators on manifolds. The Atiyah-Singer index theorem, which we study in this book, concerns such concrete Fredholm operators. Explicit consideration on concrete Fredholm operators enables us to compute the difference between dimensions of infinite dimensional vector spaces.

---

[1] There exists P' such that $P'P - I$ and $PP' - I$ are compact operators.

[2] Non-commutative differential geometry is such a subject.

Before presenting a bit more explanation in this section, we conclude the general framework by returning to the arguments (1) and (2).

Summing up the arguments (1) and (2) in a different way:

For a Fredholm linear mapping $P : \mathcal{E}^0 \to \mathcal{E}^1$, we approximate it by finite dimensional object as follows.

Namely, we decompose them into direct sums:

$$\mathcal{E}^0 = E^0 \oplus \mathcal{F}^0, \quad \mathcal{E}^1 = E^1 \oplus \mathcal{F}^1, \quad P = f \oplus P',$$

where E^0 and E^1 are finite dimensional, and P' is an isomorphism between infinite dimensional vector spaces. (We can take any such decomposition in the consideration.) Then this f is a "finite dimensional approximation" of P, and

$$\dim E^0 - \dim E^1$$

is independent of the choice of such a decomposition and determined only by P.

If we take $E^0 = \operatorname{Ker} P$ and $E^1 = \operatorname{Coker} P$, this reduces to the argument (1) before. Also, a proof of (2) is given by choosing common E^0 and E^1 for both of P and $P + Q$.

Now, we define the index.

Definition 1.2. For a Fredholm operator P, we define its **index** $\operatorname{ind} P$ by

$$\operatorname{ind} P = \dim E^0 - \dim E^1,$$

where E^0 and E^1 are finite subspaces appearing in a finite dimensional approximation $f : E^0 \to E^1$. It does not depend on the choice of finite dimensional approximations. In particular, we have

$$\operatorname{ind} P = \dim \operatorname{Ker} P - \dim \operatorname{Coker} P.$$

In the framework of Fredholm operators between Banach spaces, this definition must be fundamental. On the other hand, when we work with Hilbert spaces (hence, equipped with an inner product), we may deal $\operatorname{Ker} P$ and $\operatorname{Coker} P$ in an equal way as follows.

Firstly, we consider finite dimensional situation $f : E^0 \to E^1$, where E^0 and E^1 are equipped with inner products. Then we have the adjoint mapping $f^* : E^1 \to E^0$ and $\operatorname{Coker} f$ is isomorphic to $\operatorname{Ker} f^*$. We also have

$$\operatorname{Ker} f = \operatorname{Ker} f^* f, \quad \operatorname{Ker} f^* = \operatorname{Ker} f f^*,$$

hence,

$$\begin{aligned}\dim \operatorname{Ker} f - \dim \operatorname{Coker} f* &= \dim \operatorname{Ker} f - \dim \operatorname{Ker} f^* \\ &= \dim \operatorname{Ker} f^* f - \dim \operatorname{Ker} f f^*.\end{aligned}$$

Let us put it in a formal way.

Take a direct sum decomposition $E = E^0 \oplus E^1$. In order to keep the decomposition in mind, we introduce a self-adjoint operator ϵ_E, which is the identity on E^0 and -1 times the identity on E^1. Then E^0 and E^1 are recovered as eigenspaces of ϵ_E. We call such a structure a $\mathbb{Z}_2$-graded vector space.

Now we put

$$\tilde{f} = \begin{pmatrix} 0 & f^* \\ f & 0 \end{pmatrix},$$

then it is a symmetric transformation anti-commuting with ϵ_E. (If we work with complex vector spaces with Hermitian inner products, we have a Hermite transformation.) Namely,

$$\tilde{f}\epsilon_E = -\epsilon_E \tilde{f}.$$

We say that an endomorphism operator, which is anti-commutative with ϵ_E, has **degree 1**.

The operator

$$\tilde{f}^2 = \begin{pmatrix} f^* f & 0 \\ 0 & f f^* \end{pmatrix}$$

is a symmetric transformation commuting with ϵ_E (when we work with complex vector spaces, it is a Hermitian transformation) whose eigenvalues are all non-negative. We say that an endomorphism operator, which commutes with ϵ_E, has **degree 0**.

In this way, after introducing $\tilde{f}$, we have

$$\dim \operatorname{Ker} f - \dim \operatorname{Coker} f^* = \dim \operatorname{Ker}(\tilde{f}^2 | E^0) - \dim \operatorname{Ker}(\tilde{f}^2 | E^1).$$

Decompose the eigenvalues according to the decomposition $E = E^0 \oplus E^1$, the multiplicities are the same except for the zero eigenvalue. In fact, f and f^* give isomorphisms between these eigenspaces. Based on this fact, we also get

$$\dim \operatorname{Ker}(\tilde{f}^2 | E^0) - \dim \operatorname{Ker}(\tilde{f}^2 | E^1) = \dim E^0 - \dim E^1.$$

From the consideration above, we formulate the notion of the index as follows.

DEFINITION 1.3. Let $\mathcal{E} = \mathcal{E}^0 \oplus \mathcal{E}^1$ be a $\mathbb{Z}_2$-graded (infinite dimensional) real vector space with an inner product and $\tilde{P}$ a symmetric operator of degree 1 such that $\operatorname{Ker} \tilde{P}^2$ is finite dimensional. Then we define the **index** $\operatorname{ind} \tilde{P}$ by

$$\operatorname{ind} \tilde{P} = \dim \operatorname{Ker}(\tilde{P}^2|\mathcal{E}^0) - \dim \operatorname{Ker}(\tilde{P}^2|\mathcal{E}^1).$$

When we work with infinite dimensional complex vector spaces with a Hermitian inner product, we assume that $\tilde{P}$ is a Hermitian operator.

In the case of a bounded Fredholm operator P between Hilbert spaces, Definition 1.2 and Definition 1.3 are equivalent, provided $\tilde{P}$ is introduced in the same way. We shall, mainly, adopt Definition 1.3.

1.2. A Variant of the Index 1. In general, we may regard quantities defined for a Fredholm operator P, which are independent of the choice of finite dimensional approximations, as indices in an extended sense. In the framework of functional analysis, the indices are quantities, which are invariant under continuous deformation. Here are some variants of indices.

First of all, we present "toy models" in a finite dimensional situation.

EXAMPLE 1.4. Let g^0 and g^1 be endomorphisms of complex vector spaces E^0 and E^1, respectively. If a linear mapping $f : E^0 \to E^1$ satisfies $fg^0 = g^1 f$, we have

$$\operatorname{trace}(g^0| \operatorname{Ker} f) - \operatorname{trace}(g^1| \operatorname{Coker} f) = \operatorname{trace}(g^0|E^0) - \operatorname{trace}(g^1|E^1).$$

REMARK 1.5. We can also claim the following fact related to Example 1.4. Let E^0 and E^1 be representation spaces of a compact Lie group G, and f a G-equivariant linear mapping. Then we have

$$[\operatorname{Ker} f] - [\operatorname{Coker} f] = [E^0] - [E^1]$$

as elements in the character ring $R(G)$. Here, $R(G)$ is a module consisting of formal differences between two finite dimensional representations of G. A module, obtained from a commutative semi-group by adjoining formal differences, is sometimes called the Grothendieck group.

EXAMPLE 1.6. If E carries a Euclidean inner product and f is anti-symmetric, we have

$$\dim \operatorname{Ker} f = \dim E, \qquad \mod 2.$$

REMARK 1.7. In the case of a complex vector space, we have a similar formula, provided f is anti-symmetric with respect to a non-degenerate quadratic form. Both of them are consequences of the fact that eigenvalues appear in pairs $\pm\lambda$ except the zero eigenvalue.

Both cases in examples can be generalized to infinite dimensional Fredholm setting under similar assumptions. Namely, we can define indices using finite dimensional approximation. We will get a complex valued index in the former case and $\mathbb{Z}_2 = \mathbb{Z}/2$-valued index in the latter case.

1.3. The Index for Families of Fredholm Operators. We presented an index in the case with group actions. The index was originally defined as the difference between dimensions of $\operatorname{Ker} P$ and $\operatorname{Coker} P$. If the operator P has a symmetry under a group action, we can define a quantity, which is not just about dimensions but contains information about the symmetry under the group action.

In other words, as a principle, there must be something like $\operatorname{Ker} P - \operatorname{Coker} P$, a more direct object than just the difference between dimensions, and we extract mathematically rigorous quantities and define the indices. Such an object is supposed to vary continuously, although we do not make clear what "vary continously" means.

When we want to define the "difference" between $\operatorname{Ker} P$ and $\operatorname{Coker} P$, not just as the difference between dimensions, how should we formulate it in a rigorous mathematical way?

The question here is not just about the index and a rather general question. From an object (say, P), we would like to formulate a "naive object" (e.g., the difference of $\operatorname{Ker} P$ and $\operatorname{Coker} P$) as a mathematical concept. The problem is the following: What kind of procedures and achievement realize our desire to formulate such an object?

REMARK 1.8. When a "naive object" is supposed to be a number, then there is no problem. Also, when the "procedure" is explicitly presented (like the substitution of elements in various rings into a rational polynomial), the description of the procedure itself realizes the content of the "naive object" (like morphisms between algebraic varieties defined over rational numbers).

What we discuss here is the way of defining a notion whose "character" is not clearly known in advance.

One of the answers, which is probably due to Grothendieck, is to study all continuous families of "objects" varying with parameters, not just one object (P). If a notion (corresponding to the index) is simultaneously and systematically defined for all such families, we can discuss how the index varies under deformations of P. (The meaning of "continuity" depends case by case.) This may not be the final answer, but it is certainly necessary, at least, to study families in order to construct a satisfactory theory.

Let us consider a continuous family of Fredholm operators with a parameter space. For simplicity, we assume that $\operatorname{Ker} P$ is of constant dimension for all P belonging to the family and $\operatorname{Coker} P$ is also constant dimension. Then the totality of $\operatorname{Ker} P$ with all possible P becomes a (finite dimensional) vector bundle over Z. The same thing holds for $\operatorname{Coker} P$. The point is that these vector bundles are possibly "twisted" in a topological sense. We can take the formal difference of these two vector bundles in terms of the Grothendieck group and get more information than just as a number. This is the notion of **topological K-theory**. We define the index of a family of Fredholm operators as an element in the K-group of the parameter space Z. Even in the case that $\operatorname{Ker} P$ and $\operatorname{Coker} P$ are not of constant dimension, it is known that this framework is enough to define the index. (At least, it is true for a compact topological space Z.)

The notion of the index for families of Fredholm operators is a variant of the index, but not just that.

The space of Fredholm operators (on a fixed Hilbert space) and the K-group are intimately related through the index for families. Not only is the index of families of Fredholm operators defined as elements of the K-group, but also any element in the K-group can be realized as the index of a certain family of Fredholm operators, conversely. Moreover, elements of the K-group and homotopy classes of families of Fredholm operators are in one-to-one correspondence. Shortly, the space of Fredholm operators is the classifying space for the K-group. This implies that the K-group is precisely (sufficiently and necessarily) the receiver of the index.

REMARK 1.9. The notion of the index of families can be extended, in a similar way, to the case with group actions and the case of $\mathbb{Z}_2$-valued index.

1.4. A Variant of the Index 2. Variants of the index, which we discussed so far, are topological objects in the sense that they are invariant under continuous deformations of Fredholm operators.

But there are other quantities, which depend on finer structure than topological data and are not invariant under continuous deformations.

We shall present "toy models" in a finite dimensional situation.

EXAMPLE 1.10. For a symmetric transformation f on a real vector space E equipped with a Euclidean inner product, the determinant $\det f$ is a real number determined by f.

EXAMPLE 1.11. For a symmetric transformation f on a real vector space E equipped with a Euclidean inner product, we consider the difference between the number of positive eigenvalues and the number of negative eigenvalues. This quantity is called the **signature** of f or the quadratic form associated to f.

These two quantities are related to eigenvalues of the linear mapping f. Neither of them makes sense in infinite dimensional situations without making certain "normalizations" in some sense. If we simply take the limit of the quantities for finite dimensional approximations, it diverges.

A standard way of normalization is based on the analytic continuation of so-called ζ-function, which is defined by using infinitely many eigenvalues. It is, certainly, defined under the assumption that infinitely many eigenvalues behave in a nice way. In fact, when we work with a natural set-up arising in geometric problems, such a nice behavior is guaranteed.

These quantities are the higher objects next to the index. In other words, the index is the most basic concept when dealing with these "higher objects". The index is captured by a rather rough "sieve", namely topology, but we need a finer sieve in order to capture higher objects.

In the geometry of manifolds, the corresponding object to the determinant is the analytic torsion, and the corresponding object to the signature is the η-invariant.

Under a certain situation, the analytic torsion gives a topological invariant, but it still reflects a subtler information (Reidemeister torsion) than the case of the usual index (characteristic numbers). Moreover, there is a natural object (Quillen metric), which arises naturally in trying to understand the index more geometrically than just as a number (for example, in the case of index of a family of operators).

The η-invariant arises in study of the index for an operator, on a compact manifold with boundary, with a certain boundary condition in a global nature (Atiyah-Patodi-Singer). The signature in the finite dimensional case (Example 1.11) is an integer, but the η-invariant, which is defined through normalization, is a real number, which may not be an integer.

These quantities are variants (or relatives) of the index, but they are beyond the scope of this book, which studies linear differential operators from a viewpoint of topology.

2. What Is the Atiyah-Singer Index Theorem?

2.1. Elliptic Differential Operators. In a previous section, we discussed what kind of notion the index is.

Typical examples of Fredholm operators are elliptic differential operators on manifolds.

For example, let us consider the space of rapidly decreasing functions $\mathcal{S}(X)$ on $X = \mathbb{C}$, the complex plane and an endomorphism P given by the Cauchy-Riemann equation:

$$P : f(x + \sqrt{-1}\, y) \mapsto 2\overline{\partial} f = (\partial_x + \sqrt{-1}\, \partial_y) f.$$

We study the effect of P on each wave function, after decomposing rapidly decreasing functions into wave functions with various frequencies. Using the Fourier transformation, P becomes a multiplication operator, $p_z = p_x + \sqrt{-1}\, p_y$, acting on the space of rapidly decreasing functions $\mathcal{S}(\widehat{X})$ on the dual space $\widehat{X}$ of X. Here, p_x and p_y are the real coordinates of the dual space and p_z is the complex coordinate. This multiplication operator naturally comes out from the coefficient of the operator P and is called the **principal symbol** of P.

The multiplication by p_z is almost invertible. If $p_z \neq 0$, we can make a division by p_z. Note that Fredholm operators are those which give isomorphisms after ignoring finite dimensional ambiguity. Hence, P seems a candidate of Fredholm operators.

It, however, turns out that

$$\dim \operatorname{Ker} P = 0, \quad \dim \operatorname{Coker} P = \infty,$$

and P is not a Fredholm operator. The division by p_z cannot make sense, even after ignoring finite dimensional ambiguity. In fact, we cannot divide functions in $\mathcal{S}(\widehat{X})$, which behave like polynomials in $\overline{p_z}$, by p_z.

To investigate the behavior, around $p_z = 0$, of functions on $\widehat{X}$ corresponds to the study of wave components of low frequencies, whose effects remain at infinity on the original space X. For example, we can cut off constant functions to get compactly supported functions, which contain many such wave components. Existence of such functions is an obstruction for P to being Fredholm.

Conversely, if we restrict the behavior of functions on X at infinity, we can get a Fredholm operator.

One of such realizations is to work with a torus $\mathbb{C}/L$, which is the quotient of $X = \mathbb{C}$ by a lattice $L = \mathbb{Z} + \mathbb{Z}\tau \quad (\Im\tau > 0)$. After introducing such a compact space, the Cauchy-Riemann operator becomes Fredholm. Using Fourier expansion, we can see that Ker and Coker are one-dimensional and its index is zero.

Another way is to add a term, without containing differentiation, to the Cauchy-Riemann operator. Then the behavior at infinity is automatically controlled. We shall explain an explicit way of doing this in Chapter 3.

In general, it is not difficult to see whether differential operators with constant coefficients acting on functions on an n-dimensional torus $(\mathbb{R}/2\pi\mathbb{Z})^n$ is Fredholm, using the Fourier transformation. Write the coordinates $q^1, q^2, \ldots, q^n \mod 2\pi\mathbb{Z}$. Then a homogeneous differential operator, of order k, with constant coefficients is written as $\sigma(\partial_{q^1}, \partial_{q^2}, \ldots, \partial_{q^n})$, where σ is a homogeneous polynomial of order k with constant coefficients. The Fourier transformation maps rapidly decreasing functions on X to rapidly decreasing sequences with multi integer valued indices $p_1, p_2, \ldots, p_n$. Also, the differential operator becomes a multiplying operator by $\sigma(p_1, p_2, \ldots, p_n)$. A simple sufficient condition for this operator to be Fredholm is that $\sigma(p_1, p_2, \ldots, p_n)$ never vanishes for $(p_1, p_2, \ldots, p_n) \in \mathbb{R}^n - \{0\}$.

The definition of σ above can be extended to arbitrary linear differential operator P acting on sections of vector bundles over a closed manifold and is called the **principal symbol** of P.

If the principal symbol $\sigma(P)(v)$ has an inverse for non-zero v, P is called an **elliptic differential operator**. It is known that an elliptic differential operator of order k is Fredholm as an operator acting on smooth sections of vector bundles. As we discussed in a previous section, we need to take appropriate completions of function spaces in order to apply the general framework in Banach spaces. But completions are not unique. Rather, we could say that the essential feature of P is considered as an operator acting on smooth sections and the choice of completion is a means to investigate the operator P.

Eventually, the index is determined by the homotopy class of the principal symbol of the elliptic differential operator P of order k.

The Atiyah-Singer index theorem presents the way the index is expressed.

2.2. The Dirac Operator. It may be natural to think first order differential operators are basic among all differential operators. The Cauchy-Riemann operator is one of such operators. Are there examples of other first order elliptic differential operators? The **Laplace operator** $-(\partial_{q^k})^2$ is a typical example of second order elliptic differential operators. In fact, it is the only second order linear differential operator, up to constant multiple, acting on functions, which is invariant under parallel translations and rotations in the Euclidean space. If we can find a first order linear differential operator whose square is the Laplacian, then it must be elliptic. Although there are no such operators acting on functions, there are such operators after introducing differential operators with coefficients in matrices. It was Dirac who first noticed this fact. He studied how to deal with electrons in relativity in the framework of quantum mechanics.

Let us suppose that a linear differential operator of the form $D = \sum_k \gamma^k \partial_{q^k}$ is such an operator. Here γ^k are constant matrices. Then the coefficient matrices γ^k should satisfy the following:

$$\gamma^k \gamma^l + \gamma^l \gamma^k = -2\delta^{kl},$$

where δ^{kl} is the Kronecker delta. The algebra generated by these γ^k is called the **Clifford algebra**. In short, the **Dirac operator** is a formally self-adjoint operator associated to a representation of the Clifford algebra.

We can generalize the notion of the Dirac operator on other manifolds. In this book, we discuss the Atiyah-Singer index theorem for these Dirac operators. More precisely, we deal with Dirac operators which carry a $\mathbb{Z}_2$-grading structure. We can define the index of the Dirac operator using the square of the Dirac operator as in Definition 1.3.

REMARK 1.12. To be precise, Dirac studied an object corresponding to the differential operator $\sqrt{-1}\,\partial_{q^0} + D$ on $\mathbb{R} \times X$. Here we denote by q^0 the coordinate of $\mathbb{R}$ and by D an operator of Dirac type, namely, the square of D being the Laplacian. We assume that D acts on sections of a $\mathbb{Z}_2$-graded vector bundle W and is of degree 1. The differential operator above acts on sections of a vector bundle obtained by pulling back W to $\mathbb{R} \times X$. Denote by ϵ_W the involutive automorphism corresponding to the $\mathbb{Z}_2$-grading structure. Identifying $\sqrt{-1}\,\epsilon_W$ and γ^0 corresponding to the direction of q^0, we can construct a Clifford algebra on $\mathbb{R} \times X$. Moreover, the operator

$$\widetilde{D} = \epsilon_W(\sqrt{-1}\,\partial_{q^0} + D)$$

is an operator of Dirac type on $\mathbb{R} \times X$. The notion of operators of Dirac type does not depend on the direct product decomposion of $\mathbb{R} \times X$ into $\mathbb{R}$ and X. This fact corresponds to the fact that the equation found by Dirac is "relativistic", i.e., it is independent of the choice of axes of the time ($\mathbb{R}$) and the space (X). [3]

2.3. Quantum Mechanics and Locality. The square of an operator of Dirac type is an operator of Laplace type, which resembles the Laplace operator. On the n-dimensional Euclidean space, a typical example of operators of Laplace type is given by

$$P = -\sum_{k=1}^{n} \partial_{q^k}^2 + V(q^1, q^2, \ldots, q^n).$$

Here V is a real valued function on $\mathbb{R}^n$. These operators appear in Schrödinger equations in (time independent) quantum mechanics. The first term, which is a genuine Laplace operator, corresponds to the kinetic energy in the classical theory and the second term V corresponds to the potential energy and is called the **potential**. The sum P, which corresponds to the total energy, is called the Hamiltonian operator.

REMARK 1.13. In the setting of quantum mechanics, X is the "space", which is not involved with the time. To include the "time axis", the space $\mathbb{R} \times X$ in Remark 1.12 is the "space-time".

Denote by $K(\overline{\lambda})$ the subset of X, where V is at most $\overline{\lambda}$. Let us suppose that $K(\overline{\lambda})$ is compact.

In the classical theory, particles with energy being at most $\overline{\lambda}$ move only in the bounded region $K(\overline{\lambda})$.

In quantum mechanics, we consider that states with energy being at most $\overline{\lambda}$ are represented by eigenfunctions with eigenvalues of P being at most $\overline{\lambda}$.

[3] More precisely, in the situation of Dirac's consideration, the metric on $\mathbb{R} \times X$ is not Riemannian, but the $\mathbb{R}$ direction is negative definite. So it is a Lorentzian metric. In this setting, we write $\gamma^0 = \epsilon_W$.

As we will see in Chapter 5, eigenfunctions with eigenvalue less than $\overline{\lambda}$ decay exponentially outside of $K(\overline{\lambda})$. This is a mathematical manifestation of the correspondence between the classical theory and quantum mechanics. (They are, however, not exactly zero outside of $K(\overline{\lambda})$. This is the tunneling effect.)

In other words, eigenfunctions with eigenvalue less than $\overline{\lambda}$ are almost "localized" to $K(\overline{\lambda})$.

A typical example is the case that $V = \sum_k (q^k)^2/2$, which corresponds to oscillation of a spring following the Hoock's law. In quantum mechanics, the corresponding system is called the **harmonic oscillator**.

The harmonic oscillator is not only simple in V, but also has basic significance in quantum mechanics. Because the algebraic structure (the Heisenberg algebra) behind the harmonic oscillator is used in formulation of annihilation and creation of particles in quantum field theory, in which an arbitrary number of particles can appear. In particular, we can describe annihilation and creation of n kinds of particles by using the harmonic oscillator on $\mathbb{R}^n$. It is important to note that the classical particles moving in $\mathbb{R}^n$ and the particles above, which may be annihilated and created, are completely different objects. Let us try to describe the annihilation and creation of particles in quantum field theory on a manifold X. For simplicity, we assume only one kind of particle. Then we put a copy of $\mathbb{R}$ at each point of X. Namely, we introduce a vector bundle over X with a fiber $\mathbb{R}$. Then we can consider a harmonic oscillator on the fiber over each point. The m-th eigenfunction of the harmonic oscillator on a fiber over x is regarded as the state with m particles at x. In this physical description, we need not think of the vector bundle with a fiber $\mathbb{R}$ as a geometric reality. We only need the structure of the Heisenberg algebra behind the harmonic oscillator.

2.4. Supersymmetric Harmonic Oscillator and a Proof of the Index Theorem. We can use the localizing phenomenon for eigenfunctions of operators of Laplace type as a basic tool to compute the index of the Dirac operator. In other words, the index, itself, is also "localized" in some sense.

This fact is one of reasons why the index of the Dirac operator can be written in terms of topological invariants (characteristic classes).

These invariants are related to cut-and-paste construction of vector bundles, on which the Dirac operator acts.[4]

The most basic example of operators of Dirac type is the **supersymmetric harmonic oscillator**, which is the harmonic oscillator coupled with a Clifford algebra structure, on the Euclidean space. In mathematical language, it is the conjugate of the de Rham complex (twisted by the orientation local system) by the multiplicative action by the function $\exp \sum_k (q^k)^2/2$.

We mentioned that the harmonic oscillator appears in the framework of describing the annihilation-and-creation of arbitrary number of particles. "Particles" in this context are called bosons in physics. The Clifford algebra is used in the framework of describing the annihilation-and-creation of other kinds of particles, called fermions. The supersymmetric harmonic oscillator involves an operator, which exchanges bosons and fermions (namely, a linear combination of the annihilation operator and the creation operator). This is the reason why it contains the word "supersymmetric". A mathematical expression for this symmetry among bosons and fermions is the symmetry of non-zero eigenvalues in terms of $\mathbb{Z}_2$-graded structure, which is mentioned just before Definition 1.3.

Suppose that $\mathbb{R}^n$ is assigned to each point of X. In other words, we consider a vector bundle over X with a fiber $\mathbb{R}^n$. Then, as we have seen above, we can introduce the supersymmetric harmonic oscillator on each fiber. In order to describe the annihilation-and-creation of bosons and fermions, it is not necessary to regard the vector bundle with a fiber $\mathbb{R}^n$ as a geometric reality, but we only need the structure of the Heisenberg algebra and the Clifford algebra.

However, we may ask whether this vector bundle can be considered as a geometric reality. For example, consider the situation that the manifold X is embedded in a bigger manifold $\widetilde{X}$ as a closed submanifold. Then a neighborhood of X is diffeomorphic to the total space of the normal bundle ν of X in $\widetilde{X}$. What can be deduced from understanding the family of supersymmetric harmonic oscillators acting on each fiber of ν?

The aim of this book is to explain a proof of the Atiyah-Singer index theorem from this viewpoint, in which we regard the above family as a geometric reality. Putting physical meaning and terminology

[4]There is another way of relating the index directly to local invariants without using potentials. Namely, it is the "heat kernel method". We, however, do not explain this argument in this book.

aside, it is just the procedure of constructing an operator of Dirac type on the total space of ν.

For a given operator of Dirac type on X, we first embed X in the Euclidean space of a sufficiently large dimension. Let $\widetilde{X}$ be this Euclidean space. The basic strategy of the proof of the index theorem given in Chapter 6 is to introduce the family of supersymmetric harmonic oscillators on ν as an auxiliary tool. This "auxiliary field" enables us to reduce our index computation to an easier case, as some computation of integration becomes easier after introducing auxiliary variables.[5] In our case, the index computation reduces to the index computation of the Dirac operator on Euclidean spaces. The last step of computing the index on Euclidean spaces can be performed with the help of the Bott periodicity theorem.

In other respects, we may regard this proof as based on the phenomenon that the index of the operator of Dirac type on the Euclidean space is localized to a neighborhood of a closed manifold X.

For a family of operators of Dirac type, we can also get the formula for the index of the family in a similar way.

2.5. Elliptic Differential Operators in Geometry. A smooth manifold is different from a topological space in the sense that the notion of differentiability of functions is assigned by the definition. Thorough discussion on smooth functions must reveal its differential topological feature. A typical example is the de Rham cohomology theory, in which topological invariants are derived using the exterior differentiation.

Some of such topological invariants are known to be given by the index of a certain elliptic differential operator.

Through the index theorem, we get information on some properties of the manifold in such a case. A typical example is the Hirzebruch signature theorem.

Associated to a geometric structure, say a complex structure, on a manifold, we get an elliptic differential operator so that its index carries some geometric meaning. Such a consideration leads us to understanding the Riemann-Roch-Hirzebruch theorem for closed complex manifolds.

---

[5] It is easier to compute the integration of $e^{-x^2-y^2}$ on $\mathbb{R}^2$ than the integration of e^{-x^2} on $\mathbb{R}$.

We can also understand the Riemann-Roch-Hirzebruch theorem for families of complex manifolds in a similar way. In fact, Grothendieck had proven the Riemann-Roch-Grothendieck theorem in the realm of algebraic geometry (for arbitrary characteristics), before the proof of the index theorem for families given by Atiyah-Singer. The notion of "family" is given in the framework of fiber bundles, but Grothendieck established the extension of the Riemann-Roch theorem for morphisms, which are not necessarily the projection of fiber bundles. Historically, it is the first time the notion of K-group emerged (in the context of algebraic geometry).

2.6. Relations with K-Theory. Let us recall that the notion of K-group appears naturally in the course of defining the index of Fredholm operators.

On the other hand, the notion of K-group provides an appropriate "expression" in each step of the proof of the index theorem mentioned above. These two aspects are not independent. We can mix them up. Or it may be more adequate to claim that there is a category of manifolds in which these aspects are better to be mixed up. Like the Riemann-Roch-Grothendieck theorem really fits the category of algebraic varieties.

To begin with, how are operators of Dirac type constructed on manifolds? Having one operator of Dirac type, we can take a tensor product with arbitrary vector bundles (in the level of principal symbols) and obtain new operators of Dirac type.

- If we have an operator of Dirac type, which can be considered as a "basic object" (or a generator), on a manifold X, then all operators of Dirac type should be obtained in this way. In such a case, the manifold X is said to carry a $Spin^c$-structure. For example, complex manifolds carry $Spin^c$-structure.
- Meanwhile, the operator of Dirac type, appearing in the Hirzebruch signature theorem, always exists for any oriented manifold of even dimension. This is not a "basic object". We, however, can consider a system of all operators of Dirac type, which are obtained by tensoring vector bundles.

In the above two categories, there is a mechanism, which provides an operator of Dirac type to arbitrary vector bundle on the manifold.

Let us suppose that we have a proper continuous mapping $\phi : X \to Y$ in each category as above. We can write ϕ as a composition of an embedding and a projection mapping. For example, identifying X with the graph of ϕ, ϕ is expressed as a composition of an embedding of X to $X \times Y$ and the projection mapping from $X \times Y$ to Y. This embedding appears in the proof of the index theorem. The projection mapping appears in study of the index for families. Taking this point into account and using the mechanism of constructing operators of Dirac type from vector bundles, we can see that these mappings induce mappings between K-groups. By composing them, we get $\phi_! : K(X) \to K(Y)$. The resulting mapping does not depend on the way of decomposing ϕ into an embedding and a projection mapping and is considered as a basic homomorphism between K-groups in each category above.

A topological description of this $\phi_!$ is the differential topological version of the Riemann-Roch-Grothendieck theorem and was carried out by Atiyah and Hirzebruch.[6]

The formulation of the index theorem using K-groups indicates that an arbitrary vector bundle over X may arise as the index of a family of something parametrized by X. As if we capture a vector bundle in reality as a finite dimensional approximation of a certain infinite dimensional object.

Thinking in this way, it is not hard to prove the Bott periodicity theorem. Atiyah realized that it naturally follows from properties of indices for families, by using the supersymmetric harmonic oscillator as an "auxiliary field".

The index for a family resembles the "integration along fibers" in an analogy with the de Rham cohomology theory. We can, actually, construct a generalized cohomology theory by extending the K-group to graded objects. Then the mapping $\phi_!$, associated to a proper continuous mapping ϕ, is understood as the Gysin mapping.

The key step in constructing a generalized cohomology theory out of K-group is the Thom isomorphism theorem for K-groups. This is an extension of the Bott periodicity theorem and proven in a similar way.

2.7. The Role and the Status of the Index of Elliptic Operators. As we have seen in a previous section, the description in terms of K-groups indicates a way of unifying a series of theories around the index theorem.

[6]Atiyah and Hirzebruch constructed the mapping between K-groups over $\mathbb{Q}$, more generally, i.e., without assuming existence of differential operators.

On the other hand, however, the notion of the index of elliptic linear differential operators can be seen as the first step toward the following two directions:

(1) higher invariants, which we mentioned in §1.4.
(2) study of non-linear operators.

For the former (1), it requires more than topological information as K-group, which deals with all stable homotopy classes.

As for the latter (2), there are known non-trivial examples related to geometry in dimension at most 4. In those cases, the index theorem (for families) is a basic tool, and the non-linear theory enjoys some properties similar to “locality” (Floer theory).

The notion of the index of elliptic linear differential operators carries “topological” features from the aspect (1) and is in debt to “linearity” from the aspect (2).

It is a challenging problem to define certain higher invariants, more than topological and related to non-linear differential equations, through some normalization procedure, which mixes up two aspects (1) and (2). This does not seem just one problem, and such a framework seems to be strongly requested when we try to understand some geometric phenomena in reality.

3. 1-Dimensional Case

In this section, we explain the index theorem on 1-dimensional manifolds. In this case, a solution of an ordinary linear differential equation can be uniquely determined by initial conditions. Based on this fact, the proof of the index theorem is reduced to linear algebra.

In higher dimensions, a manifold can be “curved” in various ways and we cannot hope to make an approach to the index theorem based on existence and uniqueness of the initial value problem. And we can capture the index, which is invariant under deformation, only after extending our consideration outside of genuine solutions, e.g., eigenfunctions with non-zero eigenvalues.

The argument in the 1-dimensional case depends heavily on special features in dimension 1, and it may seem that it does not mean much, although it provides examples of the index theorem.

It is, however, possible to regard a first order linear differential equation on $X = Z \times \mathbb{R}$ or $X = Z \times S^1$ (with Z a closed manifold or requiring certain appropriate conditions at infinity) as an ordinary differential equation with coefficients in matrices of infinite size.

In such cases, the argument for ordinary differential equations presents a nice model for investigation. We do not discuss these cases in detail, but explain a bit about "toy models" using ordinary differential equations.[7]

Connected 1-dimensional manifolds are $\mathbb{R}$ and S^1.

3.1. The Index Theorem on S^1. Let $A : \mathbb{R} \to M_r(\mathbb{C})$ be a smooth mapping with values in complex matrices of size r. Here, we assume that $A(x)$ is periodic with the period R: $A(x+R) = A(x)$. For $\mathbb{C}^r$-valued functions $f(x)$ and $g(x)$, we consider the following two linear differential equations:

$$\frac{df(x)}{dx} + A(x)f(x) = 0 \tag{1.1}$$

$$-\frac{dg(x)}{dx} + A(x)^* g(x) = 0 \tag{1.2}$$

Here $A^*(x)$ is the adjoint matrix of $A(x)$. We, in practice, study differential equations on $S^1 = [0,R]/(0 \sim R)$.

If we do not require the periodicity condition, the solution uniquely exists for a given initial value at a point. Hence the dimension of the space of periodic solutions is at most r. According to changes of $A(x)$, the dimension of periodic solutions varies and takes values between 0 and r.

The index theorem, in this setting, is the following:

THEOREM 1.14 (The index theorem on S^1). *The dimensions of these two differential equations coincide.*

Before the proof, we present a condition for a solution to be periodic. We abbreviate the first equation in the form $(\nabla + A)f = 0$. Solving this equation on $[0,R]$, we get a linear mapping $\phi : \mathbb{C}^r \to \mathbb{C}^r$, which assigns the value $f(R) = v_1$ at $x = R$ of the solution to an arbitrary initial value $f(0) = v_0$ at $x = 0$. The periodic solutions are exactly those corresponding to fixed vectors of ϕ.

In the same way, we get a linear mapping $\psi : \mathbb{C}^r \to \mathbb{C}^r$ from the second equation $(-\nabla + A^*)g = 0$. The periodic solutions correspond to fixed points of ψ.

[7]In one dimension higher, Tomoyoshi Yoshida constructed a theory, whose toy model is differential equations on $S^1 \times \mathbb{R}$ or $\mathbb{R} \times \mathbb{R}$. He used it in the study of Floer homology.

Therefore the theorem follows from the fact that ϕ and ψ^{-1} are adjoint, which we see below. We have to note certain relations between two differential equations for the proof. Namely, $\nabla + A$ and $-\nabla + A^*$ are "formally adjoint" to each other.

We review formal adjoint operators.

We define a Hermitian inner product on $\mathbb{C}^r$ by

$$\langle u, v\rangle = \sum_{i=1}^{r} u_i \overline{v}_i,$$

for $u = (u_1, u_2, \ldots, u_r)$ and $v = (v_1, v_2, \ldots, v_r)$. Then the fact that the operators $\nabla + A$ and $-\nabla + A^*$ are formally adjoint means that the following equality holds:

$$\int_{x_0}^{x_1} \langle (\nabla + A)f, g\rangle dx - \int_{x_0}^{x_1} \langle f, (-\nabla + A^*)g\rangle dx = \langle f(x_1), g(x_1)\rangle - \langle f(x_0), g(x_0)\rangle.$$

If f or g vanishes at $x = x_0, x_1$, the right hand side becomes zero. Or, as we assume right now, if f and g are periodic functions with the period R, then we have

$$\int_0^R \langle (\nabla + A)f, g\rangle dx = \int_0^R \langle f, (-\nabla + A^*)g\rangle dx.$$

From these equalities, we formally derive that $\nabla + A$ and $-\nabla + A^*$ are adjoint with respect to the Hermitian inner product:

$$(f_1, f_2) = \int_0^R \langle f_1, f_2\rangle dx, \quad (g_1, g_2) = \int_0^R \langle g_1, g_2\rangle dx.$$

Now we give a proof of the index on S^1.

Proof. If functions f and g, with values in $\mathbb{C}^r$, on $\mathbb{R}$ satisfies $(\nabla + A)f = 0$ and $(-\nabla + A^*)g = 0$, we have

$$\langle f(x_1), g(x_1)\rangle = \langle f(x_0), g(x_0)\rangle.$$

Putting $x_0 = 0$ and $x_1 = R$, it implies that ϕ and ψ^{-1} are adjoint. □

3.2. The Index Theorem on $\mathbb{R}$. Next, we give a similar discussion on differential equations over $\mathbb{R}$. Solutions here are supposed to converge to zero at infinity. The goal is to compare the dimensions of the spaces of solutions of $(\nabla + A)f = 0$ and $(-\nabla + A^*)g = 0$, respectively. Then we are required to study, closely, their behavior at infinity.

Let $A : \mathbb{R} \to M_r(\mathbb{C})$ be a smooth mapping with values in complex square matrices of size r, which satisfies

$$A(x) = \begin{cases} A_0 & (x \leq R_0) \\ A_1 & (x \geq R_1) \end{cases}$$

for some complex matrices A_0 and A_1. We consider the following ordinary differential equation for a vector valued function $f(x)$, i.e., $f : \mathbb{R} \to \mathbb{C}^r$,

$$\frac{df(x)}{dx} + A(x)f(x) = 0.$$

Pick a point on $\mathbb{R}$. For any initial value, the solution exists uniquely. Hence the space of solutions is an r-dimensional complex vector space. We, however, restrict ourselves to solutions which decay at infinity.

We write the space of such solutions

$$\mathrm{Ker}(\nabla + A) = \left\{ f(x) \;\Big|\; \frac{df(x)}{dx} + A(x)f(x) = 0, \lim_{x\to\pm\infty} f(x) = 0 \right\}.$$

Here, we use the notations $\nabla = \frac{d}{dx}, A = A(x)$. Changing the function $A(x)$ on a compact interval $-R \leq x \leq R$, the dimension of this vector space may vary.

Now, we introduce another ordinary differential equation

$$-\frac{dg(x)}{dx} + A^*(x)g(x) = 0.$$

Here, $g(x)$ is a $\mathbb{C}^r$-valued function. Write

$$\mathrm{Ker}(-\nabla + A^*) = \left\{ g(x) \;\Big|\; -\frac{dg(x)}{dx} + A^*(x)g(x) = 0, \lim_{x\to\pm\infty} g(x) = 0 \right\}.$$

The dimension of this vector space may also vary according to changes of $A^*(x)$, in other words, changes of $A(x)$.

A remarkable fact is that the difference between the dimensions of $\mathrm{Ker}(\nabla + A)$ and $\mathrm{Ker}(-\nabla + A^*)$ does not change, even though $A(x)$ varies. More precisely, we have the following:

THEOREM 1.15 (the index theorem on $\mathbb{R}$). *If* 0 *is not the real part of eigenvalues of* A_0 *and* A_1*, we have*

$$\dim \mathrm{Ker}(\nabla + A) - \dim \mathrm{Ker}(-\nabla + A^*) = \frac{\mathrm{sign}(A_1)}{2} - \frac{\mathrm{sign}(A_0)}{2}.$$

Here $\mathrm{sign}(B)$ *for a square matrix* B *denotes the number of eigenvalues of* B *with positive real part subtracted by the number of those with negative real part.*

Note that the right hand side only depends on A_0 and A_1. The left hand side, the difference of dimensions of solution spaces, is the index of the operator $\nabla + A$.

Since the proof is simple, we omit it.

REMARK 1.16. As we noticed before, arguments for ordinary differential equations can be a model for a part of the argument for first order differential equations on a manifold $X = Z \times \mathbb{R}$. The above equality involving with sign is a "toy model" of the Atiyah-Patodi-Singer theory, which expresses the index with the help of η-invariant. More explanation on this topic is beyond the scope of this book.

3.3. The Index Theorem on $\mathbb{R}$ with Actions. Fix a Hermitian matrix T of size r. Let $A(x)$ be a smooth mapping, with values in complex matrices of size r, on $\mathbb{R}$. We assume that $A(x)$ takes constant values A_0 and A_1 on $x \leq R_0$ and $x \geq R_1$, respectively, and that 0 is not the real part of any eigenvalue of these matrices. We also require that $TA(-x) = -A(x)T$ holds, in particular, we have $TA_0 = -A_1 T$. Denote by τ_V and τ_W the linear mappings $f(x) \mapsto Tf(-x)$ and $g(x) \mapsto -T^* g(-x)$, respectively. Since $T = T^*$, we have $\tau_W(\nabla + A) = (\nabla + A)\tau_V$ and $\tau_V(-\nabla + A^*) = (-\nabla + A^*)\tau_W$. Write $\mathrm{Ker}(\nabla + A)$ and $\mathrm{Ker}(-\nabla + A^*)$ the spaces of solutions which converge zero at infinity. These solution spaces are preserved under the action of τ_V and τ_W.

THEOREM 1.17 (the index theorem on $\mathbb{R}$ with actions).

$$\mathrm{trace}(\tau_V | \mathrm{Ker}(\nabla + A)) - \mathrm{trace}(\tau_W | \mathrm{Ker}(-\nabla + A^*)) = \mathrm{trace}\, T.$$

A similar result also holds on S^1. The details are left to readers.

EXERCISE 1. Formulate the index theorem on S^1 with actions and prove it. (The answer is not unique.)

3.4. The mod 2 Index Theorem in Dimension 1. Let $A(x)$ be a smooth mapping with values in skew symmetric complex matrices of size r. We consider a differential equation $(\nabla + A)f = 0$ for a $\mathbb{C}^r$-valued function $f(x)$.

(1) The case over S^1. Fix $R > 0$ and a complex orthogonal matrix O of size r. We assume twisted periodicity for $A(x)$ and $f(x)$:

$$A(x + R) = OA(x)O^{-1}, \quad f(x + R) = Of(x),$$

and consider $f(x)$ as a section of a vector bundle, twisted by O, over $S^1 = \mathbb{R}/R\mathbb{Z}$.

(2) The case over $\mathbb{R}$. Fix two real numbers $R_0 < R_1$ and two skew symmetric matrices A_0 and A_1. Here we assume that 0 is not the real part of any eigenvalue of A_0 and A_1, which only exist in the case that r is odd.

We also require that $A(x)$ takes constant values A_0 and A_1 on $x \leq R_0$ and $x \geq R_1$, respectively, and that $f(x)$ converges to 0 when $|x| \to \infty$.

Under these situations, we have the following result. (The proof is also reduced to linear algebra as discussed in a previous section.)

THEOREM 1.18.

(1). *The case over* S^1. *Write* $\epsilon = 0, 1$ *according to* $\det O = (-1)^\epsilon$. *We have*

$$\dim \operatorname{Ker}(\nabla + A) \equiv r + \epsilon \pmod 2.$$

(2). *The case over* $\mathbb{R}$. *Write* $\epsilon = 0, 1$ *according to*

$$\frac{\text{the product of eigenvalues of } A_0 \text{ with positive real part}}{\operatorname{Pf}(A_0)}$$

$$= (-1)^\epsilon \frac{\text{the product of eigenvalues of } A_1 \text{ with positive real part}}{\operatorname{Pf}(A_1)}.$$

Here Pf *denotes the Pfaffian of a skew symmetric matrix. Then we have*

$$\dim \operatorname{Ker}(\nabla + A) \equiv \epsilon \pmod 2.$$

(The usual integer valued index, defined before, is zero in these cases.)

REMARK 1.19. We give a supplementary explanation on skew symmetric matrices. For a skew symmetric matrix C, the square of $\operatorname{Pf}(C)$ equals $\det(C)$. On the other hand, if 0 is not the real part of any eigenvalue of C, the square of the product of eigenvalues with positive real part equals $(-1)^{r/2}\det(C)$. Hence the ratio of them, before taking the square, is one of $\pm(\sqrt{-1}\,)^{r/2}$. If we take a conjugation of C by a complex orthogonal matrix O, this ratio is multiplied by $\det O$. This fact implies that the set of skew symmetric matrices splits into two classes.[8]

EXERCISE 2. Prove Theorem 1.18.

[8]It is related to the way two maximally isotropic subspaces intersect.

Summary.
1.1 The index of Fredholm operators is defined as the difference of the dimensions of Ker and Coker. It is also possible to define the index for Fredholm Hermitian transformations, of degree 1, on a $\mathbb{Z}_2$-graded vector space.
1.2 The index is invariant under deformations. There are variants of the index, which is invariant under deformations. In particular, we can define the index for a family of Fredholm operators, which takes values in K-group.
1.3 Elliptic linear differential operators on (closed) manifolds are Fredholm, and their indices enjoy a certain property of locality.
1.4 We can interpret the locality above using ideas in quantum mechanics. Conversely, we may put it that some arguments in quantum mechanics are justified mathematically.