Index Theory with Applications to Mathematics and Physics

David Bleecker

Bernhelm Booß–Bavnbek

Draft version of April 25, 2012.

Contents

Synopsis vii

Preface xi

Part I. Operators with Index and Homotopy Theory 1

Chapter 1. Fredholm Operators 2
1. Hierarchy of Mathematical Objects 2
2. The Concept of Fredholm Operator 3
3. Algebraic Properties. Operators of Finite Rank. The Snake Lemma 5
4. Operators of Finite Rank and the Fredholm Integral Equation 9
5. The Spectra of Bounded Linear Operators: Basic Concepts 10

Chapter 2. Analytic Methods. Compact Operators 12
1. Analytic Methods. The Adjoint Operator 12
2. Compact Operators 18
3. The Classical Integral Operators 26
4. The Fredholm Alternative and the Riesz Lemma 26
5. Sturm-Liouville Boundary Value Problems 29
6. Unbounded Operators 34
7. Trace Class and Hilbert-Schmidt Operators 53

Chapter 3. Fredholm Operator Topology 63
1. The Calkin Algebra 63
2. Perturbation Theory 65
3. Homotopy Invariance of the Index 68
4. Homotopies of Operator-Valued Functions 72
5. The Theorem of Kuiper 77
6. The Topology of $\mathcal{F}$ 81
7. The Construction of Index Bundles 83
8. The Theorem of Atiyah-Jänich 88
9. Determinant Line Bundles 91
10. Essential Unitary Equivalence and Spectral Invariants 107

Chapter 4. Wiener-Hopf Operators 120
1. The Reservoir of Examples of Fredholm Operators 120
2. Origin and Fundamental Significance of Wiener-Hopf Operators 121
3. The Characteristic Curve of a Wiener-Hopf Operator 122
4. Wiener-Hopf Operators and Harmonic Analysis 123
5. The Discrete Index Formula. The Case of Systems 125
6. The Continuous Analogue 129

Part II. Analysis on Manifolds 133

Chapter 5. Partial Differential Equations in Euclidean Space, Revisited 134
1. Linear Partial Differential Equations 134
2. Elliptic Differential Equations 137
3. Where Do Elliptic Differential Operators Arise? 139
4. Boundary-Value Conditions 141
5. Main Problems of Analysis and the Index Problem 143
6. Numerical Aspects 143
7. Elementary Examples 144

Chapter 6. Differential Operators over Manifolds 154
1. Motivation 154
2. Differentiable Manifolds - Foundations 155
3. Geometry of C^∞ Mappings 158
4. Integration on Manifolds 164
5. Exterior Differential Forms and Exterior Differentiation 168
6. Covariant Differentiation, Connections and Parallelity 172
7. Differential Operators on Manifolds and Symbols 178
8. Manifolds with Boundary 187

Chapter 7. Sobolev Spaces (Crash Course) 190
1. Motivation 190
2. Definition 191
3. The Main Theorems on Sobolev Spaces 197
4. Case Studies 200

Chapter 8. Pseudo-Differential Operators 203
1. Motivation 203
2. Canonical Pseudo-Differential Operators 207
3. Principally Classical Pseudo-Differential Operators 211
4. Algebraic Properties and Symbolic Calculus 225
5. Normalized (Global) Amplitudes 228

Chapter 9. Elliptic Operators over Closed Manifolds 234
1. Continuity of Pseudo-Differential Operators 234
2. Elliptic Operators - Regularity and Fredholm Property 236

Chapter 10. Local Elliptic Boundary-Value Systems 240
1. Differential Equations in Half-Space 240
2. Systems of Differential Equations with Constant Coefficients 244
3. Variable Coefficients 246
4. The Topology of Local Boundary Value Conditions (Case Study) 247
5. Generalizations (Heuristic) 251
6. The Poisson Principle 255
7. The Green Algebra 256
8. The Elliptic Case 258

Part III. The Atiyah-Singer Index Formula 263

Chapter 11. Introduction to Algebraic Topology (K-Theory) 265

1. Winding Numbers 265
2. The Topology of the General Linear Group 269
3. The Ring of Vector Bundles 274
4. K-Theory with Compact Support 277
5. Proof of the Periodicity Theorem of R. Bott 280

Chapter 12. The Index Formula in the Euclidean Case 286
1. Index Formula and Bott Periodicity 286
2. The Difference Bundle of an Elliptic Operator 287
3. The Index Formula 291

Chapter 13. The Index Theorem for Closed Manifolds 295
1. The Index Formula 295
2. Comparison of the Proofs: The Cobordism Proof 297
3. Comparison of the Proofs: The Imbedding Proof 299
4. Comparison of the Proofs: The Heat Equation Proof 299

Chapter 14. Applications (Survey) 305
1. Cohomological Formulation of the Index Formula 305
2. The Case of Systems (Trivial Bundles) 307
3. Examples of Vanishing Index 308
4. Euler Number and Signature 309
5. Vector Fields on Manifolds 313
6. Abelian Integrals and Riemann Surfaces 316
7. The Theorem of Riemann-Roch-Hirzebruch 320
8. The Index of Elliptic Boundary-Value Problems 323
9. Real Operators 328
10. The Lefschetz Fixed-Point Formula 328
11. Analysis on Symmetric Spaces 330
12. Further Applications 331

Part IV. Index Theory in Physics and the Local Index Theorem 333

Chapter 15. Physical Motivation and Overview 334
1. Classical Field Theory 335
2. Quantum Theory 343

Chapter 16. Geometric Preliminaries 365
1. Principal G-Bundles 365
2. Connections and Curvature 366
3. Equivariant Forms and Associated Bundles 370
4. Gauge Transformations 380
5. Curvature in Riemannian Geometry 386
6. Bochner-Weitzenböck Formulas 405
7. Characteristic Classes and Curvature Forms 414
8. Holonomy 426

Chapter 17. Gauge Theoretic Instantons 431
1. The Yang-Mills Functional 431
2. Instantons on Euclidean 4-Space 436

3. Linearization of the Manifold of Moduli of Self-dual Connections 459
4. Manifold Structure for Moduli of Self-dual connections 467

Chapter 18. The Local Index Theorem for Twisted Dirac Operators 483
1. Clifford Algebras and Spinors 483
2. Spin Structures and Twisted Dirac Operators 494
3. The Spinorial Heat Kernel 507
4. The Asymptotic Formula for the Heat Kernel 517
5. The Local Index Formula 543
6. The Index Theorem for Standard Geometric Operators 562

Chapter 19. Seiberg-Witten Theory 611
1. Background and Survey 611
2. Spin^c Structures and the Seiberg-Witten Equations 621
3. The Manifold of Moduli 629
4. Compactness of Moduli Spaces and the Definition of S-W Invariants 650

Appendix A. Fourier Series and Integrals - Fundamental Principles 669
1. Fourier Series 669
2. The Fourier Integral 671

Appendix B. Vector Bundles 676
1. Basic Definitions and First Examples 676
2. Homotopy Equivalence and Isomorphy 679
3. Clutching Construction and Suspension 681

Bibliography 686

Synopsis

Preface. Target Audience and Prerequisites. Outline of History. Further Reading. Questions of Style. Acknowledgments and Dedication

Chapter 1. Fredholm Operators. Hierarchy of Mathematical Objects. The Concept of a Bounded Fredholm Operator in Hilbert Space. Algebraic Properties. Operators of Finite Rank. The Snake Lemma of Homological Algebra. Product Formula. Operators of Finite Rank and the Fredholm Integral Equation. The Spectra of Bounded Linear Operators (Terminology)

Chapter 2. Analytic Methods. Compact Operators. Adjoint and Self-Adjoint Operators - Recalling Fischer-Riesz. Dual characterization of Fredholm operators: either by finite-dimensional kernel and cokernel, or by finite-dimensional kernels of the operator and its adjoint operator *and* closed image. Compact Operators: Spectral Decomposition, Why Compact Operators also are Called *Completely Continuous*, $\mathcal{K}$ as Two-Sided Ideal, Closure of Finite-Rank Operators, and Invariant under *. Classical Integral Operators. Fredholm Alternative and Riesz Lemma. Sturm-Liouville Boundary Value Problems. Unbounded Operators: Comprehensive Study of Linear First Order Differential Operators Over S^1: Sobolev Space, Dirac Distribution, Normalized Integration Operator as *Parametrix*, The Index Theorem on the Circle for Systems. Closed Operators, Closed Extensions, Closed (not necessarily bounded) Fredholm Operators, Composition Rule, Symmetric and Self–Adjoint Operators, Formally Self-Adjoint and Essentially Self-Adjoint. Spectral Theory. Metrics on the Space of Closed Operators. Trace Class and Hilbert-Schmidt Operators

Chapter 3. Fredholm Operator Topology. Calkin Algebra and Atkinson's Theorem. Perturbation Theory: Homotopy Invariance of the Index, Homotopies of Operator-Valued Functions, The Theorem of Kuiper. The Topology of $\mathcal{F}$: The Homotopy Type, Index Bundles, The Theorem of Atiyah-Jänich. Determinant Line Bundles: The Quillen Determinant Line Bundle, Determinants, The Segal-Furutani Construction. Spectral Invariants: Essentially Unitary Equivalence, What Is a Spectral Invariant? Eta Function, Zeta Function, Zeta Regularized Determinant

Chapter 4. Wiener-Hopf Operators. The Reservoir of Examples of Fredholm Operators. Origin and Fundamental Significance of Wiener-Hopf Operators. The *Characteristic Curve* of a Wiener-Hopf Operator. Wiener-Hopf Operators and Harmonic Analysis. The Discrete Index Formula. The Case of Systems. The Continuous Analogue

Chapter 5. Partial Differential Equations in Euclidean Space, Revisited. Review of Classical Linear Partial Differential Equations: Constant and Variable Coefficients, Wave Equation, Heat Equation, Laplace Equation, Characteristic Polynomial. Elliptic Differential Equations: Where Do Elliptic Differential Operators Arise? Boundary-Value Conditions. Main Problems of Analysis and the Index Problem. Calculations. Elementary Examples. The Hellwig-Vekua Problem with Non-Vanishing Index

Chapter 6. Differential Operators over Manifolds. Motivation. Differentiable Manifolds - Foundations: Tangent Space. Cotangent Space. Geometry of C^∞ Mappings: Embeddings. Immersions. Submersions. Embedding Theorems. Integration on Manifolds: Hypersurfaces, Riemannian Manifolds, Geodesics, Orientation. Exterior Differential Forms and Exterior Differentiation. Covariant Differentiation, Connections and Parallelity: Connections on Vector Bundles, Parallel Transport, Connections on the Tangent Bundle, Clifford Modules and Operators of Dirac Type. Differential Operators on Manifolds and Symbols: Our Data, Symbolic Calculus, Formal Adjoints. Elliptic Differential Operators. Definition and Standard Examples. Manifolds with Boundary

Chapter 7. Sobolev Spaces (Crash Course). Motivation. Equivalence of Different Local Definitions. Various Isometries. Global, Coordinate-Free Definition. Embedding Theorems: Dense Subspaces; Truncation and Mollification; Differential Embedding; Rellich Compact Embedding. Sobolev Spaces Over Half Spaces. Trace Theorem. Case Studies: Euclidean Space and Torus; Counterexamples

Chapter 8. Pseudo-Differential Operators. Motivation: Fourier Inversion; Symbolic Calculus; Quantization. *Canonical* and *Principally Classical* Pseudo-Differential Operators. Pseudo-Locality; Singular Support. Standard Examples: Differential Operators; Singular Integral Operators. Oscillatory Integrals. Kuranishi Theorem. Change of Coordinates. Pseudo-differential Operators on Manifolds. Graded *-Algebra. Invariant Principal Symbol; Exact Sequence; Non-Canonical Op as Right Inverse. Coordinate-Free (Truly Global) Approach: Normalized Fourier Transformation; Normalized Amplitudes; Normalized Invertible Op; Approximation of Differential Operators

Chapter 9. Elliptic Operators over Closed Manifolds. Continuity of Pseudo-Differential Operators between Sobolev Spaces. Parametrices for Elliptic Operators: Regularity and Fredholm Property

Chapter 10. Local Elliptic Boundary-Value Systems. Differential Equations in Half-Space. Systems of Differential Equations with Constant Coefficients. Variable Coefficients. The Topology of Local Boundary Value Conditions (Case Study)

Chapter 11. Introduction to Algebraic Topology (K-Theory). Winding Numbers. One-Dimensional Index Theorem. Counter-Intuitive Dimension Two: Disrobing Problem. The Topology of the General Linear Group. The Ring of Vector Bundles. K-Theory with Compact Support. Proof of the Bott Periodicity Theorem

Chapter 12. The Index Formula in the Euclidean Case. Index Formula and Bott Periodicity. The Difference Bundle of an Elliptic Operator. The Index Formula

Chapter 13. The Index Theorem for Closed Manifolds. Proof of the Index Theorem by Embedding: Axioms for the Analytic and the Topological Index. Comparison of the Proofs: The Cobordism Proof. Comparison of the Proofs: The Embedding Proof. Comparison of the Proofs: The Heat Equation Proof

Chapter 14. Applications (Survey). Cohomological Formulation of the Index Formula. The Case of Systems (Trivial Bundles). Examples of Vanishing Index. Euler Number and Signature. Vector Fields on Manifolds. Abelian Integrals and Riemann Surfaces. The Theorem of Riemann-Roch-Hirzebruch. The Index of Elliptic Boundary-Value Problems. Real Operators. The Lefschetz Fixed-Point Formula. Analysis on Symmetric Spaces. Further Applications

Chapter 15. Physical Motivation and Overview. Mode of Reasoning in Physics. String Theory and Quantum Gravity. The Experimental Side. Classical Field Theory: Newton-Maxwell-Lorentz, Faraday 2-Form, Abstract Flat Minkowski Space-Time, Relativistic Mass, Relativistic Kinetic Energy, Inertial System, Lorentz Transformations and Poincaré Group, Relativistic Deviation from Flatness, *Twin Paradox*, Variational Principles. Kaluza-Klein Theory: Simultaneous Geometrization of Electro-Magnetism and Gravity, Other *Grand Unified Theories*, String Theory. Quantum Theory: Photo-Electric Effect, Atomic Spectra, Quantizing Energy, State Spaces of Systems of Particles, Basic Interpretive Assumptions. Heisenberg Uncertainty Principle. Evolution with Time - The Schrödinger Picture. Nonrelativistic Schrödinger Equation and Atomic Phenomena. *Minimal Replacement* and Covariant Differentiation. Anti-Particles and Negative-Energy States. Unreasaonable Success of Standard Model. Dirac Operator vs. Klein-Gordon Equation. Feynman Diagrams

Chapter 16. Geometric Preliminaries. Principal G-Bundles. Connections and Curvature. Ricci Curvature. Equivariant Forms and Associated Bundles. Gauge Transformations. Curvature in Riemannian Geometry. Bochner-Weitzenböck Formulas. Characteristic Classes and Curvature Forms. Holonomy

Chapter 17. Gauge Theoretic Instantons. The Yang-Mills Functional. Instantons on Euclidean 4-Space. Linearization of the Manifold of Moduli of Self-dual Connections. Manifold Structure for Moduli of Self-dual connections

Chapter 18. The Local Index Theorem for Twisted Dirac Operators. Clifford Algebras and Spinors. Spin Structures and Twisted Dirac Operators. The Spinorial Heat Kernel. The Asymptotic Formula for the Heat Kernel. The Local Index Formula: The case $m = 1$ (surfaces), The case $m = 2$ (4-manifolds). The Index Theorem for Standard Geometric Operators.

Chapter 19. Seiberg-Witten Theory. Background and Survey. Spin^c Structures and the Seiberg-Witten Equations. The Manifold of Moduli. Compactness of Moduli Spaces and the Definition of S-W Invariants

Appendix A. Fourier Series and Integrals (fundamental principles). Fourier Series: The Fundamental Function Spaces on S^1; Density; Orthonormal Basis; Fourier Coefficients; Plancherel's Identity; Product and Convolution. The Fourier Integral: Different Integral Conventions; Duality Between Local and Global

- Point and Neighborhood - Multiplication and Differentiation - Bounded and Continuous; Fourier Inversion Formula; Plancherel and Poisson Summation Formulae; Parseval's Equality; Higher Dimensional Fourier Integrals

Appendix B. Vector Bundles. Basic Definitions and First Examples. Homotopy Equivalence and Isomorphy. Clutching Construction and Suspension

Bibliography. Key References. Classical and Recent Textbooks. References to Technical Details, History. Perspectives

Preface

Target Audience and Prerequisites. The mathematical philosophy of index theory and all its basic concepts, technicalities and applications are explained in Parts I-III. Those are the easy parts. They are written for upper undergraduate students or graduate students to bridge the gap between rule-based learning and first steps towards independent research. They are also recommended as general orientation to mathematics teachers and other senior mathematicians with different mathematical background. All interested can pick up a single chapter as bedside reading.

In order to enjoy reading or even work through Parts I-III, we expect the reader to be familiar with the concept of a smooth function and a complex separable Hilbert space. Nothing more. Instead of ascending systematically from simple concepts to complex ones in the classical Bourbaki style, we present a patch-work of definitions and results when needed. In each chapter we present a couple of fully comprehensible, important, deep mathematical *stories*. That, we hope, is sufficient to catch our three messages:

(1) Index theory is about *regularization*, more precisely, the index quantifies the defect of an equation, an operator, or a geometric configuration from being regular.
(2) Index theory is also about *perturbation invariance*, i.e., the index is a meaningful quantity stable under certain deformations and apt to store certain topological or geometric information.
(3) Most important for many mathematicians, the index *interlinks quite diverse mathematical fields*, each with its own very distinct research tradition.

Part IV is different. It is also self-contained. All concepts will be explained fully and rigorously, but much shorter than in the first Parts. However, this last Part is written for graduate students, PhD students and other experienced learners, interested in low-dimensional topology and gauge-theoretic particle physics. We try to explain the very place of index theory in *geometry* and for revisiting *quantum field theory*: There are thousands of other calculations, observations and experiments. But there is something special about the actual and potential contributions of index theory. Index theory is about chirality (asymmetry) of zero modes in the spectrum and classifies connections (back ground fields) and a variety of other intrinsic properties in geometry and physics. It is not just about some more calculations, some more numbers and relations.

Outline of History. When first considering infinite-dimensional linear spaces, there is the immediate realization that there are injective and surjective linear endomorphisms which are not isomorphisms, and more generally the dimension of the kernel minus that of the cokernel (i.e., the index) could be any integer. However, in the classical theory of Fredholm integral operators which goes back at least to the early 1900s (see [**Fred**]), one is dealing with compact perturbations of the identity and the index is zero. Several sources point to Fritz Noether (in his study [**No**] of singular integral operators, published in 1921), as the first to encounter the phenomenon of a nonzero index for operators naturally arising in analysis *and* to give a formula for the index in terms of a winding number constructed from data defining the operator. Over some decades, this result was generalized in various directions by G. Hellwig, I.N. Vekua and others (see [**Ve62**]), contrary to R. Courant's and D. Hilbert's expectation in [**CH**] that "linear problems of mathematical physics which are correctly posed behave like a system of N linear algebraic equations in N unknowns", i.e., they should satisfy the *Fredholm alternative* and always yield vanishing index. Meanwhile, many working mainly in abstract functional analysis were producing results, such as the stability of the index of a Fredholm operator under perturbations by compact operators or bounded operators of sufficiently small operator norm (e.g., first J.A. Dieudonné [**Di**], followed by F.W. Atkinson [**Atk**], B. Yood [**Yoo**], I.Z. Gohberg and M.G. Krein [**GK58**], etc.).

Around 1960, the time was ripe for I.M. Gelfand (see [**Ge**]) to propose that the index of an elliptic differential operator (with suitable boundary conditions in the presence of a boundary) should be expressible in terms of the coefficients of highest order part (i.e., the principal symbol) of the operator, since the lower order parts provide only compact perturbations which do not change the index. Indeed, a continuous, ellipticity-preserving deformation of the symbol should not affect the index, and so Gelfand noted that the index should only depend on a suitably defined homotopy class of the principal symbol. The hope was that the index of an elliptic operator could be computed by means of a formula involving only the topology of the underlying domain (the manifold), the bundles involved, and the symbol of the operator. In early 1962, M.F. Atiyah and I.M. Singer discovered the (elliptic) Dirac operator in the context of Riemannian geometry and were busy working at Oxford on a proof that the $\widehat{A}$-genus of a spin manifold is the index of this Dirac operator. At that time, Stephen Smale happened to pass through Oxford and turned their attention to Gelfand's general program described in [**Ge**]. Drawing on the foundational and case work of analysts (e.g., M.S. Agranovic, A.D. Dynin, L. Nirenberg, R.T. Seeley and A.I. Volpert), particularly that involving pseudo-differential operators, Atiyah and Singer could generalize F. Hirzebruch's proof of the Hirzebruch-Riemann-Roch theorem [**Hi66a**] and discovered and proved the desired index formula at Harvard in the Fall of 1962. Moreover, the Riemannian Dirac operator played a major role in establishing the general case. The details of this original proof involving cobordism actually first appeared in [**Pal65**]. A K-theoretic embedding proof was given in [**AS68a**], the first in a series of five papers. This proof was more direct and susceptible to generalizations (e.g., to G-equivariant elliptic operators in [**ASe**] and families of elliptic operators in [**AS71a**]).

The approach to proving the Index Theorem in [**AS68a**] is based on the following clever strategy, which we shall explain in detail in Chapters 11-13 of this book. The invariance of the index under homotopy implies that the index (say, the *analytic index*) of an elliptic operator is stable under rather dramatic, but continuous, changes of its principal symbol while maintaining ellipticity. Using this fact, one finds (after considerable effort) that the index (i.e., the *analytical index*) of an elliptic operator transforms predictably under various global operations such as embedding and extension. Using K-theory and Bott periodicity, a topological invariant (say, the *topological index*) with the same transformation properties under these global operations is constructed from the symbol of the elliptic operator. One then verifies that a general index function having these properties is unique, subject to normalization. To deduce the Atiyah–Singer Index Theorem (i.e., *analytic index = topological index*), it then suffices to check that the two indices are the same in the trivial case where the base manifold is just a single point. A particularly nice exposition of this approach for twisted Dirac operators over even-dimensional manifolds (avoiding many complications of the general case) is found in E. Guentner's article [**Gu93**] following an argument of P. Baum.

Not long after the K-theoretical embedding proof (and its variants), there emerged a fundamentally different means of proving the Atiyah–Singer Index Theorem, namely the *heat kernel method*. This is worked out here (see Chapter 18 in the important case of the chiral half $\mathcal{D}^+$ of a twisted Dirac operator $\mathcal{D}$. In the index theory of closed manifolds, one usually studies the index of a chiral half $\mathcal{D}^+$ instead of the total Dirac operator $\mathcal{D}$, since $\mathcal{D}$ is symmetric for compatible connections and then index $\mathcal{D} = 0$.) The heat kernel method had its origins in the late 1960s (e.g., in [**MS**], inspired by [**MiPl**] of 1949) and was pioneered in the works [**Pat**], [**Gi73**], [**ABP**]. In the final analysis, it is debatable as to whether this method is really much shorter or better. This depends on the background and taste of the beholder. Geometers and analysts (as opposed to topologists) are likely to find the heat kernel method appealing. The method not only applies to geometric operators which are expressible in terms of twisted Dirac operators, but also largely for more general elliptic pseudo-differential operators, as R.B. Melrose has done in [**Mel**]. Moreover, the heat method gives the index of a "geometric" elliptic differential operator naturally as the integral of a characteristic form (a polynomial of curvature forms) which is expressed solely in terms of the geometry of the operator itself (e.g., curvatures of metric tensors and connections). One does not destroy the geometry of the operator by using ellipticity-preserving deformations. Rather, in the heat kernel approach, the invariance of the index under changes in the geometry of the operator is a consequence of the index formula itself more than a means of proof. However, considerable analysis and effort are needed to obtain the heat kernel for $e^{-t\mathcal{D}^2}$ and to establish its asymptotic expansion as $t \to 0^+$. Also, it can be argued that in some respects the K-theoretical embedding/cobordism methods are more forceful and direct. Moreover, in [**LaMi**], we are cautioned that the index theorem for families (in its strong form) generally involves torsion elements in K-theory that are not detectable by cohomological means, and hence are not computable in terms of local densities produced by heat asymptotics. Nevertheless, when this difficulty does not arise, the K-theoretical expression for the topological index may be less appealing than the integral of a characteristic form, particularly for those who already understand and appreciate the geometrical formulation of characteristic classes. More importantly, the heat kernel approach exhibits the index as just one of a whole sequence of spectral invariants appearing as coefficients of terms of the asymptotic expansion (as $t \to 0^+$) of the trace of the relevant heat kernel. (On p. 117, we guide the reader to the literature about these particular spectral invariants and their meaning in modern physics. The required mathematics for that will be developed in Section 18.4.) All disputes aside, the student who learns *both* approaches and formulations to the index formula will be more accomplished (and probably a good deal older).

Further Reading. What the coverage of topics in this book is concerned, we hope our table of contents needs no elaboration, except to say that space limitations prevented the inclusion of some important topics (e.g., the index theorem for families, index theory for manifolds with boundary, other than the Atiyah-Patodi-Singer Theorem, Peter Kronheimer's and Thomas Mrowka's visionary work on knot homology groups from instantons, lists of all calculated spectral invariants). However, we now provide some guidance for further study. A fairly complete exposition, by Atiyah himself, of the history of index theory from 1963 to 1984 is found in Volume 3 of [**Ati**] and duplicated in Volume 4. Volumes 3, 4 and 5 contain many unsurpassed articles written by Atiyah and collaborators on index theory and its applications to gauge theory. We all owe a debt of gratitude to Herbert Schröder for the definitive guide to the literature on index theory (and its roots and offshoots) through 1994 in Chapter 5 of the excellent book [**Gi95**] of P.B. Gilkey. We have benefited greatly not only from this book, but also from the marvelous work [**LaMi**] by H.B. Lawson and M.L. Michelsohn. In that book, there are proofs of index formulas in various contexts, and numerous beautiful applications illustrating the power of Dirac operators, Clifford algebras and spinors in the geometrical analysis of manifolds, immersions, vector fields, and much more. The classical book [**Sha**] of P. Shanahan is also a masterful, elegant exposition of not only the standard index theorem, but also the G-index theorem and its numerous applications. A fundamental source on index theory for certain open manifolds and manifolds with boundary is the authoritative book [**Mel**] of R.B. Melrose. Very close to our own view upon index theory is the plea [**Fu99**] of M. Furuta for reconsidering the index theorem, with emphasis on the localization theorem. In the case of boundary-value problems for Dirac operators, we put quite some care in the writing of our [**BoWo**] jointly with K.P. Wojciechowski. The recent book [**FV**] of D. Fursaev and D. Vassilevich contains a detailed description of main spectral functions and methods of their calculation with emphasis on heat kernel asymptotics and their application in various branches of modern physics. A wealth of radically new ideas of (partly yet unproven) geometric use of instantons are given in [**KM08**] of P.B. Kronheimer and T.S. Mrowka. A taste of the recent revival of D-branes and other exotic instantons in string theory can be gained from [**GML**] of H. Ghorbani, D. Musso and A. Lerda. Indications can be found in the review [**Sa**] of F. Sannino about, how strongly coupled theories of gauge theoretic physics result in perceiving a composite universe and other new physics awaiting to be discovered.

The Question of Originality: Seeking a Balance between Mathematical Heritage and Innovation. Parts I-III and the two appendices do not aim at originality. These parts teach what mathematicians today consider general knowledge about the index theorem as one of the great achievements of 20th century mathematics. Actually, there are two novelties included which even not all experts may be aware of: The *first novelty* appears when rounding up our comprehensive presentation of the topology of the space of Fredholm operators: we do not halt with the Atiyah-Jänich Theorem and the construction of the index bundle, but also confront the student with a thorough presentation of the various definitions of determinant line bundles. This is to remind the student that index theory is not a more or less closed collection of results but a philosophy of regularization, of deformation invariance and of visionary cross connections within mathematics and between its various branches.

A *second novelty* in the first three Parts is the emphasis on global constructions, e.g., in our way of introducing and using the concept of pseudo-differential operators.

Apart from these two innovations, the student can feel protected in the first three Parts against any ambitious and possibly exaggerated, confusing and dispensable originality.

Basically, Part IV follows the same line: Happily we could avoid excessive originality in the chapters dealing with instantons and the Donaldson-Kronheimer-Seiberg-Witten results about the geometry of moduli spaces of connections. There we also summarize, refer, define, explain more or less like in the first three Parts.

However, the *core* of Part IV is different. It consists of an original, full, quite lengthy (in parts almost unbearably meticulous) proof of the *Local Index Theorem for twisted Dirac operators* in Chapter 18 and its applications to standard geometric operators. That long Chapter is thought as a new contribution to the ongoing search for a deeper understanding of the index theorem and the "best" approach to it:

Clearly, a student looking for the most general formulation of the index theorem and a proof apt for wide generalizations should concentrate on our Part III, the so-called Embedding (or K-theoretic) Proof. However, a student wanting to trace the germs of index calculations back in the geometry of the considered standard operators (all arising from various decompositions of the algebra of exterior differential forms) should consult Section 18.5 with a full proof of the *Local Index Formula* for twisted Dirac operators on spin manifolds (all terms will be explained) and Section 18.6, where we derive the *Index Theorem for Standard Geometric Operators*. These geometric index theorems are by far less general than Part III's embedding proof, but they are more geometric, and we hold, also more geometric than the *usual* heat equation proofs of the index theorem: Not striving for greatest generality, we obtain index formulas for the standard elliptic geometric operators and their twists. The standard elliptic geometric operators include the signature operator $d+\delta : (1+*)\,\Omega^*(M) = \Omega^+(M) \to \Omega^-(M) = (1-*)\,\Omega^*(M)$, the Euler-Dirac operator $d+\delta : \Omega^{even}(M) \to \Omega^{odd}(M)$, and the Dolbault-Dirac operator $\sqrt{2}\left(\bar\partial + \bar\partial^*\right) : \Omega^{-,even}(M) \to \Omega^{+,odd}(M)$ (all symbols will be defined). The index formula obtained for the above operators yields the Hirzebruch Signature Theorem, the Chern-Gauss-Bonnet Theorem, and the Hirzebruch-Riemann-Roch Theorem, respectively. While these operators generally are not globally twisted Dirac operators, locally they are expressible in terms of chiral halves of twisted Dirac operators. That applies also to the Yang-Mills operator. Thus, even if the underlying Riemannian manifold M (assumed to be oriented and of even dimension) does not admit a spin structure, we may still use the Local Index Theorem for twisted Dirac operators to compute the index density and hence the index of these operators. While it is possible to carry this out separately for each of the geometric operators, basically all of these theorems are consequences of one single index theorem for generalized Dirac operators on Clifford module bundles (all to be defined). Using the Local Index Theorem for twisted Dirac operators, we prove this index theorem first (our Theorem 18.55), and then we apply it to obtain the geometric index theorems, yielding the general Atiyah-Singer Index Theorem for practically all geometrically defined operators.

The Style. To present the rich world of index theory, we have chosen two different styles. We wrote all definitions, theorems, and proofs as concise as possible, like a flying arrow or a precise and reliable routing program that frees the reader or driver from all dispensable side information. Where possible, we begin the introduction of a new concept with a simple but generic example or a review of the local theory, immediately followed by the corresponding global or general concept. That is one half of the book, so to speak the odd numbered pages. The other half of the book consists of historical reviews, motivations, perspectives, examples, and exercises (often with extended hints). We wrote those sections in a more open web-like style.

Correspondingly, the student will meet different ways of emphasizing. Important definitions, notions, concepts are in bold face. Background information is in small, with key ideas underlined. In remarks and notes, leading terms are in italics.

Acknowledgments and Dedication. Due to circumstance, the responsible author had to finish the book alone, though on the basis of extended drafts for all chapters which had been worked out jointly with David Bleecker. The responsible author acknowledges the hints and help he received from G. Chen (Tianjin), G. Esposito (Napoli), P. Kirk (Bloomington) and M. Lesch (Bonn) in the critical phase of the final re-shuffling and updating. In particular, he thanks Roskilde University for an extraordinary leave for that work; the Università degli Studi di Napoli Federico II and the Istituto Nazionale di Fisica Nucleare (INFN, Sezione di Napoli) for their hospitality during his stay there; and H. Larsen and J. Larsen (Roskilde) and M. Lesch (Bonn) for generously providing useful LaTeX macros, hacks and environments and substantial help with the final lay-out.

Both authors agreed to dedicate this book to their teachers, to the memory of S.-S. Chern (thesis adviser of DB) on the occasion of his centenary in October 2011 and to F. Hirzebruch (thesis adviser of BBB) on the occasion of his approaching 85th birthday in October 2012.

B. Booß-Bavnbek, Roskilde (Denmark), June 2012

Part I

Operators with Index and Homotopy Theory

If we do not succeed in solving a mathematical problem, the reason frequently is our failure to recognize the more general standpoint from which the problem before us appears only as a single link in a chain of related problems.

David Hilbert, 1900

CHAPTER 1

Fredholm Operators

Synopsis. Hierarchy of Mathematical Objects. The Concept of a Bounded Fredholm Operator in Hilbert Space. The Index. Forward and Backward Shift Operators. Algebraic Properties. Operators of Finite Rank. The Snake Lemma of Homological Algebra. Product Formula. Operators of Finite Rank and the Fredholm Integral Equation. The Spectra of Bounded Linear Operators (Terminology and Basic Properties)

1. Hierarchy of Mathematical Objects

> "In the hierarchy of branches of mathematics, certain points are recognizable where there is a definite transition from one level of abstraction to a higher level. The first level of mathematical abstraction leads us to the concept of the individual numbers, as indicated for example by the Arabic numerals, without as yet any undetermined symbol representing some unspecified number. This is the stage of elementary arithmetic; in algebra we use undetermined literal symbols, but consider only individual specified combinations of these symbols. The next stage is that of analysis, and its fundamental notion is that of the arbitrary dependence of one number on another or of several others – the function. Still more sophisticated is that branch of mathematics in which the elementary concept is that of the transformation of one function into another, or, as it is also known, the operator."

Thus N. Wiener characterized the hierarchy of mathematical objects [**Wie33**, p.1]. Very roughly we can say: Classical questions of analysis are aimed mainly at investigations within the third or fourth level. This is true for real and complex analysis, as well as the functional analysis of differential operators with its focus on existence and uniqueness theorems, regularity of solutions, asymptotic or boundary behavior which are of particular interest here. Thereby research progresses naturally to operators of more complex composition and greater generality without usually changing the concerns in principle; the work remains directed mainly towards qualitative results.

In contrast it was topologists, as Michael Atiyah variously noted, who turned systematically towards quantitative questions in their topological investigations of algebraic manifolds, their determination of quantitative measures of qualitative behavior, the definition of global topological invariants, the computation of intersection numbers and dimensions. In this way, they again broadly broke through the rigid separation of the "hierarchical levels" and specifically investigated relations between these levels, mainly of the second and third level (algebraic surface – set of zeros of an algebraic function) with the first, but also of the fourth level (Laplace operators on Riemannian manifolds, Cauchy-Riemann operators, Hodge theory) with the first.

This last direction, starting with the work of Wilhelm Blaschke and William V. D. Hodge, continuing with Kunihiko Kodaira, Shiing-Shen Chern and Donald Spencer, with Henri Cartan and Jean Pierre Serre, with Friedrich Hirzebruch, Michael Atiyah and others, can perhaps be best described with the key word differential topology or analysis on manifolds. Both its relation with and distinction from analysis proper is that (from [**Ati68b**, p.57])

> "Roughly speaking we might say that the analysts were dealing with complicated operators and simple spaces (or were only asking simple questions), while the algebraic geometers and topologists were only dealing with simple operators but were studying rather general manifolds and asking more refined questions."

We can read, e.g. in [**Bri74**, p.278-283] and the literature given there, to what degree the contrast between quantitative and qualitative questions and methods must be considered a driving force in the development of mathematics beyond the realm sketched above.

Actually, in the 1920's already, mathematicians such as Fritz Noether and Torsten Carleman had developed the purely functional analytic concept of the index of an operator in connection with integral equations, and had determined its essential properties. But "although its (the theory of Fredholm operators) construction did not require the development of significantly different means, it developed very slowly and required the efforts of very many mathematicians" [**GK57**, p.185]. And although Soviet mathematicians such as Ilja V. Vekua had hit upon the index of elliptic differential equations at the beginning of the 1950's, we find no reference to these applications in the quoted principal work on Fredholm operators. In 1960 Israel M. Gelfand published a programmatic article asking for a systematic study of elliptic differential equations from this quantitative point of view. He took as a starting point the theory of Fredholm operators with its theorem on the homotopy invariance of the index (see below). Only after the subsequent work of Michail S. Agranovich, Alexander S. Dynin, Aisik I. Volpert, and finally of Michael Atiyah, Raoul Bott, Klaus Jänich and Isadore M. Singer, did it become clear that the theory of Fredholm operators is indeed fundamental for numerous quantitative computations, and a genuine link connecting the higher "hierarchical levels" with the lowest one, the numbers.

2. The Concept of Fredholm Operator

Let H be a (separable) complex Hilbert space, and let $\mathcal{B}$ denote the Banach algebra (e.g., [**Ped**, p.128], [**Rud**, p.228], [**Sche**, p.201]) of bounded linear operators $T : H \to H$ with the operator norm

$$\|T\| := \sup\{\|Tu\| \,:\, \|u\| \le 1\} < \infty,$$

where $\|\cdot\|$ denotes the norm in H induced by the inner product $\langle\,\cdot\,,\,\cdot\,\rangle$.

DEFINITION 1.1. An operator $T \in \mathcal{B}$ is called a **Fredholm operator**, if

$$\operatorname{Ker} T := \{u \in H \,:\, Tu = 0\} \text{ and } \operatorname{Coker} T := H/\operatorname{Im}(T)$$

are finite-dimensional.

This means that the homogeneous equation $Tu = 0$ has only finitely many linearly independent solutions, and to solve $Tu = v$, it is sufficient that v satisfies a finite number of linear conditions (e.g., see Exercise 2.1b below). We write $T \in \mathcal{F}$

and define the **index** of T by

$$\operatorname{index} T := \dim \operatorname{Ker} T - \dim \operatorname{Coker} T.$$

The **codimension** of $\operatorname{Im}(T) = T(H) = \{Tu : u \in H\}$ is $\dim \operatorname{Coker} T$.

REMARK 1.2. a) We can analogously define Fredholm operators $T : H \to H'$, where H and H' are Hilbert spaces, Banach spaces, or general topological vector spaces. In this case, we use the notation $\mathcal{B}(H, H')$ and $\mathcal{F}(H, H')$, corresponding to $\mathcal{B}$ and $\mathcal{F}$ above. However, in order to counteract a proliferation of notation and symbols in this section, we will deal with a single Hilbert space H and its operators as far as possible. The general case $H \neq H'$ does not require new arguments at this point. Later, however, we shall apply the theory of Fredholm operators to elliptic differential and pseudo-differential operators where a strict distinction between H and H' (namely, the compact embedding of the domain H into H', see Chapter 9) becomes decisive.
b) All results are valid for Banach spaces and a large part for Fréchet spaces also. For details see for example [**PRR**, p.182-318]. We will not use any of these but will be able to restrict ourselves entirely to the theory of Hilbert space whose treatment is in parts far simpler.
c) Motivated by analysis on symmetric spaces – with transformation group G – operators have been studied whose *index* is not a number but an element of the representation ring $R(G)$ generated by the characters of finite-dimensional representations of G [**AS68a**, p.519f] or, still more generally, is a distribution on G [**Ati74**, p.9-17]. We will not treat this largely analogous theory, nor the generalization of the Fredholm theory to the discrete situation of von Neumann algebras as it has been carried out – with real-valued index in [**Bre**, 1968/1969].

EXERCISE 1.3. Let $L^2(\mathbb{Z}_+)$ denote the space of sequences $c = (c_0, c_1, c_2, \ldots)$ of complex numbers with square-summable absolute values; i.e.,

$$\sum_{n=0}^{\infty} |c_n|^2 < \infty.$$

$L^2(\mathbb{Z}_+)$ is a Hilbert space (see Chapter A below). Show that the forward shift

$$\text{shift}^+ : (c_0, c_1, c_2, \ldots) \mapsto (0, c_0, c_1, c_2, \ldots)$$

and the backward shift

$$\text{shift}^- : (c_0, c_1, c_2, \ldots) \mapsto (c_1, c_2, c_3, \ldots)$$

are Fredholm operators with $\operatorname{index}(\text{shift}^+) = -1$ and $\operatorname{index}(\text{shift}^-) = +1$.
[Warning: Just as we can regard $L^2(\mathbb{Z}_+)$ as the limit of the finite dimensional vector spaces $\mathbb{C}^m$ (as $m \to \infty$), we can approximate shift^+ by endomorphisms of $\mathbb{C}^m$ given, relative to the standard basis, by the $m \times m$ matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Note that the kernel and cokernel of this endomorphism are one-dimensional, whence the index is zero. (See Exercise 1.4 below.).
We have yet another situation, when we consider the Hilbert space $L^2(\mathbb{Z})$ of sequences $c = (\dots, c_{-2}, c_{-1}, c_0, c_1, c_2, \dots)$ with

$$|c_0|^2 + \sum_{n=1}^{\infty} \left(|c_n|^2 + |c_{-n}|^2 \right) < \infty.$$

The corresponding shift operators are now bijective, and hence have index zero.]

3. Algebraic Properties. Operators of Finite Rank. The Snake Lemma

EXERCISE 1.4. For finite-dimensional vector spaces the notion of Fredholm operator is empty, since then every linear map is a Fredholm operator. Moreover, the index no longer depends on the explicit form of the map, but only on the dimensions of the vector spaces between which it operates. More precisely, show that every linear map $T : H \to H'$ where H and H' are finite-dimensional vector spaces has index given by

$$\operatorname{index} T = \dim H - \dim H'.$$

[Hint: One first recalls the vector-space isomorphism $H/\operatorname{Ker}(T) \cong \operatorname{Im}(T)$ and then (since H and H' are finite-dimensional) obtains the well-known identity from linear algebra

$$\dim H - \dim \operatorname{Ker} T = \dim H' - \dim \operatorname{Coker} T.]$$

[Warning: If we let the dimensions of H and H' go to ∞, we obtain only the formula $\operatorname{index} T = \infty - \infty$. Thus, we need an additional theory to give this difference a particular value.]

EXERCISE 1.5. For two Fredholm operators $F : H \to H$ and $G : H' \to H'$ consider the direct sum,

$$F \oplus G : H \oplus H' \to H \oplus H'.$$

Show that $F \oplus G$ is a Fredholm operator with

$$\operatorname{index}(F \oplus G) = \operatorname{index} F + \operatorname{index} G$$

[Hint: First verify that $\operatorname{Ker}(F \oplus G) = \operatorname{Ker} F \oplus \operatorname{Ker} G$, and the corresponding fact for $\operatorname{Im}(F \oplus G)$. Then show that

$$(H \oplus H')/(\operatorname{Im} F \oplus \operatorname{Im} G) \cong (H/\operatorname{Im} F) \oplus (H'/\operatorname{Im} G).]$$

[Warning: When $H = H'$, we can consider the sum (not direct) $F + G : H \to H$, but in general this is not a Fredholm operator; e.g., we could set $G := -F$.]

EXERCISE 1.6. Show by algebraic means, that the composition $G \circ F$ of two Fredholm operators $F : H \to H'$ and $G : H' \to H''$ is again a Fredholm operator.
[Hint: Which inequality holds between

$$\dim \operatorname{Ker} G \circ F \quad \text{and} \quad \dim \operatorname{Ker} F + \dim \operatorname{Ker} G$$

and between

$$\dim \operatorname{Coker} G \circ F \quad \text{and} \quad \dim \operatorname{Coker} F + \dim \operatorname{Coker} G \ ?]$$

[Warning: Why are these in general not equalities? Nevertheless, the chain rule $\operatorname{index} G \circ F = \operatorname{index} F + \operatorname{index} G$ can be proved, since the inequalities cancel out when we form the difference. Sometimes in mathematics a result of interest appears as an offshoot of the proof of a rather dull general statement. Such is the case when we obtain the chain rule for differentiable functions by showing that the composite of differentiable maps is again differentiable. Here, however, the proof of the index formula requires extra work, and we must either utilize the topological structure by functional analytic means (see Exercise 2.3, p.14) or refine the algebraic arguments. The latter will be done next.]

We recall a basic idea of *diagram chasing*:

DEFINITION 1.7. Let $H_1, H_2, \dots$ be a sequence (possibly finite) of vector spaces and let $T_k : H_k \to H_{k+1}$ be a linear map for each $k = 1, 2, \dots$. We call

$$H_1 \xrightarrow{T_1} H_2 \xrightarrow{T_2} H_3 \xrightarrow{T_3} \cdots$$

an **exact sequence**, if for all k

$$\operatorname{Im} T_k = \operatorname{Ker} T_{k+1}.$$

In particular, the following is a list of equivalences and implications (where 0 denotes the 0-dimensional vector space)

$$\begin{aligned}
0 \longrightarrow H_1 \xrightarrow{T_1} H_2 \text{ exact} \quad &\Leftrightarrow \quad \operatorname{Ker} T_1 = 0 \ (T_1 \text{ injective}),\\
H_1 \xrightarrow{T_1} H_2 \longrightarrow 0 \text{ exact} \quad &\Leftrightarrow \quad \operatorname{Im} T_1 = H_2 \ (T_1 \text{ surjective}),\\
0 \longrightarrow H_1 \xrightarrow{T_1} H_2 \longrightarrow 0 \text{ exact} \quad &\Leftrightarrow \quad T_1 : H_1 \cong H_2 \text{ is an isomorphism},\\
0 \longrightarrow H_1 \xrightarrow{T_1} H_2 \xrightarrow{T_2} H_3 \longrightarrow 0 \text{ exact} \quad &\Leftrightarrow \quad \begin{cases} T_1 : H_1 \cong T_1(H_1) \text{ and} \\ T_2 : \frac{H_2}{T_1(H_1)} \cong H_3 \,. \end{cases}
\end{aligned}$$

In the last case, $H_2 \cong H_1 \oplus H_3$, but not canonically so.

REMARK 1.8. By definition, the sequence

$$0 \longrightarrow \operatorname{Ker} T_1 \hookrightarrow H_1 \xrightarrow{T_1} H_2 \longrightarrow \operatorname{Coker} T_1 \longrightarrow 0$$

of linear maps between vector spaces is exact. If H_1, H_2 are Hilbert spaces and T_1 bounded, the sequence becomes an exact sequence of bounded mappings between Hilbert spaces if and only if T_1 has closed range (see also the Hint to Exercise 2.1).

The following theorem is a key device in Homological Algebra for various kinds of decomposition and additivity theorems, see also Remark 1.11.

THEOREM 1.9 (Snake Lemma). *Assume that the following diagram of vector spaces and linear maps is commutative (i.e., $jF = F'i$ and $qF' = F''p$) with exact horizontal sequences and Fredholm operators for vertical maps.*

$$\begin{array}{ccccccccc}
0 & \to & H_1 & \xrightarrow{i} & H_1' & \xrightarrow{p} & H_1'' & \to & 0\\
 & & \downarrow F & & \downarrow F' & & \downarrow F'' & & \\
0 & \to & H_2 & \xrightarrow{j} & H_2' & \xrightarrow{q} & H_2'' & \to & 0.
\end{array}$$

Then we have

$$\operatorname{index} F - \operatorname{index} F' + \operatorname{index} F'' = 0.$$

PROOF. We do the proof in two parts.
1. Here we show that the following sequence is exact:

$$\begin{aligned}
&0 \longrightarrow \operatorname{Ker} F \longrightarrow \operatorname{Ker} F' \longrightarrow \operatorname{Ker} F'' \longrightarrow \\
&\quad \longrightarrow \operatorname{Coker} F \longrightarrow \operatorname{Coker} F' \longrightarrow \operatorname{Coker} F'' \longrightarrow 0.
\end{aligned} \tag{1.1}$$

For this, we first explain how the individual maps are defined. By the commutativity of the diagram, the maps $\operatorname{Ker} F \to \operatorname{Ker} F'$ and $\operatorname{Ker} F' \to \operatorname{Ker} F''$ are given by i and p; and the maps $\operatorname{Coker} F \to \operatorname{Coker} F'$ and $\operatorname{Coker} F' \to \operatorname{Coker} F''$ are induced by j and q in the natural way (please check).

Also $\operatorname{Ker} F'' \to \operatorname{Coker} F$ is well defined: Let $u'' \in \operatorname{Ker} F''$; i.e., $u'' \in H''$ and $F''u'' = 0$. Since p is surjective we can choose $u' \in H_1'$ with $pu' = u''$. Then $F'u' \in \operatorname{Ker} q$, since $qF'u' = F''pu' = F''u'' = 0$. By exactness, we have $\operatorname{Ker} q = \operatorname{Im} j$ and a unique (by the injectivity of j) element $u \in H_2$ with $ju = F'u'$. We map $u'' \in \operatorname{Ker} F''$ to the class of u in $H_2 / \operatorname{Im} F = \operatorname{Coker} F$. It remains to show that we get the same class for another choice of u'. Thus, let $\widetilde{u}' \in H'$ be such that $p\widetilde{u}' = u''$ (p is in general not injective; hence, possibly $\widetilde{u}' \neq u'$). As above, we have a $\widetilde{u} \in H_2$ with $j\widetilde{u} = F'u'$. We must now find $u_0 \in H_1$ with $Fu_0 = u - \widetilde{u}$. Then we are done. For this, note that $j(u - \widetilde{u}) = ju - j\widetilde{u} = F'u' - F'\widetilde{u}' = F'(u' - \widetilde{u}')$. Since $pu' = p\widetilde{u}' = u''$, we have $u' - \widetilde{u}' \in \operatorname{Ker} p$. By exactness, we have $u_0 \in H_1$ with $iu_0 = u' - \widetilde{u}'$, whence $F'iu_0 = F'(u' - \widetilde{u}')$. The left side is jFu_0 and the right side is $j(u - \widetilde{u})$ from above. Hence, $Fu_0 = u - \widetilde{u}$ by the injectivity of j, as desired. We introduced the map $\operatorname{Ker} F'' \to \operatorname{Coker} F$ in such great detail in order to demonstrate what is typical for diagram chasing: it is straightforward, largely independent of tricks and ideas, readily reproduced, and hence somewhat monotonous. Therefore we will forgo showing the exactness of (1.1) except at $\operatorname{Ker} F'$, and leave the rest as an exercise. The exactness of $\operatorname{Ker} F \xrightarrow{\tilde{\imath}} \operatorname{Ker} F' \xrightarrow{\tilde{p}} \operatorname{Ker} F''$, where $\tilde{\imath}$ and $\tilde{p}$ are the restrictions of i and p, means $\operatorname{Im} \tilde{\imath} = \operatorname{Ker} \tilde{p}$. Thus, we have two inclusions to show:

$$\subseteq\text{: This is clear, since } p \circ i = 0 \text{ implies } \tilde{p} \circ \tilde{\imath} = 0.$$
$$\supseteq\text{: If } u' \in \operatorname{Ker} \tilde{p}, \text{ then } u' \in \operatorname{Ker} F' \text{ and } pu' = 0, \text{ and}$$

by the exactness of $H_1 \to H_1' \to H_1''$, we have $u \in H_1$ with $iu = u'$. It remains to show that $Fu = 0$, but this is clear, since $jFu = F'iu = F'u' = 0$ and j is injective.

Note. Before we go to part 2 of the proof, we pause for a moment: It is interesting that for $H' = H \oplus H''$, $i = j$ the inclusion, and $p = q$ the projection, we recapture the addition formula of Exercise 1.5. In this case, the exact sequence (1.1) then breaks, as shown there, into two parts

$$0 \to \operatorname{Ker} F \to \operatorname{Ker} F' \to \operatorname{Ker} F'' \to 0 \text{ and}$$
$$0 \to \operatorname{Coker} F \to \operatorname{Coker} F' \to \operatorname{Coker} F'' \to 0.$$

In the general case, however, we no longer have

$$\dim \operatorname{Ker} F' = \dim \operatorname{Ker} F + \dim \operatorname{Ker} F'' \text{ and}$$
$$\dim \operatorname{Coker} F' = \dim \operatorname{Coker} F + \dim \operatorname{Coker} F'',$$

but instead we must consider the interaction ($\operatorname{Ker} F'' \to \operatorname{Coker} F$). From a topological standpoint, the concept of the index of a Fredholm operator is a special case of the general concept of the Euler characteristic $\chi(C)$ of a complex

$$C: \; \xrightarrow{T_{k+1}} C_k \xrightarrow{T_k} C_{k-1} \xrightarrow{T_{k-1}} C_{k-2} \xrightarrow{T_{k-2}} C_{k-3} \xrightarrow{T_{k-3}} \cdots$$

of vector spaces and linear maps (with $T_k \circ T_{k+1} = 0$) with finite **Betti numbers**

$$\beta_k = \dim \left(\frac{\operatorname{Ker} T_k}{\operatorname{Im} T_{k+1}} \right).$$

Here $\operatorname{Ker} T_k / \operatorname{Im} T_{k+1}$ is called the k-th **homology space** $H_k(C)$. Assuming that all these numbers are finite, as well as the number of nonzero Betti numbers, we define

$$\chi(C) := \sum_k (-1)^k \beta_k$$

whence $\operatorname{index} F = \chi(C)$ for

$$C : 0 \longrightarrow 0 \longrightarrow H \xrightarrow{F} H \longrightarrow 0 \longrightarrow 0,$$

where $C_2 = C_1 = H$ and $C_i = 0$ otherwise; thus, $H_2(C) = \operatorname{Ker} F$ and $H_1(C) = \operatorname{Coker} F$. This is the reason why we can follow in the proof of Theorem 1.9 the well-known topological arguments (see [**Gr**, p.100f] or [**ES**, p.52f]) yielding the *addition (or pasting) theorem*

$$\chi(C) - \chi(C') + \chi(C'/C) = 0.$$

See also Section 14.4. In particular, (1.1) is only a special case of the *long exact homology sequence*

$$\begin{aligned} &\to H_{k+1}(C'/C) \to H_k(C) \to H_k(C') \\ &\to H_k(C'/C) \to H_{k-1}(C) \to H_{k-1}(C') \to, \end{aligned} \tag{1.2}$$

[**Gr**, p.57f] or [**ES**, p.125-128]. This more general description would have the advantage that we only need to prove exactness of (1.2) at three adjacent places with an argument independent of k rather than at six places as in our *simplified* approach where we restricted ourselves to complexes of length two. But back to our proof:
2. For each exact sequence

$$0 \to A_1 \to A_2 \to \cdots \to A_r \to 0, \tag{1.3}$$

of finite-dimensional vector spaces, we wish to derive the formula

$$\sum_{k=1}^{r} (-1)^k \dim A_k = 0.$$

Notice first that for r sufficiently large ($r > 3$), the formula for the alternating sum for (1.3) follows, once we know the formula holds for the exact (prove!) sequences

$$0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \operatorname{Im}(A_1 \to A_2) \longrightarrow 0$$

and

$$0 \longrightarrow \operatorname{Im}(A_2 \to A_3) \longrightarrow A_3 \longrightarrow \cdots \longrightarrow A_r \longrightarrow 0.$$

Since these sequences have length less than r, the formula is proved by induction, if we verify it for $r = 1, 2, 3$.

$$\begin{aligned} &r = 1 : \text{trivial, since then } A_1 \cong 0. \\ &r = 2 : \text{also clear, since then } A_1 \cong A_2. \\ &r = 3 : \text{clear, since } 0 \to A_1 \to A_2 \to A_3 \to 0 \text{ implies } A_3 \cong A_2/A_1, \\ &\qquad \text{whence } \dim A_3 = \dim A_2 - \dim A_1. \end{aligned} \tag{1.4}$$

Thus, part 2 is finished and combining it with part 1, the Snake Lemma is proved. Actually, we have proven much more, namely, whenever two of the three maps F, F', F'' have a finite index, the third has finite index given by the snake formula. □

Exercise 1.10. Combine Theorem 1.9 and Exercise 1.6, to show that

$$\operatorname{index} G \circ F = \operatorname{index} F + \operatorname{index} G. \tag{1.5}$$

[Hint: Consider the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & H & \xrightarrow{i} & H \oplus H' & \xrightarrow{p} & H' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle F} & & \downarrow{\scriptstyle G\circ F \oplus \mathrm{Id}} & & \downarrow{\scriptstyle G} & & \\ 0 & \longrightarrow & H' & \xrightarrow{j} & H'' \oplus H' & \xrightarrow{q} & H'' & \longrightarrow & 0 \end{array}$$

where $iu := (u, Fu)$, $jv := (Gv, v)$, $p(u,v) := Fu - v$, and $q(w,v) := w - Gv$.]

Remark 1.11. Alternative proofs of the product formula (1.5) can be found in many places. They may appear shorter. Arguing via the Snake Lemma is more lengthy, but it puts the product formula in the *correct* format of a topological composition or gluing formula

$$\tau(\Phi_1 \cup \Phi_2) = \tau(\Phi_1) * \tau(\Phi_2) * \varepsilon(\Phi_1 \cap \Phi_2), \tag{1.6}$$

where we have in (1.5) the vanishing of the typical third term on the right side, the error term $\varepsilon(\Phi_1 \cap \Phi_2)$, for $\tau := \operatorname{index}; \Phi_1, \Phi_2 \in \mathcal{F}; \cup := \circ$; and $* := +$.

A stunning impression of the intricacies of simple looking product formulas may be gained by checking the proof of the corresponding product formula for the index of closed (not necessarily bounded) densely defined Fredholm operators, see Theorem 2.45, p.43f.

4. Operators of Finite Rank and the Fredholm Integral Equation

Exercise 1.12. Show that for any operator $K : H \to H$ of **finite rank** (i.e., $\dim K(H) < \infty$), the sum $\mathrm{Id} + K$ is a Fredholm operator and

$$\operatorname{index}(\mathrm{Id} + K) = 0$$

Here, $\mathrm{Id} : H \to H$ is the identity.

[Hint: Set $h := \operatorname{Im} K$ and recall Theorem 1.9 for the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & h & \xrightarrow{i} & H & \xrightarrow{p} & H/h & \longrightarrow & 0 \\ & & \Big\downarrow (\mathrm{Id}+K)_h & & \Big\downarrow \mathrm{Id}+K & & \Big\downarrow (\mathrm{Id}+K)_{H/h} & & \\ 0 & \longrightarrow & h & \xrightarrow{j} & H & \xrightarrow{q} & H/h & \longrightarrow & 0. \end{array}$$

To see that the vertical maps are well defined, we only need $(\mathrm{Id} + K)(h) \subseteq h$. The commutativity of the diagram and the exactness of the rows are clear. Since one can show $\operatorname{Ker}(\mathrm{Id} + K) \subseteq h$ and $\dim \operatorname{Coker}(\mathrm{Id} + K) \leq \dim h$, the Snake Formula gives us the result once we show

$$\operatorname{index}(\mathrm{Id} + K)_h = 0 \tag{1.7}$$

and

$$\operatorname{index}(\mathrm{Id} + K)_{H/h} = 0. \tag{1.8}$$

But (1.7) is clear from Exercise 1.4 and (1.8) is trivial because $(\mathrm{Id} + K)_{H/h} = \mathrm{Id}_{H/h}$. Note that one could deduce that $\mathrm{Id} + K$ has finite index by using the observation at the end of the proof of Theorem 1.9.]

REMARK 1.13. One may be bothered by the way in which the proposed solution produces the result so directly from the Snake Formula by means of a trick. As a matter of fact, $\operatorname{index}(\operatorname{Id}+K)$ can be computed in a pedestrian fashion by reduction to a system of n linear equations with n unknowns where $n := \dim h$. To do this one verifies that every operator K of finite rank has the form

$$Ku = \sum_{i=1}^{n} \langle u, u_i \rangle \, v_i$$

with fixed $u_1, ..., u_n, v_1, ..., v_n \in H$ (note that every continuous linear functional is of the form $\langle \cdot, u_0 \rangle$). Whether this *direct* approach, as detailed for example in [**Sche**, Theorem 4.9](see also [**Ped**, p.110f]), is in fact more transparent than the device used with the Snake Formula, depends a little on the perspective. While in the first approach the key point (namely the use of Exercise 1.4 for equation (1.7)) is singled out and separated clearly in the remaining formal argument, we find in the second more constructive approach rather a fusion of the nucleus with its packaging. However, the use of the fairly non-trivial Riesz-Fischer Lemma is unnecessary in the case where K is given in the desired explicit form, as in the following example.

EXERCISE 1.14. Consider the **Fredholm integral equation of the second kind**

$$u(x) + \int_a^b G(x,y) u(y) \, dy = h(x)$$

with degenerate (product-) weight function (or *integral kernel*)

$$G(x,y) = \sum_{i=1}^{n} f_i(x) \overline{g_i(y)}$$

with fixed $a < b$ real and f_i, g_i square integrable on $[a,b]$. Prove the **Fredholm alternative**: Either there is a unique solution $u \in L^2[a,b]$ for every given right side $h \in L^2[a,b]$, or the homogeneous equation ($h = 0$) has a solution which does not vanish identically. Moreover, the number of linearly independent solutions of the homogeneous equation equals the number of linear conditions one needs to impose on h in order that the inhomogeneous equation be solvable.
[Hint: Consider the operator $\operatorname{Id}+K$ on the Hilbert space $L^2[a,b]$, where $Ku = \sum \langle u, g_i \rangle f_i$, and apply Exercise 1.12. For the interpretation of the dimension of the cokernel, see Exercise 2.1b below.]

5. The Spectra of Bounded Linear Operators: Basic Concepts

We close this chapter with a concept that belongs to the border region between algebraic and analytic notions, the **spectrum** of a bounded linear operator. For the corresponding definitions and elementary properties in the more general (and, actually, different) case of not necessarily bounded linear operators we refer to Definition 2.59 and Exercise 2.60, p.50.

REMARK 1.15. The following comments correspond to the sets of Table 1.1. For wanted arguments we refer to [**Ped**, Section 4.1]:
1. $\operatorname{Res}(T)$ is open in $\mathbb{C}$. We can argue as in Exercise 3.6, p.65 where one has to prove that the group of units in any Banach algebra is open.
2. $\operatorname{Spec}(T)$ is closed and bounded (i.e., compact) in $\mathbb{C}$. Actually, one has $\operatorname{Spec}(T) \subset$

$\{\lambda : |\lambda| \leq \|T\|\}$.
3. $\mathrm{Res}\,(T) \subseteq \mathrm{Fred}\,(T)$, and $\mathrm{Fred}\,(T)$ is the union of at most a countable number of open, connected components.
4. $\mathrm{Spec}_e\,(T) \subseteq \mathrm{Spec}\,(T)$, and $\mathrm{Spec}_e\,(T) = \mathrm{Spec}\,(\pi\,(T))$, where $\pi : \mathcal{B} \to \mathcal{B}/\mathcal{K}$ is the projection.
5. $\mathrm{Spec}_p\,(T)$ consists of isolated points of $\mathrm{Spec}\,(T)$. $\mathrm{Spec}_p\,(T)$ contains the limit points of $\mathrm{Spec}\,(T)$ which are Fredholm points.
6. $\mathrm{Spec}_c\,(T) \subseteq \mathrm{Spec}_e\,(T)$.
7. $\mathrm{Spec}\,(T) = \mathrm{Spec}_p\,(T) \cup \mathrm{Spec}_c\,(T) \cup \mathrm{Spec}_r\,(T)$, a union of disjoint sets.

TABLE 1.1. The spectra of bounded linear operators $T : H \to H$

	Symbol	Name	Definition
1.	$\mathrm{Res}\,(T)$	resolvent set	$\left\{z \in \mathbb{C} : (T - z\,\mathrm{Id})^{-1} \in \mathcal{B}\right\}$
2.	$\mathrm{Spec}\,(T)$	spectrum	$\mathbb{C} \setminus \mathrm{Res}\,(T)$
3.	$\mathrm{Fred}\,(T)$	Fredholm points	$\{z \in \mathbb{C} : T - z\,\mathrm{Id} \in \mathcal{F}\}$
4.	$\mathrm{Spec}_e\,(T)$	essential spectrum	$\mathbb{C} \setminus \mathrm{Fred}\,(T)$
5.	$\mathrm{Spec}_p\,(T)$	point spectrum (eigenvalues)	$\{z \in \mathbb{C} : \mathrm{Ker}\,(T - z\,\mathrm{Id}) \neq \{0\}\}$
6.	$\mathrm{Spec}_c\,(T)$	continuous spectrum	$\left\{ z \in \mathbb{C} : \left\{ \begin{array}{l} \mathrm{Ker}\,(T - z\,\mathrm{Id}) = \{0\}, \\ \mathrm{Im}\,(T - z\,\mathrm{Id}) \subsetneq H, \text{ and} \\ \overline{\mathrm{Im}\,(T - z\,\mathrm{Id})} = H \end{array} \right. \right\}$
7.	$\mathrm{Spec}_r\,(T)$	residual spectrum	$\left\{ z \in \mathbb{C} : \left\{ \begin{array}{l} \mathrm{Ker}\,(T - z\,\mathrm{Id}) = \{0\}, \\ \mathrm{codim}\left(\overline{\mathrm{Im}\,(T - z\,\mathrm{Id})}\right) > 0 \end{array} \right. \right\}$

EXAMPLE 1.16. (a) $T := \mathrm{shift}^{+}$ (see [**Jö**, 1970/1982, 5.3]) $\Rightarrow$

$$\begin{array}{l} \mathrm{Spec}_p\,(T) = \emptyset, \\ \mathrm{Spec}_r\,(T) = \{z \in \mathbb{C} : |z| < 1\} \text{ and} \\ \mathrm{Spec}_c\,(T) = \mathrm{Spec}_e\,(T) = \{z \in \mathbb{C} : |z| = 1\}. \end{array}$$

(b) T compact and self-adjoint (see [**Jö**, 1970/1982, 6.2]) $\Rightarrow$

$$\begin{array}{l} \mathrm{Spec}\,(T) \subseteq \mathbb{R}, \\ \mathrm{Spec}_p\,(T) = \{\lambda_n : n = 1, 2, \ldots\}, \\ \mathrm{Spec}_r\,(T) = \{0\} \text{ and} \\ \mathrm{Spec}_e\,(T) = \{0\}. \end{array}$$

(c) $T =$ Fourier transformation on $L^2\,(\mathbb{R})$ (see [**DM**, 1972, p.98]) $\Rightarrow$

$$\mathrm{Spec}\,(T) = \mathrm{Spec}_e\,(T) = \mathrm{Spec}_p\,(T) = \{1, i, -1, -i\}.$$

CHAPTER 2

Analytic Methods. Compact Operators

Synopsis. Adjoint and Self-Adjoint Operators - Recalling Fischer-Riesz. Dual characterization of Fredholm operators: either by finite-dimensional kernel and cokernel, or by finite-dimensional kernels of the operator and its adjoint operator *and* closed image. Compact Operators: Spectral Decomposition, Why Compact Operators also are Called *Completely Continuous*, $\mathcal{K}$ as Two-Sided Ideal, Closure of Finite-Rank Operators, and Invariant under *. Classical Integral Operators. Fredholm Alternative and Riesz Lemma. Sturm-Liouville Boundary Value Problems. Unbounded Operators: Comprehensive Study of Linear First Order Differential Operators Over S^1: Sobolev Space, Dirac Distribution, Normalized Integration Operator as *Parametrix*, The Index Theorem on the Circle for Systems. Closed Operators, Closed Extensions, Closed (not necessarily bounded) Fredholm Operators, Composition Rule, Symmetric and Self–Adjoint Operators, Formally Self-Adjoint and Essentially Self-Adjoint. Spectral Theory. Metrics on the Space of Closed Operators. Trace Class and Hilbert-Schmidt Operators

1. Analytic Methods. The Adjoint Operator

With Exercises 1.12 and 1.14, we have reached the limits of our so far purely algebraic reasoning where we could reduce everything to the elementary theory of solutions of n linear equations in n unknowns. In fact, the limit process $n \to \infty$ marks the emergence of functional analysis which went beyond the methods of linear algebra while being motivated by its questions and results. This occurred mainly in the study of integral equations.

In 1927, Ernst Hellinger and Otto Toeplitz stressed in their article *Integral equations and equations in infinitely many unknowns* in the *Enzyklopädie der mathematischen Wissenschaften* that "the essence of the theory of integral equations rests in the analogy with analytic geometry and more generally in the passage from facts of algebra to facts of analysis" [**HT**, p.1343]. They showed in a concise historical survey how the awareness of these connections progressed in the centuries since Daniel Bernoulli investigated the oscillating string as a limit case of a system of n mass points:

> "Orsus itaque sum has meditationes a corporibus duobus filo flexili in data distantia cohaerentibus; postea tria consideravi moxque quatuor, et tandem numerum eorum distantiasque qualescunque; cumque numerum corporum infinitum facerem, vidi demum naturam oscillantis catenae sive aequalis sive inaequalis crassitiei sed ubique perfecte flexilis."[1]

The passage to the limit means for the Fredholm integral equation of Exercise 1.14 that more general nondegenerate weight functions $G(x, y)$ are allowed which then can be

[1]Petropol Comm. 6 (1732/33, ed. 1738), 108-122. Our translation: "In these considerations I started with two bodies at a fixed distance and connected by an elastic string; next I considered three then four and finally an arbitrary number with arbitrary distances between them; but only when I made the number of bodies infinite did I fully comprehend the nature of an oscillating elastic chain of equal or varying thickness."

approximated by degenerate weights (e.g., polynomials). Such weights play a prominent role in applications, especially when dealing with differential equations with boundary conditions, as we will see below.

For the theory of Fredholm operators developed here we must analogously abandon the notion of an operator of finite rank and generalize it (to compact operator, see below) whereby topological, i.e. continuity considerations, become essential in connection with the limit process. This will bring out the full force of the concept of Fredholm operator. We now turn to this topic.

Closed Image of Fredholm Operators. To begin with, the reader should derive the *topological* closedness of the image of a Fredholm operator from the purely *algebraic* property of finite codimension of the image.

EXERCISE 2.1. For $F : H \to H$ a Fredholm operator, prove:
a) $\operatorname{Im} F$ is closed.
b) There is an explicit criterion for deciding when an element of H lies in $\operatorname{Im} F$. Namely, let $n = \dim \operatorname{Coker} F$; then there are $u_1, ..., u_n \in H$, such that for all $w \in H$, we have

$$w \in \operatorname{Im} F \Leftrightarrow \langle w, u_1 \rangle = \cdots = \langle w, u_n \rangle = 0.$$

[Hint for a): As a vector subspace of H, naturally $\operatorname{Im} F$ is closed under addition and multiplication by complex numbers. However, here we are interested in topological closure, namely that in passing to limit points we do not leave $\operatorname{Im} F$. This has far-reaching consequences, since only closed subspaces inherit the completeness property of the ambient Hilbert space, and hence have orthonormal bases (= complete orthonormal vector systems, see e.g., [**Ped**, p.83] or [**Sche**, p.31 and Lemma 11.9]). The trivial fact that finite-dimensional subspaces are closed can be exploited: Since $\operatorname{Coker} F = H/\operatorname{Im} F$ has finite dimension, we can find $v_1, ..., v_n \in H$ whose classes in $H/\operatorname{Im} F$ form a basis. The linear span h of $v_1, ..., v_n$ is then an algebraic complement of $\operatorname{Im} F$ in H. Consider the map $\Phi : H \oplus h \to H$ with $F'(u, v) := Fu + v$. Since Φ is linear, surjective, and (by the boundedness of F) continuous, we have that Φ is open (according to the *open mapping principle*; e.g., see [**Ped**, p.53]). It follows that $H \setminus F(H) = \Phi(H \oplus h \setminus H \oplus \{0\})$ is open.
[Hint for b): As a closed subspace of H, $\operatorname{Im} F$ is itself a Hilbert space possessing a countable orthonormal basis $w_1, w_2, \ldots$. Now set $u_i := v_i - Pv_i$, where v_i are as above $(i = 1, ..., n)$ and $P : H \to \operatorname{Im} F$ is the projection

$$Pu := \sum_{j=1}^{\infty} \langle u, w_j \rangle \, w_j .$$

Then $\{u_1, ..., u_n\}$ forms a basis for $(\operatorname{Im} F)^{\perp}$, the orthogonal complement of $\operatorname{Im} F$ in H.
For aesthetic reasons, one can orthonormalize $u_1, ..., u_n$ by the Gram-Schmidt process (i.e., without loss of generality, assume they are orthonormal). Then $u_1, ..., u_n$, $w_1, w_2, w_3, \ldots$ is a countable orthonormal basis for H.]

REMARK 2.2. Be aware that in spite of the strength of the open-mapping argument, it can not be applied to show that *any* subspace W of finite codimension $\dim H/W < \infty$ is closed. Of course, one could once again construct a bounded surjective operator $\Phi : H \oplus W \to H$, say by $\Phi(u, v) := u + v$. But, in general, $H \setminus W$ can not be obtained as the image of an open subset of $H \oplus W$ by applying

Φ. Certainly $H \setminus W \neq \Phi(H \oplus W \setminus H \oplus \{0\})$. Actually, the kernel $\operatorname{Ker}(f)$ of any unbounded linear functional provides a counterexample. It is a space of codimension 1, but it is not closed since closed $\operatorname{Ker}(f)$ would imply continuity of f in 0 and hence everywhere.

EXERCISE 2.3. Once more, prove the chain rule

$$\operatorname{index} G \circ F = \operatorname{index} F + \operatorname{index} G$$

for Fredholm operators $F : H \to H'$ and $G : H' \to H''$.
[Hint: In place of the purely algebraic argument in Exercise 1.10, use Exercise 2.1a to first prove that the images are closed, and then use the technique of orthogonal complements in Exercise 2.1b.]

How trivial or nontrivial is it to prove that operators have closed ranges? For operators with finite rank and for surjective operators it is trivial, and for Fredholm operators it was proved in Exercise 2.1a. Is it perhaps true that all bounded linear operators have closed images? As the following counterexample explicitly shows, the answer is *no*. Moreover, we will see below that all compact operators with infinite-dimensional image are counterexamples.

EXERCISE 2.4. For a Hilbert space H with orthonormal system $e_1, e_2, e_3, \ldots$, consider the contraction operator

$$Au := \sum_{j=1}^{\infty} \frac{1}{j} \langle u, e_j \rangle \, e_j .$$

Show that $\operatorname{Im} A$ is not closed.
[Hint: Clearly A is linear and bounded ($\|A\| = ?$), and furthermore, we have the criterion for $\operatorname{Im}(A)$

$$v \in \operatorname{Im} A \Leftrightarrow \sum_{j=1}^{\infty} j \, \langle v, e_j \rangle \, e_j \in H \Leftrightarrow \sum_{j=1}^{\infty} j^2 \, |\langle v, e_j \rangle|^2 < \infty .$$

It follows that for

$$v_0 := \sum_{j=1}^{\infty} \frac{1}{j\sqrt{j}} e_j$$

and

$$v_n := \sum_{j=1}^{\infty} \frac{1}{j\sqrt{j} j^{1/n}} e_j ; \quad n = 1, 2, \ldots ;$$

we get $v_0 \in H \setminus \operatorname{Im} A$ and $v_n \in \operatorname{Im} A$. (The old trick: $\sum 1/j^a$ converges for $a > 1$ (e.g., for $a = 1 + 1/n$), but diverges for $a = 1$.) To finally prove that the sequence actually converges to v_0, observe that

$$\|v_0 - v_n\|^2 = \sum_{j=1}^{\infty} \frac{1}{j^2} \frac{\left(j^{1/n} - 1\right)^2}{j \; j^{2/n}} .$$

It is clear, that for each j

$$\frac{j^{1/n} - 1}{j^{1/n}} \to 0 \quad \text{as} \quad n \to \infty,$$

and then in particular $\langle v_0 - v_n, e_j \rangle \to 0$. However, to show that $\|v_0 - v_n\|^2$ converges to 0 as $n \to \infty$, one must estimate more precisely. To do this, we exploit the fact that $\sum_{j>j_0}^{\infty} j^{-3}$ can be made smaller than any $\varepsilon > 0$ for j_0 sufficiently large, while on the other hand, choosing n sufficiently large (so large that $(1+\varepsilon)^n \geq j_0$), we have

$$\frac{j^{1/n} - 1}{j^{1/n}} < \varepsilon \text{ for } j \leq j_0.]$$

The Adjoint Operator. We will draw further conclusions from the closure of the image of a Fredholm operator and to do this we introduce adjoint operators. The purpose is to eliminate the asymmetry between kernel and cokernel or, in other words, between the theory of the homogeneous equation (questions of uniqueness of solutions) and the theory of the inhomogeneous equation (questions of existence of solutions). This is achieved by representing the cokernel of an operator as the kernel of a suitable adjoint operator.

Projective geometry deals with a comparable problem via duality: one thinks of space on the one hand as consisting of points, on the other as consisting of planes, and depending on the point of view, a straight line is the join of two points or the intersection of two planes. Analytic geometry passes from a matrix (a_{ij}) to its transpose (a_{ji}) or $(\overline{a}_{ji})$ in the complex case to technically deal with dual statements. We can do the same successfully for operators (= infinite matrices). The basic tool is the following Representation Theorem with nice proofs in [**Ped**, Proposition 3.1.9] or [**Sche**, Theorem 2.1]. Originally, it was proven independently by Frigyes Riesz and Ernst Sigismund Fischer only for $H := L^2([a,b])$.

THEOREM 2.5 (E. Fischer, F. Riesz, 1907). *Let H be a separable complex Hilbert space. To each continuous linear mapping $\varphi : H \to \mathbb{C}$ (called* functional*) there exists a unique element $u \in H$ such that $\varphi v = \langle v, u \rangle$ for all $v \in H$.*

EXERCISE 2.6. Show that on the space $\mathcal{B}(H)$ of bounded linear operators of a Hilbert space H, there is a natural isometric (anti-linear) involution

$$* : \mathcal{B}(H) \to \mathcal{B}(H)$$

which assigns to each $T \in \mathcal{B}(H)$ the **adjoint** operator $T^* \in \mathcal{B}(H)$ such that for all $u, v \in H$

$$\langle u, T^* v \rangle = \langle Tu, v \rangle \, .$$

[Hint: It is clear that T^*v is well defined for each $v \in H$, since $u \mapsto \langle Tu, v \rangle$ is a continuous linear functional on H; and so, by the preceding theorem, you can express the functional through a unique element of H, which you may denote by T^*v. The linearity of T^* is clear by construction. While proving the continuity (i.e., boundedness) of T^*, show more precisely that $\|T^*\| = \|T\|$ (i.e., that $T \mapsto T^*$ is an isometry).]

Just as easily, we have the involution property $T^{**} = T$, the composition rule $(T \circ R)^* = R^* \circ T^*$, and conjugate-linearity $(aT + bR)^* = \bar{a}T^* + \bar{b}R^*$, where the bars denote complex conjugation. Details can be found for example in [**Ped**, Theorem 3.2.3], [**Sche**, Sections 3.2 and 11.2]. Observe that for $T \in \mathcal{B}(H, H')$, where H and H' may differ, the adjoint operator T^* is in $\mathcal{B}(H', H)$.

THEOREM 2.7. *For $F \in \mathcal{F}$, the $u_1, ..., u_n$ in* Exercise 2.1b *form a basis of* $\operatorname{Ker} F^*$*, whence*

$$\operatorname{Im} F = (\operatorname{Ker} F^*)^{\perp} \text{ and } \operatorname{Coker}(F) = \operatorname{Ker}(F^*) \, .$$

PROOF. First we note that $u \in (\operatorname{Im} F)^\perp$ exactly when

$$0 = \langle u, w \rangle = \langle u, Fv \rangle = \langle F^* u, v \rangle ,$$

for all $w \in \operatorname{Im}(F)$ (i.e., for all $v \in H$); thus, $(\operatorname{Im} F)^\perp = \operatorname{Ker} F^*$. By again taking orthogonal complements, we have $(\operatorname{Ker} F^*)^\perp = (\operatorname{Im} F)^{\perp\perp} = \operatorname{Im} F$, since $\operatorname{Im} F$ is closed by Exercise 2.1a. □

Observe that the above argument remains valid for any bounded linear operator with closed range. In this case, we have the criterion that the equation $Fv = w$ is solvable exactly when $w \perp \operatorname{Ker} F^*$. This is the basic

LEMMA 2.8 (Polar Lemma). *Let H, H' be Hilbert spaces and $T \in \mathcal{B}(H, H')$. Then* $(\operatorname{Im} T)^\perp = \operatorname{Ker} T^*$. *Moreover, if* $\operatorname{Im} T$ *is closed, we have* $\operatorname{Im} T = (\operatorname{Ker} T^*)^\perp$.

In the language of *categories* and *functors* (see e.g. [**Brd**, p.176]) this can be reformulated in the following, a bit exaggerated way:

THEOREM 2.9. *The functor $H \mapsto H$, $H' \mapsto H'$, $\mathcal{B}(H, H') \ni T \mapsto T^* \in \mathcal{B}(H', H)$ is a contravariant functor on the category of (separable) Hilbert spaces and bounded operators. It preserves the norm and is exact; i.e., if the sequence*

$$H \xrightarrow{T} H' \xrightarrow{S} H'' \xrightarrow{R}$$

is exact, then the sequence

$$\xrightarrow{R^*} H'' \xrightarrow{S^*} H' \xrightarrow{T^*} H$$

is also exact.

PROOF. We show the exactness only at H', i.e., $\operatorname{Im} S^* = \operatorname{Ker} T^*$. Since $ST = 0$, we have $T^* S^* = 0$, hence $\operatorname{Im} S^*$ is contained in $\operatorname{Ker} T^*$.

To show the opposite inclusion, we notice that $\operatorname{Im} T = \operatorname{Ker} S$ is closed in H' and $\operatorname{Im} S = \operatorname{Ker} R$ is closed in H'' . So we have a decomposition of

$$H' = \operatorname{Ker} T^* \oplus \operatorname{Ker} S = (\operatorname{Im} T)^\perp \oplus \operatorname{Im} T \tag{2.1}$$

and

$$H'' = \operatorname{Im} S \oplus (\operatorname{Im} S)^\perp \tag{2.2}$$

into pairs of mutually orthogonal closed subspaces. Notice also that

$$S|_{\operatorname{Ker} T^*} : \operatorname{Ker} T^* \longrightarrow \operatorname{Im} S \tag{2.3}$$

is bounded, injective and surjective, hence its inverse is also bounded (though not necessarily a Hilbert space isomorphism, i.e. not necessarily unitary).

Now let $y \in H'$ with $y \in \operatorname{Ker} T^*$, i.e., $\langle y, y' \rangle = 0$ for all $y' \in T(H)$. We consider the mapping

$$\begin{array}{cccc} [y]: & H'' & \longrightarrow & \mathbb{C} \\ & Sy' + z' & \mapsto & \langle y', y \rangle \end{array}$$

where the splitting on the left side is according to (2.2) and the inner product on the right side is taken in the Hilbert space H'. By construction, the mapping $[y]$ is linear and vanishes on the second factor of H'' . On the first factor it is continuous because of the homeomorphism of (2.3). Hence the functional $[y]$ can be represented by an element of the Hilbert space H'' which we also will denote by $[y]$. So we have

$$\langle y' + y'', S^*([y]) \rangle = \langle Sy' + Sy'', [y] \rangle = \langle Sy', [y] \rangle = \langle y', y \rangle = \langle y' + y'', y \rangle \tag{2.4}$$

for all elements $y' + y'' \in H'$ with $y' \in (\operatorname{Im} T)^\perp$ and $y'' \in \operatorname{Im} T$ according to the decomposition (2.1). Note that the second and the third inner product in (2.4) are taken in H'' and the other inner products in H'. Equation (2.4) shows that $y = S^*([y])$. Thus $\operatorname{Ker} T^* \subseteq \operatorname{Im} S^*$, and $\operatorname{Ker} T^* = \operatorname{Im} S^*$, as desired. □

THEOREM 2.10. *A bounded linear operator F is a Fredholm operator, precisely when* $\operatorname{Ker} F$ *and* $\operatorname{Ker} F^*$ *are finite-dimensional and* $\operatorname{Im} F$ *is closed. In this case,*

$$\operatorname{index} F = \dim \operatorname{Ker} F - \dim \operatorname{Ker} F^*.$$

Thus, in particular, $\operatorname{index} F = 0$ *in case F is self-adjoint* (*i.e.,* $F^* = F$).

PROOF. Use Theorem 2.7 and Exercise 2.1a. □

REMARK 2.11. a) Note that $\operatorname{Ker} F^*F = \operatorname{Ker} F$: "$\supseteq$" is clear; for "$\subseteq$", take $u \in \operatorname{Ker} F^*F$, and then

$$\langle F^*Fu, u\rangle = \langle Fu, Fu\rangle = 0,$$

and $u \in \operatorname{Ker} F$. If $\operatorname{Im} F$ is closed (e.g., if $F \in \mathcal{F}$), then we also have

$$\operatorname{Im} F^*F = \operatorname{Im} F^*.$$

Here "$\subseteq$" is clear. To prove "$\supseteq$", consider F^*v for $v \in H$, and decompose v into orthogonal components $v = v' + v''$ with $v' \in \operatorname{Im} F$ and $v'' \in \operatorname{Ker} F^*$; then $F^*v = F^*v'$. In this way, we then have represented the kernel and cokernel of any Fredholm operator F as the kernels of the self-adjoint operators F^*F and FF^*, respectively.
b) The contraction operator A of Exercise 2.4 provides an example of a bounded, self-adjoint operator with $\operatorname{Ker} A$ ($= \operatorname{Ker} A^*$) $= \{0\}$ which is not a Fredholm operator.

COROLLARY 2.12. *Let H, H' be Hilbert spaces and $F : H \to H'$ a bounded Fredholm operator. Then $F^* : H' \to H$ is a Fredholm operator and we have*

$$\operatorname{Ker} F^* \cong \operatorname{Coker} F, \quad \operatorname{Coker} F^* \cong \operatorname{Ker} F, \quad \textit{and} \quad \operatorname{index} F^* = -\operatorname{index} F.$$

PROOF. Consider the exact sequence

$$0 \longrightarrow \operatorname{Ker} F \longrightarrow H \xrightarrow{F} H' \longrightarrow \operatorname{Coker} F \longrightarrow 0.$$

Then by Theorem 2.9, the sequence

$$0 \longrightarrow \operatorname{Coker} F \longrightarrow H' \xrightarrow{F^*} H \longrightarrow \operatorname{Ker} F \longrightarrow 0$$

is also exact, and the assertion follows. □

Positive operators. Another concept based on the scalar product is the notion of positive operators.

DEFINITION 2.13. If $C \in \mathcal{B} := \mathcal{B}(H)$ with $\langle Cx, x\rangle \geq 0$ for all $x \in H$, then C is called a **positive operator**. We denote the (convex) set of such operators by $\mathcal{B}^+$.

PROPOSITION 2.14. *Any $C \in \mathcal{B}^+$ is self-adjoint. Moreover, for any $n \in \mathbb{N}$,* $C^n \in \mathcal{B}^+$.

Proof. For all $x \in H$,

$$\langle Cx, x\rangle = \overline{\langle x, Cx\rangle} = \langle x, Cx\rangle = \langle C^* x, x\rangle \Rightarrow \langle (C - C^*)\, x, x\rangle = 0.$$

For $D := C - C^*$ and all $x, y \in H$, we then have

$$\left.\begin{array}{r} \langle D\,(x+y)\,, (x+y)\rangle = 0 \Rightarrow \langle Dx, y\rangle + \langle Dy, x\rangle = 0 \\ \langle D\,(x+iy)\,, (x+iy)\rangle = 0 \Rightarrow -i\,\langle Dx, y\rangle + i\,\langle Dy, x\rangle = 0 \end{array}\right\} \Rightarrow \langle Dx, y\rangle = 0.$$

So $D = 0$ and $C = C^*$. We now have $\left\langle C^2 x, x\right\rangle = \langle Cx, Cx\rangle \geq 0$ and for $n \geq 3$,

$$\langle C^n x, x\rangle = \left\langle C^{n-2}\,(Cx)\,, Cx\right\rangle,$$

whence the positivity of C^n follows by induction. □

For *real* Hilbert spaces H, $\langle Cx, x\rangle \geq 0$ does not imply $C = C^*$ (e.g., $\langle Ax, x\rangle = 0 \geq 0$ for any skew-symmetric A), but we assumed that H is complex.

Below, in Section 2.7 on trace class and Hilbert-Schmidt operators, we shall prove the fundamental *Square Root Lemma* (Theorem 2.66, p.54ff) for all operators belonging to $\mathcal{B}^+$ by means of completely elementary arguments.

2. Compact Operators

So far we found that the space $\mathcal{F}$ of Fredholm operators is closed under composition and passage to adjoints and particularly that all operators of the form $\mathrm{Id} + T$ belong to $\mathcal{F}$ when T is an operator of finite rank. We will increase this supply of examples, in passing to *compact operators* by taking limits. This, however, does not lead to Fredholm operators of non-zero index.

We begin with an exercise which emphasizes a simple topological property of operators of finite rank, more generally characterizes the finite-dimensional subspaces which are fundamental for the index concept, and prepares the introduction of *compact operators*.

Exercise 2.15. a) Every operator with finite rank maps the unit ball (or any bounded subset) of H to a relatively compact set.
b) If H is finite-dimensional, then the closed unit ball $B_H := \{u \in H : \|u\| \leq 1\}$ is compact.
c) If H is infinite-dimensional, then B_H is noncompact.
[Hint for a) and b): Recall the theorem of Bernhard Bolzano and Karl Weierstrass that says that every closed bounded subset of $\mathbb{R}^n$ (or $\mathbb{C}^n$) is compact.
For c): Every orthonormal system $e_1, e_2, \ldots$ in H is a sequence in B_H without a convergent subsequence. For instance, how large is $\|e_i - e_k\|$ for $i \neq k$?]

Definition 2.16. We denote by $\mathcal{K}$ (or $\mathcal{K}\,(H)$) the set of linear operators from H to H which map the open unit ball (or more generally, each bounded subset of H) to a relatively compact subset of H. Such operators are called **compact** (or sometimes **completely continuous**) operators.

By Exercise 2.15a, the compact operators form the largest class of operators that behave (in this respect) like finite rank operators, i.e., like the operators of linear algebra which are defined via matrices.

The following theorem supports establishing the relative compactness of subsets in function or mapping spaces, for instance in our proof of Rellich's compact embedding of Sobolev spaces, Theorem 7.15, p.198f. We state it and its common reformulation without proof. For a clear (but a bit lengthy) proof we refer to [**HiSch**, Satz 3.10].

THEOREM 2.17 (Arzela-Ascoli, 19??). *Let Y be a compact topological space, (X,d) a metric space, and let $C(Y,X)$ denote the metric space of continuous mappings from Y to X, equipped with the uniform metric*

$$d_\infty(f,g) := \sup\{\min\{d(f(y),g(y)),1\} : y \in Y\}.$$

Then we have for every $V \subset C(Y,X)$:

$$V \text{ is relatively compact } \iff \begin{cases} (i)\ V \text{ is uniformly continuous, and} \\ (ii)\ V(y) := \{f(y) : f \in V\} \text{ is relatively compact in } X \text{ for all } y \in Y. \end{cases}$$

Here "V is uniformly continuous" means the following: For all $y \in Y$ and $\varepsilon > 0$ there exists a neighborhood U_y of y such that $d(f(y'), f(y)) < \varepsilon$ for all $y' \in U_y$ and all $f \in V$.

The Arzela-Ascoli Theorem is applied mostly in the following form.

COROLLARY 2.18. *Let Y be a compact topological space and $V \subset C(Y,\mathbb{C})$. Then we have:*

$$V \text{ is relatively compact} \iff V \text{ is uniformly continuous and bounded.}$$

Despite the risks inherent in pictures, we can perhaps best visualize compact operators as "asymptotically" contracting maps which in the case of operators of finite rank map the ball B_H to a finite-dimensional disk, and in general to some sort of elliptical spiral as in Figure 2.1.

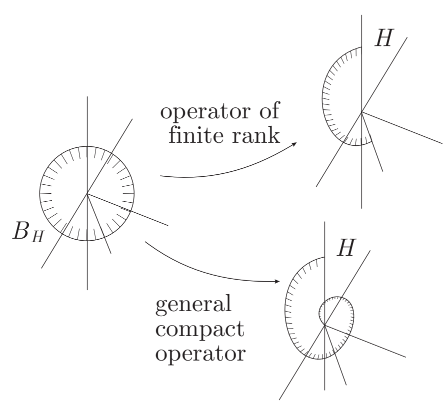


FIGURE 2.1. Compact operators visualized as elliptical spirals

David Hilbert made this visualization precise in his spectral representation of a compact operator K. Accordingly (in the *normal* case $KK^* = K^*K$) the value Ku can be expanded into a series in eigenvectors $u_1, u_2, \ldots$ with the corresponding eigenvalues as coefficients, i.e.,

$$Ku = \sum_{j=1}^{\infty} \lambda_j \langle u, u_j \rangle u_j,$$

whereby the eigenvalues accumulate at 0 and the eigenvectors form an orthonormal system. In the language of operator algebras, that means that K *vanishes at infinity.*

In direct analogy to the principal axis transformation of analytic geometry, we thus obtain for the *quadratic form* defined by K the representation

$$\langle Ku, u\rangle = \sum\nolimits_{j=1}^{\infty} \lambda_j \langle u, u_j\rangle^2 .$$

All proofs can be found in [**Jö**, 1970/1982, 6.2-6.4] or [**Ped**, Lemma 3.3.5 and Theorem 3.3.8], the historical background in [**HT**, 16, 34 and 40] or more compactly in [**Kli**, p.1064-1066].

We give an elementary proof of the spectral representation of a compact operator in the simplest case, namely when the operator is self-adjoint. Our proof is inspired by [**AkGl**, Section 55] and [**GG**, Theorem 5.1]; see also [**CodLev**, Section 7.3–4], where the very same chain of arguments is played through for the special case of self-adjoint Sturm-Liouville boundary problems.

DEFINITION 2.19. We say that an operator is **diagonalizable** (or "discrete") if there is an orthonormal basis $\{e_j : j = 1, 2, \dots\}$ for H and a bounded set $\{\lambda_j : j = 1, 2, \dots\}$ in $\mathbb{C}$ such that

$$Tu = \sum_{j=1}^{\infty} \lambda_j \langle u, e_j\rangle e_j$$

for every $u \in H$.

Note that the numbers $\langle u, e_j\rangle$ are the coordinates for u in the basis $\{e_j\}$ and that each λ_j is an eigenvalue for T corresponding to the eigenvector e_j. So, the matrix corresponding to T and the basis $\{e_j\}$ is the diagonal matrix

$$\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}.$$

We are going to prove the spectral decomposition for compact operators (first proven by David Hilbert in 1904 for Fredholm integral equations and generalized by his student Erhard Schmidt in 1905).

THEOREM 2.20 (Hilbert-Schmidt Theorem, 1904). *Every compact self-adjoint operator K is diagonalizable.*

In the proof we shall use a simple, well-known technical result (Lemma 2.22 below) which follows from the also well-known proposition:

PROPOSITION 2.21. *If a bounded operator T is self-adjoint, then*

$$\|T\| = \sup_{\|u\|=1} \left|\langle Tu, u\rangle\right| .$$

The proposition remains valid for normal operators, see for instance [**Ped**, Proposition 3.2.25] or [**Rud**, Theorem 12.25]. The number on the right is also called the **numerical radius** of T and is denoted by $|||T|||$.

PROOF. Let m denote the numerical radius of T. We deduce $m \le \|T\|$ from the Cauchy–Schwarz inequality

$$\left|\langle Tu, u\rangle\right| \le \|Tu\| \cdot \|u\| \le \|T\| \text{ for } \|u\| = 1.$$

To prove $m \geq \|T\|$ we consider arbitrary $u, v \in H$ and obtain

$$\langle T(u \pm v), u \pm v\rangle = \langle Tu, u\rangle \pm 2\Re\langle Tu, v\rangle + \langle Tv, v\rangle$$

(using the fact that T is self-adjoint), from which

$$4\Re\langle Tu, v\rangle = \langle T(u+v), u+v\rangle - \langle T(u-v), u-v\rangle\,,$$

where $\Re z$ denotes the real part of a complex number z.

To bring m into play, we recall

$$\big|\langle Tw, w\rangle\big| = \big|\langle T\frac{w}{\|w\|}, \frac{w}{\|w\|}\rangle\big|\, \|w\|^2 \leq m\,\|w\|^2 \text{ for every } w \in H.$$

We then get the estimate

$$4\Re\langle Tu, v\rangle \leq m\big(\|u+v\|^2 + \|u-v\|^2\big) \leq 2m\big(\|u\|^2 + \|v\|^2\big) \tag{2.5}$$

with the right inequality deduced from the parallelogram law

$$\|u+v\|^2 + \|u-v\|^2 = 2\,\|u\|^2 + 2\,\|v\|^2\ .$$

If we replace u by αu with $\alpha \in \mathbb{C}$ and $|\alpha| = 1$, the right side of (2.5) remains unchanged and we obtain (for $\alpha := e^{-i\theta}$ when $\langle Tu, v\rangle = |\langle Tu, v\rangle| e^{i\theta}$ for suitable real θ, hence $|\langle Tu, v\rangle| = \Re\langle T(e^{-i\theta}u), v\rangle$):

$$\big|\langle Tu, v\rangle\big| \leq \frac{m}{2}\big(\|u\|^2 + \|v\|^2\big). \tag{2.6}$$

Suppose $Tu \neq 0$. Then taking $v := \frac{\|u\|}{\|Tu\|} Tu$ in (2.6) yields

$$\|u\|\ \|Tu\| = \big|\langle Tu, v\rangle\big| \leq m\,\|u\|^2\ .$$

Hence $\|Tu\| \leq m\,\|u\|$ for all $u \in H$ and so $\|T\| \leq m$. □

Now, the key to the spectral representation is the following

LEMMA 2.22. *If $K \in \mathcal{B}(H)$ is compact and self-adjoint, then at least one of the numbers $\|K\|$ or $-\|K\|$ is an eigenvalue of K.*

PROOF. The lemma is trivial if $K = 0$. Assume $K \neq 0$. It follows from Proposition 2.21 that there exists a sequence (u_n) in H of unit vectors with $\langle Ku_n, u_n\rangle \to \lambda$, where $\lambda = \|K\|$ or $\lambda = -\|K\|$.

To prove that λ is an eigenvalue of K, we first note that

$$\begin{aligned} 0 \leq \|Ku_n - \lambda u_n\|^2 &= \|Ku_n\|^2 - 2\lambda\langle Ku_n, u_n\rangle + \lambda^2 \\ &\leq 2\lambda^2 - 2\lambda\langle Ku_n, u_n\rangle \longrightarrow 0. \end{aligned}$$

Thus

$$Ku_n - \lambda u_n \longrightarrow 0. \tag{2.7}$$

Since K is compact, there exists a subsequence $(Ku_{n'})$ of (Ku_n) which converges to some $v \in H$. Consequently, (2.7) implies that $u_{n'} \to \frac{1}{\lambda}v$, and by continuity of K,

$$v = \lim Ku_{n'} = \frac{1}{\lambda}Kv.$$

Hence $Kv = \lambda v$ and $v \neq 0$ since $\|v\| = \lim \|\lambda u_{n'}\| = |\lambda| = \|K\|$. Thus λ is an eigenvalue of K. □

Now we prove the theorem by repeated application of the preceding lemma.

Proof of Theorem 2.20. By Lemma 2.22, there exists an eigenvalue λ_1 of K and a corresponding unit eigenvector e_1 with $|\lambda_1| = \|K\|$. Set $H_1 := H$ with the closed subspace $H_2 := [e_1]^\perp$. Here

$$[a, b, c \dots] := \mathbb{C}a + \mathbb{C}b + \mathbb{C}c + \dots \tag{2.8}$$

denotes the **linear span** of vectors $a, b, c \dots$. Now, T-invariance of a subspace M for a bounded operator T implies that $M^\perp$ is T^*-invariant. Since K is self-adjoint it follows $KH_2 \subseteq H_2$. Set $K_1 := K$ and $K_2 := K|_{H_2} \in \mathcal{B}(H_2)$. Then K_2 is compact and self–adjoint.

If $K_2 \neq 0$, we repeat the previous argument. So, there exists an eigenvalue λ_2 of K_2 and a corresponding unit eigenvector e_2 with

$$|\lambda_2| = \|K_2\| \leq \|K_1\| = |\lambda_1|.$$

Clearly the pair (e_1, e_2) is orthonormal. Now $H_3 := [e_1, e_2]^\perp$ is a closed subspace of H, $H_3 \subset H_2$ and $KH_3 \subseteq H_3$. Setting $K_3 := K|_{H_3}$, the process continues. It either stops when $K_n = 0$ or else we get a sequence (λ_n) of eigenvalues of K and a corresponding set $\{e_1, e_2 \dots\}$ of mutually orthonormal eigenvectors such that

$$|\lambda_{n+1}| = \|K_{n+1}\| \leq \|K_n\| = |\lambda_n|, \quad n = 1, 2, \dots .$$

If (λ_n) is an infinite sequence, then $\lambda_n \to 0$. Indeed, assume this is not the case. Since $|\lambda_n| \geq |\lambda_{n+1}|$, there exists an $\epsilon > 0$ such that $|\lambda_n| \geq \epsilon$ for all n. Hence for $n \neq m$,

$$\|Ke_n - Ke_m\|^2 = \|\lambda_n e_n - \lambda_m e_m\|^2 = \lambda_n^2 + \lambda_m^2 > \epsilon^2 .$$

But this is impossible since (Ke_n) has a convergent subsequence due to the compactness of K.

We are now ready to prove the diagonalization of K as asserted in the theorem. Let $u \in H$ be given.

Case 1. $K_n = 0$ for some n: Since $u_n := u - \sum_{k=1}^{n-1} \langle u, e_k\rangle e_k$ is orthogonal to e_j , $j = 1, \dots, n-1$, the vector u_n belongs to H_n . Hence

$$0 = K_n u_n = Ku - \sum_{k=1}^{n-1} \lambda_k \langle u, e_k\rangle e_k .$$

Case 2. $K_n \neq 0$ for all n: From what we have seen in case 1 (and the simple $\|u_n\| \leq \|u\|$),

$$\|Ku - \sum_{k=1}^{n-1} \lambda_k \langle u, e_k\rangle e_k\| = \|K_n u_n\| \leq \|K_n\| \|u_n\| \leq |\lambda_n| \|u_n\| \leq |\lambda_n| \|u\| \longrightarrow 0,$$

which means that

$$Ku = \sum_{k=1}^{\infty} \lambda_k \langle u, e_k\rangle e_k . \qquad \square$$

Corollary 2.23. *Let K be a self-adjoint compact operator and $\rho > 0$ a real. Then the operator K has only a finite number of linearly independent eigenvectors such that the corresponding eigenvalues exceed ρ in modulus. In particular, only* 0 *can be an accumulation point of the eigenvalues, and each nonzero eigenvalue has finite multiplicity.*

In [**AkGl**, Section 52] it is shown that the preceding assertion remains valid for any compact operator; i.e., the assumption that K is self-adjoint is dispensable.

Note that an operator K

- is bounded (= continuous) if and only if for every $u \in H$
$$\lim_{|v| \longrightarrow 0} K(u+v) = K(u),$$
- but is compact, if and only if, with respect to a complete orthogonal system $e_1, e_2, \ldots$, we have
$$\lim_{i \longrightarrow \infty} \langle v_i, e_j \rangle = 0 \text{ for each } j \Rightarrow \lim_{i \to \infty} K(u+v_i) = K(u),$$
in which the hypothesis is weaker than $\|v_i\|^2 = \sum_{j=1}^{\infty} |\langle v_i, e_j \rangle|^2 \to 0$.

This describes the context in which Hilbert developed the idea of a compact operator and why he called them *completely continuous*. For our purposes the following result is sufficient.

THEOREM 2.24. *a) $\mathcal{K}$ is a "two-sided (non-trivial) ideal" in the Banach algebra $\mathcal{B}$ of bounded linear operators in a separable, infinite-dimensional Hilbert space H.*
b) $\mathcal{K}$ is closed in $\mathcal{B}$.
c) More precisely, $\mathcal{K}$ is the closure of the subset of finite-rank operators.
d) $\mathcal{K}$ is invariant under $*$; i.e., the adjoint of a compact operator is compact.

PROOF. **To (a)**: For bounded T and compact K, the operators $T \circ K$ and $K \circ T$ are compact, by definition. We therefore have $\mathcal{B} \circ \mathcal{K} \subseteq \mathcal{K}$ and $\mathcal{K} \circ \mathcal{B} \subseteq \mathcal{K}$. Moreover, for $\lambda \in \mathbb{C}$, we have λK compact. Now, let K and K' be compact. To prove that $K + K'$ is compact, we use the sequential criterion for compactness; i.e., an operator is compact, if the image of a bounded sequence of points has a convergent subsequence. Thus, let $u_1, u_2, \ldots$ be a bounded sequence in H. Then, by a two-fold selection of subsequences, we can find $u_{i_1}, u_{i_2}, \ldots$, such that $Ku_{i_1}, Ku_{i_2}, \ldots$ and $K'u_{i_1}, K'u_{i_2}, \ldots$both converge in H, whence $(K+K')\,u_{i_1}, (K+K')\,u_{i_2}, \ldots$ also converges. Finally, it is trivial that every compact operator is bounded, since the image of the unit sphere is relatively compact and hence bounded; thus, $\mathcal{K} \subseteq \mathcal{B}$. Since $\mathcal{K}$ includes the operators of finite rank (Exercise 2.15a) but not the identity (Exercise 2.15c), $\mathcal{K}$ is a nontrivial ideal, and the assertion follows.

To (b): Let $T \in \mathcal{B}$ be an operator in the closure of $\mathcal{K}$. In order to show that each open cover (say, without loss of generality, by *all* of the balls of radius $\varepsilon > 0$ [**Du**, 1966, p.298]) of the image $T(B_H)$ of the closed unit ball of H has a finite subcover, we use an "$\varepsilon/3$-proof" (as is usual in such situations): We choose $K \in \mathcal{K}$ with $\|T - K\| < \varepsilon/3$ and a finite open covering of $K(B_H)$ by balls of radius $\varepsilon/3$, with centers at $Ku_1,, \ldots, Ku_m$ where $u_1, \ldots, u_m \in B_H$. Then the ε-balls about $Tu_1, \ldots, Tu_m$ form the desired finite covering of $T(B_H)$: For each $u \in B_H$, there is some $i \in \{1, \ldots, m\}$, such that $\|Ku - Ku_i\| < \varepsilon/3$, and so
$$\|Tu - Tu_i\| \le \|Tu - Ku\| + \|Ku - Ku_i\| + \|Ku_i - Tu_i\| < \varepsilon.$$

To (c): We now ask which operators are limits in $\mathcal{B}$ of sequences of operators of finite rank. Evidently, these limits typically lie outside of the space of finite rank operators; e.g., see Exercise 2.28 below. Let $e_1, e_2, \ldots$ be a complete orthonormal system in H and let
$$Q_n : H \longrightarrow [e_1, \ldots, e_n]$$
denote the orthogonal projection from H to the linear span of the first n basis elements. This *truncation* yields, for each $T \in \mathcal{B}$, a sequence $Q_1T, Q_2T, \ldots$ of operators of finite rank which converges pointwise to T; i.e., $Q_nTu \to Tu$ for each

$u \in H$. This does not mean that the sequence $Q_1T, Q_2T, ... = (Q_nT)_1^\infty$ converges in $\mathcal{B}$ (i.e., in the operator norm) to T. Of course, not every bounded operator can be the limit of a sequence of operators of finite rank (*e.g.*, $\|Q_n \operatorname{Id} - \operatorname{Id}\| = 1$ for all n); in fact, we have already proved in (b) that limits of compact operators (in particular, finite-rank operators - see Exercise 2.15 a) must be compact. It remains to prove that, for each $K \in \mathcal{K}$, the sequence $(Q_nK)_1^\infty$ converges to K in $\mathcal{B}$. For this, we choose (for each $\varepsilon > 0$) a finite covering of $K(B_H)$ by balls of radius $\varepsilon/3$ with centers at $Ku_1, ..., Ku_m$ and some $n \in \mathbb{N}$, so large that

$$\|Ku_i - Q_nKu_i\| < \varepsilon/3 \text{ for } i = 1, ..., m.$$

This is no problem, since $(Q_n)_1^\infty$converges pointwise to the identity. For each $u \in B_H$, we then have

$$\begin{aligned}\|Ku - Q_nKu\| &\leq \|Ku - Ku_i\| + \|Ku_i - Q_nKu_i\| + \|Q_nKu_i - Q_nKu\| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3\end{aligned}$$

for some $i \in \{1, ..., m\}$. For the last term, note that $\|Q_n\| = 1$, whence

$$\|Q_nKu_i - Q_nKu\| = \|Q_n(Ku_i - Ku)\| \leq \|Ku_i - Ku\|.$$

Thus, we have proven $\|K - Q_nK\| < \varepsilon$ for all sufficiently large n, depending on ε.
To (d): Obviously, the space of operators of finite rank is invariant under $*$, since every such operator T (as in Remark 1.13, p. 10) is of the form $T = \sum_{i=1}^n \langle \cdot, u_i \rangle v_i$, where $u_1, ..., v_n \in H$, and then (verify!) $T^* = \sum_{i=1}^n \langle \cdot, v_i \rangle u_i$ also has finite rank. Now, we can reduce the general case $K \in \mathcal{K}$ to the finite-rank case. Namely, approximate K by a sequence $(T_n)_1^\infty$ of operators of finite rank whose adjoints then approximate K^*, since (by Exercise 2.6),

$$\|T_n^* - K^*\| = \left\|(T_n - K)^*\right\| = \|T_n - K\|. \qquad \square$$

Remark 2.25. The statements in (a), (b), and (d) apply also (admittedly with somewhat different proofs; e.g., see [**Rud**, Theorems 4.18, 4.19] or [**Sche**, Section 4.3]) to the more general case of Banach spaces, but not statement (c). The search for a counterexample began with a legendary treatise by Alexander Grothendieck in [**Gro**] (1955) and led Per Enflo to success, published in [**En**] (1973), incidentally supplemented by many nice examples of the correctness of (c) in special cases, in particular for almost all well-known Banach spaces; see also [**Jö**, 1970/1982, 12.4], but - surprisingly not for $\mathcal{B}(H)$: in [**Sza**] (1981), it was proved by Andrzej Szankowski that the Banach algebra $\mathcal{B}(H)$ of bounded operators in complex separable Hilbert space is *not* approximative, i.e., there exists a compact operator $k : \mathcal{B}(H) \to \mathcal{B}(H)$ which can not be approximated by a sequence $(f_j)_{j=1,2,...} : \mathcal{B}(H) \to \mathcal{B}(H)$ of operators (on the operator space $\mathcal{B}(H)$) of finite range.

Remark 2.26. Since the closure of an ideal is again an ideal (as is trivially proved) and since the operators of finite rank obviously form an ideal in $\mathcal{B}$, we note that (a) follows from (c) – admittedly, somewhat less directly than in the above proof.

Remark 2.27. For practical needs the sequence $(Q_nK)_1^\infty$ stated in the proof of (c) is a poor approximation of K by operators of finite rank. It presupposes the knowledge of K on all of H and works with a completely arbitrary orthonormal system. However K is frequently (see for example Exercises 2.28 and 2.29 below) given in a form which suggests a special approximation or which points to a distinguished orthonormal system, namely the eigenvectors of K. The *spectral representation* of Theorem 2.20 is numerically relevant, because it implies that

$$Q_n K u = K Q_n u = \sum_{j=1}^{n} \lambda_j \langle u, u_j \rangle u_j$$

which does indeed permit a *stepwise* approximation.

EXERCISE 2.28. Show that the operator A in Exercise 2.4 is compact. Namely, give an estimate for $\left\| Au - \sum_{j=1}^{n} \frac{1}{j} \langle u, e_j \rangle e_j \right\|^2$ which is independent of u, for $\|u\| < 1$. Recall the Cauchy-Schwarz inequality $|\langle u, e_j\rangle| \le \|u\| \, \|e_j\|$.

EXERCISE 2.29. Let $[a,b]$ be a compact interval in $\mathbb{R}$. Show that we obtain a compact operator K on the Hilbert space $L^2[a,b]$ for each square-integrable function G on $[a,b]\times[a,b]$ via

$$Ku(x) := \int_a^b G(x,y)u(y)\,dy,\ x \in [a,b]$$

[Hint: Approximate the weight function G by step functions, and hence K by operators with finite rank of the kind considered in Exercise 1.14 (integral operators with *degenerate* weight). For details of the argument, see [**GG**, Example II.14.3], [**Sche**, Section 11.4], [**Kato**, Example V.2.19], [**Ped**, Propositon 3.4.16], and [**ReSi72**, Theorem VI.23]. The last three references show that one can replace the interval $[a,b]$ by any locally-compact Hausdorff space with a fixed Radon integral and that the integral operator is not only compact but **Hilbert-Schmidt**, i.e., with the eigenvalues of K^*K going to zero fast enough to be summable (i.e., the trace $\mathrm{Tr}\,(K^*K) < \infty$).]

EXERCISE 2.30. For the map

$$\begin{array}{ccc} L^2([a,b]\times[a,b]) & \longrightarrow & \mathcal{K}(L^2[a,b]) \\ G & \mapsto & K \end{array},$$

defined in Exercise 2.29, show that:
a) The map is linear and injective.
b) If G is the weight function of K, then

$$\|K\| \le |G| := \left(\int_a^b \int_a^b |G(x,y)|^2 \,dxdy \right)^{1/2}.$$

c) The adjoint operator K^* has the weight function $G^*(x,y) = G(y,x)$.
d) If K_1 and K_2 are given by weight functions G_1 and G_2, then the operator $K_2 \circ K_1$ belongs to the weight function

$$G(x,y) = \int_a^b G_2(x,z)\, G_1(z,y)\; dz.$$

[Hint: Linearity is clear. For injectivity, one naturally (Lebesgue integral!) need only show that $\int_\alpha^\beta \int_\gamma^\delta G(x,y)\,dxdy = \left\langle \chi_{[\alpha,\beta]}, K\chi_{[\gamma,\delta]} \right\rangle$, where $\chi_{[\alpha,\beta]}$ and $\chi_{[\gamma,\delta]}$ are characteristic functions of subintervals $[\alpha,\beta], [\gamma,\delta] \subseteq [a,b]$; then, we have $G = 0$ when $K = 0$. Details (and the generalization to the case of unbounded intervals) are in [**Jö**, 1970/1982, 11.2]. For the proofs of (b), (c), and (d), one needs to use the theorem of Fubini on iterated integrals; e.g., the details are in [**Jö**, 1970/1982, 11.2-11.3] or [**Ped**, Propositon 3.4.16].]

3. The Classical Integral Operators

The type of operators considered here, i.e., those which are given by weight functions square-integrable on the *product space*, are nowadays called Hilbert-Schmidt operators. For a rigorous treatment of abstract Hilbert-Schmidt operators, see Section 2.7 below. By introducing them, the Hungarian mathematician Frigyes (Friedrich) Riesz (1907), not only generalized the theory of integral equations with continuous weight function created by Vito Volterra (Turin, 1896), Erik Ivar Fredholm (Stockholm, 1900) and David Hilbert (Göttingen, 1904f), but drastically simplified it at the same time. Indeed, proofs dealing with L^2-integration theory on Hilbert space compare very favorably with the cumbersome work with uniform convergence in the Banach space of continuous functions.

Of course, there do exist important integral operators whose weight functions are not square integrable: The best-known example is the Fourier transform (see Appendix A below)

$$\widehat{f}(x) = \int_{-\infty}^{\infty} e^{-ixy} f(y)\, dy.$$

It maps $L^2(R)$ bijectively onto itself, but the L^2-norm of the weight,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left|e^{-ixy}\right|^2 dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1 dxdy$$

is *strongly infinite*. Somewhere between these two simplest principal types – the compact Hilbert-Schmidt operators on the one hand and the invertible Fourier transform on the other – lie the *convolution operators*, in particular the *Wiener-Hopf operators* and other *singular integral operators*. Their weights do not belong to L^2, but frequently they are at least componentwise integrable, for example

$$\int_a^b \left|G\left(x,y\right)\right|^2 dy < \infty \text{ for } x \in \mathbb{R}.$$

All these operators are highly significant in kinematic as well as stochastic modeling and in solving a multitude of physical, technical and economical problems: as in the method of inverting differential operators into integral operators which goes back to George Green and was developed on a large scale by David Hilbert; and as in the indirect treatment of collective and statistical phenomena or more generally in probabilistic situations. We will return to a number of particularly interesting integral operators later in this Part and in the following Part. A first survey is given in Table 2.1.

update the cross references

4. The Fredholm Alternative and the Riesz Lemma

If F is a Fredholm operator on a separable Hilbert space H, then the statement "index $F = 0$" can be expressed in familiar classical terminology as: *Either* the equation $Fu = v$ has a unique solution $u \in H$ for each $v \in H$, *or* the homogeneous equation $Fu = 0$ has a nontrivial solution. In the second case, there are at most finitely many linearly independent solutions $w_1, ..., w_n$ of $Fw = 0$ and just as many linearly independent solutions $u_1, ..., u_n$ of the adjoint homogeneous equation $F^*u = 0$; the inhomogeneous equation $Fu = v$ is solvable exactly when

$$\langle v, u_1\rangle = \cdots = \langle v, u_n\rangle = 0.$$

TABLE 2.1. Some fundamental integral operators of the form $Ku(x) := \int_X G(x,y)u(y)\,dy$

No.	X	$G(x,y)$	u	Key words	Property	Literature or Orgin
1.	$[a,b]$	$C^0(X \times X)$	$C^0(X)$	Fredholm Integral	$K : C^0(X) \to C^0(X)$ compact	Fredholm (1903)
2.	compact $\subseteq \mathbb{R}^n$	″	″	Fredholm Integral	″	Hilbert (1904-12) [**Jö**, 8.1]
3.	$[a,b]$	continuous for $y \leq x$, 0 for $y > x$.	″	Volterra Integral	″	Volterra (1896)
4.	″	in $C^0(X \times X \setminus diag.)$, bounded on $X \times X$.	″	Green's operator	″	[**Jö**, 8.2]
5.	compact $\subseteq \mathbb{R}^n$	$G(x,y) = g(x,y)\lvert x-y\rvert^{-\alpha}$, g continuous, $\alpha \in (0,n)$	″	pole singularity	″	″
6.	$[a,b]$	$L^2(X \times X)$	$L^2(X)$	Hilbert-Schmidt operator	$K : L^2(X) \to L^2(X)$ compact	Riesz (1907)
7.	$\mathbb{R}^n$	″	″	″	″	[**Jö**, 11.3], here Exercises 2.23 and 2.24
8.	″	$(2\pi)^{-n/2} e^{-i(x_1y_1+\cdots+x_ny_n)}$	″	Fourier transform	isometry of $L^2(X)$	[**DM**, 2.10]; here Appendix A.2
9.	$\mathbb{R}$	$g(x,y) = k(x-y)$, $k \in L^1(\mathbb{R})$	″	convolution	see Appendix A.2	[**DM**, 2.1]; here Appendix A.2
10.	$\mathbb{R}_+$	″	″	Wiener-Hopf operator	see Section 4.7	[**DM**, 3.6]; here Section 4.7
11.	$\mathbb{R}^n$	$G(x,y) = g(x,y)\lvert x-y\rvert^{-\alpha}$, g continuous, $\alpha \in (n,\infty)$	″	singular integral operators	see Chapter 7	[**Mik**], here Chapter 7

This statement is called the **Fredholm alternative** (met before in Exercise 1.14); its equivalence with "index $F = 0$" follows from Theorem 2.7. We saw already that the Fredholm alternative holds for self-adjoint ($F^* = F$), or more generally normal ($F^*F = FF^*$), Fredholm operators. Moreover, in analogy with finite-dimensional linear algebra, it holds for operators of the form $\mathrm{Id} + T$, when T is an operator of finite rank. Even more important for many applications are the following operators for which the Fredholm alternative holds:

THEOREM 2.31 (F. Riesz, 1918). *For each compact operator K,* $\operatorname{Id}+K$ *is a Fredholm operator with vanishing index.*

Note that our proof functions only in Hilbert space or in approximative Banach space, see Remark 2.25. An alternative proof with wider applicability can be found in [**Kat**, Theorem IV.5.26, p.238f].

PROOF. Using Theorem 2.24c, approximate K by a sequence $K_1, K_2,\ldots$ of operators of finite rank and choose n with $\|K - K_n\| < 1$. Then $\operatorname{Id}+K - K_n$ is invertible. Indeed, let $Q := K - K_n$ and consider the series $\sum_{k=0}^{\infty} Q^k$. Since $\|Q\| < 1$ and

$$\left\|\sum_{k=N}^{M} Q^k\right\| \leq \sum_{k=N}^{M} \|Q^k\| \leq \sum_{k=N}^{M} \|Q\|^k ,$$

the partial sums form a Cauchy sequence in the Banach algebra $\mathcal{B}$. The series then converges, and one has

$$(\operatorname{Id}-Q)\sum_{k=0}^{\infty} Q^k = \left(\sum_{k=0}^{\infty} Q^k\right)(\operatorname{Id}-Q) = \operatorname{Id}.$$

Thus, we can write $\operatorname{Id}+K$ as a product:

$$\operatorname{Id}+K = (\operatorname{Id}+K - K_n)(\operatorname{Id}+(\operatorname{Id}+K - K_n)^{-1}K_n)$$

where the left factor is invertible and the right is $\operatorname{Id}+$ an operator of finite rank, which we know (Exercise 1.12, p.9) is a Fredholm operator with index 0. By the composition rule of Exercise 1.10 (p.9) or Exercise 2.3 (p.14), the statement is proved. □

REMARK 2.32. Note that in the above proof of the formula $\operatorname{index}(\operatorname{Id}+K) = 0$, it was not needed that K is approximated by a sequence of operators of finite rank. It was sufficient to have a *crudely approximating* operator K (with $\|K_n - K\| < 1$). Also, for the determination of the index, it was not necessary to compute or know more precisely the inverse operator $(\operatorname{Id}+K - K_n)^{-1}$. This method of proof which reduces the general case to situations permitting explicit or at least iterative solutions goes back to E. Schmidt and works only in Hilbert space. In more general cases where the theorem still holds the proof starts by showing with a compactness argument that $\operatorname{Ker}(\operatorname{Id}+K)$ (and similarly $\operatorname{Ker}(\operatorname{Id}+K^*)$) is finite dimensional. Then one needs a careful argument concerning the limit process in order to show that $\operatorname{Im}(\operatorname{Id}+K)$ is closed, and one gets only that $\operatorname{Id}+K \in \mathcal{F}$. Finally the homotopy invariant of the index (see Theorem 3.11, p.68) implies the formula $\operatorname{index}(\operatorname{Id}+K) = 0$.

EXERCISE 2.33. Formulate and prove the Fredholm alternative for the linear **Fredholm integral equation of the second kind**

$$u(x) + \int_a^b G(x,y)u(y)\,dy = h(x),$$

where $G \in L^2([a,b]\times[a,b])$; see Exercise 1.14 (p.10), Theorem 2.7 (p.15), Exercise 2.29 (p.25), and Exercise 2.30 (p.25).

5. Sturm-Liouville Boundary Value Problems

We will apply Theorem 2.31 and Exercise 2.33 to a classical boundary value problem in the theory of ordinary differential equations which is not very transparent. We start with a dynamical system with finitely many degrees of freedom described by a system of ordinary differential equations. Such ideally simple models are used in celestial mechanics for the computation of planetary orbits, for investigations of the pendulum and gyroscope, for econometric simulation of economic processes, and for the treatment of many other discrete oscillating systems (N-body problems). Many of these problems are mathematically unsolved, but one knows at least that each set of initial values determines a unique solution curve; i.e., given the differential equation (with half-way reasonable coefficients), the system is completely determined by its state at a single moment in time. The key mathematical tools are the local existence and uniqueness theorems of Augustin Cauchy, Emile Picard, Rudolf Lipschitz, Giuseppe Peano, and Ernst Lindelöf. See, for example, [**CodLev**, Chapters 1-2] or [**UN**, p.164-170]. We have a different situation with continuously distributed oscillating systems such as the oscillating string or flexible rod, electrical oscillations in wires, acoustical vibrations in tubes, heat conduction, heat propagation and other diffusion processes, particularly the statistical treatment of equilibria and motion.

For many such processes, one has partial differential equations (see Part II below) for instance of the form

$$\frac{\partial^2 U}{\partial x^2} = \rho \frac{\partial^2 U}{\partial t^2} + F(x,t)\,, \quad x \in [0,1]\,. \tag{2.9}$$

Under suitable assumptions, (2.9) can be reduced to an ordinary differential equation by *separating variables.* More explicitly, setting

$$U(x,t) = u(x)\psi(t),$$

we obtain

$$u'' + ru = f, \tag{2.10}$$

where r, f are given, and u is to be found. Hereby (2.10) usually inherits boundary conditions, e.g.,

$$u(0) = u(1) = 0 \tag{2.11}$$

from (2.9). For details and generalizations, see [**CH**, I, V.3], [**BlCs**, p.25ff, 285ff, 351ff], [**Lo**, p.108-152] or [**UN**, p.164-170].

We stick with this example which goes back to John Bernoulli's *brachistochrone* problem and more generally to the beginnings of the calculus of variations and of geometric optics by Pierre de Fermat [**Kli**, Ch. 24]. For starters let $r = 0$. Evidently the homogeneous differential equation associated with (2.10) (put $f = 0$) has only the trivial solution $u = 0$ if the boundary conditions (2.11) are to be satisfied. In this case there is a Green's function (see e.g., [**CodLev**, Theorem 7.2.2] or [**CH**, I, V.14-15]) which, for each (piecewise continuous) f, yields a solution of the differential equation (2.10) with boundary conditions (2.11) given by the formula

$$u(x) = \int_0^1 k(x,y) f(y)\, dy. \tag{2.12}$$

For the present case of (2.10) with $r = 0$ and (2.11),

$$k(x,y) = \begin{cases} x(y-1), & \text{for } x \le y, \\ y(x-1), & \text{for } x \ge y \end{cases}$$

[**CH**, I, V.15.1].

Now let $f = 0$ and let r be a positive real number. Then the general solution of (2.10) has the form

$$u(x) = c_1 e^{i\sqrt{r}x} + c_2 e^{-i\sqrt{r}x};$$

i.e., each of the two conditions of (2.11) determines a one-dimensional family of solutions where

$$c_1/c_2 = \begin{cases} -1, & \text{for } u(0) = 0, \\ -e^{-2i\sqrt{r}}, & \text{for } u(1) = 0. \end{cases}$$

The two families coincide, exactly when $\sqrt{r}$ is a multiple of π. If so, there is in addition to the zero function another solution u_0 of the homogeneous differential equation

$$u'' + ru = 0 \tag{2.13}$$

which satisfies the boundary conditions (2.11). Thus we have here – in contrast to the Uniqueness Theorem of Lipschitz – an ordinary differential equation of second order with two conditions imposed (which, however, are not concentrated at an initial point but distributed over two points) whose solution is not unique. Another peculiarity of this case is that in contrast to the Existence Theorem of Picard the inhomogeneous differential equation (2.10) subject to the conditions (2.11) need not always have a solution. This happens for example if the driving force f (the *source* term) equals the eigenfunction u_0, or more generally [**CH**, I, V.14.2] if

$$\int_0^1 f(x)u_0(x)\,dx \neq 0.$$

This is the case of resonance which means that the system becomes unstable under the influence of an exterior force. The lack of a solution does not mean that nothing happens, but from the point of view of the user it is an indication that some critical phenomenon might occur: A short in a wire, the collapse of a bridge, extreme concentration of light beams which is technically utilized in a laser. Also, mathematically speaking, the lack of a solution means only that no solution of the given or desired type exists, in our example no bounded function which has a piecewise continuous second derivative. Frequently, this is an indication that a reformulation or refinement of the question is necessary.

We note finally that in the other case, when the nontrivial solution sets with $u(0) = 0$ and those with $u(1) = 0$ are disjoint, the equations (2.10) and (2.11) always have a unique solution, and a Green function can be constructed which carries the essential information of (2.10) and (2.11) and yields the solution for each right hand side f in the integral form (2.12) [**CH**, I, V.14.1]. Combining the two cases we obtain a kind of Fredholm alternative:

Either the differential equation (2.10) together with the boundary conditions (2.11) possesses a unique solution u for every given f, or else the homogeneous equation (2.13) has a solution which does not vanish identically. In the second case the equations (2.10) and (2.11) have a solution, if and only if the orthogonality condition

$$\int_0^1 f(x)u(x)\,dx = 0$$

holds for each solution u of the homogeneous equation (2.13), where f is the right hand side of (2.10).

The analogy with the Fredholm alternative for integral equations (Exercise 2.33, p.28) is not accidental. When the solution is unique, the Green function makes the connection via formula (2.12). But even when solutions do not necessarily exist or when they are not unique, then the classical theory manages to work with *generalized Green functions*.

We will not deal with these questions in detail, but refer the reader to the quoted literature. The fundamental methodological and, in our context, particularly interesting point of view is perhaps best made precise as follows:

Exercise 2.34. Consider the differential equation

$$u'' + pu' + qu = f \tag{2.14}$$

on the interval $[0,1]$ with $p, q, f \in C^0[0,1]$ and with the boundary conditions

$$u(0) = a \text{ and } u(1) = b. \tag{2.15}$$

Show that the integral equation

$$v - Kv = g \tag{2.16}$$

is equivalent to the boundary-value problem (2.14), (2.15), if

$$\begin{aligned}
&Kv(x) := \int_0^1 G(x,y)v(y)\,dy,\ \ x \in [0,1]\,,\\
&G(x,y) := \begin{cases} y(q(x)(1-x) - p(x)), & \text{for } y \le x,\\ (1-y)(q(x)x + p(x)), & \text{for } y > x,\end{cases}\\
&g := ph' + qh - f,\ \text{ and}\\
&h := a(1-x) + bx\ \ (\text{hence, } h' = b - a).
\end{aligned}$$

[Hint: Show that every twice continuously differentiable solution of (2.14) and (2.15) yields a solution $v := u''$ of (2.16), and conversely, every continuous solution v of (2.16) gives a twice continuously differentiable solution of (2.14) and (2.15) by means of

$$\begin{aligned}
&u(x) := h(x) + \int_0^1 k(x,y)v(y)\,dy,\ \text{ where}\\
&k(x,y) := \begin{cases} x(1-y), & \text{for } x \le y,\\ y(1-x), & \text{for } x \ge y.\end{cases}
\end{aligned}$$

Details may be found in [**Jö**, 1970/1982, 9.1] (see also [**Ped**, Section 3.4.18] for a rigorous treatment of the general symmetric second-order differential equation with a certain periodic self-adjoint boundary condition). For a first calculation and in order to maintain continuity with the preliminary remarks, it is recommended that one first try $p = 0$, q positive and constant, and set $a = b = 0$, see also p.67.]

Remark 2.35. With Exercise 2.33, the *Fredholm alternative* for the boundary-value problem follows from the equivalence proved in Exercise 2.34. More precisely, $\mathrm{Id} - K$ is a Fredholm operator on the Hilbert space $L^2[0,1]$, and $\mathrm{index}(\mathrm{Id} - K) = 0$. Thus, we have a Fredholm alternative relative to $L^2[0,1]$. Actually, from the *Closed Graph Theorem* and a regularity theorem (see Chapters 9 and 10; incidentally, we see here that the Banach space theory is genuinely more difficult than the Hilbert space theory), we have that each square integrable solution of equation (2.14) is continuous, provided the right side is continuous. Thus, the Fredholm alternative in $C^0[0,1]$ holds: Either $\dim \mathrm{Ker}(\mathrm{Id} - K) = 0$ and so (because $\mathrm{index}(\mathrm{Id} - K) = 0$, and hence $\dim \mathrm{Coker}(\mathrm{Id} - K) = 0$) the equation (2.14) has a unique solution for each

$g \in C^0[0,1]$ (whence, the boundary-value problem (2.14), (2.15) also has a unique solution for each $f \in C^0\,[0,1]$ and fixed boundary values a, b), or the homogeneous equation $v - Kv = 0$ has nontrivial solutions.
In the second case, the adjoint integral equation also has a nontrivial solution; i.e., there is a $w \in L^2[0,1]$ with

$$\begin{aligned} w(x) = (1-x)\int_0^x &(q(y)y + p(y))w(y)\,dy \\ &+ x\int_x^1 (q(y)\,(1-y) - p(y))w(y)\,dy. \end{aligned}$$

We can differentiate with respect to the upper and lower bounds, obtaining that w is continuously differentiable and

$$\begin{aligned} w'(x) = &-\int_0^x (q(y)y + p(y))w(y)\,dy + (1-x)(q(x)x + p(x))w(x) \\ &+ \int_x^1 (q(y)\,(1-y) - p(y))w(y)\,dy - x(q(x)\,(1-x) - p(x))w(x) \\ &= p(x))w(x) - \int_0^x \cdots + \int_x^1 \cdots\,. \end{aligned}$$

We bring $p(x)w(x)$ to the left side, differentiate once more, and obtain

$$(w' - pw)' + qw = 0. \tag{2.17}$$

From the integral equation for w, we have $w(0) = w(1) = 0$. Hence, every solution w of the homogeneous adjoint integral equation is a solution of the formal-adjoint homogeneous differential equation (2.17) with the homogeneous boundary conditions $w(0) = w(1) = 0$. By **formal-adjoint**, we mean that for all $u, w \in C^2[0,1]$ with the homogeneous boundary condition, we have

$$\int_0^1 u\left((w' - pw)' + qw\right)\,dx = \int_0^1 (u'' + pu' + qu)\,w\,dx$$

which one can verify through integration by parts. For brevity, we have taken all functions to be real-valued.
In the second case of the Fredholm alternative, the problem (2.14) and (2.15) is solvable exactly when (2.16) is solvable; i.e., when (2.17) has a nontrivial solution w with $\langle g, w\rangle = 0$, which means that in terms of f, we have

$$\int_0^1 f(x)w(x)\,dx = aw'(0) - bw'(1).$$

Details of this argument and similar treatments of other boundary-value problems for ordinary differential equations of second order (Sturm-Liouville problems) can be found in [**CodLev**, Chapters 7 and 12] and [**Jö**, 1970/1982, 9.2].

Remark 2.36. If the boundary value problem (2.14) and (2.15) is equivalent to the integral equation (2.16), what is the special nature of the presentation (2.16) in comparison with (2.14) and (2.15)? We bring out three points:
1. The integral equation succeeds in combining two equations, the differential equation and the boundary conditions, into one.
2. Let

$$L : C^2[0,1] \longrightarrow C^0[0,1]$$

denote the differential operator defined by the left side of (2.14) and let

$$B : C^2[0,1] \longrightarrow \mathbb{C} \oplus \mathbb{C}$$

denote the boundary operator defined by the left sides of (2.15). Then Exercise 2.34 says that the operators

$$L \oplus B : C^2[0,1] \longrightarrow C^0[0,1] \oplus \mathbb{C} \oplus \mathbb{C} \text{ and}$$
$$\mathrm{Id} - K : L^2[0,1] \longrightarrow L^2[0,1]$$

are equivalent in the sense that $\mathrm{Ker}\,(L \oplus B) \cong \mathrm{Ker}\,(\mathrm{Id} - K)$ and $\mathrm{Coker}\,(L \oplus B) \cong \mathrm{Coker}\,(\mathrm{Id} - K)$. Here $\mathrm{Id} - K$ is a bounded operator on a Hilbert space to itself, while the functional analytic structure of $L \oplus B$ is much less clear. The equivalence of the differential and the integral equation is a formal one, while the equivalence of the C^0/C^2-theory and the L^2-theory is fairly elementary, but by no means obvious: While nature poses its problems usually in the spaces C^2 or $C^{2(\text{piecewise})}$, mathematicians decide freely in which spaces they want to solve these problems. Hilbert spaces are used, not because of their intrinsic beauty, but because integral equations on L^2 can be treated more efficiently and more transparently than on C^0. The *Regularity Theorem* provides the justification for this procedure and shows at the same time that the *freedom* of the mathematician is not arbitrary.
3. Numerically, the integral operator $\mathrm{Id} - K$ is dealt with by approximating the compact operator K by operators of finite rank or by approximating the weight function $G(x,y)$ of K by degenerate weights of the form $\phi(x)\psi(y)$. Jacques-Charles-Francois Sturm and Joseph Liouville first and successfully undertook the systematic investigation of boundary value problems for ordinary differential equations of second order. It is quite characteristic that they also arrived at their algebraic solution methods by an approximation principle: They investigated related difference equations and then passed to the limit (Jour. de Math. 1 (1836), 106-186 and 373-444). The difference is that the approximation of the integral equation in some cases (e.g., when G is continuous and non-negative) can be done very naturally by development into a series in eigenfunctions (analogous to principal axis transformation of quadratic forms) so that the integral equation becomes immediately clearer ([**CH**, I, III.5.1]). In contrast, the approximation of a differential equation by difference equations is done *blindly* so to speak. It requires the ingenuity of a Sturm and Liouville (or nowadays extensive free computer time) to regain the necessary information about the boundary value problem from the discrete pieces. (Of course, the *blind* approximation always works, while for many Sturm-Liouville problems no explicit eigenfunctions are known.)

In the final analysis the three viewpoints arise from the duality between local and global terms and operations. This duality pervades large parts of analysis (see the Fourier Inversion Formula in Appendix A or the Index Formula itself): While differentiation of a function is a purely local operation, the solution of a differential equation with initial or boundary conditions always requires a certain global operation. This circumstance is illustrated already by the Newtonian formula relating derivative and integral. It may also explain why the local theory of (e.g., elliptic) differential equations is so difficult (one has to do global theory anyway, namely in R^n), and why at times a purposely global approach, say starting with differential operators on closed manifolds, leads more quickly and easily to fundamental local results. We resume this thought in Part II.

6. Unbounded Operators

So far we have considered only bounded Fredholm operators, i.e., linear Fredholm operators from one separable Hilbert space H_1 to another separable Hilbert space H_2 which are continuous and defined on all of H_1. Identifying H_1 and H_2 we ended up with the space of bounded Fredholm operators as a subspace of the algebra of bounded operators $\mathcal{B}(H)$. That is the main line of presentation chosen for this book. It is sufficient for establishing the Atiyah–Singer Index Theorem. Therefore, a hurried reader may skip this section.

However, since our Definition 1.1 (p.3) of Fredholm operator is purely algebraic, we can reformulate and generalize most of the functional analytical and topological results of this book concerning bounded Fredholm operators to the unbounded case (though still assuming the linearity of the operators).

There is good reason to make this generalization: Differential operators are naturally defined on domains which are dense subspaces of the full L^2. However, there is no reasonable way to extend them to endomorphisms, acting on the full L^2 and with values in the same L^2–space. And on their domain, they are not bounded relative to the L^2–norm. In particular for a closed (not necessarily bounded) operator T (see below) there are various ways to re-write or transform T as a bounded operator, be it by Riesz transform (for self-adjoint T), Cayley transform, or simply by equipping the domain $\mathrm{Dom}(T)$ with the graph norm, to be recalled below in (2.23), p.36. For index theory, these approaches are valuable to some extent, as we shall show. However, they always distort the picture, and a treatment of unbounded operator is in order.

Exercise 2.37. Let the unit circle S^1 be parametrized by the angle $\theta \in [0, 2\pi)$. Consider the Hilbert space $L^2(S^1)$ of all square Lebesgue integrable complex–valued functions on S^1 with inner product

$$\langle u, v\rangle_{L^2} := \int_0^{2\pi} u(\theta)\overline{v(\theta)}\, d\theta \quad \text{for } u, v \in L^2(S^1) \tag{2.18}$$

and norm $\|u\|_{L^2} := \sqrt{\langle u, u\rangle_{L^2}}$ (see Exercise A.1, p.669). On the dense subspace $C^1(S^1)$ of differentiable (periodic) functions with continuous derivative (l.c., Exercise A.1e), the differentiation $d/d\theta$ defines a linear operator

$$T_0 : C^1(S^1) \ni u \mapsto u' = \frac{du}{d\theta} \in L^2(S^1) \tag{2.19}$$

with

$$\mathrm{Im}(T_0) = \{v \in C^0(S^1) : \int_0^{2\pi} v(\theta) d\theta = 0\}.$$

Show that T_0 is not continuous as a mapping in $L^2(S^1)$.
[Hint: $\|u'\|_{L^2}$ can be arbitrarily large for $\|u\|_{L^2} = 1$.]

Of course, $T_0 = d/d\theta$ is bounded if it is regarded as an operator from the Banach space $C^1(S^1)$ to the Banach space $C^0(S^1)$. Many problems in analysis, however, require exploiting Hilbert space structure for effective treatment. After all, the state space of quantum mechanics is a Hilbert space; and the Spectral Theorem, both for bounded and unbounded operators, is valid only in Hilbert space. Admittedly, most results on Fredholm operators can also be obtained in Banach space, but are prettier, more meaningful, and much simpler in Hilbert space.

These problems lead to unbounded operators, but not irrevocably: Alternatively, staying with our example, we can extend the differential operator T_0 to *the first* **Sobolev space** $W^1(S^1)$. It is the completion of the pre–Hilbert space $C^1(S^1)$ equipped with the inner product

$$\langle u, v\rangle_{W^1} := \int_0^{2\pi} u(\theta)\overline{v(\theta)}\, d\theta + \int_0^{2\pi} u'(\theta)\overline{v'(\theta)}\, d\theta \tag{2.20}$$

and hence a Hilbert space. Then the extension (easily produced in Theorem 2.40 a)

$$T : W^1(S^1) \longrightarrow L^2(S^1) \tag{2.21}$$

of T_0 becomes a bounded Fredholm operator from the (whole) Hilbert space $W^1(S^1)$ to the (different) Hilbert space $L^2(S^1)$. Thus, unbounded operators can be averted in this sense.

In Part II of this book we shall follow this approach. It is quite effective for the study of elliptic differential operators on closed manifolds, but not sufficient for the study of boundary value problems. Different boundary conditions give different extensions of the formal differential operator and different domains. Therefore, varying boundary value problems for a fixed formal differential operator can often best be treated in a shared framework, namely considering them all as densely defined unbounded operators, operating by the same formal rules in the same basic L^2 space and distinguished only by their domains. The point is that two such operators may be equal on a dense subspace, and yet be quite different. For a 1–dimensional example see our discussion of Sturm–Liouville problems in Section 2.5.

Closed Operators. We emphasized that bounded operators in Hilbert space share many properties with familiar matrix calculus in finite dimensions; that differential operators can be treated as bounded operators in suitably re-defined domains, the so-called Sobolev spaces; and that those Sobolev spaces in our Part II will be equipped with scalar products that make them separable Hilbert spaces. However, a student would be mislead, if we over-emphasize the analogy with elementary Euclidean linear algebra. To get a realistic feeling for the delicate aspects of differential operators, of partial differential equations and global analysis, the reader must become knowledgable of the basic concepts and fundamental results regarding unbounded operators. That is the goal of the following short course in closed operators and not necessarily bounded Fredholm operators. The reader should work it through, even though we can derive most of the key results of index theory without referring to that world, to that language and theory.

Definition 2.38. Let H be a Hilbert space.
a) A **densely defined operator** (not necessarily bounded) in H is a linear mapping

$$T : \mathrm{Dom}(T) \longrightarrow H\,,$$

where $\mathrm{Dom}(T)$ – the **domain** of T – is a (linear) dense subspace of H.
b) If S and T are operators in H such that $\mathrm{Dom}(S) \subseteq \mathrm{Dom}(T)$ and $Su = Tu$ for every $u \in \mathrm{Dom}(S)$, we say that T is an **extension** of S and write $S \subseteq T$.
c) For a densely defined operator T in H, we form the **adjoint operator** T^* in H by letting $\mathrm{Dom}(T^*)$ denote the subspace of elements $u \in H$ for which the functional $v \mapsto \langle Tv, u\rangle$ on $\mathrm{Dom}(T)$ is bounded (= continuous). Since $\mathrm{Dom}(T)$ is dense in H,

the functional extends by continuity to H, and thus there is a unique element T^*u in H such that

$$\langle v, T^*u\rangle = \langle Tv, u\rangle \qquad \text{for all } v \in \mathrm{Dom}(T). \tag{2.22}$$

d) A **closed operator** in H is a densely defined operator whose **graph**

$$\mathfrak{G}(T) := \{(u, Tu) : u \in \mathrm{Dom}(T)\}$$

is a closed subspace of $H \oplus H$.
e) A densely defined operator T in H is **closable** if the (norm) closure of $\mathfrak{G}(T)$ in $H \oplus H$ is the graph of an operator $\overline{T}$. In that case $\overline{T}$ is a closed operator and it is the **minimal closed extension** of T.

Exercise 2.39. a) Show that for every densely defined operator T in H we have

$$(\operatorname{Im} T)^{\perp} = \operatorname{Ker} T^* \ ,$$

exactly as for bounded operators (see proof of Theorem 2.7).
b) Show that a closed operator, while not necessarily continuous, at least has some 'decent' ([**Ped**]) limit behavior: If (u_n) is a sequence in $\mathrm{Dom}(T)$ converging to some $u \in H$ and if (Tu_n) converges to some $v \in H$, then $u \in \mathrm{Dom}(T)$ and $Tu = v$. Moreover, show that this limit behavior characterizes a closed operator.
c) Conclude that an everywhere defined closed operator is bounded.
d) More generally, let T be a closed operator in H with $\mathrm{Dom}(T) = W \subseteq H$. Show that the inner product of the Hilbert (sub)space $\mathfrak{G}(T) \subseteq H \oplus H$ induces an inner product on W which makes W a Hilbert space. Conclude that $T : W \to H$ can be considered as a bounded operator if W is equipped with the corresponding norm induced by the graph of T.
e) Show that for every densely defined operator T in H the adjoint T^* is a closed operator and we have an orthogonal decomposition

$$H \oplus H = \overline{\mathfrak{G}(T)} \oplus U\mathfrak{G}(T^*),$$

where U is the unitary anti–involution on $H \oplus H$ given by $U(w, v) = (-v, w)$. Furthermore, T is closable iff T^* is densely defined, and in that case $\overline{T} = T^{**}$.
f) Let H be a Hilbert space with $u_1, u_2, \ldots$ a complete orthonormal system (i.e., an orthonormal basis for H). Let $(\lambda_j)_{j=1,2\ldots}$ be an unbounded sequence of numbers. Find a linear subspace $D \subset H$ such that the multiplication operator M_{id} is closed, given by the domain D and the operation

$$D \ni u = \sum c_j u_j \quad \mapsto \quad M_{\mathrm{id}}(u) := \sum \lambda_j c_j u_j \, .$$

g) Prove: If T is a densely defined, closed operator in H, and T is injective with dense range, then the same properties hold for T^* and for T^{-1}, and

$$(T^*)^{-1} = (T^{-1})^* \, .$$

[Hint: Deduce (a) from (2.22). For (c), apply the closed graph theorem. In (d), the inner product in W underlying the graph norm

$$\|u\|_T := \|((u, Tu)\|_{\mathfrak{G}(T)} = \sqrt{\|u\|_H^2 + \|Tu\|_H^2}\,, u \in W \tag{2.23}$$

induced in W by T is an immediate generalization of the inner product (2.20) of the first Sobolev space. Note that the converse is not valid, i.e. not every bounded operator from such graph–norm equipped W to H can be considered as a closed operator in H with domain W. E.g., if $W \subset H$ is dense, then

$$\overline{\mathrm{Id}\,|_W} = \mathrm{Id}\,|_H\,,$$

so $\mathrm{Id}\,|_W$ is not closed.
For (e), prove first $\mathfrak{G}(T)^\perp = U\mathfrak{G}(T^*)$. Since U is unitary, it follows that T^* is closed. This gives the decomposition, but there is more to prove. Help can be drawn from [**Ped**, Theorem 5.5]. For (f) recall that each element u in H has the form

$$u = \sum d_j u_j\,,$$

where the sum converges in the norm induced by the inner product. Taking inner products with the u_j's you see that the coordinates for u are determined by $d_j = \langle u, u_j\rangle$. Note also the **Parseval identity** (a generalization of Pythagoras' theorem)

$$\|u\|^2 = \sum |d_j|^2\,,$$

obtained by computing $\langle u, u\rangle$. See also Exercise 2.50c. A safety net for (g) is provided in [**Ped**, Proposition 5.1.7].]

We shall dwell a little more on the example of the standard first order differential operator $d/d\theta$ over the circle S^1. Our goal is to prove that the operator T which acts like $d/d\theta$ on the domain $W^1(S^1)$ is a closed operator in $L^2(S^1)$. There is nothing surprising in the result; no interesting consequences are attached to it; but it permits us to introduce and illustrate some of the most basic concepts of the analysis of elliptic differential operators (to be developed below in Part II).

Recall the well-known fact that the functions

$$e_k(\theta) = \frac{1}{\sqrt{2\pi}} e^{ik\theta}\,, \qquad k \in \mathbb{Z} \tag{2.24}$$

form a complete orthonormal system (= orthonormal basis) for $L^2(S^1)$ (see also Appendix A below). It follows that the space $C^\infty(S^1)$ is dense in the space $L^2(S^1)$. Clearly, the space $C^\infty(S^1)$ can be considered as the space of smooth complex–valued functions on the real line of period 2π.

Let $u \in L^2(S^1)$. We shall use the notation

$$\widehat{u}(k) := \langle u, e_k\rangle\,, \qquad k \in \mathbb{Z} \tag{2.25}$$

for the **Fourier coefficients**.

Recall the definition of the first Sobolev space $W^1(S^1)$ from above. Note that the inner product defined in (2.20) is exactly the inner product induced by the graph $\mathfrak{G}(T) \subset H \oplus H$. For $u \in W^1(S^1)$ we write

$$\|u\|_{W^1} := \sqrt{\|u\|_{L^2}^2 + \|Tu\|_{L^2}^2}\,, \tag{2.26}$$

where $T : W^1(S^1) \to L^2(S^1)$ extends $d/d\theta$ on $C^1(S^1)$, as in the following.

THEOREM 2.40. *a) The operator $d/d\theta$ on $C^1(S^1)$ has a unique continuous (i.e., bounded) extension $T : W^1(S^1) \to L^2(S^1)$, where $W^1(S^1)$ is the first Sobolev space $W^1(S^1)$ (see* (2.20)*).*
b) The W^1–norm and ***the Fourier coefficient norm***

$$\|u\|_1 := \frac{1}{\sqrt{2\pi}} \sqrt{\sum_{k\in\mathbb{Z}} \widehat{u}(k)^2 (1 + k^2)} \tag{2.27}$$

are equivalent norms on $W^1(S^1)$.
c) The first Sobolev space $W^1(S^1)$ *is contained in the Banach space of continuous functions* $C^0(S^1)$ *and the inclusion is continuous.*
d) For each $\theta \in S^1$, *the* ***Dirac distribution***

$$\delta_\theta : C^0(S^1) \ni u \mapsto u(\theta) \in \mathbb{C} \tag{2.28}$$

extends to a continuous mapping from $W^1(S^1)$ *to* $\mathbb{C}$ *(i.e., a bounded operator of rank 1 or continuous linear functional).*
e) The ***normalized integration operator***

$$C^0(S^1) \ni v \quad \mapsto \quad \int_0^\theta v(s)ds - \frac{\theta}{2\pi}\mathcal{J}(v) \in C^1(S^1) \tag{2.29}$$

with

$$\mathcal{J}(v) := \int_0^{2\pi} v(s)\, ds\,,$$

extends to a bounded operator $S : L^2(S^1) \to W^1(S^1)$. *Moreover, up to a finite rank operator, the operator* S *is a right and left inverse of* T *(such a quasi–inverse is called a 'parametrix'):*

$$T \circ S = \mathrm{Id}_{L^2} - \frac{1}{2\pi}\mathcal{J} \quad \text{ and } \quad S \circ T = \mathrm{Id}_{W^1} - \delta_0\,. \tag{2.30}$$

f) As an operator in $L^2(S^1)$, *the* ***differentiation operator*** T *(with* $W^1(S^1) \subset L^2(S^1)$ *as its domain) is closed. Both its kernel and cokernel are one–dimensional, so its index vanishes.*
g) The range $\mathrm{Im}(T)$ *is closed in* $L^2(S^1)$.

PROOF. *of (a)*: For $u \in C^1(S^1)$ we have

$$\|Tu\|_{L^2}^2 = \langle u', u' \rangle_{L^2} \le \langle u, u \rangle_{L^2} + \langle u', u' \rangle_{L^2} = \|u\|_{W^1}^2\,.$$

So, T is continuous as a mapping from the dense subset $C^1(S^1) \subset W^1(S^1)$ to $L^2(S^1)$, and so uniquely extends to $W^1(S^1)$ as a bounded operator.
of (b): We have

$$\langle e_k, e_k \rangle_{W^1} = 1 + k^2\,,$$

hence

$$\|u\|_{W^1}^2 = \langle u, u \rangle_{W^1} = \sum_{k\in\mathbb{Z}} |\widehat{u}(k)|^2(1+k^2).$$

of (c): For $\theta \in S^1$ and $u \in C^\infty(S^1)$ we have

$$\begin{aligned}
|u(\theta)| &= \frac{1}{\sqrt{2\pi}}\left|\sum_{k\in\mathbb{Z}} \widehat{u}(k)e^{ik\theta}\right| \le \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}} |\widehat{u}(k)| \\
&= \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}} |\widehat{u}(k)|(1+k^2)^{\frac{1}{2}}(1+k^2)^{-\frac{1}{2}} \\
&\le \frac{1}{\sqrt{2\pi}}\sqrt{\sum_{k\in\mathbb{Z}} |\widehat{u}(k)|^2(1+k^2)}\sqrt{\sum_{k\in\mathbb{Z}}(1+k^2)^{-1}},
\end{aligned}$$

where the last inequality is Schwarz' inequality

$$|\langle a, b \rangle| \le \sqrt{\langle a, a \rangle} \cdot \sqrt{\langle b, b \rangle}$$

in the Hilbert space $L^2(S^1)$ (or, correspondingly in the Hilbert space ℓ^2 of square–summable sequences) for $a := \sum \widehat{u}(k)(1+k^2)^{\frac{1}{2}}e_k$ and $b := \sum(1+k^2)^{-\frac{1}{2}}e_k$. Clearly

$b \in L^2(S^1)$ since $\sum(1+k^2)^{-1} < \infty$. To see $a \in L^2(S^1)$, we apply the preceding result (b) to $u \in C^\infty(S^1) \subset W^1(S^1)$. Applying (b) once again yields

$$\sup_{\theta \in S^1} |u(\theta)| \leq C\|u\|_{W^1}$$

where the constant C does not depend on u. We extend the estimate to the whole $W^1(S^1)$ by density.

of (d): Clearly, δ_θ is continuous on $C^0(S^1)$ for each $\theta \in S^1$. Then the assertion follows from (c).

of (e): Let $v \in C^0(S^1)$ and $k \in \mathbb{Z}$, $k \neq 0$. Then

$$\begin{aligned}
\langle Sv, e_k\rangle_{L^2} &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \left(\int_0^\theta v(s)ds \right) e^{-ik\theta} d\theta - \frac{\mathcal{J}(v)}{(2\pi)^{3/2}} \int_0^{2\pi} \theta e^{-ik\theta} d\theta \\
&= \frac{-1}{\sqrt{2\pi}} \int_0^{2\pi} v(\theta) \frac{1}{-ik} e^{-ik\theta} d\theta + \frac{1}{\sqrt{2\pi}} \left[\left(\int_0^\theta v(s)ds \right) \frac{1}{-ik} e^{-ik\theta} \right]_0^{2\pi} \\
&\qquad + \frac{\mathcal{J}(v)}{(2\pi)^{3/2}} \int_0^{2\pi} \frac{1}{-ik} e^{-ik\theta} d\theta - \frac{\mathcal{J}(v)}{(2\pi)^{3/2}} \left[\theta \frac{1}{-ik} e^{-ik\theta} \right]_0^{2\pi} \\
&= \frac{-i}{k\sqrt{2\pi}} \int_0^{2\pi} v(\theta) e^{-ik\theta} d\theta + \frac{i}{k\sqrt{2\pi}} (\mathcal{J}(v) - 0) \\
&\qquad + \frac{i\mathcal{J}(v)}{k(2\pi)^{3/2}} \left[\frac{1}{-ik} e^{-ik\theta} \right]_0^{2\pi} - \frac{i\mathcal{J}(v)}{k(2\pi)^{3/2}} (2\pi - 0) \\
&= -\frac{i}{k} \widehat{v}(k),
\end{aligned}$$

because the third term vanishes and the second and forth cancel each other. Hence

$$\langle Sv, Sv\rangle_{L^2} = \sum_{k \neq 0} |\widehat{v}(k)|^2 \frac{1}{k^2} + |\langle Sv, e_0\rangle|^2 .$$

By partial integration and Schwarz' inequality we obtain

$$|\langle Sv, e_0\rangle| = \left| \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \left(\int_0^\theta v(s)ds \right) d\theta - \frac{\mathcal{J}(v)}{(2\pi)^{3/2}} \int_0^{2\pi} \theta d\theta \right| \leq C_0 \|v\|_{L^2} .$$

Clearly,

$$T(Sv)(\theta) = T\left(\int_0^\theta v(s)ds - \frac{\theta}{2\pi} \mathcal{J}(v) \right) = v(\theta) - \frac{1}{2\pi} \mathcal{J}(v). \tag{2.31}$$

We estimate

$$|\mathcal{J}(v)| \leq \sqrt{2\pi} \|v\|_{L^2}$$

and obtain

$$\|(Sv)'\|_{L^2} = \|TSv\|_{L^2} \leq \|v\|_{L^2} + \frac{1}{2\pi} |\mathcal{J}(v)| \leq C_1 \|v\|_{L^2} .$$

So,

$$\begin{aligned}
\langle Sv, Sv\rangle_{W^1} &= \|Sv\|_{L^2}^2 + \|TSv\|_{L^2}^2 \\
&= \sum\nolimits_{k \neq 0} \left(|\widehat{v}(k)|^2 \frac{1}{k^2} \right) + C_0^2 \|v\|_{L^2}^2 + C_1^2 \|v\|_{L^2}^2 \\
&\leq (1 + C_0^2 + C_1^2) \|v\|_{L^2}^2 .
\end{aligned}$$

Moreover, for $u \in C^1(S^1)$ we have

$$S\,(Tu)\,(\theta) = \int_0^\theta u'(s)ds - \frac{\theta}{2\pi}\int_0^{2\pi} u'(s)ds = u(\theta) - u(0). \tag{2.32}$$

The assertion follows by a density argument.
of (f): Let (u_n) be a sequence in $W^1(S^1)$. We assume that (u_n) and (Tu_n) converge in $L^2(S^1)$. We denote the limits by u, respectively w. To prove that T is closed we have to show (by Exercise 2.39b) that $u \in W^1(S^1)$ and $Tu = w$. From the preceding (e) we obtain

$$\|(u_n - u_m) - (u_n(0) - u_m(0))\,\|_{W^1} = \|(S \circ T)(u_n - u_m)\|_{W^1} \leq C\|T(u_n - u_m)\|_{L^2}\,.$$

This shows that the sequence $(v_n := u_n - u_n(0))$ is convergent in $W^1(S^1)$ to a limit v. So, the sequence$(u_n - v_n = u_n(0))$ has a limit in $L^2(S^1)$, hence in $\mathbb{C}$, and so also in $W^1(S^1)$. This proves that the sequence (u_n) converges in $W^1(S^1)$ to u. So, by (a), Tu_n converges to Tu and $Tu = w$.

Clearly, $\operatorname{Ker}(T) \supseteq \{\text{constants}\}$ and, by (2.32)

$$\operatorname{Ker}(T) \subseteq \operatorname{Ker}(ST) = \{\text{constants}\}.$$

So,

$$\operatorname{Ker}(T) = \{\text{constants}\} = \operatorname{Coker}(T),$$

and hence $\operatorname{index}(T) = 0$.
of (g): Let $(v_n = Tu_n)$ be a sequence in $\operatorname{Im} T$ which converges in $L^2(S^1)$ to a $v \in L^2(S^1)$. Without loss of generality we may assume that $\delta_0(u_n) = 0$ for all $n = 1, 2, \ldots$. According to (e), the operator S is bounded. So, the sequence $(u_n = S(Tu_n) + \delta_0(u_n) = S(Tu_n))$ converges in $W^1(S^1)$ to a $u \in W^1(S^1)$. Then by the continuity of T we obtain $v = Tu$. □

Remark 2.41. We gave a fully detailed proof of the preceding theorem because the 1–dimensional case study provides a preview for the analytical part of this book: each of the statements will be reproved (and the underlying concepts generalized) in Part II by replacing S^1 by an arbitrary n–dimensional compact manifold without boundary (i.e. closed manifold) and $d/d\theta$ by an arbitrary elliptic differential or pseudodifferential operator of order m, acting on sections of a complex vector bundle. More precisely, (a) will be generalized in Theorem 9.2 (p.234), (b) in Exercise 7.2b (p.192), (c) in Theorem 7.13 (p.197), (d) in Theorem 7.14 (p.198), (e) in Theorem 9.7 (p.236), (f) once again in Theorem 9.2 (p.234) and Step (iv) of the *cobordism* proof in Section 13.2 (p. 299), and (g) in Theorem 9.9a (p.236).

In (e), we have

$$(Tu)(\theta) = \frac{1}{\sqrt{2\pi}} \sum\nolimits_{k\in\mathbb{Z}} (ik)\, e^{ik\theta}\, \widehat{u}(k) \text{ and}$$
$$(Sv)(\theta) = \frac{1}{\sqrt{2\pi}} \left(\widehat{v}(0) + \sum\nolimits_{k\neq 0} \frac{-i}{k}\, e^{ik\theta}\, \widehat{v}(k)\right)$$

for $u \in \operatorname{Dom}(T) = W^1(S^1)$ and $v \in \operatorname{Dom}(S) = L^2(S^1)$. This is a special way of writing T and S as *pseudo–differential operators* with *amplitudes*

$$t(\theta, k) = ik \quad \text{and} \quad s(\theta, k) = \frac{-i}{k}\,,$$

respectively, see Chapter 8. We are only interested in the case $k \neq 0$, and note that then t and s are invertible (namely $\neq 0$) and inverses of each other.

On arbitrary closed manifolds we will not have such simple global descriptions of differential and pseudo–differential operators. However, the symbolic calculus we are going to develop below in Chapter 8 will provide us with globally defined principal symbols (= the top order homogenous part of the amplitudes). We will characterize an elliptic operator S by the invertibility of its principal symbol s and construct a parametrix T from the inverse t of s (Section 9.2, p. 236ff).

A closer look at the proof of Assertion (c) reveals that the first Sobolev space $W^1(S^1)$ coincides with the space of absolutely continuous functions on the real line with period 2π (see also [**ReSi72**, p.257 and note p.305]).

Assertion (f) can also be proved in a different way, namely by showing that $T = R^*$ with densely defined R. Actually, $R = -T$. Therefore the index of T must vanish. This argument will be made precise below. The index must also vanish for topological reasons because the dimension of S^1 is odd (see Result (a) in Section 14.3, p.308).

In continuation of Exercise 2.37, p.34, and the preceding Theorem 2.40, we close this Section with an exercise of a *system* of r, $r \in \mathbb{N}$, linear ordinary differential equations of first order on $S^1 = [0, R]/(0 \sim R)$ with $R \in \mathbb{R}, R > 0$ fixed. Like before, one may take $R = 2\pi$ and the angle θ as coordinate. Here, however, we prefer general R and a real x as coordinate for better recalling elementary results for systems of ordinary differential equations. In particular, we denote by $C^\infty(S^1, \mathbb{C}^r)$ the space of smooth r-vector valued functions on $\mathbb{R}$ of period R.

Exercise 2.42. Denote by $\mathrm{gl}(\mathbb{C}, r)$, $r \in \mathbb{N}$, the space of complex $r \times r$ matrices. Let $A : \mathbb{R} \to \mathrm{gl}(\mathbb{C}, r)$ be a smooth mapping with $A(x + R) = A(x)$ for all $x \in \mathbb{R}$. Consider the operator $(\nabla + A)|_{S^1} : C^\infty(S^1, \mathbb{C}^r) \to C^\infty(S^1, \mathbb{C}^r)$, where

$$\nabla := \frac{d}{dx} \oplus \overset{r \text{ times}}{\cdots} \oplus \frac{d}{dx} .$$

a) Show that the operator $(-\nabla + A^*)|_{S^1}$ is *formally adjoint* to $(\nabla + A)|_{S^1}$. Here $A^*(x)$ denotes the adjoint matrix of $A(x)$, $x \in \mathbb{R}$.
b) Denote by $\Phi : \mathbb{C}^r \to C^\infty([0, \infty), \mathbb{C}^r)$ the *fundamental solution* for $\nabla + A$ which assigns a solution $f \in C^\infty([0, \infty), \mathbb{C}^r)$ to each initial value $f(0) \in \mathbb{C}^r$, and define a linear mapping $\phi := \Phi|_{x=R} : \mathbb{C}^r \to \mathbb{C}^r$ by assigning $f(0) \mapsto f(R)$. Show that the periodic solutions correspond to the fixpoints of ϕ.
c) Denote by Ψ the fundamental solution for $-\nabla + A^*$ and define a corresponding linear mapping $\psi : \mathbb{C}^r \to \mathbb{C}^r$. Show that ϕ and $-\psi^{-1}$ are adjoint.
d) Prove the *Index Theorem on the Circle*, namely

$$\dim\ker(\nabla + A)|_{S^1} = \dim\ker(-\nabla + A^*)|_{S^1}, \quad \text{i.e.,} \quad \operatorname{index}(\nabla + A)|_{S^1} = 0.$$

[Hint: To a) Define a Hermitian inner product on $\mathbb{C}^r$ by

$$\langle u, v\rangle := \sum_{i=1}^{r} u_i \overline{v_i} \quad \text{for } u = (u_1, \dots), v = (v_1, \dots).$$

Then deduce that the operators $(\nabla + A)|_{S^1}$ and $(-\nabla + A^*)|_{S^1}$ are formally adjoint from

$$\int_0^R \langle(\nabla + A)f, g\rangle dx - \int_0^R \langle f, (-\nabla + A^*)g\rangle dx = \langle f(R), g(R)\rangle - \langle f(0), g(0)\rangle \overset{!}{=} 0$$

for all $f, g \in C^\infty(S^1, \mathbb{C}^r)$.
To b) Clearly a periodic solution yields a fixpoint of ϕ. A delicate argument is needed to show that each fixpoint yields a *smooth* periodic solution.
To c) To begin with, consider two functions f, g on $\mathbb{R}$ with values in $\mathbb{C}^r$ and two points $x_0 < x_1$. If $(\nabla + A)f = 0$ and $(-\nabla + A^*)g = 0$, you get

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

Put $x_0 = 0$ and $x_1 = R$ and re-write

$$\langle f(R), g(R)\rangle = \langle \phi(f(0), g(R)\rangle \quad \text{and} \quad \langle f(0), g(0)\rangle = \langle f(0), \psi^{-1} g(R)\rangle.$$

To d) 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. Notice that, according to changes of A, the dimension of the space of periodic solutions varies and takes values between 0 and r.]

NOTE. Alternatively, we could apply the heavy analysis tools of our Part II to get the mentioned Index Theorem on the Circle. *First* we would determine the principal symbol of $(\nabla + A)|_{S^1}$, e.g., by the formula of Exercise 6.34, p. 180. Our data are $x_0 \in \mathbb{R}$, $\xi \in T^*_{x_0} S^1$, and $e \in \mathbb{C}^r$. We represent $\xi = df|_{x_0}$ by a function f vanishing at x_0 like, e.g., $f(x) = \xi(x - x_0)$, and extend e to a vector-valued function g, e.g., the constant function $g := e\mathbf{1}$ on S^1. Then we find

$$\begin{aligned}
\sigma_1(\nabla + A)(x_0, \xi)(e) &= i(\nabla + A)(fg)|_{x_0} \\
&= i\nabla(\xi(x - x_0)e\mathbf{1})|_{x_0} + iA(x)(\xi(x - x_0)e\mathbf{1})|_{x_0} \\
&= i\xi e + iA(x_0)(0e) = i\xi e = (i\xi I^r)e.
\end{aligned}$$

Here $I^r \in \mathrm{GL}(r, \mathbb{C})$ denotes the identity. Whence, $\sigma_1(\nabla + A)(x_0, \xi) = i\xi I^r$ is invertible for $\xi \neq 0$, *Next*, we conclude that $(\nabla + A)|_{S^1}$ is an elliptic differential operator on a closed manifold and has finite index which depends only on the homotopy type of the principal symbol. All that will be explained in Part II. *Finally*, we notice that the choice of A does not inflict the principal symbol. We choose $A = 0$ constant and obtain $\operatorname{index} \nabla = 0$. That suffices. As mentioned above in a similar situation, this is in fine agreement with the basic topological insight that *any* homogeneous polynomial elliptic symbol over an odd-dimensional manifold can be deformed into the identity within the class of elliptic symbols.

Closed (not necessarily bounded) Fredholm Operators. It is easy to generalize the concept of *Fredholm operators* to the unbounded case.

DEFINITION 2.43. Let H be a complex separable Hilbert space. A linear (not necessarily bounded) operator F with domain $\mathrm{Dom}(F)$, null space $\mathrm{Ker}(F)$, and range $\mathrm{Im}(F)$ is called **Fredholm** if the following conditions are satisfied.

(i) $\mathrm{Dom}(F)$ is dense in H.
(ii) F is closed.
(iii) Both $\dim \mathrm{Ker}(F)$ and $\dim \mathrm{Coker}(F)$ are finite. The difference of the dimensions is called $\mathrm{index}(F)$.

Then it follows that the range $\mathrm{Im}(F)$ of F is a closed subspace of H and that $\dim \mathrm{Ker}(F^*)$ is finite. So, a closed operator F is characterized as a Fredholm operator by the same properties as in the bounded case (see also Exercise 2.39a):

(iv) For any arbitrary $v \in H$ the equation $Fu = v$, $v \in \mathrm{Dom}(F)$, is solvable if and only if v is orthogonal to every solution w of $F^* w = 0$: differently speaking, $\mathrm{Im}(F)$ is closed.

(v) Each of the equations $Fu = 0$, $F^* w = 0$ has only finitely many linearly independent solutions.

Condition (iv) can be replaced by

(iv') $\|Fu\| \geq C\|u\|$ for all $u \in \mathrm{Ker}(F)^\perp \cap \mathrm{Dom}(F)$ with a positive constant C.

We more generally show

Lemma 2.44. *Let $T : \mathrm{Dom}(T) \to H$ be a closed (by our definition densely defined) operator. Then T has closed range if and only if* (iv') *is valid.*

Proof. Let first condition (iv') be satisfied. Let $(f_n)_{n=1,2,\dots}$ be a sequence that converges to f (shortly "$f_n \to f$") with $f_n \in \mathrm{Im}(T)$. Choose a sequence $(u_n \in \mathrm{Dom}(T) \cap \mathrm{Ker}(T)^\perp)$ with $Tu_n = f_n$. We get $u_n - u_m \in \mathrm{Ker}(T)^\perp$ for all n, m, and thus by (iv')

$$\|u_n - u_m\| \leq \frac{1}{C}\|f_n - f_m\| \longrightarrow 0, \qquad n, m \longrightarrow 0.$$

Accordingly, $u_n \to u$ and $Tu_n \to f$ which implies that $u \in \mathrm{Dom}(T)$ and $Tu = f$, i.e., $f \in \mathrm{Im}(T)$, since T is closed. Thus $\mathrm{Im}(T)$ is closed.

Now let $\mathrm{Im}(T)$ be closed. Then the mapping

$$\begin{array}{ccc} \mathrm{Dom}(T) \cap \mathrm{Ker}(T)^\perp & \longrightarrow & \mathrm{Im}(T) \\ u & \mapsto & Tu \end{array}$$

is surjective and injective for simple reasons. Moreover, both spaces are closed subspaces of H, and thus can be interpreted as Hilbert spaces. The inverse of this mapping exists and is a well-defined linear transformation of $\mathrm{Im}(T)$ into $\mathrm{Ker}(T)^\perp$, with domain $\mathrm{Im}(T)$. Also this transformation is a closed linear operator of the Hilbert space $\mathrm{Im}(T)$ into the Hilbert space $\mathrm{Ker}(T)^\perp$, as follows from the fact that T is closed. Whence it must be a bounded linear transformation by standard argument. However, this boundedness clearly amounts to the condition (iv'). □

Moreover, as in Theorem 2.10, F is a Fredholm operator if and only if F^* is a Fredholm operator (proving the closedness of $\mathrm{Im}(F^*)$ looks demanding, but follows rather directly from Lemma 2.44), and (clearly) we have

$$\operatorname{index} F = \dim \operatorname{Ker} F - \dim \operatorname{Ker} F^* = -\operatorname{index} F^*.$$

In particular, $\operatorname{index} F = 0$ in case F is self–adjoint (i.e., $F^* = F$, see below).

The composition of (not necessarily bounded) Fredholm operators yields again a Fredholm operator. More precisely, we have the following composition rule.

Theorem 2.45 (I.Z. Gohberg and M.G. Krein, 1957). *If F and G are (not necessarily bounded) Fredholm operators then their product GF is densely defined with*

$$\mathrm{Dom}(GF) = \mathrm{Dom}(F) \cap F^{-1}\,\mathrm{Dom}(G) \tag{2.33}$$

and is a Fredholm operator. Moreover,

$$\operatorname{index} GF = \operatorname{index} F + \operatorname{index} G\,. \tag{2.34}$$

REMARK 2.46. The proof is considerably more involved than in the bounded case. Why so? Can't we consider the operators F, G as *bounded* Fredholm operators in the graph norm of (2.23)? Yes. And then derive that the composition is Fredholm and satisfies (2.34) from our previous product rule for bounded Fredholm operators in Exercise 1.10? Yes, indeed. The delicate claim, however, is that the domain $\mathrm{Dom}(GF)$ of the composition is *dense* in H.

In the context of index theory, the product of closed unbounded Fredholm operators shows up

(1) for elliptic differential operators over a closed (i.e., compact and without boundary) manifold M and
(2) for elliptic differential operators over a compact manifold with smooth boundary, subject to some *regular* (*elliptic*) boundary conditions.

The first case was touched upon in Exercise 2.37, p.34f and Theorem 2.40, p.37 with $M = S^1$ and is the subject of the main body of this monograph for arbitrary, more general and more specific closed M. As we shall see in Part II, an elliptic differential operator F of order $k \geq 1$, acting between sections of vector bundles E_1 and E_2 over M can be considered both as a closed unbounded operator from $L^2(E_1)$ to $L^2(E_2)$ with domain $\mathrm{Dom}(F) = W^k(E_1) \subset L^2(E_1)$ or as an indexed family of bounded operators $\{F_r : W^r(E_1) \to W^{r-k}(E_2)\}$ for all real r, where $W^r(E_1)$ denotes the rth Sobolev space (see Chapters 7 and 9). Combining F with a second elliptic differential operator G, say of order m, with $\mathrm{dom}(G) = W^m(E_2) \subset L^2(E_2)$ yields the domain

$$\begin{aligned}\mathrm{Dom}(GF) \;&=\; \mathrm{Dom}(F) \cap F^{-1}\,\mathrm{Dom}(G)\\ &= W^k(E_1) \cap F^{-1}(W^m(E_2)) \;=\; W^k(E_1) \cap W^{m+k}(E_1) \;=\; W^{m+k}(E_1),\end{aligned}$$

which is dense in $L^2(E_1)$ by definition. Here we exploited that the Sobolev spaces can be defined via elliptic operators (along the lines of Exercise 7.2b, p. 192f). Whence, in view of our set-up of Part II, there is nothing surprising in the preceding theorem for elliptic operators on closed manifolds. Even in that case, however, the result is not trivial.

The second case is much more intricate: We have to deal with a boundary condition for the combined elliptic differential operator obtained by combining two possibly radically different *regular* (*elliptic*) boundary conditions. Solely by classical analysis arguments it might be difficult to prove the regularity of the combined boundary condition.

Surprisingly, we can prove the density of $\mathrm{Dom}(GF) \subset H$ by purely (but, admittedly, somewhat wired) functional analysis arguments, following [**GK57**] and [**CorLab**, Lemma 2.3 and Theorem 2.1].

We prepare the proof by a series of small lemmata.

LEMMA 2.47. *Let D be a dense subspace of H and M_n a closed subspace of H of finite codimension $n \in \mathbb{N}$. Then there exists a bounded idempotent (= not necessarily orthogonal projection) P, such that*

$$\mathrm{Im}(P) = M_n, \qquad \dim \mathrm{Ker}(P) = n, \qquad \mathrm{Ker}(P) \subset D.$$

Recall from Remark 2.2, p.13, that subspaces of finite codimension are not necessarily closed.

PROOF. To begin with, choose a basis $u_1, \dots, u_n$ of $M_n^\perp$. Then select $v_1, \dots, v_n \in D$ sufficiently close to the start basis (say $\|u_i - v_i\| < \delta$ for all $i = 1, \dots, n$ for sufficiently small $\delta > 0$), such that

$$\det\bigl(\langle u_i, v_j\rangle\bigr)_{i,j=1,\dots,n} \neq 0. \tag{2.35}$$

In particular, this implies linear independence of $v_1, \dots, v_n$. Let Q denote the space spanned by $v_1, \dots, v_n$. Clearly $Q \subset D$. Then for every $x \in H$ there exists a unique decomposition $x = w + v$ with $w \in M_n$ and $v \in Q$. Indeed, we simply set $v := \sum_{i=1}^n \alpha_i v_i$ where the coefficients α_i are uniquely determined by the orthogonality relations

$$\bigl\langle x - \sum_{i=1}^n \alpha_i v_i\,,\, u_j\bigr\rangle = 0, \qquad j = 1, \dots, n.$$

Note that the determinant of the preceding system of linear equations does not vanish, by (2.35). Now define $Px := w$. This defines a projection operator satisfying all claims stated above. In particular, P is bounded with

$$\|Px\| = \|w\| = \|x - v\| \le \|x\| + \|v\| = \|x\| + C\sum |\alpha_i| \le (1 + C')\|x\|. \quad \square$$

The claim of following lemma becomes wrong, if we drop the assumption of finite codimension and closedness of M. Take, for instance, $H := L^2([a, b])$, D a dense subset of step functions, and $M_\infty := C^0([a, b])$. For finite codimension of M, the lemma seems obvious. However, to prove it rigorously we depend on the preceding result.

LEMMA 2.48. *Under the assumptions of Lemma 2.47, the intersection $M_n \cap D$ is dense in M_n.*

PROOF. For given $u \in M_n$ and given approximation radius $\delta > 0$, we set $\delta' := \delta\bigl(1 + \|\mathrm{Id} - P\|\bigr)^{-1}$ and select an $x \in D$ such that $\|u - x\| \le \delta'$. As before, we decompose $x = w + v$ with $w \in M_n$ and $v \in Q \subset D$, whence $w = Px \in D$, and

$$\begin{aligned}\|u - w\| &\le \|u - x\| + \|(\mathrm{Id} - P)x\| \\ &= \|u - x\| + \|(\mathrm{Id} - P)(x - u)\| \le \bigl(1 + \|\mathrm{Id} - P\|\bigr)\delta' = \delta. \quad \square\end{aligned}$$

Now we can draw the decisive consequence for the composition of (not necessarily bounded) Fredholm operators in Hilbert space.

LEMMA 2.49. *If F and G are (not necessarily bounded) Fredholm operators then their product GF is densely defined and Fredholm.*

PROOF. **1**. We first show that $\mathrm{Dom}(GF) = \{u \in \mathrm{Dom}(F) : Fu \in \mathrm{Dom}(G)\}$ (as defined in (2.33)) is dense in H. Since by assumption

$$\dim \mathrm{Ker}(F^*) = \dim \mathrm{Im}(F)^\perp < \infty,$$

the space $\mathrm{Im}(F) \cap \mathrm{Dom}(G)$ is dense in $\mathrm{Im}(F)$ by the preceding lemma. Let now $u \in \mathrm{Dom}(F)$, then there exists $\widetilde{u} \in \mathrm{Ker}(F)^\perp \cap \mathrm{Dom}(F)$ with $F\widetilde{u} = Fu$, and we have $u - \widetilde{u} \in \mathrm{Ker}(F) \subset \mathrm{Dom}(GF)$. Since $\mathrm{Im}(F) \cap \mathrm{Dom}(G)$ is dense in $\mathrm{Im}(F)$, for every $\delta' > 0$ there exists a $v \in \mathrm{Im}(F) \cap \mathrm{Dom}(G)$ such that $\|v - Fu\| < \delta'$. Let $v = Fw$ with $w \in \mathrm{Ker}(F)^\perp$, then $w \in \mathrm{Dom}(GF)$. We get $\widetilde{u} - w \in \mathrm{Ker}(F)^\perp$, and thus $\|\widetilde{u} - w\| \le C\|Fu - v\| \le C\delta'$. Set $u' := w + u - \widetilde{u}$, and choose δ' such that $C\delta' = \delta$, then we have $\|u - u'\| < \delta$ and $u' \in \mathrm{Dom}(GF)$, which means that $\mathrm{Dom}(GF)$ is dense in $\mathrm{Dom}(F)$. Since $\mathrm{Dom}(F)$ is dense in H, we see that $\mathrm{Dom}(GF)$ is dense in

H.

2. Next we show that GF is closed. Suppose the sequence $(u_n)_{n=1,2,\ldots}$ converges to u (once again, shortly "$u_n \to u$") and $GFu_n \to v$. We decompose $Fu_n = w_n + z_n$ with $w_n \in \operatorname{Ker}(G)^\perp$ and $z_n \in \operatorname{Ker}(G)$. By the closedness of $\operatorname{Ker}(G)^\perp = \operatorname{Im}(G^*)$ (see the reformulation (iv') above), we have $w_n \to w \in \operatorname{Ker}(G)^\perp$. Also *either* $z_{n_k} \to z \in \operatorname{Ker}(G)$ for a suitable subsequence (z_{n_k}) *or* $\|z_n\| \to \infty$, since $\dim \operatorname{Ker}(G) < \infty$. In the first case we get $u_{n_k} \to u$, $Fu_{n_k} \to w + z$, $GFu_{n_k} \to v$, i.e., $u \in \operatorname{Dom}(GF)$ and $GFu = v$, since F and G are closed operators. In the second case, let $x_n := u_n/\|z_n\|$, then

$$x_n \longrightarrow 0, \qquad Fx_n = \frac{w_n}{\|z_n\|} + \frac{z_n}{\|z_n\|}, \qquad GFx_n \longrightarrow 0.$$

But $z_n/\|z_n\|$ must have a convergent subsequence, and we must get $x_{n_k} \to 0$, $Fx_{n_k} \to w$ with $\|w\| = 1$, which is a contradiction, since F is closed. This proves that GF is closed.

3. Now we show that GF satisfies condition (iv') (before Lemma 2.44, p.43). Indeed, let $u_n \in \operatorname{Dom}(GF) \cap \operatorname{Ker}(GF)^\perp$, $\|u_n\| = 1$, and $GFu_n \to 0$. Then we again write $Fu_n = w_n + z_n$ with $w_n \in \operatorname{Ker}(G)^\perp$ and $z_n \in \operatorname{Ker}(G)$. We get $w_n \to 0$ by (iv') for G, and again *either* $z_{n_k} \to w$ *or* $\|z_n\| \to \infty$. In the first case we get $GFu_{n_k} \to 0$, $Gu_{n_k} \to w$, and $u_{n_k} \in \operatorname{Ker}(F)^\perp$, i.e., $u_{n_k} \to u$, $\|u\| = 1$, $u \in \operatorname{Dom}(GF) \cap \operatorname{Ker}(GF)^\perp$, and $GFu = 0$, a contradiction. In the second case, set again $x_n := u_n/\|z_n\|$. The sequence (Fx_n) must have a convergent subsequence, and thus we get

$$x_{n_k} \longrightarrow 0, \qquad Fx_{n_k} \longrightarrow w, \qquad \|w\| = 1$$

a contradiction, because F is closed. This proves (iv') for GF.

4. Finally, it is clear that

$$\dim \operatorname{Ker}(GF) \le \dim \operatorname{Ker}(F) + \dim \operatorname{Ker}(G) < \infty$$

and

$$\begin{aligned} \operatorname{codim} \operatorname{Im}(GF) &= \dim \operatorname{Ker}((GF)^*) = \dim \operatorname{Ker}(F^*G^*) \\ &\le \dim \operatorname{Ker}(F^*) + \dim \operatorname{Ker}(G^*) < \infty. \end{aligned}$$

Here we apply that F^*G^* is also closed with closed range (by the same arguments) and we have $(GF)^* = F^*G^*$. □

Proof of Theorem 2.45. The preceding lemma yields the delicate result, namely that GF is closed with closed range and that the in (2.33) defined $\operatorname{Dom}(GF)$ is dense in H. We leave it to the reader to count the dimensions for the proof of (2.34), respectively refer to [**CorLab**, p.699] for the details of that counting. □

Exercise 2.50. Find out which of the following operators are (not necessarily bounded) Fredholm operators in the sense of Definition 2.38.

a) (Bounded) Fredholm operators in the sense of Section 1.2?

b) Operators of finite rank?

c) The multiplication operator M_{id} of Exercise 2.39f for $\lambda_j := j$, $j = 1, 2 \ldots$?

d) The operator T which extends (to $W^1(S^1)$) the differentiation operator $d/d\theta$ on $C^1(S^1)$ (Exercise 2.37)?

e) The Laplace operator Δ on the unit disk in the plane (see Exercise 5.9, p. 144).

[Answer: (a) Yes. (b) Never. (c) Yes: for

$$\operatorname{Dom}(M_{\mathrm{id}}) := \left\{ \sum\nolimits_{j=1}^{\infty} c_j u_j \;:\; \sum\nolimits_{j=1}^{\infty} j^2 |c_j|^2 < +\infty \right\}$$

you obtain a densely defined closed operator which is injective and surjective. (d) Yes, if $H = L^2(S^1)$ and the domain of T is taken to be $W^1(S^1) \subset L^2(S^1)$ (as in Theorem 2.40a). (e) Yes, if the domain is extended to the second Sobolev space and restricted by elliptic boundary conditions (see Exercise 5.9, p. 144). If the full space of smooth functions on the disc is taken as the domain of the Laplacian and no boundary conditions are imposed, the kernel of the Laplacian is infinite-dimensional, consisting of all harmonic functions on the disk.]

Symmetric and Self–Adjoint Operators. As seen immediately after Definition 2.43, the index of unbounded self–adjoint Fredholm operators vanishes like in the bounded case. So, there is no immediate *index problem.* However, spectral projections of self-adjoint Fredholm operators defined by the Spectral Theorem 2.61 (below on p.51) are of high interest in index theory, in particular for specifying elliptic boundary value problems. As a matter of fact, the most interesting operators of geometry and gauge–theoretical physics are Laplacians or Dirac type operators which are symmetric operators.

Index problems arise from symmetric operators in two ways:

(1) We may occasionally split a self–adjoint Fredholm operator P into the direct sum of *chiral* non–symmetric components

$$P = \begin{pmatrix} 0 & P^- \\ P^+ & 0 \end{pmatrix}$$

with $P^- = (P^+)^*$ and investigate the index of P^+. Since

$$\operatorname{Ker}(P) = \operatorname{Ker}(P^+) \oplus \operatorname{Ker}(P^-) \text{ and } \operatorname{Ker}(P^-) \cong \left(\operatorname{Im}(P^+)\right)^{\perp},$$

the integer $\operatorname{index}(P^+) = n_+ - n_-$ gives the *chiral asymmetry* of $\operatorname{Ker}(P)$; here $n_\pm := \dim \operatorname{Ker}(P^\pm)$. This is also the way one follows in most geometric and topological applications (e.g., when determining the Euler characteristic or the signature of even–dimensional closed Riemannian manifolds, see Part III, Section 14.4, p. 309ff).

(2) Index problems arise from symmetric operators directly on compact manifolds with boundary when symmetric elliptic operators are considered and non–self–adjoint elliptic boundary conditions are imposed (see e.g. [**BoWo**, Theorem 22.24]).

For these perspectives and also for the completeness of our presentation, we summarize the basic knowledge of symmetric and self–adjoint unbounded operators and introduce to the corresponding Fredholm theory.

Definition 2.51. a) We say that a densely defined operator S in H is **symmetric** (or **formally self–adjoint**) if

$$\langle Su, v\rangle = \langle u, Sv\rangle, \qquad u,\, v \in \operatorname{Dom}(S).$$

b) Recall that a densely defined operator S in H is called **self–adjoint** if $S^* = S$ (in the sense of Definition 2.38c).
c) A densely defined symmetric operator S in H is called **essentially self–adjoint** if its closure is self–adjoint, i.e. if $\overline{S} = S^*$.

We give an interesting criterion for proving that a symmetric operator is essentially self–adjoint. See also and similarly, [**ReSi72**, Theorem VIII.3, Corollary, p. 257] and, differently, [**Miz**, Lemma 8.14] – Mizohata requires dense $\operatorname{Im}(T)$ in H and the existence of a positive constant a such that $\|Tu\| \geq a\|u\|$ for all $u \in \operatorname{Dom}(T)$.

LEMMA 2.52. *Let T be a densely defined symmetric operator in a separable complex Hilbert space H. We assume that both* $\mathrm{Im}(T+i)$ *and* $\mathrm{Im}(T-i)$ *are dense in H. Then T is essentially self–adjoint.*

PROOF. First we show that $\mathrm{Im}(\overline{T}\pm i)=H$. Let (u_n) be a sequence in $\mathrm{Dom}(\overline{T})$ and $(\overline{T}\pm i)u_n$ converges to v_0. As with T, $\overline{T}$ is also symmetric, and we have

$$\|(\overline{T}\pm i)u\|^2=\|u\|^2+\|\overline{T}u\|^2 \text{ for all } u\in \mathrm{Dom}(\overline{T}\pm i)=\mathrm{Dom}(\overline{T}),$$

since the mixed terms $\langle \pm iu, \overline{T}u\rangle$ and $\langle \overline{T}u, \pm iu\rangle$ cancel each other by symmetry. Hence

$$\|u\|\leq \|(\overline{T}\pm i)u\|\,,$$

i.e., $\overline{T}\pm i$ injective and $(\overline{T}\pm i)^{-1}$ well defined on $\mathrm{Im}(\overline{T}\pm i)$ and bounded. We conclude that (u_n) converges to some u_0 and $\overline{T}u_n$ converges too. Since $\overline{T}$ is closed, $u_0\in\mathrm{Dom}(\overline{T})$ and $(\overline{T}\pm i)u_0=v_0$. Thus $\mathrm{Im}(\overline{T}\pm i)$ is closed, so $\mathrm{Im}(\overline{T}\pm i)=H$.

Now we show that $\overline{T}$ is self–adjoint. By definition, $\overline{T}$ is the minimal closed extension, so $\overline{T}\subseteq T^*$. Therefore it suffices to show that $\mathrm{Dom}(\overline{T}^*)\subseteq\mathrm{Dom}(\overline{T})$. Then, let $u\in\mathrm{Dom}(\overline{T}^*)$. Since $\mathrm{Im}(\overline{T}+i)=H$, there is a $w\in\mathrm{Dom}(\overline{T})$ so that $(\overline{T}+i)w=(\overline{T}^*+i)u$. But $\mathrm{Dom}(\overline{T})\subseteq\mathrm{Dom}(\overline{T}^*)$, so $u-w\in\mathrm{Dom}(\overline{T}^*)$ and

$$(\overline{T}^*+i)(u-w)=0\,.$$

Since $\mathrm{Im}(\overline{T}-i)=H$, we have $\mathrm{Ker}(\overline{T}^*+i)=\{0\}$, so $u=w\in\mathrm{Dom}(\overline{T})$. □

From the first part of the preceding proof we can distil

COROLLARY 2.53. *Let T be a closed, injective operator. If T is bounded from below by the identity, then T is a semi–Fredholm operator.*

Here we used the following notation:

DEFINITION 2.54. a) A densely defined closed operator T is called a **semi–Fredholm operator** if and only if $\mathrm{Im}(T)$ is closed and either $\mathrm{Ker}(T)$ or $\mathrm{Coker}(T)$ has finite dimension.
b) A densely defined, symmetric operator T is **bounded from below by the identity** if we have

$$\|u\|^2\leq\langle Tu,u\rangle \qquad \text{for all } u\in\mathrm{Dom}(T),$$

and consequently $\|u\|\leq\|Tu\|$.

We apply the lemma for $H=L^2(S^1)$ and take for T the operator $id/d\theta$ with domain $C^\infty(S^1)$. Integration by parts shows that $id/d\theta$ is symmetric. For $e_k(\theta):=\frac{1}{\sqrt{2\pi}}e^{ik\theta}$ $(k\in\mathbb{Z})$, we have $(T\pm i)(e_k)=(-k\pm i)\,e_k$ (or $(T\pm i)(\frac{1}{-k\pm i}e_k)=e_k$) and so $\mathrm{Im}(T\pm i)$ is dense, since it contains the linear span of the complete orthonormal system $\{e_k : k\in\mathbb{Z}\}$. Thus we have proved

THEOREM 2.55. *The operator $id/d\theta$ in $L^2(S^1)$ with domain $C^\infty(S^1)$ is symmetric and essentially self–adjoint.*

We have shown in Theorem 2.40a that $d/d\theta$ on $C^1(S^1)$ extends to a continuous operator $T: W^1(S^1)\to L^2(S^1)$. Regarding T as an unbounded operator on $L^2(S^1)$ with dense domain $W^1(S^1)$, we have already shown that T is closed (see Theorem 2.40f). Indeed, T is the closure of $d/d\theta$ with domain $C^1(S^1)\subset L^2(S^1)$. Applying the preceding theorem yields

COROLLARY 2.56. *On the circle S^1, the unbounded operator T on $L^2(S^1)$ (which extends $id/d\theta$ on $C^1(S^1)$) with domain $W^1(S^1)$ is self–adjoint.*

REMARK 2.57. The arguments of the preceding proof can be generalized to prove the well–known but seldom explicitly stated fact that each symmetric elliptic differential operator on a closed Riemannian manifold (or, more generally, on a complete Riemannian manifold) is essentially self–adjoint (see though [**Shu**, Theorem 8.3], [**BlBo03**, Lemma 3.23] and, implicitly, [**Gi95**, Lemma 1.6.3] or, differently and carried through only for the Laplacian, [**Tay2**, Proposition 8.2.4]).

Note the remarkable contrast with the case where the underlying manifold has a boundary. Then there is a huge variety of domains to which a fixed symmetric elliptic differential operators can be extended such that it becomes self–adjoint; and there is a smaller, but still large variety where the extension becomes self–adjoint *and* Fredholm. For the case of a 1–dimensional manifold (= the interval) see below Exercise 2.58g–h. For the higher dimensional case see [**BoFu1**, Section 3] and, differently, [**BoWo**, Chapter 20].

In [**BoFu2**, Proposition 7.15] it is shown that the set of *all* extensions of a given symmetric elliptic differential operator T_0 of first order over a compact smooth Riemannian manifold M with boundary Σ can be naturally identified with a dense graded subspace $\beta(T_0)$ of the distribution space $W^{-1/2}(\Sigma)$. It turns out that $\beta(T_0)$ carries a natural structure of a *symplectic* Hilbert space, i.e., a (real) Hilbert space with a skew-symmetric, nondegenerate and bounded bilinear form ω, here induced from the principal symbol of T_0 over Σ in normal direction.[2] Then all self–adjoint extensions of T_0 (in the underlying L^2–space) correspond to the *Lagrangian* subspaces of $\beta(T_0)$ and the self–adjoint Fredholm extensions correspond to the Lagrangian subspaces which form a *Fredholm pair* with the canonical Lagrangian subspace (the *Cauchy data space*).

A classical source to the systematic study of all self–adjoint extensions of a given symmetric operator is the célèbre paper [**Neu**]. For supplementary studies, involving also the deficiency indices (i.e. the codimension of the range of $T \pm i$) and relations to index theory, we refer to [**AkGl**, Section 78], [**GK57**], and, more recently, [**Les**, Chapter 4].

EXERCISE 2.58. a) Show that a densely defined operator S in H is symmetric if and only if $S \subseteq S^*$. So, in particular, each self–adjoint operator is symmetric.
b) Prove another criterion for S being symmetric, namely that $\langle Su, u\rangle \in \mathbb{R}$ for every $u \in \mathrm{Dom}(S)$.
c) Let S be a symmetric operator in H. Show that the two usual conditions (i.e., $\overline{S} = \overline{S}^*$ and $\overline{S} = S^*$) for S being essentially self–adjoint are equivalent.
d) Let S be a densely defined operator in H which is essentially self–adjoint. Show that then $\overline{S}$ is the only self–adjoint extension of S. Show that the converse is also true, i.e., if S has one and only one self–adjoint extension, then S is essentially self–adjoint.
e) Let S be a densely defined, symmetric operator in H. Show that S is self–adjoint if and only if $S \pm i$ are both surjective operators.
f) Show that the operator M_{id} of Exercise 2.50c is self–adjoint.

[2] The concept of *symplectic manifolds* is mentioned in Remark 6.9, p.163, and introduced rigorously in Definition 19.21, p.619. They are one of the major objects of Seiberg-Witten Theory.

g) Show that the operator id/dx in $L^2([0,1])$ with domain $C_0^\infty([0,1])$ (= smooth functions with support in the interior $(0,1)$ of the interval) is symmetric but not essentially self–adjoint.
h) Let P, Q be continuous real–valued functions on the interval $[0,1]$. Let $u \mapsto u'$ denote the differentiation. Show that the operator

$$u \mapsto (Pu')' + Qu$$

in $L^2([0,1])$ with domain $C_0^\infty([0,1])$ is symmetric but not essentially self–adjoint.
i) Consider the multiplication operator M_{id} $(u(x) \mapsto xu(x))$ and the differential operators of (g) and (h) on $L^2(\mathbb{R})$ with domain $C_0^\infty(\mathbb{R})$ (= smooth functions with compact support). Show that these operators are essentially self–adjoint.
[Hint: For (a) use (2.22). For (b) establish

$$4\langle u, v\rangle = \sum_{k=0}^{3} i^k \langle u + i^k v, u + i^k v\rangle$$

by straightforward calculation. For (c) recall that always $S \subseteq T$ implies $T^* \subseteq S^*$; that every symmetric operator S is closable, because S is densely defined and $S \subseteq S^*$; and that $\overline{S} = S^{**}$ for each closable S. For (d) note first that every symmetric operator is closable because S is densely defined and $S \subseteq S^*$, whence $S \subseteq \overline{S} \subseteq S^*$. Then use that $S \subseteq R$ implies $R^* \subseteq S^*$. A safety net for (e) is provided e.g. in [**Ped**, Proposition 5.2.5], see also the second part of our proof of Lemma 2.52. (f) can be proved directly or by applying (e).
For (g) and (i) recall that the **support** of a function $u \in C^0(X)$ is by definition the smallest closed subset of X outside which u vanishes identically: $\operatorname{supp} u := \overline{\{z \in X : u(z) \neq 0\}}$. Here X is a topological space, $C^0(X)$ denotes the set of complex valued continuous functions on X, and $\overline{L}$ denotes the closure of $L \subset X$.
In (g) and (h) the symmetry follows from partial integration. Candidates for non–uniquely determined self–adjoint extensions are provided in (g) by the unitarily twisted periodic boundary conditions $u(1) = e^{i\varphi}u(0)$ for all $\varphi \in [0, 2\pi)$; and in (h) e.g. by the Dirichlet boundary condition $u(0) = u(1) = 0$ and the Neumann boundary condition $u'(0) = u'(1) = 0$. For(i) see also the extensive discussion in [**AkGl**, Sections 49, 77].]

Spectral Theory. The following definition extends the notion of resolvent set and spectrum commonly defined for elements in $\mathcal{B}(H)$; cf. Table 1.1, p. 11.

DEFINITION 2.59. a) For an operator T in H we define the **resolvent set** $\operatorname{Res}(T)$ as the set of all $\lambda \in \mathbb{C}$ for which the operator $T - \lambda \operatorname{Id}$ is bijective from $\operatorname{Dom}(T)$ onto H with bounded inverse.
b) The complement of the resolvent set is called the **spectrum** of T, and is denoted by $\operatorname{Spec}(T)$.
c) The function $R(\lambda) := (T - \lambda \operatorname{Id})^{-1}$ defined on $\mathbb{C} \setminus \operatorname{Spec}(T)$ with values in $\mathcal{B}(H)$ is called the **resolvent function**.

EXERCISE 2.60. a) Show that the resolvent set of any operator in H is open.
b) Let T be a closed operator in H and $\lambda_0 \in \mathbb{C}$. Assume that the resolvent $R(\lambda_0)$ exists and is compact. Show that then the spectrum of T consists entirely of countably many isolated eigenvalues with finite multiplicities and without finite accumulation point, and the resolvent $R(\lambda)$ is compact for all $\lambda \in \operatorname{Res}(T)$. If T is

self–adjoint, conclude that T is diagonalizable (= *discrete*), i.e., the eigenvectors form a basis of $\mathrm{Dom}(T)$.
c) Prove that the spectrum of a self–adjoint operator in H is a nonempty, closed subset of $\mathbb{R}$.
d) Prove the following spectral characterization of (not necessarily bounded) self–adjoint Fredholm operators: A self–adjoint operator T has discrete spectrum of finite multiplicity in a neighborhood of 0 $\iff$ $\mathrm{Ker}(T)$ is finite–dimensional and $\mathrm{Im}(T)$ is closed.
e) Let S be a self–adjoint operator in H with compact resolvent. Show that S and each bounded self–adjoint perturbation (i.e. operators of the form $S+C$ where $C \in \mathcal{B}(H)$ and $C = C^*$) are Fredholm operators in the sense of Definition 2.38.
[Hint: For (a) cf. [**Ped**, Proposition 5.2.11] or [**Rud**, Exercise 13.17, p.365]. For (b) cf. [**Kat**, Theorem III.6.29]; see also our Section 2.2 on compact operators (p. 18ff), in particular Theorem 2.20. See also our discussion of the Green's function for the Sturm-Liouville problems in the preceding section. Note that the eigenfunctions of the Green's operator and of the Sturm-Liouville problem coincide, while the eigenvalues of the Green's operator are the reciprocals of the eigenvalues of the Sturm-Liouville problem (in that case 0 is not an eigenvalue); see [**CodLev**, p.194]. For (c) cf. [**Kat**, Section V.3.5] or [**Ped**, Proposition 5.2.13]. For (d) cf. [**DunSch**, Def. XIII.6.1 and Thm. XIII.6.5]. Dunford and Schwartz show that if T is self–adjoint, then λ is an isolated point of the spectrum of T if and only if $\mathrm{Im}(T-\lambda)$ is closed. (Note that [**DunSch**] define the essential spectrum differently). For (e) apply (c) and (d).]

For a normal (not necessarily bounded) operator, there is a famous theorem which expresses the operator as an integral of the coordinate function over the operator's spectrum with respect to a projection-valued measure. It supplements the previously proven spectral decomposition for compact operators, Theorem 2.20, p.20. We shall formulate it only for self-adjoint operators.

THEOREM 2.61 (Spectral Theorem). *Let T be a densely defined self-adjoint operator in a complex separable Hilbert space H. Then there exists a uniquely determined spectral measure E on the Borel subsets of $\mathbb{R}$ such that*

(1) $T = \int_{\lambda \in \mathrm{Spec}(T)} \lambda \, dE(\lambda)$,
(2) $E(M) = 0$ *for all* $M \subset \mathbb{R}$ *with* $M \cap \mathrm{Spec}(T) = \emptyset$,
(3) $E(M) \neq 0$ *for all open* $M \subset \mathbb{R}$ *with* $M \cap \mathrm{Spec}(T) \neq \emptyset$.

Consequently, we can associate a well-defined operator $f(T)$ to T for each function f that is integrable on $\mathrm{Spec}(T)$. That result is trivial for a polynomial f (with $\mathrm{id}(T) = T$ as in (1) and $1(T) = \mathrm{Id}$), but rather advanced for such simple functions like $f := \sqrt{\cdot}$ yielding an ultra-short proof of the *Square Root Lemma*. For comparison, see our elementary but lengthy proof of that Lemma below on p.54f. Different variants of the Spectral Theorem and a variety of different proofs are in the literature. We recommend [**Ped**, Theorems 4.4.1 and 5.3.8] for the bounded and for the not necessarily bounded case.

Metrics on the Space of Closed Operators. Recall that for a fixed separable complex Hilbert space H, we denoted by $\mathcal{B}(H)$ the algebra of bounded operators from H to H. It is naturally equipped with the metric defined by the *operator norm* $\|T-S\|$.

We shall denote by $\mathcal{C}(H)$ the space of closed densely defined operators in H. Clearly the operator norm does not make sense for unbounded operators. However, for S, $T \in \mathcal{C}(H)$ the orthogonal projections $P_{\mathfrak{G}(S)}$, $P_{\mathfrak{G}(T)}$ onto the graphs of S, T in $H \oplus H$ are bounded operators and

$$\gamma(S,T) := \|P_{\mathfrak{G}(T)} - P_{\mathfrak{G}(S)}\|$$

defines a metric for $\mathcal{C}(H)$, the **projection metric**.

It is also called the **gap metric** and it is (uniformly) equivalent with the metric given by measuring the distance between the (closed) graphs. For details and the proof of the following Lemma and Theorem, we refer to [**CorLab**, Section 3].

LEMMA 2.62. *For $T \in \mathcal{C}(H)$ the orthogonal projection onto the graph of T in $H \oplus H$ can be written (where $R_T := (\mathrm{I} + T^*T)^{-1}$) as*

$$P_{\mathfrak{G}(T)} = \begin{pmatrix} R_T & R_T T^* \\ T R_T & T R_T T^* \end{pmatrix} = \begin{pmatrix} R_T & T^* R_{T^*} \\ T R_T & T T^* R_{T^*} \end{pmatrix} = \begin{pmatrix} R_T & T^* R_{T^*} \\ T R_T & \mathrm{I} - R_{T^*} \end{pmatrix}.$$

THEOREM 2.63. (Cordes, Labrousse) *a) The space $\mathcal{B}(H)$ of bounded operators on H is dense in the space $\mathcal{C}(H)$ of all closed operators in H. The topology induced by the projection ($\cong$ gap) metric on $\mathcal{B}(H)$ is equivalent to that given by the operator norm.*
b) Let $\mathcal{CF}(H)$ denote the space of closed (not necessarily bounded) Fredholm operators. Then the index is constant on the connected components of $\mathcal{CF}(H)$ and yields a bijection between the integers and the connected components.

EXERCISE 2.64. Consider the multiplication operator M_{id} of Exercise 2.50c and let P_j denote the orthogonal projection of H onto the linear span of the j-th orthonormal basis element u_j. Clearly the sequence (P_j) does not converge in $\mathcal{B}(H)$ in the operator norm. Show that, however, the sequence $(M_{\mathrm{id}} - 2jP_j)$ of self–adjoint Fredholm operators converges in $\mathcal{C}(H)$ in the projection metric to M_{id}. [Hint: On the subset of self–adjoint (not necessarily bounded) operators in the space $\mathcal{C}(H)$, the projection metric is uniformly equivalent to the metric γ given by

$$\gamma(T_1, T_2) := \|(T_1 + i)^{-1} - (T_2 + i)^{-1}\|,$$

(cf. [**BLP**, Theorem 1.1]).]

REMARK 2.65. The results by Heinz Cordes and Jean–Philippe Labrousse may appear to be rather counter-intuitive. For (a), it is worth mentioning that the operator-norm distance and the projection metric on the set of bounded operators are equivalent, but not uniformly equivalent since the operator norm is complete, while the projection metric is not complete on the set of bounded operators. Actually, this is the point of the first part of (a), see also the preceding exercise.

Assertion (b) says two things: (i) that the index is a homotopy invariant, i.e. two Fredholm operators have the same index if they can be connected by a continuous curve in $\mathcal{CF}(H)$; (ii) that two Fredholm operators having the same index always can be connected by a continuous curve in $\mathcal{CF}(H)$. Both results are also true in the category of bounded Fredholm operators. Actually, topologically much farther reaching results for bounded Fredholm operators are shown in Chapter 3. For investigations of the topology of the subspace of self–adjoint (not necessarily bounded) Fredholm operators we refer to [**BLP**] and [**Les04**].

The delicacy of Assertion (b) is partly due to the delicacy of varying domains. However, if we fix a self–adjoint operator T with compact resolvent and dense domain $D \subseteq H$ and make D into a Hilbert space by the operator norm (along the lines of Exercise 2.39d), one may investigate all closed Fredholm operators in H with that domain D. A reasonable guess is that this space can be identified with the full space of bounded Fredholm operators by identifying D with H, and that this bijection is a homeomorphism.

7. Trace Class and Hilbert-Schmidt Operators

Here we give a rigorous definition and exposition of the fundamentals of trace class and Hilbert-Schmidt operators on a general (complex, infinite-dimensional) Hilbert space H. For an advanced reader, our presentation may seem a bit convoluted with all the small definitions, lemmata, propositions and theorems patched together. We refer such a reader to [**Ped**, Section 3.4] where all we need is done simply and directly. In this section, however, we prefer to confront our primary readership with all the details of the involved calculations instead of hiding them in general structural concepts and theorems.

We have already seen the important special case of Hilbert-Schmidt integral operators with square-integrable kernels on function spaces (see Exercises 2.29 and 2.30, p. 25). We will also need trace class operators later in Section 3.9 when we give a construction of the determinant line bundle over the space of Fredholm operators with index 0. A connection between determinants and traces is seen in the formula $\det\left(e^A\right) = e^{\operatorname{Tr} A}$ for a $A \in \mathrm{GL}(N, \mathbb{C})$, or the related formula (see Proposition 3.43, p. 96)

$$\det(\mathrm{Id} + A) = \sum_{k=0}^{N} \operatorname{Tr}\left(\Lambda^k A\right), \tag{2.36}$$

where $\Lambda^k A$ is the extension of A to the k-th exterior product $\Lambda^k(\mathbb{C}^n)$ of $\mathbb{C}^n$. For those not familiar with exterior products, let $\lambda_1(A), \ldots, \lambda_N(A)$ be the eigenvalues of A, repeated according to algebraic (as opposed to geometric) multiplicity. In view of Jordan canonical form, we have the formula (equivalent to (2.36))

$$\det(\mathrm{Id} + A) = \prod_{j=1}^{N} (1 + \lambda_j(A)) = \sum_{k=0}^{N} \left(\sum_{\langle i \rangle_k} \lambda_{i_1}(A) \cdots \lambda_{i_k}(A) \right) \tag{2.37}$$

where $\sum_{\langle i \rangle_k}$ denotes the sum over all indices $0 < i_1 < i_2 < \cdots < i_k \leq N$; this inner sum is in fact $\operatorname{Tr}\left(\Lambda^k A\right)$. The $k = 1$ term is the trace of A, namely

$$\operatorname{Tr}(A) = \sum_{i=1}^{N} \lambda_i(A), \tag{2.38}$$

and in general the k-th term is the elementary symmetric polynomial of degree k in the $\lambda_i(A)$, which again is the same as $\operatorname{Tr}\left(\Lambda^k A\right)$. The formula (2.36) extends to operators on a Hilbert space H when A is trace class. Indeed, it is essentially the Definition 3.44 (p. 98) of $\det(\mathrm{Id} + A)$ that we adopt. The equivalence of (2.36) and (2.37) for trace class operators A is true, but not trivial (see [**Sim**]). Roughly, $A \in \mathcal{B}(H)$ is trace class if the square-roots $\sqrt{\lambda_n}$ of the eigenvalues $\lambda_n \geq 0$ of A^*A are summable (i.e., $\sum_n \sqrt{\lambda_n} < \infty$). However, for this one would need to assume that there exists a complete eigenbasis of A^*A, and thus some assumptions on A (e.g., compactness) would have to be made. We would rather have the compactness of A emerge from our definition of *trace class* than assume compactness as part of the definition. In the course leading to the definition we adopt, we begin with some useful, basic definitions and results of independent interest.

The following *Square Root Lemma* is fundamental for establishing the *Polar Decomposition* of bounded operators in Proposition 2.77, which plays a prominent role in establishing the fact that the set $\mathcal{I}_1$ of trace class operators is closed under addition (see below). It is a simple consequence of the Spectral Theorem 2.61, p.51. Since we have not proven the Spectral Theorem, we give an elementary, though elaborate (i.e., a bit lengthy) proof of the Square Root Lemma.

THEOREM 2.66 (Square Root Lemma). *If $C \in \mathcal{B}^+$, then there is a unique $S \in \mathcal{B}^+$, such that $S^2 = C$. Denoting S by $\sqrt{C}$, the map $C \to \sqrt{C}$ is continuous in operator norm.*

PROOF. The power series for $\sqrt{1-x}$ about $x=0$ is

$$p(x) := \sum_{n=0}^{\infty} a_n x^n = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \prod_{k=0}^{n-1} \left(\tfrac{1}{2} - k\right) x^n. \tag{2.39}$$

For $n \geq 1$, the coefficients a_n are all negative. Thus,

$$\sum_{n=1}^{\infty} |a_n| = \lim_{x\to 1^-} \sum_{n=1}^{\infty} |a_n x^n| = \lim_{x \to 1^-} \left(1 - \sqrt{1-x}\right) = 1.$$

explain W's M-test

The Weierstrass M-test then implies that $p(x)$ converges uniformly and absolutely for $\|x\| \leq 1$ to $\sqrt{1-x}$. Let

$$\mathcal{B}_1^+ := \left\{B \in \mathcal{B}^+ : \|B\| \leq 1\right\}.$$

Since

$$A \in \mathcal{B}_1^+ \Rightarrow \|a_n A^n\| \leq |a_n| \|A\|^n \leq |a_n| \text{ and } \sum_{n=1}^{\infty} |a_n| < \infty,$$

the Weierstrass M-test implies that $\sum_{n=0}^{\infty} a_n A^n$ converges uniformly on $\mathcal{B}_1^+$ to a continuous function $p : \mathcal{B}_1^+ \to \mathcal{B}$, namely

$$p(A) := \sum_{n=0}^{\infty} a_n A^n \quad (\text{for } A \in \mathcal{B}_1^+).$$

The convergence of $\sum_{n=0}^{\infty} \|a_n A^n\|$ also implies that the formal squaring and rearrangement of the series $p(A)$ yields the expected result

$$p(A)^2 = \mathrm{Id} - A \quad (\text{for } A \in \mathcal{B}_1^+). \tag{2.40}$$

Note that $p(A) \geq 0$, since (using $\|A\| \leq 1$)

$$\left\langle \sum_{n=1}^{k} |a_n| A^n x, x \right\rangle \leq \sum_{n=1}^{k} |a_n| \|A^n x\| \|x\| \leq \sum_{n=1}^{k} |a_n| \|A\|^n \|x\|^2 \leq \|x\|^2,$$

which implies that

$$\begin{aligned}\langle p(A)x, x\rangle &= \lim_{k\to\infty} \left\langle \left(\mathrm{Id} - \sum_{n=1}^{k} |a_n| A^n\right) x, x \right\rangle \\ &= \|x\|^2 - \lim_{k\to\infty} \left\langle \sum_{n=1}^{k} |a_n| A^n x, x\right\rangle \geq 0.\end{aligned}$$

If $C \in \mathcal{B}_1^+$, we have $\mathrm{Id} - C \in \mathcal{B}_1^+$, since

$$\langle (\mathrm{Id} - C)x, x\rangle = \langle x, x\rangle - \langle Cx, x\rangle \geq \|x\|^2 - \|Cx\| \|x\| \geq \|x\|^2 (1 - \|C\|) \geq 0,$$

and then using Proposition 2.21 (p. 20), we get

$$\begin{aligned}\|\mathrm{Id} - C\| &= \sup\{\langle (\mathrm{Id} - C)x, x\rangle : x \in H, \ \|x\| = 1\} \\ &= \sup\{1 - \langle Cx, x\rangle : x \in H, \ \|x\| = 1\} \leq 1.\end{aligned}$$

Thus in (2.40) we may take $A = \operatorname{Id} - C \in \mathcal{B}_1^+$ and $(p(\operatorname{Id} - C))^2 = p(A)^2 = \operatorname{Id} - A = C$. Hence, for $C \in \mathcal{B}_1^+$ and $\|C\| \leq 1$, a positive square root of C is $p(\operatorname{Id} - C)$. For arbitrary $C \in \mathcal{B}^+$ with $C \neq 0$, note that $\operatorname{Id} - C/\|C\| \in \mathcal{B}_1^+$ and

$$\left(\sqrt{\|C\|}\, p\left(\operatorname{Id} - \frac{C}{\|C\|}\right)\right)^2 = \|C\| \frac{C}{\|C\|} = C.$$

Then there is a positive square root $S(C)$ of $C \in \mathcal{B}^+$, namely

$$S(C) := \begin{cases} p(\operatorname{Id} - C), & \text{if } \|C\| \leq 1, \\ \sqrt{\|C\|}\, p\left(\operatorname{Id} - \frac{C}{\|C\|}\right), & \text{if } \|C\| > 0. \end{cases}$$

Since $\sqrt{\alpha} p(1 - x/\alpha) = \sqrt{\alpha}\sqrt{1 - (1 - x/\alpha)} = \sqrt{\alpha}\sqrt{x/\alpha} = \sqrt{x}$, for $0 \leq x \leq \alpha \leq 1$, the two formulas agree for $0 < \|C\| \leq 1$. The first formula yields the continuity of S on $\mathcal{B}_1^+$, in particular at $C = 0$. Since $C \mapsto \sqrt{\|C\|}$ and $C \mapsto \operatorname{Id} - C/\|C\|$ are continuous functions on $\mathcal{B} \setminus \{0\}$, the second formula yields the continuity of S on the rest of $\mathcal{B}^+$.

We now prove the uniqueness property. Suppose that R is another positive square root of C. Then $RC = R^3 = R^2 R = CR$, so that R commutes with C and hence with $S(C)$ which (aside from the trivial case $C = 0$) can be written as a series of powers of $\operatorname{Id} - C/\|C\|$. To show that $R = S(C)$, we show $\operatorname{Ker}(S(C) - R)^\perp = \{0\}$. Since $RS(C) = S(C)R$, we have $(S(C) + R)(S(C) - R) = 0$ which implies that

$$(S(C) - R)(H) \subseteq \operatorname{Ker}(S(C) + R), \text{ and so } \overline{(S(C) - R)(H)} \subseteq \operatorname{Ker}(S(C) + R).$$

Since $S(C) - R$ is self-adjoint, we then have

$$\operatorname{Ker}(S(C) - R)^\perp = \overline{(S(C) - R)^*(H)} = \overline{(S(C) - R)(H)} \subseteq \operatorname{Ker}(S(C) + R).$$

Thus,

$$\begin{aligned} x \in \operatorname{Ker}(S(C) - R)^\perp &\Rightarrow \langle S(C)x, x\rangle + \langle Rx, x\rangle = \langle (S(C) + R)x, x\rangle = 0 \\ &\Rightarrow \langle S(C)x, x\rangle = 0 \text{ and } \langle Rx, x\rangle = 0, \end{aligned}$$

since $S(C) \geq 0$ and $R \geq 0$. Now,

$$\langle Rx, x\rangle = 0 \Rightarrow 0 = \left\langle S(R)^2 x, x\right\rangle = \|S(R)x\|^2 \Rightarrow Rx = S(R)^2 x = 0,$$

and similarly $\langle S(C)x, x\rangle = 0 \Rightarrow S(C)x = 0$. Thus,

$$(S(C) - R)(x) = S(C)x - Rx = 0 - 0 = 0,$$

and so $x \in \operatorname{Ker}(S(C) - R) \cap \operatorname{Ker}(S(C) - R)^\perp = \{0\}$. □

COROLLARY 2.67. *Every $A \in \mathcal{B}$ is a $\mathbb{C}$-linear combination of two self-adjoint operators, and any self-adjoint operator is a $\mathbb{C}$-linear combination of two unitary operators.*

PROOF. Note that $A + A^*$ and $i(A - A^*)$ are self-adjoint, and

$$A = \tfrac{1}{2}(A + A^*) - \tfrac{i}{2}(i(A - A^*)).$$

If $B \in \mathcal{B}$ is self-adjoint and $\|B\| = 1$, then $\operatorname{Id} - B^2 \geq 0$, and $\sqrt{\operatorname{Id} - B^2}$ makes sense. Note that

$$B = \tfrac{1}{2}\left(B + i\sqrt{\operatorname{Id} - B^2}\right) + \tfrac{1}{2}\left(B - i\sqrt{\operatorname{Id} - B^2}\right), \tag{2.41}$$

where $B \pm i\sqrt{\mathrm{Id} - B^2}$ is unitary since

$$\begin{aligned}&\left(B \pm i\sqrt{\mathrm{Id} - B^2}\right)^* \left(B \pm i\sqrt{\mathrm{Id} - B^2}\right)\\ &= \left(B \mp i\sqrt{\mathrm{Id} - B^2}\right)\left(B \pm i\sqrt{\mathrm{Id} - B^2}\right) = B^2 + \mathrm{Id} - B^2 = \mathrm{Id}\,.\end{aligned}$$

If $0 \neq B \in \mathcal{B}$ is self-adjoint, we can replace B by $B/\|B\|$ in (2.41) and multiply by $\|B\|$. If $B = 0$, then $B = 0\,\mathrm{Id} + 0\,\mathrm{Id}$. □

Definition 2.68. $A \in \mathcal{B}$ is **trace class**, written $A \in \mathcal{I}_1$, if there is a complete orthonormal system $\{e_0, e_1, \ldots\}$, such that for the operator $|A| := \sqrt{A^*A}$, we have

$$\operatorname{Tr}|A| := \sum\nolimits_{i=0}^{\infty} \langle |A|\, e_i, e_i\rangle = \sum\nolimits_{i=0}^{\infty} \left\| |A|^{\frac{1}{2}}\, e_i \right\|^2 < \infty.$$

More generally, for $p \in \mathbb{N}$, $A \in \mathcal{I}_p$ if $|A|^p \in \mathcal{I}_1$. If $A \in \mathcal{I}_2$, then A is called **Hilbert–Schmidt**.

Remark 2.69. Note that $\operatorname{Tr}|A|$ is independent of the choice of $\{e_0, e_1, \ldots\}$. Indeed, if $\{f_0, f_1, \ldots\}$ is another complete orthonormal system, then

$$\begin{aligned}\sum\nolimits_{i=0}^{\infty} \langle |A|\, e_i, e_i\rangle &= \sum\nolimits_{i=0}^{\infty} \left\| |A|^{\frac{1}{2}}\, e_i \right\|^2 = \sum\nolimits_{i=0}^{\infty} \left(\sum\nolimits_{j=0}^{\infty} \left| \left\langle |A|^{\frac{1}{2}}\, e_i, f_j \right\rangle \right|^2 \right)\\ &= \sum\nolimits_{j=0}^{\infty} \sum\nolimits_{i=0}^{\infty} \left(\left| \left\langle |A|^{\frac{1}{2}}\, f_j, e_i \right\rangle \right|^2 \right) = \sum\nolimits_{j=0}^{\infty} \langle |A|\, f_i, f_i\rangle\,.\end{aligned}$$

Remark 2.70. Note that $p \leq p' \Rightarrow \mathcal{I}_p \subset \mathcal{I}_{p'}$, since

$$\begin{aligned}&\sum\nolimits_{i=0}^{\infty} \left\langle |A|^{p'}\, e_i, e_i \right\rangle = \sum\nolimits_{i=0}^{\infty} \left\| |A|^{\frac{1}{2}p'}\, e_i \right\|^2 = \sum\nolimits_{i=0}^{\infty} \left(\left\| |A|^{\frac{1}{2}(p'-p)}\, |A|^{\frac{1}{2}p}\, e_i \right\| \right)^2\\ &\leq \sum\nolimits_{i=0}^{\infty} \left(\left\| |A|^{\frac{1}{2}(p'-p)} \right\| \left\| |A|^{\frac{1}{2}p}\, e_i \right\| \right)^2 = \left\| |A|^{\frac{1}{2}(p'-p)} \right\|^2 \sum\nolimits_{i=0}^{\infty} \left\| |A|^{\frac{1}{2}p}\, e_i \right\|^2\\ &= \left\| |A|^{\frac{1}{2}(p'-p)} \right\|^2 \operatorname{Tr}|A|^p\,.\end{aligned}$$

Recall that $\mathcal{K}$ denotes the ideal of compact operators in $\mathcal{B}$. Let $\mathcal{B}_f$ denote the ideal of finite rank operators. We will eventually show (see Proposition 2.75) that

$$\mathcal{B}_f \subset \mathcal{I}_p \subset \mathcal{I}_{p'} \subset \mathcal{K}, \text{ for } 1 \leq p \leq p',$$

but first we prove

Proposition 2.71. *Any trace class operator is compact; i.e, $\mathcal{I}_1 \subset \mathcal{K}$.*

Proof. For any $x \in H$, we have

$$Ax = A\left(\sum\nolimits_{i=0}^{\infty} \langle x, e_i\rangle\, e_i\right) = \sum\nolimits_{i=0}^{\infty} \langle x, e_i\rangle\, Ae_i.$$

Thus, we have (as always) that A is the *pointwise* limit of finite rank operators

$$A = \sum\nolimits_{i=0}^{\infty} \langle \cdot, e_i\rangle\, Ae_i = \sum\nolimits_{i=0}^{\infty} e_i^* \otimes Ae_i = \lim_{n\to\infty} \sum\nolimits_{i=0}^{n-1} e_i^* \otimes Ae_i.$$

To show that A is compact, it suffices to show convergence in *norm*; i.e.,

$$\lim_{n\to\infty} \sup\{||(A - \sum\nolimits_{i=0}^{n-1} e_i^* \otimes Ae_i)\,(x)\,|| : x \in H, \|x\|^2 = 1\} \overset{!}{=} 0. \tag{2.42}$$

First, note that $A \in \mathcal{I}_1 \Rightarrow |A|^2 \in \mathcal{I}_1$, since

$$\begin{aligned}&\sum\nolimits_{i=0}^{\infty} \left\langle |A|^2 e_i, e_i \right\rangle = \sum\nolimits_{i=0}^{\infty} \||A| e_i\|^2 = \sum\nolimits_{i=0}^{\infty} \left(\left\| |A|^{\frac{1}{2}} |A|^{\frac{1}{2}} e_i \right\| \right)^2 \\ &\leq \sum\nolimits_{i=0}^{\infty} \left(\left\| |A|^{\frac{1}{2}} \right\| \left\| |A|^{\frac{1}{2}} e_i \right\| \right)^2 = \left\| |A|^{\frac{1}{2}} \right\|^2 \sum\nolimits_{i=0}^{\infty} \left\| |A|^{\frac{1}{2}} e_i \right\|^2 \\ &= \left\| |A|^{\frac{1}{2}} \right\|^2 \operatorname{Tr} |A| .\end{aligned}$$

In the sup of (2.42), we may assume $x \in H_n := \operatorname{span}(e_0, \ldots, e_{n-1})^{\perp}$, since

$$(A - \sum\nolimits_{i=0}^{n-1} e_i^* \otimes Ae_i)|_{H_n^{\perp}} = 0.$$

Thus, it suffices to show that

$$\lim_{n\to\infty} \sup \left\{ \|Ax\| : x \in H_n, \|x\|^2 = 1 \right\} = 0.$$

Note that any $x \in H_n$ with $\|x\|^2 = 1$ may serve as f_n in a complete orthonormal extension $\{f_i\}_{i=0}^{\infty}$ of $e_0, \ldots, e_{n-1}$, and so we have

$$\sum\nolimits_{i=0}^{n-1} \|Ae_i\|^2 + \|Ax\|^2 \leq \sum\nolimits_{i=0}^{\infty} \|Af_i\|^2 = \sum\nolimits_{i=0}^{\infty} \||A| f_i\|^2 = \operatorname{Tr}\left(|A|^2\right).$$

Thus, as desired,

$$\begin{aligned}&x \in H_n, \|x\|^2 = 1 \Rightarrow \|Ax\|^2 \leq \operatorname{Tr}\left(|A|^2\right) - \sum\nolimits_{i=0}^{n-1} \|Ae_i\|^2 \Rightarrow \\ &\lim_{n\to\infty} \sup \left\{ \|Ax\|^2 : x \in H_n, \|x\|^2 = 1 \right\} = \operatorname{Tr} |A|^2 - \lim_{n\to\infty} \sum\nolimits_{i=0}^{n-1} \|Ae_i\|^2 = 0. \quad \square\end{aligned}$$

If $A \in \mathcal{I}_1$, then $|A| \in \mathcal{I}_1$ and hence $|A|$ is compact, self-adjoint and positive. By the Hilbert-Schmidt Theorem 2.20, there is a complete orthonormal system for $(\operatorname{Ker} A)^{\perp}$, say $\{e_i : 0 \leq i < N\}$, where N may be finite or ∞, such that $|A| e_i = \mu_i e_i$ with $\mu_i > 0$, $\mu_0 \geq \mu_1 \geq \mu_2 \geq \ldots \geq 0$ (repeated according to multiplicity), and $\lim_{i\to\infty} \mu_i = 0$ if $N = \infty$. The positive eigenvalues μ_i of $|A|$ are known as the **singular values** of A. For $x \in H$,

$$\begin{aligned}&Ax = A\left(\sum\nolimits_{i=0}^{N} \langle x, e_i\rangle e_i\right) = \sum\nolimits_{i=0}^{N} \langle x, e_i\rangle Ae_i = \sum\nolimits_{i=0}^{N} \mu_i \langle x, e_i\rangle \left(\mu_i^{-1} Ae_i\right) \\ &= \sum\nolimits_{i=0}^{N} \mu_i \langle x, e_i\rangle f_i, \text{ where } f_i := \mu_i^{-1} Ae_i.\end{aligned}$$

Thus, we have the so-called **canonical expansion** of A

$$A = \sum\nolimits_{i=0}^{N} \mu_i \langle \cdot, e_i\rangle f_i, \tag{2.43}$$

which (assuming $N = \infty$) converges in norm to the compact operator A, since

$$\begin{aligned}&\left\| A - \sum\nolimits_{i=0}^{n} \mu_i \langle \cdot, e_i\rangle f_i \right\| = \sup \left\{ \left\| Ax - \sum\nolimits_{i=0}^{n} \mu_i \langle x, e_i\rangle f_i \right\| : x \in H, \ \|x\| = 1 \right\} \\ &= \sup \left\{ \|Ax\| : x \in \operatorname{span}\{e_0, \ldots e_n\}^{\perp}, \ \|x\| = 1 \right\} \\ &= \sup \left\{ \|(|A| x)\| : x \in \operatorname{span}\{e_0, \ldots e_n\}^{\perp}, \ \|x\| = 1 \right\} = \mu_{n+1} \to 0 \text{ as } n \to \infty.\end{aligned}$$

We have

$$\begin{aligned}&\langle f_i, f_j\rangle = \left\langle \mu_i^{-1} Ae_i, \mu_j^{-1} Ae_j \right\rangle = \mu_i^{-1} \mu_j^{-1} \langle A^* Ae_i, e_j\rangle = \delta_{ij}, \text{ and} \\ &(AA^*) f_i = (AA^*) \mu_i^{-1} Ae_i = \mu_i^{-1} A (A^* Ae_i) = \mu_i^{-1} A \left(\mu_i^2 e_i\right) = \mu_i^2 \left(\mu_i^{-1} Ae_i\right) = \mu_i^2 f_i.\end{aligned}$$

It follows that $\{f_i : 0 \le i < N+1\}$ is a complete orthonormal system for $(\operatorname{Ker} A^*)^\perp = \overline{\operatorname{Im} A}$, and the μ_i are also the singular values of A^*, which (noting that $f_i = \mu_i^{-1} A e_i \Rightarrow e_i = \mu_i^{-1} A^* f_i$) has the canonical expansion

$$A^* = \sum_{i=0}^{N} \mu_i \langle \cdot, f_i \rangle e_i.$$

The following definition then yields $\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}$, for $A \in \mathcal{I}_1$.

Definition 2.72. For $A \in \mathcal{I}_1$ with *canonical expansion* $A = \sum_{i=0}^{N} \mu_i \langle \cdot, e_i \rangle f_i$, we define

$$\operatorname{Tr} A := \sum_{j=0}^{N} \langle Ae_j, e_j \rangle = \sum_{j=0}^{N} \left\langle \sum_{i=0}^{N} \mu_i \langle e_j, e_i \rangle f_i, e_j \right\rangle = \sum_{j=0}^{N} \mu_j \langle f_j, e_j \rangle .$$

This sum is absolutely convergent, since $|\mu_j \langle f_j, e_j \rangle| \le \mu_j \|f_j\| \|e_j\| = \mu_j$ and

$$\sum_{j=0}^{N} \mu_j = \sum_{j=0}^{N} \langle |A| e_j, e_j \rangle = \operatorname{Tr} |A| < \infty.$$

Remark 2.73. If $A \in \mathcal{I}_1$ and A is self adjoint, then $A = \sum_{i=0}^{N} \lambda_i \langle \cdot, e_i \rangle e_i$ is the *canonical expansion of* A, *where* $\{e_i\}$ is a complete orthonormal system for $(\operatorname{Ker} A)^\perp$ with $Ae_i = \lambda_i e_i$. In this case,

$$\operatorname{Tr} A = \sum_{j=0}^{N} \lambda_j \langle e_j, e_j \rangle = \sum_{j=0}^{N} \lambda_j,$$

which converges absolutely and agrees with (2.38) when $N < \infty$.

Proposition 2.74. *If* $\{g_i\}$ *is any complete orthonormal system for* H *and* $A \in \mathcal{I}_1$, *then*

$$\sum_{j=0}^{\infty} \langle Ag_j, g_j \rangle = \operatorname{Tr} A,$$

and the sum is absolutely convergent.

Proof. We first verify that the sum is absolutely convergent:

$$\begin{aligned}
&\sum_{j=0}^{\infty} |\langle Ag_j, g_j \rangle| = \sum_{j=0}^{\infty} \left| \left\langle \sum_{i=1}^{N} \mu_i \langle g_j, e_i \rangle f_i, g_j \right\rangle \right| \\
&\le \sum_{j=0}^{\infty} \sum_{i=1}^{N} \mu_i |\langle g_j, e_i \rangle \langle f_i, g_j \rangle| \\
&= \sum_{i=1}^{N} \mu_i \sum_{j=0}^{\infty} |\langle g_j, e_i \rangle \langle f_i, g_j \rangle| \\
&\le \sum_{i=1}^{N} \mu_i \left(\sum_{j=0}^{\infty} |\langle g_j, e_i \rangle|^2 \right)^{\frac{1}{2}} \left(\sum_{j=0}^{\infty} |\langle f_i, g_j \rangle|^2 \right)^{\frac{1}{2}} \\
&= \sum_{i=1}^{N} \mu_i \|e_i\| \|f_i\| = \sum_{i=1}^{N} \mu_i = \operatorname{Tr}(|A|) = \|A\|_1 .
\end{aligned}$$

The absolute convergence of $\sum_{j=0}^{\infty}\sum_{i=1}^{N}\mu_i\left|\langle g_j,e_i\rangle\langle f_i,g_j\rangle\right|$ just shown allows the interchange of the sums over i and j in the following:

$$\begin{aligned}\sum_{j=0}^{\infty}\langle Ag_j,g_j\rangle &= \sum_{j=0}^{\infty}\left\langle\sum_{i=1}^{N}\mu_i\langle g_j,e_i\rangle f_i,g_j\right\rangle\\ &= \sum_{j=0}^{\infty}\sum_{i=1}^{N}\mu_i\langle g_j,e_i\rangle\langle f_i,g_j\rangle\\ &= \sum_{i=1}^{N}\sum_{j=0}^{\infty}\mu_i\langle g_j,e_i\rangle\langle f_i,g_j\rangle\\ &= \sum_{i=1}^{N}\mu_i\langle f_i,e_i\rangle = \operatorname{Tr}A.\end{aligned}$$ □

PROPOSITION 2.75. *For any $p,p'\in\mathbb{N}$ with $p\le p'$, we have*

$$\mathcal{B}_f\subset\mathcal{I}_p\subset\mathcal{I}_{p'}\subset\mathcal{K}.$$

PROOF. Since $\mathcal{B}_f\subset\mathcal{I}_p$ is clear, it suffices (by Remark 2.70) to prove that $\mathcal{I}_p\subset\mathcal{K}$ for any $p\in\mathbb{N}$. For $A\in\mathcal{I}_p$, we have $|A|^p\in\mathcal{I}_1$. Thus, $|A|^p$ is compact. Let $A=\sum_{i=0}^{N}\mu_i\langle\cdot,e_i\rangle e_i$ be the canonical expansion of $|A|^p$. Then the canonical expansion of $|A|$ is $\sum_{i=0}^{N}\mu_i^{1/p}\langle\cdot,e_i\rangle e_i$. Since $\mu_i\to 0$, we have $\mu_i^{1/p}\to 0$, and so $|A|$ is compact. This implies that A is compact. Indeed, since $\||A|x\|=\|Ax\|$, if $\{|A|x_n\}$ has a convergent subsequence $\{|A|x_{n_i}\}$ for any bounded sequence $\{x_n\}$, then $\{Ax_n\}$ has the convergent subsequence $\{Ax_{n_i}\}$ since it will also be a Cauchy sequence: $\left\|Ax_{n_i}-Ax_{n_j}\right\|=\left\||A|x_{n_i}-|A|x_{n_j}\right\|\to 0$ as $i,j\to\infty$. □

The set $\mathcal{I}_1$ is clearly closed under scalar multiplication, but *less* clearly under addition, as the following exercise suggests.

EXERCISE 2.76. Show that there are 2×2 matrices A and B for which $|A+B|\not\leq|A|+|B|$ (i.e. $|A|+|B|-|A+B|$ is not positive).

To show that nevertheless $\mathcal{I}_1$ is closed under addition, it is convenient to first introduce polar decomposition.

PROPOSITION 2.77 (Polar Decomposition). *Any $A\in\mathcal{B}$ can be uniquely expressed in the form $A=UP$, where*

$P\in\mathcal{B}$ is self-adjoint and positive (i.e., $\langle Px,x\rangle\ge 0$ for all $x\in H$) with $P\left((\operatorname{Ker}A)^{\perp}\right)\subseteq(\operatorname{Ker}A)^{\perp}$, and

$U\in\mathcal{B}$, $\operatorname{Ker}U=\operatorname{Ker}A$, and $\langle U(x),U(y)\rangle=\langle x,y\rangle$ for $x,y\in(\operatorname{Ker}A)^{\perp}$ (i.e., $U|_{(\operatorname{Ker}A)^{\perp}}$ is an isometry of $(\operatorname{Ker}A)^{\perp}$ onto $U(H)$).

Indeed,

$$P=\sqrt{A^*A}\quad\text{and}\quad U=0_{\operatorname{Ker}A}\oplus\left(A\circ\left(\sqrt{A^*A}|_{(\operatorname{Ker}A)^{\perp}}\right)^{-1}\right). \tag{2.44}$$

PROOF. We first establish uniqueness of the decomposition $A=UP$. Note that $U^*U\left((\operatorname{Ker}A)^{\perp}\right)\subseteq(\operatorname{Ker}A)^{\perp}$, since

$$x\in(\operatorname{Ker}A)^{\perp},\ y\in\operatorname{Ker}A\Rightarrow\langle U^*U(x),y\rangle=\langle Ux,Uy\rangle=0.$$

Also,

$$x,y\in(\operatorname{Ker}A)^{\perp}\Rightarrow\langle U^*Ux,y\rangle=\langle Ux,Uy\rangle=\langle x,y\rangle\Rightarrow U^*U|_{(\operatorname{Ker}A)^{\perp}}=\operatorname{Id}_{(\operatorname{Ker}A)^{\perp}}.$$

For $x \in \operatorname{Ker} A$, $(A^*A)\,x = 0 = P^2x$, and for $x \in (\operatorname{Ker} A)^\perp$ we have

$$A^*Ax = (UP)^*\,(UP)\,x = P^*\,(U^*U)\,(Px) = P^*Px = P^2x,$$

since $Px \in (\operatorname{Ker} A)^\perp$ and $U^*U|_{(\operatorname{Ker} A)^\perp} = \operatorname{Id}_{(\operatorname{Ker} A)^\perp}$. Thus, $A^*A = P^2$ and so $P = \sqrt{A^*A}$ by Theorem 2.66. Then $A = UP \Rightarrow U|_{(\operatorname{Ker} A)^\perp} = A \circ \left(P|_{(\operatorname{Ker} A)^\perp}\right)^{-1}$, and since $U|_{\operatorname{Ker} A} = 0$, we have the uniqueness. Defining P and U by (2.44), we have $A = UP$, since

$$x \in \operatorname{Ker} A \Leftrightarrow \langle Px, Px\rangle = \left\langle P^2x, x\right\rangle = \langle A^*Ax, x\rangle = 0 \Rightarrow UPx = 0 = Ax, \text{ and}$$
$$x \in (\operatorname{Ker} A)^\perp \Rightarrow UPx = A\left(\left(P|_{(\operatorname{Ker} A)^\perp}\right)^{-1}\right)(Px) = Ax.$$

By definition, $\sqrt{A^*A}$ is self-adjoint and positive. Also, since $(A^*A)\,|_{(\operatorname{Ker} A)^\perp} \in \mathcal{B}\left((\operatorname{Ker} A)^\perp\right)$ is positive, by the uniqueness of positive square roots, we have

$$\sqrt{A^*A} = 0_{\operatorname{Ker} A} \oplus \sqrt{(A^*A)\,|_{(\operatorname{Ker} A)^\perp}}\ ,$$

whence $\sqrt{A^*A}\left((\operatorname{Ker} A)^\perp\right) \subseteq (\operatorname{Ker} A)^\perp$. For $x, y \in (\operatorname{Ker} A)^\perp$, we have

$$\begin{aligned}
\langle U(x), U(y)\rangle &= \left\langle A \circ \left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} x, A \circ \left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} y\right\rangle \\
&= \left\langle \left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} A^*A\left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} x, y\right\rangle \\
&= \left\langle \left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} \left(\sqrt{A^*A}|_{(\operatorname{Ker} A)^\perp}\right)^{-1} A^*Ax, y\right\rangle \\
&= \left\langle \left((A^*A)\,|_{(\operatorname{Ker} A)^\perp}\right)^{-1} A^*Ax, y\right\rangle = \langle x, y\rangle\,.
\end{aligned}$$

By definition, $\operatorname{Ker} A \subseteq \operatorname{Ker} U$ and since we have just seen that $\operatorname{Ker}\left(U|_{(\operatorname{Ker} A)^\perp}\right) = 0$, we have $\operatorname{Ker} A = \operatorname{Ker} U$. □

Exercise 2.78. (a) Use polar decompositions of A, B and $A + B$ to show that for $p = 1$ and $p = 2$

$$A, B \in \mathcal{I}_p \Rightarrow (\operatorname{Tr}|A + B|^p)^{\frac{1}{p}} \leq (\operatorname{Tr}|A|^p)^{\frac{1}{p}} + (\operatorname{Tr}|B|^p)^{\frac{1}{p}}\,, \tag{2.45}$$

and hence that $A + B \in \mathcal{I}_p$. By Proposition 2.74 $\operatorname{Tr} : \mathcal{I}_1 \to \mathbb{C}$ is then linear. [Hint: The case $p = 2$ is a bit less tricky. As a last resort, see [**ReSi72**, p. 208] for the case $p = 1$.]
(b) Show that if $A \in \mathcal{I}_p$ for $p \in \mathbb{N}$ and $B \in \mathcal{B}$, then A^*, AB and BA are in $\mathcal{I}_p$. Moreover, for $A \in \mathcal{I}_1$, prove that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. [Hint: By Corollary 2.67 and Part (a), we may assume that B is unitary.]

While (2.45) is valid for all p, here we only need it for $p = 1$ and 2. In general, $\mathcal{I}_p$ is a normed linear space, the **Schatten class**, with norm $\|A\|_p := (\operatorname{Tr}|A|^p)^{\frac{1}{p}}$. In particular, there is a norm (the *trace norm*) on $\mathcal{I}_1$, given by

$$\|A\|_1 := \operatorname{Tr}(|A|) = \sum\nolimits_{i=0}^{N} \mu_i, \tag{2.46}$$

where $A = \sum_{i=0}^{N} \mu_i \langle \cdot, e_i \rangle f_i$ is the canonical expansion of A. We have

$$\|Ax\|^2 = \sum_{i=0}^{N} \mu_i^2 |\langle x, e_i \rangle|^2 \le \mu_0^2 \|x\|^2 ,$$

with equality for $x = e_0$. Thus,

$$\|A\| = \mu_0 \le \sum_{i=0}^{N} \mu_i = \|A\|_1 .$$

Note that Part (b) of Exercise 2.78 yields that $\mathcal{I}_p$ is an *ideal* of algebra $\mathcal{B}$, whence the use of the symbol "$\mathcal{I}$". When $p = 2$, we now show that $\|A\|_2$ is in fact the norm associated with an inner product. Note that $\operatorname{Tr}(D^*C)$ exists for $C, D \in \mathcal{I}_2$, since

$$\begin{aligned}
&4D^*C = (C^* + D^*)(C + D) + i(C^* - iD^*)(C + iD) \\
&- (C^* - D^*)(C - D) - i(C^* + iD^*)(C - iD) \\
&= \sum_{k=0}^{3} i^k \left(C + i^k D\right)^* \left(C + i^k D\right) = \sum_{k=0}^{3} i^k \left|C + i^k D\right|^2 .
\end{aligned}$$

Then by Exercise 2.78, $\left|C + i^k D\right|^2 \in \mathcal{I}_1$ and $D^*C \in \mathcal{I}_1$. Thus, we have an inner product $\langle \cdot, \cdot \rangle_{\operatorname{Tr}}$ on $\mathcal{I}_2$ given by

$$\langle C, D \rangle_{\operatorname{Tr}} := \operatorname{Tr}(D^*C) \text{ for } C, D \in \mathcal{I}_2,$$

with norm $\sqrt{\langle C, C \rangle_{\operatorname{Tr}}} = \left(\operatorname{Tr}|C|^2\right)^{\frac{1}{2}} = \|C\|_2$.

Proposition 2.79. *For $A \in \mathcal{I}_1$ and $B \in \mathcal{B}$, we have $BA \in \mathcal{I}_1$* (by Exercise 2.78) *and*

$$\|BA\|_1 \le \|B\| \, \|A\|_1 .$$

Proof. Let $A = U|A|$ denote the polar decomposition of $A \in \mathcal{I}_1$. Since $|A|^{\frac{1}{2}} \in \mathcal{I}_2$,

$$\begin{aligned}
|\operatorname{Tr}(BA)|^2 = |\operatorname{Tr}(BU|A|)|^2 &= \left|\operatorname{Tr}\left(BU|A|^{\frac{1}{2}}|A|^{\frac{1}{2}}\right)\right|^2 \\
&= \left|\left\langle \left(|A|^{\frac{1}{2}}\right)^*, BU|A|^{\frac{1}{2}} \right\rangle_{\operatorname{Tr}}\right|^2 = \left|\left\langle |A|^{\frac{1}{2}}, BU|A|^{\frac{1}{2}} \right\rangle_{\operatorname{Tr}}\right|^2 .
\end{aligned}$$

By the Cauchy-Schwarz inequality for $\langle \cdot, \cdot \rangle_{\operatorname{Tr}}$,

$$\begin{aligned}
&\left|\left\langle |A|^{\frac{1}{2}}, BU|A|^{\frac{1}{2}} \right\rangle_{\operatorname{Tr}}\right|^2 \le \left\||A|^{\frac{1}{2}}\right\|_2^2 \left\|BU|A|^{\frac{1}{2}}\right\|_2^2 = \operatorname{Tr}(|A|) \left\|BU|A|^{\frac{1}{2}}\right\|_2^2 \\
&= \|A\|_1 \operatorname{Tr}\left(\left(BU|A|^{\frac{1}{2}}\right)^* BU|A|^{\frac{1}{2}}\right) = \|A\|_1 \operatorname{Tr}\left(|A|^{\frac{1}{2}} U^*B^*BU|A|^{\frac{1}{2}}\right).
\end{aligned}$$

For a complete orthonormal system $e_1, e_2, \ldots$, we have

$$\begin{aligned}
&\operatorname{Tr}\left(|A|^{\frac{1}{2}} U^*B^*BU|A|^{\frac{1}{2}}\right) = \sum_{i=0}^{\infty} \left\langle |A|^{\frac{1}{2}} U^*B^*BU|A|^{\frac{1}{2}} e_i, e_i \right\rangle \\
&= \sum_{i=0}^{\infty} \left\langle U^*B^*BU|A|^{\frac{1}{2}} e_i, |A|^{\frac{1}{2}} e_i \right\rangle \le \sum_{i=0}^{\infty} \|U^*B^*BU\| \left\langle |A|^{\frac{1}{2}} e_i, |A|^{\frac{1}{2}} e_i \right\rangle \\
&= \|U^*B^*BU\| \left(\sum_{i=0}^{\infty} \langle |A| e_i, e_i \rangle\right) = \|A\|_1 \|B\|^2 .
\end{aligned}$$

Thus, combining the above, $|\operatorname{Tr}(BA)|^2 \le \|A\|_1^2 \|B\|^2$ or $|\operatorname{Tr}(BA)| \le \|A\|_1 \|B\|$. Letting $BA = W|BA|$ denote the polar decomposition of BA, we then have

$$\|BA\|_1 = \operatorname{Tr}|BA| = |\operatorname{Tr}(W^*BA)| \le \|A\|_1 \|W^*B\| = \|A\|_1 \|B\| . \qquad \square$$

Recall that $\mathcal{B}_f \subset \mathcal{B}$ denotes the subspace of finite-rank operators. Relative to the operator norm, the closure of $\mathcal{B}_f$ is the space $\mathcal{K}$ of compact operators. Since there are compact operators which are not trace class (e.g., $\sum_{n=1}^{\infty} \frac{1}{n} \langle \cdot, e_n \rangle e_n$), $\mathcal{I}_1$ is not a closed subspace of $\mathcal{B}$ in the operator norm. However, we have

PROPOSITION 2.80. *$(\mathcal{I}_1, \|\cdot\|_1)$ is a Banach space. The set $\mathcal{B}_f$ of finite-rank operators is $\|\cdot\|_1$-dense in $\mathcal{I}_1$; i.e., $(\mathcal{I}_1, \|\cdot\|_1)$ is the $\|\cdot\|_1$-completion of $\mathcal{B}_f$.*

PROOF. Let (A_n) be a Cauchy sequence in $(\mathcal{I}_1, \|\cdot\|_1)$, then (A_n) is a Cauchy sequence of compact operators in $(\mathcal{B}, \|\cdot\|)$ with limit $A \in \mathcal{K}$. We need to show $A \in \mathcal{I}_1$with $\|A - A_n\|_1 \to 0$. Since (A_n) is a Cauchy relative to $\|\cdot\|_1$, the sequence $(\|A_n\|_1)$ is bounded ($\|A_n\|_1 \leq \|A_n - A_m\|_1 + \|A_m\|_1 \leq \varepsilon + \|A_m\|_1$ for $n \geq m$, where m is chosen sufficiently large). Let $A = \sum_{i=0}^{N} \mu_i \langle \cdot, e_i \rangle f_i$ denote the canonical expansion of A. We first show $\|A\|_1 = \sum_{i=0}^{N} \mu_i < \infty$. Since the case $N < \infty$ is clear, let $N = \infty$. If $\|A_n - A\| \to 0$, then $\|A_n^* A_n - A^* A\| \to 0$, since

$$\begin{aligned} &\|A^*A - A_n^* A_n\| \leq \|(A^* - A_n^*) A + A_n^* (A - A_n)\| \\ &\leq \|A^* - A_n^*\| \|A\| + \|A_n^*\| \|A - A_n\| \leq (2\|A\| + 1) \|A - A_n\|, \end{aligned}$$

for n sufficiently large. Using this and Theorem 2.66, we obtain

$$\|A_n - A\| \to 0 \Rightarrow \||A| - |A_n|\| = \left\| \sqrt{A^*A} - \sqrt{A_n^* A_n} \right\| \to 0. \tag{2.47}$$

Then for each (finite) $m \in \mathbb{N}$,

$$\begin{aligned} \sum\nolimits_{i=1}^{m} \langle |A| e_i, e_i \rangle &= \sum\nolimits_{i=1}^{m} \lim_{n\to\infty} \langle |A_n| e_i, e_i \rangle \\ &= \lim_{n\to\infty} \sum\nolimits_{i=1}^{m} \langle |A_n| e_i, e_i \rangle \leq \|A_n\|_1 . \end{aligned}$$

Hence $A \in \mathcal{I}_1$, since

$$\|A\|_1 = \lim_{m\to\infty} \sum\nolimits_{i=1}^{m} \langle |A| e_i, e_i \rangle \leq \sup_{n\in\mathbb{N}} \{\|A_n\|_1\} < \infty.$$

By (2.47), we have (as $p \to \infty$)

$$\|A - A_p\| \to 0 \Leftrightarrow \|(A - A_n) - (A_p - A_n)\| \to 0 \Rightarrow |A_p - A_n| \to |A - A_n|,$$

and so

$$\begin{aligned} &\sum\nolimits_{i=1}^{m} \langle |A - A_n| e_i, e_i \rangle = \lim_{p\to\infty} \sum\nolimits_{i=1}^{m} \langle |A_p - A_n| e_i, e_i \rangle \leq \lim_{p\to\infty} \|A_p - A_n\|_1 \\ &\Rightarrow \|A - A_n\|_1 = \lim_{m\to\infty} \sum\nolimits_{i=1}^{m} \langle |A - A_n| e_i, e_i \rangle \leq \lim_{p\to\infty} \|A_p - A_n\|_1 . \end{aligned}$$

Given any $\varepsilon > 0$, for n and p sufficiently large we have $\|A_p - A_n\|_1 \leq \varepsilon$. Hence, $\|A - A_n\|_1 \leq \varepsilon$ for n sufficiently large (i.e., $\lim_{n\to\infty} \|A - A_n\|_1 = 0$). As for the density of $\mathcal{B}_f$ in $\mathcal{I}_1$, let $A = \sum_{i=0}^{\infty} \mu_i \langle \cdot, e_i \rangle f_i$ denote the canonical expansion of $A \in \mathcal{I}_1 - \mathcal{B}_f$. Then as $n \to \infty$,

$$\left\| A - \sum\nolimits_{i=0}^{n} \mu_i \langle \cdot, e_i \rangle f_i \right\|_1 = \left\| \sum\nolimits_{i=n+1}^{\infty} \mu_i \langle \cdot, e_i \rangle f_i \right\|_1 \leq \sum\nolimits_{i=n+1}^{\infty} \mu_i \to 0. \qquad \square$$

CHAPTER 3

Fredholm Operator Topology

Synopsis. Calkin Algebra and Atkinson's Theorem. Perturbation Theory: Homotopy Invariance of the Index, Homotopies of Operator-Valued Functions, The Theorem of Kuiper. The Topology of $\mathcal{F}$: The Homotopy Type, Index Bundles, The Theorem of Atiyah-Jänich. Determinant Line Bundles: The Quillen Determinant Line Bundle, Determinants, The Segal-Furutani Construction. Spectral Invariants: Essentially Unitary Equivalence, What Is a Spectral Invariant? Eta Function, Zeta Function, Zeta Regularized Determinant

1. The Calkin Algebra

So far, we have introduced compact operators for purely practical reasons: Within pure mathematics, they came from the search for a (closed) class of operators that exhibit properties analogous to those of the operators of finite rank. In applied mathematics, they enter through the theory of integral equations associated with the study of oscillations. Actually, the compact operators have yet a deeper significance in the representation of Fredholm operators.

We recall our notation: H is a complex, separable Hilbert space; $\mathcal{B}$ denotes the Banach algebra of bounded linear operators on H (in modern terminology, $\mathcal{B}$ is even a C^*-algebra; see Exercise 2.6, where one needs to verify the additional axiom $\|T^*T\| = \|T\|^2$); $\mathcal{K} \subseteq \mathcal{B}$ denotes the closed two-sided ideal of compact operators (see Theorem 2.24); and $\mathcal{F} \subseteq \mathcal{B}$ denotes the space of Fredholm operators. We begin with a simple exercise.

Exercise 3.1 (J. W. Calkin, 1941). Show that the quotient space $\mathcal{B}/\mathcal{K}$, consisting of equivalence classes $\pi(T) := \{T - K : K \in \mathcal{K}\}$, where $T \in \mathcal{B}$, forms a Banach algebra.
[Hint: Since $\mathcal{K}$ is a linear subspace, clearly $\mathcal{B}/\mathcal{K}$ is a vector space. To prove that $\mathcal{B}/\mathcal{K}$ is an algebra, one must use the fact that $\mathcal{K}$ is a two-sided ideal. Then show that since $\mathcal{K}$ is closed, $\mathcal{B}/\mathcal{K}$ can be made into a Banach space by defining a norm on $\mathcal{B}/\mathcal{K}$ by

$$\|\pi(T)\| := \inf\{\|T - K\| : K \in \mathcal{K}\} = \inf\{\|R\| : R \in \pi(T)\}.$$

It remains to show that

$$\|\pi(\mathrm{Id})\| = 1 \text{ and } \|\pi(T)\pi(S)\| \leq \|\pi(T)\| \, \|\pi(S)\|.$$

To prove the left equation, assume that there is a $K \in \mathcal{K}$ with $\|\mathrm{Id} - K\| < 1$ and show that K is invertible (using the argument in the proof of Theorem 2.31 involving geometric series); this contradicts the compactness of K. To prove the right inequality, apply the trick

$$\inf_{K \in \mathcal{K}} \|TS - K\| \leq \inf_{K_1, K_2 \in \mathcal{K}} \|(T - K_1)(S - K_2)\|. \quad]$$

THEOREM 3.2. (F. V. Atkinson, 1951) *If* $(\mathcal{B}/\mathcal{K})^\times$ *is the group of units (i.e., elements which are invertible with respect to multiplication) of* $\mathcal{B}/\mathcal{K}$ *and* $\pi : \mathcal{B} \to \mathcal{B}/\mathcal{K}$ *is the natural projection, then we have*

$$\mathcal{F} = \pi^{-1}((\mathcal{B}/\mathcal{K})^\times).$$

EXERCISE 3.3. Show that this theorem of Frederick Valentine Atkinson can also be written as: An operator $T \in \mathcal{B}$ is a Fredholm operator exactly when there are $S \in \mathcal{B}$ and $K_1, K_2 \in \mathcal{K}$, such that $ST = \mathrm{Id} + K_1$ and $TS = \mathrm{Id} + K_2$. Such an S is called a **parametrix** (or **quasi-inverse**) for T. One also says that T is **essentially invertible**; i.e., invertible modulo $\mathcal{K}$.

EXERCISE 3.4. Suppose that K_1 and K_2 in Exercise 3.3 are trace class (see Section 2.7) and self-adjoint. Show that

$$\operatorname{index} T = \operatorname{Tr} K_1 - \operatorname{Tr} K_2. \tag{3.1}$$

[Hint. Using $K_1 = ST - \mathrm{Id}$ and $K_2 = TS - \mathrm{Id}$, show that $TK_1 = K_2T$ and $SK_2 = K_1S$. Using this, verify that if v (resp. w) is an eigenvector of K_1 (resp. K_2) with eigenvalue λ, then Tv (resp. Sw) is an eigenvector of K_2 (resp. K_1) with eigenvalue λ. If V_λ (resp. W_λ) is the eigenspace of K_1 (resp. K_2) for eigenvalue λ, then check that $ST|_{V_\lambda} = (1+\lambda)\,\mathrm{Id}_{V_\lambda}$ and $TS|_{W_\lambda} = (1+\lambda)\,\mathrm{Id}_{W_\lambda}$. Conclude that $T|_{V_\lambda} : V_\lambda \cong W_\lambda$ for $\lambda \neq -1$. Also check that $\operatorname{Ker} T = V_{-1}$ and $\operatorname{Coker} T \cong W_{-1} = \operatorname{Ker} S$. Verify that

$$\operatorname{Tr} K_2 - \operatorname{Tr} K_1 = \sum\nolimits_\lambda \lambda \dim W_\lambda - \sum\nolimits_\lambda \lambda \dim V_\lambda,$$

and all but two desirable terms cancel in the difference of these absolutely convergent sums.]

PROOF OF THEOREM 3.2. For "$\subseteq$", let $F \in \mathcal{F}$. We show that $\pi(F)$ is invertible. For this, consider the operator $F^*F + P$, where

$$P : H \longrightarrow \operatorname{Ker} F \text{ is orthogonal projection.}$$

In Remark 2.11 (p.17), we have already shown that $\operatorname{Ker} F^*F = \operatorname{Ker} F$ and $\operatorname{Im} F^*F = \operatorname{Im} F^*$; thus, $F^*F + P$ is bijective and hence invertible in $\mathcal{B}$. Since P is compact (being of finite rank), it follows that $\pi(F^*F) = \pi(F^*)\pi(F)$ is invertible in $\mathcal{B}/\mathcal{K}$. Similarly, one shows with the help of the orthogonal projection

$$Q : H \longrightarrow \operatorname{Ker} F^*$$

that $FF^* + Q$ in $\mathcal{B}$ and $\pi(F)\pi(F^*)$ in $\mathcal{B}/\mathcal{K}$ are invertible. With a left-inverse for $\pi(F^*)\pi(F)$ and a right-inverse for $\pi(F)\pi(F^*)$, it follows that $\pi(F)$ is invertible in $\mathcal{B}/\mathcal{K}$.

For "$\supseteq$", let $T \in \mathcal{B}$ with $\pi(T)$ invertible in $\mathcal{B}/\mathcal{K}$; i.e., there is $S \in \mathcal{B}$ such that TS and ST lie in $\pi(\mathrm{Id})$. Now, $\pi(\mathrm{Id}) = \{\mathrm{Id} + K : K \in \mathcal{K})\}$ consists of Fredholm operators by Theorem 2.31 (indeed, of index zero, but that does not concern us here). In particular, we then have that $\operatorname{Ker} ST$ and $\operatorname{Coker} TS$ are finite-dimensional. Since

$$\operatorname{Ker} T \subseteq \operatorname{Ker} ST \text{ and } \operatorname{Im} T \supseteq \operatorname{Im} TS,$$

it follows that $T \in \mathcal{F}$. □

REMARK 3.5. The trick in the first part of the above proof consists of first considering F^*F and FF^* (whose invertibility modulo $\mathcal{K}$ is trivial) rather than F, and only then drawing conclusions about $\pi(F)$. This has the advantage that one need not explicitly exhibit the *parametrix* (i.e., inverse modulo $\mathcal{K}$) for F. An explicit, if somewhat cumbersome, proof of the theorem of Atkinson can be found in [**Sche**, Theorems 5.4 and 5.5].

EXERCISE 3.6. Show that the set of Fredholm operators is open in the Banach algebra of bounded linear operators on a fixed Hilbert space H. [Hint: Because of the continuity of π (π is even contracting), it suffices to show that $(\mathcal{B}/\mathcal{K})^\times$ is open in $\mathcal{B}/\mathcal{K}$. For this, show in general that the group of units $\mathcal{A}^\times$ in any Banach algebra A is open; more precisely, show that about each $a \in \mathcal{A}^\times$ there is a ball of radius $1/\left\|a^{-1}\right\|$ contained in $\mathcal{A}$. For this, apply again the geometric series argument in the proof of Theorem 2.31 or from Exercise 3.1.]

EXERCISE 3.7. Conclude from the theorem of Atkinson that the space of Fredholm operators is closed under composition, the adjoint operation, and addition of compact operators. Show that such a conclusion is not circular, since the earlier proofs of the same results (e.g., Exercise 1.6 (p.5) and Theorem 2.10, (p.17)) were not needed in the proof of Atkinson's theorem.

EXERCISE 3.8. Illustrate Atkinson's theorem with the shift$^+$ operator (in Exercise 1.3, p. 4) on $L^2(\mathbb{Z}_+)$. In particular, show that the similarly defined shift$^-$ is a *parametrix* (= an inverse modulo $\mathcal{K}$) for shift$^+$. Which compact operators do we get for $(\mathrm{shift}^- \circ \mathrm{shift}^+) - \mathrm{Id}$ and for $(\mathrm{shift}^+ \circ \mathrm{shift}^-) - \mathrm{Id}$?

EXERCISE 3.9. Using the theorem of Riesz (Theorem 2.31, p.28), show that each parametrix G for a Fredholm operator F is itself a Fredholm operator, and we have $\operatorname{index} G = -\operatorname{index} F$.

EXERCISE 3.10. From Exercise 3.7, we know already that F is closed under addition of compact operators. Now show that the index is invariant:

$$\operatorname{index}(F+K) = \operatorname{index} F \text{ for all } F \in \mathcal{F} \text{ and } K \in \mathcal{K}.$$

[Hint: Show that each parametrix for F is also a parametrix for $F+K$, and apply Exercise 3.9.]

2. Perturbation Theory

The result of Exercise 3.10, which we obtained as an easy corollary of the Theorem 2.31 (p.28) of Riesz and of Theorem 3.2, is also due to Frederick Valentine Atkinson. It represents a fundamental result of perturbation theory which asks how the properties of a complicated system are related to those of an easy, ideal system close by whose properties are more easily computed or known. The idea comes from the variational calculus which asks the opposite way, namely determining optimal shapes of curves and surfaces (e.g., minimizing some energy functionals) by comparison with less advantageous neighbors, formalized in the famous Euler-Lagrange Equations and developed further in Morse Theory. In that context, the basic idea of homotopy was expressed by the young Giuseppe Lodovico (Luigi) Lagrangia (Lagrange) in his [**Lag**, Second letter to Euler, 12 August, 1755], to us the birth certificate of deformation theory and differential topology:

> "Differentiale ipsius y quatenus hic differentiatur, x manente, pro habendo maximo, minimove formulae datae valore, ad distinctionem aliarum eiusdem y differentiarum, quae in illa jam ingrediuntur, denotabo per δ; sic et δdy est differentia ipsius dy, dum

> y crescunt quantitate δy; idem dic generaliter de valore δFy [Fy mihi est functio quaecumque (emphasized by the authors) y]."[1]

Perturbation theory in a wider sense arose in celestial mechanics which tries to determine the deviations of planetary orbits from the *unperturbed* Keplerian paths due to the gravitational forces of other celestial bodies. While the methods used there point in a different direction, it is the perturbation theory of Lord Rayleigh (concerned with continuously extended oscillating systems) which leads frequently and typically to operators perturbed by the addition of a compact operator. This happens for example, when in elasticity the passage is made from constant mass density to variable density. See for example [**CH**, I, V.13]. That these are as a rule compact perturbations, is due to the fact that in the underlying partial and ordinary differential equations the terms of highest order remain unchanged and only the coefficients of the derivatives of lower order are modified. A theorem of Franz Rellich (see below Theorem 7.15, p. 198) explains why this produces compact perturbations.

Quantum mechanics poses farther reaching perturbation problems which are in parts mathematically unsolved. An example is the quantitative determination of energy levels of complicated systems of quantum mechanics.

The oscillations and motions of quantum mechanical systems are largely determined by the eigenvalues and eigenfunctions of the corresponding operators. Therefore, *perturbation theory* usually amounts to applying approximation methods to solving the eigenvalue problem of a complicated linear operator $T+K$ which differs *little* from a simpler T with a solved eigenproblem. Perturbation theory becomes spectral theory which studies the different constituents of the *spectrum* of an operator. A reference is the comprehensive exposition in [**Kat**].[2]

We will not pursue the physical applications any further here, since there is abundant motivation for perturbation theory within mathematics. Consider for instance the above mentioned calculus of variations of which local perturbations are an actual principle, or geometric questions which ask how much a curve (asymptotically or in its shape) or a surface, etc., changes if relevant parameters in their equations are modified. In particular, we are interested in the degree to which our quantitative invariants dim Ker, dim Coker, and index are independent of "small" perturbations. Here "small" does not exclusively mean that the dimension of the

---

[1]Our translation: "I shall denote the (peculiar) derivative of y, which is here to differentiate to obtain the largest or smallest value of a given formula while x remains unchanged, by δ - to distinguish it from the other differentiations of that y which already enter that formula; in such a way δdy denotes the difference just of dy, when (all) the y increase by a value δy; likewise speak generally of the value δFy [to me, Fy is an arbitrary function (emphasized by the authors) of y]."

[2]Kato's perturbation theory is incomparably deeper than our investigation: While we consider a single invariant, the index, Kato's theory is concerned with countably many real parameters associated with the power series expansion of the eigenvalues of a perturbed (symmetric) operator $T + cK$ where the parameters depend analytically on the perturbation. Just as one can classify symmetric matrices in linear algebra

- -according to their rank
- -projectively, according to their index of inertia (*Sylvester index*) , and
- -orthogonally, according to their diagonal elements (after *principal axis transformation*)

we have in the perturbation theory of operators in Hilbert space several levels of stability: index/essential spectrum (see below)/perturbation parameters of the power series expansion. In the crude mirror of finite-dimensional linear algebra, Kato's theory is closest to the principal axis transformation, while we restrict ourselves in index theory to consideration of the rank.

image of the perturbing operator is small, as with operators of finite rank and in a sense with compact operators, but may mean the perturbation is small in operator norm.

To get a feeling for the complications, we put together a list of established results:

1. The group of invertible elements of a Banach algebra is open, by Exercise 3.6. In particular, for each invertible, bounded, linear operator T, there is an ε ($:= \|T^{-1}\|^{-1}$) such that for all $S \in \mathcal{B}$ with $\|S\| < \varepsilon$, we have:

(i) $T + S \in \mathcal{F}$
(ii) $\operatorname{index}(T+S) = \operatorname{index} T$ $(= 0)$
(iii) $\dim \operatorname{Ker}(T+S) = \dim \operatorname{Ker} T$ $(= 0)$
(iv) $\dim \operatorname{Coker}(T+S) = \dim \operatorname{Coker} T$ $(= 0)$.

2. Further, by Exercise 3.10, for all $T \in \mathcal{F}$ and $K \in \mathcal{K}$

(i) $T + K \in \mathcal{F}$
(ii) $\operatorname{index}(T+K) = \operatorname{index} T$.

3. On the other hand, one can always find a perturbation of the identity by a compact operator K such that

$$\dim \operatorname{Ker}(\operatorname{Id} - K) > 10^{80}$$

making $\operatorname{Ker}(\operatorname{Id} - K)$ unimaginably large, since its dimension could not be matched by the atoms in a universe of "only" 10^{11} galaxies. Namely, select an orthonormal basis for H and define K as the orthogonal projection onto the linear span of the first $10^{80} + 1$ basis elements. However, $\operatorname{index}(\operatorname{Id} - K) = \operatorname{index} \operatorname{Id}$ $(= 0)$ by the Riesz Theorem (Theorem 2.31, p. 28).

4. In each arbitrarily small neighborhood of the zero operator there are Fredholm operators (namely, iterates of the shift operators multiplied by a small constant ε) with any large or small index; e.g.,

$$\operatorname{index}(0 + \varepsilon(\operatorname{shift}^+)^k) = k.$$

Hence, in the neighborhood of 0, the index behaves (metaphorically) as a holomorphic function in the neighborhood of an essential singularity (Theorem of Felix Casorati and Karl Weierstrass).

5. For the boundary-value problem

$$u'' + ru = 0, \ u(0) = u(1) = 0, \ r \in \mathbb{R}, \ r > 0,$$

treated in Chapter 2 (see Exercise 2.34, p.31), or the equivalent problem

$$v - rKv = 0,$$

where

$$Kv = (1-x)\int_0^x yv(y)\,dy + x\int_x^1 (1-y)v(y)dy,$$

it was already shown that

$$\dim \operatorname{Ker}(\operatorname{Id} - rK) = \begin{cases} 1, & \text{for } r = n^2\pi^2 \text{ and } n \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

While $\dim \operatorname{Ker}(\operatorname{Id} - rK)$ is not perturbable if it is zero (this is also clear because $\operatorname{Id} - rK$ is then invertible by the Riesz Theorem (Theorem 2.31, p. 28), whence Exercise 3.6 or the above result 1 applies), it is very prone to change when $r = n^2\pi^2$

– however, only in one direction: the dimension can only decrease. In other words, $\dim \operatorname{Ker}(\operatorname{Id} - rK)$ is upper semi-continuous; i.e.,

$$\dim \operatorname{Ker}(\operatorname{Id} - rK) \leq \dim \operatorname{Ker}(\operatorname{Id} - r_0 K) \text{ for all } r \text{ sufficiently close to } r_0.$$

6. Closed (not necessarily bounded) Fredholm operators with compact resolvent (typically elliptic differential operators of positive order on closed manifolds or on compact manifolds with smooth boundary subject to suitable boundary conditions) have *either* discrete spectrum *or* the whole set $\mathbb{C}$ as essential spectrum. Non-vanishing index implies the second case, by Exercise 3.10, p.65.

The following theorem shows, for arbitrary small (in the operator norm sense) perturbations what we already proved in Exercise 3.10 for compact perturbations: Even though the dimensions of the kernel of an operator and of its adjoint are not invariant under perturbations, the two jump by the same amount, so that their difference (the index) remains constant. The *perturbation-invariance* of the index is its most remarkable property. Together with the composition rule (Exercise 1.10, p.9, or Exercise 2.3, p.14), it shows that the index has properties analogous to homotopy invariants in *algebraic topology* such as the Euler characteristic $\chi(M)$ of a compact manifold M. Indeed, $\chi(M)$ is in fact the index of a certain operator (namely, $d+\delta$ from the space of even differential forms to the space of odd differential forms on M.

3. Homotopy Invariance of the Index

After these heuristic considerations, we now come to the aforementioned main theorem.

THEOREM 3.11 (J. Dieudonné, 1943). *With regard to the operatornorm topology, the mapping* $\operatorname{index} : \mathcal{F} \to \mathbb{Z}$ *is locally constant.*

PROOF. By Exercise 3.6, we already know that there is a neighborhood in $\mathcal{B}$ which is contained in $\mathcal{F}$. We now amplify the argument used there: By Theorem 3.2 (p. 64), we first choose a parametrix G for F; i.e., a $G \in B$ such that

$$FG = \operatorname{Id} + K_1 \quad \text{and} \quad GF = \operatorname{Id} + K_2,$$

where $K_1, K_2 \in \mathcal{K}$. We now show that for all $T \in \mathcal{B}$ with $\|T\| < \|G\|^{-1}$, we have $F + T \in \mathcal{F}$. Recall the *geometric series* argument (see the hint for Exercise 3.6), whereby the operators $\operatorname{Id} + TG$ and $\operatorname{Id} + GT$ are invertible, since $\|TG\|$ and $\|GT\|$ are less than 1. Thus $(\operatorname{Id} + GT)^{-1} G$ is a left inverse of $F + T$ modulo $\mathcal{K}$, since

$$\begin{aligned} (\operatorname{Id} + GT)^{-1} G\,(F + T) &= (\operatorname{Id} + GT)^{-1}\,(\operatorname{Id} + K_2 + GT) \\ &= \operatorname{Id} + (\operatorname{Id} + GT)^{-1} K_2. \end{aligned} \tag{3.2}$$

Similarly, $G(\operatorname{Id} + TG)^{-1}$ is a right inverse of $F + T$ modulo $\mathcal{K}$. Thus, by Theorem 3.2, we have $F + T \in \mathcal{F}$. Applying the composition rule (Exercise 2.3, p.14)) and the Riesz Theorem (Theorem 2.31, p. 28), we easily obtain from (3.2) the index formula

$$\operatorname{index}\big((\operatorname{Id} + GT)^{-1}\big) + \operatorname{index} G + \operatorname{index}(F + T) = 0.$$

Hence $\operatorname{index}(F + T) = \operatorname{index} F$, since the index of an invertible operator vanishes, and $\operatorname{index} G = -\operatorname{index} F$ by Exercise 3.9. □

REMARK 3.12. a) As mentioned in Theorem 2.63, the index remains locally constant (and, in fact, distinguishes the connected components) also in the unbounded case, even for varying domains.
b) The following example shows that the index of Fredholm operators in a Fréchet space is not a homotopy invariant.

EXAMPLE 3.13. As explained in Appendix A, the functions $z \mapsto z^k$, $k \in \mathbb{Z}$ form an orthonormal basis for the Hilbert space $L^2(S^1)$ of square integrable complex valued functions on the circle $S^1 = \{z : |z| = 1\}$. Let H_+ denote the (closed) subspace spanned by all z^k with k non-negative. Let $P_+ : L^2(S^1) \to H_+$ denote the orthogonal projection. Then each $f \in C^0(S^1)$ induces a bounded operator

$$T_f := P_+ \circ M_f|_{H_+} : H_+ \longrightarrow H_+,$$

where M_f denotes multiplication by f. We shall see below in Exercise 4.3 that $T_f \in \mathcal{B}(H_+)$ depends continuously of f. Moreover, Theorem 4.4 implies that T_f is a Fredholm operator if and only if $f(z) \neq 0$ for all $z \in S^1$. In that case we have $\operatorname{index} T_f = -W(f, 0)$, where $W(f, 0)$ denotes the winding number of f around the origin (e.g., $W(z^k, 0) = k$ for $k \in \mathbb{Z}$).
On the Fréchet space $C^\infty_+(S^1) := P_+\big(C^\infty(S^1)\big)$, each $f \in C^\infty(S^1)$ correspondingly induces a continuous linear operator

$$\tau_f := \pi_+ \circ M_f|_{C^\infty_+(S^1)} : C^\infty_+(S^1) \longrightarrow C^\infty_+(S^1),$$

where π_+ denotes the restriction of P_+ to $C^\infty_+(S^1)$. Similarly as in the preceding case, the operator τ_f depends continuously on f (in the respective Fréchet spaces). However, the domain of τ_f is much smaller than the domain of T_f, and so are the secondary sets. In particular, we now have that τ_f is Fredholm if and only if f has not more than a finite number of zeros, each of finite order. If f has no zeros at all, we still have $\operatorname{index} \tau_f = \operatorname{index} T_f = -W(f, 0)$. Now, τ_{z^k} can continuously be deformed into the identity within the space of Fredholm operators in our Fréchet space, e.g., by $f_t(z) := z^k - 2t$, $t \in [0, 1]$, since $z^k - 2$ clearly is homotopic to a constant function.

EXERCISE 3.14. Return to the space $\mathcal{F}$ of all bounded Fredholm operators in a fixed Hilbert space. Show that the index is constant on the connected components of $\mathcal{F}$.

EXERCISE 3.15. Show that

$$\dim \operatorname{Ker} : \mathcal{F} \longrightarrow \mathbb{N} \cup \{0\}$$

is upper semi-continuous; i.e., $\dim \operatorname{Ker} F$ does not *expand* suddenly when F is changed continuously. But it may well *shrink* suddenly. As an example, consider a continuous path F_t, $t \in [0, 1]$, in $\mathcal{F}$ connecting an invertible operator F_0 (with $\dim \operatorname{Ker} F_0 = 0$) to a non-invertible operator F_1 (with $\dim \operatorname{Ker} F_1 > 0$). More precisely, show that

$$\dim \operatorname{Ker}(F + T) \leq \dim \operatorname{Ker} F$$

for $F \in \mathcal{F}$ and $\|T\|$ sufficiently small.
[Hint: Show that $\operatorname{Ker}(F + T) \cap (\operatorname{Ker} F)^\perp = \{0\}$. Details are found in [**Sche**, Proof of Theorem 5.11].]

As an aside, we mention that Dieudonné (in [**Di43**]) proved Theorem 3.11 only implicitly without use of the index concept. The numerous interrelations of this theorem can be seen from the fact that by now a number of quite diverse proofs exist. All proofs have in common the reduction to the geometric series argument or the openness of the group of units of a Banach algebra.

This idea is most apparent in [**Do**, p.36f. and pp.133-148] where it is first shown that $A^\times/A_0^\times$ is discrete where A is an arbitrary Banach algebra (with identity), $A^\times$ its group of units, and $A_0^\times$ is the connected component containing the identity. There is an abstract index

$$i : \mathcal{A}^\times \longrightarrow \mathcal{A}^\times/\mathcal{A}_0^\times$$

defined in a natural way and whose continuity and hence local invariance is clear from the definition. The main task consists in making the connection between this ideally simple algebraic object and the *real* index. Less algebraic proofs can be found in [**Jö**, 1970/1982, 5.4], where the reduction to the openness of $\mathcal{B}^\times$ is achieved in a sequence of explicit extensions and projections which are computed in detail. The trick of fixing one dimension is carried out particularly elegantly in [**Ati69**, p.104]. As in [**Jö**] and in contrast to our proof above which is inspired by [**CaSc**, 12/06] and [**Sche**, Theorem 5.11], the Atiyah proof does not use the nontrivial theorems of Riesz and Atkinson and thus may be the most transparent proof on the whole. We will render it next.

Alternative Proof of Theorem 3.11. Let $e_0, e_1, ...$ be a complete orthonormal system for the Hilbert space H. We take H_n to be the closure of the linear span of the e_i with $i \geq n$, and we let P_n denote the orthogonal projection of H onto H_n.
Step 1: Clearly P_n is self-adjoint and $P_n \in \mathcal{F}$, since $\operatorname{Ker} P_n$ and $\operatorname{Coker} P_n$ are finite-dimensional. Hence $\operatorname{index} P_n = 0$, and for each $F \in \mathcal{F}$, we then have

$$\operatorname{index} P_n F = \operatorname{index} P_n + \operatorname{index} F = \operatorname{index} F.$$

Step 2: Since $\dim \operatorname{Coker} F < \infty$ ($F \in \mathcal{F}$), we have n_0 such that $e_0, e_1, ..., e_{n_0-1}$ together with $F(H)$ span H; in particular, for all $n \geq n_0$,

$$P_n F(H) = H_n \text{ and } \dim \operatorname{Coker} P_n F = n.$$

(Incidentally, we see that $\dim \operatorname{Coker} P_n F$ and also $\dim \operatorname{Ker} P_n F$ can be made arbitrarily large with n.)
Step 3: Although the function $\dim \operatorname{Ker}$ is only semi-continuous on $\mathcal{F}$, we claim that for G sufficiently near to F and n sufficiently large (as in Step 2)

$$\dim \operatorname{Ker} P_n G = \dim \operatorname{Ker} P_n F \text{ and } \dim \operatorname{Coker} P_n G = \dim \operatorname{Coker} P_n F.$$

For $G \in \mathcal{B}$ and $p : H \to \operatorname{Ker} P_n F$ the projection, consider the operator

$$\widehat{G} : H \longrightarrow H_n \oplus \operatorname{Ker} P_n F \text{ given by } \widehat{G}u := (P_n G u, pu).$$

If $G = F$, then $\widehat{F}$ is bijective, and hence has a bounded inverse by the Open Mapping Principle. Identifying $H_n \oplus \operatorname{Ker} P_n F$ with H, we then have $\widehat{F} \in \mathcal{B}^\times$. By the familiar argument in the hint to Exercise 3.6 (p. 65), there is a neighborhood of $\widehat{F}$ contained in $\mathcal{B}^\times$(the units, i.e., the invertible operators belonging to the algebra $\mathcal{B}$). Since "$G \mapsto \widehat{G}$" is continuous, there is also a neighborhood V of F such that for all $G \in V$ the operator $\widehat{G}$ is an isomorphism. From the surjectivity of $\widehat{G}$, it follows that $P_n G(H) = H_n$, whence $\dim \operatorname{Coker} P_n G = n = \dim \operatorname{Coker} P_n F$. Moreover, $\operatorname{Ker} P_n G = \widehat{G}^{-1}(\operatorname{Ker} P_n F)$, since by definition of $\widehat{G}$, a point u is mapped to $\operatorname{Ker} P_n F$ by $\widehat{G}$ exactly when the first component (i.e., $P_n G u$) of $\widehat{G}u$ vanishes. Since $\widehat{G}$ is an isomorphism, we then also have

$$\dim \operatorname{Ker} P_n G = \dim \widehat{G}^{-1}(\operatorname{Ker} P_n F) = \dim \operatorname{Ker} P_n F,$$

which establishes the above claim.

Summary: We have shown that for each $F \in \mathcal{F}$, there is a natural number n and $\eta > 0$ such that for all $G \in \mathcal{B}$ with $\|F - G\| < \eta$, we have

$$\operatorname{index} F = \operatorname{index} P_n F = \operatorname{index} P_n G = \operatorname{index} G.$$

Here, one could replace "index" by "Ker" or "Coker" in the inner equality, but not in the outer equalities. Moreover, $\operatorname{Im}(P_n F) = \operatorname{Im}(P_n G) = H_n$ and $\operatorname{Ker} P_n G = \widehat{G}^{-1}(\operatorname{Ker} P_n F)$. □

Exercise 3.16. Formulate Theorem 3.11 for a continuous family of Fredholm operators, by which we mean a continuous map $G : X \to \mathcal{F}$, where X is any topological space. More precisely, show that for all $x_0 \in X$, there is a neighborhood U and a natural number n, such that for all $x \in U$,

$$\operatorname{Im} P_n G(x) = H_n.$$

Then prove that the function

$$\dim \operatorname{Ker} P_n G : X \longrightarrow \mathbb{N} \cup \{0\}$$

is constant (say k) on U, and that there are k continuous functions

$$f_i : X \longrightarrow H, \ i = 1, ..., k,$$

such that for all $x \in U$, the points $f_1(x), ..., f_k(x)$ form a basis of $\operatorname{Ker}(P_n G(x))$. [Hint: Imitate the preceding alternate proof, replacing G by $G(x)$ and F by $G(x_0)$. Let $f_1(x_0), ..., f_k(x_0)$ be a basis of $\operatorname{Ker} P_n G(x_0)$ and set $f_i(x) := \widehat{G}(x)^{-1}(f_i(x_0))$.]

The concept of a continuous family of operators comes from classical analysis in the investigation of operators depending on a parameter or families of operators. In the simplest examples, the parameter space X is the unit interval, all $\mathbb{R}$, or a bordered domain in a higher-dimensional Euclidean space (e.g., the domain of permissible control variables). With somewhat more complex problems of analysis (e.g., as in the study of elliptic boundary value problems), we are quickly forced to consider families with more general parameter spaces: An elliptic differential equation defines a continuous family of Fredholm operators, where the parameter space is the sphere bundle of covectors of the underlying manifold restricted to the boundary; see Part 2, Chapter 10. We will first treat these questions not from the standpoint of applications, but rather out of natural topological-geometric considerations, namely interest in deformation invariants[3].

[3]Motivated by problems of optics (and also questions in astronomy, surveying, and architecture), the projective geometry of the 17th century originated in the idea of searching for properties of geometric figures which are invariant under transformations (central projection and cross section). In addition to these linear transformations, the concept of a deformation (i.e., the continuous change of a mathematical object) existed, for instance, when Johannes Kepler in 1604 noted that if the plane is compactified, then ellipse, hyperbola, parabola and circle can be transformed into one another by a continuous relocation of the foci (See [**Kli**, p.299]). But it was not until well into the 19th and 20th centuries that deformation invariants were found for a greater variety of mathematical objects. These include homology and cohomology theories, as presented axiomatically in [**ES**] for example, as well as the so-called K-theory, another branch of algebraic topology which was developed by Michael F. Atiyah and Friedrich Hirzebruch and is specifically aimed at the needs of analysis; see Part III below.

We assign to each compact parameter space X a group and to each continuous family of Fredholm operators

$$G : X \longrightarrow \mathcal{F}$$

we assign a group element, which is indeed invariant under deformation. This means that another continuous family of Fredholm operators $G' : X \to \mathcal{F}$ is assigned to the same group element, if G and G' are homotopic. By this, we mean that G can be continuously deformed to G' (see Chapter 11); i.e., there is a continuous family g of Fredholm operators parametrized by the product space $X \times I$, where $I = [0, 1]$,

$$g : X \times I \longrightarrow \mathcal{F}, \text{ such that } g|_{X\times\{0\}} = G \text{ and } g|_{X\times\{1\}} = G'.$$

If X consists of a single point, then a continuous family of Fredholm operators is just a single Fredholm operator, and the homotopy of G and G' clearly means that G and G' lie in the same (path) component of $\mathcal{F}$. In this way, we can answer fundamental questions concerning the nature of the connected components of F (e.g., via approximation theory).

Before we study continuous families of Fredholm operators (i.e., the geometry of $\mathcal{F}$ or the group of units $(\mathcal{B}/\mathcal{K})^\times$ by Theorem 3.11), we first turn to a simpler problem, the geometric investigation of the group of units $\mathcal{B}^\times$.

4. Homotopies of Operator-Valued Functions

This section and the following sections of this chapter are central for understanding index theory. They may be skipped in first reading and then later read in conjunction with Part III. In particular, here we use some concepts from topology which will be made precise only in Part III below.

Exercise 3.17. Let X and Y be topological spaces and f, g and h continuous maps from X to Y. Show that if f is homotopic to g and g is homotopic to h, then f is homotopic to h.

We recall some definitions and introduce some notation.

1. Two continuous maps f and g from X to Y are called **homotopic** (written $f \sim g$), if one can continuously deform one into the other; i.e., there is a continuous map $F : X \times I \to Y$ (where $I = [0, 1]$) such that

$$F \circ i_0 = f \text{ and } F \circ i_1 = g,$$

where $i_t : X \to X \times \{t\}$ is the canonical inclusion ($t \in I$). We write F_t for $F \circ i_t$, and can then roughly regard F as a 1-parameter family (over I) of continuous maps from X to Y. We call F a homotopy of f to g.

2. By the transitivity (Exercise 3.17) and obvious symmetry and reflexivity of the relation *homotopic*, it follows that the **homotopy classes**

$$\bar{f} := \{g \in C(X, Y) : g \sim f\}$$

where $C(X, Y)$ is the set of continuous functions from X to Y, and the homotopy set

$$[X, Y] := \left\{\bar{f} : f \in C(X, Y)\right\} \tag{3.3}$$

are well defined. Note that $[\text{point}, Y]$ corresponds to the set of pathwise connected components of Y.

3. Two topological spaces X and Y are **homeomorphic**, if there is a bijective map $f : X \to Y$ which is continuous in both directions.

4. Two topological spaces are **homotopy equivalent**, if there are continuous maps $f : X \to Y$ and $g : Y \to X$ such that $f \circ g \sim \mathrm{Id}$ and $g \circ f \sim \mathrm{Id}$. Clearly, the real line $\mathbb{R}$ and the plane $\mathbb{R}^2$ are homotopy equivalent and have the same *cardinality* (i.e., there is a bijection between them), but they are not homeomorphic, as we show in Part 3 (or just note that the removal of a point disconnects $\mathbb{R}$ but not $\mathbb{R}^2$.

5. Y is called a **retract** of X, if $Y \subseteq X$ and there is a continuous map $f : X \to Y$ with $f|_Y = \mathrm{Id}$. Such an f is called a **retraction**. If, in addition, $i \circ f \sim \mathrm{Id}$, where $i : Y \to X$ is inclusion, then Y is called a **deformation retract** of X, and X and Y are homotopy equivalent. Each $P \in X$ (or rather $\{P\}$) trivially constitutes a retract of X (but the sphere, as the boundary of the solid ball, is not a retract of the ball – see Part 3 below). If $\{P\}$ is a deformation retract of X, then X is called **contractible**. The shape of X must then be *starlike* in a certain sense.

Here we will study the homotopy type of operator spaces. Let $\mathcal{B}^\times(H)$ denote the group of invertible operators on a Hilbert space H that we allow to be a finite-dimensional complex vector space, say $\mathbb{C}^N$ in which case $\mathcal{B}^\times(H) = \mathrm{GL}(N, \mathbb{C})$.

Exercise 3.18. Investigate the group $\mathcal{B}^\times(H \times H)$ (or $GL(2N, \mathbb{C})$) of invertible operators on the product space $H \times H$, which can be written as 2×2 (block) matrices. For $R, S \in \mathcal{B}^\times(H)$ – or more generally for $R, S : X \to \mathcal{B}^\times(H)$ continuous with X a given topological space – show that

$$\begin{bmatrix} SR & 0 \\ 0 & \mathrm{Id} \end{bmatrix} \sim \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix}.$$

[Hint: Consider the map $F : X \times [0, \pi/2] \to \mathcal{B}^\times(H \times H)$, which is given by

$$F_t := \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & \mathrm{Id} \end{bmatrix} \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & \mathrm{Id} \end{bmatrix}.$$

Here, we have for brevity written $\cos t$ and $\sin t$ for the operators $\cos t \,\mathrm{Id}$ and $\sin t \,\mathrm{Id}$ $\in \mathcal{B}(H)$. Show that the image of F really lies in $\mathcal{B}^\times(H \times H)$ – not entirely in $\mathcal{B}^\times(H) \times \mathcal{B}^\times(H)$ – and investigate F_0 and $F_{\pi/2}$. Why can we use an interval $[a, b]$ $(a \neq b)$ of $\mathbb{R}$ different from I in the definition of a homotopy?]

The trigonometric functions which appeared in the preceding problem are typical of homotopy investigations of linear spaces in which rotations and compressions or dilations are the most important deformations. This considerably simplifies the explicit statement of homotopies. Of course, it does not simplify the demonstration of the non-existence of a homotopy since this forces one to consider <u>all</u> homotopies, a task which in general is solvable only with the <u>crude</u> means of algebraic topology; see Chapter 11 below.

Recall from linear algebra the fact that the group $\mathrm{GL}(N, \mathbb{C})$ of invertible complex $N \times N$-matrices contains the compact subgroup $\mathrm{U}(N)$, where $\mathrm{U}(N)$ consists of the unitary matrices of rank N; i.e.,

$$\mathrm{U}(N) := \left\{ A \in \mathrm{GL}(N, \mathbb{C}) : A^* = A^{-1} \right\},$$

where A^* is the conjugate transpose of A. Such A are matrices of $\mathbb{C}$-linear transformations of $\mathbb{C}^N$ which preserve the usual Hermitian inner product in $\mathbb{C}^N$. For a (complex) Hilbert space H, we have the group $\mathrm{U}(H) := \left\{ T \in \mathcal{B}(H) : T^* = T^{-1} \right\}$.

Exercise 3.19. If $R, S \in \mathrm{U}(H)$, show that the homotopy used in Exercise 3.18 does not leave $\mathrm{U}(H \times H)$.

EXERCISE 3.20. Regard $S^1 := \{z : z \in \mathbb{C} \text{ and } |z| = 1\}$ as a subset of $\mathbb{C} \times \{0\}$ (via $z \mapsto (z, 0)$) and choose $a \in S^1$ (e.g., $a = (1, 0)$). Construct a continuous map

$$g : S^1 \longrightarrow \mathrm{U}(2)$$

with the properties
(i) $(g(z))(z) = a$ for all $z \in S^1$
(ii) $g \sim f$, where $f(z) = \mathrm{Id}$ for all $z \in S^1$.
[Warning: The exercise would be trivial and solvable without using the 2nd dimension (i.e., within U(1) rather than U(2)) if S^1 were contractible. In that case the maps $f : z \mapsto 1$ and $g : z \mapsto az^{-1}$ would be homotopic as maps from S^1 to S^1 (or equivalently U(1)).]
[Hint: Reduce to Exercise 3.19 by setting

$$g(z) = \begin{bmatrix} az^{-1} & 0 \\ 0 & za^{-1} \end{bmatrix}. \quad]$$

We show that $\mathrm{U}(N)$ and $\mathrm{GL}(N, \mathbb{C})$ are pathwise connected as follows. The polar decomposition theorem of linear algebra states that every $g \in \mathrm{GL}(N, \mathbb{C})$ is (uniquely) a product AP, where $A \in \mathrm{U}(N)$ and $P \in H_N^+ :=$ the convex space of positive-definite (and hence invertible), $N \times N$ Hermitian matrices. Thus, if $\mathrm{U}(N)$ is path-connected, the multiplication map

$$\mathrm{U}(N) \times H_N^+ \longrightarrow \mathrm{GL}(N, \mathbb{C})$$

exhibits $\mathrm{GL}(N, \mathbb{C})$ as the continuous image of the path-connected space $\mathrm{U}(N) \times H_N^+$, and so the path-connectedness of $\mathrm{GL}(N, \mathbb{C})$ follows from that of $\mathrm{U}(N)$. To show that $\mathrm{U}(N)$ is connected, we may proceed as follows. If $e_N = (0, \ldots, 0, 1) \in \mathbb{C}^N$, then the map

$$f : \mathrm{U}(N) \longrightarrow S^{2N-1} \text{ given by } f(A) = Ae_N$$

is a continuous, open surjection with *fibers* $f^{-1}(f(A)) = A\mathrm{U}(N-1)$, homeomorphic to $\mathrm{U}(N-1)$. Suppose that $\mathrm{U}(N)$ is not connected. If $\mathrm{U}(N) = V_1 \cup V_2$ where V_1 and V_2 are nonvoid, open disjoint sets, then as S^{2N-1} is connected, $f(V_1) \cap f(V_2) \neq \phi$, say $Ae_N \in f(V_1) \cap f(V_2)$. Thus

$$A\mathrm{U}(N-1) = (V_1 \cap A\mathrm{U}(N-1)) \cup (V_2 \cap A\mathrm{U}(N-1)),$$

which implies that $A\mathrm{U}(N-1)$ is not connected. Continuing, we arrive at the contradiction that U(1) (a circle) is not connected. If $\mathfrak{u}(N)$ is the real vector space of skew-Hermitian matrices, the exponential map $\exp : \mathfrak{u}(N) \to \mathrm{U}(N)$ has differential Id at I_N and hence is a local homeomorphism about $0 \in \mathfrak{u}(N)$. It follows that $\mathrm{U}(N)$ is locally path connected. Finally, a connected, locally path-connected space is path-connected, since the path-components are then open and disjoint. In Part III below, we further investigate the homotopy type of $\mathrm{U}(N)$ which is only partially *known*. In contrast, we can show for *infinite* dimensional H that $\mathrm{U}(H)$ is contractible; see Remark 3.24 following Theorem 3.22 below.

THEOREM 3.21. *The group $\mathcal{B}^\times$ of invertible bounded linear operators on a Hilbert space H is pathwise connected.*

It is not true that the group of units of a Banach algebra is pathwise connected. The group of units $(\mathcal{B}/\mathcal{K})^\times$ in the Calkin algebra is a counterexample; its connected components – the connected components of Fredholm operators – are mapped bijectively to $\mathbb{Z}$ by the index, as the following paragraph shows.

This theorem is usually proved by means of deeper results of spectral theory (see, e.g., [**Do**, 1972, p.134ff]). One first shows that every unitary operator U has a spectral decomposition $U = e^{iA} = \cos A + i \sin A$ where A is a self-adjoint operator. Then, by the Spectral Theorem 2.61,

$$t \mapsto U_t := e^{itA}, \ t \in I$$

is a continuous path in $\mathrm{U}(H)$ from Id to U. (If one is willing to use spectral theory, this argument can replace the one above for the connectivity of $\mathrm{U}(N)$.) One shows further that each invertible operator R can be factored as $R = UB$ where U is unitary, and $B = \sqrt{R^* R}$ is self-adjoint, positive and invertible. Then one again connects U with Id using U_t and B with Id with the path

$$t \mapsto B_t := t \,\mathrm{Id} + (1-t)B, \ t \in [0,1]$$

which does not go outside $\mathcal{B}^\times$ by the Spectral Theorem, since B is positive and invertible. In this fashion $t \mapsto U_t B_t$ defines a path from R to Id.

Following an idea of Nicolaas Kuiper, here we provide a completely elementary proof of the theorem which perhaps is not as elegant as the proof outlined above and which (as most elementary proofs) requires more calculation and perhaps some more geometric imagination. The decisive advantage for us is that the elementary proof generalizes effortlessly to a proof of Kuiper's Theorem (Theorem 3.22) according to which $[X, \mathcal{B}^\times] = 0$, even if X does not consist of a single point as in Theorem 3.21 but is an arbitrary compact topological space. While the content of Theorem 3.21 remains unchanged in passing from $\mathbb{C}^N$ to the infinite-dimensional Hilbert space H, Theorem 3.22 exhibits a fundamental difference (see the Bott Periodicity Theorem in Chapter 11) between the linear algebra of finite-dimensional vector spaces and the functional analysis of Hilbert space. This aspect we can bring out clearly in the following proof of Theorem 3.21.

PROOF OF THEOREM 3.21. Let $R_0 \in \mathcal{B}^\times$. We seek a continuous path in $\mathcal{B}^\times$ connecting R_0 with Id. We proceed in two stages: In the first stage, we connect R_0 with an operator R_2 which is the identity on a cleverly constructed infinite dimensional subspace. In the second stage, we connect R_2 with the identity of H.
Stage 1, step 1: We begin by recursively constructing a sequence of unit vectors $a_1, a_2, ... \in H$ and a sequence of 2-dimensional subspaces $A_1, A_2, ... \subseteq H$, such that

$$\begin{aligned} & A_i \perp A_j \text{ for } i \neq j, \text{ and} \\ & a_i \in A_i, \ R_0 a_i \in A_i \text{ for all } i = 1, 2, \ldots . \end{aligned}$$

Start with any unit vector $a_1 \in H$ and a 2-dimensional subspace A_1 which contains a_1 and $R_0 a_1$. Then choose a unit vector

$$a_2 \in A_1^\perp \cap R_0^{-1}\left(A_1^\perp\right),$$

and let A_2 be a 2-dimensional subspace with $A_2 \perp A_1$ and containing a_2 and $R_0 a_2$; note that $a_2 \in R_0^{-1}\left(A_1^\perp\right) \Rightarrow R_0 a_2 \in A_1^\perp$. Then proceed with

$$a_3 \in A_1^\perp \cap A_2^\perp \cap R_0^{-1}\left(A_1^\perp\right) \cap R_0^{-1}\left(A_2^\perp\right), \text{ etc.}$$

The construction never breaks down, since the intersection of finitely many subspaces of finite codimension in H (recall $\dim H = \infty$) is never trivial.

Stage 1, step 2: Now we deform the operator R_0 to R_1 so that $R_1 a_i$ is a unit vector in the direction of $R_0 a_i$. Thus, define (for $t \in I$)

$$R_t u := \begin{cases} R_0 u, & \text{for } u \in \left(\bigoplus_{i=1}^{\infty} A_i\right)^{\perp}, \\ \left((1-t) + \frac{t}{|R_0 a_i|}\right) R_0 u, & \text{for } u \in A_i \end{cases}$$

Stage 1, Step 3: Deform the operator R_1 to an operator R_2 with the desired property

$$R_2 a_i = a_i \text{ for all } i.$$

This will be done by constructing a suitable curve $T_t \in \mathrm{U}(H)$ $(t \in [0,1])$ with $T_0 = \mathrm{Id}$, $T_t(A_i) = A_i$ and $T_1(R_1 a_i) = a_i$. With $R_2 := T_1 R_1$, we then will have $R_2 a_i = a_i$. Moreover, by construction, T_t will leave all vectors in $\left(\bigoplus_{i=1}^{\infty} A_i\right)^{\perp}$ fixed. The simple geometric construction of T_t (in which we have given up spatial intuition, since a complex plane, of $\mathbb{C}$-dimension 2, has $\mathbb{R}$-dimension 4) can be reduced to Exercise 3.20. Indeed, we map each A_i by an isometry α_i onto $\mathbb{C}^2$, in such a way that the complex line $\{\lambda R_1 a_i : \lambda \in \mathbb{C}\}$ is mapped to $\mathbb{C} \times \{0\}$. Then let $g_i : S^1 \to \mathrm{U}(2)$ be a map with the properties (guaranteed by Exercise 3.20):

(*i*) $(g_i(z))(z) = (1,0)$, for all $z \in S^1 \subset \mathbb{C} \times \{0\}$
(*ii*) There is a continuous map $F_i : S^1 \times I \longrightarrow \mathrm{U}(2)$ with $F_i(\cdot, 1) = g_i$ and $F_i(z,0) = \mathrm{Id}_{\mathbb{C}^2}$ for all z.

Let $B_t \in \mathrm{U}(2)$ $(t \in [0,1])$ be a curve chosen so that $B_0 = \mathrm{Id}$ and $B_1 \alpha_i(a_i) = (1,0) \in S^1 \subset \mathbb{C} \times \{0\}$. For $t \in I$, we now set

$$T_t u := \begin{cases} u, & \text{for } u \in \left(\bigoplus_{i=1}^{\infty} A_i\right)^{\perp}, \\ \alpha_i^{-1} B_t^{-1} F_i\left(\alpha_i\left(R_1 a_i\right), t\right) \alpha_i u, & \text{for } u \in A_i. \end{cases}$$

Then $T_0 = \mathrm{Id}$, and since $\alpha_i(R_1 a_i) \in S^1 \subset \mathbb{C} \times \{0\}$, we have

$$\begin{aligned} T_1(R_1 a_i) &= \alpha_i^{-1} B_1^{-1} F_i(\alpha_i R_1 a_i, 1) \alpha_i R_1 a_i \\ &= \alpha_i^{-1} B_1^{-1}(1,0) = \alpha_i^{-1}(\alpha_i(a_i)) = a_i. \end{aligned}$$

Thus,

$$t \quad \mapsto \quad R_{1+t} := T_t \circ R_1$$

is a continuous path in $\mathcal{B}^{\times}$ from R_1 to R_2 with $R_2|_{H'} = \mathrm{Id}$, where $H' \subseteq H$ denotes the infinite-dimensional closed subspace spanned by $a_1, a_2, \ldots$.

Stage 2, step 1: Relative to the decomposition $H = H_1 \oplus H'$, where $H_1 := (H')^{\perp}$ denotes the orthogonal complement of H' in H, R_2 has the form

$$\begin{bmatrix} Q & 0 \\ * & \mathrm{Id} \end{bmatrix},$$

where $Q \in \mathcal{B}^{\times}(H_1)$ and the perturbation term $*$ can be deformed to zero by a continuous path in $\mathcal{B}^{\times}(H)$

$$R_{2+t} = \begin{bmatrix} Q & 0 \\ (1-t)* & \mathrm{Id} \end{bmatrix}, \ t \in I; \text{ with } R_3 = \begin{bmatrix} Q & 0 \\ 0 & \mathrm{Id} \end{bmatrix}.$$

Stage 2, step 2: By the classical argument (which one uses in set theory to count $\mathbb{Q}$ or to demonstrate the equipotence of $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$), we can decompose H'

into an infinite sum of Hilbert spaces $H_2, H_3, ...$ Explicitly, let $a_1, a_2, ...$ form an orthonormal basis of H'. Decompose $\mathbb{N}$ into the infinite disjoint subsets

$$\mathbb{N}_j := \left\{2^{j-2}(2n-1) : n \in N)\right\},\ j = 2, 3, 4, ...$$

and take H_j to be the closed subspace spanned by the a_i with $i \in \mathbb{N}_j$. In this way we have $H = \bigoplus_{j=1}^{\infty} H_j$ (recall $H_1 = (H')^{\perp}$) and

$$R_3 = Q \oplus \mathrm{Id} \oplus \mathrm{Id} \oplus \cdots := \begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & \mathrm{Id} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathrm{Id} \end{bmatrix}.$$

If we now identify H_j and H_1 for $j \neq 1$ (all infinite-dimensional, separable, Hilbert spaces are trivially isomorphic), we can also write:

$$R_3 = Q \oplus \begin{bmatrix} QQ^{-1} & 0 \\ 0 & \mathrm{Id} \end{bmatrix} \oplus \begin{bmatrix} QQ^{-1} & 0 \\ 0 & \mathrm{Id} \end{bmatrix} \oplus \cdots .$$

With the rotation of Exercise 3.18, we obtain a continuous path in $\mathcal{B}^{\times}(H_1) \times \mathcal{B}^{\times}(H_1 \times H_1) \times \cdots \subseteq \mathcal{B}^{\times}(H)$ from R_3 to an operator

$$\begin{aligned} R_4 &= Q \oplus \left(Q^{-1} \oplus Q\right) \oplus \left(Q^{-1} \oplus Q\right) \oplus \cdots \\ &= \left(Q \oplus Q^{-1}\right) \oplus \left(Q \oplus Q^{-1}\right) \oplus \cdots \end{aligned}$$

and with one more rotation (this time in $\mathcal{B}^{\times}(H_1 \times H_1) \times \mathcal{B}^{\times}(H_1 \times H_1) \times \cdots)$ to a continuous path from R_4 to $R_5 = \mathrm{Id} \oplus \mathrm{Id} \oplus \cdots = \mathrm{Id}_H$. □

5. The Theorem of Kuiper

The idea of the last step of the preceding proof can be found in Albert Solomonovich Schwarz [**Schw64**] and in Klaus Jänich [**Ja**]. It bares a *secret* which separates fundamentally (from the topological standpoint) the linear functional analysis in Hilbert space from the linear algebra of finite dimensional vector spaces: It is the possibility of (figuratively speaking) escaping any squeeze by moving aside into a new dimension. If we had decomposed H into only finitely many components $H_1 \oplus \cdots \oplus H_m$, we would have gotten stuck in H_m, either with the homotopy from R_3 to R_4 or with the homotopy from R_4 to R_5 (depending on whether m is even or odd). We meet a similar phenomenon when investigating the geometry of unitary matrices where, e.g., the homotopy set $[S^i, \mathrm{U}(N)]$ (the well-known homotopy groups $\pi^i(\mathrm{U}(N))$, where S^i denotes the i-sphere of unit vectors in $\mathbb{R}^{i+1}$) is determined for $2N \geq i+1$ by the Periodicity Theorem of Raoul Bott but not known for all smaller values of N. Details are in Chapter 11.

Finally we wish to remark that generally in topology low-dimensional structures, particularly 3- and 4-dimensional manifolds (recall the key words Vaughan Jones and *knot theory*, Grigori Perelman's proof of the *Poincaré conjecture* or Friedrich Hirzebruch's work on the *signature*) are among the most difficult areas, while analogous questions for higher-dimensional objects were either solved decades ago or at least pose no fundamentally new problems. This is the background which may help understand the following basic theorem.

THEOREM 3.22. (N. Kuiper, 1964) *For any compact, topological space X, the homotopy set $[X, \mathcal{B}^{\times}(H)]$ consists of a single element, where $\mathcal{B}^{\times}(H)$ is the group of bounded, invertible operators in the Hilbert space H.*

REMARK 3.23. Just like Theorem 3.21 (X = point), Theorem 3.22 holds for nonseparable Hilbert spaces (see [**Il**, No. 284/02f.] and for real Hilbert spaces (see [**Kui**, p.19-30]). However, according to our earlier convention, we restrict ourselves always to separable complex Hilbert space.

REMARK 3.24. It is a corollary of Theorem 3.22 that $\mathcal{B}^\times(H)$ is contractible. This would be completely trivial if Theorem 3.22 were valid for non-compact X, for example for $X = \mathcal{B}^\times(H)$. But it is not this simple. Still there is a way (by studying the *nerve* of an open cover of $\mathcal{B}^\times(H)$ as in Stage 0 of the following proof) of reducing the question of contractibility to Theorem 3.22. See [**Kui**, p.27f.] and [**Il**, No. 284/01f.].

REMARK 3.25. Contrary to the topological investigation of the general linear group $\mathrm{GL}(N,\mathbb{C})$, there is no gain in restricting attention first to the group $\mathrm{U}(H)$ of unitary operators ($TT^* = T^*T = \mathrm{Id}$). In the *classical case* (see Chapter 11) the advantage results from the compactness of $\mathrm{U}(N) := \mathrm{U}(\mathbb{C}^N)$ but for $\dim H = \infty$, $\mathrm{U}(H)$ is not compact.

PROOF. We adopt the proof of Theorem 3.21 with the necessary modifications and several additional observations. In order to make a reasonable analogy between a continuous family $f : X \to \mathcal{B}^\times(H)$ of operators and a single operator $R \in \mathcal{B}^\times(H)$, say $R : \{point\} \to \mathcal{B}^\times(H)$, we must guarantee that $\operatorname{Im} f$ is at least contained in a finite-dimensional subspace of $\mathcal{B}(H)$. Thus, before we consider the analogy, we establish

Stage 0: Each continuous $f_0 : X \to \mathcal{B}^\times(H)$ with X compact is homotopic to a continuous map $f_1 : X \to \mathcal{B}^\times(H)$ with $f_1(X) \subseteq V$, where V is a finite-dimensional subspace of $\mathcal{B}(H)$.

To prove this, we use the openness of $\mathcal{B}^\times(H)$ (see Exercise 3.6, p. 3.6) and place an open ball contained in $\mathcal{B}^\times(H)$ about each operator $T \in f_0(X)$. This gives us an open cover $\mathcal{U}$ of $f_0(X)$. For safety reasons (see below), we replace each ball $U \in \mathcal{U}$ by a ball U' with the same center, but with 1/3 the radius. Clearly, $\mathcal{U}' := \{U' : U \in \mathcal{U}\}$ is an open cover of $f_0(X)$, since each operator in the image of f_0 is the center of some U'. As the continuous image of a compact set, $f_0(X)$ is also compact and it is covered by a finite subset of $\mathcal{U}'$. Thus, we have $f_0(X) \subseteq U_*$, where $U_* = \bigcup_{i=1}^N K(T_i, \varepsilon_i)$ is the union of the finite set of open balls

$$K(T_i, \varepsilon_i) := \{T \in \mathcal{B}(H) : \|T - T_i\| < \varepsilon_i\},\ i = 1, ..., N.$$

The balls are contained in $\mathcal{B}^\times(H)$ and are so small that $K(T_i, 3\varepsilon_i)$ is also contained in $\mathcal{B}^\times(H)$.

Then we are essentially done with the verification of stage 0: Although U_* is still infinite-dimensional, it is visibly contractible to a simplicial complex with vertices $T_1, ..., T_N$, meaning a structure that lies entirely in the subspace of dimension $\leq N$ of $\mathcal{B}(H)$ spanned by $T_1, ..., T_N$, see Figure 3.1.

Since intuition (particularly for infinite-dimensional spaces) can be deceiving, we will write out this argument precisely: For each $t \in [0,1]$ and $T \in U_*$, we define an operator

$$g_t(T) := (1-t)T + t\sum_{i=1}^N \phi_i(T)T_i,$$

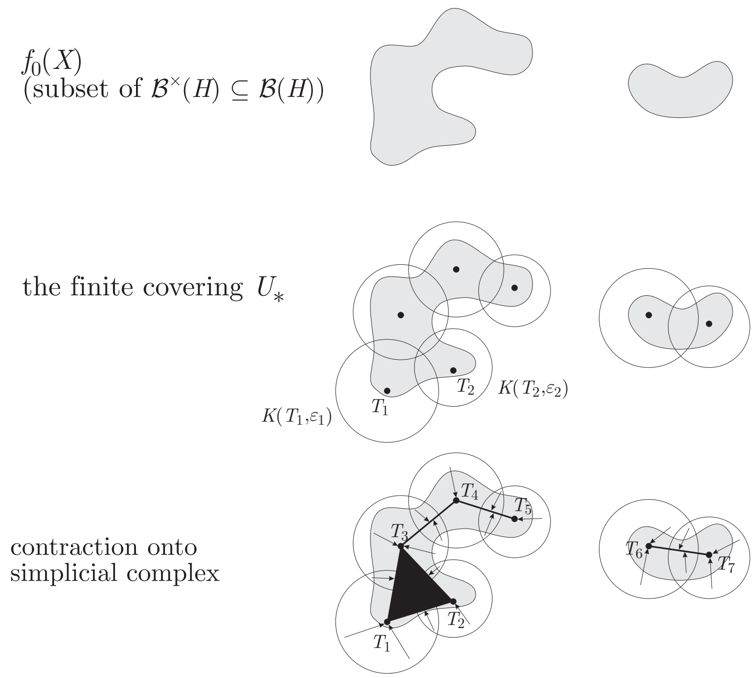


FIGURE 3.1. Two cases of visible contraction to a simplicial complex

where ϕ_i $(i = 1, ..., N)$ is a *partition of unity* subordinate to U_*; i.e., $\phi_i : U_* \to \mathbb{R}$ is continuous and

$$\begin{array}{ll} (i) & \text{support } \phi_i \subseteq \overline{K(T_i, \varepsilon_i)};\ \text{i.e., } \phi_i(T) = 0 \text{ for } \|T - T_i\| \geq \varepsilon_i, \\ (ii) & 0 \leq \phi_i(T) \leq 1 \\ (iii) & \sum_{i=1}^N \phi_i(T) = 1 \text{ for } T \in U_* \end{array}$$

For example, set

$$\phi_i(T) := \frac{\psi_i(T)}{\sum_{i=1}^N \psi_i(T)} \text{ for } T \in U_*, \text{ where}$$

$$\psi_i(T) := \begin{cases} \varepsilon_i - \|T - T_i\|, & \text{for } T \in K(T_i, \varepsilon_i), \\ 0, & \text{otherwise.} \end{cases}$$

In this way, $g_0 : U_* \to \mathcal{B}^\times(H)$ is the inclusion and $g_1 : U_* \to \mathcal{B}^\times(H)$ is a retraction of U_* onto a simplicial complex (composed of points, segments, triangles, tetrahedra, and corresponding higher-dimensional simplices) with vertices $T_1, ..., T_N$. To be sure that the homotopy of g_0 to g_1 does not leave $\mathcal{B}^\times(H)$, we use an "$\varepsilon/3$-argument" (see Figure 3.2): Let $T \in U_*$. In considering $g_t(T)$, we are (because of (i)) only interested in those summation indices i for which $T \in K(T_i, \varepsilon_i)$. We compare these balls and let $K(T_m, \varepsilon_m)$ be the one with the largest radius. By the triangle inequality, each of these balls is contained in $K(T_m, 3\varepsilon_m)$. Thus, the convex hull of T and *these* T_i (and hence $g_t(T)$, $0 \leq t \leq 1$) is contained in the ball $K(T_m, 3\varepsilon_m)$, which by construction lies in $\mathcal{B}^\times(H)$.

With $f_t := g_t f_0$, $t \in I$, we obtain a homotopy in $\mathcal{B}^\times(H)$ of f_0 to an $f : X \to \mathcal{B}^\times(H)$ with the desired properties. This completes the essential work in carrying

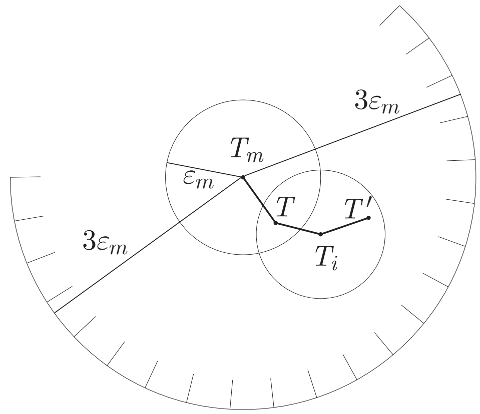


FIGURE 3.2. Keeping the homotopy of g_0 to g_1 in $\mathcal{B}^\times(H)$ by an $\varepsilon/3$-argument

over the proof of Theorem 3.21 to the more general situation of Theorem 3.22. Indeed, the rest of the proof now is quite analogous; we no longer consider R_0 as a single element of $\mathcal{B}^\times(H)$, but rather as the intersection of the finite-dimensional subspace V (in which $f_1(X)$ lies) with $\mathcal{B}^\times(H)$. Then we need only show that $\{\mathrm{Id}\}$ is a deformation retract of span $(V, \mathrm{Id}) \cap \mathcal{B}^\times(H)$, just as we showed in the proof of Theorem 3.21 that R_0 and Id are connected by a continuous path in $\mathcal{B}^\times(H)$, and we are done. The proofs are identical in principle. In the particulars, we need the following modifications:

Stage 1, step 1: We construct as above a sequence of unit vectors $a_1, a_2, \ldots \in H$ and a sequence of pairwise orthogonal $(n+1)$-dimensional subspaces $A_1, A_2, \ldots \subset H$ such that $a_i \in A_i$ and $Ra_i \in A_i$ for all $R \in V$ and $i = 1, 2, \ldots$; here, $n := \dim V$ and V is the vector space constructed in stage 0, except for containing in addition the operator Id; i.e., replace V by span (V, Id).

Stage 1, steps 2 and 3: Now we show that the canonical inclusion

$$\gamma_0 : V \cap \mathcal{B}^\times(H) \hookrightarrow \mathcal{B}^\times(H)$$

is homotopic to a map

$$\gamma_2 : V \cap \mathcal{B}^\times(H) \longrightarrow \mathcal{B}^\times(H)$$
$$\text{with } \gamma_2(R)(a_i) = a_i \text{ for } i = 1, 2, \ldots \text{ and } R \in V \cap \mathcal{B}^\times(H).$$

As before, in stage 1 step 2 (p. 76), we first go from γ_0 to γ_1 with

$$\gamma_1(R)\,u := \begin{cases} Ru, & \text{for } u \in \left(\bigoplus_{i=1}^\infty A_i\right)^\perp, \\ \frac{Ru}{|Ra_i|}, & \text{for } u \in A_i. \end{cases}$$

For step 3, one must again take up the rotation argument of Exercise 3.20 and generalize it somewhat by induction: We regard

$$S^{2n-1} = \left\{ z = (z_1, \ldots, z_n) \in \mathbb{C}^n : |z_1|^2 + \ldots + |z_n|^2 = 1) \right\}$$

as a subset of $\mathbb{C}^n \times \{0\} \subseteq \mathbb{C}^{n+1}$ and construct, for each $a \in S^{2n-1}$, a continuous map $g : S^{2n-1} \to U(n+1)$ with the properties

$$(i)\ g(z)(z) = a \text{ for all } z \in S^{2n-1} \text{ and}$$
$$(ii)\ g \sim h, \text{ where } h(z) = \mathrm{Id}_{\mathbb{C}^{n+1}} \text{ for all } z \in S^{2n-1}$$

(see also [**Il**, No. 284/04]). We reduce our problem to this situation by mapping each A_i by an isometry α_i onto $\mathbb{C}^{n+1}$ so that $\{Ta_i : T \in V\}$ is mapped into $\mathbb{C}^n \times \{0\}$. Then, if $F_i : S^{2n-1} \times I \to \mathrm{U}(n+1)$ is the corresponding homotopy for $a := \alpha_i(a_i)$, then the following is a homotopy from γ_1 to γ_2 with the desired properties:

$$\gamma_{1+t}(T)\,u := \begin{cases} \gamma_1(T)\,u, & \text{for } u \in \left(\bigoplus_{i=1}^{\infty} A_i\right)^{\perp}, \\ \alpha_i^{-1} F_i\left(\alpha_i \gamma_1(T)\,a_i, t\right) \alpha_i \gamma_1(T)\,u, & \text{for } u \in A_i. \end{cases}$$

In particular, we have an infinite-dimensional H' (with orthonormal basis $a_1, a_2, \ldots$) such that $\gamma_2|_{H'} = \mathrm{Id}$.

Stage 2: In the proof of Theorem 3.21, we have already implicitly shown that $\mathcal{B}^{\times}((H')^{\perp}) \times \mathrm{Id}_{H'}$, is a deformation retract of the group of all invertible operators on H which are the identity on H' (step 1), and that $\mathcal{B}^{\times}((H')^{\perp}) \times \mathrm{Id}_{H'}$, can be contracted to $\{\mathrm{Id}_H\}$ (step 2). The proof of Theorem 3.22 is now complete. □

EXERCISE 3.26. Prove the Theorem of Dixmier (and Douady) that shows that $\mathcal{U}(H)$ is contractible in the strong operator topology. Recall or check, e.g., in [**Ped**, Section 4.6, p.171]: The **strong topology** on $\mathcal{B}(H)$ is induced by the family of seminorms of the form $T \mapsto \|Tx\|$ for various $x \in H$. Since $\|Tx\| \le \|T\|\,\|x\|$, we immediately observe that the strong topology is weaker than the norm topology. Kuiper's theorem is about the norm topology and is much harder (although there are now, depending on taste, more conceptual proofs than ours which is Kuiper's original one). The idea is very simple.

[Hint, following [**Dix**, Lemma 10.8.2]: First realize H as $L^2[0,1]$ and consider the strongly continuous family P_t of orthogonal projections onto the subspaces $L^2[0,t]$, $t \in [0,1]$. Of course P_t is given by multiplication by the characteristic function of $[0,t]$. Now for $t > 0$, let V_t denote the obvious isometry of $L^2[0,t]$ onto $L^2[0,1]$, namely $V_t(x)(s) = \sqrt{t}x(ts)$ for $s \in [0,1]$ and $x \in L^2[0,t]$, and let $V_0 = 0$. Then $t \mapsto V_t$ is strongly continuous. The contraction of $\mathcal{U}(H)$ to 1 is given by $(u,t) \mapsto (1 - P_t) + P_t (V_t)^* u V_t P_t$. At time $t = 1$, this is just the identity map $u \mapsto u$ on $\mathcal{U}(H)$, while at time $t = 0$, this is the map $u \mapsto 1$. Check that $V_t^*(y)(\sigma) = \frac{1}{\sqrt{t}} y(\frac{\sigma}{t})$ for $\sigma \in [0,t]$, and verify the unitarity of $(1 - P_t) + P_t (V_t)^* u V_t P_t$ for $t \in (0,1)$.]

6. The Topology of $\mathcal{F}$

Like the preceding one, this section is central for index theory but may be skipped in first reading and read later in connection with the study of the topology of the general linear group (Bott's Periodicity Theorem, Chapter 11) and the topological interpretation of elliptic boundary value problems (Chapter 10 and Section 14.8).

Our point of departure is the principal theorem on the *homotopy invariance* of the index saying that the index is defined in a neighborhood of a Fredholm operator and is constant there, and consequently on each connected component (Theorem 3.11, p.68). Regarding the topology of $\mathcal{F}$, the space of Fredholm operators on a given Hilbert space H, the following questions arise:

1. What is the *number* of connected components of $\mathcal{F}$?
2. What can be said about the *structure* of the individual connected components? How many *holes* are there of each type?

Conceptually imagine a serving of Swiss cheese. Not only do we note the numbers of slices (Question 1) but also the kind of holes (Question 2), those that cross the cheese like a channel and those which are enclosed like air bubbles.

We already have a partial answer for Question 1. There are at least $\mathbb{Z}$ connected components, since the right-handed and left-handed shift, together with its iterates (the natural powers), show that every integer can be the index of a Fredholm operator. We denote the path connected components of $\mathcal{F}$ by $[point, \mathcal{F}]$. Then the map index : $[point, \mathcal{F}] \to \mathbb{Z}$ is well-defined by Theorem 3.11, p.68, (homotopy invariance of the index) and surjective. In fact, the map is injective also as we will prove in this chapter. This answers Question 1 completely. In particular, it is utterly impossible that $\mathcal{F}$ or the group of units $(\mathcal{B}/\mathcal{K})^\times$ of the quotient algebra is contractible, in contrast to $\mathcal{B}^\times$ (Theorem 3.21, p.74). In addition, the individual connected components of $\mathcal{F}$ differ as to homotopy type (Question 2) radically from $\mathcal{B}^\times$, which has an extremely simple structure according to the Theorem of Kuiper (Theorem 3.22, p.77). In fact, we will show that the *holes* of $\mathcal{F}$, its *fissures*, can be arbitrarily complex in some sense, that $\mathcal{F}$ is a kind of model for *all possible* topological structures (distinguishable via the *functor* K; see below). We can sketch this aspect before starting with formal definitions and theorems, as follows: In algebraic topology, for instance in the "K-theory" treated in Part 3, one has methods which assign to certain topological spaces (compact or triangulable spaces, differentiable manifolds, etc.) certain algebraic objects (groups, rings, algebras, etc.) and to continuous (differentiable etc.) maps certain homomorphisms between the associated algebraic objects. In this fashion certain topological-geometric phenomena, which are very difficult to distinguish in concrete visualization, can be reduced to algebraic terms (*made discrete*, F. Waldhausen) for which there is a well-developed formalism that is easier to comprehend. For example, it is immediate that there is no epimorphism of the group $\mathbb{Z}$ onto the group $\mathbb{Z} \oplus \mathbb{Z}$, while the non-existence of a retraction of the n-dimensional ball B^n onto its boundary S^{n-1} is not immediately clear (see Theorem 11.21). The construction of the *index bundle* which will be treated in this chapter is the basis of a further step from algebraic topology to functional analysis or to *elliptic topology* (Atiyah). We will be able to explain this term only in the following chapters which deal with the connection between elliptic differential equations and Fredholm operators. In this way, deep *geometric* questions in the proof of Bott's Periodicity Theorem (generalization of the concept *winding number*) find an *algebraic* formalism in K-theory. Thereby the *Bott isomorphism* $K(X \times \mathbb{R}^2) \to K(X)$ is described by means of families of Fredholm operators, and in this form can be understood more easily and elementarily by a symmetric use of classical results of functional analysis (see Chapter 11). This is an example where the interpretation in *functional analysis* makes the understanding of geometric or algebraic situations easier or perhaps only possible.

Conversely, the topological and algebraic problems and methods serve analysis: for example, when the index or the index bundle yield algebraic-topological invariants for Fredholm operators or families of Fredholm operators which turn up concretely in problems of analysis.

7. The Construction of Index Bundles

We now come to the construction of *index bundles*. Let $T : X \to \mathcal{F}$ be a continuous family of Fredholm operators in a Hilbert space H, where X is a compact topological space. If X is connected, Theorem 3.11 (p.68) gives

$$\operatorname{index} T_x = \operatorname{index} T_{x'}, \text{ for all } x, x' \in X.$$

In this way, we can assign to each T an integer, which is independent of possible *small continuous perturbations* of T. For X connected, we therefore have a map

$$\operatorname{index} : [X, \mathcal{F}] \longrightarrow \mathbb{Z},$$

which is well defined on the homotopy set $[X, \mathcal{F}]$; see (3.3), p.72. Actually, we can extract much more information out of T.

Exercise 3.27. Show that for each continuous family $T : X \to \mathcal{F}$ with constant kernel dimension, i.e., $\dim \operatorname{Ker} T_x = \dim \operatorname{Ker} T_{x'}$, for all $x, x' \in X$, one can assign a vector bundle $\operatorname{Ker} T$ over X, in a natural way. Here, a **vector bundle** E over X is a *continuous locally-trivial family* of complex vector spaces E_x of finite dimension, parameterized by the *base space* X. For the details of the definition, we refer to the Appendix.
[Hint: Set $\operatorname{Ker} T := \cup_{x \in X} \{x\} \times \operatorname{Ker} T_x$, and give this the induced topology that it inherits as a subset of $X \times H$. Then show, as in the alternative proof (see p.70) of Theorem 3.11, the property of local triviality; i.e., locally there is a basis of $\operatorname{Ker} T_x$ which depends continuously on x. See also Exercise B.2c (p. 678) of the Appendix, where $X = S^1$.]

Exercise 3.28. Under the same assumptions as in Exercise 3.27, show that there are naturally defined vector bundles $\operatorname{Ker} T^*$ and the *quotient bundle* $\operatorname{Coker} T$. Moreover, show that these bundles are *isomorphic* (see Appendix).

We denote the set of all isomorphism classes of vector bundles over X by $\operatorname{Vect}(X)$. By Exercises 3.27 and 3.28, we have a map ι from the set $C(X, \mathcal{F})$ of continuous families of Fredholm operators with constant kernel dimension to $\operatorname{Vect}(X) \times \operatorname{Vect}(X)$:

$$\iota : C(X, \mathcal{F}) \longrightarrow \operatorname{Vect}(X) \times \operatorname{Vect}(X) \text{ given by}$$
$$\iota(T) := ([\operatorname{Ker} T], [\operatorname{Coker} T]).$$

If X consists of a single point, then a vector bundle over X is simply a single vector space, and $\operatorname{Vect}(X)$ can be identified with $\mathbb{Z}_+$, since vector spaces are isomorphic precisely when their dimensions are equal; symbolically, $[\cdot] = \dim(\cdot)$. In this case, we then have the maps

$$\mathcal{F} \xrightarrow{\iota} \mathbb{Z}_+ \times \mathbb{Z}_+ \xrightarrow{\delta} \mathbb{Z} \text{ given by}$$
$$T \overset{\iota}{\mapsto} ([\operatorname{Ker} T], [\operatorname{Coker} T]) \overset{\delta}{\mapsto} [\operatorname{Ker} T] - [\operatorname{Coker} T],$$

where δ is the difference mapping and $\delta \circ \iota = \operatorname{index}$.

In the more general case where X is not a single point, we can formally write such a difference $[\operatorname{Ker} T] - [\operatorname{Coker} T]$, at least when the family T has constant kernel dimension. Admittedly, this is meaningless for the moment: While one can naturally introduce an addition $\oplus$ in $\operatorname{Vect}(X)$ by forming the direct sum pointwise (see Exercise B.4, p. 679, of the Appendix), this only makes $\operatorname{Vect}(X)$ a semigroup. However, one can go from the abelian semigroup $\operatorname{Vect}(X)$ to an abelian group (just as from $\mathbb{Z}_+$ to $\mathbb{Z}$), which we denote by $K(X)$ and define as follows. An equivalence relation on the product space $\mathrm{Vect}(X) \times \mathrm{Vect}(X)$ is defined by means of

$$(E, F) \sim (E \oplus G, F \oplus G) \text{ for } G \in \mathrm{Vect}(X),$$

and then

$$K(X) := (\mathrm{Vect}(X) \times \mathrm{Vect}(X))/\sim .$$

The equivalence class of the pair (E, F) is then written as $E - F$; these we call *difference bundles*.

The details of this construction, and its basic importance for algebraic topology, is explained in Section 11.3. Here we only need to establish that the difference map δ extends from $\mathbb{Z}_+ \times \mathbb{Z}_+$ to $\mathrm{Vect}(X) \times \mathrm{Vect}(X)$ in a canonical way so that its values form an abelian group $K(X)$. (To be careful, X is always compact in this chapter, but many of the steps carry over easily to more general cases.)
[Warning: In spite of the close analogy between the construction of $K(X)$ and $\mathbb{Z}$, notice that (for $X \neq$ point) the natural map

$$\mathrm{Vect}(X) \longrightarrow K(X), \text{ given by } E \mapsto E - 0$$

need not be injective. In Chapter 11, we will get to know vector bundles E and F over the two-dimensional sphere S which are not isomorphic, but when the same vector bundle G is added to each of them, the results are isomorphic. (Visually, one may note that the tangent bundle of S^2 and the real two-dimensional trivial bundle over S^2 are not isomorphic, but the addition of a trivial one-dimensional bundle to each yields isomorphic bundles; see the Appendix, Exercise B.11a, p. 683).]

With the help of the maps ι and δ just introduced, we can now deduce from Exercise 3.27 and Exercise 3.28:

EXERCISE 3.29. To each continuous family $T : X \to F$ of Fredholm operators, with constant kernel dimension, there can be assigned a difference bundle (**index bundle**) $\mathrm{index}\, T := [\mathrm{Ker}\, T] - [\mathrm{Coker}\, T] \in K(X)$. Moreover, for a point space X (T is then a single Fredholm operator and $K(X) \cong \mathbb{Z}$), the concepts of *index* and *index bundle* coincide.
[Hint: Set $\mathrm{index} := \delta \circ \iota$.]

Now we will show that the unrealistic condition that the kernel dimension be constant can be dropped, and the index bundle is invariant under continuous deformation just as the index (Theorem 3.11, p.68).

THEOREM 3.30. *Let X be compact. For each continuous family $T : X \to \mathcal{F}$ of Fredholm operators in a Hilbert space H, there is an index bundle*

$$\mathrm{index}\, T \in K(X)$$

assigned in a canonical way.

PROOF. As in the alternative proof (see p.70) of Theorem 3.11, we first choose an orthonormal system $e_0, e_1, \ldots$ for H and consider the Fredholm operator $P_n T_x$ (which has the same index as T_x), where P_n is again the projection of H onto the closed subspace H_n spanned by $e_n, e_{n+1}, \ldots$. Since X is compact, we can choose n such that

$$\mathrm{Im}(P_n T_x) = H_n \text{ for all } x \in X, \text{ and then}$$
$$\dim \mathrm{Ker}\, P_n T_x = \dim \mathrm{Ker}\, P_n T_{x'} \text{ for all } x, x' \in X.$$

Indeed, for each $y \in X$, there is a natural number n_y and a neighborhood U_y so that for all $x \in U_y$, T_x will be sufficiently close to T_y to ensure (by means of the alternative proof, p.70) that $\dim \operatorname{Ker} P_n T_x = \dim \operatorname{Ker} P_n T_y$ for all $n \geq n_y$. Then we pass to a finite subcover $\{U_y : y \in Y\}$, where Y is a finite subset of X, and set $n := \max(n_y : y \in Y)$. Relative to the family $P_n T$, we can therefore (as in Exercise 3.29) set

$$\begin{aligned}\operatorname{index} T &:= \operatorname{index} P_n T = [\operatorname{Ker} P_n T] - [\operatorname{Coker} P_n T] \\ &= [\operatorname{Ker} P_n T] - [X \times (H_n)^{\perp}].\end{aligned}$$

Thus, we have assigned an index bundle in $K(X)$ to T. Indeed, we have expressed the index bundle in *normal form*, in the sense that the bundle subtracted is trivial.

We must show that the definition is independent of the sufficiently large natural number n. Without loss of generality, replace n by $n+1$. Then, by construction,

$$[\operatorname{Coker} P_{n+1} T] = [X \times (H_{n+1})^{\perp}] = [\operatorname{Coker} P_n T] \oplus [X \times \mathbb{C} e_n].$$

To calculate $[\operatorname{Ker} P_{n+1} T]$, we note that for all $x \in X$,

$$P_n T_x|_{(\operatorname{Ker} P_n T_x)^{\perp}} : (\operatorname{Ker} P_n T_x)^{\perp} \longrightarrow H_n$$

is bijective. Hence, by the closure of H_n and the *open mapping principle*, there is a bounded inverse operator

$$\widetilde{T}_x : H_n \longrightarrow (\operatorname{Ker} P_n T_x)^{\perp} \subseteq H.$$

We set

$$\begin{aligned}&v_x := \widetilde{T}_x(e_n) \in (\operatorname{Ker} P_n T_x)^{\perp}, \text{ so that} \\ &P_n T_x v_x = e_n \text{ for all } x \in X.\end{aligned}$$

Then for all $x \in X$, we have

$$\begin{aligned}\operatorname{Ker} P_{n+1} T_x &= \operatorname{Ker} P_n T_x \oplus \mathbb{C} v_x \text{ and} \\ [\operatorname{Ker} P_{n+1} T_x] &= [\operatorname{Ker} P_n T] \oplus [\{(x, z v_x) : x \in X, z \in \mathbb{C}\}].\end{aligned}$$

As with T, the family $\widetilde{T}$ is also continuous (prove!). Thus, $\widetilde{T}$ yields an isomorphism of bundles

$$X \times \mathbb{C} e_n \longrightarrow \{(x, z v_x) : x \in X, z \in \mathbb{C}\},$$

(i.e., $x \mapsto v_x$ is a continuous, nowhere zero section over X in $X \times H$). Thus, the index bundles $\operatorname{index}(P_{n+1} T)$ and $\operatorname{index}(P_n T)$ are equal in $K(X)$.

Finally, we must show the independence of the choice of orthonormal basis. For this problem we use the *independence of* n just shown. Let $\widetilde{e}_0, \widetilde{e}_1, \ldots$ be another complete orthonormal system for H, and let $\widetilde{P}_m : H \to \widetilde{H}_m$ denote the orthogonal projection onto the closed subspace $\widetilde{H}_m$ spanned by $\widetilde{e}_m, \widetilde{e}_{m+1}, \ldots$. There is $\widetilde{m}_0 \in \mathbb{N}$ such that for all $m \geq \widetilde{m}_0$, we have $\operatorname{Im}\left(\widetilde{P}_m T_x\right) = \widetilde{H}_m$. As H_n and $\widetilde{H}_m$ both have finite codimension, so does $H_n \cap \widetilde{H}_m$. Since $\operatorname{Ker} P_n T$ is the same if we replace $e_n, e_{n+1}, \ldots$ by another orthonormal system for H_n, we may assume that $e_n, e_{n+1}, \ldots$ is adapted to $H_n \cap \widetilde{H}_m$ in the sense that for some k, we have that $e_{n+k}, e_{n+k+1}, \ldots$ is a complete orthonormal system for $H_n \cap \widetilde{H}_m$. Similarly, we may assume that $\widetilde{e}_m, \widetilde{e}_{m+1}, \ldots$ is chosen so that $\left(\widetilde{e}_{m+\widetilde{k}}, \widetilde{e}_{m+\widetilde{k}+1}, \ldots\right) = (e_{n+k}, e_{n+k+1}, \ldots)$ for some $\widetilde{k}$. Then $(e_{n+k}, e_{n+k+1}, \ldots)$ is a common tail of both sequences $e_n, e_{n+1}, \ldots$ and $\widetilde{e}_m, \widetilde{e}_{m+1}, \ldots$. The orthogonal projection onto the span of this common tail (i.e., onto $H_n \cap \widetilde{H}_m$) is $P_{n+k} = \widetilde{P}_{m+\widetilde{k}}$. Thus, by the above

$$\operatorname{index} P_n T = \operatorname{index} P_{n+k} T = \operatorname{index} \widetilde{P}_{m+\widetilde{k}} T = \operatorname{index} \widetilde{P}_m T. \qquad \square$$

REMARK 3.31. For simplicity, we have assumed that the continuous family $T : X \to \mathcal{F}$ of Fredholm operators is such that T is defined on the same Hilbert space for all $x \in X$. In applications, we actually often have a different situation. For example, in analysis the arising Hilbert spaces are function spaces with values in a vector space V_x which can vary with the parameter $x \in X$. Then T_x is a Fredholm operator on the Hilbert space $H \otimes V_x$; i.e., the Hilbert space on which the continuous family of Fredholm operators operates is not constant. We will encounter such examples mainly in Chapter 10, but also in Part 3, when we interpret homomorphisms of K-theory, in particular products of algebraic topology, with tools of analysis. What can be done in such cases? In general, using the Theorem of Kuiper, one can show that every Hilbert space bundle is trivial (i.e., isomorphic to a product bundle $X \times H$, where H is a single Hilbert space). This follows from [**Ste**, p.54f] and the contractibility of the X *structure group* $\mathcal{B}^\times$, and more simply from $[A, \mathcal{B}^\times] = 1$ for all compact A (Theorem 3.22, p.77) following [**Ste**, p.148f] (e.g., in the case that X is a compact, triangulable manifold). For the construction of index bundles, it suffices to have only a local product structure which is already part of the definition of a Hilbert bundle and this always holds (by the *naturality* of definitions) in our applications.

REMARK 3.32. In the proof of Theorem 3.30, we have strongly used the fact that H is a Hilbert space. We treated the general case with variable $\operatorname{Ker} T_x$ by reducing it, through composition with an orthogonal projection, to the simple special case of constant kernel dimension. Instead, we can study the family T on the quotient spaces $H/\operatorname{Ker} T_x$ which produces the constant kernel dimension 0, or we can make the operators T_x surjective by extending their domain of definition (to $H \oplus \operatorname{Ker}(T_x^*)$ say) which produces constant cokernel dimension 0. Details for these two alternatives, which are already indicated in the usual proofs of the homotopy invariance of the index (e.g., see [**Jö**, 1970/1982, 5.4] and our comments after Theorem 3.11), can be found for the first alternative in [**Ati67a**, p.155-158] and for the second alternative, for instance if H is a Fréchet space and T is an elliptic operator (see below) in [**AS71a**, p.122-127].

[Warning: In connection with *Wiener-Hopf operators*, we will later meet families of Fredholm operators $T : X \to \mathcal{F}$ for which it is quite possible that $\operatorname{index} T = 0$ for all $x \in X$, while the global index bundle $\operatorname{index} T \in K(X)$ does not vanish. The index bundle is simply a much sharper invariant than just an integer. Thus one must be very careful if one wishes to infer properties of the index bundle from those of the index. The next exercise may be comforting.]

EXERCISE 3.33. Show that for each continuous Fredholm family $T : X \to \mathcal{F}$, we have $\operatorname{index} T^* = -\operatorname{index} T$, where $T^* : X \to \mathcal{F}$ is the adjoint family $(T^*)_x := (T_x)^*$ for $x \in X$.

THEOREM 3.34. *The construction of the index bundle in* Theorem 3.30 *depends only on the homotopy class of the given family of Fredholm operators, and for X compact yields a homomorphism of semigroups*

$$\operatorname{index} : [X, \mathcal{F}] \longrightarrow K(X).$$

PROOF. **1**. We show first the *homotopy-invariance* of the index bundle. Let $T : X \times I \to \mathcal{F}$ be a homotopy between the families of Fredholm operators $T_0 := Ti_0$ and $T_1 := Ti_1$ parametrized by X, where

$$i_t : X \longrightarrow X \times \{t\} \hookrightarrow X \times I, \ t \in I$$

are the mutually homotopic natural inclusions. The functoriality of index bundles (Exercise 3.37) then yields

$$\operatorname{index} T_0 = i_0^*(\operatorname{index} T) = i_1^*(\operatorname{index} T) = \operatorname{index} T_1,$$

where the middle equality follows from the homotopy invariance of $K(X)$; namely, $f^* = g^*$ for $f \sim g : Y \to X$ (Theorem B.5, p. 680, Appendix).

2. For the *homomorphism property* of the index, we point out first that for two Fredholm families $S, T : X \to \mathcal{F}$, we have a well defined product family $TS : X \to \mathcal{F}$ given by composition in $\mathcal{F}$ (i.e., $(TS)_x := T_x S_x$), and that $[X, \mathcal{F}]$ is then a semigroup. On the other hand, $K(X)$ has an addition defined via the direct sum of vector bundles which makes it a group (Section 11.3). The index is a semigroup homomorphism now, since

$$\begin{aligned} \operatorname{index} TS &= \operatorname{index} TS + \operatorname{index} \mathrm{Id} = \operatorname{index}(TS \oplus \mathrm{Id}) \\ &= \operatorname{index}(S \oplus T) = \operatorname{index} S + \operatorname{index} T. \end{aligned}$$

In the second and fourth equalities, we have used the fact that (by the construction in Theorem 3.30) the index bundle, for a family of Fredholm operators on a product space $H \times H$ which can be written in the form

$$S \oplus T := \begin{bmatrix} S & 0 \\ 0 & T \end{bmatrix},$$

coincides with the direct sum of the index bundles of the two diagonal elements. Moreover, for the third equality, we have applied the usual trick of homotopy theory, where $TS \oplus \mathrm{Id}$ can be deformed into $S \oplus T$ in $\mathcal{F}(H \times H)$. This is clear by the definition of this homotopy in Exercise 3.18 (p. 73). Note that we do not need $T, S \in \mathcal{B}^\times(H)$ since we are deforming in $\mathcal{F}(H \times H)$, as opposed to $\mathcal{B}^\times(H \times H)$. □

REMARK 3.35. Note that in the proof of homotopy invariance, we did not need to again investigate the topology of $\mathcal{F}$, but rather everything followed from the entirely different aspect of *homotopy invariance* of vector bundles. Why did it take more work to prove the homotopy-invariance of the index of a single operator in the proof of Theorem 3.11, p.68, while in the proof of Theorem 3.34 we did not use this result, even though we can deduce it for $X = \{point\}$. The solution of this paradox lies in the fact that the actual generalization of the homotopy invariance of the index is already contained in the construction of index bundles in Theorem 3.30.

EXERCISE 3.36. Show that the index bundle of a continuous family of self-adjoint Fredholm operators, $T : X \to \mathcal{F}$ with $T_x = T_x^*$ for all $x \in X$ is zero; i.e., $\operatorname{index} T = 0$.
[Warning: One cannot deduce this from Exercise 3.33, since it is possible that $K(X)$ has a finite cyclic (*torsion-*) factor; e.g., possibly $a + a = 0$, but $a \neq 0$.]
[Hint: Consider the homotopy $tT + i(1-t)\,\mathrm{Id}$. Show that a self-adjoint operator A has a real spectrum; i.e.,

$$A - z\,\mathrm{Id} \in \mathcal{B}^\times \text{ for } z \in \mathbb{C} - \mathbb{R} :$$

Step 1. $u \in \operatorname{Ker}(A - z\,\mathrm{Id})$ implies $zu = Au = A^*u = \bar{z}u$, and so $u = 0$ since $z \neq \bar{z}$.
Step 2. If v is in the orthogonal complement of $\operatorname{Im}(A - z\,\mathrm{Id})$, then $\langle Au - zu, v\rangle = 0$ and so

$$\langle u, Av\rangle = \langle Au, v\rangle = \langle zu, v\rangle = \langle u, \bar{z}v\rangle \text{ for all } u \in H.$$

Thus, $Av = \bar{z}v$, and $v = 0$ by step 1. Hence, $\operatorname{Im}(A - z\,\mathrm{Id})$ is dense in H.
Step 3. Let $v \in H$ and let $v_1, v_2, \ldots \in \operatorname{Im}(A - z\,\mathrm{Id})$ be a sequence converging to v. Then show that the sequence of unique preimages $u_1, u_2, \ldots$ is a Cauchy sequence, and set $u := \lim u_i$. Then note that $Au - zu = v$.]

We will now investigate the construction in the Theorem 3.30 more generally, and show that it has all the properties that one might reasonably expect. We begin with the following exercise.

Exercise 3.37. Show that our definition of index bundle is functorial: Let $f : Y \to X$ be a continuous map (X and Y compact) and let $T : X \to \mathcal{F}$ be a continuous family of Fredholm operators. Then (see Exercise B.1d, p. 676, of the Appendix), we have $\operatorname{index} Tf = f^*(\operatorname{index} T)$.
[Hint: If $e_1, e_2, \ldots$ is an orthonormal basis for H and the projection P_n is chosen so that $\dim \operatorname{Ker} P_n T_x$ is independent of x, then $\dim \operatorname{Ker} P_n T_{f(y)}$ is also independent of $y \in Y$. Thus, one does not need to choose another projection P_n to exhibit the index bundle for the family $Tf : Y \to \mathcal{F}$.]

Exercise 3.38. The proof of Theorem 3.34 yields (for $X = \textit{point}$) a further proof of the composition rule $\operatorname{index} TS = \operatorname{index} S + \operatorname{index} T$ for Fredholm operators. We have already seen three proofs, namely Exercise 1.10 (p.9), Exercise 2.3 (p.14), and Exercise 3.7 (p.65). Do these three proofs carry over without difficulty to the case of families of Fredholm operators? Does one need a homotopy argument each time? What is the real relationship between the four proofs?

Exercise 3.39. Show that the set

$$\mathcal{F}_0 := \{T \in \mathcal{F} : \operatorname{index} T = 0\}$$

is pathwise-connected.
[Hint: For $T \in \mathcal{F}_0$ choose an isomorphism

$$\phi : \operatorname{Ker} T \longrightarrow (\operatorname{Im} T)^\perp$$

(vector spaces of the same finite dimension!) and set

$$\Phi := \begin{cases} \phi, & \text{on } \operatorname{Ker} T, \\ 0, & \text{on } (\operatorname{Ker} T)^\perp . \end{cases}$$

By construction, we have $T + \Phi \in \mathcal{B}^\times$ and $T + t\Phi \in \mathcal{F}$ for $t \in I$ (Why? What kind of operator is Φ?). Now apply Theorem 3.21.]

8. The Theorem of Atiyah-Jänich

The sets $\mathcal{F}_i := \{T \in \mathcal{F} : \operatorname{index} T = i\}$ are bijectively (modulo compact operators) mapped onto each other in a continuous fashion by shift operators (see Exercise 1.3, p.4). Thus, Exercise 3.39 with the homotopy-invariance of the index already proved in Theorem 3.11 (p. 68) gives us a bijection from the pathwise-connected components of $\mathcal{F}$ to $\mathbb{Z}$. Naturally, this result still does not say much about the topology of $\mathcal{F}$, and we do not want to carry it out in detail. Much more informative is the following theorem which gives the result,

$$\text{index} : [\{point\}, \mathcal{F}] \longrightarrow \mathbb{Z} \text{ bijective,}$$

as the special case $X = \{point\}$.

THEOREM 3.40 (M. F. Atiyah, K. Jänich 1964). *We have an isomorphism*

$$\text{index} : [X, \mathcal{F}] \longrightarrow K(X).$$

PROOF. We show that the natural sequence of semigroups

$$[X, \mathcal{B}^\times] \longrightarrow [X, \mathcal{F}] \overset{\text{index}}{\longrightarrow} K(X) \longrightarrow 0$$

is exact. From this result of M. Atiyah and K. Jänich along with the theorem of N. Kuiper (Theorem 3.22, p.77), the present theorem follows. (For the concept of *exactness*, see the material before Theorem 1.9, p. 6).

Step 1. The index bundle of a continuous family $T : X \to \mathcal{B}^\times$ is trivially zero.

Step 2. Now let $T : X \to \mathcal{F}$ be a Fredholm family with vanishing index bundle. In Exercise 3.39, we have stated what one must do in the case where $X = \{point\}$; i.e., T is a single Fredholm operator. Namely, one chooses an isomorphism

$$\phi : \operatorname{Ker} T \longrightarrow (\operatorname{Im} T)^\perp$$

and then

$$\Phi := \begin{cases} \phi, & \text{on } \operatorname{Ker} T, \\ 0, & \text{on } (\operatorname{Ker} T)^\perp \end{cases}$$

is an operator of finite rank with $T + t\Phi$ the desired homotopy in $\mathcal{F}$ to $T + \Phi \in \mathcal{B}$. To generalize this to $X \neq \{point\}$, we now must deal with two difficulties:

1. $\bigcup_{x \in X} \operatorname{Ker} T_x$ is not always a vector bundle.
2. In general, there is no canonical choice of $\phi_x : \operatorname{Ker} T_x \longrightarrow (\operatorname{Im} T)^\perp$ which depends continuously on x.

The first problem, we can quickly solve: We choose again an orthogonal projection $P_n : H \to H_n$, such that $\dim \operatorname{Ker} P_n T_x$ is constant and

$$\operatorname{index} T = [\operatorname{Ker} P_n T_x] - [X \times (H_n)^\perp] \in K(X).$$

Then "index $T = 0$" means

$$[\operatorname{Ker} P_n T] = [X \times (H_n)^\perp] \text{ in } K(X),$$

which in turn means (see Section 11.3) that, for N sufficiently large, there is an isomorphism of vector bundles

$$\phi : (\operatorname{Ker} P_n T) \oplus (X \times \mathbb{C}^N) \cong (X \times (H_n)^\perp) \oplus (X \times \mathbb{C}^N).$$

This means that by *augmenting* with the trivial bundle $X \times \mathbb{C}^N$, we can also cure the second difficulty in principle. Actually we can avoid this K-theoretical argument. Indeed, we do not need to *augment*, if we take n to be large enough: As shown in the proof of Theorem 3.30 (p.84), we have

$$\operatorname{Ker} P_{n+1} T \cong \operatorname{Ker} P_n T \oplus (X \times \mathbb{C}).$$

Thus, for $m := n + N$, the map ϕ can be regarded as an isomorphism

$$\operatorname{Ker} P_m T \longrightarrow X \times (H_m)^\perp.$$

In this way, the construction of ϕ in Exercise 3.39 carries over pointwise, whence we obtain a homotopy of Fredholm families between

$$P_mT : X \longrightarrow \mathcal{F}(H) \text{ and } P_mT + \Phi : X \longrightarrow \mathcal{B}^\times(H).$$

Since the index of the orthogonal projection $P_m : H \to H_m$ vanishes (and hence coincides with the index of Id), we can connect the *constants* P_m with Id in $\mathcal{F}(H)$ (Exercise 3.39), and hence also connect P_mT with T in a further homotopy of maps $X \to \mathcal{F}(H)$. Each Fredholm family with vanishing index bundle is therefore homotopic to a continuous family of invertible operators, and, in conjunction with step 1, exactness in the middle of the short sequence is now proved.

Step 3: We must now show the surjectivity of index : $[X, \mathcal{F}] \to K(X)$. By the construction in Exercise B.10 (p. 683) of the Appendix (see also Section 11.3), every element of $K(X)$ can be written in the form $[E] - [X \times \mathbb{C}^k]$, where E is a vector bundle over X and $k \in \mathbb{Z}_+$. Hence we will be finished with the proof if, for each vector bundle E, we can find a continuous family S of (surjective) Fredholm operators on a suitable Hilbert space such that index $S = [E]$. Then, the homomorphism of the index bundle construction (Theorem 3.30) yields

$$\operatorname{index}\left(\operatorname{shift}^+\right)^k S = [E] - [X \times \mathbb{C}^k],$$

where shift (as in Exercise 1.3, p. 1.3) is the displacement to the right relative to an orthonormal basis of the Hilbert space. Note that, since $\operatorname{index} \operatorname{shift}^+ = -1$, the constant family $x \to \operatorname{shift}^+$ gives the index bundle $-[X \times \mathbb{C}]$.

Thus, let E be a vector bundle over X. If X consists of a single point, then we just complete an orthonormal basis of E (regarded as a subspace of an infinite-dimensional separable Hilbert space) to an orthonormal basis of the whole Hilbert space and set $S := (\operatorname{shift}^-)^{\dim E}$. Now, we consider the general case. By Exercise 11.8 (p. 275) of the Appendix, there is a vector bundle F over X and a finite-dimensional (complex) vector space V such that $E \oplus F \cong X \times V$. Let $\pi : V \to E$ denote the projection.

Let H be an arbitrary Hilbert space. We will see that it is simplest to consider the desired operators S_x, $x \in X$, to be defined on the space $\operatorname{Hom}(V, H)$ of linear transformations from V to H. We choose for the vector space V (which we can identify with $\mathbb{C}^N$, $N = \dim V$, via a basis) a Hermitian scalar product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$. For every pair $(f, u) \in V \times H$, we have an element of $\operatorname{Hom}(V, H)$ defined by

$$v \longrightarrow \langle f, v \rangle\, u,\ v \in V,$$

which we will denote by $f \otimes u$; recall the isomorphism $\operatorname{Hom}(V, H) = V' \otimes H$ of linear algebra, where $V'(\cong V)$ is the vector space of linear maps from V to $\mathbb{C}$. Then, we have

$$\operatorname{Hom}(V, H) = \left\{ \sum\nolimits_{i=1}^m f_i \otimes u_i : m \in \mathbb{N},\ f_i \in V,\ u_i \in H \right\},$$

where

$$zf \otimes u = f \otimes zu, z \in \mathbb{C} \text{ and}$$
$$(f + f') \otimes (u + u') = f \otimes u + f \otimes u' + f' \otimes u + f' \otimes u'.$$

We define a scalar product

$$\left\langle \sum\nolimits_{i=1}^m f_i \otimes u_i, \sum\nolimits_{j=1}^m f_j' \otimes u_j' \right\rangle := \sum\nolimits_{i,j=1}^m \left\langle f_i, f_j' \right\rangle \left\langle u_i, u_j' \right\rangle,$$

which makes $\mathrm{Hom}(V,H)$ a Hilbert space (since V is finite-dimensional, nothing can go wrong; addition and scalar multiplication by complex numbers are defined naturally). One easily checks that if $f_1,...,f_N$ is an orthonormal basis for V and $e_0,e_1,e_2,\ldots$ is an orthonormal basis for H, then

$$\{f_j \otimes e_i\}_{j=1,\ldots,N,\ i=0,1,2,\ldots}$$

is a orthonormal basis for $\mathrm{Hom}(V,H)$.

Each bounded linear operator T on H and endomorphism τ of V clearly define a bounded linear operator $\tau \otimes T$ on $\mathrm{Hom}(V,H)$ by means of

$$(\tau \otimes T)(f \otimes u) := \tau(f) \otimes T(u)\,.$$

For a fixed chosen basis $e_0,e_1,e_2,\ldots$ of H, we set

$$S_x := (\pi_x \otimes \mathrm{shift}^-) + (\mathrm{Id}_V - \pi_x) \otimes \mathrm{Id}_H\,.$$

Then,

$$S_x(f \otimes e_i) = \begin{cases} \pi_x(f) \otimes e_{i-1} + (f - \pi_x(f)) \otimes e_i, & \text{for } i \geq 1,\\ (f - \pi_x(f)) \otimes e_0, & \text{for } i = 0, \end{cases}$$

and in particular for $f \in E_x$, we have $S_x(f \otimes e_0) = 0$. Hence, we have $\mathrm{Im}\, S_x = \mathrm{Hom}(V,H)$ and $\mathrm{Ker}\, S_x = \{f \otimes e_0 : f \in E_x\}$. Thus, $\mathrm{Ker}\, S_x$ is isomorphic to E_x in a natural way, and $\mathrm{index}\, S = [\mathrm{Ker}\, S] - 0 = [E]$. □

The preceding theorems are significant on various levels: In the following chapters, in dealing topologically with boundary value problems as well as in proving analytically the Periodicity Theorem of the topology of linear groups, we will repeatedly use Theorems 3.30 and 3.34, i.e., the construction of the index bundle and its elementary properties, but we will not use explicitly Theorem 3.40, our actual *main theorem*. Nevertheless, the last theorem has fundamental significance for our topic, as it provides the reasons for the theoretical relevance of the notion *index bundle* and explains why this concept proved suitable to express deep relations in analysis as well as topology.

9. Determinant Line Bundles

A primary motivation for the study of determinant line bundles originated from the desire of quantum physicists to produce a gauge-invariant volume element for the purpose of computing (via functional integration) Greens functions for Dirac operators coupled to gauge potentials. An obstruction to doing this is the nontriviality of the so-called determinant bundle for a certain family of Fredholm operators, namely the family of Dirac operators parametrized by the quotient space of gauge potentials modulo gauge transformations. This obstruction signals the presence of so-called anomalies that arise when physicists attempt to quantize the classical field theory.

While it would be premature to go into the mathematically murky details of this now, in this section we develop the notion of the determinant line bundle of a continuous family of Fredholm operators. Moreover, in the case of a family with index 0, we assist the reading in showing (see Exercise 3.51 and Corollary 3.52) that it is the pull-back of a *universal* determinant line bundle that we construct over $\mathcal{F}_0$, the space of Fredholm operators of index 0. The simplest way of describing this line bundle (often referred to as the *Quillen determinant line bundle*, which arose in [**Q85**]) is to declare the fiber above the point $T \in \mathcal{F}_0$ to be

$$\Lambda^{d_T}(\ker T)^* \otimes \Lambda^{d_T}(\ker T^*), \quad \text{where } d_T := \dim\ker T = \dim\ker T^*.$$

However, the fact that these fibers may be pieced together to form a genuine line bundle is not trivial since d_T varies with T, and d_T is an unbounded function of $T \in \mathcal{F}_0$. In doing this, we adopt a novel approach due primarily to Graeme Segal (in [**Seg90**]). Various ways of describing this universal bundle are summarized in Theorem 3.53.

The Quillen Determinant Line Bundle. For compact X and a continuous family $T : X \to \mathcal{F}$ of Fredholm operators in a fixed Hilbert space H, we have seen that there is an index bundle

$$\operatorname{index} T \in K(X)$$

assigned in a canonical way. While the determinant of an operator on a Hilbert space H exists only in restrictive circumstances, we now show that there is a well-defined complex line bundle $\det T \to X$, known as the *determinant line bundle* of T. For any $\alpha \in K(X)$, we will first define an isomorphism class $\det \alpha \in K(X)$ which is represented by a line bundle. For a continuous family $T : X \to \mathcal{F}$, we then can (and do) simply define $\det T$ to be $\det(\operatorname{index} T)$, where $\operatorname{index} T \in K(X)$ is given in the Atiyah–Jänich Theorem 3.40. We know that $\alpha = [E] - [F] \in K(X)$ for $E, F \in \operatorname{Vect}(X)$. For a finite-dimensional vector bundle $V \to X$, it is convenient to use the notation $\Lambda^{\max}(V) = \Lambda^{\dim V}(V)$. We claim that $\left[\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F)\right] \in K(X)$ only depends on α, in the strong sense that if $\alpha = [E] - [F] = [E'] - [F']$, then

$$\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F) \quad \cong \quad \Lambda^{\max}(E')^* \otimes \Lambda^{\max}(F') \tag{3.4}$$

(not just $\left[\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F)\right] = \left[\Lambda^{\max}(E')^* \otimes \Lambda^{\max}(F')\right]$ in $K(X)$). Indeed, using the fact that $\Lambda^{\max}(V \oplus W) = \Lambda^{\max}(V) \otimes \Lambda^{\max}(W)$, we have

$$\begin{aligned}
&[E] - [F] \quad = \quad [E'] - [F'] \qquad \Longleftrightarrow \qquad E \oplus F' \quad \cong \quad E' \oplus F \\
&\Longrightarrow \Lambda^{\max}(E \oplus F') \quad \cong \quad \Lambda^{\max}(E' \oplus F) \\
&\Longrightarrow \Lambda^{\max}(E) \otimes \Lambda^{\max}(F') \quad \cong \quad \Lambda^{\max}(E') \otimes \Lambda^{\max}(F).
\end{aligned}$$

Tensoring both sides with $\Lambda^{\max}(E')^* \otimes \Lambda^{\max}(E)^*$ and noting that for a line bundle L, $L \otimes L^*$ is isomorphic to the trivial bundle $X \times \mathbb{C}$, we then obtain (3.4). We can now make the following

Definition 3.41. For X compact and $\alpha \in K(X)$, let E and F be vector bundles over X with $\alpha = [E] - [F]$. We define the **determinant** of α by

$$\det \alpha = \left[\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F)\right] \in K(X).$$

By (3.4) the bundle $\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F)$ is well defined (independent of the choice of E and F). By an abuse of notation, we also denote this isomorphism class by $\det \alpha$. For a continuous family $T : X \to \mathcal{F}$,

$$\det T := \det(\operatorname{index} T).$$

Note that while $\det T$ has been defined, this does not directly imply that

$$\bigcup_{x \in X} \Lambda^{\max}(\operatorname{Ker}(T_x))^* \otimes \Lambda^{\max}(\operatorname{Ker}(T_x^*)) \tag{3.5}$$

can be given the structure of a (global) vector bundle over X. Usually, doubts about this are glibly deflected by saying: while $\operatorname{Ker}(T)$ and $\operatorname{Ker}(T^*)$ are not defined globally in general, $\dim \operatorname{Ker}(T_x)$ and $\dim \operatorname{Ker}(T_x^*)$ jump by the same amount if x

varies. It is useful (and comforting) to show that for *some* genuine vector bundles E and F over X with $\operatorname{index} T = [E] - [F]$, the fiber $\Lambda^{\max}(E_x)^* \otimes \Lambda^{\max}(F_x)$ of the manifestly well defined bundle $\Lambda^{\max}(E)^* \otimes \Lambda^{\max}(F) \to X$ is isomorphic to $\Lambda^{\max}(\operatorname{Ker}(T_x))^* \otimes \Lambda^{\max}(\operatorname{Ker}(T_x^*))$ in a *natural* fashion. Indeed, as in the proof of Theorem 3.30 (p. 84), let $P_n : H \to H_n = \{e_0, \ldots, e_n\}^\perp$ be an orthogonal projection so that $P_n T_x H = H_n$ for all $x \in X$. We then have bundles $E = \operatorname{Ker} P_n T$ and $F = \operatorname{Ker}(P_n T)^* = X \times H_n^\perp$, and a *natural* isomorphism

$$\Lambda^{\max}(\operatorname{Ker} P_n T_x)^* \otimes \Lambda^{\max}\left(H_n^\perp\right) \quad \cong \quad \Lambda^{\max}(\operatorname{Ker}(T_x))^* \otimes \Lambda^{\max}(\operatorname{Ker}(T_x^*))$$

is supplied by taking G to be T_x in the following

PROPOSITION 3.42. *Suppose that $G \in \mathcal{F}$ and $P_n G(H) = H_n$. Then there is an isomorphism (depending only on the choice P_n)*

$$\Psi_{G,n} : \Lambda^{\max}(\operatorname{Ker} P_n G)^* \otimes \Lambda^{\max}\left(H_n^\perp\right) \quad \cong \quad \Lambda^{\max}(\operatorname{Ker} G)^* \otimes \Lambda^{\max}(\operatorname{Ker} G^*) . \tag{3.6}$$

PROOF. First note that we have an isomorphism

$$\widetilde{G} := G|_{\operatorname{Ker}(G)^\perp} : (\operatorname{Ker} G)^\perp \longrightarrow G(H),$$

and hence

$$\begin{aligned} \operatorname{Ker} P_n G = G^{-1} H_n^\perp &= \operatorname{Ker}(G) \oplus \left(\operatorname{Ker}(G)^\perp \cap G^{-1}\left(H_n^\perp\right)\right) \\ &= \operatorname{Ker}(G) \oplus \widetilde{G}^{-1}\left(H_n^\perp \cap G(H)\right). \end{aligned} \tag{3.7}$$

The orthogonal projection $Q_n : H \to H_n^\perp$ is $\operatorname{Id}_H - P_n$. We are given that $P_n G(H) = H_n$. This implies that $Q_n|_{G(H)^\perp} : G(H)^\perp \to H_n^\perp$ is injective, because

$$\begin{aligned} & v \in G(H)^\perp,\ Q_n(v) = 0 \Rightarrow v \in H_n \cap G(H)^\perp \\ & \Longrightarrow v = P_n(w) \text{ for some } w \in G(H), \text{ say } w = G(u) \\ & \Longrightarrow G(u) = w = P_n(w) + Q_n(w) = v + Q_n(w) \\ & \Longrightarrow v = G(u) - Q_n(w) \\ & \Longrightarrow \langle v, v\rangle = \langle v, G(u) - Q_n(w)\rangle = \langle v, G(u)\rangle - \langle v, Q_n(w)\rangle = 0, \end{aligned}$$

since $v \in G(H)^\perp$ and $\langle v, Q_n(w)\rangle = \langle P_n(w), Q_n(w)\rangle = 0$. We claim

$$H_n^\perp = \left(H_n^\perp \cap G(H)\right) \oplus Q_n\left(G(H)^\perp\right). \tag{3.8}$$

First note that $\left(H_n^\perp \cap G(H)\right) \cap Q_n\left(G(H)^\perp\right) = \{0\}$, since

$$\begin{aligned} & u \in \left(H_n^\perp \cap G(H)\right) \cap Q_n\left(G(H)^\perp\right) \\ & \Longrightarrow u = G(w) \in H_n^\perp \text{ for some } w \in H, \text{ and } u = Q_n(v) \text{ for } v \in G(H)^\perp \\ & \Longrightarrow \langle u, u\rangle = \langle Q_n(v), G(w)\rangle = \langle v - P_n(v), G(w)\rangle \\ & = \langle v, G(w)\rangle - \langle P_n(v), G(w)\rangle = 0 - \langle P_n(v), G(w)\rangle = 0, \end{aligned}$$

since $v \in G(H)^\perp$ and $G(w) \in H_n^\perp$. The same proof yields the general fact that for two subspaces V and W of an inner product space, the orthogonal projection of $V^\perp$ onto W is orthogonal to $V \cap W$. To obtain (3.8), it now suffices to show that

$$\dim\left(H_n^\perp \cap G(H)\right) \geq \dim H_n^\perp - \dim Q_n\left(G(H)^\perp\right).$$

For this, note that

$$\begin{aligned}&\dim\left(H_n^{\perp}\cap G(H)\right)\geq\dim\left(H_n^{\perp}\right)-\operatorname{codim}(G(H))\\&=\dim\left(H_n^{\perp}\right)-\dim G(H)^{\perp}=\dim\left(H_n^{\perp}\right)-\dim Q_n\left(G(H)^{\perp}\right),\end{aligned}$$

since we have shown that $Q_n|_{G(H)^{\perp}}$ is injective. This also follows from

$$\begin{aligned}&\operatorname{index} G=\operatorname{index} G+\operatorname{index} P_n=\operatorname{index} P_nG\\&\Longrightarrow\dim\operatorname{Ker} G-\dim G(H)^{\perp}=\dim\operatorname{Ker} P_nG-\dim H_n^{\perp}\\&\Longrightarrow\dim G(H)^{\perp}=\dim\left(H_n^{\perp}\cap G(H)\right)-\dim H_n^{\perp}.\end{aligned}$$

Let

$$\widetilde{G}_n:=\widetilde{G}|_{\widetilde{G}^{-1}(H_n^{\perp}\cap G(H))}:\widetilde{G}^{-1}\left(H_n^{\perp}\cap G(H)\right)\cong H_n^{\perp}\cap G(H),\tag{3.9}$$

and

$$q_{G,n}:=Q_n|_{G(H)^{\perp}}:G(H)^{\perp}\cong Q_n\left(G(H)^{\perp}\right).$$

By (3.8) and (3.7), we have the isomorphisms

$$\begin{aligned}&\Lambda^{\max}(\widetilde{G}_n^{-1})^*:\Lambda^{\max}(\widetilde{G}^{-1}\left(H_n^{\perp}\cap G(H)\right))^*\cong\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)^*,\\&\Lambda^{\max}(q_{G,n}^{-1})^*:\Lambda^{\max}(Q_n\left(G(H)^{\perp}\right))^*\cong\Lambda^{\max}\left(G(H)^{\perp}\right)^*,\ \text{and}\\&\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)^*\otimes\Lambda^{\max}\left(\left(H_n^{\perp}\cap G(H)\right)\right)\cong\mathbb{C}.\end{aligned}$$

Then we obtain (3.6), as follows:

$$\begin{aligned}&\Lambda^{\max}\left(\operatorname{Ker} P_nG\right)^*\otimes\Lambda^{\max}\left(H_n^{\perp}\right)\\&=\Lambda^{\max}(\operatorname{Ker}(G)\oplus\widetilde{G}^{-1}\left(H_n^{\perp}\cap G(H)\right))^*\\&\otimes\Lambda^{\max}\left(\left(H_n^{\perp}\cap G(H)\right)\oplus Q_n\left(G(H)^{\perp}\right)\right)\\&=\Lambda^{\max}\left(\operatorname{Ker} G\right)^*\otimes\Lambda^{\max}(\widetilde{G}^{-1}\left(H_n^{\perp}\cap G(H)\right))^*\\&\otimes\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)\otimes\Lambda^{\max}\left(Q_n\left(G(H)^{\perp}\right)\right).\end{aligned}$$

Via $\mathrm{Id}\otimes\Lambda^{\max}(\widetilde{G}_n^{-1})^*\otimes\mathrm{Id}\otimes\Lambda^{\max}(q_{G,n}^{-1})$, this last tensor product is

$$\begin{aligned}&\cong\Lambda^{\max}\left(\operatorname{Ker} G\right)^*\otimes\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)^*\\&\otimes\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)\otimes\Lambda^{\max}\left(G(H)^{\perp}\right)\\&\cong\Lambda^{\max}\left(\operatorname{Ker} G\right)^*\otimes\Lambda^{\max}\left(G(H)^{\perp}\right).\end{aligned}$$

In other words, for

$$\begin{aligned}&\alpha\in\Lambda^{\max}\left(\operatorname{Ker} G\right)^*,\ \beta_n\in\Lambda^{\max}\left(\widetilde{G}^{-1}\left(H_n^{\perp}\cap G(H)\right)\right)^*\\&b_n\in\Lambda^{\max}\left(H_n^{\perp}\cap G(H)\right)\ \text{and}\ a_n\in\Lambda^{\max}\left(Q_n\left(G(H)^{\perp}\right)\right),\end{aligned}$$

we have $(\alpha\otimes\beta_n)\otimes(b_n\otimes a_n)\in\Lambda^{\max}\left(\operatorname{Ker} P_nG\right)^*\otimes\Lambda^{\max}\left(H_n^{\perp}\right)$, and we define

$$\Psi_{G,n}\left((\alpha\otimes\beta_n)\otimes(b_n\otimes a_n)\right):=\left\langle\Lambda^{\max}(\widetilde{G}_n^{-1})^*\left(\beta_n\right),b_n\right\rangle\alpha\otimes\Lambda^{\max}(q_{G,n}^{-1})\left(a_n\right).$$

Note that

$$\Psi_{G,n}^{-1}(\alpha \otimes a) = (\alpha \otimes \beta_n) \otimes (b_n \otimes \Lambda^{\max}(q_{G,n})a),$$

where β_n and b_n are chosen so that $\left\langle \Lambda^{\max}(\widetilde{G}_n^{-1})^*(\beta_n), b_n \right\rangle = 1$; i.e., b_n is dual to $\Lambda^{\max}(\widetilde{G}_n^{-1})^*(\beta_n)$, or equivalently, β_n is the dual of $\Lambda^{\max}(\widetilde{G}_n^{-1})(b_n)$. □

Thus, the set (3.5) can be given the structure of a genuine line bundle, namely that which is induced by the bijections

$$\Lambda^{\max}(\operatorname{Ker}(T_x))^* \otimes \Lambda^{\max}(\operatorname{Ker}(T_x^*)) \leftrightarrow \Lambda^{\max}(\operatorname{Ker} P_n T_x)^* \otimes \Lambda^{\max}\left(H_n^{\perp}\right),$$

and this line bundle structure is unique up to isomorphism. The set (3.5) then serves as a standard representative of the isomorphism class $\det T$.

What we have done so far is sufficient for many purposes, but we will go on to construct a restricted version, the so-called **Quillen determinant line bundle** $q: \mathcal{Q} \to \mathcal{F}$ and show that $\det T \to X$ is isomorphic to the pull-back via T of $\mathcal{Q}$, at least in the case $\operatorname{index} T = 0$. The fiber of $\mathcal{Q}$ above $F \in \mathcal{F}$, is simply

$$q^{-1}(F) = \Lambda^{\dim \operatorname{Ker} F}(\operatorname{Ker} F)^* \otimes \Lambda^{\dim \operatorname{Ker} F^*}(\operatorname{Ker} F^*),$$

which is clearly a complex line. However, as with the set (3.5), it is not immediately clear that there are suitable local trivalizations for

$$\mathcal{Q} := \bigcup_{F \in \mathcal{F}} q^{-1}(F) \tag{3.10}$$

with continuous transition functions, because $\operatorname{Ker} F$ (or $\operatorname{Ker} F^*$) is not the fiber of a vector bundle over $\mathcal{F}$ about points where F (or F^*) is not surjective. It is true that for any $F \in \mathcal{F}$, there is some n_F such that $P_n F(H) = H_n^{\perp}$ for some suitable neighborhood, say U_F, of F in $\mathcal{F}$. Moreover, one can construct a trivialization for $q|_{U_F} : q^{-1}(U_F) \to U_F$. However, unlike the case of a family $T : X \to \mathcal{F}$ with X *compact* where we had fixed n for which $P_n T_x(H) = H_n^{\perp}$ for *all* $x \in X$, note now that n_F is an *unbounded* function of $F \in \mathcal{F}$ (noncompact). At the very least, this causes difficulties in defining the transition functions and exhibiting their continuity. Instead of attempting this, we opt for an interesting, instructive alternative construction, using transition functions that are determinants of the form $\det(\operatorname{Id} + A)$ where A is a trace class operator, defined below. This idea is based on informal notes of Graeme Segal (see [**Seg90**]), with extensions elaborated upon by Kenro Furutani in [**Fur**], to whom we are indebted. One drawback is that this construction is only for $\mathcal{Q}|_{\mathcal{F}_0}$, where $\mathcal{F}_0 := \{A \in \mathcal{F} : \operatorname{index} A = 0\}$, and we have yet to find a graceful way to similarly construct $\mathcal{Q}$ over the components of $\mathcal{F}$ with nonzero index. This may not be of crucial importance, since the usual convention is that if $\operatorname{index} T \neq 0$, then $\det T = 0$ if $\det T$ is defined at all (e.g., for $T : \mathbb{R}^n \to \mathbb{R}^m$, we have $\operatorname{index} T = n - m$ and $\det T$ is undefined if $n \neq m$). Of course, we always can identify the connected components of $\mathcal{F}$ (distinguished by the index), e.g., by shift operators after fixing a basis for the underlying Hilbert space H. The bundle $\mathcal{Q}$ then constructed via pull-back would, however, depend on the choice of the identifications. Another potential difficulty in the Segal approach is that one needs to know the functional analysis of trace class and Hilbert-Schmidt operators (which is covered in Section 2.7), and determinants which we consider next.

Determinants. For a complex vector space V with $d = \dim V$ finite, let $\Lambda^k(V)$ denote the k-th exterior product of V ($k = 0, \ldots, d$), where $\Lambda^0(V) = \mathbb{C}$. If $T \in \operatorname{End}(V)$ (i.e., $T : V \to V$ is linear), then T induces $\Lambda^k T \in \operatorname{End}\left(\Lambda^k(V)\right)$, determined by

$$\left(\Lambda^k T\right)(v_1 \wedge \cdots \wedge v_k) = Tv_1 \wedge \cdots \wedge Tv_k \text{ for } k > 0, \tag{3.11}$$

and $\Lambda^0 T = \operatorname{Id}_{\mathbb{C}}$. Since $\dim \Lambda^d(V) = 1$, $\Lambda^d T$ is multiplication by a scalar, and this scalar is $\det T$. Indeed $\det T$ can be defined in this way, and it is straightforward to show that this agrees with the usual definition in terms of the matrix of T relative to a basis (e.g., just take $k = d$ and let $v_1, \cdots, v_d$ be a basis in (3.11)). Note that under the natural isomorphisms

$$\operatorname{End}\left(\Lambda^d(V)\right) \cong \Lambda^d(V)^* \otimes \Lambda^d(V) \cong \mathbb{C},$$

and $\Lambda^d T \in \operatorname{End}\left(\Lambda^d(V)\right)$ corresponds to $\det T$. There are immediate problems with the notion $\det T$ when $\dim V = \infty$, but $\det T$ can be defined when V is a Hilbert space and T is sufficiently close to Id (e.g., $T - \operatorname{Id}$ is trace class), as we will show. We begin with the following proposition whose proof can be shortened somewhat if one presupposes that A can be put in Jordan canonical form.

Proposition 3.43. *If V is a vector space with $\dim V = d < \infty$, we have*

$$\det(\operatorname{Id} + A) = \sum_{k=0}^{d} \operatorname{Tr}\left(\Lambda^k A\right).$$

Proof. We define

$$\epsilon_{i_1 \ldots i_k}^{j_1 \ldots j_k} = \begin{cases} 1, & \text{if } j_1 \ldots j_k \text{ is an even permutation of } i_1 \ldots i_k, \\ -1, & \text{if } j_1 \ldots j_k \text{ is an odd permutation of } i_1 \ldots i_k, \\ 0, & \text{otherwise,} \end{cases}$$

and $\epsilon_{i_1 \ldots i_d} = \epsilon_{i_1 \ldots i_d}^{1 \ldots d}$.

In what follows, we use the multi-index notation

$$(i)_k := (i_1, \ldots, i_k) \text{ and } (i)'_{d-k} = (i_{k+1}, \ldots, i_d),$$

where $i_1, \ldots, i_d$ run independently from 1 to d. We also use the abbreviated notations

$$\wedge_{(i)_k} v := v_{i_1} \wedge \cdots \wedge v_{i_k},\ \wedge_{(i)'_{n-k}} v = v_{i_{k+1}} \wedge \cdots \wedge v_{i_d},\ \text{and}$$

$$\epsilon_{(i)_k}^{(j)_k} = \epsilon_{i_1 \ldots i_k}^{j_1 \ldots j_k} \text{ and } \epsilon_{(i)_d} = \epsilon_{i_1 \ldots i_d} = \epsilon_{i_1 \ldots i_d}^{1 \ldots d}.$$

Assuming that $v_1, \ldots, v_d$ is an orthonormal basis of V, we compute

$$\begin{aligned}
&\left(\Lambda^d (\mathrm{Id} + A)\right)(v_1 \wedge \ldots \wedge v_d) = (v_1 + Av_1) \wedge \ldots \wedge (v_d + Av_d) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \left(\wedge_{(i)_k} (Av)\right) \wedge \left(\wedge_{(i)'_{d-k}} v\right) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \left(\frac{1}{k!} \sum\nolimits_{(j)_k} \left\langle \wedge_{(i)_k} (Av), \wedge_{(j)_k} v \right\rangle \left(\wedge_{(j)_k} v\right)\right) \wedge \left(\wedge_{(i)'_{d-k}} v\right) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \left(\frac{1}{k!} \sum\nolimits_{(j)_k} \left\langle \wedge_{(i)_k} (Av), \wedge_{(j)_k} v \right\rangle\right) \left(\wedge_{(j)_k} v\right) \wedge \left(\wedge_{(i)'_{d-k}} v\right) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \left(\left(\frac{1}{k!} \sum\nolimits_{(j)_k} \left\langle \wedge_{(i)_k} (Av), \wedge_{(j)_k} v \right\rangle\right) \epsilon^{(j)_k}_{(i)_k} \left(\wedge_{(i)_d} v\right)\right) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \left(\left(\frac{1}{k!} \sum\nolimits_{(j)_k} \left\langle \wedge_{(j)_k} (Av), \wedge_{(j)_k} v \right\rangle\right) \left(\wedge_{(i)_d} v\right)\right) \\
&= \frac{1}{d!} \sum\nolimits_{(i)_d} \epsilon_{(i)_d} \sum\nolimits_{k=0}^{d} \mathrm{Tr}\left(\Lambda^k A\right) \left(\wedge_{(i)_d} v\right) \\
&= \left(\sum\nolimits_{k=0}^{d} \mathrm{Tr}\left(\Lambda^k A\right)\right)(v_1 \wedge \ldots \wedge v_d). \qquad \square
\end{aligned}$$

If H is a separable (complex) Hilbert space with $\dim H = \infty$, and A is a suitable operator on H, one is tempted to define

$$\det(\mathrm{Id} + A) = \sum\nolimits_{k=0}^{\infty} \mathrm{Tr}\left(\Lambda^k A\right). \tag{3.12}$$

This is certainly finite if A has finite rank, but the terms on the right side need not exist, even if A is compact. If there is a complete orthonormal system $\{e_0, e_1, \ldots\}$ with $Ae_i = \lambda_i e_i$, $\lambda_i \in \mathbb{C}$, then it is reasonable to define

$$\det(\mathrm{Id} + A) = \prod\nolimits_{i=0}^{\infty} (1 + \lambda_i),$$

provided the infinite product converges (i.e., the partial products $\prod_{i=0}^{n} (1 + \lambda_i)$ converge). It is well known that this is the case if $\sum_{i=0}^{\infty} |\lambda_i| < \infty$; i.e., if A is trace class. However, we will proceed somewhat differently, in the spirit of Proposition 3.43 and (3.12).

Since H is a fixed separable (complex) Hilbert space with $\dim H = \infty$, throughout this section we use the notation

$$\begin{aligned}
&\mathcal{B} = \mathcal{B}(H) := \left\{A \in \mathrm{End}(H) : \|A\| = \sup\nolimits_{\|x\|=1} \{\|Ax\|\} < \infty\right\} \text{ and} \\
&\mathcal{B}^{\times} = \mathcal{B}^{\times}(H) := \left\{A \in \mathcal{B} : A^{-1} \in \mathcal{B}\right\},
\end{aligned}$$

and recall from Section 2.7 that $\mathcal{I}_1$ denotes the Banach space (and ideal in $\mathcal{B}$) of trace class operators, with norm $\|A\|_1 := \mathrm{Tr}\,|A| = \mathrm{Tr}\sqrt{A^* A}$; see (2.46) (p.60) and Propositions 2.79 and 2.80 (p.62). Let $\Lambda^k : \mathcal{B}(H) \to \mathcal{B}\left(\Lambda^k H\right)$ denote the continuous linear map determined by

$$\Lambda^k(A)(v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k.$$

Note that $\Lambda^k(AB) = \Lambda^k(A)\Lambda^k(B)$, and $\Lambda^k(A^*) = \Lambda^k(A)^*$ since

$$\begin{aligned}
&\left\langle \Lambda^k(A)(v_1 \wedge \cdots \wedge v_k), w_1 \wedge \cdots \wedge w_k \right\rangle = \langle Av_1 \wedge \cdots \wedge Av_k, w_1 \wedge \cdots \wedge w_k \rangle \\
&= \frac{1}{k!} \sum \varepsilon_{i_1 \cdots i_k} \langle Av_1, w_{i_1} \rangle \cdots \langle Av_k, w_{i_k} \rangle \\
&= \frac{1}{k!} \sum\nolimits_{(i)_k} \varepsilon_{i_1 \cdots i_k} \langle v_1, A^* w_{i_1} \rangle \cdots \langle v_k, A^* w_{i_k} \rangle \\
&= \left\langle v_1 \wedge \cdots \wedge v_k, \Lambda^k(A^*)(w_1 \wedge \cdots \wedge w_k) \right\rangle.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left|\Lambda^k(A)\right| &= \sqrt{\Lambda^k(A)^* \Lambda^k(A)} = \sqrt{\Lambda^k(A^*)\Lambda^k(A)} = \sqrt{\Lambda^k(A^*A)} \\
&= \sqrt{\Lambda^k\left(|A|^2\right)} = \sqrt{\Lambda^k(|A|)\Lambda^k(|A|)} = \Lambda^k(|A|).
\end{aligned}$$

Hence, the singular values of $\Lambda^k(A)$ coincide with those of $\Lambda^k(|A|)$, namely the products $\mu_{i_1} \cdots \mu_{i_k}$, where the μ_i are the singular values of A. Moreover,

$$\begin{aligned}
\left\|\Lambda^k(A)\right\|_1 &= \operatorname{Tr}\left(\Lambda^k(|A|)\right) = \sum\nolimits_{i_1 < \cdots < i_k} \mu_{i_1} \cdots \mu_{i_k} \le \frac{1}{k!} \sum\nolimits_{i_1, \cdots, i_k} \mu_{i_1} \cdots \mu_{i_k} \\
&= \frac{1}{k!} \operatorname{Tr}(|A|)^k = \frac{1}{k!} \left(\|A\|_1\right)^k,
\end{aligned}$$

and for any $z \in \mathbb{C}$,

$$\begin{aligned}
&\left|\sum\nolimits_{k=0}^{\infty} \operatorname{Tr}\left(\Lambda^k(A)\right) z^k\right| \le \sum\nolimits_{k=0}^{\infty} \left|\operatorname{Tr}\left(\Lambda^k(A)\right)\right| |z|^k \le \sum\nolimits_{k=0}^{\infty} \operatorname{Tr}\left(\left|\Lambda^k(A)\right|\right) |z|^k \\
&= \sum\nolimits_{k=0}^{\infty} \left\|\Lambda^k(A)\right\|_1 |z|^k \le \sum\nolimits_{k=0}^{\infty} \frac{1}{k!} \left(|z| \|A\|_1\right)^k \le e^{|z| \|A\|_1}.
\end{aligned}$$

Since $\sum_{k=0}^{\infty} \operatorname{Tr}\left(\Lambda^k(A)\right) z^k$ exists for any z (in particular $z = 1$), we may make the following definition which agrees with the finite-dimensional case.

Definition 3.44. For $A \in \mathcal{I}_1$, we define

$$\det(\operatorname{Id} + A) = \sum\nolimits_{k=0}^{\infty} \operatorname{Tr}\left(\Lambda^k(A)\right),$$

which is known as the **Fredholm determinant** of $\operatorname{Id} + A$.

Proposition 3.45. *For* $A, B \in \mathcal{I}_1$,

$$|\det(\operatorname{Id} + A) - \det(\operatorname{Id} + B)| \le \|A - B\|_1 \exp\left(\|A\|_1 + \|B\|_1\right),$$

whence $\det(\operatorname{Id} + (\cdot)) : \mathcal{I}_1 \to \mathbb{C}$ *is a continuous function.*

Proof. Since Tr is linear and $\operatorname{Tr}\left(\Lambda^0(A)\right) = \operatorname{Tr}\left(\Lambda^0(B)\right) = \operatorname{Tr}(\operatorname{Id}_{\mathbb{C}}) = 1$,

$$|\det(\operatorname{Id} + A) - \det(\operatorname{Id} + B)| \le \sum\nolimits_{k=1}^{\infty} \left|\operatorname{Tr}\left(\Lambda^k(A) - \Lambda^k(B)\right)\right|.$$

Observe that

$$\begin{aligned}
&\left\langle \left(\Lambda^k(A) - \Lambda^k(B)\right)(v_1 \wedge \cdots \wedge v_k), v_1 \wedge \cdots \wedge v_k \right\rangle \\
&= \sum_{p=0}^{k-1} \begin{pmatrix} \langle Bv_1 \wedge \cdots \wedge Bv_p, v_1 \wedge \cdots \wedge v_p \rangle \cdot \\ \cdot \langle (A - B) v_{p+1}, v_{p+1} \rangle \cdot \\ \cdot \langle Av_{p+2} \wedge \cdots \wedge Av_k, v_{p+2} \wedge \cdots \wedge v_k \rangle \end{pmatrix},
\end{aligned}$$

since

$$\begin{aligned}
&\left(\Lambda^k(A)-\Lambda^k(B)\right)(v_1\wedge\cdots\wedge v_k)\\
&=Av_1\wedge\cdots\wedge Av_k-Bv_1\wedge Av_2\cdots\wedge Av_k\\
&+Bv_1\wedge Av_2\cdots\wedge Av_k-Bv_1\wedge\cdots\wedge Bv_k\\
&=(A-B)\,v_1\wedge Av_2\wedge\cdots\wedge Av_k+Bv_1\wedge Av_2\wedge\cdots\wedge Av_k\\
&-Bv_1\wedge Bv_2\wedge Av_3\cdots\wedge A_k+Bv_1\wedge Bv_2\wedge Av_3\cdots\wedge A_k-Bv_1\wedge\cdots\wedge Bv_k\\
&=(A-B)\,v_1\wedge Av_2\wedge\cdots\wedge Av_k+Bv_1\wedge(A-B)\,v_2\wedge Av_3\cdots\wedge A_k\\
&+Bv_1\wedge Bv_2\wedge Av_3\cdots\wedge A_k-Bv_1\wedge\cdots\wedge Bv_k\\
&=\cdots=\sum_{p=0}^{k-1}Bv_1\wedge\cdots\wedge Bv_p\wedge(A-B)\,v_{p+1}\wedge Av_{p+2}\wedge\cdots\wedge Av_k.
\end{aligned}$$

Thus, using $\left\|\Lambda^k(A)\right\|_1=\frac{1}{k!}\left(\|A\|_1\right)^k$,

$$\begin{aligned}
&\left|\operatorname{Tr}\left(\Lambda^k(A)-\Lambda^k(B)\right)\right|\\
&\le\left|\operatorname{Tr}(A-B)\right|\sum_{p=0}^{k-1}\frac{p!\,(k-1-p)!}{(k-1)!}\operatorname{Tr}\left|\Lambda^p(B)\right|\operatorname{Tr}\left|\Lambda^{k-1-p}(A)\right|\\
&\le\frac{\operatorname{Tr}|A-B|}{(k-1)!}\sum_{p=0}^{k-1}\|B\|_1^p\,\|A\|_1^{k-1-p}\le\frac{\|A-B\|_1}{(k-1)!}\sum_{p=0}^{k-1}\|B\|_1^p\,\|A\|_1^{k-1-p}\\
&\le\frac{\|A-B\|_1}{(k-1)!}\left(\|B\|_1+\|A\|_1\right)^{k-1},\text{ and so}
\end{aligned}$$

$$\begin{aligned}
&\left|\det(\operatorname{Id}+A)-\det(\operatorname{Id}+B)\right|\le\sum\nolimits_{k=1}^{\infty}\left|\operatorname{Tr}\left(\Lambda^k(A)-\Lambda^k(B)\right)\right|\\
&\le\|A-B\|_1\sum\nolimits_{k=1}^{\infty}\frac{1}{(k-1)!}\left(\|B\|_1+\|A\|_1\right)^{k-1}\\
&\le\|A-B\|_1\exp\left(\|B\|_1+\|A\|_1\right).
\end{aligned}$$

□

PROPOSITION 3.46. *For $A,B\in\mathcal{I}_1$, we have*

$$\det\left((\operatorname{Id}+A)(\operatorname{Id}+B)\right)=\det(\operatorname{Id}+A)\det(\operatorname{Id}+B).$$

PROOF. Note that $\det\left((\operatorname{Id}+A)(\operatorname{Id}+B)\right)$ exists, since $(\operatorname{Id}+A)(\operatorname{Id}+B)=\operatorname{Id}+A+B+AB$ and $A+B+AB\in\mathcal{I}_1$ by Exercise 2.78 (p.60) and Proposition 2.79 (p.61). Let $A=\sum_{i=0}^{N_A}\mu_i(A)\langle\cdot,e_i(A)\rangle f_i(A)$ and $B=\sum_{i=0}^{N_B}\mu_i(B)\langle\cdot,e_i(B)\rangle f_i(B)$ denote the canonical expansions (see (2.43), p.57) of A and B. Let

$$A_n:=\sum\nolimits_{i=0}^{n}\mu_i(A)\langle\cdot,e_i(A)\rangle f_i(A)\ \text{ if } n<N_A$$

and $A_n=A$ if $n\ge N_A<\infty$. Define B_n similarly. Let

$$V_n=\operatorname{span}\{e_i(A),f_i(A),e_i(B),f_i(B):i\le n\}$$

and note that $A_n|_{V_n},B_n|_{V_n}\in\operatorname{End}(V_n)$. Since $\dim V_n<\infty$, we have

$$\det\left((\operatorname{Id}+A_n|_{V_n})(\operatorname{Id}+B_n|_{V_n})\right)=\det(\operatorname{Id}+A_n|_{V_n})\det(\operatorname{Id}+B_n|_{V_n}).$$

As $A_n \to A$, $B_n \to B$, and $A_nB_n \to AB$ in $(\mathcal{I}_1, \|\cdot\|_1)$ by Proposition 2.79, we have

$$\det((\mathrm{Id}+A)(\mathrm{Id}+B)) = \det(\mathrm{Id}+A+B+AB)$$
$$= \det\left(\lim_{n\to\infty}(\mathrm{Id}+A_n+B_n+A_nB_n)\right) = \lim_{n\to\infty}\det(\mathrm{Id}+A_n+B_n+A_nB_n),$$

since $\det(\mathrm{Id}+(\cdot))$ is continuous on $(\mathcal{I}_1, \|\cdot\|_1)$ by Proposition 3.45. Since $A_n = 0$ on $V_n^{\perp}$,

$$\det(\mathrm{Id}+A_n) = \sum_{k=0}^{\infty}\mathrm{Tr}\left(\Lambda^k(A_n)\right) = \sum_{k=0}^{\infty}\mathrm{Tr}\left(\Lambda^k(A_n|_{V_n})\right) = \det(1+A_n|_{V_n}),$$

and similarly for $A_n + B_n + A_nB_n$. Thus,

$$\begin{aligned}\det(\mathrm{Id}+A_n+B_n+A_nB_n) &= \det(\mathrm{Id}+A_n|_{V_n}+B_n|_{V_n}+A_n|_{V_n}B_n|_{V_n})\\ &= \det((\mathrm{Id}+A_n|_{V_n})(\mathrm{Id}+B_n|_{V_n}))\\ &= \det(\mathrm{Id}+A_n)\det(\mathrm{Id}+B_n),\ \text{and}\end{aligned}$$

$$\begin{aligned}\det((\mathrm{Id}+A)(\mathrm{Id}+B)) &= \lim_{n\to\infty}\det(\mathrm{Id}+A_n+B_n+A_nB_n)\\ &= \lim_{n\to\infty}\det(\mathrm{Id}+A_n)\det(\mathrm{Id}+B_n)\\ &= \det(\mathrm{Id}+A)\det(\mathrm{Id}+B).\end{aligned}$$ □

Corollary 3.47. *If* $A \in \mathcal{I}_1$, *then* $\mathrm{Id}+A$ *is invertible* $\Leftrightarrow$ $\det(\mathrm{Id}+A) \neq 0$.

Proof. Note that

$$(\mathrm{Id}+A)(\mathrm{Id}+B) = \mathrm{Id}+A+B+AB = \mathrm{Id} \Leftrightarrow B = -A(\mathrm{Id}+A)^{-1}.$$

For $B := -A(\mathrm{Id}+A)^{-1}$, $B \in \mathcal{I}_1$ by Proposition 2.79. Then $\det(\mathrm{Id}+A) \neq 0$, since

$$\det(\mathrm{Id}+A)\det(\mathrm{Id}+B) = \det((\mathrm{Id}+A)(\mathrm{Id}+B)) = \det\mathrm{Id} = 1.$$

If $\mathrm{Id}+A$ is not invertible, then $-1 \in \mathrm{Spec}(A) := \{\lambda \in \mathbb{C} : \lambda\,\mathrm{Id} - A \notin \mathcal{B}^{\times}\}$, and we know that there is a unit eigenvector e with $Ae = -e$ (since $\mathrm{Id}+A \in \mathcal{F}_0$, we have $-1 \notin \mathrm{Spec}_e(A)$ and, actually, $-1 \in \mathrm{Spec}_p(A)$). Let $P \in \mathcal{B}$ denote the orthogonal projection onto $\mathrm{span}(e)$ given by $P(x) := \langle x, e\rangle e$. Note that $AP = -P$. The orthogonal projection $Q = I - P$ onto $\mathrm{span}(e)^{\perp}$, obeys $P + Q = \mathrm{Id}$, and $AQAP = -AQP = 0$. Thus,

$$(\mathrm{Id}+AQ)(\mathrm{Id}+AP) = \mathrm{Id}+A(Q+P)+AQAP\,\mathrm{Id} = \mathrm{Id}+A.$$

Since $\mathrm{Id}+AP = \mathrm{Id}-P$ and $\det(\mathrm{Id}-P) = \sum_{k=0}^{\infty}\mathrm{Tr}\,\Lambda^k(-P) = 1-1 = 0$,

$$\det(\mathrm{Id}+A) = \det(\mathrm{Id}+AQ)\det(\mathrm{Id}+AP) = 0.$$ □

Remark 3.48. So, the zeros of the function $A \mapsto \det(\mathrm{Id}+A)$ arise exactly where $\dim\mathrm{Ker}\,A > 0$. This gives a hint of the intimate relation between index theory and the geometric study of determinants. It may as well legitimize our construction of the determinant bundle via the index bundle. For much deeper relations see, e.g., [**Nas**, Chapter X] where index theory yields an obstruction to the existence of a gauge invariant determinant for classical Dirac operators coupled to a background field (a connection).

The Segal-Furutani Construction. We now begin the alternative construction of the restriction $q : \mathcal{Q}|_{\mathcal{F}_0} \to \mathcal{F}_0$ based on work of G. Segal [**Seg90**] with contributions of K. Furutani [**Fur**]. We refer to [**Boh**, Appendix] for a short cocycle definition of the Segal determinant bundle. For a fixed Hilbert space H, recall that $\mathcal{B}^\times$ denotes the group of invertible elements of the ring $\mathcal{B} := \mathcal{B}(H)$ and $\mathcal{I}_1$ denotes the ideal of trace class operators.

PROPOSITION 3.49. *The space $\mathcal{F}_0$ of all Fredholm operators of index 0 on H equals $\mathcal{B}^\times + \mathcal{I}_1$.*

PROOF. Let $T \in \mathcal{F}_0$, and let $\sigma_T : H \to \operatorname{Ker} T$ denote the orthogonal projection. Since the index of T is 0, there is an isomorphism $L : \operatorname{Ker} T \cong (\operatorname{Im} T)^\perp$. Then $L \circ \sigma_T$ is a finite rank (and hence trace class) operator, and $T + L \circ \sigma_T \in \mathcal{B}^\times$. Indeed, $\operatorname{Ker}(T + L \circ \sigma_T) = \{0\}$, since

$$\begin{aligned} &0 = (T + L \circ \sigma_T)(v) = T(v) + (L \circ \sigma_T)(v) \in \operatorname{Im} T \oplus (\operatorname{Im} T)^\perp \\ &\Longrightarrow T(v) = 0 \text{ and } L(\sigma_T(v)) = 0 \Longrightarrow v \in \operatorname{Ker} T \text{ and } v = \sigma_T(v) = 0. \end{aligned}$$

Moreover, $T + L \circ \sigma_T$ is surjective, since

$$\begin{aligned} &(T + L \circ \sigma_T)(\operatorname{Ker} T) = (L \circ \sigma_T)(\operatorname{Ker} T) = (\operatorname{Im} T)^\perp \quad \text{and} \\ &(T + L \circ \sigma_T)\left((\operatorname{Ker} T)^\perp\right) = T\left((\operatorname{Ker} T)^\perp\right) = \operatorname{Im} T. \end{aligned}$$

Thus, $T + L \circ \sigma_T \in \mathcal{B}^\times$ and $T \in \mathcal{B}^\times - L \circ \sigma_T \subset \mathcal{B}^\times + \mathcal{I}_1$. Conversely, any element of $\mathcal{B}^\times + \mathcal{I}_1$ is of the form $C + K$, where $C \in \mathcal{B}^\times$ and $K \in \mathcal{I}_1 \subset \mathcal{K}$ by Proposition 2.75. Then $C + K \in \mathcal{F}$ by Exercise 3.7 (p.65) which makes use of Theorem 3.2 (Atkinson), p.64. By Exercise 3.10 (p. 65), $\operatorname{index}(C + K) = \operatorname{index} C = 0$, and so $C + K \in \mathcal{F}_0$. □

For $A \in \mathcal{I}_1$, let

$$\mathcal{U}_A := \left\{T \in \mathcal{B} : T + A \in \mathcal{B}^\times\right\} = \mathcal{B}^\times - A \subset \mathcal{B}^\times + \mathcal{I}_1.$$

In other words, $\mathcal{U}_A$ consists of all perturbations of invertible operators by $-A \in \mathcal{I}_1$. Then $\{\mathcal{U}_A : A \in \mathcal{I}_1\}$ is an open cover of $\mathcal{F}_0 = \mathcal{B}^\times + \mathcal{I}_1$ in the topology of the operator norm. To see that $\mathcal{B}^\times + \mathcal{I}_1 = \cup_{A \in \mathcal{I}_1} \mathcal{U}_A$, note that if $S \in \mathcal{B}^\times + \mathcal{I}_1$, then $S = G - A$ for some $A \in \mathcal{I}_1$ and $G \in \mathcal{B}^\times$. Thus, $S + A = G \in \mathcal{B}^\times$ and so $S \in \mathcal{U}_A$. Note that for $T_0 \in \mathcal{F}_0$, we have $T_0 + L_{T_0} \circ \sigma_{T_0} \in \mathcal{B}^\times$ and so $T_0 \in \mathcal{U}_{L_{T_0} \circ \sigma_{T_0}}$. For $A, B \in \mathcal{I}_1$ and $T \in \mathcal{U}_A \cap \mathcal{U}_B$, let

$$g_{AB}(T) := \det\left((T + A)^{-1}(T + B)\right) \in \mathbb{C}. \tag{3.13}$$

This is defined, since composing both sides of $T + B = (T + A) + (B - A)$ on the left with $(T + A)^{-1}$ yields

$$(T + A)^{-1}(T + B) = \operatorname{Id} + (T + A)^{-1}(B - A) \in \operatorname{Id} + \mathcal{I}_1. \tag{3.14}$$

Then Definition 3.44 applies. Note that (3.14) also implies that g_{AB} is continuous. Indeed, $\mathcal{U}_A \cap \mathcal{U}_B \to \mathcal{I}_1$, given by

$$T \mapsto (T + A)^{-1}(B - A),$$

is continuous since $T \mapsto (T + A)^{-1} \in \mathcal{B}$ is continuous in the operator norm topology, and by Proposition 2.79 the composition multiplication $\mathcal{B} \times \mathcal{I}_1 \to \mathcal{I}_1$ is (jointly) continuous. Then Proposition 3.45 yields the continuity of

$$T \mapsto g_{AB}(T) = \det\Big(\mathrm{Id} + (T+A)^{-1}(B-A)\Big).$$

For $A, B, C \in \mathcal{I}_1$and $T \in \mathcal{U}_A \cap \mathcal{U}_B \cap \mathcal{U}_C$, we have the cocycle condition

$$g_{AC}(T) = g_{AB}(T)\, g_{BC}(T).$$

Indeed, using Proposition 3.46

$$\begin{aligned} g_{AB}(T)\, g_{BC}(T) &= \det\Big((T+A)^{-1}(T+B)\Big) \det\Big((T+B)^{-1}(T+C)\Big) \\ &= \det\Big((T+A)^{-1}(T+B)(T+B)^{-1}(T+C)\Big) = \det\Big((T+A)^{-1}(T+C)\Big). \end{aligned}$$

Let $\pi : \mathcal{S} \to \mathcal{F}_0$ denote the line bundle defined by $\{g_{AB}\}$; i.e., $\mathcal{S}$ is the disjoint union of $\{\mathcal{U}_A \times \mathbb{C} : A \in \mathcal{I}_1\}$, but with the identifications

$$\begin{aligned} &(T, z_A) \in \mathcal{U}_A \times \mathbb{C} \sim (T, z_B) \in \mathcal{U}_B \times \mathbb{C} \\ &\Leftrightarrow z_A = g_{AB}(T)\, z_B = \det\Big((T+A)^{-1}(T+B)\Big) z_B. \end{aligned}$$

In particular, for $T \in \mathcal{B}^\times + \mathcal{I}_1$, say $T - A \in \mathcal{B}^\times$ for $A \in \mathcal{I}_1$, the fiber $\pi^{-1}(T)$ can be written as

$$\pi^{-1}(T) = [T, z_A] = \{(T, z_B) \in \mathcal{U}_B \times \mathbb{C} : B \in \mathcal{I}_1 \text{ and } z_A = g_{AB}(T)\, z_B\}.$$

We now show that $\mathcal{Q}|_{\mathcal{F}_0}$ (with $\mathcal{Q}$ defined in (3.10), p.95) can be made into a genuine line bundle by exhibiting identification of $\mathcal{Q}|_{\mathcal{F}_0}$ with $\mathcal{S}$. Let $T \in \mathcal{F}_0$. As in Proposition 3.49, we have the orthogonal projection $\sigma_T : H \to \operatorname{Ker} T$ and some isomorphism $L : \operatorname{Ker} T \cong \operatorname{Ker} T^*$. Let $e_1, \ldots, e_d$ be a basis for $\operatorname{Ker} T$, and let $e_1^*, \ldots, e_d^*$ denote the dual basis for $(\operatorname{Ker} T)^*$. Define

$$\phi_A : \mathcal{U}_A \times \mathbb{C} \longrightarrow \mathcal{Q}|_{\mathcal{F}_0} := \bigcup_{T \in \mathcal{U}_A} \Lambda^d\left((\operatorname{Ker} T)^*\right) \otimes \Lambda^d(\operatorname{Ker} T^*) \text{ by} \tag{3.15}$$

$$\begin{aligned} &\phi_A(T, z_A) := z_A \det\Big((T + L \circ \sigma_T)^{-1}(T+A)\Big) \cdot \\ &e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d), \text{ for } T \in \mathcal{U}_A,\ z_A \in \mathbb{C}. \end{aligned}$$

We now show that $\phi_A(T, z_A)$ is independent of the choice of basis $e_1, \ldots, e_d$; later we show that it is also independent of the choice of L. For a new basis $F(e_1), \ldots, F(e_d)$ where $F \in \mathrm{GL}(\operatorname{Ker} T)$, note that we have

$$(L \circ F)(e_1) \wedge \cdots \wedge (L \circ F)(e_d) = (\det F)\,(L(e_1) \wedge \cdots \wedge L(e_d)).$$

If $F(e_j) = \sum_i F_j^i e_i$, then

$$\delta_{ij} = F(e_i)^*(F(e_j)) = F(e_i)^*\Big(\sum\nolimits_k F_j^k e_k\Big) = \sum\nolimits_k F(e_i)^*(e_k)\, F_j^k.$$

Since $\delta_{ij} = \sum_k \left(F^{-1}\right)_k^i F_j^k$, we have

$$\begin{aligned} F(e_i)^*(e_k) &= \left(F^{-1}\right)_k^i = \Big(\sum\nolimits_j \left(F^{-1}\right)_j^i e_j^*\Big)(e_k), \text{ and so} \\ F(e_i)^* &= \sum\nolimits_j \left(F^{-1}\right)_j^i e_j^*. \end{aligned}$$

Thus,

$$F(e_1)^* \wedge \cdots \wedge F(e_d)^* = \det\left(F^{-1}\right) e_1^* \wedge \cdots \wedge e_d^*.$$

Then ϕ_T is independent of the choice of basis $e_1, \ldots, e_d$, since

$$\begin{aligned}
&F(e_1)^* \wedge \cdots \wedge F(e_d)^* \otimes (L \circ F)(e_1) \wedge \cdots \wedge (L \circ F)(e_d) \\
&= \left(\det\left(F^{-1}\right) \det F\right) e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d) \\
&= e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d).
\end{aligned}$$

We now show that ϕ_T is also independent of L. Suppose that $L' : \operatorname{Ker} T \cong \operatorname{Ker} T^*$ is another isomorphism. Then

$$\begin{aligned}
&L'(e_1) \wedge \cdots \wedge L'(e_d) \\
&= \det\left(L' \circ L^{-1}\right) L(e_1) \wedge \cdots \wedge L(e_d).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\det\Big(\left(T + L' \circ \sigma_T\right)^{-1}(T + A)\Big) e_1^* \wedge \cdots \wedge e_d^* \otimes L'(e_1) \wedge \cdots \wedge L'(e_d) \\
&= \det\Big(\left(T + L' \circ \sigma_T\right)^{-1}(T + A)\Big) \det\left(L' \circ L^{-1}\right) \cdot \\
&\cdot e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d).
\end{aligned}$$

Hence, we must show that

$$\begin{aligned}
&\det\Big(\left(T + L' \circ \sigma_T\right)^{-1}(T + A)\Big) \det\Big(L' \circ \left(\tau_T \circ L\right)^{-1}\Big) \\
&= \det\Big(\left(T + L \circ \sigma_T\right)^{-1}(T + A)\Big), \text{ or equivalently,}
\end{aligned}$$

$$\det\left(L' \circ L^{-1}\right) = \det\Big(\left(T + L' \circ \sigma_T\right) \circ \left(T + L \circ \sigma_T\right)^{-1}\Big). \tag{3.16}$$

For this, note that for $v \in (\operatorname{Ker} T)^{\perp}$,

$$\left(T + L \circ \sigma_T\right)(v) = T(v) = \left(T + L' \circ \sigma_T\right)(v),$$

and for $v \in \operatorname{Ker} T$,

$$\left(T + L \circ \sigma_T\right)(v) = L(v), \text{ while } \left(T + L' \circ \sigma_T\right)(v) = L'(v).$$

Hence, (3.16) holds, since

$$\left(T + L' \circ \sigma_T\right)\left(T + L \circ \sigma_T\right)^{-1} = \operatorname{Id}_{(\operatorname{Ker} T)^{\perp}} \oplus \left(L' \circ L^{-1}\right).$$

For $T \in \mathcal{U}_A \cap \mathcal{U}_B$, we have

$$\phi_A(T, z_A) = \phi_B(T, z_B) \Leftrightarrow z_A = g_{AB}(T) z_B, \text{ since}$$

$$\begin{aligned}
&z_A \det\Big(\left(T + L \circ \sigma_T\right)^{-1}(T + A)\Big) \\
&= g_{AB}(T) z_B \det\Big(\left(T + L \circ \sigma_T\right)^{-1}(T + A)\Big) \\
&= z_B \det\Big(\left(T + L \circ \sigma_T\right)^{-1}(T + A)\Big) \det\Big((T + A)^{-1}(T + B)\Big) \\
&= z_B \det\Big(\left(T + L \circ \sigma_T\right)^{-1}(T + B)\Big).
\end{aligned}$$

Thus, we have a well defined bijection

$$\Phi : \mathcal{S} \longrightarrow \mathcal{Q}|_{\mathcal{F}_0} \text{ given by } \Phi([T, z_A]) := \phi_A(T). \tag{3.17}$$

So far we have not given $\mathcal{Q}|_{\mathcal{F}_0}$ a topology. At this point, the easiest way to give $\mathcal{Q}|_{\mathcal{F}_0}$ a topology is to assert that Φ is a homeomorphism, since the line bundle $\mathcal{S}$

has a topology given by the topologies on the $\mathcal{U}_A \times \mathbb{C}$, after taking the quotient by the equivalence relation.

We summarize what has been done thus far.

PROPOSITION 3.50. *If $\mathcal{Q}|_{\mathcal{F}_0}$ is given the unique topology so that the bijection $\Phi : \mathcal{S} \to \mathcal{Q}|_{\mathcal{F}_0}$ in (3.17) is a homeomorphism, then $\mathcal{Q}|_{\mathcal{F}_0}$ inherits the structure of a complex line bundle from $\mathcal{S}$. Explicitly, the local trivializations of $\mathcal{Q}|_{\mathcal{F}_0}$ are given by the maps*

$$\phi_A : \mathcal{U}_A \times \mathbb{C} \longrightarrow \mathcal{Q}|_{\mathcal{U}_A},$$

found in (3.15), and the transition functions are $g_{AB} : \mathcal{U}_A \cap \mathcal{U}_B \to \mathbb{C}$, given by

$$g_{AB}(T) := \det\Big((T+A)^{-1}(T+B)\Big).$$

EXERCISE 3.51. Use Proposition 3.42, p.93, and the fact that for any $T \in \mathcal{F}$, there is some n_T such that $P_{n_T} T(H) = H_{n_T}^{\perp}$ for some neighborhood of T in $\mathcal{F}$, in order to *directly* give $\mathcal{Q}|_{\mathcal{F}_0}$ a possibly different alternate line bundle structure. Then show that this alternate structure is in fact equivalent to that induced by $\Phi : \mathcal{S} \to \mathcal{Q}|_{\mathcal{F}_0}$.

COROLLARY 3.52. *For a continuous family $T = \{T_x\}_{x\in X} : X \to \mathcal{F}_0$, the determinant bundle $\det T \to X$ is the pull-back $T^*(\mathcal{Q}|_{\mathcal{F}_0})$ of the line bundle $\mathcal{Q}|_{\mathcal{F}_0} \to \mathcal{F}_0$. Here*

$$\det T := \det(\operatorname{index} T).$$

PROOF. For $x \in X$, we have

$$T^*(\mathcal{Q}|_{\mathcal{F}_0})_x = (\mathcal{Q}|_{\mathcal{F}_0})_{T_x} = \Lambda^{\max}(\operatorname{Ker}(T_x))^* \otimes \Lambda^{\max}(\operatorname{Ker}(T_x^*)) = (\det T)_x.$$

Thus, the fibers of $T^*(\mathcal{Q}|_{\mathcal{F}_0})$ coincide with those of $\det T$. This does not yet prove that $\det T = T^*(\mathcal{Q}|_{\mathcal{F}_0})$ as line bundles, but this is a consequence of the preceding Exercise 3.51 if we give $\mathcal{Q}|_{\mathcal{F}_0}$ its alternate structure. Without giving Exercise 3.51 away, let $P_n : H \to H_n := \{e_0, \dots, e_n\}^{\perp}$ be an orthogonal projection so that $P_n T_x H = H_n$ for all $x \in X$, and note that

$$\det T \cong \det P_n T = \Lambda^{\max}(\operatorname{Ker} P_n T)^* \otimes \Lambda^{\max}(H_n^{\perp}) = (P_n T)^*(\mathcal{Q}|_{\mathcal{F}_0}). \qquad \square$$

Besides $\mathcal{Q}|_{\mathcal{F}_0}$, we now show that there is another way to interpret the line bundle $\mathcal{S}$. This is also essentially due to Graeme Segal; see [**Seg90**]. For $T \in \mathcal{F}_0 = \mathcal{B}^\times + \mathcal{I}_1$, let

$$\mathcal{F}_T := T + \mathcal{I}_1 \text{ and } \mathcal{F}_T^\times := \mathcal{F}_T \cap \mathcal{B}^\times \neq \emptyset.$$

(Note that there appears to be a notational problem in using $\mathcal{F}_0$ and $\mathcal{F}_T$, but since $T \in \mathcal{F}_0$, we never have $T = 0$.) In particular, $\mathcal{F}_{\mathrm{Id}} = \mathrm{Id} + \mathcal{I}_1$ is the set of operators on which the Fredholm determinant (see Definition 3.44) is defined. For $S \in \mathcal{F}_T^\times$, define a continuous bijection

$$\Phi_S : \mathcal{F}_T \longrightarrow \mathcal{F}_{\mathrm{Id}} \text{ by } \Phi_S(R) = S^{-1}R, \text{ where } R \in \mathcal{F}_T.$$

Note that $S^{-1}R \in \mathcal{F}_{\mathrm{Id}}$, since

$$R = T + B,\ S = T + B' \text{ (where } B,\ B' \in \mathcal{I}_1)$$

$$\Longrightarrow S^{-1}R = (T+B')^{-1}(T+B) = \mathrm{Id} + (T+B')^{-1}(B - B') \in \mathrm{Id} + \mathcal{I}_1 = \mathcal{F}_{\mathrm{Id}}.$$

Also, Φ_S is clearly 1-1, and we see that Φ_S is onto, as follows. For $C = I + B'' \in \mathcal{F}_{\mathrm{Id}}$, we have (by Proposition 2.79)

$$SC = S(I + B'') = S + SB'' \in \mathcal{F}_T^\times + \mathcal{I}_1 \subseteq \mathcal{B}^\times + \mathcal{I}_1.$$

Since $C = S^{-1}SC = \Phi_S(SC)$, Φ_S is onto. On $\mathcal{F}_{\mathrm{Id}} \times \mathbb{C}$, we have an equivalence relation

$$(R, z) \sim (R', z') \Leftrightarrow z \det R = z' \det R'.$$

Alternatively, there is a map

$$\kappa^0 : \mathcal{F}_{\mathrm{Id}} \times \mathbb{C} \longrightarrow \mathbb{C} \text{ given by } \kappa^0(R, z) := z \det R,$$

and the (huge) equivalence classes are just the preimages of points in $\mathbb{C}$. We denote the set of equivalence classes $[R, z]$ by

$$E_{\mathrm{Id}} := \{[R, z] : (R, z) \in \mathcal{F}_{\mathrm{Id}} \times \mathbb{C}\} = \left\{ \left(\kappa^0\right)^{-1}(w) : w \in \mathbb{C} \right\}.$$

Since $\kappa^0(\mathrm{Id}, z) = z$, there is a bijection

$$\kappa : E_{\mathrm{Id}} \longrightarrow \mathbb{C} \;\text{ given by }\; \kappa([R, z]) = \kappa^0(R, z) = z \det R,$$

and E_{Id} inherits a vector space structure from $\mathbb{C}$ via κ. For $S \in \mathcal{F}_T^\times$, we have a map

$$\begin{aligned} &\kappa^0_{T,S} := \kappa^0 \circ (\Phi_S \times \mathrm{Id}) : \mathcal{F}_T \times \mathbb{C} \longrightarrow \mathbb{C}, \text{ or (for } A \in \mathcal{I}_1) \\ &\kappa^0_{T,S}(T + A, z) = \kappa^0\left(S^{-1}(T + A), z\right) = z \det\left(S^{-1}(T + A)\right), \end{aligned}$$

which defines an equivalence relation on $\mathcal{F}_T \times \mathbb{C}$. Note that for $S' \in \mathcal{F}_T^\times$, we have

$$\begin{aligned} &\kappa^0_{T,S}(T + A, z) = z \det\left(S^{-1}(T + A)\right) = z \det\left(S^{-1}S'S'^{-1}(T + A)\right) \\ &= \det\left(S^{-1}S'\right) z \det\left(S'^{-1}(T + A)\right) = \det\left(S^{-1}S'\right) \kappa^0_{T,S'}(T + A, z), \text{ and so} \end{aligned}$$

$$\kappa^0_{T,S} = \det\left(S^{-1}S'\right) \kappa^0_{T,S'}. \tag{3.18}$$

Since $\det\left(S^{-1}S'\right) \neq 0$, $\kappa^0_{T,S'}$ defines the same equivalence relation, say $\sim_T$, on $\mathcal{F}_T \times \mathbb{C}$. For $T \in \mathcal{F}_0 = \mathcal{B}^\times + \mathcal{I}_1$, let

$$\begin{aligned} &E_T := (\mathcal{F}_T \times \mathbb{C}) / \sim_T = \{[R, z] : (R, z) \in \mathcal{F}_T \times \mathbb{C}\}, \\ &E := \left\{(T, [R, z]) : T \in \mathcal{B}^\times + \mathcal{I}_1 \text{ and } [R, z] \in E_T\right\}, \text{ and let} \\ &p : E \longrightarrow \mathcal{B}^\times + \mathcal{I}_1 \text{ be given by } p(T, [R, z]) = T. \end{aligned}$$

Note that for $S \in \mathcal{F}_T^\times$, $\Phi_S : \mathcal{F}_T \longrightarrow \mathcal{F}_{\mathrm{Id}}$ induces

$$[\Phi_S] : E_T \longrightarrow E_{\mathrm{Id}} \;\text{ defined by }\; [\Phi_S]([T + A, z]) = [\Phi_S(T + A), z] = \left[S^{-1}(T + A), z\right].$$

Then $[\Phi_S] : E_T \to E_{\mathrm{Id}}$ defines a vector space structure on E_T which is independent of the choice of $S \in \mathcal{F}_T^\times$. However, the identification $\kappa_{T,S} : E_T \to \mathbb{C}$, given by

$$\kappa_{T,S}([R, z]) := \kappa([\Phi_S][R, z]) = \kappa([\Phi_S(R), z]) = z \det\left(S^{-1}R\right),$$

depends on S. Using (3.18), we have

$$\kappa_{T,S} = \det\left(S^{-1}S'\right) \kappa_{T,S'}.$$

For $A \in \mathcal{I}_1$, recall that $\mathcal{U}_A := \mathcal{B}^\times - A$ and that $\{\mathcal{U}_A : A \in \mathcal{I}_1\}$ is an open cover of $\mathcal{B}^\times + \mathcal{I}_1$. Although $\kappa_{T,S} : E_T \to \mathbb{C}$ depends on $S \in \mathcal{F}_T^\times$, for each $T \in \mathcal{U}_A$, there is a natural choice for $S \in \mathcal{F}_T^\times$, namely $S = T + A$. Then we may define

$$\psi_A : p^{-1}(\mathcal{U}_A) \longrightarrow \mathcal{U}_A \times \mathbb{C} \text{ by (where } [R, z] \in E_T, T \in \mathcal{U}_A)$$

$$\psi_A(T, [R, z]) := \left(T, \kappa_{T,(T+A)}([R, z])\right) = \left(T, z \det\left((T + A)^{-1} R\right)\right). \tag{3.19}$$

For $B \in \mathcal{I}_1$, note that from $\kappa_{T,S} = \det\left(S^{-1}S'\right)\kappa_{T,S'}$ and (3.13), we have

$$\kappa_{T,(T+A)} = \det\left((T+A)^{-1}(T+B)\right)\kappa_{T,(T+B)} = g_{AB}(T)\,\kappa_{T,(T+B)}.$$

Thus, for $z_A = \kappa_{T,(T+A)}([R,z])$ and $z_B = \kappa_{T,(T+B)}([R,z])$, we have

$$\left(\psi_A \circ \psi_B^{-1}\right)(T, z_B) = \psi_{AA}(T,[R,z]) = (T, z_A)\,, \text{ where}$$
$$z_A = \kappa_{T,(T+A)}([R,z]) = g_{AB}(T)\,\kappa_{T,(T+B)}([R,z]) = g_{AB}(T)\,z_B.$$

Hence, the transition functions for the bundle $p : E \to \mathcal{B}^\times + \mathcal{I}_1$ relative to the trivializations Ψ_A are the same as those for the bundle $\pi : \mathcal{S} \to \mathcal{B}^\times + \mathcal{I}_1$, and so the Ψ_A induce an isomorphism

$$\Psi : E \cong \mathcal{S}.$$

By Proposition 3.50, the bundle $\mathcal{S}|_{\mathcal{F}_0}$ is isomorphic to the Quillen bundle $\mathcal{Q}|_{\mathcal{F}_0}$. In summary we have

Theorem 3.53. *There are three isomorphic complex line bundles over the component $\mathcal{F}_0$ of Fredholm operators with index zero on a fixed Hilbert space H, namely*

$$\pi : \mathcal{S} \longrightarrow \mathcal{F}_0, \quad q : \mathcal{Q}|_{\mathcal{F}_0} \longrightarrow \mathcal{F}_0 \quad \text{and } p : E \longrightarrow \mathcal{F}_0.$$

For $A \in \mathcal{I}_1$, and $T \in \mathcal{U}_A = \mathcal{B}^\times - A$, the point in $q^{-1}(T)$ corresponding to $[T, z_A] \in \pi^{-1}(T)$ is

$$z_A \det\left((T + L\circ\sigma_T)^{-1}(T+A)\right) e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d)\,, \tag{3.20}$$

which is independent of the choice of isomorphism $L : \operatorname{Ker} T \cong \operatorname{Ker} T^*$ and the choice of basis $\{e_1, \ldots, e_d\}$ of $\operatorname{Ker} T$; here σ_T is the orthogonal projection of H onto $\operatorname{Ker} T$. The point in $p^{-1}(T)$ corresponding to $[T, z_A] \in \pi^{-1}(T)$ is

$$(T, [T + A, z_A]) \in p^{-1}(T)\,,$$

since (by (3.19))

$$\psi_A(T, [T+A, z_A]) = \left(T, \kappa_{T,(T+A)}([T+A, z_A])\right)$$
$$= \left(T, z_A \det\left((T+A)^{-1}(T+A)\right)\right) = (T, z_A)\,.$$

Remark 3.54. Recall that

$$[T+A, z_A] = \left\{(R,z) : z_A = z\det\left((T+A)^{-1}R\right) \text{ and } R \in T + \mathcal{I}_1\right\}. \tag{3.21}$$

Taking $R = T + L\circ\sigma_T \in (T + \mathcal{I}_1) \cap \mathcal{B}^\times$ in (3.21), we obtain

$$(T, [T+A, z_A]) = \left(T, [R, z_A/\det\left((T+A)^{-1}R\right)]\right)$$
$$= \left(T, [R, z_A \det\left(R^{-1}(T+A)\right)]\right).$$

The $z_A \det\left(R^{-1}(T+A)\right)$ in this last expression is precisely the factor multiplying $e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d)$ in (3.20), but neither (3.21) nor (3.20) is determined by this factor alone.

In the fiber $p^{-1}(T)$, there are standard elements

$$[T,1] = \left\{(R,z) \in \mathcal{F}_T \times \mathbb{C} : z\det S^{-1}R = \det S^{-1}T \text{ for all } S \in \mathcal{F}_T^\times\right\} \text{ and}$$
$$[T,0] = \left\{(R,z) \in \mathcal{F}_T \times \mathbb{C} : z\det S^{-1}R = 0 \text{ for all } S \in \mathcal{F}_T^\times\right\}.$$

Actually $[T,0]$ is just the zero element in $p^{-1}(T)$. For $T \in \mathcal{F}_0$,

$$[T,1] = [T,0] \Leftrightarrow \det S^{-1}T = 0 \text{ for all } S \in \mathcal{F}_T^{\times} \Leftrightarrow S^{-1}T \notin \mathcal{B}^{\times} \text{ for all } S \in \mathcal{F}_T^{\times}$$
$$\Leftrightarrow T \notin \mathcal{F}_T^{\times} \text{ (i.e., } T \text{ is not invertible).}$$

The **canonical section** $\sigma : \mathcal{F}_0 \to E$ is defined by $\sigma(T) = [T,1]$. We have just seen that $\sigma(T) \neq 0 \Leftrightarrow T \in \mathcal{B}^{\times}$. On $\mathcal{U}_A$, we have (see (3.19))

$$\psi_A(\sigma(T)) = \psi_A([T,1]) = \left(T, \det\left((T+A)^{-1}T\right)\right).$$

In other words, the local representative $\sigma_A : \mathcal{U}_A \to \mathbb{C}$ of σ is given by

$$\sigma_A(T) = \det\left((T+A)^{-1}T\right) = \det\left(\mathrm{Id} - (T+A)^{-1}A\right).$$

Setting $z_A = \det\left((T+A)^{-1}T\right)$ in (3.20) and noting that

$$\det\left((T + L\circ\sigma_T)^{-1}(T+A)\right)\det\left((T+A)^{-1}T\right) = \det\left((T+A)^{-1}T\right)$$

(by Proposition 3.46, p.99), we obtain the Quillen version of the canonical section, namely $\sigma_q : \mathcal{F}_0 \to \mathcal{Q}|_{\mathcal{F}_0}$ given by

$$\sigma_q(T) = \det\left((T + L\circ\sigma_T)^{-1}T\right)e_1^* \wedge \cdots \wedge e_d^* \otimes L(e_1) \wedge \cdots \wedge L(e_d),$$

where again $L : \operatorname{Ker} T \to \operatorname{Ker} T^*$ is an arbitrary isomorphism and $\{e_1, \ldots, e_d\}$ is an arbitrary basis of $\operatorname{Ker} T$.

EXERCISE 3.55. Show directly that for $T \in \mathcal{F}_0$,

$$\sigma_q(T) = \begin{cases} 0 \in \Lambda^d(\operatorname{Ker} T)^* \otimes \Lambda^d(\operatorname{Ker} T^*), & \text{if } T \notin \mathcal{B}^{\times}, \\ 1 \otimes 1 \in \mathbb{C} \otimes \mathbb{C}, & \text{if } T \in \mathcal{B}^{\times}. \end{cases}$$

From this, it would seem that σ_q is discontinuous. Is this really the case? One may wish to read the relevant discussion and footnotes in [**Nas**, p.276].

REMARK 3.56. At the end of the next Section, we provide a brief glimpse of the fascinating relations (discovered in [**ScWo00**]) between the rich structure of the determinant line bundle (and the canonical section) to the corresponding object in physics, namely the zeta-function regularized determinant of Dirac operators. Presently, we will only note the following recent development. On the infinitedimensional manifold $\mathcal{F}_0$, we have the Quillen-Segal determinant line bundle E. This is a rather complicated object, say of *degree* one. On the space $\Omega\mathcal{F}_{\mathrm{Id}}$ of loops $\{T : S^1 \to \mathcal{F}_{Id}\}$, i.e., $T(\theta) = \mathrm{Id} + A_\theta$ invertible for $\theta \in S^1$ with $A_\theta \in \mathcal{I}_1$, we have $U(1)$-valued functions, $\det(T)$ (after normalization) which are simpler objects, say of *degree* zero. Currently, a very active research field is the investigation of similar *transgressions* between *an object of degree* $k+1$ *on a base* and a corresponding *object of degree* k *on the free loop space of the base.*

10. Essential Unitary Equivalence and Spectral Invariants

Our main focus in this book is on the index, which is stable under rather general deformations of the operator in question. However, in this section we wish to broaden the perspective somewhat by indicating some of the finer attributes of operators which are captured by more sensitive quantities constructed from their spectra; i.e., *spectral invariants.*

The power and limitations of the homotopy theoretic technique in analysis, specifically in the theory of Fredholm operators, are demonstrated in results of Lawrence Brown, Lewis Coburn, Ronald Douglas, Peter Fillmore, William Helton, Roger Howe and others. We will review them briefly; for details see the collection [**Fi**, 1973].

1. Let $\mathcal{B}$ denote the Banach algebra of linear bounded operators on the Hilbert space H with the closed ideal $\mathcal{K}$ of compact operators and the canonical projection $\pi : \mathcal{B} \to \mathcal{B}/\mathcal{K}$. For $S, T \in \mathcal{B}$, we define (in $\mathcal{B}/\mathcal{K}$) *essential unitary equivalence* or *unitary equivalence mod* $\mathcal{K}$, denoted by $\pi(S) \approx \pi(T)$, by either of the following equivalent ([**Fi**, 1973, p.77]) conditions:

(i) There is a unitary operator $U \in \mathcal{B}$ (i.e., $U^* = U^{-1}$), such that $S - UTU^* \in \mathcal{K}$.

(ii) There is a unitary element v in $\mathcal{B}/\mathcal{K}$ such that $\pi(S) = v\pi(T)v^*$.

Now let $S \in \mathcal{F}$; i.e., $\pi(S)$ is an invertible element in $\mathcal{B}/\mathcal{K}$ (see Theorem 3.2, p. 64). What is the relationship in $\mathcal{F}$ between

- the topological relation $S \sim T$ (S and T are *homotopic*; i.e., S and T can be connected by a continuous path in $\mathcal{F}$, or equivalently: $\operatorname{index} S = \operatorname{index} T$) and
- the numerical-analytic relation $\pi(S) \approx \pi(T)$; i.e., S and T are modulo K *unitarily equivalent*?

Since $\mathcal{B}$ and even more $\mathcal{U}$ are connected (Theorem 3.21), the homotopy equivalence follows trivially from the unitary equivalence modulo $\mathcal{K}$. The converse is not immediately clear. Rather, by looking, for example, at the homotopic operators Id and $-\operatorname{Id}$, it is apparent that Fredholm operators from the same path-connected component of $\mathcal{F}$ may well differ by "more" than a compact operator. However, Brown, Douglas and Fillmore showed in 1970: In the group of unitary elements of $\mathcal{B}/\mathcal{K}$ the relations "$S \sim T$" and "$\pi(S) \approx \pi(T)$" coincide; i.e., *essentially unitary* operators in Hilbert space can be joined by a continuous path in $\mathcal{F}$ (i.e., they have the same index) if and only if they are *unitarily equivalent modulo* $\mathcal{K}$. In still another way:

By [**Fi**, 1973, p.71], the classes of unitarily equivalent unitary elements of $\mathcal{B}/\mathcal{K}$ form an infinite cyclic group with representatives

$$\pi(\mathrm{Id}) \text{ or } \pi((\mathrm{shift}^+)^n) \text{ or } \pi((\mathrm{shift}^-)^n),\ n \in \mathbb{N}.$$

For all *essentially-unitary* operators $T \in \mathcal{B}$ (i.e., $\pi(T)$ is unitary in $\mathcal{B}/\mathcal{K}$), the *essential spectrum* $\operatorname{Spec}_e(T) = \operatorname{Spec}\pi(T)$ is contained in $S^1 = \{z \in \mathbb{C} : |z| = 1)\}$, since $\|\pi(T)\| = 1$ (see Exercise 3.6, p.65). Since $\operatorname{index}(T - z\,\mathrm{Id}) = 0$ whenever $|z| > 1$, the homotopy invariance of the index implies that the inclusion $\operatorname{Spec}_e(T) \subseteq S^1$ is proper, only if $\operatorname{index} T = 0$. Thus each class of unitarily equivalent unitary elements in $\mathcal{B}/\mathcal{K}$ is characterized by the index, where for fixed T the index of $T - z\,\mathrm{Id}$ is a map on $\mathbb{C} - S^1$ into $\mathbb{Z}$ with value 0 everywhere outside S^1 and constant value $n = \operatorname{index} T$ inside S^1 as depicted in Figure 3.3. In geometric language: The index is a *complete* unitary invariant for the unitary elements in $\mathcal{B}/\mathcal{K}$.

2. For normal operators ($T \in \mathcal{N} \Leftrightarrow T^*T - TT^* = 0$), however, the essential spectrum is the *complete* unitary invariant[4], while the index is identically 0 in

[4] I.D. Berg, Trans. Amer. Math. Soc. 160 (1971), 365-371, shows that every normal operator can be diagonalized by a compact perturbation, and that $S, T \in \mathcal{N}$ are unitarily equivalent modulo $\mathcal{K}$ (i.e., there is a unitary operator U with $S - UTU^* \in \mathcal{K}$), if and only if $\operatorname{Spec}_e(T) = \operatorname{Spec}_e(S)$. For self-adjoint S, T this result is due to John von Neumann, whose point of departure was a lemma

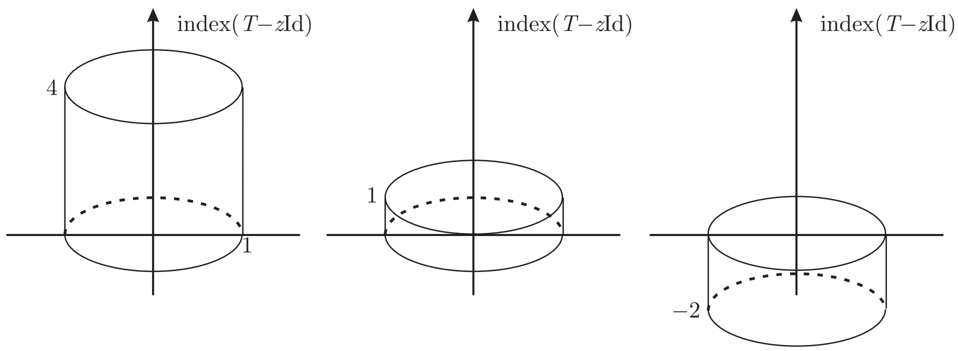


FIGURE 3.3. Three classes of unitarily equivalent elements in $\mathcal{B}/\mathcal{K}$, distinguished by the height $n = 4, 1, -2$ of the $|z| < 1$-towers

the complement of the essential spectrum (see Remark 2.11, p. 17). The class of operators which is the most natural next object of study after the normal and Fredholm operators are (working modulo $\mathcal{K}$ at any rate) the *essentially normal* operators, i.e., the operators $T \in \mathcal{B}$ for which $\pi(T)$ is normal, i.e., $TT^* - T^*T \in \mathcal{K}$. This includes in particular the compact perturbations of normal operators (nothing new since their index vanishes) and the essentially-unitary operators. We have the following theorem of Brown-Douglas-Fillmore [**BDF**, 1975]: Two essentially normal operators S and T are unitarily equivalent modulo $\mathcal{K}$, if and only if they have the same essential spectrum X and if on every connected component of $\mathbb{C} \setminus X$, we have $\operatorname{index}(S - z\,\mathrm{Id}) = \operatorname{index}(T - z\,\mathrm{Id})$ [**Fi**, 1973, p.73-122].

The proof of this theorem with its dependence on delicate questions of topological algebra and K-theory is by no means trivial. Let us consider the situation once more: The essential spectrum X of an essentially normal operator T is a compact subset of $\mathbb{C}$. The expression $\operatorname{index}(T - z\,\mathrm{Id})$ defines a $\mathbb{Z}$-valued function on $\mathbb{C} - X$. The connected components of $\mathbb{C} - X$ with non-vanishing index are (so to speak) the *obstructions* to the normality of T.

3. While pathwise connectedness in $\mathcal{F}$ yields nothing but a decomposition into the $\mathbb{Z}$ connected components (a classification by index $T - z\,\mathrm{Id}$ at the point $z = 0$) we have in 2 above a finer classification by *index towers* on the connected components of the complement of the essential spectrum as depicted in Figure 3.4.

We state some concrete consequences which form a transition to the next chapter.

(i) An essentially normal operator T with $\operatorname{index}(T - z\,\mathrm{Id}) = 0$ for all z outside the essential spectrum of T belongs to $\mathcal{N} + \mathcal{K}$, i.e., can be written as a sum of a normal and a compact operator [**Fi**, p.118]

(ii) The family $\mathcal{N} + \mathcal{K}$ is topologically closed (in the operator norm); thus elements of the complement of $\mathcal{N} + \mathcal{K}$ in $\mathcal{B}$ (even if their index vanishes) cannot be approximated by a sequence in $\mathcal{N} + \mathcal{K}$ [**Fi**, p.119]. Unfortunately, there is not yet an elementary proof for this remarkable result.

(iii) It is another very interesting fact, that essentially normal operators whose essential spectrum is described by the image of a simple, closed curve are, modulo $\mathcal{K}$,

by Hermann Weyl (1909) saying that the accumulation points of the spectrum of a self-adjoint operator remain unchanged under perturbation by a compact operator.

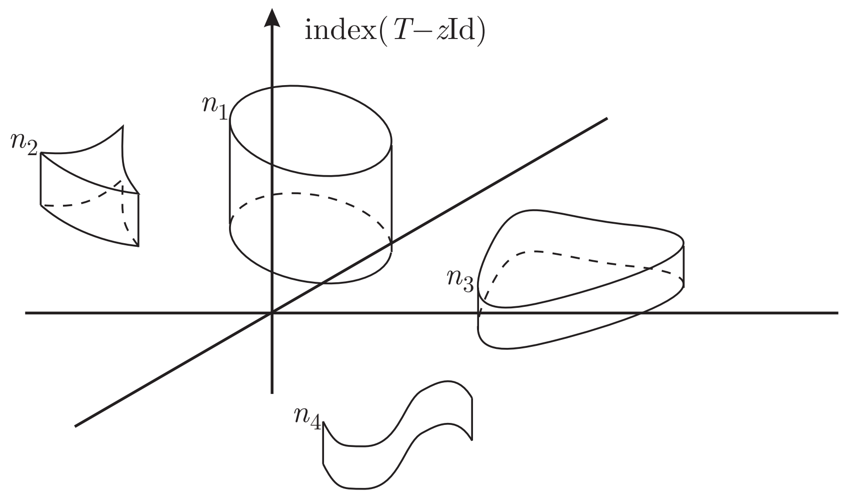


FIGURE 3.4. Index towers on the connected components of the complement of the essential spectrum

unitarily equivalent to Wiener-Hopf operators with the same *characteristic curve*. Details are in the next chapter and in [**Fi**, p.73].

What Is a Spectral Invariant? Let $\mathcal{A}$ be a set of operators (possibly unbounded) on a Hilbert space, and let $\operatorname{spec}(\mathcal{A},\mathcal{A}^*) := \{(\operatorname{Spec}(T), \operatorname{Spec}(T^*)) : T \in \mathcal{A}\}$. Roughly speaking, a **spectral invariant** of $\mathcal{A}$ with values in a set $\mathcal{C}$ (typically $\mathbb{Z}$, $\mathbb{R}$ or $\mathbb{C}$) is a function $\Phi : \operatorname{spec}(\mathcal{A},\mathcal{A}^*) \to \mathcal{C}$ which is preserved (i.e., invariant) under a given set (often a group) $\mathcal{G}$ of transformations $g : \mathcal{A} \to \mathcal{A}$; i.e., $\Phi(g(T)) = \Phi(T)$. We say that a spectral invariant Φ' is *finer* than Φ if $\Phi'(S) = \Phi'(T) \Rightarrow \Phi(S) = \Phi(T)$, but not conversely. One might expect that finer invariants are better, but they may be more difficult to compute, and a courser one may solve the problem at hand.

The main spectral invariant that we have considered thus far is

$$\operatorname{index} : \operatorname{spec}(\mathcal{F},\mathcal{F}^*) \longrightarrow \mathbb{Z}, \text{ given by } \operatorname{index} T = \dim\operatorname{Ker} T - \dim\operatorname{Ker} T^*.$$

This case brings out the point that $\operatorname{spec}(T)$ and $\operatorname{spec}(T^*)$ should include information on the *multiplicities* of eigenvalues of T and T^* (e.g., if $\operatorname{Ker} T \neq \{0\}$ or $\operatorname{Ker} T^* \neq \{0\}$, the multiplicity of the eigenvalue 0). For the index, the set $\mathcal{G}$ could be the additive group $\mathcal{K}$ of compact operators K with $K(T) := T + K$. Exercise 3.10 (p. 65) shows that the index is a spectral invariant of $\mathcal{F}$ under translations by $\mathcal{K}$. More generally, by the homotopy invariance of the index (Theorem 3.11, p.68), the index is a spectral invariant of $\mathcal{F}$ under any set of transformations on $\mathcal{F}$ that map each component of $\mathcal{F}$ into itself.

Note that *any* function $\Phi : \operatorname{spec}(\mathcal{A},\mathcal{A}^*) \to \mathcal{C}$ is a spectral invariant under a set of transformations that leaves $\operatorname{spec} T$ and $\operatorname{spec} T^*$ invariant. For example, this is the case if $\mathcal{G} := \mathcal{B}^\times$, acting via conjugation (i.e., $g(T) = gTg^{-1}$), since

$$gTg^{-1} - \lambda\operatorname{Id} = g(T - \lambda\operatorname{Id})g^{-1} \text{ (for } g \in \mathcal{B}^\times)$$

shows that the resolvent set of gTg^{-1} is the same as that for T (see Definition 2.59). Of course, we may form a semi-direct product $\mathcal{B}^\times \oslash \mathcal{K}$ and let it act on $\mathcal{F}$

via $(g, K) \cdot (T) = gTg^{-1} + K$. Clearly the index is still a spectral invariant on $\mathcal{F}$ under this larger group. The result of Brown, Douglas and Fillmore in 1 above that two essentially unitary operators have the same index if and only if they are unitarily equivalent modulo $\mathcal{K}$, can be interpreted as the statement that there is no spectral invariant for essentially unitary operators which is finer than the index, under the subgroup $\mathcal{U} \oslash \mathcal{K} \subset \mathcal{B}^{\times} \oslash \mathcal{K}$. However, the result in 2 says that this is very far from the case when we enlarge the class of operators from the essentially unitary operators to the essentially normal operators.

In addition to the index, many other spectral invariants have arisen, especially for (unbounded) differential operators such as the Laplace and Dirac operators. Referring to Theorem 2.40 (p. 37), consider the unbounded operator $D_0 = -iT$ in the Hilbert space $L^2(S^1)$, where $T|_{C^0(S^1)}$ denotes differentiation and $i = \sqrt{-1}$. Recall that T, and hence D_0, has dense domain $W^1(S^1) \subset L^2(S^1)$. Moreover, D_0 is self-adjoint by Corollary 2.56, p. 49, and index $D_0 = 0$. For $e_k(\theta) := \frac{1}{\sqrt{2\pi}} e^{ik\theta}$, we have $D_0 e_k = k e_k$ and the standard complete orthonormal system $\{e_k\}_{k \in \mathbb{Z}}$ of $L^2(S^1)$ consists of normalized eigenvectors of D_0 with simple eigenvalues constituting $\operatorname{Spec}(D) = \mathbb{Z}$. The operator D_0 is essentially the so-called *Dirac operator for the circle* with the trivial spin structure. More generally, standard (first-order) Dirac operators can be defined on spinor fields that live on oriented Riemannian n-manifolds M with spin structures, and indeed certain *"twisted" operators of Dirac type* do not require spin structures. We will consider them later in some detail. As with the primordial example D_0 on $L^2(S^1)$, the spectra of operators of Dirac type, say D, over compact spin manifolds have a discrete real spectrum of eigenvalues (not necessarily simple) which is unbounded above and below. If the eigenvalues of D are ordered so that $|\lambda_1| \leq |\lambda_2| \leq \cdots$, then there is some constant C (depending on M) such that (see [**Gi95**, Lemma 1.12.6, p.113])

$$|\lambda_k| \sim C k^{1/n}, \text{ where } n = \dim M.$$

The Eta Function. The **eta function** for D is a $\mathbb{C}$-valued function of $s \in \mathbb{C}$ defined, for $\Re s$ sufficiently large, by

$$\eta_D(s) := \sum\nolimits_{\lambda \in (\operatorname{spec} D) \setminus \{0\}} (\operatorname{sign} \lambda)\, m_\lambda\, |\lambda|^{-s},$$

where m_λ is the multiplicity of the eigenvalue λ. Note that $\eta_D(s)$ is a measure of the spectral asymmetry of $\operatorname{spec} D$ in the sense that if $m_\lambda = m_{-\lambda}$, then $\eta_D(s) = 0$. This is the case for D_0:

$$\eta_{D_0}(s) = \sum\nolimits_{k \in \mathbb{Z} \setminus \{0\}} \frac{k}{|k|} |k|^{-s} = \sum\nolimits_{k=1}^{\infty} (1 - 1) |k|^{-s} = 0.$$

It is known ([**Gi95**, Lemma 1.13.1, p.114]) that $\Gamma((s+1)/2)\, \eta_D(s)$ extends to a meromorphic function (possibly 0) defined on $\mathbb{C}$, all of whose poles (if any) are simple and located at points of the form $(n+1-k)/2$, $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$. Generally, the **reduced eta invariant** of D (not to be confused with the eta function of D or with $\eta_D(0)$, also called the eta-invariant) is defined by

$$\widetilde{\eta}_D := \tfrac{1}{2}\left(\eta_D(0) + \dim \operatorname{Ker}(D)\right) \operatorname{mod} \mathbb{Z}.$$

The reduced eta invariant makes a natural appearance as a boundary term for the Atiyah-Patodi-Singer index formula ([**APS73**]) for operators of Dirac type, with certain boundary conditions, on manifolds of even dimension with boundary. In this case the D in $\widetilde{\eta}_D$ is an induced tangential Dirac operator on the *odd*-dimensional *boundary* of the manifold.

EXAMPLE 3.57. Although $\eta_{D_0}(s) \equiv 0$, consider $D_a := D_0 - a$ for $a \in \mathbb{R}$. If $a \in \mathbb{Z}$, then $\mathrm{Spec}D_a = \mathrm{Spec}D_0$ and $\eta_{D_a}(s) \equiv 0$. Thus, assume that $a \notin \mathbb{Z}$, so that $0 \notin \mathrm{Spec}D_a = \{k - a : k \in \mathbb{Z}\}$. Then

$$\eta_{D_a}(s) = \sum\nolimits_{k\in\mathbb{Z}} \frac{k-a}{|k-a|} \, |k-a|^{-s}, \text{ for } \Re s > 1. \tag{3.22}$$

Note that $\eta_{D_a}(s)$ is periodic of period 1 in the variable a. Also,

$$\begin{aligned}\eta_{D_{-a}}(s) &= \sum\nolimits_{k\in\mathbb{Z}} \frac{k+a}{|k+a|} \, |k+a|^{-s} = \sum\nolimits_{k\in\mathbb{Z}} \frac{-k+a}{|-k+a|} \, |-k+a|^{-s} \\ &= \sum\nolimits_{k\in\mathbb{Z}} \frac{-(k-a)}{|k-a|} \, |k-a|^{-s} = -\eta_{D_a}(s),\end{aligned}$$

whence $\eta_{D_a}(s)$ is odd in a as well. We will find that $\eta_{D_a}(0)$ is defined, but not by the above sum in (3.22). Using $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\, dt$, we have

$$\int_0^\infty t^{(s-1)/2} \lambda e^{-\lambda^2 t}\, dt = \Gamma\left(\tfrac{1}{2}(s+1)\right) \frac{\lambda}{|\lambda|} \, |\lambda|^{-s} \text{ for } \Re s > -1 \text{ and } \lambda \in \mathbb{R}\setminus\{0\}.$$

Indeed, using the change of variable $\tau = \lambda^2 t$,

$$\begin{aligned}&\int_0^\infty t^{(s-1)/2} \lambda e^{-\lambda^2 t}\, dt = \lambda \int_0^\infty \left(\lambda^{-2}\tau\right)^{(s-1)/2} e^{-\tau}\, \lambda^{-2} d\tau \\ &= \lambda^{-1} \int_0^\infty |\lambda|^{1-s} \tau^{(s-1)/2} e^{-\tau} d\tau = \lambda^{-1} \int_0^\infty |\lambda|^{1-s} \tau^{\frac{1}{2}(s+1)-1} e^{-\tau} d\tau \\ &= \Gamma\left(\tfrac{1}{2}(s+1)\right) \frac{\lambda}{|\lambda|} \, |\lambda|^{-s}.\end{aligned}$$

Thus, with $\lambda = k - a \neq 0$,

$$\int_0^\infty t^{(s-1)/2} (k-a)\, e^{-(k-a)^2 t}\, dt = \Gamma\left(\tfrac{1}{2}(s+1)\right) \frac{k-a}{|k-a|} \, |k-a|^{-s},$$

and so for $\Re s > 1$,

$$\begin{aligned}&\Gamma\left(\tfrac{1}{2}(s+1)\right) \eta_{D_a}(s) = \Gamma\left(\tfrac{1}{2}(s+1)\right) \sum\nolimits_{k\in\mathbb{Z}} \frac{k-a}{|k-a|} \, |k-a|^{-s} \\ &= \sum\nolimits_{k\in\mathbb{Z}} \int_0^\infty t^{(s-1)/2} (k-a)\, e^{-(k-a)^2 t}\, dt \\ &= \int_0^\infty t^{(s-1)/2} \sum\nolimits_{k\in\mathbb{Z}} (k-a)\, e^{-(k-a)^2 t}\, dt.\end{aligned}$$

Since the final expression is analytic in s for $\Re s > -1$, it is the analytic continuation of $\Gamma\left(\frac{1}{2}(s+1)\right) \eta_{D_a}(s)$ which was originally defined only for $\Re s > 1$. We claim that for $t > 0$,

$$\sum\nolimits_{k\in\mathbb{Z}} (k-a)\, e^{-(k-a)^2 t} = -2\left(\frac{\pi}{t}\right)^{3/2} \sum\nolimits_{k=1}^\infty k e^{-\frac{\pi^2 k^2}{t}} \sin(2\pi k a).$$

This is a consequence of the Poisson Summation Formula

$$\sum\nolimits_{k\in\mathbb{Z}} f(k) = \sqrt{2\pi} \sum\nolimits_{k\in\mathbb{Z}} \widehat{f}(2\pi k)$$

which holds for rapidly decreasing functions $f \in C_{\downarrow}^{\infty}(\mathbb{R})$ (and under less stringent conditions; see [**BlCs**, p.445]). We apply this to

$$f(x) = (x-a)\, e^{-(x-a)^2 t}, \text{ for which } \widehat{f}(\xi) = \frac{-i\xi}{(2t)^{3/2}} e^{-\xi^2/4t} e^{-ia\xi}.$$

Then

$$\begin{aligned}
&\sum\nolimits_{k\in\mathbb{Z}} (k-a)\, e^{-(k-a)^2 t} \\
&= \sqrt{2\pi} \sum\nolimits_{k\in\mathbb{Z}} \widehat{f}(2\pi k) = \sqrt{2\pi} \sum\nolimits_{k\in\mathbb{Z}} \frac{-i2\pi k}{(2t)^{3/2}} e^{-(2\pi k)^2/4t} e^{-i2\pi a k} \\
&= -i \left(\frac{2\pi}{2t}\right)^{3/2} \sum\nolimits_{k=1}^{\infty} k e^{-\pi^2 k^2/t} \left(e^{-i2\pi a k} - e^{i2\pi a k}\right) \\
&= -2 \left(\frac{\pi}{t}\right)^{3/2} \sum\nolimits_{k=1}^{\infty} k e^{-\pi^2 k^2/t} \sin(2\pi a k).
\end{aligned}$$

Hence, for $1 < \Re s < 2$

$$\begin{aligned}
&\Gamma\left(\tfrac{1}{2}(s+1)\right) \eta_{D_a}(s) = \Gamma\left(\tfrac{1}{2}(s+1)\right) \sum\nolimits_{k\in\mathbb{Z}} \frac{k-a}{|k-a|} |k-a|^{-s} \\
&= \sum\nolimits_{k\in\mathbb{Z}} \int_0^\infty t^{(s-1)/2} (k-a)\, e^{-(k-a)^2 t}\, dt = \int_0^\infty t^{(s-1)/2} \sum\nolimits_{k\in\mathbb{Z}} (k-a)\, e^{-(k-a)^2 t}\, dt \\
&= -2 \int_0^\infty t^{(s-1)/2} \left(\frac{\pi}{t}\right)^{3/2} \sum\nolimits_{k=1}^{\infty} k e^{-\pi^2 k^2/t} \sin(2\pi a k)\, dt \\
&= -2\pi^{\frac{3}{2}} \sum\nolimits_{k=1}^{\infty} k \left(\int_0^\infty t^{\frac{1}{2}s-2} e^{-\pi^2 k^2/t}\, dt\right) \sin(2\pi a k).
\end{aligned}$$

This last expression is analytic for $\Re s < 2$. Thus, it is the analytic continuation of $\Gamma\left(\frac{1}{2}(s+1)\right) \eta_{D_a}(s)$ for $\Re s < 2$. For $s = 0$, we get

$$\begin{aligned}
\sqrt{\pi}\eta_{D_a}(0) &= -2\pi^{\frac{3}{2}} \sum\nolimits_{k=1}^{\infty} k \left(\int_0^\infty t^{-2} e^{-\pi^2 k^2/t}\, dt\right) \sin(2\pi a k) \\
&= -2\pi^{\frac{3}{2}} \sum\nolimits_{k=1}^{\infty} k \frac{1}{\pi^2 k^2} \sin(2\pi a k) \\
&= -\frac{2}{\sqrt{\pi}} \sum\nolimits_{k=1}^{\infty} \frac{1}{k} \sin(2\pi a k).
\end{aligned}$$

Since $4\int_0^{1/2} (2x-1) \sin(2k\pi x)\, dx = \frac{-2}{k\pi}$, we have that $-\frac{2}{\pi}\sum_{k=1}^{\infty} \frac{1}{k} \sin(2\pi x k)$ is the Fourier sine series of $2x-1$ on $\left[0, \frac{1}{2}\right]$. We then have

$$\eta_{D_a}(0) = -\frac{2}{\pi} \sum\nolimits_{k=1}^{\infty} \frac{1}{k} \sin(2\pi a k) = 2a - 1 \text{ (for } 0 < a \le \tfrac{1}{2}).$$

Since $\eta_{D_0}(s) = 0$ and $\eta_{D_a}(s)$ is odd and periodic in a of period 1, $\eta_{D_a}(0) = 2a-1$ for $\frac{1}{2} < a < 1$ and $\eta_{D_1}(0) = 0$. Thus, $\eta_{D_a}(0)$ is the periodic extension (of period 1) of

$$a \mapsto \begin{cases} 0, & \text{for } a = 0 \text{ or } 1, \\ 2a-1, & \text{for } 0 < a < 1. \end{cases}$$

Thus $\eta_{D_a}(0)$ is a discontinuous function with a jump -2 as a crosses each integer. The reduced eta invariant of D_a is then

$$\begin{aligned} \widetilde{\eta}_{D_a} \quad &= \quad \tfrac{1}{2}\left(\eta_{D_a}(0) + \dim \operatorname{Ker}(D_a)\right) \bmod \mathbb{Z} \quad = \quad \tfrac{1}{2}(2a-1) \bmod \mathbb{Z} \\ &= \left(a - \tfrac{1}{2}\right) \bmod \mathbb{Z} \quad = \quad \left(a + \tfrac{1}{2}\right) \bmod \mathbb{Z} \qquad \text{for all } a \in \mathbb{R}. \end{aligned}$$

It is easy to check the relation

$$D_a \quad := \quad -i\frac{d}{dx} - a \quad = \quad e^{ixa} D_0 e^{-ixa} \qquad \text{for all } a \in \mathbb{R}. \tag{3.23}$$

For integer a, (3.23) can be read as a special kind of unitary equivalence between the operators D_a and D_0 with $U_a := e^{ixa}$ unitary operator on $L^2(S^1)$ and $U_a^* = U_{-a}$. Note that U_a and U_a^* keep the domains of the operators D_0 and D_a (namely the first Sobolev space $W^1(S^1)$, defined in(2.20) on page 35) invariant. That explains $\operatorname{Spec} D_a = \operatorname{Spec} D_0$. If $a \in \mathbb{R} \setminus \mathbb{Z}$, the transformation U_a is still unitary with $U_a^* = U_{-a}$, but it does not keep $W^1(S^1) \supset C^0(S^1)$ invariant. Hence we obtain a different spectral situation.

At this place, we shall not re-formulate (3.23) in the language of essential unitary equivalence. We shall turn back to the example later when we discuss the symbolic calculus in Part II, Chapter 7, pages 190ff.

The Zeta Function. In contrast to the eta function, the zeta function is typically defined for certain unbounded operators P, such as Laplacians or squares of Dirac operators, with discrete spectrum which is positive (or more generally in an unbounded wedge containing the positive real axis). (However, see [**ScWo00**] in which zeta functions for operators of Dirac type are defined.) In the context of Laplacians for Riemannian manifolds the zeta function made an early appearance in the seminal paper [**MiPl**] of S. Minakshisundaram and Å. Pleijel. Let $\{\lambda_k\}_{k=1}^{\infty}$, denote the eigenvalues of P with positive real part, ordered so that $0 < \Re\lambda_1 \leq \Re\lambda_2 \leq \cdots$ (repeated according to multiplicity). We define **zeta function of** P by

$$\zeta_P(s) := \sum\nolimits_{k=1}^{\infty} \lambda_k^{-s}.$$

If P is a self-adjoint, elliptic differential operator of order d over a compact n-manifold, this sum converges to an analytic function for $\Re s > n/d$, since $|\lambda_k| \sim Ck^{d/n}$ (see again [**Gi95**, Lemma 1.12.6, p.113]). It turns out that $\zeta_P(s)$ extends to a meromorphic function (still called the zeta function of P and still denoted by $\zeta_P(s)$) on all of $\mathbb{C}$. Assuming further that P is positive semi-definite, all of the poles of $\zeta_P(s)$ are simple and they form a subset of $\{(n-k-1)/d : k \in \mathbb{N}\}$; see [**Gi95**, Theorem 1.12.5, p.112].

Example 3.58. For $P = D_0^2 = -d^2/d\theta^2$, we have eigenvalues k^2, for $k = 0, 1, 2, \ldots$, each of multiplicity 2. Thus, in this case

$$\zeta_{D_0^2}(s) = 2\sum\nolimits_{k=1}^{\infty} k^{-2s} = 2\zeta(2s),$$

where ζ is the well-known Riemann zeta function. Since $\zeta(z)$ is known to be analytic except for a simple pole at $z = 1$ with residue 1, we have that $\zeta_{D_0^2}(s)$ is analytic except for a simple pole at $s = 1/2 = n/d$ with residue 1.

Closely related to the zeta function is the **trace of the heat kernel** for P, namely

$$\operatorname{Tr}\left(e^{-tP}\right) := \sum\nolimits_{k=n_P}^{\infty} e^{-\lambda_k t}.$$

Here the sum is over *all* of the eigenvalues of P, say

$$\lambda_{n_P} \le \lambda_{n_P+1} \le \cdots \le \lambda_1 \le \lambda_2 \le \cdots ,$$

not just the positive ones $\lambda_1 \le \lambda_2 \le \cdots$. If π_+ is the projection onto the closed subspace spanned by the eigenspaces of P with positive eigenvalues $\lambda_1 \le \lambda_2 \le \cdots$, then the so-called **renormalized heat trace** for P is

$$\text{(3.24)} \qquad \operatorname{Tr}\left(e^{-tP\pi_+}\right) = \sum\nolimits_{k=1}^{\infty} e^{-\lambda_k t} = \operatorname{Tr}\left(e^{-tP}\right) - \sum\nolimits_{k=n_P}^{0} e^{-\lambda_k t}.$$

For $\tau = \lambda t$,

$$\int_0^\infty t^{s-1} e^{-\lambda t}\, dt = \int_0^\infty (\tau/\lambda)^{s-1} e^{-\tau}\, \lambda^{-1} d\tau = \lambda^{-s} \int_0^\infty \tau^{s-1} e^{-\tau}\, d\tau = \lambda^{-s}\Gamma(s).$$

Thus, for $\Re s > n/d$,

$$\begin{aligned} \zeta_P(s) &= \sum\nolimits_{k=1}^{\infty} \lambda_k^{-s} = \Gamma(s)^{-1} \sum\nolimits_{k=1}^{\infty} \int_0^\infty t^{s-1} e^{-\lambda_k t}\, dt \\ \text{(3.25)} \qquad &= \Gamma(s)^{-1} \int_0^\infty t^{s-1} \sum\nolimits_{k=1}^{\infty} e^{-\lambda_k t}\, dt = \Gamma(s)^{-1} \int_0^\infty t^{s-1} \operatorname{Tr}\left(e^{-tP\pi_+}\right) dt. \end{aligned}$$

For a suitable function $f(t)$, the function $s \mapsto \int_0^\infty t^{s-1} f(t)\, dt$ is the so-called **Mellin transform** of $f(t)$. Thus, (3.25) says that $\zeta_P(s)$ is the Mellin transform of the renormalized heat trace $\operatorname{Tr}\left(e^{-tP\pi_+}\right)$. The residues of the poles of $\Gamma(s)\zeta_P(s)$ are spectral invariants in that they depend only on the spectrum of P. For any $\varepsilon > 0$,

$$\Gamma(s)\zeta_P(s) = \left(\int_0^\varepsilon t^{s-1} \operatorname{Tr}\left(e^{-tP\pi_+}\right) dt + \int_\varepsilon^\infty t^{s-1} \operatorname{Tr}\left(e^{-tP\pi_+}\right) dt \right).$$

Since the second integral is analytic, the residues of $\Gamma(s)\zeta_P(s)$ only depend on the behavior of $\operatorname{Tr}\left(e^{-tP\pi_+}\right)$ for *small* $t > 0$. As we will do later, at least for certain natural geometric operators P, it is possible to develop an asymptotic expansion

$$\begin{aligned} \text{(3.26)} \quad \operatorname{Tr}\left(e^{-tP}\right) \sim \frac{1}{t^{n/d}} \Big(& a_0(P) + a_1(P) t^{1/d} + \cdots + a_N(P) t^{N/d} \\ & + \mathrm{O}\left(t^{(N+1)/d}\right)\Big) \text{ as } t \to 0^+, \end{aligned}$$

where the $a_k(P)$ are integrals of certain functions on M, which are expressible in terms of the coefficients of P and their derivatives. Note that

$$\sum\nolimits_{k=n_P}^{0} e^{-\lambda_k t} = \sum\nolimits_{m=0}^{\infty} \left(\tfrac{1}{m!} \sum\nolimits_{k=n_P}^{0} (-\lambda_k)^m \right) t^m.$$

By (3.24), the coefficients of the asymptotic expansion (as $t \to 0^+$) of $\operatorname{Tr}\left(e^{-tP\pi_+}\right)$ will have coefficients $\tilde{a}_k(P)$ which generally differ from $a_k(P)$ when $(k-n)/d \in \{0,1,2,\ldots\}$, namely

$$\tilde{a}_k(P) = a_k(P) - \tfrac{1}{m!} \sum\nolimits_{k=n_P}^{0} (-\lambda_k)^m \ \text{ if } m = (k-n)/d \in \{0,1,2,\ldots\}.$$

Note that for $0 \le k \le N$,

$$\int_0^1 t^{s-1} t^{(k-n)/d}\, dt = \int_0^1 t^{s-(n-k)/d-1}\, dt = \frac{1}{s-(n-k)/d} \text{ for } s > (n-N)/d.$$

It follows that the residue of $\Gamma(s)\zeta_P(s)$ at the point $s = (n-k)/d$ is $\tilde{a}_k(P)$; i.e.,

$$\text{(3.27)} \qquad \operatorname{Res}_{s=(n-k)/d}\left(\Gamma(s)\zeta_P(s)\right) = \tilde{a}_k(P),\ k \in \{0,1,2,\ldots\}.$$

It is well-known that about $s=0$, $\Gamma(s)$ has the initial Laurent expansion $\Gamma(s) = s^{-1}+\gamma+\cdots$, where $\gamma = -\lim_{n\to\infty}\left(\log n - \sum_{k=1}^{n}\frac{1}{k}\right) = 0.5772\ldots$ is Euler's constant. Thus, $\zeta_P(s)$ is regular at $s=0$, and

$$\zeta_P(0) = \operatorname{Res}_{s=0}(\Gamma(s)\,\zeta_P(s)) = \tilde{a}_n(P)\,. \tag{3.28}$$

Note that $\tilde{a}_n(P)$ is the term in (3.26) which is t-independent. In the case that P is the square of an operator D of Dirac type on a manifold of even dimension, $\tilde{a}_n(P)$ is the index of the restriction, say D_+, of D to the space of positive spinor fields with the space of negative spinor fields as the codomain (see Chapter 18 in Part IV). Moreover, when $P=D^2$, $\tilde{a}_n(P) = \operatorname{index} D_+$ is a topological invariant. Sometimes (but not always!) $\tilde{a}_n(P)$ is a topological invariant even when P is not of the form D^2, as in the following.

EXAMPLE 3.59. There is a notion of a Laplace operator Δ defined on the space $C^2(M,\mathbb{R})$ of functions on manifold M with a Riemannian metric. In the case of a smooth, compact surface M embedded in $\mathbb{R}^3$ and $f\in C^2(M,\mathbb{R})$, one can define Δf as the restriction to M of the ordinary Laplacian for $\mathbb{R}^3$ of the extension, say $\bar{f}$, of f to a neighborhood of obtained by constantly extending f along line segments normal to M; i.e., $\Delta f := \left(\Delta\bar{f}\right)|_M$. Then $P=-\Delta$ has a discrete spectrum $0=\lambda_0<\lambda_1\leq\lambda_2\leq\ldots$(see [**Gi95**, Lemma 1.6.3]). As $t\to 0^+$, we have the asymptotic expansion

$$\begin{aligned} 1+\operatorname{Tr}\left(e^{-tP\pi_+}\right) &= \operatorname{Tr}\left(e^{-tP}\right) = \operatorname{Tr}\left(e^{t\Delta}\right) = \sum\nolimits_{k=1}^{\infty} e^{-\lambda_k t} \\ &\sim \frac{1}{4\pi t}\left(\int_M dA + \left(\tfrac{1}{3}\int_M K\,dA\right)t + \left(\tfrac{1}{15}\int_M K^2\,dA\right)t^2+\cdots\right), \end{aligned} \tag{3.29}$$

where K is the Gaussian curvature and dA is the element of area (see the seminal paper [**MS**] of H. P. McKean, Jr. and I. M. Singer). Note that $\int_M dA$ is simply the area of M, which is then a spectral invariant of Δ; i.e., two surfaces with the same spectrum for Δ must have the same area. For an eigenfunction u_k with $-\Delta u_k = \lambda_k u_k$, $w(p,t) := \cos\left(\sqrt{\lambda_k}t\right)u_k(p)$ is clearly a solution of the wave equation $w_{tt} = \Delta w$. Hence, the λ_k are proportional to the frequencies $\lambda_k/2\pi$ of possible fundamental harmonic tones emitted from the surface. In this sense, spectral invariants of Δ are quantities that can be *heard*, since they are determined by the set of these tones. In particular, (3.29) implies that the area of M can be heard. We can also hear the total Gaussian curvature $\int_M K\,dA$ whose significance is explained as follows. The Gaussian curvature at $(0,0,0)$ of the surface $z=\frac{1}{2}\left(k_1x^2+k_2y^2\right)$ is k_1k_2, which is negative for a hyperbolic paraboloid (saddle) and positive for an elliptic paraboloid. At an arbitrary point p of a surface M in $\mathbb{R}^3$, K is defined the same way by means of the best quadratic approximation to M in a coordinate system centered at p and adapted to M with the z-axis normal to M at p. The Gauss-Bonnet Theorem asserts that $\int_M K\,dA = 2\pi(2-2g)$, where the so-called **genus** g is the number of holes of M (e.g., $g=1$ for a torus, and $g=0$ for a sphere). Intuitively, the more holes M has, the more negative Gaussian curvature M has. Thus, a torus and a sphere not only look different, but they also sound different, even if they have the same area. Incidentally, $2-2g$ is the **Euler characteristic**

$$\chi(M) = \#faces - \#edges + \#vertices,$$

for a triangulation of M. At any rate, (3.29) tells us that

$$\int_M K\,dA = 2\pi\,(2-2g) = 2\pi\chi(M) \tag{3.30}$$

can be heard. Moreover, $\int_M K^2\,dA$ can be heard, as well as all of the higher order terms in (3.29) which involve derivatives of K. These terms are computable, but with efforts that soon exceed the rewards, especially in higher dimensions.

EXAMPLE 3.60. Let $0 = \lambda_0 < \lambda_1 \le \lambda_2 \le \ldots$. denote the eigenvalues of $-\Delta$ for a compact Riemannian manifold M as in Example 3.59. Let $C_0^\infty(\mathbb{R})$ denote the space of compactly supported, $\mathbb{R}$-valued C^∞functions on $\mathbb{R}$. Let $W : C_0^\infty(\mathbb{R}) \to \mathbb{R}$ denote the linear functional defined by

$$W(f) := \sum\nolimits_{k=0}^{\infty} \int_{-\infty}^{\infty} f(t)\cos(\sqrt{\lambda_k}t)\,dt, \text{ for } f \in C_{\downarrow}^{\infty}(\mathbb{R}).$$

Even though the sum $\sum_{k=0}^{\infty}\cos(\sqrt{\lambda_k}t)$ may not converge, one writes

$$W = \sum\nolimits_{k=0}^{\infty}\cos(\sqrt{\lambda_k}t) \text{ in the distributional sense,}$$

since W is a **distribution** (a continuous linear functional on $C_0^\infty(\mathbb{R})$ with the topology of uniform convergence of each derivative on each compact subset of $\mathbb{R}$). The distribution W has the interpretation as the trace of the **wave kernel** $\operatorname{Tr}\cos\left(t\sqrt{-\Delta}\right)$ as opposed to the heat kernel. We say that $x \in \mathbb{R}$ is in the **singular support** of W (denoted by sing supp W) if there is no open interval I about x and $F \in C^\infty(I)$, such that $W(f) = \int_I F(x)f(x)dx$ for all $f \in C_0^\infty(\mathbb{R})$ with $f|_{\mathbb{R}\setminus I} = 0$. It is clear that $0 \in$ sing supp W. At least for generic M, it has been proven (see [**DG**] for much more) that $(\operatorname{sing\,supp} W)\setminus\{0\}$ is the closure of the **length spectrum** of M which is the set of multiples of lengths of smoothly closed **geodesics** (curves whose sufficiently short subarcs are of minimal length between their fixed endpoints) of M. In other words, at least generically, the closure of the **length spectrum** of M is a spectral invariant.

REMARK 3.61. From the above examples, one may have the impression that the spectrum of Δ for a compact Riemannian manifold contains so much information that it might even determine M up to isometry. The first counterexample was discovered by John Milnor [**Mil64**] who found that the quotients of $\mathbb{R}^{16}$ by the lattices $E_8 \times E_8$ and E_{16} provide two flat tori which are **isospectral** (i.e., have the same spectrum for Δ) but not isometric to each other. Since then a large variety of counterexamples have been found, including one-parameter families of non-isometric isospectral deformations. Moreover, non-isometric isospectral surfaces were first found in 1992 (see [**GWW**]), which led to a negative answer to the query of Mark Kac, "Can one hear the shape of a drum?" (see [**Kac66**]).

For a survey of this and many other related topics, consult Chapter 9 of the truly monumental book [**Berg**] of Marcel Berger. A short list of spectral invariants derived from first coefficients of the asymptotic expansion of the heat kernel can be found in [**Esp98**]. Much longer lists are given in [**Gi04, Gi08**]. Of special interest to mathematicians are the monographs and reviews by E. Elizalde [**El**] and Dmitri Vassilevich and collaborators [**FV, Va**] which discuss main spectral functions appearing in the context of modern physics, in particular quantum field theory.

As will become clear from the asymptotic formula for the heat kernel (elaborated below in Section 18.4, p. 517ff), in general, information about the whole spectrum can not be gained from the heat kernel asymptotics alone but requires insight into the derivatives of the heat kernel and other tools. For the zeta-regularized determinant, this is explained in the following section. For details see also our reviews [**BlBo03**, Section 3.2] and [**Bo02**]. Bauer et al. [**BFI**] give an interesting review of related work on *homogeneous* spaces where new relations for the Hurwitz zeta-function are obtained and representations and characters of the underlying symmetry group enter into the calculations.

The Zeta Regularized Determinant. While a self-adjoint, elliptic differential operator P of order d over a compact n-manifold, with spectrum bounded below, is far from possessing a Fredholm determinant in the sense of Definition 3.44 (p. 98), there is the so-called zeta regularized determinant of P defined and motivated as follows. Note that for $\Re s > n/d$, we have

$$\zeta_P'(s) = \sum\nolimits_{k=1}^{\infty} \tfrac{d}{ds}\left(\lambda_k^{-s}\right) = \sum\nolimits_{k=1}^{\infty} \tfrac{d}{ds} e^{-s\log\lambda_k} = \sum\nolimits_{k=1}^{\infty} -\lambda_k^{-s}\log\lambda_k .$$

If one sets $s = 0$, then the right side becomes the *formal* undefined expression $-\sum_{k=1}^{\infty}\log\lambda_k$ which can formally be rewritten as other undefined expressions:

$$-\sum\nolimits_{k=1}^{\infty}\log\lambda_k = -\log\left(\prod\nolimits_{k=1}^{\infty}\lambda_k\right) = -\log\left(\det P|_{\pi_+ H}\right),$$

where $\pi_+ H$ is the projection of H onto the closure H_+ of the span of the eigenspaces of P with positive eigenvalues. Although as it stands, $-\log\left(\det P|_{H_+}\right)$ is a purely formal meaningless expression, we can *define* $-\log\left(\det P|_{H_+}\right)$ to be $\zeta_P'(0)$, which does exist since the meromorphic extension $\zeta_P(s)$ is regular at $s = 0$. Then the **zeta determinant** of $P|_{H_+}$ is defined to be

$$\det\nolimits_\zeta\left(P|_{H_+}\right) := e^{-\zeta_P'(0)}.$$

This notion of determinant appeared in the 1971 paper [**RaSi71**] of D. Ray and I. M. Singer. If 0 is not an eigenvalue of P, then it is natural to define

$$\det\nolimits_\zeta(P) := \lambda_{n_P}\cdots\lambda_0 e^{-\zeta_P'(0)},$$

where we recall that $\lambda_{n_P} \le \cdots \le \lambda_0$ are the non-positive eigenvalues of P. If 0 is an eigenvalue of P, the consensus seems to be to eliminate it by restricting P to the $(\operatorname{Ker} P)^{\perp}$. Of course, it would be nice to have a way of computing $\zeta_P'(0)$. By (3.25),

$$\int_0^\infty t^{s-1}\operatorname{Tr}\left(e^{-tP\pi_+}\right)dt = \Gamma(s)\,\zeta_P(s) = \left(s^{-1}+\gamma+\cdots\right)\left(\zeta_P(0)+\zeta_P'(0)\,s+\cdots\right)$$
$$= \zeta_P(0)\,s^{-1} + \left(\zeta_P'(0)+\gamma\zeta_P(0)\right) + \cdots .$$

By (3.28), we then have the (rather intractable) formula

$$\zeta_P'(0) = -\gamma\tilde{a}_n(P) + \lim_{s\to 0}\left(\int_0^\infty t^{s-1}\operatorname{Tr}\left(e^{-tP\pi_+}\right)dt - \tilde{a}_n(P)\,s^{-1}\right). \tag{3.31}$$

Although $\zeta_P'(0)$ is not locally computable, as with $a_n(P)$ (or $\tilde{a}_n(P)$ when P is semi-definite), it is a sensitive spectral invariant with important applications not only to quantum physics in relation to anomalies (see [**Nas**, Chapter X]), but it has also been used in other contexts, e.g., to show the compactness of the space of non-isometric compact surfaces with a given spectrum for Δ (see [**OPS**]).

At first sight, Fredholm determinants and determinant line bundles of the previous section and the zeta-regularized determinants discussed here seem to have little in common. It is a very remarkable development arising from the work of researchers such as Jinsung Park, Simon Scott and Krzysztof Wojciechowski (see e.g., [**PaWo02**], [**Sco02**] and [**ScWo00**]) that there are relations between quotients of the respective determinants. While we cannot go into the details here, perhaps the reader can experience the flavor of such relations by simply looking at one of them, say the following formula which is explained and proved in [**Sco02**]:

$$\frac{\det_\zeta\left(\Delta_{P_1}\right)}{\det_\zeta\left(\Delta_{P_2}\right)} = \frac{\det_F\left(S\left(P_1\right)^* S\left(P_1\right)\right)}{\det_F\left(S\left(P_2\right)^* S\left(P_2\right)\right)}.$$

Here (for $i = 1$ or 2), Δ_{P_i} is essentially a Dirac Laplacian (i.e., $D_{P_i}^* D_{P_i}$ where D_{P_i} is an operator of Dirac type on a manifold with boundary), P_i is a suitable boundary condition, and $S\left(P_i\right)^* S\left(P_i\right)$ is a boundary *Laplacian*, involving a generalized scattering operator $S\left(P_i\right)$. Moreover, such formulas have interpretations in the context of determinant line bundles over suitable spaces of boundary conditions.

It was remarked above that usually the finer a spectral invariant is, the more difficult it is to compute. In order of increasing computational difficulty, we generally have: the index, the reduced eta invariant, the eta invariant, and the zeta-determinant which seems to be the most delicate and informative of the four thus far.

CHAPTER 4

Wiener-Hopf Operators

Synopsis. The Reservoir of Examples of Fredholm Operators. Origin and Fundamental Significance of Wiener-Hopf Operators. The *Characteristic Curve* of a Wiener-Hopf Operator. Wiener-Hopf Operators and Harmonic Analysis. The Discrete Index Formula. The Case of Systems. The Continuous Analogue

1. The Reservoir of Examples of Fredholm Operators

We already proved some deep theorems on Fredholm operators, but our supply of examples is still very small, even trivial, as we only studied the following types of Fredholm operators:

(1) The identity operator Id.
(2) The shift operator shift^+ (with respect to an orthonormal basis); see Example 1.3.
(3) The Riesz operators $\text{Id} + K$, where K is an operator with finite rank or, more generally, a Hilbert-Schmidt integral operator of the form

$$(Ku)(x) := \int_X G(x,y)u(y)\,dy$$

with square integrable weight function G; see Exercise 2.29.
(4) The differentiation operator on the first Sobolev space $W^1(S^1)$ and its parametrix; see Theorem 2.40 and Exercise 2.42.

All the other Fredholm operators which appeared so far were elementary function-analytic modifications of the above three basic types. For instance, the left-handed shift is the adjoint of the right-handed shift, i.e., $\text{shift}^- = (\text{shift}^+)^*$. Further, the (unitary) Fourier transformation (Exercise A.2e) $F : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ can be written as a direct sum

$$F = i\,\text{Id} \;\oplus\; (-\,\text{Id}) \;\oplus\; (-i\,\text{Id}) \;\oplus\; \text{Id}$$

by decomposing $L^2(\mathbb{R})$ into a direct sum of four closed subspaces H_1, H_2, H_3, H_4, which are the eigenspaces of F for the eigenvalues i^n. (A proof which explicitly exhibits the eigenfunctions, the *Hermite functions*, can be found in [**DM**, p.97-101].) Viewed in this fashion, from the standpoint of our abstract operator theory on Hilbert space, the Fourier transformation is nothing but a *trivial* modification of the identity.

We will now enlarge our supply of examples by a class of operators which is connected to all four of the basic types (the relation to differentiation were disclosed by Louis Boutet de Monvel in his profound study of local elliptic boundary value problems in [**Bou**]): The Wiener-Hopf operators of the form $\text{Id} + K$, defined on the

Hilbert space $L^2[0,\infty]$, where

$$Ku(x) := \int_0^\infty k(x-y)u(y)\,dy, \text{ for } x \geq 0 \text{ and } k \in L^1(\mathbb{R}).$$

We will give some background information before developing the mathematical theory of these operators and the analogous *discrete* operators

$$S : L^2[\mathbb{Z}_+] \longrightarrow L^2[\mathbb{Z}_+], \text{ where}$$
$$(Su)_n := \sum\nolimits_{k\geq 0} f_{n-k}u_k,$$

whereby the f_n are, for example, the Fourier coefficients of a continuous function $f \in C^0(S^1)$ on the circle S^1.

2. Origin and Fundamental Significance of Wiener-Hopf Operators

Norbert Wiener wrote (1954) in his autobiography:

> "However, the best of the work which he (Eberhard Hopf) and I undertook together concerned a differential equation occurring in the study of the radiation equilibrium of the stars. Inside a star there is a region where electrons and atomic nuclei coexist with light quanta, the material of which radiation is made. Outside the star we have radiation alone, or at least radiation accompanied by a much more diluted form of matter. The various types of particles which form light and matter exist in a sort of balance with one another, which changes abruptly when we pass beyond the surface of the star. It is easy to set up the equations for this equilibrium, but it is not easy to find a general method for the solution of these equations.
>
> The equations for radiation equilibrium in the stars belong to a type now known by Eberhard Hopf's name and mine. They are closely related to other equations which arise when two different physical regimes are joined across a sharp edge or a boundary, as for example in the atomic bomb, which is essentially the model of a star in which the surface of the bomb marks the change between an inner regime and an outer regime; and, accordingly, various important problems concerning the bomb receive their natural expression in Hopf-Wiener equations. The question of the bursting size of the bomb turns out to be one of these.
>
> From my point of view, the most striking use of Hopf-Wiener equations is to be found where the boundary between the two regimes is in time and not in space. One regime represents the state of the world up to a given time and the other regime the state after that time. This is the precisely appropriate tool for certain aspects of the theory of prediction, in which a knowledge of the past is used to determine the future. There are however many more general problems of instrumentation which can be solved by the same technique operating in time. Among these is the wave-filter problem, which consists in taking a message which has been corrupted by a simultaneous noise and reconstructing the pure message to the best of our ability.
>
> Both prediction problems and filtering problems were of importance in the last war and remain of importance in the new technology which

has followed it. Prediction problems came up in the control of anti-aircraft fire, for an anti-aircraft gunner must shoot ahead of his plane as does a duck shooter. Filter problems were of repeated use in radar design, and both filter and prediction problems are important in the modern statistical techniques of meteorology." (N.W.: I am a Mathematician, Victor Follancz Ltd., London, 1956.)

Here we cannot treat all three main areas of application of Wiener-Hopf operators mentioned by Norbert Wiener, namely

(i) the analysis of boundary-value problems
(ii) filter problems in information theory, and
(iii) time series analysis in statistics.

We have to concentrate on the aspect (i) (see Chapters 10 and 11, where we intend to clarify the connection with topological-geometric questions). But it is useful for this purpose to have an idea of the other applications, since it simplifies the transfer of the methods in (ii) and (iii) to our area (i).

3. The Characteristic Curve of a Wiener-Hopf Operator

From information sciences we are interested in the stance taken by electrical engineers: The computation in $\mathbb{C}$ and the Fourier analysis of electric oscillations with the classification of filters (or more generally control circuits) by the geometric shape of the *characteristic curve* as depicted in Figure 4.1. Imagine a *filter* K acting on an input signal u resulting in an output

$$Ku(x) = \int_{-\infty}^{\infty} k(x-y)u(y)\,dy.$$

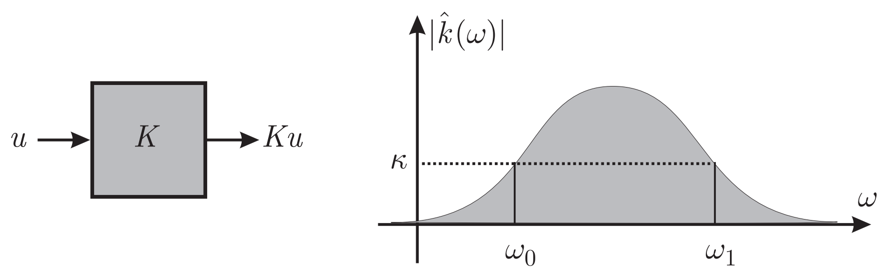


FIGURE 4.1. Scheme of a filter (LEFT) and specifying a transmission region in the corresponding amplitude ratio curve (RIGHT)

Such *linear*, *time independent* and (if $k(x) = 0$ for $x < 0$) *purely past-dependent* filters are good models for many devices of physics and technology. The information scientist *measures* such channels of information by processing a pure sine wave $u(x) = e^{i\omega x}$ through the filter

$$Ku(x) = \int_{-\infty}^{\infty} k(x-y)e^{i\omega y}dy = (\text{with } z = x-y) = -\int_{\infty}^{-\infty} k(z)e^{i\omega(x-z)}dz$$
$$= e^{i\omega x}\int_{-\infty}^{\infty} k(z)e^{-i\omega z}dz = e^{i\omega x}\widehat{k}(\omega),$$

and sketching the *characteristic values* $\widehat{k}(\omega)$ as a function of the *phase* ω or the *frequency* $1/\omega$. Note that here and throughout the rest of this chapter, we define $\widehat{k}(\omega) := \int_{-\infty}^{\infty} k(z)e^{-i\omega z}dz$ without the factor $1/\sqrt{2\pi}$ which would only serve as a distraction in the current context.

The *amplitude ratio* $|\widehat{k}(\omega)|$ is only one measure for the *linear distortion* indicating its *reinforcement* or *weakening*. From it the *transmission region* $[\omega_0, \omega_1]$ may be found via the condition $|\widehat{k}(\omega)| \geq \kappa$. However, a true *harmonic analysis* is achieved only if the *phase shift*, i.e., the argument of the complex number $\widehat{k}(\omega)$, is taken into account. The *nonlinear* distortion is given essentially by the shape of the curve $\{\widehat{k}(\omega) : \omega \in \mathbb{R}\}$, see Figure 4.2. This *characteristic curve*, *filter characteristic* or *periodogram* coincides under certain conditions with the essential spectrum of the operator K; see Theorem 4.11 (p. 129) below.

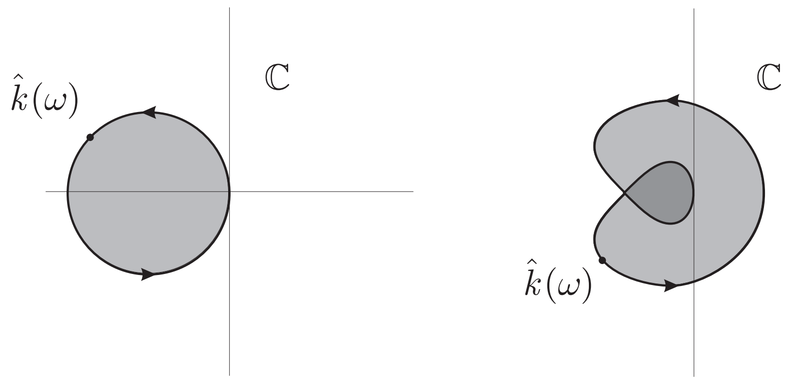


FIGURE 4.2. *Left*: *characteristic curve* of a filter without *feedback*. *Right*: *feedback* is possible in a certain region. Note: the *linear distortion* of both filters may be the same

For details, in particular for the relationship with the general theory of electric circuits we refer to [**DM**, p.170-176] and the literature quoted there.

4. Wiener-Hopf Operators and Harmonic Analysis

We have just pointed out how complex analysis, with its varied geometric-topological aspects, enters markedly into operator theory through information theory. Roughly, the real reason is that the Fourier transform of a square-integrable function k which vanishes identically on the left half-line is holomorphic on the upper half-plane $\mathbb{C}_+$, see Figure 4.3 and [**DM**, p.161f]. Formulated differently, the reason is that the situation of singularities in $\mathbb{C}$ of certain functions associated with dynamical systems carries information about the asymptotic behavior of the oscillating system; see Chapter 10 below and the literature listed there.

The methods of complex analysis thus introduced are based on the idea (founded in the notion of a holomorphic function expandable in a power series) of quantities which vary smoothly and continuously and which are ultimately completely determined through the knowledge of the function value and those of the derivatives at a single point. In contrast, the statistical theory of time series analysis rests on the theory of real functions and thus enters into functional analysis an experience of

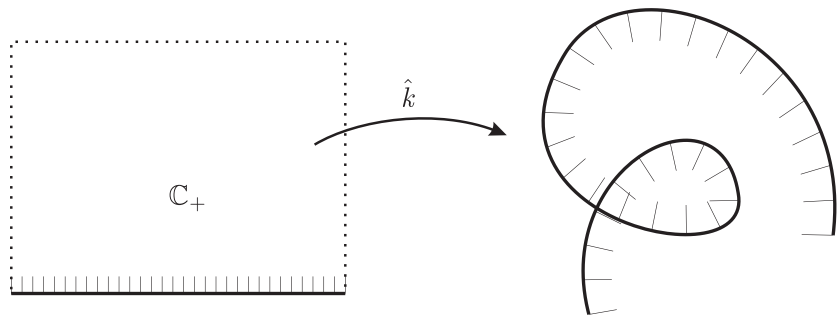


FIGURE 4.3. Fourier transform of a square-integrable function which vanishes identically on the left half-line

dealing mathematically (in the framework of harmonic analysis) with curves which are pieced together from unrelated parts.

With the terminology of the preceding section, we have (roughly) that every operator on the past of $u(x)$ which is linear and invariant under translation of the time origin can be represented as a filter

$$Ku(x) = \int_0^\infty k(y)u(x-y)\,dy$$

or as the limit of a sequence of such operators. If K is defined in this fashion as a linear statistical prediction operator, for example, then the method of *least squares* yields an *optimality criterion* of minimizing

$$\int_{-\infty}^\infty |u\,(x+a) - (Ku)\,(x)|^2\,dx,$$

where a is a given prediction period and the function k which defines K is sought. However, in a statistical theory no statements are made about single occurrence but only about large numbers of such. Correspondingly, the prediction or extrapolation based on a single *time series* u (the determination of k from a single u) does not make any sense. The optimality criterion itself must be interpreted statistically, and the goodness of the operator must be measured not by a single sample but by its average effect. Hence the *stochastic processes* which appear are classified by their *autocorrelation*

$$\phi\,(a) := \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^T u\,(x-a)\,\overline{u\,(x)}\,dx,\ \ a\in\mathbb{R}.$$

When passing from u to the function ϕ, a certain part of the *information content* of the time series u is isolated, while for the rest the specific features of u are ignored. For a class of time series with known autocorrelation ϕ, the optimality criterion can be written as a Wiener-Hopf equation

$$\phi\,(x+a) - \int_0^\infty \phi\,(x-y)\,k\,(y)\,dy = 0,\ \ x\geq 0,$$

where ϕ and a are given and k is sought.

These methods have become standard fare in the statistical time series analysis through the pioneering works [**Kol**, 1943] and [**Wie49**], and can be found in any of the textbooks on statistics and probability theory, frequently under the title *Spectral theory* *of stochastic processes*. A survey with an abundance of examples from economics and technology is provided by [**IL**, 1967], which includes *non-stationary processes* also. The details of these methods are not always interesting from our point of view (computation of the index of Fredholm operators). Conversely, the computation of the index is as a rule uninteresting for correlation theory, since the Wiener-Hopf operators which show up usually have vanishing index, see [**IL**, 1967, p.75]; but also note [**Zy**, 1970, p.147f] who warns about the *illusion of an easy computability of the optimal kernel functions* and points out the *large computational effort necessary for the determination of the correlation functions...* in spite of their uniqueness and explicit solvability in principle. He suggests *adaptive algorithms* as an alternative. These are associated with other types of Fredholm operators, and the uniqueness of the solution is lost. In our context, we want to retain the probabilistic method which roughly consists in forming averages by means of Lebesgue integration, and in compressing and selecting information. The relevant information is then that which (as autocorrelation and the prediction operator itself) yields statements on the kind of connections and transitions between one curve segment (time series) and the next (*transition probabilities*). This is exactly the same strategy that is practical in algebraic topology which investigates, again roughly how geometric structures are composed of simpler *pieces* (see Part III). On this background, the explanation takes shape of why the Wiener-Hopf operators, which originated in boundary value problems of analysis and gained significance in probability theory, more recently turned out to be relevant for the representation of operations in K-theory (see Section 11.5). It is simply because they are (as all Fredholm operators) a functional analytic tool in the treatment of *seams*, *transitions*, and *relations*.

5. The Discrete Index Formula. The Case of Systems

In Appendix A, we become familiar with the Hilbert space $L^2(S^1)$ of measurable, square-integrable functions on the circle, and we cited the fact that the functions $z \mapsto z^n$, $n \in \mathbb{Z}$ form an orthonormal basis for $L^2(S^1)$.

Exercise 4.1. Let H_n denote the subspace of $L^2(S^1)$ spanned by z^k with $k \geq n \in \mathbb{N}$. Show that the functions $z^0, z^1, ..., z^{n-1}$ form a basis of the orthogonal complement $(H_n)^\perp$ of H_n in H_0.

Exercise 4.2. Let P denote the orthogonal projection $L^2(S^1) \to H_0$, and let f be a continuous complex-valued function on the circle; i.e., $f \in C^0(S^1)$.
(a) Show that $T_f := PM_f|_{H_0}$ defines a bounded linear operator on the Hilbert space H_0, where M_f denotes multiplication by f.
(b) Verify that for $u \in H_0$ and $n \in \mathbb{Z}_+$

$$(T_f u)\widehat{\ }(n) = \sum\nolimits_{k=0}^{\infty} \widehat{f}(n-k)\widehat{u}(k),$$

where $\widehat{f}(m) := \langle f, z^m \rangle$ is the m-th Fourier coefficient of f (see Appendix A). T_f is the (discrete) **Wiener-Hopf Operator** assigned to f.

Exercise 4.3. Show that $f \mapsto T_f$ defines a continuous linear map

$$T : C^0(S^1) \longrightarrow \mathcal{B}(H_0),$$

where the Banach algebra $C^0(S^1)$ has norm $\|f\| := \sup\left\{|f(z)| : z \in S^1\right\}$. [Hint: $\|T_f\| \leq \|f\|$. Incidentally, is T a Banach algebra homomorphism; i.e., does it respect the ring structure? See Step 2 in the proof of Theorem 4.4 below.]

THEOREM 4.4 (*Discrete* Gohberg-Krein Index Formula, 1956). *If* $f \in C^0(S^1)$ *and* $f(z) \neq 0$ *for all* $z \in S^1$, *then*
(a) $T_f : H_0 \to H_0$ *is a Fredholm operator,*
(b) $\operatorname{index} T_f = -W(f,0)$. *For the definition of winding number* $W(f,0)$, *see Section 11.1.*

PROOF. We begin with **(a)**.
Step 1: Let $\mathcal{B} := \mathcal{B}(H_0)$ denote the Banach algebra of bounded linear operators on the Hilbert space H_0, and let $\mathcal{K} \subseteq \mathcal{B}$ denote the closed ideal of compact operators on H_0 with $\pi : \mathcal{B} \to \mathcal{B}/\mathcal{K}$ the canonical projection onto the quotient algebra; see Chapter 2 also. From Exercise 4.3, it follows that $\pi \circ T : C^0(S^1) \to \mathcal{B}/\mathcal{K}$ is linear and continuous.
Step 2: Let $C^\vee$ denote the subalgebra of $C^0(S^1)$ consisting of the continuous functions representable by finite Fourier series. Let $f, g \in C^\vee$, say

$$f(z) = \sum\nolimits_{k=-n}^{n} \widehat{f}(k) z^k, \; g(z) = \sum\nolimits_{k=-m}^{m} \widehat{g}(k) z^k,$$

for $n, m \in \mathbb{N}$. In Appendix A, the Fourier coefficients of fg are already calculated:

$$\widehat{fg}\,(j) = \sum\nolimits_{k=-\infty}^{\infty} \widehat{f}(j-k)\widehat{g}(k), \; j \in \mathbb{Z},$$

where the sum is actually taken over only finitely many k. Thus we have (see also Exercise 4.2b)

$$T_f T_g(z^k) = T_{fg}(z^k) \text{ for } k \geq m+n.$$

The operators $T_f T_g$ and T_{fg} coincide on the subspace H_{m+n} of H_0. Since the codimension of H_{m+n} in H_0 is finite ($= m+n$), this means that $T_f T_g - T_{fg}$ is an operator of finite rank, and hence is compact. While T is not a homomorphism of Banach algebras (give a counterexample with $f := ...$ and $g := ...$), by passing to the quotient algebra $\mathcal{B}/\mathcal{K}$, we have

$$\pi T(fg) = \pi T(f) \pi T(g).$$

Thus, $\pi \circ T$ is a homomorphism, when restricted to the subalgebra $C^\vee$.
Step 3: By the Approximation Theorem of Karl Weierstrass (see Chapter A or, for a direct proof, [**DM**, p.49]), each continuous function on a compact interval can be uniformly approximated (i.e., in the sup-norm) by polynomials, and even more so by rational functions. Thus, $C^\vee$ is dense in $C^0(S^1)$. Since $\pi \circ T$ is continuous, the multiplicative property carries over; i.e., $\pi \circ T : C^0(S^1) \to \mathcal{B}/\mathcal{K}$ is a homomorphism of Banach algebras.
Step 4: Since $\pi T(1) = 1$ (where the 1 on the left is the constant function $z \mapsto 1$ and the 1 on the right is the class $\{\operatorname{Id} + K : K \in \mathcal{K}\}$, it follows that $\pi \circ T$ takes invertible functions into invertible elements of $\mathcal{B}/\mathcal{K}$. Hence, if $f(z) \neq 0$ for all $z \in S^1$, then $\pi(T_f)$ is invertible in $\mathcal{B}/\mathcal{K}$, and so T_f is a Fredholm operator by the Theorem of Atkinson (Theorem 3.2, p. 64).
We now prove **(b)**.
We begin with the simplest case, the function $f(z) = z^m$. Relative to the canonical orthonormal basis of H_0 consisting of the functions z^n, $n \in \mathbb{N}$, the Wiener-Hopf operator T_{z^m} assigned to f (Exercise 4.2b) has the form of the one-sided shift operator $(\text{shift}^+)^m$ for $m \geq 0$ and $(\text{shift}^-)^{|m|}$ for $m < 0$. By Exercise 1.3, we then have $\operatorname{index} T_{z^m} = -m$. From the continuity of T (Exercise 4.3) and the continuity (homotopy invariance or local constancy) of the index (see Theorem 3.11, p. 68), it follows from (a) that $\operatorname{index} T_g = -m$ for any $g \in C^0(S^1)$ with values in $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$

which can be connected to the function z^m by a continuous path of functions in $C^0(S^1)$ with values in $\mathbb{C}^\times$. Now, the winding number of the curve $S^1 \to \mathbb{C}$ (defined by z^m) about the point 0 is m. Since curves in $\mathbb{C}^\times$ are homotopic through curves in $\mathbb{C}^\times$ exactly when they have the same winding number (see Section 11.1), we have $\operatorname{index} T_g = -W(g, 0)$, and the index formula is proved. □

EXERCISE 4.5. In the construction of the index bundle (Theorem 3.30), we have seen that for each prescribed orthonormal basis $e_0, e_1, e_2, \ldots$ and Fredholm operator $S \in \mathcal{F}(H_0)$, there is an $n \in \mathbb{N}$ such that

$$P_n S : H_0 \longrightarrow H_n$$

is surjective, where P_n is the orthogonal projection of H_0 onto the closed subspace H_n spanned by the basis elements $e_n, e_{n+1}, e_{n+2}, \ldots$. Now show that (in the case of Wiener-Hopf operators) to each $f \in C^0(S^1)$ with $f(S^1) \subseteq \mathbb{C} \setminus \{0\}$, one can explicitly give an n for which $P_n T_f : H_0 \to H_n$ will be surjective.
[Hint: One naturally exploits the fact that we deal not with an arbitrary Hilbert space, but rather with function spaces, where there is an additional *structure*: Approximate the function $z \mapsto 1/f(z)$ by a finite Fourier series

$$g(z) = \sum_{k=-n}^{n} \widehat{g}(k)\, z^k$$

with n chosen large enough so that $\sup\left\{|f(z)\, g(z) - 1| : z \in S^1\right\} < 1$.]

Theorem 4.4 and Exercise 4.5 demand a detailed topological discussion, in relation to Chapter 1 and in view of Part III. However the families of Wiener-Hopf operators which we will encounter in the following are not of such elementary type. So we need some generalizations.

The first generalization is apparent, if we interpret the Wiener-Hopf operator T_f as a prediction operator for a time series $\ldots, u_{-4}, u_{-3}, u_{-2}, u_{-1}, u_0$ of (say geophysical) measurements, as depicted in Figure 4.4.

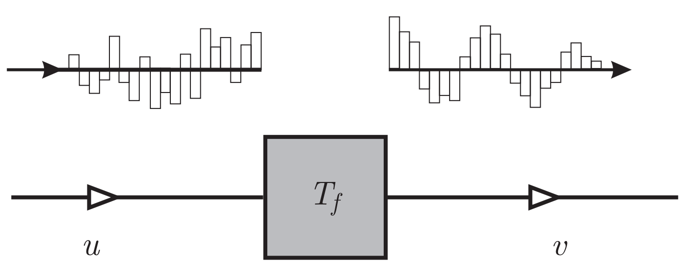


FIGURE 4.4. Interpretation of a Wiener-Hopf operator as a prediction operator

Here, e.g., $v_n = \sum_{k=0}^{\infty} f_{n-k} u_{-k}$, with $u_{-k} = \widehat{u}(k)$ the given time series, $f_{n-k} = \widehat{f}(n-k)$ the *weighting* and $v_n = (T_f u)^\wedge(n)$ the predicted time series.

From the standpoint of the statistician, it is now perfectly obvious (even if one is interested in the weather in Frankfurt exclusively) that the inclusion of additional series of meteorological measurements (from Iceland or the Azores, say) can result in more information than the most sophisticated evaluation of a single series of data (of Frankfurt, for example) could provide. While for a single time series, the *weights* f_{n-k} are numbers, they must be matrices in the statistical analysis of multiple time series. Hence, if we deal with an N-fold time series, the condition $f(z) \neq 0$ which implies the Fredholm property must be replaced by $\det(f(z)) \neq 0$, where $f(z) \in \mathrm{GL}\,(N, \mathbb{C})$.

Exercise 4.6. Let H be a Hilbert space of complex-valued functions (e.g., $H = L^2(S^1)$ or other examples in Appendix A). Show that the well-known notion of tensor product from multilinear algebra for finite-dimensional vector spaces also yields a sensible definition $H \otimes \mathbb{C}^N$. Convince yourself that $H \otimes \mathbb{C}^N$ is again a Hilbert space and (for the concrete examples) is related to the scalar-valued function space H, in such a way that one can regard $H \otimes \mathbb{C}^N$ as being the *corresponding* function space with values in $\mathbb{C}^N$.
[Hint: Compare the analogous considerations in the proof of Theorem 3.40 with regard to the Hilbert space $\mathrm{Hom}(\mathbb{C}^N, H)$ isomorphic to $H \otimes \mathbb{C}^N$. How does one obtain a basis for $H \otimes \mathbb{C}^N$ from bases of H and $\mathbb{C}^N$? Details of the algebraic construction are in [**UN**, 1970, p. 116f], and the peculiarities of infinite-dimensional spaces (which are indeed no problem, when one factor of the tensor product is finite dimensional) are found in [**Do**, 1972, p.31 and 79f].]

Exercise 4.7. For a continuous map $f : S^1 \to \mathrm{GL}(N, \mathbb{C})$, define the Wiener-Hopf operator

$$T_f := PM_f|_{H_0} \otimes \mathbb{C}^N : H_0 \otimes \mathbb{C}^N \longrightarrow H_0 \otimes \mathbb{C}^N,$$

where $P : H \otimes \mathbb{C}^N \to H_0 \otimes \mathbb{C}^N$ is the projection, and M_f is multiplication by the matrix function f. Show:
a) T_f is a Fredholm operator,
b) $\operatorname{index} T_f$ depends only on the homotopy class of f in the homotopy set

$$[S^1, \mathrm{GL}(N, \mathbb{C})].$$

[Hint: Repeat the arguments from Exercises 4.2a and 4.3, and Theorem 4.4. Because of (b), we can identify $\operatorname{index} T_f \in \mathbb{Z}$ with the element $[f]$ in the fundamental group $\pi_1(\mathrm{GL}(N, \mathbb{C})) \cong \mathbb{Z}$ that f represents. Each continuous map of S^1 into $\mathrm{GL}(N, \mathbb{C})$ is homotopic to a continuous map of S^1 into the space of invertible diagonal matrices of rank N. Therefore, set

$$[f] := -W(\det f, 0),$$

where $\det f(z)$ is the determinant of the matrix $f(z)$. See also under Section 11.2.]

Exercise 4.8. In the next generalization, let X be a compact parameter space. Assign to each continuous map $f : S^1 \times X \to \mathrm{GL}(N, \mathbb{C})$ a Fredholm family $T_f : X \to \mathcal{F}$ and also an index bundle $\operatorname{index} T_f \in K(X)$. Show that $\operatorname{index} T_f$ only depends on the homotopy class of f.
[Hint: Note that $f(z, x)$ is an invertible matrix that depends continuously on the variables z and x. Apply Exercise 4.7, noting that we obtain $T_{f(\cdot,x)} \in \mathcal{F}$, for each $x \in X$. Here $\mathcal{F}$ is the space of Fredholm operators on the Hilbert space $H \otimes \mathbb{C}^N$. Show that $T_{f(\cdot,x)}$ depends continuously on x, and then apply the construction from Theorem 3.30, p.84.]

Exercise 4.9. For a further generalization let E be a complex vector bundle over X of fiber dimension N. Figuratively speaking, one allows the vector space $\mathbb{C}^N$ to change from point to point. Given a function $f(z, x) \in \mathrm{Iso}\,(E_x, E_x)$ which

depends continuously on z and x and therefore defines a family of automorphisms of the vector bundle E, construct a family of Fredholm operators (in the variable Hilbert space $H \otimes E$), and finally an index bundle $\operatorname{index} T_f \in K(X)$ that again only depends on the homotopy class of f.
[Hint: See Theorem 3.30, Remark 3.31(p. 86), where we may take the base X to be *sufficiently nice* (e.g., triangulable). Question: Do we really need the Theorem of Kuiper in this Exercise (as in Remark 3.31) or can we proceed directly because of the particular structure of the problem? See [**Ati69**, p.115].]

6. The Continuous Analogue

In connection with local elliptic boundary-value problems (Chapter 10) and topological investigations of the general linear group $GL(N, \mathbb{C})$ (Chapter 11, the Periodicity Theorem of Raoul Bott), we will return to the preceding construction. For the moment, we will only consider the continuous analog of Theorem 4.4:

EXERCISE 4.10. Let $L^1(\mathbb{R})$ denote the space of measurable, absolutely integrable functions. Show that each $\phi \in L^1(\mathbb{R})$ defines a bounded linear operator $K_\phi : L^2(\mathbb{R}_+) \to L^2(\mathbb{R}_+)$, via

$$(K_\phi u)(x) := \int_0^\infty \phi(x-y)u(y)\,dy,\ x \in \mathbb{R}_+.$$

[Hint: Regard $L^2(\mathbb{R}_+)$ as a subspace of $L^2(\mathbb{R})$, and then apply the results of Chapter A on the *convolution*. For detailed estimates, see [**Ti**, 1937, p.90f].]

THEOREM 4.11. *Let $\phi \in L^1(\mathbb{R})$ with $\widehat{\phi}(t) + 1 \neq 0$ for all $t \in \mathbb{R}$ and let K_ϕ be as in Exercise 4.10. Then*

$$\operatorname{Id} + K_\phi : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$$

is a Fredholm operator, and we have

$$\operatorname{index}(\operatorname{Id} + K_\phi) = W(\widehat{\phi} + 1, 0),$$

where $W(\widehat{\phi}+1, 0)$ is the winding number of the oriented curve $t \mapsto \widehat{\phi}(t) + 1$ $(t \in \mathbb{R})$ about the origin (*see Section 11.1*).

REMARK 4.12. In this index formula, one always must be aware of the dependence of the orientation in the definition of the winding number (for us, $W(z, 0) = 1$, for $z(t) = e^{i2\pi t}$, $t \in [0, 1]$) and the orientation in the Fourier transformation (for us, $\widehat{\phi}(x) = \int_{-\infty}^{\infty} e^{-ixy}\phi(y)\,dy$). If one removes the minus sign in the exponent (e.g., as does Mark Krein), then one obtains a minus sign in the index formula.

REMARK 4.13. More exactly, for any $\phi \in L^1(\mathbb{R})$:
(i) $\operatorname{Spec}_e(K_\phi) = \left\{\widehat{\phi}(t) : t \in \mathbb{R}\right\}$
(ii) $\operatorname{index}(z \operatorname{Id} - K_\phi) = W(\widehat{\phi}, z)$ for $z \in \operatorname{Spec}_e(K_\phi)$
(iii) $z \operatorname{Id} - K_\phi$ is $\begin{cases} \text{surjective for index } z \operatorname{Id} - K_\phi \geq 0 \\ \text{injective for index } z \operatorname{Id} - K_\phi \leq 0. \end{cases}$
Proofs for these results discovered by Mark Krein are found in [**Jö**, 1970/1982, 13.4], for example.

REMARK 4.14. If we regard $\operatorname{Id} + K_\phi$ as a map of $L^1(\mathbb{R})$, then under the assumptions of Theorem 4.11, we have that $\operatorname{Id} + K_\phi$ is an isomorphism [**Wie33**]. We then have no index problem.

Proof. Instead of presenting a complete proof, we will comment on the very different ways one can prove Theorem 4.11.

Approach 1: Reduce to Theorem 4.4 with the Cayley transformation $\kappa(z) := \frac{z-i}{z+i}$, which maps the upper half-plane conformally onto the open unit disk, as depicted in Figure 4.5.

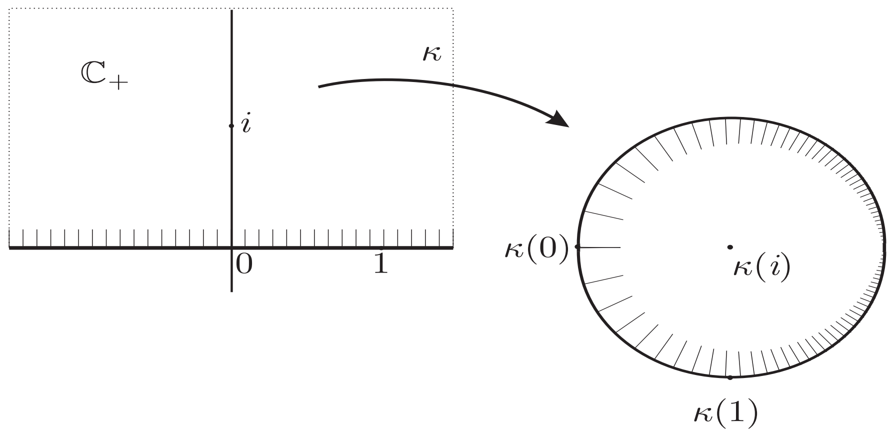


Figure 4.5. The Cayley transform $\kappa : \mathbb{C}_+ \to D^2$

For $v \in L^2(S^1)$,

$$(Uv)(x) := \sqrt{2}\,\frac{v(\kappa(x))}{x+i},\ x \in \mathbb{R}$$

defines an isometry from $L^2(S^1)$ to $L^2(\mathbb{R})$, which carries the Hilbert space

$$H_0(S^1) := \left\{v \in L^2(S^1) : \widehat{v}(n) = 0 \text{ for } n < 0\right\}$$

to the Hilbert space

$$H_0(\mathbb{R}) := \left\{u \in L^2(\mathbb{R}) : \widehat{u}|_{(-\infty,0)} = 0\right\}. \tag{4.1}$$

Equivalently, $H_0(\mathbb{R})$ consists of the square-integrable functions on $\mathbb{R}$ which can be analytically continued to the lower half-plane $\mathbb{C}_-$; e.g., see [**De**, 1967, p.82-84].

Instead of working with the projection $P : L^2(S^1) \to H_0(S^1)$ (see Exercise 4.2), we utilize the corresponding projection $Q : L^2(\mathbb{R}) \to H_0(\mathbb{R})$, where $Q = UPU^{-1}$. To each continuous $\mathbb{C}$-valued function $f \in C^0(\mathbb{R})$ of the form $f = c + \widehat{\phi}$, where $c \in \mathbb{C}$ and $\phi \in L^1(\mathbb{R})$, we assign a (continuous) Wiener-Hopf operator $W_f := Q(M_f)|_{H_0(\mathbb{R})} : H_0(\mathbb{R}) \to H_0(\mathbb{R})$, where M_f means multiplication by f. (In contrast to such *continuous* Wiener-Hopf operators, one often refers to the discrete Wiener-Hopf operators as Toeplitz operators). We now have defined three different operators:

- the discrete Wiener-Hopf operator T_g, $g \in C^0(S^1)$,
- the convolution operator K_ϕ, $\phi \in L^1(\mathbb{R})$, and
- the continuous Wiener-Hopf operator W_f for $f = c + \phi$.

From the properties of U, it then follows [**De**, p.91] that

(i) $T_g = U^{-1} W_f U$, if $f = g \circ \kappa$, and
(ii) $\widehat{W_f}(u) = \widehat{f} * \widehat{u}$, for all $u \in H_0(\mathbb{R})$; i.e.,
$\widehat{W_f}(u) = c\widehat{u}(x) + \int_0^\infty \phi(x-y)\widehat{u}(y)\,dy,\ x \in \mathbb{R}_+, c \in \mathbb{C}$ and $\phi \in L^1(\mathbb{R})$.

Denoting the Fourier transform by $F : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, we can write (ii) as

$$\text{(ii}'\text{)} \; FW_f = (c\,\mathrm{Id} + K_\phi)F, \text{ where } f = c + F\phi.$$

Fact (i) expresses the unitary equivalence of discrete and continuous Wiener-Hopf operators, which is trivial by the definition of W_f here. Fact (ii) requires some caution with the Fourier transformation: In Appendix A, we deal only with the *harmonic analysis* of periodic processes or of processes which (in some sense) abate with increasing or decreasing time. However, in the kinematic and statistical analysis of most natural, physical, technical, economic, etc. processes, the classical machinery is, in fact, not sufficient, since these processes oscillate about some *mean* without being strictly periodic. The formally analogous Fourier analysis requires *functions* on $\mathbb{R}$ which are identically 0 away from a point but are so *strongly infinite* at this one point that the integral over all of $\mathbb{R}$ does not vanish. Physicists, such as Paul Dirac, used this idea in their computations long before Norbert Wiener rigorously proved the necessary generalizations of harmonic analysis [**Wie33**], and which Laurent Schwartz later placed on an even broader foundation with his theory of distributions. In the sense of distributions the Fourier transformation of the constant function 1 is just the Dirac distribution δ at the point 0. See [**Hö63**, p.21f] or [**Schw50**, II, p.11]. Theorem 4.11 follows immediately from Theorem 4.4 with (i) and (ii′).

Approach 2: When Allen Devinatz proved the unitary equivalence between discrete and continuous Wiener-Hopf operators, he showed more than is actually necessary for the proof of Theorem 4.11. Alternatively, one can reduce Theorem 4.11 to Theorem 4.4 in a pedestrian fashion via approximating $f - 1$ by functions which are identically zero outside a bounded interval. If g is such a function, then $g + 1$ can be considered a continuous periodic function, i.e., an element of $C^0(S^1)$. The approximation is done in such a way that $g(z) + 1 \neq 0$ for all $z \in S^1$ and then Theorem 4.4 applies to $g + 1$.

We now proceed as in the passage from Fourier series to Fourier integrals (see Appendix A) whereby the convergence questions must be considered very carefully. For an indication of the computations involved, see e.g. [**GRS**, p.129-132].

Approach 3: One can avoid Theorem 4.4 and basically give a new proof (e.g. [**Kr**, 1958, Theorem 9.2], [**Jö**, 1970/1982 , 13.4], also for *systems* (i.e., matrix valued f, [**GK58**, Theorem 4.1]. These proofs employ the famous *factorization method* introduced by Eberhard Hopf and Norbert Wiener in their original paper (Über eine Klasse singulärer Integralgleichungen. Sitzber. Preuss. Akad. Wiss., Sitzung der phys.-math. Klasse, Berlin 1931, 696-706). Its idea and essential content are presented very comprehensibly in [**Wie49**, p.153-157] (Norman Levinson's heuristic addendum), [**UN**, 1970, p.47-48] (Friedrich Sommer's survey paper on complex analysis) or [**DM**, 1972, 176-184]. The last source contains an altogether good introduction to the function theoretic properties of *Hardy functions* and the elements of our spaces $H_0(S^1)$ and $H_0(\mathbb{R})$.

Approach 4: One can use projection methods of a more general sort which encompass the discrete as well as the continuous case. A detailed exposition is found in [**GF**, 1974]. Here, as in Approach 3, the stress is on finding explicit solutions so that statements about the index enter more frequently in the *opposite direction*, since "the applicability of one or another projection method to the Wiener-Hopf integral equation is determined by the index" (l.c., p. 9l). Complete proofs of our Theorem 4.11, in the manner of these projection methods, are in [**Prö74**, 2.1.5, 2.4.1, 3.2]. Here Theorem 4.11 is not only proved for square integrable functions, but at once for broad varieties of more general function spaces, as do the authors of Approach 3. □

Finally we return once more to the discrete Wiener-Hopf operators whose totality $\left\{T_f : f \in C^0(S^1)\right\}$ we will denote by $\mathcal{T}$ after adjoining the compact operators on H_0.

EXERCISE 4.15. Show that the following is an exact sequence of Banach spaces

$$0 \longrightarrow \mathcal{K}(H_0) \longrightarrow \mathcal{T} \longrightarrow C^0(S^1) \longrightarrow 0,$$

where $\mathcal{K}(H_0)$ denotes the space of compact operators on the Hilbert space $H_0(S^1)$. How are the arrows defined?

[Hint. One best begins with the fact, established in the proof of Theorem 4.4, that the commutator ideal $\left\{T_\phi T_\psi - T_{\phi\psi} : \phi, \psi \in C^0\left(S^1\right)\right\}$ is contained in $\mathcal{K}(H_0)$. Show then that the quotient algebra $\mathcal{T}/\mathcal{K}(H_0)$ is mapped isometrically via T onto $C^0\left(S^1\right)$, where the maps of the short sequence are defined using algebraic generalities. The proof is not entirely simple. One may consult [**Do**, 1972, p.184]. Note the similarity to the exact symbol sequence in the theory of partial differential equations (see Part II below). Also compare it to the *tensorial sequence* in the case of systems [**Do**, 1972, p.202f].]

With these classes of Wiener-Hopf operators, we have greatly enlarged our reservoir of examples. One can even show that, up to *unitary equivalence modulo unitary operators* (see Section 3.10, p. 107), every *essentially normal* operator R on a separable Hilbert space can be written in the form of a Wiener-Hopf operator. More precisely:

1. If the essential spectrum of R has the form of a simple, closed curve (say the image of the circle S^1 under an orientation-preserving, continuous, embedding $\eta : S^1 \to \mathbb{C}$) and $\operatorname{index}(R - z\,\mathrm{Id}) = n$ for z interior to curve, then R is unitarily equivalent to a compact perturbation of the multiplication operator M_η or the discrete Wiener-Hopf operator $T_{\eta\circ\kappa^{-n}}$, where $\kappa(z) := z$ [**Fi**, 1973, p.73].

2. Even when the essential spectrum cannot be parametrized so nicely, classes of generalized Wiener-Hopf operators (namely on *generalized Hardy spaces*, where the domain of holomorphy need not be the upper half-plane or the open unit disk, but may be any bounded region of $\mathbb{C}$) exist from which a *model* of R can be patched together; see [**Fi**, 1973, p.122] and the original papers quoted there.

Part II

Analysis on Manifolds

But, we ask, will not the growth of mathematical knowledge eventually make it impossible for a single researcher to embrace all parts of this knowledge? In answer let me point out how thoroughly, by the very nature of the mathematical sciences, any true progress brings with it the discovery of more incisive tools and simpler methods which at the same time facilitate the understanding of earlier theories and eliminate older more awkward developments. By acquiring these sharper tools and simpler methods the individual researcher succeeds more easily in orienting himself in the different branches of mathematics. In no other science is this possible to the same degree.

David Hilbert, 1900

CHAPTER 5

Partial Differential Equations in Euclidean Space, Revisited

Synopsis. Review of Classical Linear Partial Differential Equations: Constant and Variable Coefficients, Wave Equation, Heat Equation, Laplace Equation, Characteristic Polynomium. Elliptic Differential Equations: Where Do Elliptic Differential Operators Arise? Boundary-Value Conditions. Main Problems of Analysis and the Index Problem. Calculations. Elementary Examples. The Hellwig-Vekua Problem with Non-Vanishing Index

Elliptic operators on sections of complex vector bundles over manifolds provide a primary source of Fredholm operators. In this Part, we explain how this happens. Before that, we recall a bit of general knowledge about elementary geometric aspects of partial differential equations in the plane or in n-dimensional Euclidean space.

1. Linear Partial Differential Equations

The theory of partial differential equations serves the characterization of motions and equilibria *with infinitesimal interactions* and constitutes *the mathematics of all quantities varying in space and time* (Norbert Wiener) which makes up a good part of mathematical physics and of applied mathematics altogether.

We distinguish ordinary and partial differential equations. In ordinary differential equations, the unknown is a function or a system of functions which depend on a single independent variable. In most applications, this variable is time. In partial differential equations, the one or more unknown functions depend on several variables. In applications, these variables are usually the coordinates of a point in space, but one of them may be time. A differential equation expresses relations between measurable quantities and their changes in space and/or time (rates of change). In geometric language (see Figure 5.1), solving an ordinary differential equation means finding a curve and solving a partial differential equation means finding a family of curves or a surface or a manifold of higher dimension, whereby the curvatures of the curves or surfaces must satisfy the conditions expressed by the differential equation.

In this Chapter, we restrict ourselves here to the treatment of linear differential equations of the form

$$Pu = f$$

where u and f are infinitely differentiable complex-valued function on $\mathbb{R}$ and

$$Pu(x) := \sum\nolimits_{\alpha} a_{\alpha}(x)(D^{\alpha}u)(x),\ x \in \mathbb{R}.$$

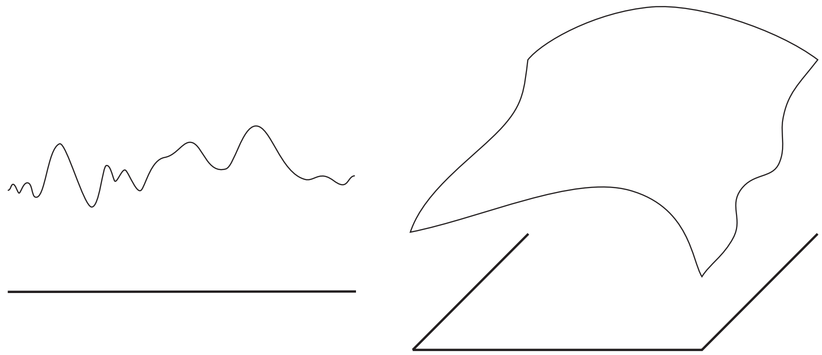


FIGURE 5.1. Finding a curve (ode task, *left*) and finding a family of curves etc. (pde task, *right*)

Here, $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{Z}_+ \times \overset{n \text{ times}}{\cdots} \times \mathbb{Z}_+$ is a multi-index to specify the partial derivative; e.g.

$$D^{(1,0,\dots,0)} := \frac{1}{i}\frac{\partial}{\partial x_1},\ D^{(2,0,\dots,0)} := \left(\frac{1}{i}\right)^2 \frac{\partial^2}{\partial x_1^2},\ \text{and}$$
$$D^\alpha := \left(\frac{1}{i}\right)^{|a|} \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots \partial x_n^{\alpha_n}},\ \text{where}\ |\alpha| := \alpha_1 + \cdots + \alpha_n.$$

REMARK 5.1. It's also convenient to carry the factor of $i^{-|\alpha|}$ when integrating Hermitian inner products by parts. Then integration by parts can be done symmetrically. For example, when $n = 1$,

$$\int_a^b \left(\tfrac{d}{dx} f\right) g = -\int_a^b f\ \overline{\tfrac{d}{dx} g} + \text{boundary terms, while}$$
$$\int_a^b (Df)\, g = \int_a^b f\ \overline{Dg} + \text{boundary terms for } D = \tfrac{1}{i}\tfrac{d}{dx}.$$

In this way we achieve that the differential operators D^α are formally self-adjoint (see Exercise 6.38, p. 181) and yield better expressions under Fourier transformation (note that under Fourier transform the operator $\frac{1}{i}$ is converted into a simple multiplication operator, see Exercise A.2, p. 672). A drawback of the factor is that we have to re-define the D^α and the principal symbol (see (6.23), (6.27) below on p.183ff) for real differential operators of odd order to stay in the real category. But here we follow the notation of main stream analysis, which seems unaware of this drawback, possibly out of a former neglect of operators of first order in their community. Conversely, we emphasize that *real differential operators of first order* (typically of Dirac type) are of interest in index theory, see our Section 14.9 and the literture given there.

The coefficients a_α are always taken to be infinitely differentiable; moreover, $a_\alpha = 0$ for all but finitely many α. P is then called a **differential operator of order** $\max\{|\alpha| : a_\alpha \neq 0\}$. We give the space $C^\infty(\mathbb{R}^n)$ of infinitely differentiable (complex-valued) functions on $\mathbb{R}^n$ the topology defined by the following family of semi-norms (for $k \in \mathbb{Z}_+$, $K \subset \mathbb{R}^n$ compact):

$$\|f\|_{k,K} := \sum_{|\alpha| \leq k} \sup \{|D^\alpha f(x)| : x \in K\}.$$

Accordingly, a sequence $f_1, f_2, \ldots$ of C^∞ functions converges to the constant function 0, if and only if the functions f and all their derivatives converge to 0 uniformly on each compact subset of $\mathbb{R}^n$.

EXERCISE 5.2. Show that a linear (for simplicity, assume scalar) differential operator P is a continuous, linear, and local map $P : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$; here *local* means that

$$\begin{aligned} \operatorname{supp} Pf \ &\subseteq \operatorname{supp} f \qquad \text{for all } f \in C^\infty(\mathbb{R}^n), \text{ where} \\ \operatorname{supp} f &:= \overline{f^{-1}(\mathbb{R} \setminus \{0\})} = \text{the closure of } \{x \in \mathbb{R}^n : f(x) \neq 0\}. \end{aligned}$$

REMARK 5.3. Conversely, one can show that every continuous, linear, local map $P : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is a differential operator (if one allows the order to be infinite and only finite on compact subsets). In fact, for each $x \in \mathbb{R}^n$, the map $f \mapsto (Pf)(x)$ is a continuous linear form on $C^\infty(\mathbb{R}^n)$ with one-point support $\{x\}$, whence [**Schw50**, I, Ch. III, Theorem XXXV] it is a finite linear combination of derivatives (in the distributional sense) of the Dirac δ at x. As x varies in $\mathbb{R}^n$, one can piece together these distributions to obtain the desired differential operator with C^∞ coefficients. The details are in [**CaSc**, 1-03f.]. In 1960, Peetre showed that one can drop the continuity assumption. One can find the completely elementary proof, avoiding distribution theory, in [**Na**, p.172-175].

EXERCISE 5.4. Show that the linear differential operators with coefficients in $C^\infty(\mathbb{R}^n)$ form a non-commutative algebra. Verify that the commutator $PQ - QP$ is a differential operator of order at most $m + m' - 1$, if P has order m and Q has order m'.

EXERCISE 5.5. Show that the space of linear differential operators with constant coefficients forms a commutative subalgebra which is isomorphic to the polynomial algebra $\mathbb{C}[\xi_1, \ldots, \xi_n]$ in the variables $\xi_1, \ldots, \xi_n$.

EXERCISE 5.6. Study the connection between the following partial differential equations appearing most frequently in mathematical physics texts:
a) The **wave equation**

$$\frac{\partial^2 u}{\partial t^2} - a^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} \right) = f(x_1, x_2, x_3, t)$$

is the differential equation for the spreading of vibrations in a homogeneous medium, where the right side vanishes if no force intervenes, and u denotes the displacement (e.g., of a vibrating membrane).
b) The **heat equation** (which governs many other *diffusion processes*) in a homogeneous isotropic body is

$$\frac{\partial u}{\partial t} - a^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} \right) = f(x_1, x_2, x_3, t).$$

Here the right side vanishes when no sources or sinks are present; u is the temperature.

c) The **potential** (or **Poisson**) equation for the potential of an electric field (for example) is

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = -4\pi f(x_1, x_2, x_3),$$

where f is the given charge density and u is the potential whose negative gradient is the electric field.

One can easily classify the (scalar) second order linear differential equations in several independent variables. For the corresponding differential operator

$$P = \sum\nolimits_{|\alpha| \leq 2} a_\alpha D^\alpha$$

and a point $x \in \mathbb{R}$, consider the **characteristic form**

$$\sum\nolimits_{|\alpha|=2} a_\alpha (x)\, \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

which is a quadratic form in $\xi_1, ..., \xi_n$ since $|a| = \sum_{j=1}^n \alpha_j = 2$. In analogy with the classification of conic sections in affine geometry, P is called **elliptic** at the point x, if the form is definite in the sense that

$$\sum\nolimits_{|\alpha|=2} a_\alpha (x)\, \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n} \neq 0 \text{ for } (\xi_1, ..., \xi_n) \in \mathbb{R}^n \setminus \{0\}.$$

In this case, by a change of variable $(\xi_i) \to (\eta_i)$ (not necessarily orthogonal), one can express the form (at x) as

$$\pm \left(\eta_1^2 + \cdots + \eta_n^2\right)$$

We call P **hyperbolic** at x, if the characteristic form can be expressed as

$$\eta_1^2 + \cdots + \eta_{n-1}^2 - \eta_n^2$$

by a change of variables; and P is **parabolic** at x, if we can express the form as

$$\eta_1^2 + \cdots + \eta_{n-1}^2$$

The wave equation is then hyperbolic ($n = 4$), the heat equation is parabolic ($n = 4$), and the potential equation is elliptic ($n = 3$).

2. Elliptic Differential Equations

Roughly speaking the elliptic differential equations of second order differ from the other classical types, in that there is no distinguished coordinate (e.g., *time*). More precisely, if P (of order k) is not elliptic at x_0, then in general (i.e., except in certain degenerate cases which can cause difficulties when using Hamilton-Jacobi methods) there is a function $f \in C^\infty(\mathbb{R}^n)$ with $f(x_0) = 0$ and

$$\sum\nolimits_{|\alpha|=k} a_\alpha (x_0) \left(\frac{\partial f}{\partial x_1}(x_0)\right)^{\alpha_1} \cdots \left(\frac{\partial f}{\partial x_n}(x_0)\right)^{\alpha_n} = 0,$$

where the gradient $\left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)(x_0) \in \mathbb{R} \setminus \{0\}$; i.e., the directional derivatives of f at x_0 do not all vanish. By the Implicit Function Theorem, the set $S := \{x : f(x) = 0)\}$ is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ in a neighborhood of x_0 (see Figure 5.2).

The manifold S is called **characteristic** for the differential operator P at the point x_0. Solutions which are otherwise smooth can have jumps of their second derivatives only along these **characteristic surfaces**. (In physics the characteristic surfaces are possible *wave fronts*.) Furthermore, one obtains from them certain

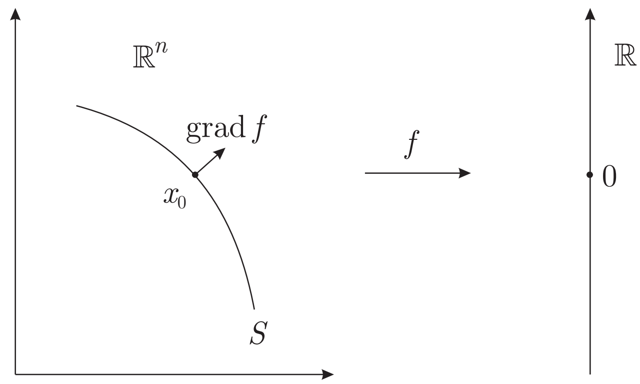


FIGURE 5.2. Distinguished coordinate and characteristic surface at x_0 for non-elliptic differential equation

curves along which (*separation of variables*) the partial differential equation reduces to a simpler differential equation of first order, the so-called *transport equation.* For these reasons the study of characteristic surfaces is a central task in the theory of non-elliptic differential equations.

However, we shall deal with elliptic differential operators which have no (real) characteristic manifolds. The above may well be the reason why, in the theory of elliptic differential equations, it is not initial value problems but boundary value problems and problems on compact curved manifolds involving global questions which are at the center of interest. Slightly exaggerated: Since elliptic operators *look locally the same in all directions* (there are no characteristic manifolds, no *distinguished directions* etc.), and since the local solvability presents no problems (according to [**Hö63**, Theorem 7.2.1] there are no local singularities which could cause trouble globally). Also since there are *sufficiently many* local solutions (e.g., the *large* spaces of harmonic functions for the Laplace operator and holomorphic functions for the Cauchy-Riemann operator), interesting global problems can be formulated immediately and at times solved. We will come back to this philosophy later.

We refer to [**ABP**] for the connection between elliptic equations and parabolic initial value problems, which we will discuss more closely below in Section 18.3 for twisted Dirac operators, see in particular Proposition 18.29, p.510. The solutions of the parabolic heat equation, with arbitrary initial values, solve the potential equation asymptotically. This fact is made the starting point for the heat equation proof of the Atiyah-Singer Index Formula, explained in detail in [**BeGeVe, Gi95, Yosh**].

Until now we have considered only single differential equations (i.e., scalar differential operators). The treatment of simultaneous differential equations (where the interacting unknowns cannot be decoupled) requires the concept of **vectorial differential operators**. These are operators of the form $P = \sum_\alpha a_\alpha(x)D^\alpha$, where (for each $x \in \mathbb{R}^n$), $a_\alpha(x)$ is a linear map from a complex vector space V to a complex vector space W. Relative to bases of V and W, one can regard the $a_\alpha(x)$ as matrices with complex entries. The differential equation $Pu = f$, where u and f are C^∞ vector-valued functions on $\mathbb{R}^n$ ($u(x) \in V \cong \mathbb{C}^N$ and $f(x) \in W \cong \mathbb{C}^M$), can be regarded as a system of M differential equations in N unknown functions.

Then, we have

$$P : C^\infty(\mathbb{R}^n, \mathbb{C}^N) \to C^\infty(\mathbb{R}^n, \mathbb{C}^M),$$

where $C^\infty(\mathbb{R}^n, \mathbb{C}^N)$ denotes the C^∞ functions from $\mathbb{R}^n$ to $\mathbb{C}^N$.

EXERCISE 5.7. To what extent do the previous exercises carry over to vectorial differential operators? A differential operator P of order k is said to be **elliptic**, if (for all $x \in \mathbb{R}^n$ and $(\xi_1, ..., \xi_n) \in \mathbb{R}^n \setminus \{0\}$) the **characteristic polynomial** or the **principal part**

$$\sum\nolimits_{|\alpha|=k} a_\alpha(x)\, \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

is an isomorphism from V to W; in particular, $\dim V = \dim W$.

In the following paragraphs we will investigate the concept of ellipticity more fully; in particular we will work out the geometric meaning of the *principal part.* Here, we give only a few hints for why one is interested in elliptic differential operators and what kinds of related questions come to the forefront.

3. Where Do Elliptic Differential Operators Arise?

Linear elliptic differential operators emerge in many different contexts.

(o) Linear elliptic differential *equations* of order 1 can *not* arise in more than two variables: An $\mathbb{R}$-linear mapping $\mathbb{R}^n \to \mathbb{C} \cong \mathbb{R}^2$, $(\xi_1, \ldots, \xi_n) \mapsto \sum \alpha_j \xi_j$ with complex α_j can not be injective for $n > 2$. That may explain why the main stream of the partial differential equations community was late to show interest in the study of geometric defined differential operators of first order and Dirac type (see our Parts III-IV): "First order?! That's not a challenge." Perhaps, they were right regarding equations, but terribly wrong regarding systems.

(i) Modeling of equilibrium states of oscillating systems. A typical example from mathematical physics is the Laplace equation of potential theory; see Exercise 5.6c above. More complicated problems, require more complicated operators: Operators with variable coefficients which only pointwise resemble Laplace operators (e.g., when the material is not isotropic); operators of higher order; and operators on those function spaces, where the individual functions are not the concrete distributions of a continuous quantity (e.g., temperature), but for instance, the probability amplitudes (*wave functions*) describing a discrete quantum mechanical system consisting of single electrons, atoms or molecules.

(ii) Investigation of classical operators on more complicated geometric surfaces. In analogy with the Laplace operator $\Delta = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_n^2$, one can construct an operator on any Riemannian manifold; see Chapter 6. Various properties of such operators depend on the form of the manifold and serve to classify such surfaces and manifolds to some degree see Chapter 14 below.

(iii) Probabilistic characterization of diffusion processes. In contrast to discrete decay processes with transition probabilities, e.g. on lattices, we deal here with infinitesimal descriptions of flows and other processes, whereby the transition probabilities are given in the form of vector fields. Depending on the model, the random growth, the mean exit time (for problems with boundary), the expectation of some other quantity, etc. appear as solutions of *characteristic operators* which are associated with the *Markov process* via some infinitesimal consideration. Conceptually, imagine a particle which performs a *symmetrical random motion* on the lattice points of $\mathbb{Z}^n$ by moving in equal time intervals one unit to one of the $2n$

neighboring lattice points with transition probability $1/2n$ always, i.e., the transition probability *equidistributed* and *history independent.* If f is a *payoff* function defined on the lattice points, then the expectation of the payoff after one time unit is given by the mean

$$Pf(x) := \frac{1}{2n} \sum\nolimits_{k=1}^{n} \left(f(x+e_k) + f(x-e_k)\right)$$

where the random motion placed the particle one unit ago at the point $x \in \mathbb{Z}^n$, and $e_1, ..., e_n$ are the canonical basis vectors of $\mathbb{R}^n$. The linear operator $P - \mathrm{Id}$ is then a discrete analogue of the operator $\frac{1}{2}\Delta$ in that one can show that the *statistical* operator $P - \mathrm{Id}$ yields the half the Laplace operator, when the distances between lattice points approach zero. The reason is the identity

$$(\Delta f)(x) = \sum\nolimits_{k=1}^{n} \lim_{h \to 0} \frac{1}{h^2} \left(f(x+he_k) - 2f(x) + f(x-he_k)\right)$$

which holds for sufficiently smooth functions. In this fashion, the Laplace operator is linked with the Wiener process which models the random motion of very small particles suspended in some fluid. The Wiener process is characterized probabilistically by the fact that the random change $x(t+s) - x(t)$ of a trajectory x possesses a normal distribution with a particularly simple density function. Other probability distributions yield different characteristic operators, but again elliptic ones if the underlying random process is a *diffusion process.* A very elementary and clear exposition can be found in [**DY**]. Further details are in [**KR**].

(iv) Branching of solutions of nonlinear differential equations. It should be noted that physical, biological, or social systems rarely contain intrinsic justifications for the linearity assumption of mathematical models. The supposition that the effect on a system under study is exactly proportional to the effect contradicts the presence of friction and, more generally, the laws of thermodynamics. Linear models are therefore used exclusively for pragmatic reasons, "either in order to facilitate computation or on account of the present imperfection of engineering techniques of realization (of models) [**Wie49**, p.12]. There are a multitude of situations which unquestionably warrant the use of linear models, for example, in the theory of elasticity of materials, whose deformations are nearly proportional to the forces acting on them, or for many questions of stability theory and of control theory, for which the underlying machinery has been made fairly linear by man. On the other hand, some situations require non-linear modeling, since the essential phenomenon of *branching of solutions* cannot be described in any other way. (Some examples from mechanics are the bending of a straight rod under a constant force, the buckling of a flexible plate, the oscillations of a satellite in its orbital plane, and the surface waves of a heavy fluid.)

These facts in no way render the study of linear models superfluous. Rather it is true that very many nonlinear systems can be approximated by so-called *implicit operators* which are linear, and in many cases also elliptic differential operators. In these cases the *index* of the implicit linear elliptic differential operator plays an important role for the derivation of the *branching equation.* The following example illustrates why the theory of the branching of solutions of a nonlinear equation, with an *analytic variety* as solution manifold, is a natural analogue of the Fredholm theory with affine spaces as solution manifolds. Consider the nonlinear operator $(x, \lambda) \mapsto Tx - \lambda x$ on $H \times \mathbb{R}$ where H is a Hilbert space and T a (linear) compact operator. The solution set $\{(x, \lambda) : Tx - \lambda x = 0\}$ consists of the $\mathbb{R}$-axis $\{0\} \times \mathbb{R}$ and the *kernels* Ker $\{T - \lambda \operatorname{Id}\} \times \{\lambda\}$ of the operators $T - \lambda \operatorname{Id}$, which are Fredholm for $\lambda \neq 0$, depicted in Figure 5.3. Here the jumps of the kernel dimension of $T - \lambda \operatorname{Id}$ (i.e., the eigenvalues of T) are of special interest. See [**VT**, Chs. VII/VIII, esp. Sect. 27] and [**Ize**] for this rapidly developing theory.

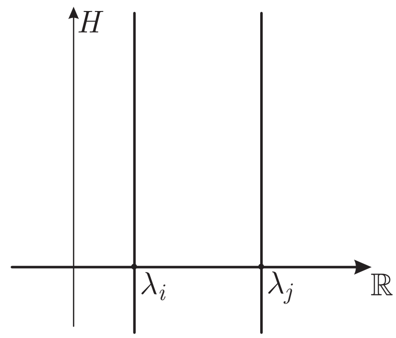


FIGURE 5.3. Analytic variety $\{(x, \lambda) : Tx - \lambda x = 0\}$ as solution manifold of the simple non-linear operator $(x, \lambda) \mapsto Tx - \lambda x$ for $T \in \mathcal{K}(H)$

(v) Problems of optimization theory. Frequently elliptic differential equations are solved by solving the *associated* variational problem, i.e., a problem of optimization. Conversely, many complicated problems of optimization, particularly those occurring in *control theory*, can be reduced to elliptic differential equations and in this way made clearer and more accessible for particular questions. A comprehensive exposition of this aspect can be found in [**Mo**].

(vi) Non-elliptic boundary value problems. Another area of applications is the treatment of systems of non-elliptic differential equations which sometimes can be represented as a family of elliptic differential operators in space coordinates parametrized by time. This is true for instance for the important type of parabolic differential equations which describes a multitude of spacial growth and differentiation processes. Here the connection between parabolic initial value problems and families of elliptic operators is well researched (see above Section 5.2).

4. Boundary-Value Conditions

Notice that in (i), (iii) and (iv) boundary-value conditions play essential role, while in (ii) interesting and deep results can be found considering operators on *closed manifolds* (see Chapter 6 below), thereby avoiding the analytic difficulties of boundary-value problems. We will see below how closely connected boundary-value problems are with problems on closed manifolds. In fact, in the geometric expressions of K-theory, every boundary-value problem on a region X with boundary has a corresponding problem on the boundary ∂X of X and a problem on the *double* $X \cup_{\partial X} X$ of X (see Figure 5.4 and Section 14.8 below). Conversely, elliptic operators over a closed manifold reflect in this fashion how complicated manifolds are built from *macromolecules*, the classical regions with boundary of Euclidean space $\mathbb{R}$.

Warning: Conceptually, the term *boundary-value problems* first brings to mind the boundary-value problems of the theory of elasticity, where an oscillating membrane is held fast along its border. This is mathematically the Dirichlet problem

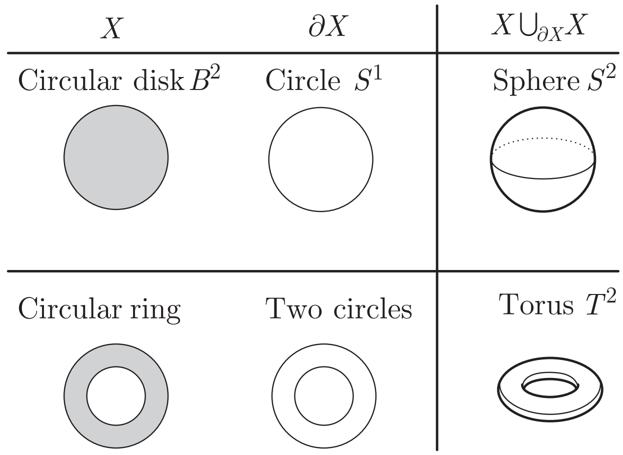


FIGURE 5.4. Correspondence between manifolds with boundary, boundaries, and closed doubles

$u|_{\partial X} = 0$, or more generally $u|_{\partial X} = g$, where g is a function on ∂X. But in many applications, we deal with much more general types of boundary-value conditions. Good examples for all that can occur on the boundary of a region are furnished by the theory of diffusion processes described in (iii). We list just a few of the simplest phenomena following [**DY**, p.137-139], see also Figure 5.5:

(I) **Backward jump** of the particle upon reaching the boundary to a fixed point x inside X, possibly according to a certain probability distribution π generally depending on the boundary point y.

(II) **Absorption**: The particle stays for good at the boundary point first reached.

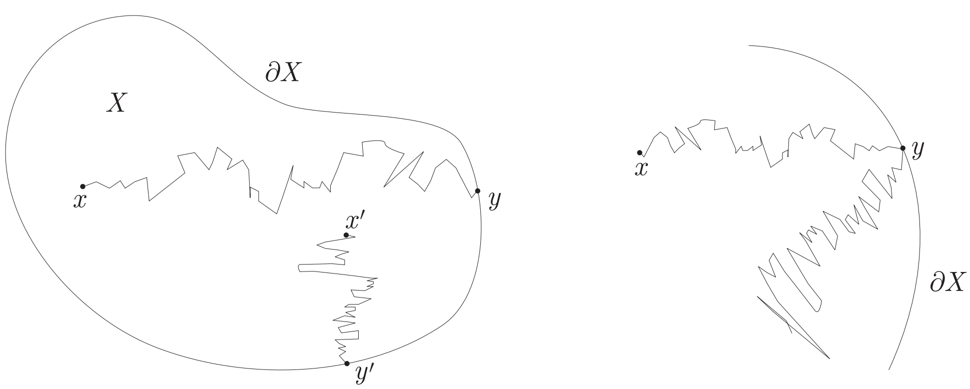


FIGURE 5.5. Boundary occurrences in the theory of diffusion processes: extinction (*left*) and reflection (*right*)

(III) **Extinction**: The particle is annihilated upon first reaching the boundary.
(IV) **Reflection**: Symmetric reflection of the trajectory in the boundary.

For us, these different boundary-value problems only serve as a supply of conceptual examples, and we will not pursue them further. But we wish to stress that it is lastly the investigation and classification of the various boundary-value problems (just like the investigation and classification of various manifolds) that yield the most interesting results. A simple but meaningful example is the Hellwig-Vekua-Theorem (Theorem 5.11, p. 146).

5. Main Problems of Analysis and the Index Problem

Let X be a region in $\mathbb{R}^n$ (or a C^∞ manifold; see below) and a differential operator on X with C^∞ coefficients. Consider the equation $Pu = f$, where u and f are functions (not necessarily C^∞) on X. Somewhat vaguely, we can (following [**Hö71c**]) formulate the following questions:

(i) Under which conditions on P and X can one obtain local or global existence results?

(ii) Given X and P, how are the singularities of u and of f related? Lars Hörmander shows in detail [loc. cit.] that these questions "are in fact so closely related that they can be considered different forms of the same problem." We are interested in the index of elliptic problems that is in questions of type (i). We will show in Part II that, for $Pu = f$ to have a solution at all, every elliptic problem with *suitable* boundary-value conditions must possess an index, i.e., a finite number of linearly independent solutions of the homogeneous equation ($f = 0$) and a finite number of linear conditions for f. In Part III, we will introduce methods for computing the index from the coefficients of P and from numerical invariants of the structure of X, and conversely, for representing topological invariants of manifolds as indices of elliptic operators.

6. Numerical Aspects

> "Much of the modern work in partial differential equations looks highly esoteric, and only a few years ago such work would have been considered of no interest for applications, where one wants a solution expressed in a workable form, say by a sufficiently simple formula. The advent of the modern computing machines has changed this. If a problem involving a differential equation is sufficiently understood theoretically, then, in principle at least, a numerical solution can be obtained on a machine. If the mathematics of the problem is not understood, then the biggest machine and an unlimited number of machine-hours may fail to yield a solution." (COSRIMS Report of the National Science Foundation, 1969).

Sometimes the choice of numerical methods can cleverly be based on previous knowledge of the index. While dealing with Wiener-Hopf operators in Chapter 4, we pointed out such results of I. Z. Gohberg and A. Feldman (after Theorem 4.11, p.129). Similarly, in the numerical treatment of non-linear problems, different methods have been recommended, depending on the index of the associated linear problem ([**VT**], [**Ize**]). Volker Strassen and other mathematicians uncovered the importance of the Theorem of Riemann-Roch (see Section 14.7) and of other quantitative (index-) formulas of algebraic geometry for basic questions of computational complexity (e.g., for the calculation of the computational steps needed for inverting a matrix). Thus, there is an indirect relevance of index calculations on computer oriented numerical mathematics in this setting as well.

7. Elementary Examples

After these general remarks we will work out in detail some elementary examples.

EXERCISE 5.8. Investigate the (trivially elliptic) ordinary differential operator on the unit interval $I = [0, 1]$ with boundary $\partial I = \{0, 1\}$, defined by

$$\begin{array}{rcccc} P & : & C^\infty(I) \times C^\infty(I) & \to & C^\infty(I) \times C^\infty(I), \\ & & (f, g) & \mapsto & (f', -g'), \end{array}$$

with three choices of boundary conditions $C^\infty(I) \times C^\infty(I) \to C^\infty(\partial I) \cong \mathbb{C} \times \mathbb{C}$

$$\begin{aligned} &(i)\ R_1 : (f, g) \mapsto (f - g)\,|_{\partial I} \\ &(ii)\ R_2 : (f, g) \mapsto f|_{\partial I} \\ &(iii)\ R_3 : (f, g) \mapsto (f + g')|_{\partial I}. \end{aligned}$$

Determine the index of the operators (for $i = 1, ..., 3$)

$$(P, R_i) : C^\infty(I) \times C^\infty(I) \to C^\infty(I) \times C^\infty(I) \times C^\infty(\partial I).$$

[Hint: Clearly, dim $\operatorname{Ker}(P, R_i) = 1$. To determine the cokernel, one writes $F, G \in C^\infty(I)$ and $h = (h_0, h_1) \in \mathbb{C} \times \mathbb{C}$, obtaining

$$f(t) = \int_0^t F(\tau) d\tau + c_1, \quad g(t) = -\int_0^t G(\tau) d\tau + c_2$$

and two more equations for the boundary condition. The dimension of $\operatorname{Coker}(P, R_i)$ is then the number of linearly independent conditions on F, G, and h which must be imposed in order to eliminate the constants of integration. For each of the three boundary conditions, check that there is only one condition on the triple (F, G, h), namely $h_0 = h_1 - \int_0^1 F(\tau) d\tau - \int_0^1 G(\tau) d\tau$; $h_0 = h_1 - \int_0^1 F(\tau) d\tau$; resp., $h_0 = h_1 - \int_0^1 F(\tau) d\tau - G(0) + G(1)$. Conclude that the index vanishes in all three cases.]

For a more comprehensive treatment of the existence and uniqueness of boundary-value problems for ordinary differential equations (including systems), we refer to [**CodLev**] and [**Ha**, p.322-403]. Does the index always vanish?

We now consider the Laplace operator $\Delta := \partial^2/\partial x^2 + \partial^2/\partial y^2$, as a linear elliptic differential operator from $C^\infty(X)$ to $C^\infty(X)$, where X is the unit disk $\{z = x + iy : |z| \le 1\} \subset \mathbb{C}$ with boundary $\partial X := \{z \in \mathbb{C} : |z| = 1\}$.

EXERCISE 5.9. For the boundary-value problem (named after Peter Gustav Dirichlet) with boundary condition

$$R : C^\infty(X) \to C^\infty(\partial X), \text{ with } R(u) = u|_{\partial X},$$

show that

$$\begin{aligned} &\text{a) } \operatorname{Ker}(\Delta, R) = \{0\} \text{ and} \\ &\text{b) } \operatorname{Im}(\Delta, R)^\perp = \{0\}, \end{aligned}$$

where $\perp$ is orthogonal complement in $L^2(X) \times L^2(\partial X)$.[1] In particular, it follows that $\operatorname{index}(\Delta, R) = 0$.
[Hint for (a): $\operatorname{Ker}(\Delta, R)$ consists of functions of the form $u + iv$, where u and v are real-valued. Since the coefficients of the operators Δ and R are real, we may assume $v = 0$ without loss of generality. Thus, consider a real solution u with $\Delta u = 0$ in X and $u = 0$ on ∂X. Then (where $\nabla u := (\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y})$)

$$0 = -\int_X u\Delta u\, dxdy = \int_X |\nabla u|^2\, dxdy, \tag{5.1}$$

whence $\nabla u = 0$, noting that u is real. Thus, is constant, and indeed zero since $u = 0$ on ∂X. The trick lies in the equality (5.1), an integration by parts which is perhaps most simply derived from the integral theorem of George Gabriel Stokes in the calculus of differential forms (see Exercise 6.17, p. 169 and [**UN**, p.133f]). Stokes' formula reads $\int_X d\omega = \int_{\partial X} \omega$, where ω is a 1-form. We set $\omega := u \wedge *du$, where $*$ denotes the Hodge star operator (defined here via $*du = *(u_x dx + u_y dy) \overset{!}{=} u_x dy - u_y dx$, again, see Exercise 6.17) and obtain

$$d\omega = du \wedge *du + u \wedge d * du = |\nabla u|^2\, dx \wedge dy + (u\Delta u)\, dx \wedge dy.$$

Using Stokes' formula and $u|_{\partial X} = 0$, we have

$$\int_X |\nabla u|^2\, dxdy + \int_X (u\Delta u)\, dxdy = \int_X d\omega = \int_{\partial X} \omega = \int_{\partial X} u \wedge *du = 0.$$

From this and $\Delta u = 0$, conclude that $\nabla u = 0$ and u is constant.]
[Hint for (b): Choose $L \in C^\infty(X)$ and $l \in C^\infty(\partial X)$ with (L, l) orthogonal to $\operatorname{Im}(\Delta, R)$, whence (relative to the usual measures on X and ∂X)

$$\int_X (\Delta u)\, L + \int_{\partial X} ul = 0 \text{ for all } u \in C^\infty(X). \tag{5.2}$$

Use a 2-fold integration by parts (in the exterior calculus) to obtain

$$\int_X u\Delta L - \int_X (\Delta u)\, L = \int_X (u(d * dL) - d\,(*du)\, L) = \int_{\partial X} (u * dL - L * du)\,. \tag{5.3}$$

First consider u with support $\operatorname{supp}(u) :=$ the closure of $\{z \in X : u(z) \neq 0\}$ contained in the interior of X. Then

$$\int_X u\Delta L = \int_X (\Delta u)\, L = -\int_{\partial X} ul = 0,$$

and so $\Delta L = 0$. Now for $u \in C^\infty(X)$ apply (5.2) and (5.3) to deduce that

$$\begin{aligned}\int_{\partial X} ul = -\int_X (\Delta u)\, L &= \int_{\partial X} (u * dL - L * du) \\ &= \int_{\partial X} \left(u\left(x\tfrac{\partial L}{\partial x} + y\tfrac{\partial L}{\partial y}\right) - L\left(x\tfrac{\partial u}{\partial x} + y\tfrac{\partial u}{\partial y}\right)\right).\end{aligned}$$

Conclude that $l = x\frac{\partial L}{\partial x} + y\frac{\partial L}{\partial y}$ and $L|_{\partial X} = 0$, and finally apply (a). Details are in [**Hö63**, p.264].]

[1] Here, consider that the intersection of the orthogonal complement of $\operatorname{Im}(\Delta, R)$ relative to the usual inner product in $L^2(X) \times L^2(\partial X)$ with the space $C^\infty(X) \times C^\infty(\partial X)$ is isomorphic to $\operatorname{Coker}(\Delta, R)$. This is true, since the image of the natural Sobolev extension of (Δ, R) is closed in the L^2-norm, and its L^2-orthogonal complement is contained in $C^\infty(X) \times C^\infty(\partial X)$.

REMARK 5.10. The preceding result $\operatorname{index}(\Delta, R) = 0$ (for $Ru = u|_{\partial X}$) can also be obtained by proving the symmetry of Δ and that the L^2 extension on the domain defined by $Ru = 0$ is a self-adjoint Fredholm extension.

We now consider a C^∞ vector field $\nu : \partial X \to \mathbb{C}$ on the boundary $\partial X = \{z : |z| = 1\}$. For $u \in C^\infty(X)$, $z \in \partial X$, and $\nu(z) = \alpha(z) + i\beta(z)$, one defines the *directional derivative* of the function u relative to the vector field ν at the point z to be the number

$$\tfrac{\partial u}{\partial \nu}(z) := \alpha(z)\tfrac{\partial u}{\partial x}(z) + \beta(z)\tfrac{\partial u}{\partial y}(z)\,.$$

From the standpoint of differential geometry it is better either to denote the vector field by $\frac{\partial}{\partial \nu}$ or to write the directional derivative as simply as $\nu\,[u]\,(z)$, since $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ can be regarded as vector fields; see Chapter 6 below. The pair $(\Delta, \frac{\partial}{\partial \nu})$ defines a linear operator

$$(\Delta, \tfrac{\partial}{\partial \nu}) : C^\infty(X) \to C^\infty(X) \oplus C^\infty(\partial X) \text{ given by } u \mapsto (\Delta u, \tfrac{\partial u}{\partial \nu}).$$

THEOREM 5.11 (G. Hellwig, I. N. Vekua, both 1952). *For $p \in \mathbb{Z}$ and $\nu(z) := z^p$ as depicted in Figure 5.6, we have that $(\Delta, \frac{\partial}{\partial \nu})$ is an operator with finite-dimensional kernel and cokernel, and*

$$\operatorname{index}(\Delta, \tfrac{\partial}{\partial \nu}) = 2\,(1-p)\,.$$

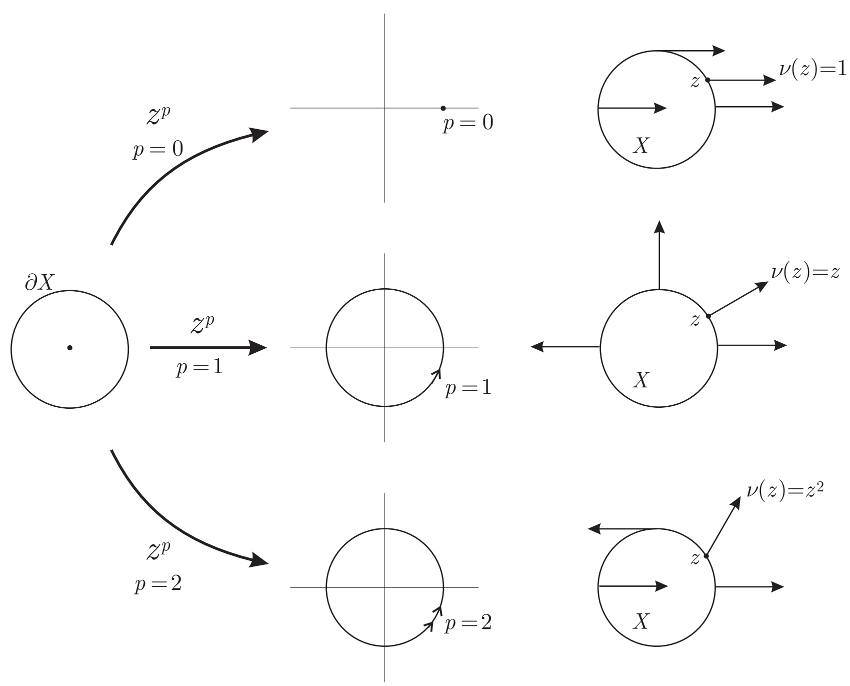


FIGURE 5.6. The vector field $\nu : \partial X \to \mathbb{C}$ with winding number $p = 0, 1, 2$

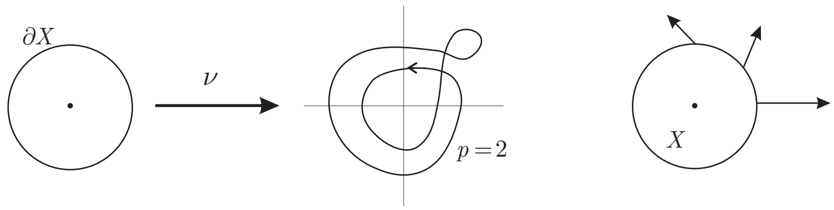


FIGURE 5.7. Another vector field ν with winding number 2

REMARK 5.12. The theorem of G. Hellwig and I.N. Vekua remains true, if we replace z^p by any nonvanishing *vector field* $\nu : \partial X \to \mathbb{C} \setminus \{0\}$ with *winding number* p as in Figure 5.7.

Moreover, in place of the disk, we can take X to be any simply connected (i.e., without *holes*) domain in $\mathbb{C}$ with a *smooth* boundary ∂X; see Chapter 6 below. The reason is the homotopy invariance of the index (see Theorem 3.11, p. 68) which holds for elliptic differential operators on closed manifolds and on compact manifolds with smooth boundary when admissable boundary conditions are imposed (see Section 14.8, p.323 below) .

REMARK 5.13. One encounters the number $2(1-p)$ also in the theory of Riemann surfaces of genus p; e.g., as the Euler characteristic of a closed surface or, deeper, in the theorem of Bernhard Riemann and Gustav Roch (see Section 14.7, p.320). This is no accident, but rather it is connected with the relation between elliptic boundary-value problems and elliptic operators on closed manifolds, as mentioned above on p. 139. Specifically, there is a relation between the index of $(\Delta, \frac{\partial}{\partial \nu})$ and the index of the Cauchy-Riemann operator for complex line bundles over $S^2 = \mathbb{P}^1(\mathbb{C})$ with Chern number $1-p$ (e.g., see Example 5.16 for a start).

REMARK 5.14. Motivated by the method of replacing a differential equation by difference equations, David Hilbert and Richard Courant expected "linear problems of mathematical physics which are correctly posed to behave like a system of N linear algebraic equations in N unknowns... If for a correctly posed problem in linear differential equations the corresponding homogeneous problem possesses only the trivial solution zero, then a uniquely determined solution of the general inhomogeneous system exists. However, if the homogeneous problem has a non-trivial solution, the solvability of the non-homogeneous system requires the fulfillment of certain additional conditions." This is the *heuristic principle* which [**CH**, II, p. 179/231] saw in the Fredholm Alternative (see Chapter 2). Günter Hellwig [**Hel**] (nicely explained in [**Haa**]) in the real setting and Ilya Nestorovich Vekua [**Ve56**] in complex setting disproved it with their independently found example where the principle fails for $p \neq 1$.
We remark that in addition to these *oblique-angle* boundary-value problems, *coupled* oscillation equations, as well as restrictions of boundary-value problems, even with vanishing index, to suitable half-spaces, furnish further elementary examples for index $\neq 0$. The simplest example of a system of first order differential operators on the disc is provided in Exercise 5.16, p. 151 below. A world of more advanced, and for differential geometry much more meaningful examples, is approached by the Atiyah–Patodi–Singer Index Theorem, see Section 14.8 below.

Proof of Theorem 5.11. (after [**Hö63**, p.266f]): Since the coefficients of the differential operators $(\Delta, \frac{\partial}{\partial\nu})$ are real, we may restrict ourselves to real functions. Thus, $u \in C^\infty(X)$ denotes a single real-valued function, rather than a complex-valued function (i.e., a pair $u_1 + iu_2$ of real-valued functions u_1 and u_2). **Ad** $\mathrm{Ker}(\Delta, \frac{\partial}{\partial\nu})$: It is well-known that $\mathrm{Ker}(\Delta)$ consists of real (or imaginary) parts of holomorphic functions on X (e.g., see [**Ah**, p.175f]). Such functions are called harmonic. Hence, $u \in \mathrm{Ker}(\Delta)$, exactly when $u = \Re(f)$ where $f = u + iv$ is holomorphic; i.e., the Cauchy-Riemann equation $\frac{\partial f}{\partial \bar z} = 0$ holds, where $\frac{\partial}{\partial \bar z} := \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$. Explicitly,

$$0 = \frac{\partial f}{\partial \bar z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(u + iv) = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right).$$

Every holomorphic (= complex differentiable) function f is twice complex differentiable and its derivative is given by

$$\begin{aligned}\frac{\partial f}{\partial z} &:= \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u + iv)\\ &= \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + i\left(-\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.\end{aligned}$$

In this way we have a holomorphic function $\phi := f'$ for each $u \in \mathrm{Ker}(\Delta)$. Since

$$\frac{\partial u}{\partial \nu} = \Re\left(z^p\right)\frac{\partial u}{\partial x} + \mathrm{Im}\left(z^p\right)\frac{\partial u}{\partial y} = \Re\left(\left(\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}\right)z^p\right) = \Re\left(\phi\left(z\right)z^p\right),$$

the boundary condition $\frac{\partial u}{\partial \nu} = 0$ $(\nu = z^p)$ then means that the real part $\Re(\phi(z)z^p)$ vanishes for $|z| = 1$. For $p > 0$, $\phi(z)z^p$ is holomorphic as well as ϕ, and hence for $\phi(z) := \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}$, we have

$$\begin{aligned}&u \in \mathrm{Ker}\left(\Delta, \tfrac{\partial}{\partial\nu}\right) \text{ with } \nu := z^p,\ p \geq 0\\ &\Rightarrow \Re\left(\phi(z)z^p\right) \in \mathrm{Ker}\left(\Delta, R\right) \text{ where } R\left(\cdot\right) := \left(\cdot\right)|_{\partial X}.\end{aligned}$$

Thus, we have associated the *oblique-angle* boundary-value problem for u with a Dirichlet boundary-value problem for $\Re\left(\phi(z)z^p\right)$, which has only the trivial solution by Exercise 5.9a. Since $\phi(z)z^p$ is holomorphic with $\Re\left(\phi(z)z^p\right) = 0$, the partial derivatives of the imaginary part vanish, and so there is a constant $C \in \mathbb{R}$ such that $\phi(z)z^p = iC$ for all $z \in X$. If $p > 0$, then we have $C = 0$ (set $z = 0$). Hence $\phi = 0$, and (by the definition of ϕ) the function u is constant (i.e., $\dim\mathrm{Ker}\left(\Delta, \frac{\partial}{\partial\nu}\right) = 1$). If $p = 0$, then $\phi(z) = iC$; and so $u(x, y) = -Cy + \widetilde{C}$, whence $\dim\mathrm{Ker}\left(\Delta, \frac{\partial}{\partial\nu}\right) = 2$ in this case.

We now come to the case $p < 0$, which curiously is not immediately reducible to the case $q > 0$ where $q := -p$. One can try to look for a solution by simply turning $\frac{\partial}{\partial\nu}$ around to $-\frac{\partial}{\partial\nu}$ as illustrated in Figure 5.8. However, this is futile since the winding numbers of ν and $-\nu$ about 0 are the same. Besides, if $\nu_p(z) = z^p$, we do *not* have $\frac{\partial}{\partial\nu_{-p}} = -\frac{\partial}{\partial\nu_p}$. In order to reduce the boundary-value problem with $p < 0$ to the elementary Dirichlet problem, we must now go through a more careful argument. Note that $\phi(z)z^p$ can have a pole at $z = 0$, whence $\Re\left(\phi(z)z^p\right)$ is not necessarily harmonic. We write the holomorphic function $\phi(z)$ as a finite Taylor series

$$\phi(z) = \sum\nolimits_{j=0}^{q} a_j z^j + g\left(z\right)z^{q+1}$$

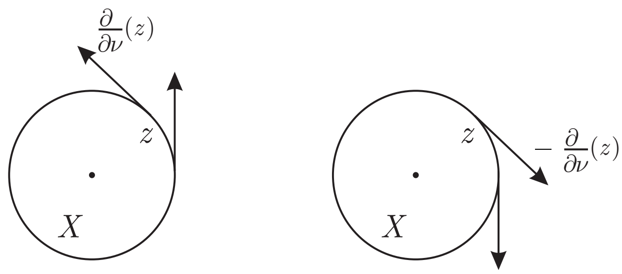


FIGURE 5.8. Replacing ν by $-\nu$ does not change the winding number

where $q := -p$ and g is holomorphic. We define a holomorphic function ψ by

$$\psi(z) := g(z)\, z + \sum_{j=0}^{q-1} \bar{a}_j z^{q-j},$$

with $\psi(0) = 0$. Then one can write

$$\phi(z)z^p = a_q + \sum_{j=0}^{q-1} \left(a_j z^{j-q} - \bar{a}_j z^{q-j}\right) + \psi(z). \tag{5.4}$$

The boundary condition $\frac{\partial u}{\partial \nu} = 0$ implies $\Re\left(\phi(z)z^p\right) = 0$ for $|z| = 1$. By (5.4), we have $0 = \Re\left(\phi(z)z^p\right) = \Re(\psi(z) + a_q)$ for $|z| = 1$ since then $z^{-1} = \bar{z}$. Since ψ is holomorphic, we have again arrived at a Dirichlet boundary value problem; this time for the function $\Re(\psi(z) + a_q)$. From Exercise 5.9a, it follows again that $\psi(z) + a_q$ is an imaginary constant, whence $\psi(z) = \psi(0) = 0$ and a_q is pure imaginary. We have

$$\phi(z) = \phi(z)z^p z^q = a_q z^q + \sum_{j=0}^{q-1} \left(a_j z^j - \bar{a}_j z^{2q-j}\right)$$

for arbitrary $a_0, a_1, ..., a_{q-1} \in \mathbb{C}$ and $a_q \in i\mathbb{R}$. As a vector space over $\mathbb{R}$, the set

$$\left\{ \tfrac{\partial u}{\partial x} - i \tfrac{\partial u}{\partial y} : u \in \operatorname{Ker}\left(\Delta, \tfrac{\partial}{\partial \nu}\right) \right\}$$

has dimension $2q + 1$; here we have restricted ourselves to real u, according to our convention above. Since u is uniquely determined by ϕ up to an additive constant, it follows that for $\nu = z^p$ and $p < 0$,

$$\dim \operatorname{Ker}\left(\Delta, \tfrac{\partial}{\partial \nu}\right) = 2q + 2 = 2 - 2p.$$

Ad Coker $\left(\Delta, \frac{\partial}{\partial \nu}\right)$: As Exercise 5.9b shows, the equation $\Delta u = F$ has a solution for each $F \in C^\infty(X)$. In view of this we can show

$$\operatorname{Coker}\left(\Delta, \tfrac{\partial}{\partial \nu}\right) = \frac{C^\infty(X) \times C^\infty(\partial X)}{\operatorname{Im}\left(\Delta, \frac{\partial}{\partial \nu}\right)} \cong \frac{C^\infty(\partial X)}{\frac{\partial}{\partial \nu}\left(\operatorname{Ker}\Delta\right)}$$

as follows. We assign to each representative pair $(F, h) \in C^\infty(X) \times C^\infty(\partial X)$ the class of $h - \frac{\partial u}{\partial \nu} \in C^\infty(\partial X)$, where u is chosen so that $\Delta u = F$. This map is clearly well defined on the quotient space of pairs, and the inverse map is given by $h \mapsto (0, h)$. Hence, we have found a representation for Coker$(\Delta, \frac{\partial}{\partial \nu})$ in terms of the *boundary functions* $\left\{\frac{\partial u}{\partial \nu} : u \in \operatorname{Ker}\Delta\right\}$, rather than the cumbersome pairs in Im $\left(\Delta, \frac{\partial}{\partial \nu}\right)$. (This trick can always be applied for the boundary-value problems (P, R), when the operator P is surjective.)

We therefore investigate the existence of solutions of the equation $\Delta u = 0$ with the *inhomogeneous* boundary condition $\frac{\partial u}{\partial \nu} = h$, where h is a given C^∞ function on

∂X. According to the trick introduced in the first part of our proof, it is equivalent to ask for the existence of a holomorphic function ϕ with the boundary condition $\Re(\phi(z)z^p) = h$, $|z| = 1$, i.e., for a solution of a Dirichlet problem for $\Re(\phi(z)z^p)$. By Exercise 5.9b, there is a unique harmonic function which restricts to h on the boundary ∂X; hence, we have a (unique up to an additive imaginary constant) holomorphic function θ with $\Re\theta(z) = h$ for $|z| = 1$.

In the case $p \leq 0$, the boundary problem for ϕ is always solvable; namely, set $\phi(z) := z^{-p}\theta(z)$. Hence, we have

$$\dim \operatorname{Coker}\left(\Delta, \tfrac{\partial}{\partial\nu}\right) = 0 \text{ for } \nu(z) = z^p \text{ and } p \leq 0.$$

For $p > 0$, we can construct a solution of the boundary-value problem for ϕ from θ, if and only if there is a constant $C \in \mathbb{R}$, such that $(\theta(z) - iC)/z^p$ is holomorphic (i.e., the holomorphic function $\theta(z) - iC$ has a zero of order at least p at $z = 0$). Using the Cauchy Integral Formula, these conditions on the derivatives of θ at $z = 0$ correspond to conditions on line integrals around ∂X. In this way, we have $2p - 1$ linear (real) equations that h must satisfy in order that the boundary-value problem have a solution. We summarize our results in Table 5.1 ($\nu(z) = z^p$) and Figure 5.9. □

TABLE 5.1. Dimensions of kernel and cokernel for varying p

p	$\dim \operatorname{Ker}\left(\Delta, \frac{\partial}{\partial\nu}\right)$	$\dim \operatorname{Coker}\left(\Delta, \frac{\partial}{\partial\nu}\right)$	$\operatorname{index}\left(\Delta, \frac{\partial}{\partial\nu}\right)$
> 0	1	$2p - 1$	$2 - 2p$
≤ 0	$2 - 2p$	0	$2 - 2p$

Warning 1: We already noted in the proof the peculiarity that the case $p < 0$ cannot simply be played back to the case $p > 0$. This is reflected here in the asymmetry of the dimensions of kernel and cokernel and the index. It simply reflects the fact that there are *more* rational functions with prescribed poles than there are polynomials with *corresponding* zeros. See also Section 14.7, the Riemann-Roch Theorem.

Warning 2: In contrast to the Dirichlet Problem, which we could solve via integration by parts (i.e., via Stokes' Theorem), the above proof is function-theoretic in nature and cannot be used in higher dimensions. This is no loss in our special case, since the index of the *oblique-angle* boundary-value problem must vanish anyhow in higher dimensions for topological reasons; see [**Hö63**, p.265f] or Section 14.8 below. The actual mathematical challenge of the function-theoretic proof arises less from the restriction $\dim X = 2$ than from a certain arbitrariness, namely the tricks and devices of the definitions of the auxiliary functions ϕ, ψ, θ, by means of which the oblique-angle problem is reduced to the Dirichlet problem. Is there not a canonical, straightforward general method for finding the index of a boundary value problem? We will return to this question below (Section 14.8).

Warning 3: The theory of ordinary differential equations easily conveys the impression that partial differential equations also possess a *general solution* in the form of a functional relation between the unknown function (*quantity*) u, the independent variables x and some arbitrary constants or functions, and that every

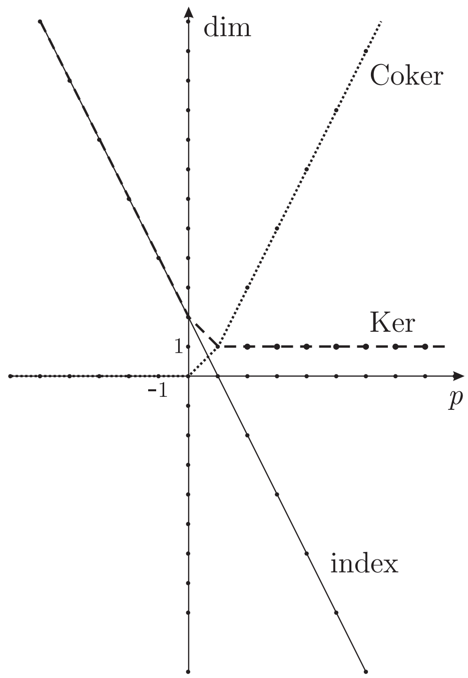


FIGURE 5.9. The dimensions of kernel and cokernel and the index of the Laplacian with boundary condition given by $\nu(z) = z^p$ for varying p

particular solution is obtained by substituting certain constants or functions f, h, etc. for the arbitrary constants and functions. (Corresponding to the higher degree of freedom in partial differential equations, we deal not only with constants of integration but with arbitrary functions.) The preceding calculations, regarding the boundary value problem of the Laplace operator, clearly indicate how limited this notion is which was conceived in the 18th century on the basis of geometric intuition and physical considerations. The classical recipe of first searching for general solutions and only at the end determining the arbitrary constants and functions fails. For example, the specific form of boundary conditions must enter the analysis to begin with.

EXERCISE 5.15. Without using Theorem 5.11, show that $\mathrm{Index}(\Delta, \frac{\partial}{\partial\nu}) = 0$ for $\nu := z$. This boundary-value problem, where $\frac{\partial}{\partial\nu}$ is the field normal to the boundary ∂X is named after Carl Neumann. From the topological viewpoint it is equivalent (modulo constant functions) to the Dirichlet boundary-value problem defined by a tangent vector field, see Figure 5.10.

EXERCISE 5.16. Let $X := \{z = x + iy : |z| < 1\}$ be the unit disk and define an operator

$$T : C^\infty(X) \times C^\infty(X) \to C^\infty(X) \oplus C^\infty(X) \oplus C^\infty(\partial X) \text{ by}$$
$$T(u,v) := \left(\frac{\partial u}{\partial \bar z}, \frac{\partial v}{\partial z}, (u-v)\,|_{\partial X}\right),$$

FIGURE 5.10. Topological equivalence of Neumann and Dirichlet boundary condition

where $C^\infty(X) := C^\infty(X,\mathbb{C})$ (i.e., we are back to complex-valued functions), $\frac{\partial}{\partial z} = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y})$ is complex differentiation and $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$ denotes the Cauchy-Riemann differential operator *formally adjoint* to $\frac{\partial}{\partial z}$. Prove that $\operatorname{index}(T) = 1$. [Hint: Show first that $\dim(\operatorname{Ker} T) = 1$: Suppose that $(u,v) \in \operatorname{Ker} T$. Then $\frac{\partial u}{\partial \bar{z}} = 0$ and $\frac{\partial v}{\partial z} = 0$ in which case u is holomorphic and v is conjugate-holomorphic (i.e., $\bar{v}$ is holomorphic). In particular, u and v are harmonic. Then since $(u-v)\,|_{\partial X} = 0$, we have $u = v$ on X. Why? Now $u'(z) = \frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0$, and so u and v are the same constant function. Then show $\operatorname{Coker}(T) = \{0\}$, or more precisely $(\operatorname{Im} T)^\perp = \{0\}$; see the footnote to Exercise 5.9. For this, choose arbitrary $f, g \in C^\infty(X)$ and $h \in C^\infty(\partial X)$ and prove that f, g and h must identically vanish, if

$$\int_X \left(\tfrac{\partial u}{\partial \bar{z}} f + \tfrac{\partial v}{\partial z} g\right) + \int_{\partial X} (u-v)h = 0 \text{ for all } u, v \in C^\infty(X). \tag{5.5}$$

Note that for $P, Q \in C^\infty(X)$

$$\begin{aligned} d\,(Pdz + Qd\bar{z}) &= \tfrac{\partial P}{\partial \bar{z}} d\bar{z} \wedge dz + \tfrac{\partial Q}{\partial z} dz \wedge d\bar{z} = \left(\tfrac{\partial P}{\partial \bar{z}} - \tfrac{\partial Q}{\partial z}\right) d\bar{z} \wedge dz \\ &= \left(\tfrac{\partial P}{\partial \bar{z}} - \tfrac{\partial Q}{\partial z}\right)(dx - idy) \wedge (dx + idy) = 2i \left(\tfrac{\partial P}{\partial \bar{z}} - \tfrac{\partial Q}{\partial z}\right) dx \wedge dy \end{aligned}$$

Thus you obtain the complex version of Stokes' Theorem,

$$\int_{\partial X} Pdz + Qd\bar{z} = \int_X d\,(Pdz + Qd\bar{z}) = 2i \int_X \left(\tfrac{\partial P}{\partial \bar{z}} - \tfrac{\partial Q}{\partial z}\right).$$

From this, you get

$$\begin{aligned} &\int_X \tfrac{\partial u}{\partial \bar{z}} f = \int_X \tfrac{\partial}{\partial \bar{z}}(uf) - \int_X u \tfrac{\partial f}{\partial \bar{z}} = \frac{1}{2i} \int_{\partial X} uf\,dz - \int_X u \tfrac{\partial f}{\partial \bar{z}} \text{ and} \\ &\int_X \tfrac{\partial v}{\partial z} g = \int_X \tfrac{\partial}{\partial z}(vg) - \int_X v \tfrac{\partial g}{\partial z} = \frac{-1}{2i} \int_{\partial X} vgd\bar{z} - \int_X v \tfrac{\partial g}{\partial z}. \end{aligned}$$

Hence,

$$\int_X \left(\tfrac{\partial u}{\partial \bar{z}} f + \tfrac{\partial v}{\partial z} g\right) = -\int_X \left(u \tfrac{\partial f}{\partial \bar{z}} + v \tfrac{\partial g}{\partial z}\right) + \frac{1}{2i} \int_{\partial X} (uf\,dz - vgd\bar{z}).$$

Assuming (5.5), you have

$$\begin{aligned} 0 &= \int_X \left(\tfrac{\partial u}{\partial \bar z} f + \tfrac{\partial v}{\partial z} g\right) + \int_{\partial X} (u-v)h \\ &= -\int_X \left(u \tfrac{\partial f}{\partial \bar z} + v \tfrac{\partial g}{\partial z}\right) + \frac{1}{2i}\int_{\partial X} (uf\,dz - vg\,d\bar z) + \int_{\partial X} (u-v)h \end{aligned}$$

By considering u and v with compact support inside the open disk, you can deduce that $\frac{\partial f}{\partial \bar z} = 0$ and $\frac{\partial g}{\partial z} = 0$ (i.e., f and g are analytic and conjugate analytic respectively). Thus, (5.5) implies

$$0 = \frac{1}{2i}\int_{\partial X} (uf\,dz - vg\,d\bar z) + \int_{\partial X} (u-v)h,$$

for all $u, v \in C^\infty(X)$. Choosing $v = u$, you have

$$\begin{aligned} 0 &= \frac{1}{2i}\int_{\partial X} u\,(f\,dz - g\,d\bar z) \text{ for all } u \Rightarrow f\,dz = g\,d\bar z \text{ on } \partial X \\ &\Rightarrow f\left(e^{i\theta}\right) ie^{i\theta}d\theta = -g\left(e^{i\theta}\right) ie^{-i\theta}d\theta \Rightarrow f\left(e^{i\theta}\right)e^{i\theta} = -g\left(e^{i\theta}\right)e^{-i\theta}. \end{aligned}$$

However, since f is analytic, the Fourier series of $f\left(e^{i\theta}\right)e^{i\theta}$ has a nonzero coefficient for $e^{im\theta}$ only when $m > 0$, and since g is conjugate analytic, $g\left(e^{i\theta}\right)e^{-i\theta}$ only has a nonzero coefficient for $e^{im\theta}$ only when $m < 0$. Thus, $f = g = 0$. Choosing $v = -u$, (5.5) then yields

$$0 = \int_{\partial X} (u-v)h = 2\int_{\partial X} uh \text{ for all } u \in C^\infty(X) \Rightarrow h = 0.$$

REMARK 5.17. In engineering one calls a system of separate differential equations

$$\begin{aligned} Pu &= f \\ Qv &= g, \end{aligned}$$

which are related by a *transfer condition* $R(u,v) = h$, a *coupling problem*; when the domains of u and v are different, but have a common boundary (or boundary part) on which the transfer condition is defined, then we have a *transmission problem*; e.g., see [**Bo72**, p.7f]. Thus, we may think of T as an operator for a problem on the spherical surface $X \cup_{\partial X} X$ (see Exercise 6.46 below) with different behavior on the upper and lower hemispheres, but with a fixed coupling along the equator.

CHAPTER 6

Differential Operators over Manifolds

Synopsis. Motivation. Differentiable Manifolds - Foundations: Implicit Function Theorem, Tangent Space, Cotangent Space. Geometry of C^∞ Mappings: Embeddings, Immersions, Submersions, Embedding Theorems. Integration on Manifolds: Hypersurfaces, Riemannian Manifolds, Geodesics, Orientation. Exterior Differential Forms and Exterior Differentiation. Covariant Differentiation, Connections and Parallelity: Connections on Vector Bundles, Parallel Transport, Connections on the Tangent Bundle, Clifford Modules and Operators of Dirac Type. Differential Operators on Manifolds and Symbols: Our Data, Symbolic Calculus, Formal Adjoints. Elliptic Differential Operators: Definition and Standard Examples. Manifolds with Boundary

1. Motivation

For many decades now, workers in differential geometry and mathematical physics have been increasingly concerned with differential operators (exterior differentiation, connections, Laplacians, Dirac operators, etc.) associated to underlying Riemannian or space–time manifolds. Of particular interest is the interplay between the spectral decomposition of such operators and the geometry/topology of the underlying manifold. This has become a large, diverse field involving index theory, the distribution of eigenvalues, zero sets of eigenfunctions, Green functions, heat and wave kernels, families of elliptic operators and their determinants, canonical sections, etc.. Moreover, John Donaldson's analysis of moduli of solutions of the nonlinear Yang–Mills equations and Seiberg–Witten theory have led to profound insights into the classification of four–manifolds, which were not accessible by techniques that are effective in higher dimensions. We shall touch upon many of these topics, but we focus on index theory and its applications (in Parts III-IV). In this Chapter and the following of this Part, we shall present an elementary introduction into the basic notions, concepts, and tools of global analysis.

We begin with the concept of a closed manifold. It allows us to generalize and simultaneously drastically simplify the index problem by eliminating boundary conditions. For example, the homogeneous Laplace equation $\Delta u = 0$ on the disk has infinitely many linearly independent solutions (e.g., $(x+iy)^n$), while the corresponding Laplace equation on the sphere has a one-dimensional solution space consisting of the constant functions. In this respect the notion of differentiable manifold, does not make the mathematics more complicated, but is a genuine first approximation to the difficult boundary value problems in Euclidean space $\mathbb{R}^n$.[1]

[1] The development of mathematics shows again and again how, in the growth of knowledge, the conceptual and non-conceptual form a unit, alternating, and fading into one another. A most striking example is furnished by the famous four-color problem, which characteristically still presents many puzzles in the plane, even after its computer aided solution, while the corresponding questions for closed manifolds have long been disposed of. "Most of the early attempts at solving

But also from the point of view of immediate applications, the geometric concept of a manifold played an important role. In fact, space-time problems defined initially and canonically in Euclidean space frequently do not have unrestricted independent variables, but these variables are restricted by side conditions to certain submanifolds of Euclidean space. Examples are the *constraints* in mechanics; the path equations of electrodynamics into which enter essentially the shape and surface of the *conductor*; or the symmetry conditions of elementary particle physics which replace the high dimensional Euclidean state spaces by low dimensional state spaces in the form of manifolds.

2. Differentiable Manifolds - Foundations

We begin with a compilation of the basic notions and elementary relations of the concept of a *differentiable manifold.* As a general reference, we refer to [**ST, BJ, Brd**].

Exercise 6.1. Recall the following two classical theorems of differential calculus, which form the foundation of the concept of a differentiable manifold.

a. (Inverse Function Theorem). If $f = (f_1, ..., f_n)$ is a C^∞ map from $\mathbb{R}^n$ to $\mathbb{R}^n$ whose $n \times n$ Jacobian matrix $[(\partial f_i/\partial x_j)(p)]$ has rank n at $p \in \mathbb{R}^n$ (i.e., its determinant is nonzero), then there is a neighborhood U of p in $\mathbb{R}^n$ which is diffeomorphic by f to a neighborhood V of $f(p)$ in $\mathbb{R}^n$.

b. (Implicit Function Theorem). Let $\mathcal{O} \subset \mathbb{R}^m$ be open and $f = (f_1, ..., f_n)$ be a C^∞ map from $\mathcal{O}$ to $\mathbb{R}^n$ $(m > n)$, whose $m \times n$ Jacobian matrix $[(\partial f_i/\partial x_j)(p)]$ at the point $p = (p_1, ..., p_m) \in U$ has maximal rank n. Thus, for some permutation of the coordinates x_j, the first $n \times n$ submatrix has rank n, i.e., we have

$$\det[(\partial f_i/\partial x_j)(p)]_{i,j=1,...,n} \neq 0.$$

Then the *isolevel set* $\{x \in \mathcal{O} : f(x) = f(p)\}$ can be *parametrized* locally. More precisely (see Figure 6.1), there is a differentiable map (*implicit function*) $g = (g_1, ..., g_n)$ defined in a neighborhood V of $(p_{n+1}, ..., p_m) \in \mathbb{R}^{m-n}$ with values in a neighborhood W of $(p_1, ..., p_n)$, such that $W \times V \subset \mathcal{O}$ and for all $x \in W \times V$ we have:

$$f(x) = f(p) \Leftrightarrow x_i = g_i(x_{n+1}, ..., x_m) \text{ for all } i \in \{1, ..., n\}.$$

A **topological manifold** without boundary is a locally Euclidean, Hausdorff topological space X. By locally Euclidean, we mean that for some $n \in \mathbb{N}$, each point of X has a neighborhood U which is homeomorphic via some function $u : U \to V$ to an open subset V of $\mathbb{R}^n$. The function u is called a **chart** for X. One might concretely think of geography, where a curved and uneven piece of the earth's surface is mapped onto a flat piece of paper. A chart is also known as a **local coordinate system**, when one wishes to stress the computational point of view. A set $\mathcal{A}$ of charts, whose domains of definition form an open covering of X, is called an **atlas**.

---

this problem were based on direct attack, and they not only failed, but did not even contribute any useful mathematics." Only "a new and highly indirect approach to the coloring problem based on a generalization of Kirchhoff's laws of circuit theory in a completely unforeseen direction", proved to be "successful in understanding a variety of combinatorial problems." (Gian-Carlo Rota, The Mathematical Sciences: A Report, 1969. Reprinted with the permisssion of the National Academy of Sciences)

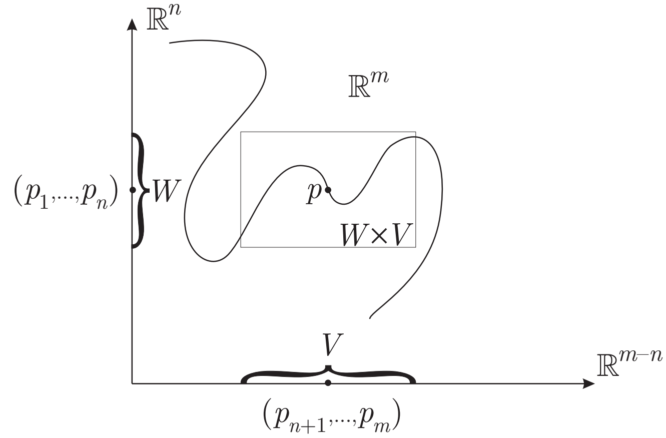


FIGURE 6.1. The Implicit Function Theorem, recalled, prametrizing locally (over V) the isolevel set $\{x : f(x) = f(p)\}$

The number n, the *local dimension*, is constant on each connected component of X. In our applications, n will not vary from component to component; so we may speak of the **dimension** of the manifold X.

Now let X be an n-dimensional manifold and $\mathcal{A}$ an atlas for X. Of geometric and analytic interest is the study of the coordinate changes $u \circ v^{-1}$, for two charts $u, v \in \mathcal{A}$ whose domains have non-void intersection. The change $u \circ v^{-1}$ is a continuous function from one open subset of $\mathbb{R}^n$ to another. This is trivial, by definition. However, in $\mathbb{R}^n$ one has much richer structure, which permits to impose further restrictions on the coordinate changes: The atlas $\mathcal{A}$ is called a **C^∞-atlas**, if all coordinate changes are C^∞ maps.

EXERCISE 6.2. Show: Each atlas $\mathcal{A}$ for X induces on an open subset W of X an atlas $\mathcal{A}|_W := \{u|_W : u \in \mathcal{A}\}$. If $\mathcal{A}$ is a C^∞-atlas, then so is $\mathcal{A}|_W$.

EXERCISE 6.3. Show that for each C^∞-atlas $\mathcal{A}$ on an n-dimensional manifold X, there is a commutative subalgebra (consisting of "C^∞ functions on X") of the algebra $C^0(X)$ of continuous complex-valued functions on X, namely

$$C^\infty(X) := \left\{\phi \in C^0(X) : \phi \circ u^{-1} \text{ is } C^\infty \text{ on } \operatorname{Im}(u) \text{ for all } u \in \mathcal{A}\right\}.$$

Moreover, show that the usual properties hold; e.g.,
a) For $\phi \in C^\infty(X)$, we have $\phi|_W \in C^\infty(W)$, where $W \subseteq X$ is open and $C^\infty(W)$ corresponds to the C^∞ atlas $\mathcal{A}|_W$.
b) Suppose ϕ is a fixed complex-valued function that is *locally* smooth (i.e., $\phi|_W \in C^\infty(W)$ for all $W \in \mathcal{W}$, where $\mathcal{W}$ is an open covering of X). Then $\phi \in C^\infty(X)$.
c) For $\phi \in C^\infty(X)$ and $\psi \in C^\infty(\mathbb{C})$, we have $\psi \circ \phi \in C^\infty(X)$.
d) The constant functions on X are in $C^\infty(X)$.

A C^∞ **manifold** is a topological manifold X with a "C^∞-structure" $C^\infty(X)$, defined by a C^∞-atlas $\mathcal{A}$.

A C^∞ **map** from a C^∞-manifold X to a C^∞-manifold Y is a function $f : X \to Y$ with $\phi \circ f \in C^\infty(X)$ for all $\phi \in C^\infty(Y)$. One can express this condition in terms of local coordinates as follows: For each u in the atlas for X and v in the atlas for Y, we have $v \circ f \circ u^{-1}$ is C^∞, as a map from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^m$, where

$n = \dim X$ and $m = \dim Y$. We denote the set of all C^∞ maps from X to Y by $C^\infty(X,Y)$.

The C^∞ manifolds X and Y are called **diffeomorphic** if there is a **diffeomorphism** from X to Y; i.e., a bijective C^∞ map $f \in C^\infty(X,Y)$ whose inverse is in $C^\infty(Y,X)$. If $X = Y$, one also calls such a map an **automorphism**.

For technical reasons one frequently requires that a C^∞ manifold be **paracompact**. This means that every covering of X by open subsets $\{U_j\}_{j\in J}$ possesses a **locally finite refinement** $\{V_k\}_{k\in K}$, in the sense that each V_k is an open subset of some $U_{j(k)}$, and for each fixed $x \in X$ there is a neighborhood O_x of x, such that the set $\{k \in K : O_x \cap V_k \neq \emptyset\}$ is finite. Note that by defining $W_j := \cup_{j(k)=j} V_k$ (possibly void), we can obtain a locally finite refinement $\{W_j\}_{j\in J}$ of $\{U_j\}_{j\in J}$ without changing the index set J. Unlike the stronger condition of compactness, we do not require that K is finite. It is well-known that every metric space is paracompact [**Du**, p.186], and every paracompact space is normal [**Du**, p.163].

THEOREM 6.4. *Let $\{U_j\}_{j\in J}$ be an open covering of a paracompact C^∞ n-manifold X. Then there is a "C^∞ partition of unity" subordinate to $\{U_j\}_{j\in J}$, namely, a family $\{\varphi_j \in C^\infty(X)\}_{j\in J}$ such that the following hold:*

(i) $\varphi_j \geq 0$,

(ii) *We have* $\operatorname{supp}\varphi_j :=$ *closure of* $\{y \in X : \varphi_j(y) \neq 0\} \subseteq U_j$, *and the family* $\{\operatorname{supp}\varphi_j : j \in J\}$ *is locally finite; i.e., for each $x \in X$, there is a neighborhood V_x such that $\{j \in J : V_x \cap \operatorname{supp}\varphi_j \neq \emptyset\}$ is finite, and*

(iii) $\sum_{j\in J} \varphi_j(x) = 1$ *for all $x \in X$.*

PROOF. If we can produce a C^∞ partition of unity subordinate to a refinement of $\{U_j\}_{j\in J}$, then it is subordinate to $\{U_j\}_{j\in J}$ itself. We can produce a refinement of $\{U_j\}_{j\in J}$ consisting of open subsets each of which are contained within the domain of a coordinate chart which maps the open subset to a bounded subset of $\mathbb{R}^n$. Since X is paracompact, we may then assume (without loss of generality) that the covering $\{U_j\}_{j\in J}$ is locally finite and for each $j \in J$, U_j is a subset of the domain $\widetilde{U}_j$ of a coordinate chart $u_j : \widetilde{U}_j \to \mathbb{R}^n$ such that $u_j(U_j)$ has compact closure in $\mathbb{R}^n$.
Step 1. We select an open neighborhood W_x about each point $x \in X$, such that the closure $\overline{W}_x$ is contained in some U_j. There is a subset $Y \subseteq X$, such that $\{W_y : y \in Y\}$ is a locally finite covering of X. By defining

$$V_j := \cup\left\{W_y : y \in Y,\ \overline{W}_y \subseteq U_j\right\},$$

we have a covering $\{V_j\}_{j\in J}$ of X. We also have $\overline{V}_j \subseteq U_j$ by the local finiteness of $\{W_y : y \in Y\}$. Indeed, suppose that $x \in \overline{V}_j$, then every neighborhood O_x of x intersects some W_y with $\overline{W}_y \subseteq U_j$. But there is some O_x such that only finitely many of such W_y, say $W_{y_1}, \ldots, W_{y_m}$, intersect O_x. Since the union of *finitely* many closed sets is closed,

$$x \in \overline{W_{y_1} \cup \ldots \cup W_{y_m}} \subseteq \overline{W}_{y_1} \cup \ldots \cup \overline{W}_{y_m} \subseteq U_j.$$

Step 2. For each $j \in J$, below we will construct a function $\psi_j \in C^\infty(M,\mathbb{R})$, which is positive on $V_j \subset U_j$ and identically zero on $X \setminus U_j$. Then, as $\{U_j\}_{j\in J}$ is locally finite, about each point of X there is a neighborhood on which all but a finite number of the ψ_j are identically 0, and so $\psi := \sum_j \psi_j$ is C^∞ and positive (since $\cup_{j\in J} V_j = X$). Set $\varphi_j := \psi_j/\psi$ and note that conditions (i), (ii) and (iii)

hold. With the help of the coordinate function $u_j : \widetilde{U}_j \to \mathbb{R}^n$, we carry out the construction of ψ_j as follows. Define $\eta \in C^\infty(\mathbb{R}^n)$ by

$$\eta(x) := \begin{cases} \exp\left(\frac{-1}{1-|x|^2}\right), & \text{for } |x| < 1, \\ 0, & \text{for } |x| \geq 1. \end{cases}$$

Since $u_j(\overline{V}_j)$ is a closed subset of the compact subset $\overline{u_j(U_j)} \subset \mathbb{R}^n$, $u_j(\overline{V}_j)$ is also compact. Let $\delta > 0$ be the distance from the compact set $u_j(\overline{V}_j)$ to the closed subset $\mathbb{R}^n - u_j(U_j)$. Cover $u_j(\overline{V}_j)$ with finitely many open balls of radius less than δ and centers $a_1, \ldots, a_m \in u_j(\overline{V}_j)$. Finally, for any $p \in X$, let

$$\psi_j(p) := \begin{cases} \sum_{k=1}^m \eta\left(\frac{u_j(p)-a_k}{\delta}\right), & p \in U_j, \\ 0, & p \in X \setminus U_j. \end{cases}$$

Note that $\psi_j > 0$ on V_j since $\eta\left(\frac{x-a_k}{\delta}\right) > 0$ for $|x - a_k| < \delta$, and supp $\psi_j \subseteq U_j$ since $\sum_{k=1}^m \eta\left(\frac{x-a_k}{\delta}\right)$ is 0 for x in a neighborhood of $\mathbb{R}^n - u_j(U_j)$. □

REMARK 6.5. The preceding proof is typical for many *non-constructive*, set-theoretic arguments in analysis. Actually, one almost always has canonically given charts relative to which an *explicit* partition of unity can be provided.

REMARK 6.6. The C^∞ partition of unity is an important tool which is used to globally piece together locally given data in a smooth way. Although in some applications we deal with analytic manifolds, generally we will stay in the category of C^∞ manifolds, since there clearly is no analytic version of Theorem 6.4.

3. Geometry of $\mathbf{C^\infty}$ Mappings

In elementary differential calculus, many geometrical questions concerning functions (the location of extreme values, inflection points, etc.) can be answered by investigating the derivatives (i.e., linear approximations) of the function. By means of linear algebra, one can also study C^∞ mappings between manifolds. The essential concepts for this are:

1. The directional derivative. Let x be a point of a C^∞ manifold X, $\varphi \in C^\infty(X)$, and $c : \mathbb{R} \to X$ a C^∞ map (a C^∞ curve) with $c(0) = x$. Then the **directional derivative** of the function φ in the direction of the curve c is defined to be $(\varphi \circ c)'(0)$, the derivative of $\varphi \circ c$ at 0 in the sense of elementary calculus. Two such curves are **equivalent**, when the directional derivatives of each function relative to the two curves are the same. We denote such an equivalence class by

$$\dot{c}(0) := \left\{\widetilde{c} \in C^\infty(\mathbb{R}, X) : \widetilde{c}(0) = x \text{ and } \forall_{\varphi \in C^\infty(X)} (\varphi \circ \widetilde{c})'(0) = (\varphi \circ c)'(0)\right\}.$$

Note that $\dot{c}(0)$ only depends on how c is defined near 0.

2. The tangent space. The set of directional derivatives, and hence the set of equivalence classes of curves, forms a real vector space,

$$T_xX := (TX)_x := \{\dot{c}(0) : c \in C^\infty(\mathbb{R}, X),\ c(0) = x\}$$

called the **tangent space** of X at x. Clearly, the multiplication of the directional derivative $c'(0)$ by a real number λ is given by a λ-fold increase in the speed; i.e.,

$$\lambda\dot{c}(0) := \dot{\tilde{c}}(0), \text{ where } \tilde{c}(t) := c(\lambda t), \text{ for } t \in \mathbb{R}.$$

Also, for two curves c_1 and c_2, we can add $\dot{c}_1(0)$ and $\dot{c}_2(0)$ by setting

$$\dot{c}_1(0) + \dot{c}_2(0) := \left(u^{-1}(u \circ c_1 + u \circ c_2)\right)'(0),$$

where $u : U \to \mathbb{R}^n$ $(n := \dim X)$ is a chart with $u(x) = 0 \in \mathbb{R}^n$ and we arbitrarily redefine c_1 and c_2 outside of a neighborhood of $0 \in \mathbb{R}$ so that $c_1(\mathbb{R}) \cup c_2(\mathbb{R}) \subset U$. One can verify that these operations are well-defined and the axioms for a vector space hold. Also one may check that $\dim(T_x X) = \dim X$. For this, one chooses a C^∞ chart $u : U \to \mathbb{R}^n$, from the open neighborhood U of x to an open subset of $\mathbb{R}^n$. Then, for each positively directed coordinate line through $u(x)$, there is a corresponding C^∞ curve in X, as depicted in Figure 6.2. The corresponding directional derivatives are denoted by $\frac{\partial}{\partial u_1}|_x, \ldots, \frac{\partial}{\partial u_n}|_x$ and these form a basis for $T_x X$.

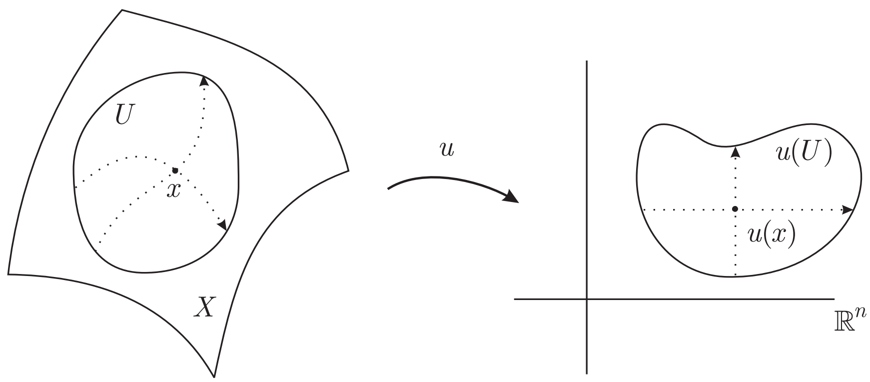


FIGURE 6.2. Coordinate lines through $u(x)$ and the corresponding curves in X

3. The tangent bundle. The disjoint union $TX := \cup_{x \in X} T_x X$ of tangent spaces has the structure of a real C^∞ vector bundle over X, namely the **tangent bundle**; see Appendix, Exercise B.11, p. 683. A map $s : X \to TX$ is a **section**, if $s(x) \in T_x X$ for all $x \in X$. Let $x \in X$ and let $c_x \in C^\infty(\mathbb{R}, X)$ be a curve representing $s(x)$; i.e., $c_x(0) = x$ and $s(x) = \dot{c}_x(0)$. Then the directional derivative of $\varphi \in C^\infty(X)$ at x in the direction $s(x)$ is

$$s(x)[\varphi] := (c_x \circ \varphi)'(0).$$

If the function

$$s[\varphi] : X \to \mathbb{R} \text{ given by } s[\varphi](x) := s(x)[\varphi]$$

is in $C^\infty(X)$ for each $\varphi \in C^\infty(X)$, then $s : X \to TX$ is a C^∞ section of TX or a C^∞ **vector field**.

4. The differential. A C^∞ map $f : X \to Y$ determines a linear map (the **differential** of f at x)

$$f_{*x} : T_x X \to T_{f(x)} Y \quad \text{given by} \quad f_{*x}(\dot{c}_x(0)) = (f \dot{\circ} c_x)(0),$$

where $\dot{c}_x(0) \in T_x X$. Sometimes we write $f_*|_x$ instead of f_{*x} to clarify that the differential f_* is evaluated at x. Let $u = (u_1, \ldots, u_m) : U \to \mathbb{R}^m$ and $v = (v_1, \ldots, v_n) : V \to \mathbb{R}^n$ be coordinates about x and $f(x)$ respectively, and let

$$(f_1, \ldots, f_n) := v \circ f \circ u^{-1} : u(f^{-1}(V) \cap U) \to \mathbb{R}^n.$$

The matrix of f_{*x} with respect to the coordinate bases $\left(\left.\frac{\partial}{\partial u_1}\right|_x, \ldots, \left.\frac{\partial}{\partial u_m}\right|_x\right)$ and $\left(\left.\frac{\partial}{\partial v_1}\right|_{f(x)}, \ldots, \left.\frac{\partial}{\partial v_n}\right|_{f(x)}\right)$ is given by

$$\left[\left.\frac{\partial}{\partial u_j}\right|_x [f_i]\right] = \left[\frac{\partial f_i}{\partial u_j}(u(x))\right],$$

which is the $n \times m$ Jacobian matrix of $(f_1, \ldots, f_n)$. Note that $f_* : TX \to TY$ is a bundle map (linear in the fibres, as explained in the Appendix p.676). Moreover, f is called an **immersion** if f_* is injective, an **embedding** if f and f_* are injective, and a **submersion** if f_* is surjective at each $x \in X$.

5. Submanifolds. A subset $Y \subseteq X$ is called a **submanifold**, if Y is a C^∞ manifold with the induced topology and the inclusion $i : Y \to X$ is a C^∞ embedding. (One can also consider the image sets of immersions as *submanifolds with self-intersections*, a concept that we will not pursue further.) For all $y \in Y$, T_yY is a linear subspace of T_yX in a natural way.

The following theorem illustrates some closely related representations of C^∞ manifolds.

THEOREM 6.7. *Let X and Y be C^∞ manifolds of dimensions m and n $(m > n)$.*
a) (***Definition of manifolds through equations***) *Let $q \in Y$ and $f : X \to Y$ a C^∞ map with f_{*x} surjective for all $x \in f^{-1}\{q\}$. Then $f^{-1}\{q\}$ has the structure of an $(m-n)$-dimensional submanifold of X, in a natural way.*
b) (***Representation of submanifolds of*** $\mathbb{R}^N$) *If Y is a submanifold of $\mathbb{R}^N$ and $y \in Y$, then (for a certain renumbering of the Euclidean coordinates $x_1, ..., x_N$), the projection of Y to the n-dimensional subspace $\{(x_1, ..., x_n, 0, \ldots, 0)\} \cong \mathbb{R}^n$ is a local coordinate system $v : U \to \mathbb{R}^n$ for Y in some neighborhood $U \subseteq Y$ of y. Moreover, there is a neighborhood V of y in $\mathbb{R}^N$ as in Figure 6.3, such that $Y \cap V$ is the set of points satisfying the following system of equations for unique functions $g_{n+1}, \ldots, g_N \in C^\infty(v(U))$:*

$$x_{n+1} = g_{n+1}(x_1, ..., x_n), \ldots, \; x_N = g_N(x_1, ..., x_n).$$

c) (***Embedding Theorem***) *Every closed C^∞ manifold X can be embedded in $\mathbb{R}^N$ for N sufficiently large.*

REMARK 6.8. One might consider the special cases of part a) of Theorem 6.7 where the sphere is represented as a level set of the distance function, or where the matrix manifold $\mathrm{SL}(n, \mathbb{R})$ is represented as a level set of the determinant function. One can visualize (b) with $Y = S^2$ and $N = 3$. The proof of (a) and (b) follows without difficulty from the Implicit Function Theorem (Exercise 6.1b). The complete and elementary proofs for (b) and (c) can be found in [**Wa**, p.35-43] or [**Brd**, p.91-92] where it is proved that every C^∞ n-manifold X can be embedded in $\mathbb{R}^{2n+1}$. Actually, H. Whitney proved that every C^∞ n-manifold X can be embedded in $\mathbb{R}^{2n}$; see [**Wh**]. Assuming moreover that X is orientable, X can be embedded in $\mathbb{R}^{2n-1}$. For $n \neq 4$, this was done in [**HH**]. Decades later, the case $n = 4$ was finally settled as a consequence of work by J. Boéchat, A. Haefliger and S. K. Donaldson; see the posting "All smooth orientable 4-manifolds embedd in $\mathbb{R}^7$" by Paulo Ney de Souza at http://math.berkeley.edu/˜desouza and the references therein.

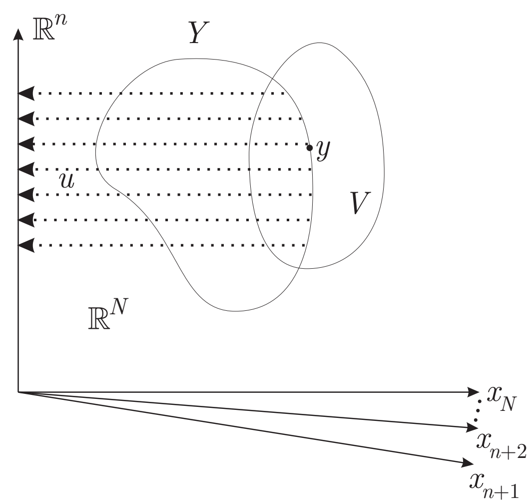


FIGURE 6.3. Local representation of a submanifold Y (depicted as a curve) by projection

We will not repeat the quoted proofs here, since we aim not at minimizing the dimension of the *receiving* Euclidean space. Instead of that we give an ultra-short proof, following [**Fu99**, Section 2.2, pp.30-33]. This proof yields what we want, namely an embedding into a finite-dimensional Euclidean space, but eventually of quite high dimension. The basic idea goes back to work in algebraic geometry by Kunihiko Kodaira, Fritz Hirzebruch and others, namely to *fill* large spaces of functions or sections until an *ample* level is reached where one gets something manageable or trivial.

PROOF OF THEOREM 6.7C. We begin with an elemenary set-theoretical argument, to illustrate the idea of *filling*: Any manifold X can be *embedded* in a huge Euclidean space of highly infinite dimension, for suitable definition of the terms *topology*, *differential* and *embedding* for infinite-dimensional manifolds. Indeed, consider the natural mapping

$$\begin{array}{cccc} \iota & : & X & \longrightarrow & \mathbb{R}^{C^\infty(X)} \\ & & x & \mapsto & C^\infty(X) \ni f \mapsto f(x) \in \mathbb{R}, \end{array}$$

where $\mathbb{R}^{C^\infty(X)}$ denotes the set of all mappings from the space $C^\infty(X)$ of smooth (here real-valued) functions to $\mathbb{R}$, or, differently put, the direct product of copies of $\mathbb{R}$ over all elements of $C^\infty(X)$. This is a really huge Euclidean space. Each single $f \in C^\infty(X)$ may be perceived as a *coordinate function*, namely a reader or *parser* of all $x \in X$. We give a formal argument for the fact that the preceding map ι is an embedding. First we address the immersiveness. An analog

$$\begin{array}{cccc} \iota_* & : & TX & \longrightarrow & T\mathbb{R}^{C^\infty(X)} \\ & & \dot{c}_x(0) & \mapsto & C^\infty(X) \ni f \mapsto (f \dot{\circ} c_x)(0) \in T\mathbb{R}, \end{array}$$

of the differential for ι is defined by taking the differential of each *coordinate*. Then the mapping ι_* maps any non-zero tangen vector $v = \dot{c}_x(0)$ to a non-vanishing vector in $T\mathbb{R}^{C^\infty(X)}$. More precisely, for a non-zero tangent vector v, there is a smooth function g so that the derivative of g in the direction of v is not zero. This implies that the coordinate function, corresponding to g, of $\iota_*(v)$ is non-zero.

Now, we check the injectivity of the mapping ι. For a pair x_0, x_1 of distinct points, there is a smooth function f so that the values $f(x_0)$ and $f(x_1)$ are distinct. This implies that the points $\iota(x_0), \iota(x_1) \in \mathbb{R}^{C^\infty(X)}$ take different values of the coordinate function f, and we are done.

However, we want an embedding in a *finite*-dimensional Euclidean space. To do that, we trimm the full set $C^\infty(X)$ of coordinate functions down to a finite number. The set of the coordinate functions to be selected must be suffciently *ample* to yield still an embedding (the idea is very similar to the task of Exercise B.10).

Let's begin anew. Denote by $\mathbb{P}(TX)$ the set of all possible tangent directions over all points on X and call its elements *tangent lines* (as in our Exercise B.2 in the Appendix). The space $\mathbb{P}(TX)$ is a fiber bundle over the *compact* base space X. (The concept of a fiber bundle embraces the concept of vector bundles of our Appendix B, and the concept of principal G-bundles of our Definition 16.1). Its fiber is a real projective space, which is compact. Hence, the total space is also compact. For any tangent line ℓ at any point x, we can choose a smooth function f_ℓ, whose derivative in the direction of ℓ is not zero. Then there is an open neighborhood U_ℓ of ℓ in $\mathbb{P}(TX)$ such that all derivatives of f_ℓ for all directions in U_ℓ are not zero. Since $\mathbb{P}(TX)$ is compact, we can cover it with finitely many open subsets $U_{\ell_1}, U_{\ell_2}, \ldots, U_{\ell_s}$. Then the mapping $F := (f_{\ell_1}, f_{\ell_2}, \ldots, f_{\ell_s})$ gives an immersion.

Unfortunately, the mapping F is not necessarily injective. However, since the mapping F is an immersion, its restriction to a suffciently small neighborhood U_x of an arbitrary $x \in X$ is injective. We form the open subset $\bigcup_{x \in X} U_x \subset X \times X$. Its complement K is compact. By definition, for each pair $y = (x_0, x_1) \in K$, the points x_0 and x_1 are distinct. So there is an $f_y \in C^\infty(X)$ with $f_y(x_1) \neq f_y(x_0)$. Then there is an open neighborhood V_y of y in $X \times X$ such that $f_y(x_1') \neq f_y(x_0')$ for all $(x_0', x_1') \in V_y$. We select finitely many such open subsets $V_{y_1}, V_{y_2}, \ldots, V_{y_u}$ to cover the compact K. Now we set $G := (f_{y_1}, f_{y_2}, \ldots, f_{y_u})$. Then the mapping $(F, G) : X \to \mathbb{R}^{s+u}$ gives a desired embedding. □

6. The cotangent bundle. Let X be a C^∞ manifold with $x \in X$. In place of the tangent space T_xX, one can consider its dual space T_x^*X (other valid notations are $(T_xX)^*$, $(TX)_x^*$, and $(T^*X)_x$) of linear maps from T_xX to $\mathbb{R}$. An element of T_x^*X can be identified with the differential (see 4. above) at x of a real-valued function $\varphi \in C^\infty(X, \mathbb{R})$, namely $\varphi_{*x} : T_xX \to T_{\varphi(x)}\mathbb{R} \cong \mathbb{R}$. The notation $(d\varphi)_x$ or $d\varphi|_x$ is also used for φ_{*x}. If $u = (u_1, ..., u_n) : U \to X$ is a chart for X on a neighborhood U of x, then the differentials $(du_1)_x, ..., (du_n)_x$ form a basis for T_x^*X. One can also give the disjoint union $T^*X = \cup_{x \in X} T_x^*X$ a bundle structure, and indeed, T^*X is exactly the *dual bundle* of TX; see Appendix, Exercise B.4, p. 679. For short, T^*X is called the **dual tangent bundle**, **covariant bundle**, or most commonly, the **cotangent bundle**.

Under a coordinate change $v = \kappa \circ u$, the differentials change *covariantly*,

$$dv_i = \sum_{j=1}^{n} \tfrac{\partial \kappa_i}{\partial x_j} du_j$$

while the tangent vectors transform *contravariantly* by means of the Jacobian of κ^{-1}. From the standpoint of category theory, however, the tangent bundle is covariant since $f : X \to Y$ yields a well-defined bundle map $f_* : TX \to TY$ whereas there is a well defined bundle map $\widetilde{f} : T^*X \to T^*Y$ only when f is a diffeomorphism. Then we may define

$$\left(\widetilde{f}\,(\alpha_x)\right)(Z) := \alpha_x(f^{-1})_*\,(Z) \qquad \text{for } \alpha_x \in T_x^*X \text{ and } Z \in T_{f(x)}Y. \tag{6.1}$$

The situation for induced maps on sections is different, as we now explain. The space of C^∞ sections of TX is denoted by $C^\infty(TX)$ and such a section is known as a **vector field** on X. A vector field on X generally does not push forward to a well-defined vector field on Y unless $f : X \to Y$ is a diffeomorphism. Indeed, if f is not onto, the purported push-forward will not be defined everywhere, while if f is not 1-1, the purported push-forward may be ill-defined on $f(X)$. The sections in $C^\infty(T^*X)$ (also denoted by $\Omega^1(X)$) are known as 1-forms on X. For any $f \in C^\infty(X, Y)$ (not necessarily a diffeomorphism), there is a well-defined map

$$\begin{aligned} &f^* : C^\infty(T^*Y) \to C^\infty(T^*X) \quad \text{ given by} \\ &f^*(\mu)(Z_x) := \mu_{f(x)}(f_* Z_x) \quad \text{ for } \mu \in C^\infty(T^*Y) \text{ and } Z_x \in T_xX, \end{aligned}$$

which is known as the **pull-back** of 1-forms induced by f. Moreover, there is a pull-back $f^* : \Omega^k(Y) \to \Omega^k(X)$ of k-forms (see Appendix B, Exercise 6.17, p. 169) defined in the same way.

If X is a submanifold of Y with the embedding $f : X \to Y$, then although f is not necessarily a diffeomorphism, we can still easily define $f^* : T^*Y|_{f(X)} \to T^*X$ via $f^*\left(\alpha_{f(x)}\right)(Z_x) := \alpha_{f(x)}\left(f_*(Z_x)\right)$ for $\alpha_{f(x)} \in T^*Y|_{f(X)}$ and $Z_x \in T_xX$, and then $\operatorname{Ker} f^*$ is a subbundle of $T^*Y|_{f(X)}$ known as the **normal bundle** of the embedding. Put differently, the normal bundle consists of those covectors at points of $f(X)$, which annihilate all vectors tangent to $f(X)$.

REMARK 6.9. Cotangent bundles of manifolds arise naturally in abstract formulations of classical mechanics and analytical mechanics, e.g., in the Hamiltonian formulation of classical mechanics, which provides one of the major motivations for the field: The set of all possible configurations of a system is modelled as a manifold, and this manifold's cotangent bundle describes the phase space of the system. Locally, i.e., over a coordinate patch $u(U) \subset X$ for a chart $u : U \to X$, we have $T^*X|_{u(U)} \cong \mathbb{R}^{2n}$ with the *canonical* (once the chart is chosen) coordinates $x_1, \ldots, x_n, du_1, \ldots, du_n$, traditionally called the pairing of space and impulse coordinates. For two such (x, μ) and (y, ν) we set

$$\omega\big((x,\mu),(y,\nu)\big) := \langle (x,\mu), J(y,\nu)\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathbb{R}^{2n}$ and $J := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ the usual skew-symmetric $2n \times 2n$ block matrix. Then ω is a symplectic form (bilinear, skew-symmetric and nondegenerate) for $T^*X|_{u(U)}$. Actually, the whole bundle T^*X can be considered as a symplectic manifold by defining (rather trivially) an exterior nondegenerate skew-symmetric differential 2-form ω on it with $d\omega = 0$, see also our Definition 19.21, p.619. A very readable introduction to local symplectic geometry and Hamilton-Jacobi theory is given in [**GS**, Chapter 5, pp.55-66, and Chapter 9, pp.97-106]. In the symbolic calculus of elliptic operators, the *symplectic cone* $\mathring{T}^*X := T^*X \setminus X$ plays a fundamental role. It consists of the punctured cotangent spaces.

4. Integration on Manifolds

Hypersurfaces. Suppose that X is an n-dimensional submanifold of $\mathbb{R}^{n+1}$ (i.e., a **hypersurface**), and moreover assume that X is the boundary of a bounded open subset of $\mathbb{R}^{n+1}$. From the notion of integration on $\mathbb{R}^{n+1}$ where one has a canonical volume element, we have a *surface element* on X, whence integration over X is well-defined.

Riemannian Manifolds. In principal one can use the same recipe for a compact C^∞ Riemannian manifold X.

DEFINITION 6.10. A C^∞ manifold X is **Riemannian**, if it has been given a **Riemannian metric** (**tensor**), namely for all $x \in X$ the tangent space T_xX is equipped with a fixed Euclidean metric tensor $\langle \cdot\cdot, \cdot\cdot \rangle$ (positive, symmetric, nondegenerate $\mathbb{R}$-valued, bilinear form), such that for two C^∞ sections s_1 and s_2 of the tangent bundle TX, the function $\langle s_1, s_2 \rangle$ is in $C^\infty(X)$.

It may be helpful to understand the concept of a Riemannian metric in local coordinates. Thus, let $x \in X$ and $u = (u^1, \dots, u^n) : U \to \mathbb{R}^n$ be coordinates with $X \supset U \ni x$, whence $\left(\frac{\partial}{\partial u^1}|_x, \dots, \frac{\partial}{\partial u^n}|_x\right)$ a basis for T_xX. In these coordinates, a metric is represented by a positive definite, symmetric matrix $\left(g_{ij}(x)\right)_{i,j=1,\dots,n}$, (i.e., $g_{ij}(x) = g_{ji}(x)$ for all i, j and $\sum_{i,j} g_{ij}(x)\xi^i\xi^j > 0$ for all $\mathbb{R}^n \ni \xi = (\xi^1, \dots, \xi^n) \neq 0$), where the coefficients depend smoothly on x.

The inner product of two tangent vectors $A, B \in T_xX$ with coordinate representations $(a^1, \dots, a^n)$ and $(b^1, \dots, b^n)$ (i.e., $A = \sum_i a^i \frac{\partial}{\partial u^i}$ and $B = \sum_j b^j \frac{\partial}{\partial u^j}$) then is

$$\langle A, B \rangle_x = \sum_{1 \le i,j \le n} g_{ij}(x) a^i b^j .$$

In particular, $\left\langle \frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j} \right\rangle_x = g_{ij}(x)$. Similarly, the length of A is given by $\|A\| := \sqrt{\langle A, A \rangle_x}$.

With the help of a C^∞ partition of unity (see Theorem 6.4, p. 157), one can furnish every paracompact manifold with a Riemannian metric. Indeed, let $\{U_j\}_{j \in J}$ be a locally finite covering of the n-manifold X by domains of coordinate charts $u_j : U_j \to \mathbb{R}^n$, say $u_j = \left(u_j^1, \dots, u_j^n\right)$, and let $\{\varphi_j \in C^\infty(X)\}_{j \in J}$ be a partition of unity subordinate to $\{U_j\}_{j \in J}$, then, for $x \in X$ and $A, B \in T_xX$,

$$g_x(A, B) := \sum\nolimits_{j \in J \text{ with } x \in U_j} \left(\varphi_j(x) \sum\nolimits_{k=1}^n d\left(u_j^k\right)_x (A)\, d\left(u_j^k\right)_x (B) \right)$$

defines a Riemannian metric. Since there is no j-th term if $x \notin U_j$, the sum over J is really finite on a neighborhood of each point.

On any submanifold X of the Euclidean space $\mathbb{R}^N$ there is a natural Riemannian metric induced by restricting the Euclidean inner product on $\mathbb{R}^N$ to TX. Since we have seen that any n-manifold X can be realized as a submanifold of $\mathbb{R}^N$ for N sufficiently large (Theorem 6.7c), we have another (less elementary) existence proof for Riemannian metrics.

Geodesics. For a general metric space, a geodesic is defined as a curve which realizes the shortest distance between any two sufficiently close points lying on it. For Riemannian manifolds we can be more explicit. As in metric space, we ask that geodesics are only *locally* the shortest distance between points. Additionally we ask that they are parameterized with *constant velocity*, i.e., proportionally to arc length.

It is very fortunate that (locally) minimizing the *energy* will also minimize the *length* - and give the wanted parametrisation for free. The details of the argument can be found in [**Jo**, Section 1.4] and any other textbook on Riemannian geometry.

Let $[t_0, t_1]$ be a closed interval in $\mathbb{R}$ and $c : [a, b] \to X$ a smooth curve. The **length** and the **energy** of c then are defined as

$$L(c) := \int_{t_0}^{t_1} \|\dot{c}(t)\| dt \quad \text{ and } \quad E(c) := \frac{1}{2}\int_{t_0}^{t_1} \|\dot{c}(t)\|^2 dt.$$

In physics, the massless term $E(c)$ is usually called *action* of c where c is considered as the orbit of a mass point. In local coordinates, we write $\dot{c}(t) = \sum_i \gamma^i(t) \frac{\partial}{\partial u^i}|_{c(t)}$ and obtain

$$L(c) = \int_{t_0}^{t_1} \Big(\sum_{i,j} g_{ij}(c(t))\gamma^i(t)\gamma^j(t)\Big)^{\frac{1}{2}} dt, \quad E(c) = \frac{1}{2}\int_{t_0}^{t_1} \sum_{i,j} g_{ij}(c(t))\gamma^i(t)\gamma^j(t) dt.$$

We also remark for later technical purposes that the length of a (continuous and) piecewise smooth curve may be defined as the sum of the lengths of the smooth pieces, and the same holds for the energy.

On a Riemannian manifold X, the **distance** $\operatorname{dist}(x, x')$ between two points x, x' can be defined as

$$\inf\{L(c) : c : [t_0, t_1] \to X \text{ piecewise smooth curve with } c(t_0) = x, c(t_1) = x'\}.$$

If X is connected, it is also pathwise connected, i.e., any two points $x, x' \in X$ can be connected by a path, actually by a piecewise smooth path. [Prove it by decomposing X into the open (!) set X_x of all $p \in X$ which can be connected with x by a piecewise smooth path, and the open (!) complement $X \setminus X_x$ consisting of the union of all similarly defined sets X_q with $q \notin X_x$. Since X_x is not empty, the complement must be.] So, $\operatorname{dist} : X \times X \to [0, \infty)$ is well defined, and one checks easily that it is a metric for X.

Definition 6.11. A smooth curve $c : [t_0, t_1] \to X$ which is a critical point of the energy functional is called a **geodesic**.

Recall that the *Euler-Lagrange equations* of a functional

$$I(c) := \int_{t_0}^{t_1} f\left(t, c^1(t), \dots, c^n(t), \dot{c}^1(t), \dots, \dot{c}^n(t)\right) dt$$

for $c = (c^1, \dots, c^n) : [t_0, t_1] \to \mathbb{R}^n$, are given by

$$\frac{d}{dt}\frac{\partial f}{\partial \dot{c}^i} - \frac{\partial f}{\partial c^i} = 0, \quad i = 1, \dots, n.$$

Then the critical points of our energy functional $E(c)$ are given by the system of n second order differential equations

$$\ddot{c}^i(t) + \sum_{j,k} \Gamma^i_{jk}(c(t))\dot{c}^j(t)\dot{c}^k(t) = 0, \quad i = 1, \dots, n$$

with

$$\Gamma^i_{jk} := \frac{1}{2}\sum_{\ell} g^{i\ell}(g_{j\ell,k} + g_{k\ell,j} - g_{jk,\ell}),$$

where

$$(g^{ij})_{i,j=1,\dots,n} := (g_{ij})^{-1}, \quad (\text{i.e., } \sum_{\ell} g^{i\ell}g_{\ell j} = \delta_{ij}), \quad \text{and} \quad g_{j\ell,k} := \frac{\partial}{\partial u^k} g_{j\ell}.$$

The expressions Γ^i_{jk} are called **Christoffel symbols**. Christoffel symbols play a prominent role in all concrete calculations with connections. In that context, they show up below in Section 6.6, Equation (6.15), and become central in Section 16.5, in Equation (18.62) of Section 18.4, p.525, and in analyzing Equations (19.108) in Section 19.4, p.662f in our Part IV, beginning with Equations (16.47), (16.48) on p.392. The details of the preceding deduction can be found in [**Jo**, Lemma 1.4.4] and many other places.

From the Local Existence and Uniqueness Theorem for systems of ordinary differential equations we obtain

PROPOSITION 6.12. *Let X be a Riemannian manifold, $x \in X, v \in T_xX$. Then there exist $\varepsilon > 0$ and precisely one geodesic $c : [0,\varepsilon] \to X$ (to be denoted by c_v) with $c(0) = x,\ \dot{c}(0) = v$.*

One can show (and make precise) that, in addition, c_v depends smoothly on x and v. We then define

DEFINITION 6.13. Let X be a Riemannian manifold, $x \in X$.
a) The mapping

$$\begin{array}{cccc} \exp_x & : & V_x & \to & X \\ & & v & \mapsto & c_v(1) \end{array} \quad \text{with } V_x := \{v \in T_xX : c_v \text{ is defined on } [0,1]\}$$

is called the **exponential map** of X at x.
b) The **point injectivity radius** of x is

$$\rho(x) := \sup\bigl\{\rho > 0 : \exp_x \text{ is defined and injective on } \{v \in T_xX : \|v\| \le \rho\}\bigr\}.$$

c) The **injectivity radius** of X is $\rho(X) := \inf\{\rho(x) : x \in X\}$.

For example, the injectivity radius of the sphere $X := S^n$ is π, since the exponential map of any point x maps the open ball of radius π in T_xX injectively onto the complement of the antipodal point of x.

3. Orientation and Integrability. A C^∞ manifold X is **oriented** when an atlas for X has been chosen such that the Jacobian matrix for each coordinate change is positive (i.e., $det(\partial\kappa_i/\partial x_j) > 0$, for charts u and v, with $u = \kappa \circ v$ on the intersection of their domains). More simply, without recourse to differential calculus, we can also express this as follows. The bases of a finite-dimensional vector space are divided into two **orientation classes**; two bases belong to the same class, if their transformation matrix has positive determinant. By means of an orientation of a C^∞ manifold, one may select two classes of bases from each tangent space (*positive* and *negative* bases with regard to the fixed orientation) in such a way that in a neighborhood of each point the choice is given by a continuous or differentiable choice of basis.

The familiar Möbius band is an example of a non-orientable manifold. A submanifold Y of an orientable manifold X (even of codimension only 1) is therefore not always orientable. On the other hand, Y is automatically orientable if it is the boundary of an open, oriented manifold X. Then at each point $y \in Y$ one can define a basis of T_yY to be positively oriented when a positive basis of T_xX is obtained by adjoining an *outward pointing* vector. Since we will only be interested in *bounding* manifolds with classically defined surface elements in most of our applications, we state without proof that with the help of a suitable partition of unity and local charts in which the metric can be expressed in terms of curvilinear coordinates, the concept of integration of functions on $\mathbb{R}^n$ carries over to the case of Riemannian manifolds. Orientability is required for the integration of n-forms, because the *sign* of an n-form will make sense since it will not vary under an orientation-preserving change of chart (see Exercise 6.17, p. 169). Actually, the integration of $\mathbb{R}$-valued functions (as opposed to n-forms) requires only a measure or density (absolute value of an n-form) since the sign of such a function is already unambiguous.

Definition 6.14. Let X be a smooth oriented paracompact Riemannian manifold with $\dim X = n$. We denote the metric tensor by g.
a) Let $\{U_j\}_{j\in J}$ be an open cover of X and $\{x_j = (x_j^1, \dots, x_j^n) : U_j \to \mathbb{R}^n\}$ local, positively oriented coordinates on U_j. Then each

$$\nu_{j,g} := \sqrt{|\det(g_{ik})|}\,dx_j^1 \wedge \cdots \wedge dx_j^n$$

defines a Lebesgue measure on each U_j and hence all together a Lebesgue measure ν_g on X that is called the **volume form** of X.
b) Let $\pi : E \to X$ be a smooth complex vector bundle over X with Hermitian metric $\langle \cdot, \cdot \rangle_h$. Set $|e|_h := \sqrt{\langle e, e\rangle_h}$ for $e \in E$. An L^2**-section** of E is a Lebesgue measurable map $\psi : X \to E$ (i.e., $\psi^{-1}(U)$ is Lebesgue measurable for any open subset $U \subset E$) such that

(i) $\pi \circ \psi(x) = x$ for almost all $x \in X$ except possibly a negligible set.
(ii) The function $x \mapsto |\psi(x)|_h$ belongs to $L^2(X, \mathbb{R})$.

The space of L^2-sections of E is denoted by $L^2(E)$.

In the notation of Exercise 6.17 of the following section, we have $\nu_g = *_0(1) \in \Omega^n(X)$. That yields a coordinate-free definition of the volume form. It is worth mentioning that the volume $\mathrm{vol}(X) := \int_X \nu_g \in [0, \infty) \cup \{\infty\}$ is well defined. We leave it to the reader to check that $L^2(E)$ is a Hilbert space with respect to the scalar product

$$(\varphi, \psi)_0 := \int_X \langle \varphi(x), \psi(x)\rangle_{E_x,h}\, \nu_g \qquad \varphi, \psi \in L^2(E).$$

Remark 6.15. For some computations on manifolds it is impractical and confusing to constantly revert back to local coordinates. In such cases *intrinsic* coordinate-invariant concepts of integration like the preceding definition of the volume form by the linear star operator are welcome, and they require in part weaker hypotheses. For the integration of n-forms, the existence of a volume element is essential, or, more generally in modern terminology, the existence of a distinguished n-form where $n = \dim X$. See, e.g., [**ST**, p.134/150].

Remark 6.16. For the time being, we will use only the above integration aspects of the Riemannian metric, and thus only scratch the surface of *Riemannian geometry* which unfolds in the great classic theorems on parallel displacement (connections), curvature and rigidity with their varied computations (see [**Berg**]). Our Part IV deals extensively with such topics, in particular with connections on principal G-bundles, rigorously defined in Section 16.2, p.366. Already in Section 6.6, p.172ff, connections on vector bundles will be introduced and used for coordinate free integration in our definition of Sobolev spaces and pseudo-differential operators in Chapters 7 and 8. We have noted (see (3.30), p. 117) that there is a close relationship (the Gauss-Bonnet Theorem) between the integral of the Gaussian curvature and the topological form of a surface, namely the genus or Euler characteristic. See also Chapters 13/14 below in Part III, and Part IV where the higher-dimensional Gauss-Bonnet-Chern Theorem is proved using the local index theorem (see Theorem 18.63, p. 580).

5. Exterior Differential Forms and Exterior Differentiation

It is possible to construct (see Exercise 6.17 below), from the tangent bundles, bundles of *exterior differential forms* by means of multilinear algebra. These are an important tool in describing physical laws mainly in the areas of electromagnetism and special relativity. This is the case when empirical relationships are to be expressed in terms of an integral in such a way that the physicist or engineer can pursue qualitative and quantitative changes resulting from modifications of the integrand or the domain of integration. For this reason, exterior bundles are of interest also in differential topology; see Chapter 14.

We briefly summarize (details are found in [**Brd**, p.260f], [**KN63**, p.17f], and [**UN**, p.111-161] and the literature given there): For a real n-dimensional vector space V, we form the vector space $\Lambda^p(V)$ of p-fold **skew-symmetric tensors** (or p-**vectors**); these are the multilinear maps

$$V^* \times \overset{p \text{ times}}{\cdots} \times V^* \to \mathbb{R},\ p \in \mathbb{N},\ V^* := L(V, \mathbb{R})$$

which change, under a permutation of the arguments, by a factor equal to the sign of the permutation. One sets $\Lambda^0(V) := \mathbb{R}$ and obtains $\Lambda^1(V) = V$, $\Lambda^{n-1}(V) \cong V$, $\Lambda^n(V) \cong \mathbb{R}$ and $\Lambda^p(V) = \{0\}$ for $p > n$. For $v \in \Lambda^p(V)$ and $w \in \Lambda^q(V)$, we define $v \wedge w \in \Lambda^{p+q}(V)$ by

$$(v \wedge w)(a_1, ..., a_{p+q}) := \frac{1}{p!q!} \sum\nolimits_\sigma \operatorname{sgn}(\sigma)\,(v \otimes w)\,(a_{\sigma(1)}, ..., a_{\sigma(p+q)})$$

(sum over all permutations), which gives the **exterior multiplication** $\Lambda^p(V) \times \Lambda^q(V) \to \Lambda^{p+q}(V)$. This multiplication makes $\Lambda^\bullet(V) := \sum_{p=0}^n \Lambda^p(V)$ a graded algebra, the **exterior algebra** of V.

If $e_1, ..., e_n$ is a basis of V, then the $\binom{n}{p}$ forms $e_{i_1} \wedge \cdots \wedge e_{i_p}$ with $1 \le i_1 < \cdots < i_p \le n$ yield a basis for $\Lambda^p(V)$. With this property, $\Lambda^p(V)$ is occasionally defined (in order to avoid the suggestive but tedious definition via maps) as the space of p-**vectors**: the space of formal linear combinations of the p-tuples of basis vectors $e_{i_1} \wedge \cdots \wedge e_{i_p}$ with only the relation $e_{\sigma(i_1)} \wedge \cdots \wedge e_{\sigma(i_p)} = \operatorname{sgn}(\sigma) e_{i_1} \wedge \cdots \wedge e_{i_p}$.

A scalar product (= inner product) for V induces a scalar product $\langle \cdot, \cdot \rangle$ for $\Lambda^p(V)$, and declaring an orthonormal basis $e_1, ..., e_n$ of V to be positively oriented yields an explicit isomorphism $\Lambda^n(V) \cong \mathbb{R}$ via $e_1 \wedge \cdots \wedge e_n \mapsto 1$, which only depends on the chosen orientation and scalar product. The linear **star operator**

$$\begin{array}{ccccc} *_p & : & \Lambda^p(V) & \to & \Lambda^{n-p}(V) \\ \text{is generated} & \text{by} & e_{i_1} \wedge \cdots \wedge e_{i_p} & \mapsto & e_{j_1} \wedge \cdots \wedge e_{j_{n-p}}, \end{array} \tag{6.2}$$

where $j_1 \dots j_{n-p}$ is selected such that $e_{i_1}, \dots e_{i_p}, e_{j_1} \dots e_{j_{n-p}}$ is a positive basis of V.

Since the star operator is supposed to be linear, it is determined by its values on some basis (6.2). It is characterized by the property $u \wedge *_p v = \langle u, v \rangle\, e_1 \wedge \cdots \wedge e_n$

for all $u, v \in \Lambda^p(V)$. In particular,

$$*_0(1) = e_1 \wedge \cdots \wedge e_n \tag{6.3}$$

$$*_n(e_1 \wedge \cdots \wedge e_n) = 1, \tag{6.4}$$

if $e_1, \ldots, e_n$ is a positive basis. From the rules of multilinear algebra, it easily follows that if $A \in \mathrm{End}(V)$, and if $f_1, \ldots, f_p \in V$, then

$$*_p(Af_1 \wedge \cdots \wedge Af_p) = (\det A) *_p (f_1 \wedge \cdots \wedge f_p).$$

In particular, this implies that the star operator does not depend on the choice of positive orthonormal basis in V, as any two such bases are related by a linear transformation with determinant 1. For a negative basis instead of a positive one, one gets a minus sign on the right hand sides of (6.2), (6.3), (6.4).

EXERCISE 6.17. Let X be a compact C^∞ manifold of dimension n with or without boundary, with metric tensor g.
a) Show that the family of vector spaces $\Lambda^p(T_x^*X)$, $x \in X$, yields a real vector bundle of fiber dimension $\binom{n}{p}$ over X in a natural way. We denote this bundle by $\Lambda^p(T^*X)$ or shortly $\Lambda^p(X)$. Correspondingly, define a bundle $\Lambda^\bullet(T^*X)$ by summation. Check that g induces a smoothly varying inner product on the fibres of $\Lambda^p(T^*X)$ and $\Lambda^\bullet(T^*X)$.
b) Customarily, one writes $\Omega^p(X) := C^\infty(\Lambda^p(T^*X))$, which is the space of **exterior differential p-forms** on X and $\Omega^\bullet(X) = \sum_{p=0}^n \Omega^p(X)$. For a C^∞ function f, consider the differential df (see also Section 6.6.2) and show that the operator $d : \Omega^0(X) \to \Omega^1(X)$ uniquely extends to a linear differential operator $d : \Omega^p(X) \to \Omega^{p+1}(X)$ of first order (see Chapter 6) for each p, such that $d^2 := d \circ d = 0$ and

$$d\,(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^p\, \alpha \wedge d\beta$$

for all $\alpha \in \Omega^p(X)$ and $\beta \in \Omega^q(X)$.
c) Once we have an orientation of X, the definition of the linear star operator carries over from the vector spaces $\Lambda^p(T_x^*)$ to the vector bundles $\Lambda^p(T^*X)$ and, finally, to the vector spaces $\Omega^\bullet(X)$, then called **Hodge star operator**. Apply (6.3) to obtain

$$*_0(1) = \sqrt{|\det{(g_{ik})}|}dx^1 \wedge \cdots \wedge dx^n \tag{6.5}$$

in local coordinates.
Prove the *Hodge duality* $*_p : \Omega^p(X) \cong \Omega^{n-p}(X)$. Assume n even, and check, whether $* := \oplus_{p=0}^n *_p$ is an involution. If not, how should one modify $*$ to obtain an involution?
d) Prove that for a compact, oriented, n-dimensional, Riemannian manifold X with boundary ∂X, we have Stokes' Theorem

$$\int_X d\omega = \int_{\partial X} \omega, \text{ for } \omega \in \Omega^p(X).$$

[Hint for a): In principle, use the same mechanism as in Exercise B.4. Note that for charts u and w for X in a neighborhood of $x \in X$, we have the simple transformation rules, e.g., for a 1-form $v \in \Omega^1(X)$,

$$v(x) = \sum\nolimits_{j=1}^n a_j(x)\; du^j\big|_x = \sum\nolimits_{i=1}^n b_i(x)\; dw^i\big|_x\,, \text{ where}$$

$$b_i(x) := \sum\nolimits_{j=1}^n a_j(x) \frac{\partial \left(u \circ w^{-1}\right)^j}{\partial x^i} \left(w\left(x\right)\right).$$

For b): d is characterized by the Leibniz rule

$$d(v \wedge w) = dv \wedge w + (-1)^p v \wedge dw, \text{ for } v \in \Omega^p(X),\ w \in \Omega^q(X).$$

How is d written in local coordinates?
For c): Note that $*_{n-p}*_p = (-1)^p \operatorname{Id}_{\Omega^p(X)}$. That yields Hodge duality, but noinvolution $*$. Assuming n even, say $n = 2m$, does not help. However, try $\tau_p := i^{m+p(p-1)}*_p$. Then

$$\tau_{2m-p} \circ \tau_p = i^{m+(2m-p)(2m-p-1)} i^{m+p(p-1)} (-1)^p,$$

and so $\tau := \oplus_{p=0}^{2m} \tau_p$ is an involution ($\tau^2 = 1 = \operatorname{Id}_{\Omega^\bullet(X)}$). The ± 1 eigenspaces $\Omega^\pm$ of τ are crucial for defining the Hirzebruch signature operator, see Sections 13.2, p.298, 14.4, p.312, and 18.6, Theorem 18.59, p.573, where some of the preceding tasks are executed, both in greater generality - and in more detail.
For d): See [**GP**, p.182-187]. Incidentally, here one really needs the orientation.]

Remark 6.18. In algebraic terms, the exterior algebra $\Lambda^\bullet(T^*X)$ is the universal unital algebra generated by T^*X subject to the relations $\xi \wedge \xi = 0$ for $\xi \in T^*X$. To sum up, the wedge product extends to the full bundle $\Lambda^\bullet(T^*X)$ of exterior algebras and to the space $\Omega^\bullet(X)$ of exterior differential forms on X. The mapping $w\wedge : \Omega^\bullet(X) \to \Omega^\bullet(X)$ is called (left) *exterior multiplication*.

There is also an *interior multiplication* for skew-symmetric tensors, exterior algebras and exterior differential forms. Exterior multiplication adds an index if possible and increases the degree of forms; interior multiplication cancels an index and decreases the degree of forms. More precisely, let V be a real vector space equipped with an inner product $\langle \cdot, \cdot \rangle$ and $\dim V = n$. Then there is a linear mapping $w\llcorner\colon \Lambda^p(V) \to \Lambda^{p-1}(V)$ for each $w \in V$, defined via

$$w\llcorner(v_1 \wedge \cdots \wedge v_p) := \sum_{j=1}^{p} \langle v_j, w \rangle v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_p, \tag{6.6}$$

where $\widehat{v_j}$ means that the factor v_j is omitted.

Relative to the induced inner product on $\Lambda^\bullet(V)$, the mapping $w\wedge$ and $w\llcorner$ are adjoints (see also Section 18.1 and Proposition 18.10, p.489). Like exterior multiplication, the mapping $w\llcorner$ induces endomorphisms of $\Lambda^\bullet(T^*X)$ and $\Omega^\bullet(X)$, called the (left) **interior multiplication**.

The geometric meaning of interior multiplication is explained in Sections 6.7, p.181 and 14.4, p.311 for the principal symbol of the codifferential δ (=the adjoint of d, see below), in Section 14.5, p.316 for discussing the number of vector fields, and in Section 18.10, p.489 for spinor representations.

The operator $d : \Omega^p(X) \to \Omega^{p+1}(X)$ is known as the **exterior derivative operator** or **exterior differentiation**. In local coordinates, $x^1, \ldots, x^n$ defined on a coordinate neighborhood $U \subseteq M$, a form $\alpha \in \Omega^p(X)$ can be written as

$$\alpha = \frac{1}{p!} \sum_{1 \le i_1, \ldots, i_p \le 1}^{n} \alpha_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p},$$

where the $\alpha_{i_1\cdots i_p} \in C^\infty(U)$ are antisymmetric in the indices $i_1, \ldots, i_p$. On U,

$$\begin{aligned} d\alpha &= \frac{1}{p!} \sum_{1 \le i_1, \ldots, i_p \le n}^{n} d\left(\alpha_{i_1\cdots i_p}\right) dx^{i_1} \wedge \cdots \wedge dx^{i_p} \\ &= \frac{1}{p!} \sum_{1 \le i_1, \ldots, i_p \le n}^{n} \tfrac{\partial}{\partial x^i}\left(\alpha_{i_1\cdots i_p}\right) dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}. \end{aligned}$$

However, since $d : \Omega^p(X) \to \Omega^{p+1}(X)$ is uniquely determined by the coordinate-free operation $d : \Omega^0(X) \to \Omega^1(X)$, one should be able to express d in a coordinate-free manner. For this purpose (and because it is an important and basic notion), we introduce *Lie differentiation*.

Let $A \in C^\infty(TX)$ be a vector field on X. The theory of systems of first-order ordinary differential equations guarantees that for each point $p \in X$, there is $\varepsilon > 0$ and a curve $\alpha_p : (-\varepsilon, \varepsilon) \to X$ such that $\dot\alpha_p'(t) = A_{\alpha_p(t)}$. This curve α is known as the **integral curve** of A through p. Furthermore, ε can be chosen so that for all q in some neighborhood U of p, the integral curve $\alpha_q : (-\varepsilon, \varepsilon) \to X$ exists, and there is a well-defined C^∞ map

$$\alpha : U \times (-\varepsilon, \varepsilon) \to X \quad \text{ given by } \alpha(q,t) := \alpha_q(t),$$

such that $\alpha_t := \alpha(\cdot, t) : U \to \alpha(U,t)$ is a diffeomorphism for each $t \in (-\varepsilon, \varepsilon)$. Moreover, $\alpha_{t+s} = \alpha_t \circ \alpha_s$ whenever both sides are defined. In particular $\alpha_t^{-1} = \alpha_{-t}$ on $\alpha(U,t) \cap U$ which is nonvoid for small t. Given a second vector field $B \in C^\infty(TX)$, we have $\left(\alpha_t^{-1}\right)_* \left(B_{\alpha_t(p)}\right) \in T_pX$ and so the curve

$$t \mapsto \left(\alpha_t^{-1}\right)_* \left(B_{\alpha_t(p)}\right) = (\alpha_{-t})_* \left(B_{\alpha_t(p)}\right)$$

lies in the *single* vector space T_pX. It then makes sense to differentiate this curve at $t = 0$ to obtain a vector in T_pX which is known as the **Lie derivative** of B with respect to A at p, namely

$$\left(\mathcal{L}_A B\right)_p := \tfrac{d}{dt} \left(\alpha_{-t}\right)_* \left(B_{\alpha_t(p)}\right)\Big|_{t=0}.$$

It is also called the **Lie bracket** of A and B at p. We explain why: The assignment $p \mapsto \left(\mathcal{L}_A B\right)_p$ defines a vector field $\mathcal{L}_A B \in C^\infty(TX)$. If A and B are vector fields on $\mathbb{R}^n$ and p, $\delta p \in \mathbb{R}^n$, then (where $\approx$ denotes equality modulo terms of first-order in t) we have

$$(\alpha_{-t})_*(\delta p) \approx \delta p - t\,(dA)_p(\delta p) \qquad \text{and} \qquad B_{\alpha_t(p)} \approx B_{p+tA_p} \approx B_p + t\,(dB)_p(A_p).$$

Thus,

$$\begin{aligned} &(\alpha_{-t})_*\left(B_{\alpha_t(p)}\right) \approx B_{\alpha_t(p)} - t\,(dA)_p\left(B_{\alpha_t(p)}\right) \approx B_p + t\,(dB)_p(A_p) - t\,(dA)_p(B_p) \\ &\Rightarrow \left(\mathcal{L}_A B\right)_p = (dB)_p(A_p) - (dA)_p(B_p) = A_p[B] - B_p[A]. \end{aligned}$$

Hence, $\mathcal{L}_A B = -\mathcal{L}_B A$ and viewing the vector fields A and B as differential operators on functions f, we have

$$\left(\mathcal{L}_A B\right)_p(f) = (df)_p\left(A_p[B] - B_p[A]\right) = A_p[B[f]] - B_p[A[f]],$$

and so as a differential operator $\mathcal{L}_A B$ is the commutator of A and B, i.e.,

$$\mathcal{L}_A B = A \circ B - B \circ A = [A, B].$$

In local coordinates $x^1, \ldots, x^n$, we have (automatically summing over repeated indices)

$$\begin{aligned}\mathcal{L}_A B &= \mathcal{L}_A B\left[x^i\right]\partial_{x^i} = \left(A\left[B\left[x^i\right]\right] - B\left[A\left[x^i\right]\right]\right)\partial_{x^i}\\ &= \left(A^j\partial_{x^j}B^i - B^j\partial_{x^j}A^i\right)\partial_{x^i}.\end{aligned}$$

Now we show how the exterior derivative can be expressed by the Lie derivative. For a one-form ω, we then have the coordinate-free relation

$$d\omega\left(A,B\right) = A\left[\omega\left(B\right)\right] - B\left[\omega\left(A\right)\right] - \omega\left(\left[A,B\right]\right), \text{ since} \tag{6.7}$$

$$\begin{aligned}&A\left[\omega\left(B\right)\right] - B\left[\omega\left(A\right)\right] - \omega\left(\left[A,B\right]\right)\\ &\quad= A\left[\omega_j B^j\right] - B\left[\omega_j A^j\right] - \omega\left(A^i\partial_{x^i}\left(B^j\right)\partial_{x^j} - B^i\partial_{x^i}\left(A^j\right)\partial_{x^j}\right)\\ &\quad= A^i\left(B^j\partial_{x^i}\omega_j + \omega_j\partial_{x^i}B^j\right) - B^i\left(A^j\partial_{x^i}\omega_j + \omega_j\partial_{x^i}A^j\right)\\ &\quad- \omega_j A^i\partial_{x^i}B^j + \omega_j B^i\partial_{x^i}A^j\\ &\quad= \partial_{x^i}\omega_j\left(A^iB^j - B^iA^j\right) = \partial_{x^i}\left(\omega_j\right)\left(dx^i\wedge dx^j\right)\left(A,B\right) = d\omega\left(A,B\right).\end{aligned}$$

EXERCISE 6.19. For $\psi \in \Omega^2(X)$ and vector fields $A, B, C \in C^\infty(TX)$, show that

$$\begin{aligned}d\psi\left(A,B,C\right) &= A\left[\psi\left(B,C\right)\right] + B\left[\psi\left(C,A\right)\right] + C\left[\psi\left(A,B\right)\right]\\ &\quad- \psi\left(\left[A,B\right],C\right) - \psi\left(\left[C,A\right],B\right) - \psi\left(\left[B,C\right],A\right)\\ &= \mathfrak{S}\left(A\left[\psi\left(B,C\right)\right] - \psi\left(\left[A,B\right],C\right)\right),\end{aligned}$$

where $\mathfrak{S}$ denotes the sum over all cyclic permutations of (A, B, C). [Hint. Consider the case $\psi = \omega\wedge\varphi$, for $\omega, \varphi \in \Omega^1(X)$, and use $d\psi = d\omega\wedge\varphi - d\varphi\wedge\omega$. The general case follows by linearity, since any 2-form is locally a sum of wedges of 1-forms.]

More generally, one can show by induction that for $\psi \in \Omega^k(X)$

$$\begin{aligned}d\psi\left(A_1,\ldots,A_{k+1}\right) &= \sum_{i=1}^{k+1}\left(-1\right)^{i+1}A_i\left[\psi\left(A_1,\ldots,\widehat{A_i},\ldots,A_{k+1}\right)\right]\\ &+ \sum_{1\le i<j\le k+1}\left(-1\right)^{i+j}\psi\left(\left[A_i,A_j\right],A_1,\ldots,\widehat{A_i},\ldots,\widehat{A_j},\ldots,A_{k+1}\right),\end{aligned} \tag{6.8}$$

where $\widehat{A_i}$ indicates that A_i is omitted. For a proof, see [**KN63**, p.36]. Note that the extra numerical factor of $1/\left(k+1\right)$ in the formula of [**KN63**, p.36] is ultimately due to their convention that $\left(dx^1\wedge\cdots\wedge dx^n\right)\left(\partial_{x^1},\ldots,\partial_{x^n}\right) = 1/n!$ (see [**KN63**, p.7]), while our convention is that $\left(dx^1\wedge\cdots\wedge dx^n\right)\left(\partial_{x^1},\ldots,\partial_{x^n}\right) = 1$.

An extensive discussion of Lie brackets, and the general theory of Lie derivatives can be found in [**Bl81**, Chapter 0] and [**KMS**, Chapters I-II]. In particular, see the last reference (Sections II.7.6-II.7.9, pp.63-66) for the place of interior multiplication and for further relations between exterior differentiation and Lie derivation.

6. Covariant Differentiation, Connections and Parallelity

Children of our motorized time are familiar with the concepts of speed and accelleration and able to clearly distinguish between them: A car can move on a straight highway with high speed, constant velocity and no accelleration; and it can move after a stop light with low velocity, but high accelleration. So much for small children. When they grow older and have learned about the interpretation of force as a product of mass and accelleration (Newton's Second Law), most of them will fall back to the pre-Newtonian identification of velocity and accelleration. Ask them to draw the trajectory of a thrown ball or rock! Most will correctly draw a parabola which is a good approximation for a rock that is thrown for short distances. But then ask them to mark the acting forces by directed arrows along the trajectory! Most will draw tangent vectors of varying length (the impulses) instead of the solely vertically acting constant gravitation. Some smarties would talk away the apparent contradiction by referring to *resulting force* or to *air resistance*. The smartest of all of them was the Greek philosopher and polymath Aristotle (384 - 322 BCE) who derived a straight trajectory of finite length for the thrown stone until the impuls of the initial throw was *consumed*, followed by vertical fall-down. To his student Alexander, the later famous military leader, he explained the visible deviation of his theoretical trajectory from observed orbits by the complexity of full reality, air resistance, wind influence, imperfect shape of the thrown object etc.

A modern geometer may have two comments to that continuing confusion of concepts. (I) Analyzing a single trajectory is not very challenging. Elementary calculus yields simple definitions of the velocity $\dot{c}(t) \in \mathbb{R}^3$ and the accelleration $\ddot{c}(t) \in \mathbb{R}^3$ of a sufficiently smooth path $c : [t_0, t_1] \to \mathbb{R}^3$ at a point $c(t) \in \mathbb{R}^3$ for $t \in [t_0, t_1]$. The student will see at once that the vectors $\dot{c}(t)$ and $\ddot{c}(t)$ have different direction, in general, and even can be perpendicular to each other in natural parametrisation, namely $\ddot{c}(t)$ pointing to the center of the curvature. Moreover, writing the equations of motion for the curve c with given initial position, velocity and accelleration $c(t_0), \dot{c}(t_0), \ddot{c}(t_0)$ in (x, y, z) coordinates gives a simple one-dimensional problem. Only the vertical z-coordinate is relevant. One can neglect the y-coordinate for a plane movement and the movement in x direction is not accellerated in the absence of forces in that direction. In flat $\mathbb{R}^3$, we can do without distinguishing between the spaces of state (configuration), velocity (tangent) and accelleration (forces) as long as we do correct calculations. At first, the geometer may wonder about the success of classical mechanics with concepts that belong to different categories but are commonly put in the same (Euclidean) space. Thinking about it, the geometer will explain the astonishing correctness of sloppy and vague physics terminology by the flatness of Euclidean space. There is nothing to worry about.

(II) In a second comment, the geometer would admit that there is a lot to worry about. Recall Section 6.3, where it is natural to distinguish between the points of a manifold X and equivalence classes of paths making its tangent bundle TX. We did it in a coordinate-free manner, admitting non-Euclidean X, rigorously, canonically and without special choices or ambiguities. Later we made choices, to put a Riemannian stucture, i.e., smoothly varying metrics on the tangent spaces. But the basic concept of TX was canonical. Now, similarly, we might wish to define a second derivative - canonically. That is impossible. Why? Consider a vector field $s \in C^\infty(X, TX)$, i.e., a section in the bundle $TX \to X$. It specifies a direction, i.e., a tangent vector $s(x) \in T_xX$ in each point $x \in X$. We may consider the vector field s as a field of velocities. To get something like a second derivative, an accelleration or a force we would specify a direction $v \in T_xX$ and take the limit

$$\nabla_v s|_x \overset{!?}{:=} \lim_{t \to 0} \frac{s(x + tv) - s(x)}{t}\,. \tag{6.9}$$

That looks familiar - except for two problems: We have to define a translation "$+tv$" for small real t yielding a point $x + tv$ in the neigborhood of x. That can be done using a Riemannian metric for X (in contrast to the fact that the concept of tangent vector, tangent space and tangent bundle was defined fully invariantly and without reference to a metric). The second problem is more serious: There is no canonical way in TX how to compare two tangent vectors at different base points. We have to make choices. We have to make parallel translations and to specify the ways to do them for a given bundle. That is the idea of a connection.

In this Part II, we use only a very simple concept of connection, first, and most general for real and complex vector bundles, and then more specifically for the tangent bundle. We would like to emphasize that the concept of a connection has many more ramifications than the few dry definition terms we give in this section. It has become the central concept in modern low-dimensional geometry and, as well, in gauge-theoretic quantum field theory and particle physics. That will be explained in our Part IV.

Connections on Vector Bundles.

DEFINITION 6.20. Let X be a smooth manifold and E a smooth real or complex vector bundle over X. A **connection** (or **covariant differentiation operator**) ∇^E on E is an $\mathbb{R}$-linear first order differential operator $\nabla^E : C^\infty(X,E) \to C^\infty(X, T^*X \otimes E)$ satisfying the Leibniz rule

$$\nabla^E(fs) = df \otimes s + f\nabla^E s \tag{6.10}$$

for all functions $f \in C^\infty(X,\mathbb{R})$ and sections $s \in C^\infty(X,E)$.

Recall that the bundles $T^*X \otimes E$ and $\mathrm{Hom}(TX,E)$ are isomorphic (as real vector bundles, see also Exercise B.4, p.679 for the complex category). Whence, we can consider ∇^E as a mapping

$$C^\infty(X,E) \times C^\infty(X,TX) \ni (s,v) \quad \mapsto \quad \nabla^E_v s := \big(\nabla^E(s)\big)(v) \in C^\infty(X,E)$$

with the following properties for $v,w \in C^\infty(X,TX)$, $s,r \in C^\infty(X,E)$, $f \in C^\infty(X,\mathbb{R})$ and $c \in \mathbb{R}$:

tensorial in v:

$$\nabla^E_{v+w} s = \nabla^E_v s + \nabla^E_w s \quad \text{and} \quad \nabla^E_{fv} s = f\nabla^E_v s; \tag{6.11}$$

linear in s:

$$\nabla^E_v(s+r) = \nabla^E_v s + \nabla^E_v r \quad \text{and} \quad \nabla^E_v cs = c\nabla^E_v s; \tag{6.12}$$

product rule:

$$\nabla^E_v(fs) = (df)\big(\nabla^E_v s\big) + f\big(\nabla^E_v s\big). \tag{6.13}$$

EXERCISE 6.21. a) Set $X := \mathbb{R}^n$, E the (trivial) product bundle $X \times \mathbb{C}^N$ (denoted by $\mathbb{C}^N_X$ in Exercise B.1a of the Appendix B, p.676) and show that the attempted definition of (6.9) yields a connection.
b) Show that each complex or real vector bundle E over a smooth manifold X admits a connection.
[Hint for (a): Check that all the properties of the preceding list are satisfied.
Hint for (b): Choose a locally finite covering of X by charts, and choose trivializations of E; apply (a); assemble the local connections to a global operator; and check the properties.]

Parallel Transport. Closely related to the formal definition of a connection is the concept of parallel transport of geometric information.

DEFINITION 6.22. Let ∇^E be a connection for a vector bundle $\pi_E : E \to X$. A section $s \in C^\infty(X, E)$ is called **parallel** relative to ∇^E along a smooth path $c : [t_0, t_1] \to X$ in X, if

$$\nabla^E_{\dot{c}(t)} s(c(t)) = 0 \quad \text{for all } t \in (t_0, t_1). \tag{6.14}$$

EXERCISE 6.23. Let X, E, ∇^E, c be as in the preceding definition. Define a **parallel translation**

$$\tau^E_{c,t} : E_{c(t)} \longrightarrow E_{c(t_0)} \quad \text{for } t \in (t_0, t_1).$$

[Hint: Begin with local coordinates around $x_0 := c(t_0)$. So, choose a coordinate patch $X \supset U \ni x_0$ and coordinates $x^1, \dots, x^n$ yielding coordinate vector fields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ in $TX|_U$. For simplicity, consider E as a real vector bundle of fiber dimension N. Through the identification $E|_U \cong U \times \mathbb{R}^N$ you obtain a basis $s_1, \dots, s_N$ of sections of $E|_U$. For the given connection ∇^E, define the so-called **Christoffel symbols** Γ^k_{ij} $(j, k = 1, \dots, N, i = 1, \dots, n)$ by the condition

$$\nabla^E_{\frac{\partial}{\partial x^i}} s_j =: \sum_{k=1}^N \Gamma^k_{ij} s_k \,. \tag{6.15}$$

See also our geometric interpretation of the Christoffel symbols in Equations (16.47) and (16.48) of Part IV below on pp.392ff. Let now $s \in C^\infty(X, E)$. Locally, you may write $s(y) = \sum_k a^k(y) s_k(y)$. Putting $s(t) := s(c(t))$, you define a section of E along c. Furthermore, let $\dot{c}(t) =: \sum \gamma^i \frac{\partial}{\partial x^i}$. Then by (6.11)-(6.13) and (6.15), find

$$\begin{aligned}\nabla^E_{\dot{c}(t)} s(t) &= \sum_k a^k(t) s_k(c(t)) + \sum_{i,k} \gamma^i(t) a^k(t) \big(\nabla^E_{\frac{\partial}{\partial x^i}} s_k\big)(c(t)) \\ &= \sum_k a^k(t) s_k(c(t)) + \sum_{i,j,k} \gamma^i(t) a^k(t) \Gamma^j_{ik}(c(t)) s_j(c(t)).\end{aligned}$$

Note that $\nabla^E_{\dot{c}(t)} s(t)$ depends only on the values of s along the curve c, and not on all the values of s in a neighborhood of the trajectory $\{c(t) : t \in [t_0, t_1]\}$. Our Equation (6.14) thus represents a linear system of first order ordinary differential equations for the coefficients $a^1(t), ..., a^N(t)$ of the section $s(t)$ along c you are looking for. Therefore, for given initial values $s(t_0) \in E_{c(t_0)}$, you obtain a unique solution of (6.14). This gives you an isomorphism $E_{c(t_0)} \xrightarrow{\cong} E_{c(t)}$ for all $t \in [t_0, t_1]$. Take the inverse as the wanted parallel translation.]

Thus, if $x, x' \in X$, the fibers of E above x and x', E_x and $E_{x'}$, respectively, can be identified by choosing a curve c from x to x' $(x = c(0), x' = c(1))$ and moving each $s_0 \in E_x$ along c to $E_{x'}$ by parallel translation. This identification depends only on the choice of the curve c. Now assume that X is a compact manifold with Riemannian metric g and injectivity radius $\rho > 0$ (see Definition 6.13, p. 166). Then we have a geodesic with respect to the Riemannian metric g as canonical curve, which is uniquely determined by the endpoints x, x', if $\operatorname{dist}(x, x') < \rho$.

DEFINITION 6.24. Let X be a compact Riemannian manifold with injectivity radius $\rho > 0$ and let $E \to X$ be a vector bundle equipped with a connection ∇^E.

For $x, x' \in X$, with $\mathrm{dist}(x, x') < \rho$, we denote the parallel translation relative to ∇^E along the unique geodesic from x' to x with minimal length $\mathrm{dist}(x, x')$ by

$$\tau^E_{x,x'} : E_{x'} \longrightarrow E_x.$$

REMARK 6.25. a) In the preceding definition, we obtain parallel translation on a Riemannian manifold from a connection. Conversely, we can regain a connection from parallel translation by the recipe of (6.9) at the beginning of this section.
b) To explain the name *connection*, we refer to our Section 16.2, p.366ff, where we develop the topic *Connections and Curvature*. In particular, we refer to Figure 16.1, p.367. More precisely (replacing P by E and M by X in the figure), consider the tangent space T_pE at the point $p \in E$ to the total space E of a vector bundle $\pi : E \to X$. Inside T_pE, there is a distinguished subspace, namely the tangent space to the fiber E_x containing p ($x = \pi(p)$). This space is called the *vertical* space V_p. However, there is no distinguished *horizontal* space H_p complementary to V_p, i.e., satisfying $T_pE = V_p \oplus H_p$. If we have a covariant derivative ∇^E, however, we can parallely transport p for each $v \in T_xX$ along a curve $c_v(t)$ with $c_v(0) = x, \dot{c}_v(0) = v$. Thus, for each v, we obtain a curve $p_v(t)$ in E. The subspace of T_pE spanned by all tangent vectors to E at p of the form $\frac{d}{dt}p_v(t)|_{t=0}$ then is a suitable choice of a horizontal space H_p. In this manner, one obtains a rule how the fibers in neighboring points are *connected* with each other.

DEFINITION 6.26. Let $\pi : E \to X$ be a real or complex vector bundle on the differentiable manifold X with Euclidean, respectively Hermitian bundle metric $\langle \cdot, \cdot \rangle$. A connection ∇ on E is called **metric** (also **Riemannian**, respectively **Hermitian**), if

$$d\langle s, r\rangle = \langle \nabla s, r\rangle + \langle s, \nabla r\rangle \quad \text{for all } s, r \in C^\infty(X, E). \tag{6.16}$$

A metric connection thus has to be compatible with the metric. To read (6.16) correctly, you notice that $\langle s, r\rangle$ is a smooth function on X, whence $d\langle s, r\rangle \in C^\infty(T^*X)$. Applying the left side of (6.16) to a tangent vector field $v \in C^\infty(TX)$ yields a smooth function on X. Similarly, each of the two terms on the right side of (6.16) yields a function when the vector field v is inserted into $\nabla s, \nabla r \in \mathrm{Hom}(TX, E)$.

Connections on the Tangent Bundle. Connections on the tangent bundle TX are particularly important:

DEFINITION 6.27. Let ∇ be a connection on the tangent bundle TX of a differentiable manifold X.
a) A curve $c :\to X$ is called **autoparallel** or **geodesic** with respect to ∇, if $\nabla_{\dot{c}} c \equiv 0$, i.e. if the tangent field $\dot{c}$ of c is parallel along c.
b) The **torsion tensor** of ∇ is defined as

$$T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y] \quad (X, Y \in C^\infty(TX)).$$

c) The connection ∇ is called **torsion free**, if $T \equiv 0$.

A classical result of Riemannian geometry (proved in [**Jo**, Section 3.3] and many other places) is now

THEOREM 6.28. *On each Riemannian manifold X, there is precisely one metric and torsion free connection θ (on TX). It is determined by the formula*

$$\langle \theta_X Y, Z\rangle = \frac{1}{2}\Big(X\langle Y,Z\rangle - Z\langle X,Y\rangle + Y\langle Z,X\rangle - \langle X,[Y,Z]\rangle + \langle Z,[X,Y]\rangle + \langle Y,[Z,X]\rangle \Big). \tag{6.17}$$

The connection θ determined by (6.17) is called the **Levi-Civita connection** of X. In the sequel, θ (or θ^g) will always denote the Levi-Civita connection.

Clifford Modules and Operators of Dirac Type. There are many different concepts of a Dirac operator in global analysis: classical and twisted Dirac operators on spin manifolds; operators of Dirac type with a square with scalar principal symbol; generalized (or compatible) Dirac operators defined by arbitrary (or compatible) connections on bundles of Clifford modules over Riemannian manifolds; full and split (odd-parity) Dirac operators; boundary Dirac operators; etc. The concepts depend on various geometrical features like dimension parity, orientation and chirality, almost complex structure, and suitable boundary. Each definition has its own merits and range of application and we will return to them.

Let X be a compact smooth oriented manifold (with or without boundary) with Riemannian metric g. Let $\dim X = n$. Let S be a complex vector bundle over X of Clifford modules; i.e., we have a representation

$$\mathbf{c} : \mathrm{C}\ell(X) \longrightarrow \mathrm{Hom}(S,S)$$

with

$$\mathbf{c}(v)^2 = -\|v\|^2 \,\mathrm{Id}_{S_x} \qquad \text{for } v \in TX_x \text{ and } x \in X. \tag{6.18}$$

Recall that the Clifford bundle $\mathrm{C}\ell(X)$ consists of the Clifford algebras $\mathrm{C}\ell(TX_x, g_x)$, $x \in X$, which are associative algebras with unit generated by TX_x and subject to the relation $v\cdot w + w\cdot v = -2g_x(v,w)$. We shall call **c left Clifford multiplication** and occasionally write

$$\mathbf{c} : C^\infty(X, TX \otimes S) \longrightarrow C^\infty(X,S).$$

We may assume that S is equipped with a Hermitian metric which makes Clifford multiplication skew-adjoint, i.e. $\mathbf{c}(v)^* = -\mathbf{c}(v)$ for all $v \in TX_x$.

DEFINITION 6.29. A connection $\nabla^S : C^\infty(X,S) \longrightarrow C^\infty(X, T^*X \otimes S)$ for S will be called **compatible** with the Clifford module structure of S, if $\nabla^S \mathbf{c} = 0$, i.e., ∇^S is a **module derivation** with

$$(\nabla^S \mathbf{c})(v)(s) = \nabla^S(\mathbf{c}(v)s) - \mathbf{c}(\theta^g v)s - \mathbf{c}(v)(\nabla^S s) = 0, \tag{6.19}$$

where θ^g denotes the Levi-Civita connection on X.

Patching locally constructed spin connections together proves

THEOREM 6.30 (Branson, Gilkey [**BrGi**]). *There exist compatible connections on S which extend the Riemannian connection on X to S.*

DEFINITION 6.31. Let $D : C^\infty(X,S) \longrightarrow C^\infty(X,S)$ be a linear differential operator of first order operating on smooth sections of a $\mathrm{C}\ell(X)$-module S.
a) We call D an **operator of Dirac type**, if it can be written as

$$D = \mathbf{c} \circ J \circ \nabla^S, \tag{6.20}$$

where ∇^S is a (not necessarily compatible) connection and $J : C^\infty(X; T^*X \otimes S) \cong C^\infty(X; TX \otimes S)$ denotes the canonical identification. In terms of a local orthonormal frame $v_1, \ldots, v_n$ of TX we then have

$$Ds|_x = \sum_{\nu=1}^{n} \mathbf{c}(v_\nu)(\nabla^S_{v_\nu} s)|_x. \tag{6.21}$$

b) We call D a **(compatible) Dirac operator**, if it can be written as $D = \mathbf{c} \circ J \circ \nabla^S$, where ∇^S is a compatible connection.

Note. The Dolbeault complex (to be studied extensively in Parts III-IV) is an example of a non-compatible Dirac operator.

As we shall see below, all (total) Dirac operators are elliptic and formally self-adjoint with a *Green's formula*

$$(Ds, s') - (s, Ds') = -\int_Y G(y)\langle s|_Y, s'|_Y\rangle, \tag{6.22}$$

where $G(y) := \mathbf{c}(\mathbf{n})$ denotes Clifford multiplication by the inward unit tangent vector over the (possibly empty) boundary $Y = \partial X$.

For even n the splitting $C\ell(X) = C\ell^+(X) \oplus C\ell^-(X)$ of the Clifford bundles induces a corresponding splitting of $S = S^+ \oplus S^-$ and a *chiral decomposition*

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

The **partial (chiral) Dirac operators** $D^\pm$ are especially interesting in index theory since they are also elliptic, but in general not self-adjoint and provide interesting integer-valued invariants as their indices. In this book, we shall come back to Dirac type operators incessantly.

7. Differential Operators on Manifolds and Symbols

We shall define linear differential operators on differentiable manifolds, acting on sections of complex vector bundles.

Our Data. Let X be a C^∞ n-manifold. Let $\pi_E : E \to X$ be a C^∞ complex vector bundle over X of fiber-dimension N, i.e., a family of N-dimensional complex vector spaces E_x, with the parameter x ranging over X, whose disjoint union carries C^∞ structure in a natural way; see Appendix, Exercise B.11b, p. 683. We denote the linear space of C^∞ sections of E by $C^\infty(X, E)$ or shortly $C^\infty(E) := \{s : X \to E : \pi_E \circ s = \mathrm{id}_X\}$. Unless otherwise stated, we remain in the C^∞ category. For example, let E denote the *trivial* product bundle $X \times \mathbb{C}^N$, where we write $\mathbb{C}^N_X$ when we wish to emphasize the bundle point of view. Then $C^\infty(E)$ denotes the space of $\mathbb{C}^N$-valued C^∞ functions on X. Strictly speaking when $N = 1$, we should write $C^\infty(\mathbb{C}_X)$ instead of $C^\infty(X)$, but when the context is clear this is not necessary.

Since manifolds and vector bundles can be locally described in terms of coordinate functions, the following definition makes sense:

Definition 6.32. Let $\pi_E : E \to X$, $\pi_F : F \to X$ be two C^∞ complex vector bundle over X of fiber-dimension N and M. A **linear differential operator** P of integer **order** $k \geq 0$ from E to F is a linear mapping $P : C^\infty(E) \to C^\infty(F)$ that satisfies the following conditions:
(i) For any element $s \in C^\infty(E)$, the support of Ps is contained in the support of s.

(ii) Via local coordinates, the mapping P can be represented as a vectorial differential operator (see before Exercise 5.7, p. 139) in which there derivatives of order $\leq k$ appear, but not of order $> k$. More precisely, for all coordinate neighborhoods $U \subset X$ and trivializations $\tau_E : E|_U \cong U \times \mathbb{C}^N$ and $\tau_F : F|_U \cong U \times \mathbb{C}^M$, the mapping P can be locally expressed in the form

$$P[s](x) = \tau_F^{-1}\big(\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha (\tau_E \circ s)|_x \big),\ x \in U, \tag{6.23}$$

where α ranges over all multi-indices $(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+ \times \overset{n}{\cdots} \times \mathbb{Z}_+$ with $|\alpha| := \alpha_1 + \cdots + \alpha_n$, $a_\alpha \in C^\infty(X, \mathrm{Hom}(\mathbb{C}^N, \mathbb{C}^M))$, and

$$D^\alpha := i^{-|\alpha|} \frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \cdots \partial {x_n}^{\alpha_n}},$$

where $x_1, \ldots, x_n$ denote the chosen local coordinates on U.

NOTE (1). The splitting of our definition in two parts is for convenience only. Actually, condition (i) makes it sufficient to check condition (ii) solely for s with compact support contained in the coordinate patch U. That would be a weaker condition. However, to argue along that line excludes analytic differential operators in the case that they should be applied only on analytic sections. Mathematically, the splitting is redundant: (i) can be deduced from (ii) as in Exercise 5.2, p.136. For the deduction of (ii) from (i), see Remark 5.3, p.136.

NOTE (2). The reason for inclusion of the factor $i^{-|\alpha|}$ is explained in Remark 5.1, p.135. There we also emphasize why this tradition, originating from mathematicians working in analysis, is a bit unfortunate for topologists, geometers and physicists when they are interested in differential operators of first order on *real* vector bundles.

We write $P \in \mathrm{Diff}_k(E, F)$ for linear differential operators of order $\leq k$. A **real** differential operator of order k is defined similarly with $\mathbb{C}$ replaced by $\mathbb{R}$.

Symbolic calculus. *Fourier analysis* makes it natural to replace the differential expressions by multiplication, e.g., to replace the differentiation D^α by the monomial ξ^α with $\xi \in \mathbb{R}^n$. We can do that for all differential expressions in (6.23) and obtain a polynomial in ξ at each $x_0 \in X$. That polynomial depends heavily on the choice of coordinates. To reduce that dependence, we can define the polynomial in cotangent variables, i.e., choosing $\xi \in \mathring{T}^*_{x_0} X = T^*_{x_0} X \setminus \{0\}$. The expression defined in that way is called the *total* or *complete symbol*.

Workers in analysis use the complete symbol often. It is intimately related to the concepts of *quantization*, see below. However, roughly speaking, it carries too much information. Consequently, it is not defined independently of the choice of coordinates, in general. In the geometric tradition of, e.g., classifying conic sections or more advanced curves and surfaces, it is obvious that one has to select the relevant information. For index theory, that is the principal symbol, i.e., the leading term of the total symbol. Fortunately, it will turn out that the definition of the principal symbol does not depend on the choice of local coordinates. Whence, it is a genuinely *geometric* object.

Let $\pi : \mathring{T}^* X \to X$ denote the *dotted* cotangent bundle of X, i.e., the bundle with the symplectic cone $\mathring{T}^* X := T^* X \setminus X$ as total space. Let $\pi^* E \to \mathring{T}^* X$ and $\pi^* F \to \mathring{T}^* X$ denote the pull-backs of the vector bundles $\pi_E : E \to X$ and

$\pi_F : F \to X$ via π. We then have a vector bundle $\text{Hom}(\pi^*E, \pi^*F) \to \mathring{T}^*X$. We define a section $\sigma(P)$ of this bundle for each linear differential operator P, acting between sections of E and F, i.e., $\sigma \in C^\infty(\text{Hom}(\pi^*E, \pi^*F))$. That is our view of the *principal symbol* (similarly, e.g., [**Wells**, p.115-116]).

DEFINITION 6.33. For a linear differential operator $P : C^\infty(X, E) \to C^\infty(X; F)$ of order $k \geq 0$ we define the **symbol** (also called the **principal** or the **leading symbol** to emphasize that only the principal or leading terms of P enter into the definition) via the formula

$$\sigma(P)\,(x, \xi_x)\,(e) = \frac{i^k}{k!} P\left(\varphi^k g\right)_x, \tag{6.24}$$

where $x \in X$; $\xi_x \in T_x^*X$, $\xi_x \neq 0$; $e \in E_x$; φ is a real-valued C^∞ function on X with $d\varphi_x = \xi_x$ and $\varphi(x) = 0$; and $g \in C^\infty(E)$ with $g(x) = e$. (Such choices are always possible.)

For $\xi_x \in T_x^*X$, note that the fiber $(\pi^*E)_{\xi_x}$ may be identified with E_x and we will do so; we write the identification as $(\pi^*E)_{\xi_x} \sim E_x$.

NOTE. Most workers in analysis prefer to define the principal symbol in local coordinates, i.e., taking the *characteristic polynomial* when the operator is given by (6.23), see the following Exercise 6.34b. For the further treatment of elliptic differential operators, this definition is somewhat opaque, since in that way the principal symbol is defined in a piecewise unrelated fashion. Then, it may look like a mystery that the locally given *characteristic polynomials* collectively define a bundle homomorphism and that the principal symbol of a differential operator on a manifold admits a geometric interpretation. Of course, basically it does not make a big difference whether we introduce the principal symbol without coordinates or with coordinates, as the following exercise shows. Even in our preferred coordinate-free first way one has to use coordinates to show that (6.24) in Definition 6.33 is well defined and yields a smooth section of the homorphism bundle.

EXERCISE 6.34. a) Show that the point $\sigma(P)\,(x, \xi_x)\,(e) \in F_x$ in Equation (6.24) in the preceding Definition is well defined, i.e., it does not depend on the choices of functions and sections representing cotangent vectors and fiber points. Moreover, show that the definition yields a smooth section of the homomorphism bundle.
b) Choose a coordinate patch $U \subset X$ and local coordinates $x^1, \ldots, x^n$ on U. Write $\xi_x = \xi_1 dx^1 + \cdots + \xi_n dx^n$ with $(\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$, and write the product $\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$ as ξ^α. For $e \in (\pi^*E)_{\xi_x} \sim E_x$, define

$$\begin{aligned}\sigma(P)(\xi_x)(e) :&= \tau_F^{-1}\Big(x, \sum_{|\alpha|=k} a^\alpha(x)\big(\tau_E(x,e)\big)\xi^\alpha\Big) \in F_x \sim (\pi^*F)_{\xi_x} \\ &= \sum_{|\alpha|=k} \tau_F^{-1}\Big(x, a^{(\alpha_1,\ldots,\alpha_n)}(x)\big(\tau_E(x,e)\big)\Big)\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n},\end{aligned}$$

where τ_E, τ_F are local bundle trivializations like in (6.23). Making identifications $(\pi^*E)_{\xi_x} \sim E_x \sim \mathbb{C}^N$ (via τ_E) and $(\pi^*F)_{\xi_x} \sim F_x \sim \mathbb{C}^M$ (via τ_F), we can write this more transparently as

$$\sigma(P)(\xi_x) = \sum_{|\alpha|=k} a^{(\alpha_1,\ldots,\alpha_n)}(x)\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n} \in \text{Hom}(\mathbb{C}^N, \mathbb{C}^M). \tag{6.25}$$

Show that in this way $\sigma(P) \in C^\infty\big(\mathrm{Hom}(\pi^*E, \pi^*F)\big)$ is well-defined, i.e., independent of the choice of local coordinates and trivializations τ_E and τ_F.
[Hint: You can prove (b) by (a), i.e., writing the mapping in coordinate-free manner; and you can prove (a) by (b), i.e. checking all choices and transformations. In both cases, you have to do some calculations at some point. *There are no free lunches.* The result is *not* obvious. Indeed, if we had summed over α with $|\alpha| = k-1$, the resulting so-called *subprincipal symbol* is not well-defined.]

In this way, we have defined a linear map $\sigma : \mathrm{Diff}_k(E,F) \to \mathrm{Smbl}_k(E,F)$, where

$$\text{(6.26)}\quad \mathrm{Smbl}_k(E,F) := \{\sigma \in \mathrm{Hom}(\pi^*E, \pi^*F) : \sigma(x, \lambda v) = \lambda^k \sigma(x,v) \text{ for all } (x,v) \in \mathring{T}^*X \text{ and } \lambda > 0\}.$$

Exercise 6.35. Show that the sequence of vector spaces

$$0 \to \mathrm{Diff}_{k-1}(E,F) \xrightarrow{j} \mathrm{Diff}_k(E,F) \xrightarrow{\sigma} \mathrm{Smbl}_k(E,F)$$

is exact, where j denotes the natural inclusion.
[Hint: This is clear from the representation in coordinates; see Exercise 6.34b. Replacing $\mathrm{Smbl}_k(E,F)$ by the subspace of *polynomial* symbols (see [**Pal65**, p.63]), one gets a surjective symbol map, and the exact sequence may be extended by zero on the right.]

Exercise 6.36. Show that for $P \in \mathrm{Diff}_k(E,F)$ and $Q \in \mathrm{Diff}_j(F,G)$, the operator QP is in $\mathrm{Diff}_{k+j}(E,G)$ with $\sigma(QP) = \sigma(Q) \circ \sigma(P)$.
[Hint: Carry out the proof using the chain rule first for the local vector-valued differential operators (see p. 139), and then generalize.]

Formal Adjoints. In the following, we assume that X is compact, oriented, closed (i.e., without boundary) and Riemannian, and that the vector bundle E is equipped with a Hermitian metric; i.e., each fiber E_x has a non-degenerate, conjugate-symmetric bilinear form $(\cdot\cdot, \cdot\cdot)_{E_x}$ which is C^∞ in the sense that $(e_1, e_2)_E \in C^\infty(X)$ for any two sections $e_1, e_2 \in C^\infty(E)$. Whence, we can form the integral $\int_X (e_1, e_2)_E$, obtaining a Hermitian bilinear form on the vector space $C^\infty(E)$. Assume that the vector bundle F is also given a Hermitian metric.

Definition 6.37. In analogy with Hilbert space theory (see Chapter 2), we say that two operators $P \in \mathrm{Diff}_k(E,F)$ and $P^* \in \mathrm{Diff}_k(F,E)$ are **formally adjoint** (or **formal adjoints**), if $\int_X (Pe, f)_F = \int_X (e, P^*f)_E$ for all sections $e \in C^\infty(E)$ and $f \in C^\infty(F)$.

Note. The definition extends to a compact manifold X with smooth boundary $\Sigma = \partial X$. In that case we require the symmetry condition for sections e, f with support in the interior $\mathring{X} := X \setminus \Sigma$ of X.

Exercise 6.38. Show:
a) There is at most one formally adjoint differential operator P^* for a given P.
b) $(P+Q)^* = P^* + Q^*$, $(P \circ Q)^* = Q^* \circ P^*$, and $P^{**} = P$.

Exercise 6.39. Show that for each $P \in \mathrm{Diff}_k(E,F)$ there is an adjoint differential operator $P^* \in \mathrm{Diff}_k(F,E)$ such that $\sigma(P^*) = \sigma(P)^*$, where $\sigma(P)^* : \pi^*(F) \to \pi^*(E)$ is the homomorphism pointwise adjoint to $\sigma(P)$.
[Hint: 1. Begin with the special case $k = 0$, where $P \in \mathrm{Diff}_0(E,F)$ is given by a

vector bundle homomorphism $h : E \to F$ (i.e., a family of linear maps $h : E_x \to F_x$ parametrized smoothly by $x \in X$). Then $P(e)(x) = h_x(e(x))$ and $\sigma(P)(x,v) = h_x$, where $e \in C^\infty(E)$ and $v \in \mathring{T}_x X)$. Let $h_x^* : F_x \to E_x$ be the linear map adjoint to h_x relative to the Hermitian metrics on E_x and F_x. In this case, for $f \in C^\infty(F)$

$$(P^* f)(x) = h_x^*(f(x)) \quad \text{and} \quad \sigma(P^*)(x,v) = h_x^*.$$

Thus, the statement is proven for this trivial case.
2. For each $\chi \in C^\infty(TX)$ define an operator $P \in \mathrm{Diff}_1(\mathbb{C}_X, \mathbb{C}_X)$ by

$$P\varphi := \tfrac{1}{i}\chi\,[\varphi] = \tfrac{1}{i} d\varphi\,(\chi)\,,$$

where at any point x, $\chi\,[\varphi]\,(x) = d\varphi_x\,(\chi)$ is the derivative of φ in the direction of $\chi|_x$. Then

$$\sigma\,(P)\,(x,v) = v\,(\chi|_x)\,, \text{ where } v \in \mathring{T}_x X = T_x^* X \setminus \{0\}\,.$$

Furthermore, by the Stokes Theorem in the classical Green form (e.g., see [**GP**, p.152 and 182-187] or the Cartan calculus in our Exercise 6.17, p. 169 below, which we have already used in Exercise 5.9 (p. 144) and which also applies here), we have for all $\varphi, \psi \in C^\infty(X)$

$$\int_X \left(\chi\,[\varphi]\,\bar{\psi} + \varphi\overline{\mathrm{div}\,(\psi\chi)}\right) = 0,$$

where $\mathrm{div}\,(\psi\chi) \in C^\infty(X)$ is the divergence of the vector field $\psi\chi$. Thus, $P^*\psi = \frac{1}{i}\,\mathrm{div}(\psi\chi)$. One further checks that

$$(\sigma(P^*)(x,v))\,(z_x) = (\sigma(P)(x,v))\,(z_x) = v\,(\chi|_x)\,z_x,\;\; z_x \in (\mathbb{C}_X)_x\,.$$

Since $v\,(\chi|_x)$ is real and hence self-adjoint as a linear map from $\mathbb{C}$ to $\mathbb{C}$, we have $e(P^*) = e(P)^*$.
3. One may now show that every global differential operator can be constructed from the two preceding types via sums and compositions (locally, this is entirely trivial), and thus Exercise 6.39 reduces to Exercise 6.38.]

Remark 6.40. In contrast to Exercise 6.38, the solution of the preceding Exercise 6.38 is not so trivial, even though we only applied Stokes' theorem in the weak form. Alternatively, one can first assign to each vector-valued differential operator $P : C^\infty(\mathbb{C}_U) \to C^\infty(\mathbb{C}_U)$ given (over an open set $U \subset \mathbb{R}^n$) by

$$Pu = \sum\nolimits_{|\alpha|\le k} a_\alpha\,(x)\,D^\alpha u$$

the operator $P^* : C^\infty(\mathbb{C}_U) \to C^\infty(\mathbb{C}_U)$ given by

$$P^* v := \sum\nolimits_{|\alpha|\le k} D^\alpha(a_\alpha^* v),$$

where $a_\alpha^*(x)$ is the adjoint (conjugate transpose) of the $N \times N$ matrix $a_\alpha(x)$. Using integration-by-parts, it then follows at once that

$$\int_U (Pu, v) = \int_U (u, P^* v) \text{ for all } u, v \in C_0^\infty(\mathbb{C}_U)\,, \text{ where}$$

$$C_0^\infty(\mathbb{C}_U) := \{w \in C^\infty(\mathbb{C}_U) : \mathrm{supp}\,w \text{ is a compact subset of } U\}$$

and where $(\cdot,\cdot)$ is the canonical Hermitian scalar product on $\mathbb{C}^N$. The major work consists of globalizing this result; see [**Pal65**, p.70-75], [**Na**, p.181-183], or [**Wells**, p.117 f].

Elliptic Differential Operators. Definition and Standard Examples. Geometrically defined linear differential operators on closed manifolds (i.e., operators of Laplace type and operators of Dirac type, see below) are marked by two features, namely the algebraic symmetry and regularity of their expression and the finite number of linearly independent solutions. Geometers have always noticed these two features and exploited them. They have been pleased with the easiness and transparency of manipulations, and were enthusiastic when recognising geometric or topological invariants in the dimensions of the solution spaces.

As seen from analysis, these two features are interrelated: algebraic regularity of the principal symbol of a differential operator over a closed manifold implies that the dimension of the kernel of the operator is finite. For that, the key notion is the ellipticity of the principal symbol. That notion will be explained now. In Chapter 9, we deduce the regularity (=smoothness) of the solutions and the Fredholm properties from the ellipticity. Then in Part III, we prove the Atiyah-Singer Index Theorem for elliptic operators on closed manifolds. Roughly speaking, it gives a thorough explanation for the astonishing and previously as somewhat mysterious perceived interrelations between algebraic symmetries of a formal expression (the principal symbol of a variety of geometric defined elliptic operators) and geometric features of the underlying manifold. In Part IV, much wider implications are drawn for law-dimensional topology and gauge-theoretic physics of the same *philosophy*, namely exploiting symmetries and regularities of formal expressions for sensing asymmetries and irregularities of related geometric or physical objects.

DEFINITION 6.41. Let $P : C^\infty(E) \to C^\infty(F)$ be a differential operator with principal symbol $\sigma(P) : T^*(X) \to \mathrm{Hom}(\pi^*E, \pi^*F)$. If $\sigma(P)(\xi_x)$ is an isomorphism for all *nonzero* $\xi_x \in T^*_x X$, then P is called an **elliptic differential operator**.

Whence, a differential operator P is elliptic, if in each chart all of the locally-defined vectorial differential operators are elliptic (see Exercise 5.7); i.e., for each local representation of P over the chart domain U, the **characteristic polynomial of the principal part** (associated to the terms of highest order, i.e., what we call the principal symbol)

$$p_k\,(x,\xi) := \sum_{|\alpha|=k} a_\alpha\,(x)\,\xi^\alpha \tag{6.27}$$

is an invertible linear map for all x in U and all $\xi \in \mathbb{R}^n \setminus \{0\}$.

EXERCISE 6.42. Show P elliptic $\Rightarrow P^*$ elliptic.

We consider some standard examples. Even though $C^\infty(E)$ and $C^\infty(F)$ are not Hilbert spaces (and hence P is not Fredholm), here we take the index of P to be $\dim\ker P - \dim\ker P^*$, where P^* denotes the formal adjoint of P. This is the same as the usual index of a Fredholm extension of P to a suitable Sobolev space, as is explained in Chapters 7 and 9 below.

EXERCISE 6.43. Check the principal symbols of Table 6.1 and derive the ellipticity of all the listed standard operators.
[Hint to (1): This operator (and its counterpart on S^1 for periodic matrices A) was studied in Exercise 2.42, p.41f. Its principal symbol was calculated on p.42 (Note). To (2): Let ∂_{x^i} be a shorthand notation for $\frac{\partial}{\partial x^i}$. Regard the domain $C^\infty(\mathbb{R}^n,\mathbb{C})$ as the space of sections of the trivial bundle $\mathbb{R}^n \times \mathbb{C} \to \mathbb{R}^n$ and write the spaces down where the principal symbol of the Laplacian $\Delta = \partial^2_{x^1} + \cdots + \partial^2_{x^n}$ acts. Do it slowly

TABLE 6.1. The principal symbol of elliptic standard operators

	Operator P	Domain	$\sigma(P)(x,\xi)$
1	ODE system $P := \nabla + A$, $A \in C^\infty(\mathbb{R}, \mathrm{gl}(r,\mathbb{C})), \nabla = \frac{d}{dt} I^r$	$C^\infty(\mathbb{R},\mathbb{C})$	$i\xi I^r$
2	Laplace operator $P := \Delta = \partial^2_{x^1} + \cdots + \partial^2_{x^n}$	$C^\infty(\mathbb{R}^n,\mathbb{C})$	$-(\xi_1^2 + \cdots + \xi_n^2)$
3	Cauchy-Riemann operator $P := \frac{\partial}{\partial \bar z} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$	$C^\infty(\mathbb{R}^2,\mathbb{C})$	$\frac{1}{2} i(\xi_1 + i\xi_2)$
4	Euler operator $P := d + \delta$	$\Omega^\bullet(X), \Omega^{\mathrm{ev}}(X)$	$i(\xi \wedge\ - \xi \llcorner)$
5	Dirac type operator $P := D = \mathbf{c} \circ \nabla^S$	$C^\infty(X,S)$	$\mathbf{c}(i\xi)$
6	Dirac Laplacian $P := D^2$ Riemannian metric g	$C^\infty(X,S)$	$-\Vert\xi\Vert^2_{g(x)}\, I_{S_x}$

with the previous notations: Then you have

$$\begin{array}{cccc} \sigma(\Delta) & : & \mathring{T}^*\mathbb{R}^n & \longrightarrow & \mathrm{Hom}\big(\pi^*(\mathbb{R}^n\times\mathbb{C}), \pi^*(\mathbb{R}^n\times\mathbb{C})\big) \\ & & (x, \xi_1 dx^1 + \cdots + \xi_n dx^n) & \mapsto & -(\xi_1^2 + \cdots + \xi_n^2) \in \mathrm{End}\left(\pi^*(\mathbb{R}^n\times\mathbb{C})_{x,\xi}\right), \end{array}$$

regarded as multiplication on $\pi^*(\mathbb{R}^n\times\mathbb{C})_{x,\xi} \cong \mathbb{C}$ by a real.
To (3): The Cauchy-Riemann operator is a first-order operator on the same space of sections as in Example (2), but with a different symbol

$$\sigma(\frac{\partial}{\partial \bar z})(x, \xi_1 dx^1 + \xi_2 dx^2) = \frac{1}{2} i(\xi_1 + i\xi_2),$$

regarded as complex multiplication on $\pi^*(\mathbb{R}^n\times\mathbb{C})_{x,\xi} \cong \mathbb{C}$. Similarly, you have the complex differentiation operator $\frac{\partial}{\partial \bar z} = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y})$. Note that $\Delta = 4\frac{\partial^2}{\partial \bar z \partial z}$ and at $\xi_1 dx^1 + \xi_2 dx^2$ you confirm

$$\sigma(\Delta)(x, \xi_1 dx^1 + \xi_2 dx^2) = 4\frac{1}{2} i(\xi_1 + i\xi_2)\frac{1}{2} i(\xi_1 - i\xi_2) = -(\xi_1^2 + \xi_2^2).$$

There is a type of exterior derivative on $\Omega^{0,0}(\mathbb{R}^2,\mathbb{C}) := C^\infty(\mathbb{R}^2,\mathbb{C})$, namely the *Dolbeaut operator*

$$\begin{array}{cccc} \bar\partial : & \Omega^{0,0}(\mathbb{R}^2,\mathbb{C}) & \longrightarrow & \Omega^{0,1}(\mathbb{R}^2,\mathbb{C}) := \{h d\bar z : h \in C^\infty(\mathbb{R}^2,\mathbb{C})\}, \\ & f & \mapsto & \frac{\partial f}{\partial \bar z} d\bar z \end{array}.$$

Here $\Omega^{0,1}(\mathbb{R}^2,\mathbb{C})$ is called the space of *complex forms of type* $(0,1)$. For a compact Riemann surface S a strictly analogous operator $\bar\partial : \Omega^{0,0}(S,\mathbb{C}) \longrightarrow \Omega^{0,1}(S,\mathbb{C})$ can be defined. Its index yields the classical Riemann-Roch Theorem. For any higher-dimensional compact, complex manifold X of $\dim_{\mathbb{C}} X = m$, there is a Dolbeaut operator complex $\bar\partial : \Omega^{0,k}(X,\mathbb{C}) \to \Omega^{0,k+1}(X,\mathbb{C})$, $k = 0, 1, \ldots m$ which can be *rolled up* to give an elliptic operator

$$\bar\partial : \bigoplus_{k \text{ even}}^{m} \Omega^{0,k}(X,\mathbb{C}) \to \bigoplus_{k \text{ odd}}^{m} \Omega^{0,k}(X,\mathbb{C}).$$

The Hirzebruch-Riemann-Roch Theorem expresses index $\bar\partial$, called the *arithmetic genus* of X in terms of so-called Chern numbers. We shall explain all that below in Section 14.7, p.320ff, and present a comprehensive generalization in Section 18.6, see in particular p.583ff.
To (4): Recall from Section 6.5

- the definition of the bundle $\Lambda^k(X) \to X$ of complex exterior k-covectors over the compact, orientable C^∞ Riemannian n-manifold X with metric tensor g;
- the wedge product $\wedge$, the interior product $\llcorner$ and the (Hodge) star operator $*$ on $\Lambda^\bullet(X)$;
- the space $\Omega^k(X) := C^\infty(\Lambda(X))$ of C^∞ sections of $\Lambda^k(X)$, namely the space of $\mathbb{C}$-valued k-forms on X;
- the exterior derivative $d : \Omega^k(X) \to \Omega^{k+1}(X)$ and the codifferential $\delta : \Omega^{k+1}(X) \to \Omega^k(X)$ which is the formal adjoint of d; i.e.,
$$(d\alpha, gb) = \int_X \langle d\alpha, \beta\rangle_g \nu_g = \int_X \langle \alpha, \delta\beta\rangle_g \nu_g = (\alpha, \delta\beta),$$
where ν_g denotes the volume form and $\langle \cdot, \cdot\rangle_g$ denotes the inner product on $\Lambda^\bullet(X)$ induced by g; and
- set again $\Omega^\bullet(X) = \oplus_{k=0}^n \Omega^k(X)$.

Next, you bring δ in a manageable form. Try (here just for fun, but later highly usable)
$$\delta = -(-1)^{nk} *_{n-k} d *_{k+1} : \Omega^{k+1}(X) \to \Omega^k(X).$$
Then you are ready to calculate the principal symbol of the first order operator $d + \delta : \Omega^\bullet(X) \to \Omega^\bullet(X)$. One way to do it would be the following exercise in the manipulation of exterior forms (do it):

Choose a local system of coordinates $x = (x^1, \ldots, x^n)$ on X. Then $\{dx^1, \ldots, dx^n\}$ is the corresponding local frame for T^*X and $\{dx^\alpha := dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_k}\}_{\alpha \in J}$ a local frame for the bundle $\Lambda^\bullet(X)$, where J denotes the set of all coordinate selections $\alpha = (\alpha_1, \ldots, \alpha_k)$ with $k \le n$ and $1 \le \alpha_1 < \cdots < \alpha_k \le n$). For given $x \in X$ and $e \in \Lambda^\bullet(X)_x$ expand $e = \omega(x)$ with $\omega = \sum_{\alpha \in J} f_\alpha dx^\alpha$ close to x with smooth functions $\{f_\alpha\}_{\alpha \in J}$. Check that
$$d\omega = d\Big(\sum_{\alpha \in J} f_\alpha dx^\alpha\Big) = \sum_{\alpha \in J} \sum_{j=1}^n \frac{\partial}{\partial x^j} f_\alpha dx^j \wedge dx^\alpha .$$
For $\xi = \xi_1 dx^1 + \cdots + \xi_n dx^n$ you will find $\sigma(d)(x,\xi)(e) = i\xi \wedge e$. Then $\sigma(\delta)(x,\xi)(e) = -i\xi \llcorner e$ since $\delta = d^*$ and the principal symbol of the adjoint operator is the adjoint of the principal symbol (as you have shown in Exercise 6.39, p.181). Check that the mapping $\xi\llcorner$ is the dual of $\xi\wedge$ by applying the mappings to an orthonormal local frame. In this way you obtain
$$\begin{array}{rcccc} \sigma(d+\delta) & : & \mathring{T}^*X & \longrightarrow & \mathrm{Hom}\big(\pi^*\Lambda^\bullet(X), \pi^*\Lambda^\bullet(X)\big) \\ & & (x,\xi) & \mapsto & (\pi^*\Lambda^\bullet(X))_{(x,\xi)} \ni e \mapsto i(\xi \wedge e - \xi \llcorner e). \end{array}$$
Note that $\sigma(d+\delta)_\xi$ is invertible ($\xi \neq 0$), and hence $d+\delta$ is elliptic. To see that, please check $\sigma(d+\delta)_\xi \circ \sigma(d+\delta)_\xi = -\|\xi\|^2 I_{\pi^*\Lambda^\bullet(X)_\xi} \big(= \sigma((d+\delta)^2)_\xi\big)$. You can find the details below in Section 14.4, p.309f. The second order operator $(d+\delta)^2 = d\delta + \delta d$ is called the Beltrami Laplacian. Since $d + \delta$ is formally self-adjoint, its index is zero. By definition, elements of $\ker(d+\delta)$ are known as harmonic forms. Hodge theory tells us that the algebra of harmonic forms, say $\mathcal{H}^\bullet(X)$, is isomorphic to the cohomology algebra $H^*(X;\mathbb{C})$ where wedge product of harmonic forms corresponds to cup product in $H^*(x;\mathbb{C})$ (see [**GH**, p.43f and p.59-61] for a quick orientation). To get something more interesting, you must restrict $d+\delta$ to the even or to the odd forms. The restricted operators are still elliptic with symbols that are inverses

modulo a factor of $-\|\xi\|^2$. Their index is not necessarily zero. Indeed, restricting to the even forms yields an operator with the Euler characteristic of X as index. That will be explained in the mentioned Section 14.4.
To (5): The claim follows immediately from the definition of Dirac type operators (see Definition 6.31, p.177 and the elemenary properties of Clifford multiplication. You can use either the global definition of (6.20) or the local definition of (6.21). Ellipticity follows at once. Also it follows that the principal symbol is symmetric, if the defining connection is metric (i.e., compatible with the metric structure of the bundle, as explained in Definition 6.26, p.176).

Actually, the standard operators $\nabla + A, \bar{\partial}$ and $d+\delta$ of (1), (3) and (4) are all special instances of Dirac type operators. You could have obtained your results for these operators by clever Clifford multiplication alone! Indeed, take for instance $S := \Lambda^\bullet(T^*X)$ and let $\mathbf{c}$ be the extension of $\mathbf{c}(\xi)(s) = \xi \wedge s - \xi \llcorner s$ for $\xi \in T_x^*X \subset C\ell(T_x^*X, g_x)$ and $s \in \Lambda^\bullet(T_x^*X)$ (here we identify TX and T^*X). The extension is guaranteed by the fact that $\mathbf{c}(\xi)^2 = -g_x(\xi,\xi)\,\mathrm{I}$.
To (6): Apply Exercise 6.36 and simple Clifford multiplication. There is a zoo of Laplacians, depending on the assumptions about the underlying manifold X and the bundle S and the connection ∇^S. Special cases (the *connection Laplacian* and the *Hodge Laplacian*) are discussed in our Section 16.6, p.405ff.]

What is the meaning of the geometric description of the symbol of a differential, and in particular elliptic operator? This is the basic question which will concern us from now on. Recall from Remark 5.3 the geometric characterization of differential operators as support compressing. Unfortunately, that property can not be quantified or graded. Conversely, the symbol can express both quantitative and qualitative aspects. That for a first vague answer.

For a second vague answer, imagine an electron microscope aimed at a point x of our manifold. Under enlargement, both the neighborhood of x and the vector bundle above it become linear, of course, while the hole in the cotangent bundle becomes large like a sphere. With even greater magnification, we see, instead of a highly complicated differential operator on infinite dimensional spaces of cross-sections (the object of study of functional analysis), families of maps on spheres S^{n-1} into the general linear group $GL(N,\mathbb{C})$ (the object of study of linear algebra and of topology). The Atiyah-Singer Index Formula is (crudely) a manifestation of this change of the planes of investigation.

The term *symbol* suggests the *symbolic method* of Oliver Heaviside, which owes its power to the transition between the plane of operator theory and the plane of polynomial algebra. For this reason it is characterized in [**CH**, II, p.187/518] pointedly as the *separation of the algebraic part from the mathematical-conceptual part.*

Initially, the relationship between the study of operators and the study of their symbols was penetrated more deeply not by mathematicians, but by physicists with the ideas of Erwin Schrödinger and Paul Adrien Maurice Dirac on the *quantization* of classical mechanical systems according to which classical mechanics deals with symbols and quantum mechanics with operators (for quantum mechanical systems with spin, these are differential operators for nontrivial vector bundles). Thus first one considers a problem of classical physics (mechanics, electrodynamics), establishes the classical Hamiltonian function, changes to position and momentum coordinates, and obtains, e.g., for the harmonic oscillator the function

$$h(x,p) = \frac{1}{2m}p^2 + \frac{k}{2}x^2.$$

(**Interpretation**: Consider on the real axis the motion of a particle of mass m with the kinetic energy $\frac{m}{2}\dot{x}^2 = \frac{1}{2m}p^2$ where $x(t)$ is the location at time t and $p = m\dot{x}$ is the momentum. Then one supposes that, particle moves in a force field whose potential energy is $\frac{k}{2}x^2$.) According to the quantization rule, we now choose a suitable Hilbert space, as rule an L^2 (or Sobolev space as in Chapter 7 below), replace the position coordinate x by multiplication by x and the momentum coordinate by the differential operator $\frac{\hbar}{i}\frac{d}{dx}$. We then obtain (in the Schrödinger representation) the operator H with

$$Hf = -\frac{\hbar^2}{2m}\frac{d^2 f}{dx^2} + \frac{k}{2}x^2 f,$$

with the 2-symbol

$$\sigma_2\left(H\right)\left(x,\xi\right) = \frac{1}{2m}\xi^2.$$

The *total* symbol (which is not invariant and therefore not defined here) is

$$\sigma_{2,1,0}\left(H\right)\left(x,\xi\right) = \frac{1}{2m}\xi^2 + \frac{k}{2}x^2.$$

So much for this simple example; see e.g. [**He**, II, p.257-293], who strongly advocates this point of view. We shall return to this topic in our Part IV.

8. Manifolds with Boundary

Instead of modeling a manifold locally on open subsets of E, we can also work with charts that map open subsets of the topological space X homeomorphically onto open subsets of the half-space $\mathbb{R}^n_+ := \{(x_1, ..., x_n) \in \mathbb{R}^n : x_n \geq 0\}$ such that the *coordinate changes* are again C^∞. In this way, we introduce the concept of a C^∞ **manifold with boundary**, in the same way as we have done above for (unbounded) manifolds. We call $x \in X$ an **interior point**, if it has a neighborhood which is mapped by a chart onto an open subset of $\mathbb{R}^n$ (i.e., contained in the interior of $\mathbb{R}^n_+$). On the other hand, if there is a chart mapping x to a point on the boundary of $\mathbb{R}^n_+$, then x is called a **boundary point** of X; we write ∂X for the set of all boundary points. As examples, the closed solid sphere or torus are three-dimensional C^∞ manifolds with boundaries being the 2-sphere or 2-torus, respectively.

Exercise 6.44. Let X be a C^∞ manifold with boundary.
a) Carry the concepts $C^\infty(X)$, TX, T^*X, orientation, Riemannian metric, etc., over to this case.
b) Construct a C^∞ atlas for ∂X from a C^∞ atlas for X, showing ∂X is a C^∞ manifold of dimension $n-1$, when $\dim X = n$. Show that $\partial\partial X = \emptyset$ and that ∂X inherits Riemannian structure and orientation from those on X.

For short, we write $Y = \partial X$ here. Since each C^∞ path in Y is a C^∞ path in X, we have a canonical embedding of TY into $TX|_Y$. Over Y, we have the following diagram of tangent and cotangent bundles

$$\begin{array}{ccc} TY & \cong & T^*Y \\ \cap & & \cap \\ \left(TX\right)|_Y & \cong & \left(T^*X\right)|_Y \end{array}$$

the left inclusion is canonical; the other isomorphisms and the right inclusion depend on the choice of a Riemannian metric.

EXERCISE 6.45. With the help of a Riemannian metric $(\cdot,\cdot)$, define a *normal field* $\nu \in C^\infty(TX|_Y)$ such that $(\nu(y), w) = 0$ and $(\nu(y), \nu(y)) = 1$ for all $y \in Y$ and $w \in T_yY$, as depicted in Figure 6.4. Show that there are two such normal fields, and characterize the *inner* one via the condition $d\varphi\,(v(y)) \geq 0$ for all real-valued $\varphi \in C^\infty(X)$ which are positive except at y. Characterize the *dual normal field* $v^* \in C^\infty\left((T^*X)|_Y\right)$.

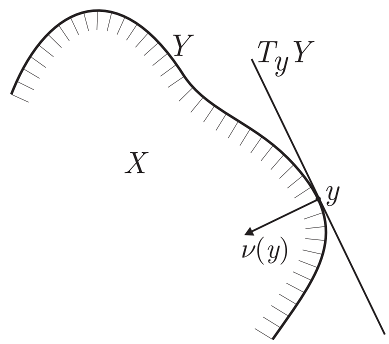


FIGURE 6.4. The inner normal field over the boundary ∂X of a Riemannian manifold X

EXERCISE 6.46. Let X_1 and X_2 be C^∞ manifolds with boundaries, and let $f : Y_1 \to Y_2$ be a diffeomorphism of their boundaries. Show that one can construct a C^∞ manifold $X_1 \cup_f X_2$ in a canonical way by identifying the boundaries of X_1 and X_2 via f as in Figure 6.5.

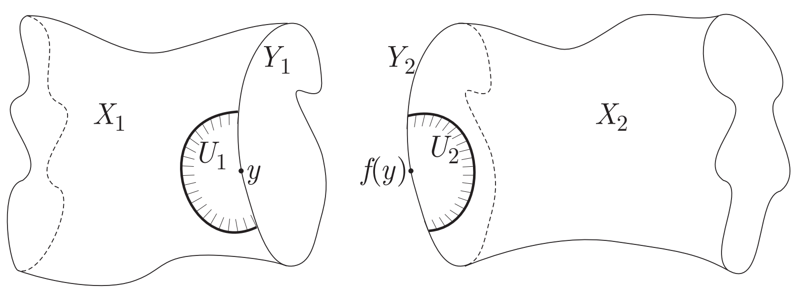


FIGURE 6.5. Gluing two manifolds X_1, X_2 along their boundaries

[Hint: Form the disjoint union of $X_1 - Y_1$, $X_2 - Y_2$, and $\{(y, f\,(y)) : y \in Y_1\}$. It is clear what a coordinate neighborhood of $x \in X_i - Y_i$ $(i = 1, 2)$ should be. For $x = (y, f(y))$, choose a neighborhood U_1 of y in X_1, neighborhood U_2 of $f(y)$ in X_2 with $f\,(U_1 \cap Y_1) = U_2 \cap Y_2$. Then U_1 and U_2 form a neighborhood of $(y, f(y))$, see Figure 6.6. By this recipe, define an atlas for $X_1 \cup_f X_2$ from the C^∞ atlases for X_1 and X_2, such that $X_1 \cup_f X_2$ becomes a topological manifold. It is not entirely easy to prove that the coordinate changes are C^∞. Without loss of generality, assume $X_1 = X_2 =: X$ and $f = \mathrm{Id}$; in our applications, we always have this situation. In order to avoid the difficulties with the corners originating from the way in which the charts are joined, choose a Riemannian metric and extend the above-mentioned normal vector field (Exercise 6.45) to a neighborhood of $\partial X := Y$. The integral curves of the vector field provide a diffeomorphism (*collar*) of $Y \times [0, 1]$ with a neighborhood of Y in X. The differentiable *doubling* of $Y \times [0, 1]$ along $Y \times \{0\}$ is trivial; see also [**BJ**, 13.5-13.11].]

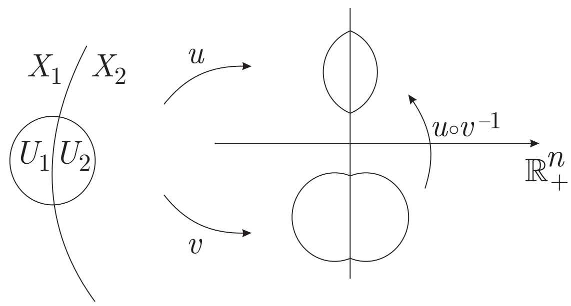


FIGURE 6.6. Defining an atlas for $X_1 \cup_f X_2$ from charts for X_1, X_2

REMARK 6.47. The definition of differential operators on manifolds with boundary does not require any modification in relation to the case discussed above. The following difference is essential, however: While every differential operator over a *closed* (i.e., compact, without boundary) manifold has a formal adjoint by Exercise 6.39, one always has an extra term for manifolds with boundary

$$\int_X (Pe, f) - \int_X (e, P^* f) = \int_{\partial X} (\cdots)$$

involving a differential operator over the boundary; e.g., see our discussion above on the Sturm-Liouville boundary-value problem (Section 2.5), or more generally [**Pal65**, p.73-75] and [**BoWo**, Proposition 3.4]. We will come back to this problem in Chapter 10. Thus, one of the advantages of the computations on compact manifolds without boundary is the existence of formal adjoints. In passing to the boundary-value problems of interest in applications, Exercise 6.46 comes into play. More generally, in Chapter 10 and Section 14.8, we will assign to each boundary-value problem over X *associated* operators over the two closed manifolds ∂X and $X \cup_{\partial X} X$. Incidentally, these are manifolds which can be rather complicated topologically, even in the classical case when X is a bounded domain in $\mathbb{R}^n$. (See the *Heegaard diagrams*, whereby each three-dimensional oriented manifold can represented as such a doubling with boundary diffeomorphism, not necessarily the identity [**ST**, p.219].)

CHAPTER 7

Sobolev Spaces (Crash Course)

Synopsis. Motivation. Equivalence of Different Local Definitions. Various Isometries. Global, Coordinate-Free Definition. Embedding Theorems: Dense Subspaces; Truncation and Mollification; Differential Embedding; Rellich Compact Embedding. Sobolev Spaces Over Half Spaces. Trace Theorem. Case Studies: Euclidean Space and Torus; Counterexamples

1. Motivation

In this book, the concept of Sobolev spaces will be used to solve two different problems. The first question is: *How can we fit the analytic concept of (elliptic) differential operators into the framework of functional analysis?* That question is dealt with extensively in many textbooks of modern analysis. One of the answers to that question is, that Sobolev spaces permit a transition from Banach spaces (natural domains of differential operators) to easy Hilbert spaces and provide a link to the theory of bounded Fredholm operators in Hilbert space, as developed above in our Part I. We shall explain all the needed definitions and results in the sequel. That is the easy question and we can be short dealing with it. The second question is much deeper: *How can we fit the geometric concept of a connection into the framework of functional analysis?*. Recall from Section 6.6, Definition 6.31 and our Table 6.1, that all geometric standard operators are related to connections, namely as operators of Dirac type, respectively, Dirac Laplacians (= squares). We will show the fundamental role of connections in gauge theoretic physics and low-dimensional topology in our Part IV. Then the key problem is to develop an all embracing view of the space of all suitable connections in a given geometric or physics context. Such a view is provided by the concept of a manifold. To establish the manifold character of spaces of connections, we need linearisation and parametrisation tools. The most important are infinite-dimensional analogues to the Implicit Function Theorem (IFT). There we shall use Sobolev spaces once again. But then, our point will not be the rather trivial aspect of Sobolev spaces, cultivated in this and the following sections and perhaps overemphasised in analysis main stream literature (namely, the simple replacement of Banach space theory by Hilbert space theory), but the replacement of Fréchet spaces (where we have no IFT) by Banach spaces which will do.

Like we have two different motivations for the introduction of Sobolev spaces, we have also two different ways of doing it. As so often in global analysis, we have the heritage from calssical analysis with its proficiency in making calculations in coordinates. In this section, we shall follow that tradition, making easier reading for a student or teacher who feels safer with coordinates, does not care so much about global geometric meaning and accepts the arbitrary and tiresome coordinate shifts. In Section 17.4, p.468 we present an alternative, namely the natural (coordinate-free) Definition 17.17 of much more delicate families of Sobolev Banach spaces. The global approach is mandatory in Part IV but would also be quite appropriate in this section for readers who seek a really simple, coordinate-free presentation and are not afraid of abstract geometric concepts, see Remark 7.8 below.

Let us leave the connections for later and return to our first question regarding differential operators and there relations to Hilbert spaces. There is first of all the L^2 concept of a Lebesgue measurable square integrable function, which can transferred naturally to sections in a Hermitian vector bundle E on a Riemannian manifold X: A section $u : X \to E$ (not necessarily continuous) represents an element of $L^2(E)$ if $\int_X \langle u, u\rangle\, \nu_g < \infty$. Here $\langle\cdot,\cdot\rangle$ is a Hermitian metric for the vector bundle E, whence $\langle u, u\rangle$ is a $\mathbb{R}$-valued function on X which is integrated with respect to the volume element ν_g defined by the metric tensor g of X (see Section 6.4 above). In this way, $L^2(E)$ becomes a Hilbert space with the usual identification of sections that differ on a set of measure 0.

The traditional way of fitting a differential operator $P \in \mathrm{Diff}_k(E, F)$ into the well-understood and powerful Hilbert space theory consists in considering P as a map on $L^2(E)$ to $L^2(F)$, restricted to functions with sufficient differentiability. We proceeded this way above, when introducing the notion of *formally adjoint operators* (Section 6.7). But even restricted to these subspaces, the differential operators are not continuous in the norm topology of L^2. A simple example is the operator d/dt which maps the sequence $\frac{1}{n}\sin nt$ (converging to 0 in $L^2(S^1)$) to the sequence $\cos nt$ which does not converge in $L^2(S^1)$. This circumstance leads to the extensive field of classical mathematical research on *unbounded* linear operators. We gave a taste in Section 2.6.

Thanks to Sergey Lvovich Sobolev, we now have a more potent tool for the definition of Hilbert spaces with various differentiability properties, namely the Sobolev spaces $W^s(X)$. More common *notations* for these spaces are $H^s(X)$ or $L_2^s(X)$. We prefer the W^s of the Russian literature where the "W" reminds at weak solutions, namely in distributional sense, to be explained below in Exercise 7.7a. In a topology book like ours, $H^s(X)$ is reserved for cohomology. The notation $L_2^s(X)$ would be correct in emphasising that we model the Sobolev spaces after the Hilbert space $L_2(X)$. However, it is a bit heavy and, moreover, we have so many spaces of linear mappings, carrying an L. So we had better stick to $W^s(X)$.

These spaces have gained great importance in the theory of partial differential equations, especially for existence questions where precise statements on the *regularity* of solutions are desired, but often cannot be expressed in the language of C^k-Banach spaces. Since introducing Sobolev spaces by coordinates is an established part of now classical analysis, we may keep it short and refer to the abundant textbook *literature*: the classic [**Ad**], [**BJS**, Chs. III/IV], [**Hö63**, p.33-63], [**LiMa**, p.1-118], [**Na**, p.184-200], [**Pal65**, p.125-174], and [**Yo**, p.55 and 173f], and the more recent [**GS**, p.44ff] and [**Gru**, Sections 4.2, 6.2, 6.3 and 8.2].

2. Definition

Customary Definitions in Coordinates. In the following, we put together some of the various customary equivalent definitions of Sobolev spaces. Here, we restrict ourselves to the *case of functions* and where $s \geq 0$. In the framework of distribution theory, the spaces W^s can be treated clearly and uniformly also for $s < 0$; e.g., see the carefully written [**GS**] and [**Gru**].

The basic idea of Sobolev spaces is very simple: As explained above, we wish a common functional analytic frame for linear differential operators. Let us try:

DEFINITION 7.1. Let $m \geq 0$ be an integer. We define the **Sobolev space** $W^m(\mathbb{R}^n)$ as the intersection of the maximal domains of all elementary formally self-adjoint differential operators

$$D^\alpha := (-i)^{|\alpha|} \frac{\partial^{\alpha_1}}{\partial x_{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_{\alpha_n}}, \ \ |\alpha| := \alpha_1 + \cdots + \alpha_n \leq m,$$

of order $\leq m$, i.e., to consist of all $u \in L^2(\mathbb{R}^n)$ such that $D^\alpha u \in L^2(\mathbb{R}^n)$ for all multiindices α with $|\alpha| \leq m$.

Recall that for $u \in L^2(\mathbb{R}^n)$, we mean by $D^\alpha u \in L^2(\mathbb{R}^n)$ that there exists a $v \in L^2$ such that the *distribution* $D^\alpha u$ acts like v on all *test functions* w, i.e.,

$$\begin{aligned}\big(D^\alpha u\big)(w) \ := \ \langle u, (D^\alpha)^* w\rangle_0 \ = \ \langle u, D^\alpha w\rangle_0 \ := \int_{\mathbb{R}^n} u(x)\, \overline{D^\alpha w(x)}\, dx \\ \overset{!}{=} \int_{\mathbb{R}^n} v(x)\, \overline{w(x)}\, dx \ = \ \langle v, w\rangle_0 \qquad \text{for all } w \in C_0^\infty(\mathbb{R}^n).\end{aligned}$$

By Fourier analysis (Differentiation-Multiplication Conversion and Plancherel Formula, Exercises A.5b,d of Appendix A), this is equivalent to requiring $\xi^\alpha\, \widehat{u}(\xi) \in L^2$ for $|\alpha| \leq m$, or, what is the same, $(1+|\xi|)^m\, \widehat{u} \in L^2$. The following exercise makes you familiar with the arguments.

EXERCISE 7.2. Show that, for $u \in C_0^\infty(\mathbb{R}^n)$, the two following norms $|\cdot|_m$ and $\|\cdot\|_m$ are equivalent ($m \in \mathbb{N}$):

a)

$$|u|_m := \left(\sum\nolimits_{|\alpha|\leq m} |D^\alpha u|_0^2\right)^{1/2}, \qquad \text{where } |u|_0^2 := \langle u, u\rangle_0 = \int_{\mathbb{R}^n} u\,(x)\, \overline{u\,(x)}\, dx.$$

b) For $\Delta = -\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right)$,

$$\|u\|_m := \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^m\ |\widehat{u}\,(\xi)|^2\ d\xi\right)^{1/2} = \left(\int_{\mathbb{R}^n} \langle (1+\Delta)^m u, u\rangle\ dx\right)^{1/2},$$

where the last equality is due to Exercise A.5(b,d), p. 675 in Appendix A.
[Hint: Start with the Fourier differentiation formula (see Appendix A), giving

$$|u|_m^2 := \int_{\mathbb{R}^n} \left(\sum\nolimits_{|\alpha|\leq m} (\xi^\alpha)^2\right) |\widehat{u}\,(\xi)|^2\ d\xi.$$

Then prove that for some constant c,

$$\left(1+|\xi|^2\right)^m \leq \sum\nolimits_{|\alpha|\leq m} (\xi^\alpha)^2 \leq c\left(1+|\xi|^2\right)^m$$

and deduce $\|u\|_m \leq |u|_m \leq \sqrt{c}\, \|u\|_m$.]

This leads to the following more general definition.

EXERCISE 7.3. For real non-negative s, define the **Sobolev space**

$$W^s(\mathbb{R}^n) := \left\{u \in L^2(\mathbb{R}^n) : \xi \mapsto (1+|\xi|^2)^{s/2}\, \widehat{u}(\xi) \in L^2(\mathbb{R}^n)\right\}$$

and show:

a) For all $s \in \mathbb{N}$, the space $W^s(\mathbb{R})$ is the completion of $C_0^\infty(\mathbb{R}^n)$ relative to the

(equivalent, by Exercise 7.2) s-norms $|\cdot|_s$ and $\|\cdot\|_s$, and whence a Banach space.
b) By considering scalar products which induce the respective norms, $W^s(\mathbb{R}^n)$ becomes a Hilbert space.
c) The following inclusions are defined in a natural way, and are continuous and dense (where $W^s = W^s(\mathbb{R}^n)$ for short)

$$C_0^\infty(\mathbb{R}^n) \subset W^\infty := \bigcap_{s=0}^{\infty} W^s \subset \ldots \subset W^{s+t} \subset \ldots \subset W^s \subset \ldots \subset W^0 := L^2(\mathbb{R}^n).$$

See also Theorem 7.13, p. 197.
[Hint: For a: Investigate Cauchy sequences in $C_0^\infty(\mathbb{R}^n)$ relative to $|\cdot|_s$.
For b: For natural s, this is clear by (a). For arbitrary real s, see the classical [**Hö63**, p.37 and 45f] or [**LiMa**, p.35-37] or the more recent [**GS**] and [**Gru**].
For c: For the inclusions, note the monotonicity of $\left(1+|\xi|^2\right)^s$ in s. For the proof that $C_0^\infty(\mathbb{R}^n)$ is dense in $W^s(\mathbb{R}^n)$, note that

- $C_0^\infty(\mathbb{R}^n)$ is dense in the Schwartz space $C_\downarrow^\infty(\mathbb{R}^n)$ of rapidly decreasing functions;
- the Fourier transform is an isometric isomorphism
$$\mathcal{F} : W^s(\mathbb{R}^n) \longrightarrow L^2\big(\mathbb{R}^n, \left(1+|\xi|^2\right)^s d\xi\big); \text{ and}$$
- $\mathcal{F}^{-1}\big(C_\downarrow^\infty(\mathbb{R}^n)\big) = C_\downarrow^\infty(\mathbb{R}^n)$.

Since $C_\downarrow^\infty(\mathbb{R}^n)$ is clearly dense in $L^2\big(\mathbb{R}^n, \left(1+|\xi|^2\right)^s d\xi\big)$, it follows that $C_\downarrow^\infty(\mathbb{R}^n)$ is dense in $W^s(\mathbb{R}^n)$, and you are done.
You may try another more direct proof of the density: Note that the space $W_K^s(\mathbb{R}^n)$ of functions with support in compact $K \subset \mathbb{R}^n$ is dense in $W^s(\mathbb{R}^n)$. So, given $u \in W_K^s(\mathbb{R}^n)$, how can you construct a sequence $u_\nu \in C_0^\infty(\mathbb{R}^n)$ with $\operatorname{supp} u_\nu \subset \widetilde{K}$ (another compact subset with $K \subset \operatorname{Int} \widetilde{K}$), such that $\|u_\nu - u\|_s \to 0$ for $\nu \to \infty$? Exploit the convolution (see Definition A.4c, p.674): You can choose compactly supported standard test function $\varphi \in C_0^\infty(\mathbb{R}^n)$, real valued with $\varphi \geq 0$, $\operatorname{supp} \varphi \subset \{|x| < 1\}$ and $\int_{\mathbb{R}^n} \varphi(x)dx = 1$. Then put

$$\varphi_\varepsilon(x) := \varepsilon^{-n}\varphi(x/\varepsilon) \text{ and } u_\varepsilon(x) := (\varphi_\varepsilon * u)(x) = \langle u(y), \varphi_\varepsilon(x-y)\rangle_0, \qquad \varepsilon > 0.$$

Note that $\operatorname{supp} \varphi_\varepsilon \subset \{|x| < \varepsilon\}$ and $\int \varphi_\varepsilon(x)dx = 1$. To show that $u_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ is an easy exercise in the differentiability of integrals with parameters. The interesting part is to prove that $\lim_{\varepsilon\to 0+} \|u_\varepsilon - u\|_s = 0$. For that, work with the Fourier transform: Since $\widehat{\varphi}_\varepsilon(\xi) = \widehat{\varphi}(\varepsilon\xi)$ and $\widehat{\varphi}(0) = 1$, the question reduces to establishing the relation

$$\lim_{\varepsilon\to 0+} \int |\widehat{\varphi}(\varepsilon\xi) - 1|^2 \, |\widehat{u}(\xi)|^2 \left(\xi_1^{2s} + \cdots + \xi_n^{2s}\right) d\xi \;=\; 0,$$

which is evident from the dominated convergence theorem.]

Remark 7.4. a) The family $\left(\varphi_\varepsilon\right)_{0<\varepsilon<1}$ (more precisely the family of linear scalar operators $\left(J_\varepsilon\right)_{0<\varepsilon<1}$ with $J_\varepsilon(\cdot) := \varphi_\varepsilon * \cdot$, mapping integrable functions to smooth functions) is called a **mollifying family** because J_ε smoothes out the asperities and we have $J_\varepsilon \to \mathrm{Id}$ in the appropriate norm.
b) There is a second basic technique for dealing with Sobolev spaces, namely **truncation**. The essentials are contained in the following result: Let $u \in W^k(\mathbb{R}^n)$. Consider for each $R > 0$ a smooth bump (cut-off) function $\chi_R \in C_0^\infty(\mathbb{R}^n)$ with $\chi_R(x) \equiv 1$ for $|x| \leq R$, $\chi_R(x) \equiv 0$ for $|x| \geq R+1$, but sufficiently moderate, say

with differential $|d\chi_R(x)| \le 2$ for all $x \in \mathbb{R}^n$. Then $\chi_r \cdot u \in W^k(\mathbb{R}^n)$ for all $R > 0$ and, moreover,

$$\chi_R \cdot u \overset{W^k}{\longrightarrow} f, \qquad \text{as } R \longrightarrow \infty.$$

We leave the proof to the reader (else see [**Nic**, Lemma 9.2.9]).

EXERCISE 7.5. Show that for $s \in \mathbb{R}$, the formally self-adjoint operator (in fact, a pseudo-differential operator, see the following chapter)

$$(\Lambda_s u)(x) := \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \left(1 + |\xi|^2\right)^{s/2} \widehat{u}(\xi)\, đ\xi,$$

with $đx = (2\pi)^{-n/2} dx_1 \cdots dx_n$, defines an isomorphism (in particular, isometry) $\Lambda_s : W^{t+s}(\mathbb{R}^n) \to W^t(\mathbb{R}^n)$ of Hilbert spaces for $t \ge 0$ and $t + s \ge 0$.
[Hint: Note that Parseval's Formula (Appendix A, Exercise A.5d, p. 675) implies the equality $\|u\|_s = \|\Lambda_s u\|_0$ for $u \in W^s(\mathbb{R}^n)$, and that the family $\{\Lambda_s : s \in \mathbb{R}\}$ forms a group since $\Lambda_s \circ \Lambda_r = \Lambda_{s+r}$.]

REMARK 7.6. A common abbreviation of the expression $\left(1 + |\xi|^2\right)^{\frac{1}{2}}$ is the symbol $\langle\xi\rangle$. Whence, in the notation of pseudo-differential operators of Chapter 8, we can write $\Lambda_s = \mathrm{Op}(\langle\xi\rangle^s)$.

Global and Coordinate-Free Definitions. From the preceding presentation of Sobolev spaces in local *Euclidean coordinates*, the reader can catch the basic idea, namely that Sobolev spaces are closures of spaces of differentiable functions with regard to the L^2-norms of the highest derivative. However, it seems to us that the power of the concept of Sobolev spaces becomes clearer in global and coordinate-free presentation. We shall give several choices.

EXERCISE 7.7. Let X be a compact, oriented, C^∞ Riemannian n-manifold (without boundary), and let E be a C^∞ Hermitian vector bundle over X of fiber dimension N.
a) For a positive integer s, define the **Sobolev space** (to begin with, only the underlying vector space)

$$W^s(E) := \{u \in L^2(E) : \text{for each } P \in \mathrm{Diff}_s(E,E), \text{ there is } v \in L^2(E) \\ \text{such that } \langle u, Pw\rangle_0 = \langle v, w\rangle_0 \text{ for all } w \in C^\infty(E)\}. \tag{7.1}$$

Show that a section $u \in L^2(E)$ lies in $W^s(E)$, exactly when, for each local representation of u in the form $u(x) = \sum u_i(x) e_i(x)$ relative to a local chart and local basis $e_1, ..., e_N$ of E, we have $\varphi u_i \in W^s(\mathbb{R}^n)$, for all C^∞ functions φ with support in the domain of the chart.
b) Define $W^s(E)$ for $s \in \mathbb{R}^+$, using this local recipe.
[Hint: For a: Definition 7.1 and Exercise 7.2a. Note that v is uniquely determined by P. One says that v arises by *weak application* of the formal adjoint operator P^* on u (*differentiation in the distributional sense*, i.e., "$P^*u = v$" $\iff$ P^*u and v act identically on all *test vector valued functions* $w \in C^\infty(E)$ with $\langle P^*u, w\rangle_0 = \langle u, Pw\rangle_0$).
For b: The crucial point is the independence of the *set* $W^s(E)$ of the choice of charts, local trivializations, the smoothing functions. Take care with the coordinate changes: It is not entirely trivial that each diffeomorphism $\kappa : U \to V$ between open subsets of $\mathbb{R}^n$ induces (via $v \mapsto v \circ \kappa$) an isomorphism $W^s_K(\mathbb{R}^n) \to W^s_{\kappa^{-1}(K)}(\mathbb{R}^n)$,

where $K \subset V$ is compact and $W^s_K(\mathbb{R}^n) := \{v \in W^s(\mathbb{R}^n) : \operatorname{supp} v \subseteq K\}$. An elementary proof for this is found in [**Hö63**, p.57-59]. In our context it is simpler to jump forward to Theorem 8.18, p. 220, where we shall show the invariance of the space $\mathrm{L}^s_{\mathrm{pc}}$ of (principally classical) pseudo-differential operators under a coordinate change (but only for $s \in \mathbb{Z}^+$). However, the proof goes through smoothly for $s \in \mathbb{R}^+$. Then define $W^s(E)$ as in (a), where $P \in \mathrm{L}^s_{\mathrm{pc}}(E,E)$. Instead of coordinate invariance, which is self-evident, one must show, as in (a), that one obtains elements of $W^s(\mathbb{R}^n)$ locally. Details are found in [**Hö66b**, p.169f] or [**Ni**, p.151f].]

Remark 7.8. (a) For a fixed choice of atlas, local trivializations of the bundle E, and an appropriate C^∞ partition of unity, one obtains a norm and scalar product which makes $W^s(E)$ a Hilbert space. Without such choices, we must do with a **Hilbertable** space ([**Pal65**]'s notation).
(b) Instead of arguing with *all* elements in $\mathrm{Diff}_s(E,E)$ we can define $W^s(E)$ both as set and as Hilbert space by specifying a single **generating operator** $\Lambda_{E,s}$, actually a (principally classical) pseudo-differential operator (belonging to $\mathrm{L}^s_{\mathrm{pc}}(E,E)$, a space to be defined below in Section 8.3): Let $\{U_j\}_{j\in J}$ be a locally finite covering of the underlying n-manifold X by domains of coordinate charts $\kappa_j : U_j \to \mathbb{R}^n$ and local trivializations $\tau_j : E|_{U_j} \to U_j \times \mathbb{C}^N$, and let $\{\varphi_j \in C^\infty(X)\}_{j\in J}$ be a partition of unity subordinate to $\{U_j\}_{j\in J}$. Let $u \in L^2(E)$, i.e., $\int_X \langle u,u\rangle_h \nu_g < \infty$ where ν_g denotes the volume element for the Riemannian metric g on X and $\langle\cdot,\cdot\rangle_h$ the Hermitian product on the vector bundle E, see Definition 6.14b, p.167. We set

$$\Lambda_{E,s} u := \sum_{j\in J} \varphi_j \cdot \left(\tau_j^{-1} \circ \left(\Lambda_s^N \overline{u_j}\right) \circ \kappa_j\right), \tag{7.2}$$

where $\overline{u_j} := \begin{cases} \tau_j \circ u \circ \kappa_j^{-1}, & \text{on } \kappa_j(U_j), \\ 0, & \text{on } \mathbb{R}^n \setminus \kappa_j(U_j), \end{cases}$ and $\Lambda_s^N := \Lambda_s \oplus \cdots \oplus \Lambda_s$ with $\Lambda_s : W^s(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ as defined in Exercise 7.5. The process used in Equation (7.2) to construct the operator $\Lambda_{E,s}$ from the operators Λ_s^N given in local coordinates is called *gluing together*. Similarly, we defined differential operators globally via local coordinates in Definition 6.32, p.178 and shall define pseudo-differential operators globally via local coordinates in Definition 8.12, p.212. We set

$$W^s(E) := \Lambda_{E,s}^{-1}\left(L^2(E)\right) \quad \text{and} \quad \langle u,v\rangle_s := \langle \Lambda_{E,s} u, \Lambda_{E,s} v\rangle_0 \tag{7.3}$$

for $u, v \in W^s(E)$. One checks that the Equations (7.1) and (7.3) yield the same vector space $W^s(E)$. Contrary to the definition of the Sobolev space by (7.1), the preceding definition based on the generating $\Lambda_{E,s}$ yields a scalar product at once and makes $W^s(E)$ a Hilbert space. As before, however, the inner product is not canonical but depends also here on the choice of coordinates etc, entering into the definition of $\Lambda_{E,s}$.
Warning: While we had $\Lambda_{-s} \circ \Lambda_s = \mathrm{Id}$ on $W^s(\mathbb{R}^n)$ we get now an error term

$$R_s := \Lambda_{E,-s} \circ \Lambda_{E,s} - \mathrm{Id}_{W^s(E)}, \tag{7.4}$$

which is by definition an integral operator $(R_s u)(x) = \int_X K_s(x, x-y) u(y) dy$ with smooth kernel K_s and therefore compact in $\mathcal{B}(W^s(E), W^s(E))$ (to be proved like in Exercise 2.29, p.25) and extendable to the whole $L^2(E)$ and transforming it into $C^\infty(E)$.
(c) However, there are other choices to fix the scalar product: As with Exercise

7.2b, on manifolds one may define a Laplace operator Δ which is an elliptic, self-adjoint, positive-definite second-order differential operator. For a natural s (and also for real s, via the Spectral Theorem 2.61, p.51), we then may explicitly set

$$\|u\|_s := \langle (\mathrm{Id} + \Delta)^s u, u\rangle_0 , \qquad u \in C^\infty(E).$$

(d) By [**AS68a**, p.511] (the idea goes back to [**Mag**, p.134-197], see also [**LiMa**, p.42]), one can proceed in this way even further, if the vector bundle E is furnished with a C^∞ *connection* ∇^E (defined and discussed in Section 6.6, 172ff). Consider $\nabla^E : C^\infty(E) \to C^\infty(E \otimes T^*X)$ as differential operator. By composition with its formally adjoint $\nabla^* : C^\infty(E \otimes T^*X) \to C^\infty(E)$, we obtain a positive, semi-definite, formally self-adjoint, *Laplacian*, namely $\Delta := \nabla^* \circ \nabla$.

Following up on this remark, we may replace our *conventional* introduction of the Sobolev spaces via lengthy and, in principle, artificial coordinate transformations, by an alternative *geometric* (in particular, coordinate-free) definition, namely by specifying one single *generating* differential operator ∇^j.

Definition 7.9. We equip the bundle $E \to X$ with a Hermitian structure h and a metric covariant differentiation operator $\nabla^E : C^\infty(E) \to C^\infty(T^*X \otimes E)$. By also employing a Riemannian metric g and Levi-Civita connection θ^g on X, we obtain for any $k = 0, 1, 2, \ldots$ a connection in the tensor products $(\otimes^k T^*X) \otimes E$

$$\nabla^{k,E} : C^\infty\left((\otimes^k T^*X) \otimes E\right) \to C^\infty\left((\otimes^{k+1} T^*X) \otimes E\right).$$

For each $j = 1, 2, \ldots$ we write shortly $\nabla^j : C^\infty(E) \to C^\infty\left((\otimes^j T^*X) \otimes E\right)$ for the composition

$$\begin{aligned}\nabla^j : C^\infty(E) \xrightarrow{\nabla^E} C^\infty(T^*X \otimes E) &\xrightarrow{\nabla^{1,E}} C^\infty(T^*X \otimes T^*X \otimes E)\\ &\xrightarrow{\nabla^{2,E}} \quad \cdots \quad \xrightarrow{\nabla^{j-1,E}} C^\infty\left((\otimes^j T^*X) \otimes E\right).\end{aligned}$$

a) For $u, v \in C^\infty(E)$ and $m > 0$, we then set

$$(u, v)_m := \sum_{j=0}^{m} \int_X \langle \nabla^j u, \nabla^j v\rangle \nu_g , \quad \text{and} \quad \|u\|_m := \sqrt{(u,u)_m}\,,$$

where ν_g denotes the volume element for g, and the inner product $\langle \nabla^j u, \nabla^j v\rangle$ is the natural one constructed from the one induced by g on $\otimes^j T^*X$ and the Hermitian structure on E.

b) The Sobolev space $W^m(E)$ is the completion of $C^\infty(E)$ with the norm $\|\cdot\|_m$.

Differently put, $W^m(E)$ is the space of sections $u \in L^2(E)$ such that for all $j = 1, \ldots, m$ there exists $v_j \in L^2\left((\otimes^j T^*X) \otimes E\right)$ with $\nabla^j u = v_j$ *weakly*, i.e.,

$$\int_X \langle v, w\rangle\, \nu_g \;=\; \int_X \langle u, \left(\nabla^j\right)^* w\rangle\, \nu_g, \qquad \text{for all } w \in C^\infty\left((\otimes^j T^*X) \otimes E\right).$$

We shall come back to this global type of definition in Section 17.4, Definition 17.17, p.468. For now, we leave it to the reader to check the accordance between our two Definitions, i.e., the conventional definition by combining Exercises **??** and 7.7, and the preceding global Definition 7.9.

Remark 7.10. Not the underlying set of the Sobolev space $W^m(E)$ introduced here, but its scalar product depends (as before, but differently) on several choices: the metrics on X and E and the connection on E. It was pointed out in [**Nic**, Theorem 9.2.24] that the identity map between two such versions of $W^m(E)$ is a Banach space isomorphism, if we restrict ourselves to *compatible* connections (i.e., *metric* in the sense of Definition 6.26, p.176). When we expand our definition to non-compact X this dependence is very dramatic and has to be seriously taken into consideration.

REMARK 7.11. In the literature and in our Part IV, much more general Sobolev spaces (*Bessel potentials*, etc.) are treated. For these, one starts with L^p theory instead of the Hilbert spaces L^2, and works with *weights* other than our $(1+|\xi|^2)^{s/2}$. It is interesting that in "the study of classes of differential equations with variable coefficients which are defined by conditions on their highest-order part" (Hörmander) only those W-spaces play a role which are distinguished in a way by their invariance on manifolds (*translation invariance* of L^2 and *diagonalizability* of the derivative by means of Fourier transformation). Indeed, in the present part of our book, the L^2-modelled Sobolev spaces suffice. To describe the manifold structure of moduli of self-dual connections, however, we shall, as mentioned before, define wider Sobolev spaces in Section 17.4, in immediate generalization of the preceding Definition 7.9.

Sobolev Spaces Over Half-Spaces. In many applications it is natural to consider manifolds with boundary, modelled on half-spaces. For a comprehensive treatment we refer to [**BoWo**]. For now, the following Exercise may suffice.

EXERCISE 7.12. Let m be a positive integer.
a) For $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_n \geq 0\}$, define (as in Exercise 7.7a) the space

$$W^m(\mathbb{R}^n_+) := \{u \in L^2(\mathbb{R}^n_+) : \text{for each } P \in \mathrm{Diff}_m(\mathbb{R}^n_+), \text{ there is } v \in L^2(\mathbb{R}^n_+) \text{ such that } \langle u, Pw\rangle_0 = \langle v, w\rangle_0 \text{ for all } w \in C^\infty\left(\mathbb{R}^n_+\right)\}.$$

Prove that

$$W^m(\mathbb{R}^n_+) = \{u \in L^2(\mathbb{R}^n_+) : \text{there is } v \in W^m(\mathbb{R}^n) \text{ with } v|_{\mathbb{R}^n_+} = u\}.$$

Show that $W^m(\mathbb{R}^n_+)$ is a Hilbert space.
b) Define the space $W^m(X)$ for a compact, orientable manifold X with boundary via *localization*, and carry over Exercise 7.3c.
[Hint: For a: Set $\|u\|_m := \inf\{\|v\|_m : v \in W^m(\mathbb{R}^n) \text{ and } v|_{\mathbb{R}^n_+} = u\}$. Be careful with restricting to the half-space: For $m \neq 0$, one must distinguish between $W^m(\mathbb{R}^n_+)$ and $W^m_{\mathbb{R}^n_+}(\mathbb{R}^n)$, the space of W^m functions with support in $\mathbb{R}^n_+$; see [**Hö63**, p.51-54]. For b: The invariance under diffeomorphism is trivial here, since (without any loss in the applications; see Chapter 10) we only consider whole numbers m. See also [**Hö63**, p.60f].]

3. The Main Theorems on Sobolev Spaces

Here, we discuss briefly (partly without full proofs, for which we refer to the literature) the three main theorems. The content of these results is illustrated in Section 7.4, *Case Studies*.

We begin with a regularity theorem showing how one can pass from Hilbert space results given in the language of Sobolev spaces to results in classical form.

THEOREM 7.13 (S. L. Sobolev, 1938). *Let X be a compact C^∞ manifold (with or without boundary) and $s > 0$. Define the* **strength** *of the Sobolev space $W^s(X)$*

by

$$\operatorname{str}(W^s(X)) := s - \frac{\dim X}{2}\,.$$

Then $W^s(X) \subset C^k(X)$ *for all* $k < \operatorname{str}(W^s(X))$*, and the embedding is continuous with the* ***Sobolev inequality***

$$\|u\|_{C^k} \le \varepsilon\,\|u\|_{W^s} + C_\varepsilon\,\|u\|_{L^2} \qquad \text{for } u \in W^s(X) \text{ and } k < \operatorname{str} W^s(X),$$

where $\varepsilon > 0$ *can be made arbitrarily small, if* C_ε *is sufficiently large.*

Note. Whence, the strength of $W^s(X)$ is a measure for the regularity of its elements: the bigger the strength the more regular are the functions in that space and thus it consists of *fewer* functions or, rather, *elements* or *classes*. More precisely, an element u of $W^s(X) \subset L^2(X)$ is a class of functions which agree almost everywhere. The theorem means this: In each class $u \in W^s(X)$, there is a representative in $C^k(X)$, and each sequence of elements in $W^s(X)$ which converges in the norm of $W^s(X)$ yields a sequence of representatives in $C^k(X)$ which converges in the norm of the Banach space $C^k(X)$.

Proof. For the euclidean case, we give a taste of the proof below in Theorem 7.16. Else see [**BJS**, p.167] and [**Pal65**, p.159f] for $X = T^n := S^1 \times \cdots \times S^1$, and [**Pal65**, p.169] for the transition to arbitrary X and to sections of vector bundles. For the Sobolev inequality, see [**Ad**, 75-76/97f], where X is a codimension 0 submanifold (with boundary) of $\mathbb{R}^n$. □

The treatment of boundary value problems with Hilbert space methods is problematic, since in a *fixed* L^2 space a function is only uniquely defined modulo its values on sets of measure zero such as the boundary. Thus, the restriction of such a function to the boundary is completely arbitrary. However, the following restriction theorem, which also goes back to S.L. Sobolev, is helpful (e.g., for $Y = \partial X$ and $m = 1$).

Theorem 7.14. *Let* X *be a compact,* C^∞ *manifold* (*possibly with boundary*) *with a compact submanifold* Y *of codimension* m*, and let* E *be a vector bundle over* X*. Then, for each integer* $s > m/2$*, the canonical restriction map* $C^\infty(E) \to C^\infty(E|_Y)$ *extends to a continuous, linear, surjective map* $W^s(E) \to W^{s-\frac{m}{2}}(E|_Y)$*.*

Proof. For the periodic case $X = T^n$, $Y = T^{n-1}$ and E the trivial line bundle, we give a full proof below in Theorem 7.17. For the hyperplane problem $W(\mathbb{R}^n_+) \to W^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$, see [**Hö63**, p.54f] or [**LiMa**, p.38]; for $Y = \partial X$, see [**LiMa**, p.44-48]; for $X = T^n$ and $Y = T^{n-m}$ and the general case, see [**Pal65**, p.161 f]. □

Finally, the following lemma, named after Franz Rellich and proved by him in a different formulation, brings in compact operators (see our Chapters 2 ff.), and will furnish a further connection with the Fredholm theory of elliptic operators.

Theorem 7.15 (F. Rellich, 1930). *If* X *is a compact,* C^∞ *manifold (possibly with boundary) and* E *is a complex vector bundle over* X*, then the inclusion* $W^m(E) \hookrightarrow W^s(E)$ *is compact for* $m > s \ge 0$*.*

There are many different proofs in the literature: See [**Ad**, p.144], when X is a codimension 0 submanifold of $\mathbb{R}^n$. For $X = T^n := S^1 \times \cdots \times S^1$, see [**BJS**, p.169f] or [**Pal65**, p.158f], similarly the very clear presentation [**Gru**, Theorem

8.2], and [**Pal65**, p.168] for the general case. We shall give an explicit proof in the special scalar and compact-supported euclidean case of $W^1_K(\mathbb{R}^n) \hookrightarrow L^2(\mathbb{R}^n)$ (closely following [**Nic**, Theorem 9.2.14]). One may wonder about a shorter, structural and more general proof. Inspiration may be found in [**Shu**, Theorem 7.4]). See also our Remark 7.8b with the compact error term $R_s := \Lambda_{E,-s}\Lambda_{E,s} - \mathrm{Id}_{W^s(E)}$ of (7.4) on p.195.

Explicit proof in scalar, Euclidean, compactly supported case. Let $R > 0$ and let us show that the inclusion $W^1_{B_R}(\mathbb{R}^n) \hookrightarrow L^2(\mathbb{R}^n)$ is compact. Let (u_μ) be a bounded sequence in $W^1(\mathbb{R}^n)$ supported in the ball $B_R = B_R(0) := \{|x| \le R\}$. We have to show that the sequence contains a subsequence convergent in L^2. The proof will be carried out in two steps.

Step 1. We will prove that for every $0 < \delta < 1$ the mollified sequence (see Remark 7.4) $(u_{\mu,\delta} := \varphi_\delta * u_\mu)$ admits a subsequence *uniformly convergent* on $B_{R+1} := B_{R+1}(0)$.

To prove this we will apply the Arzela-Ascoli Theorem (Theorem 2.17 and Corollary 2.18, p.19f.). Whence, we shall show that

(i) $\exists_{C=C(\delta)}\, \forall_\mu\, \forall_{|x|\le R+1}\, |u_{\mu,\delta}(x)| < C$,
(ii) $\forall_\mu\, \forall_{|x|,|x'|\le R+1}\, |u_{\mu,\delta}(x) - u_{\mu,\delta}(x')| < C|x - x'|$.

Indeed,

$$
\begin{aligned}
|\left(\varphi_\delta * u_\mu\right)(x)| &\le \delta^{-n} \int_{|y-x|\le\delta} \varphi\left(\frac{x-y}{\delta}\right) |u_\mu(y)|\, dy \le \delta^{-n} \int_{B_\delta(x)} |u_\mu(y)|\, dy \\
&\overset{\text{Schwarz}}{\le} \; C(n)\delta^{-N}\, \|u_\mu\|_{L^2(\mathbb{R}^n)} \cdot \mathrm{vol}(B_\delta)^{1/2} \;\le\; C(\delta)\, \|u_\mu\|_{W^1(\mathbb{R}^n)},
\end{aligned}
$$

by the continuous embedding of $W^1 \hookrightarrow L^2$, Exercise 7.3c.
Similarly,

$$
\begin{aligned}
|u_{\mu,\delta}(x) - u_{\mu,\delta}(x')| &\le \int_{B_{R+1}} |\varphi_\delta(x-y) - \varphi_\delta(x'-y)| \cdot |u_\mu(y)|\, dy \\
&\le C(\delta)\cdot|x-x'| \int_{B_{R+1}} |u_\mu(y)|\, dy \le C(\delta)\cdot|x-x'|\, \|u_\mu\|_{W^1(\mathbb{R}^n)}.
\end{aligned}
$$

Step 1 is completed.
Step 2. So, for each $\delta \in (0,1)$ we have a uniformly convergent subsequence $(v_{\nu,\delta})$ of the mollifier sequence $(u_{\mu,\delta})$. Using the diagonal procedure for $\delta = 1/\nu \to 0$ and $\nu \to \infty$, we pick for each $\nu \in \mathbb{N}$ the function $v_{\nu,1/\nu}$ and denote the corresponding element of the original sequence (u_μ) by $u'_\nu = u_{\mu'}$, i.e., the element $u_{\mu'}$ that yields exactly $v_{\nu,1/\nu} = \varphi_{1/\nu} * u_{\mu'}$ by convolution with $\varphi_{1/\nu}$. Note that the sequence $\left(v_{\nu,1/\nu}\right)$ is uniformly convergent on B_R by Step 1 and $\lim_{\nu\to\infty}\|v_{\nu,1/\nu} - u'_\nu\|_{W^1(\mathbb{R}^n)} = 0$ under mollification.

We claim the subsequence $(u'_\nu) \subset (u_\mu)$ is convergent in $L^2(B_R)$. Indeed, for all natural ν and ρ

$$
\|u'_\nu - u'_\rho\|_{L^2(B_R)} \;\le\; \|u'_\nu - v_{\nu,1/\nu}\|_{L^2(B_R)} + \|v_{\nu,1/\nu} - v_{\rho,1/\rho}\|_{L^2(B_R)} + \|v_{\rho,1/\rho} - u'_\rho\|_{L^2(B_R)}.
$$

Each of the three terms on the right tends to 0 for $\nu \to \infty$ since the embedding $W^1 \hookrightarrow L^2$ is continuous (once again, Exercise 7.3c). Hence the subsequence (u'_ν) of our original bounded sequence (u_μ) is a Cauchy sequence in $L^2(B_R)$ and thus it converges. The compactness theorem is proved. □

4. Case Studies

To illustrate the preceding theorems, we consider some simple special cases. To begin with we prove a simple euclidean version of Theorem 7.13.

THEOREM 7.16. *If $s > n/2$, then each $u \in W^s(\mathbb{R}^n)$ is bounded and continuous, and the inclusion $W^s(\mathbb{R}^n) \hookrightarrow C^0(\mathbb{R}^n)$ is continuous.*

PROOF. By the Fourier Inversion Formula and the Integrable-Continuous Conversion (Exercise A.5, p.675 in Appendix A), it suffices to prove $\widehat{u}$ is in $L^1(\mathbb{R}^n)$. Indeed, we get

$$\begin{aligned}(7.5)\qquad \int_{\mathbb{R}^n} |\widehat{u}(\xi)|\, d\xi &\le \int_{\mathbb{R}^n} |\widehat{u}(\xi)| \left(1+|\xi|^2\right)^{s/2} \left(1+|\xi|^2\right)^{-s/2} d\xi \\ &\le \left(\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(1+|\xi|^2\right)^{s} d\xi\right)^{1/2} \left(\int_{\mathbb{R}^n} \left(1+|\xi|^2\right)^{-s} d\xi\right)^{1/2}.\end{aligned}$$

Here we have used the Schwarz Inequality $\langle a,b\rangle \le \langle a,a\rangle^{1/2} \langle b,b\rangle^{1/2}$; note that the first factor is finite by assumption and the latter factor is finite precisely for $s > n/2$. By the Riemann-Lebesgue Lemma, we can even conclude that $u(x)$ vanishes at infinity. To prove the continuity of the inclusion, let $u \in C_0^\infty(\mathbb{R}^n)$. We have for $x \in \mathbb{R}^n$, with our convention $đ\xi := (2\pi)^{-n/2} d\xi$,

$$|u(x)| = \left|\int_{\mathbb{R}^n} e^{ix\xi}\widehat{u}(\xi)\, đ\xi\right| \le \int_{\mathbb{R}^n} |\widehat{u}(\xi)|\, đ\xi.$$

Estimate (7.5) and the definition in Exercise 7.2b then yield

$$\sup_{x\in\mathbb{R}^n} |u(x)| \le K \|u\|_{W^s(\mathbb{R}^n)},$$

where the constant K (e.g., $(2\pi)^{-n/2} \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^{-s} d\xi\right)^{1/2}$) does not depend on u. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $W^s(\mathbb{R}^n)$ (see Exercise 7.3c, p. 192), we are done. □

THEOREM 7.17. *We write the n-dimensional torus T^n in the form $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$. Then the restriction map $C^\infty(T^n) \to C^\infty(T^{n-1})$ (induced by the projection $(y,\theta) \mapsto y$ of T^n onto T^{n-1}) extends to a continuous linear map $W^s(T^n) \to W^{s-\frac{1}{2}}(T^{n-1})$, for $s \ge 1/2$.*

PROOF. We follow [**Pal65**, p.143-162], while previously Peter Lax[1] had remarked that, in the periodic case, certain technical difficulties vanish:

Step 1: $C^\infty(T^n)$ consists of functions on $\mathbb{R}^n$ which are periodic of period 2π in each variable. The functions

$$e_\nu(x) := (2\pi)^{-n/2}\, e^{2\pi i\langle \nu,x\rangle},\ \nu \in \mathbb{Z}^n$$

form a complete orthonormal system for $L(T^n)$; see the theory of Fourier series (Appendix A). By definition (see Exercise 7.2b), we have $u \in W^s(T^n)$ exactly when

$$\|u\|_s^2 := \sum\nolimits_{\nu\in\mathbb{Z}^n} |\widehat{u}(\nu)|^2 (1+|\nu|^2)^s < \infty,$$

[1]P. Lax, Comm. Pure Appl. Math. **8** (1955), 615-633.

where the ν-th *Fourier coefficient* $\widehat{u}(\nu)$ is given by

$$\widehat{u}(\nu) := (2\pi)^{-n/2} \int_{T^n} u(x)\, e^{-2\pi i \langle \nu, x\rangle} dx = \langle u, e_\nu \rangle_{L^2(T^n)}\,, \text{ with}$$
$$\langle \nu, x\rangle := \nu_1 x_1 + \cdots + \nu_n x_n \text{ and } |\nu|^2 = \langle \nu, \nu\rangle\,.$$

Then the series $\sum_\nu \widehat{u}(\nu) e_\nu$ converges absolutely to u in the $W^s(T^n)$ topology.

Step 2: For $x = (y,\theta) \in T^{n-1} \times S^1$ and $u \in C^\infty(T^n)$, we have $u(y,\theta) = \sum_{(\lambda,\mu)} \widehat{u}(\lambda,\mu) e_\lambda(y) e_\mu(\theta)$, where the sum is over all $(\lambda,\mu) \in \mathbb{Z}^{n-1} \times \mathbb{Z}$ and the convergence on T^n is uniform. Since $e_\mu(0) = (2\pi)^{-1/2}$, it follows that

$$u(y,0) = (2\pi)^{-1/2} \sum\nolimits_{\lambda \in \mathbb{Z}^{n-1}} e_\lambda(y) \sum\nolimits_{\mu \in \mathbb{Z}} \widehat{u}(\lambda,\mu),$$

where the series converges uniformly on T^n. Thus, we have

$$(u|_{T^{n-1}})\widehat{\ }(\lambda) = (2\pi)^{-1/2} \sum\nolimits_{\mu \in \mathbb{Z}} \widehat{u}(\lambda,\mu).$$

Step 3: We essentially follow [**Pal65**, p.143f] (but avoid a minor error, the last inequality on p. 143). For $s \geq \frac{1}{2}$, $b \geq 1$ and $a : \mathbb{Z}^+ \to \mathbb{R}_+$, we set

$$x_\mu = b^{-1/4} \left(1 + \tfrac{\mu^2}{b}\right)^{-s/2} \text{ and } y_\mu = a_\mu b^{1/4} \left(1 + \tfrac{\mu^2}{b}\right)^{s/2}.$$

The Schwarz inequality $\left(\sum_{\mu \in \mathbb{Z}} x_\mu y_\mu\right)^2 \leq \sum_{\mu \in \mathbb{Z}} x_\mu^2 \sum_{\mu \in \mathbb{Z}} y_\mu^2$ yields

$$\left(\sum\nolimits_{\mu \in \mathbb{Z}} a_\mu\right)^2 \leq \sum\nolimits_{\mu \in \mathbb{Z}} b^{-1/2} \left(1 + \tfrac{\mu^2}{b}\right)^{-s} \sum\nolimits_{\mu \in \mathbb{Z}} a_\mu^2 b^{1/2} \left(1 + \tfrac{\mu^2}{b}\right)^{s}$$

or

$$\left(\sum\nolimits_{\mu \in \mathbb{Z}} a_\mu\right)^2 b^{s-\frac{1}{2}} \leq \sum\nolimits_{\mu \in \mathbb{Z}} b^{-1/2} \left(1 + \tfrac{\mu^2}{b}\right)^{-s} \sum\nolimits_{\mu \in \mathbb{Z}} a_\mu^2 \left(b + \mu^2\right)^{s}.$$

By integral comparison,

$$\begin{aligned}\sum\nolimits_{\mu \in \mathbb{Z}} b^{-1/2} \left(1 + \tfrac{\mu^2}{b}\right)^{-s} &\leq b^{-1/2} + 2 \int_0^\infty b^{-1/2} \left(1 + \tfrac{x^2}{b}\right)^{-s} dx \\ &= b^{-1/2} + 2 \int_0^\infty \left(1 + y^2\right)^{-s} dx = b^{-1/2} + \frac{\sqrt{\pi}\,\Gamma\left(s - \frac{1}{2}\right)}{\Gamma(s)}.\end{aligned}$$

For $C_s := 1 + \frac{\sqrt{\pi}\Gamma\left(s-\frac{1}{2}\right)}{\Gamma(s)}$, we have (since $b \geq 1$)

$$\left(\sum\nolimits_{\mu \in \mathbb{Z}} a_\mu\right)^2 b^{s-\frac{1}{2}} \leq C_s \sum\nolimits_{\mu \in \mathbb{Z}} a_\mu^2 \left(b + \mu^2\right)^{s}.$$

We set $a_\mu := |\widehat{u}(\lambda,\mu)|$ and $b := 1 + |\lambda|^2$ and then obtain

$$\left(\sum\nolimits_{\mu \in \mathbb{Z}} |\widehat{u}(\lambda,\mu)|\right)^2 \left(1 + |\lambda|^2\right)^{s-\frac{1}{2}} \leq C_s \sum\nolimits_{\mu \in \mathbb{Z}} |\widehat{u}(\lambda,\mu)|^2 \left(1 + |\lambda|^2 + \mu^2\right)^{s},$$

Thus, with the above Step 2, we have

$$\begin{aligned}\|(u|_{T^{n-1}})\|_{s-\frac{1}{2}}^2 &= \sum\nolimits_{\lambda \in \mathbb{Z}^{n-1}} |(u|_{T^{n-1}})\widehat{\ }(\lambda)|^2 (1 + |\lambda|^2)^{s-\frac{1}{2}} \\ &= \sum\nolimits_{\lambda \in \mathbb{Z}^{n-1}} \left|(2\pi)^{-1} \sum\nolimits_{\mu \in \mathbb{Z}} \widehat{u}(\lambda,\mu)\right|^2 (1 + |\lambda|^2)^{s-\frac{1}{2}} \\ &\leq (2\pi)^{-1} \sum\nolimits_{\lambda \in \mathbb{Z}^{n-1}} \left(\sum\nolimits_{\mu \in \mathbb{Z}} |\widehat{u}(\lambda,\mu)|\right)^2 (1 + |\lambda|^2)^{s-\frac{1}{2}} \\ &\leq (2\pi)^{-1} C_s \sum\nolimits_{\lambda \in \mathbb{Z}^{n-1}} \sum\nolimits_{\mu \in \mathbb{Z}} |\widehat{u}(\lambda,\mu)|^2 \left(1 + |\lambda|^2 + \mu^2\right)^{s}.\end{aligned}$$

Thus,

$$\left\|(u|_{T^{n-1}})\right\|_{s-\frac{1}{2}} \leq \sqrt{C_s/2\pi}\,\|u\|_s\,,$$

and we are done, since $C^\infty(T^n)$ is dense in $W^s(T^n)$. □

The following case study gives insight into the possible loss of differentiability under restrictions of Sobolev spaces, in contrast to the gain of differentiability in the C^k theory:

THEOREM 7.18. *There is no continuous, linear map $W^s(\mathbb{R}^n) \to W^s(\mathbb{R}^{n-1})$ which extends the restriction map $C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^{n-1})$ defined by $u \mapsto u(\cdot, 0)$.*

PROOF. On the ball $B^n := \{x \in \mathbb{R}^n : |x| \leq 1\}$, the function $|x|^\alpha$ is integrable, if $\alpha > -n$, since then in polar coordinates, we have

$$\int_{B^n} |x|^\alpha \, dx \leq C \int_0^1 r^{\alpha+n-1} dr < \infty.$$

Now consider the function $u(x) := |x|^\alpha \chi(x)$, where χ is a C^∞ function with compact support and $\chi(x) = 1$ for all $x \in B^n$. For $\alpha = -1/2$ and $n = 2$, we have $u \in L^2(\mathbb{R}^2)$, but $u(\cdot, 0) \notin L^2(\mathbb{R})$, since $\int_0^1 x^{-1} dx = \infty$. Thus, the theorem is verified for $s = 0$, since there is a sequence $\{u_\nu\}_{\nu=1}^\infty$ with $u_\nu \in C_0^\infty(\mathbb{R}^2)$ which converges in $W^0(\mathbb{R}^2)$ $(= L^2(\mathbb{R}^2))$ to u, but $\{u_\nu(\cdot, 0)\}_{\nu=1}^\infty$ is not a Cauchy sequence in $W^0(\mathbb{R})$ $(= L^2(\mathbb{R}))$. One can also easily construct counterexamples for $s > 0$, since by the above argument it follows that the function u (defined there) lies in $W^1(\mathbb{R}^n)$ exactly when $2\alpha > 2 - n$ (see Exercise 7.2a). For example, $u(x) := |x|^{-1/4} \chi(x)$ is an element of $W^1(\mathbb{R}^3)$, but $u(\cdot, \cdot, 0) \notin W^1(\mathbb{R}^2)$ and $u(\cdot, 0, 0) \notin W^1(\mathbb{R})$. If a sequence $u_\nu \in C_0^\infty(\mathbb{R}^3)$ converges to $u \in W^1(\mathbb{R}^3)$ and the restrictions $u_\nu(\cdot, 0, 0)$ were to converge in $W^1(\mathbb{R})$, then the limit in $W^1(\mathbb{R})$ would also be in $C^0(\mathbb{R})$ by Theorem 7.16; this contradicts the form of u. □

THEOREM 7.19. *Without the assumption that X is compact, Theorem 7.15 above is false.*

PROOF. For each $n \in \mathbb{N}$, one constructs $u_n \in W^1(\mathbb{R})$ with $\|u_n\|_1 < 3$, as in Figure 7.1. However, we have $\|u_n - u_{2n}\|_0^2 \geq 2n\left(1/\sqrt{n} - 1/\sqrt{2n}\right)^2 = \left(\sqrt{2} - 1\right)^2$,

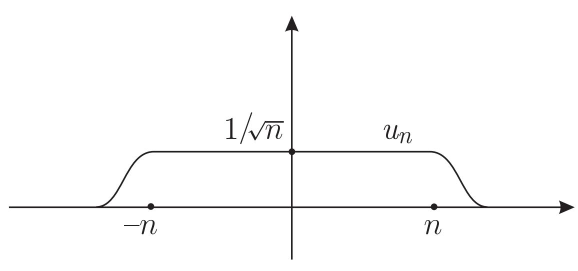


FIGURE 7.1. No *Rellich* Theorem for non-compact $X := \mathbb{R}$

independent of n. Thus, the u_n lie in a bounded subset of $W^1(\mathbb{R})$, but there is no subsequence convergent in $W^0(\mathbb{R}) = L^2(\mathbb{R})$. □

CHAPTER 8

Pseudo-Differential Operators

Synopsis. Motivation: Fourier Inversion; Symbolic Calculus; Quantization. *Canonical* and *Principally Classical* Pseudo-Differential Operators. Pseudo-Locality; Singular Support. Standard Examples: Differential Operators; Singular Integral Operators. Oscillatory Integrals. Kuranishi Theorem. Change of Coordinates. Pseudo-differential Operators on Manifolds. Graded *-Algebra. Invariant Principal Symbol; Exact Sequence; Non-Canonical Op-Construction as Right Inverse. Coordinate-Free (Truly Global) Approach: Normalized Fourier Transformation; Normalized Amplitudes; Normalized Invertible Op-Construction; Approximation of Differential Operators

1. Motivation

For better reading by a traditionally educated student, this chapter is based on a *local* definition of pseudo-differential operators and then generalizing to operators acting on sections of vector bundles over manifolds by charts and local trivializations of the bundles. That approach has its merits since it has become the standard way of introducing pseudo-differential operators and since it admits some easy elementary calculations. However, for *geometrically defined operators*, the arbitrary character of the coordinate shifts does not facilitate calculations and even can block for natural constructions (like the product of pseudo-differential operators in non-trivial cases).

Whence, a self-confident reader may skip the first four sections of this chapter and advance directly to Section 8.5 where we give a *coordinate-free* description of pseudo-differential operators. As we shall see in our Part III, that global description is more powerful for the investigation of the *analytical index* under *embedding*. That said, the reader must be reminded that also our global description of pseudo-differential operators is *neither really invariant* nor *canonical*: While it does not depend on coordinates, it depends heavily on other choices, namely the choice of metric structures on the manifold and the involved bundles and on the choice of connections for that bundles.

We will now turn to a class of operators which, roughly speaking (details below), are locally presentable in the form

$$(Pu)\,(x) := \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p(x,\xi)\widehat{u}(\xi)\, đ\xi,$$

where we use the convenient shorthand

$$đ\xi = (2\pi)^{-n/2}\, d\xi$$

for Lebesgue measure on $\mathbb{R}^n$ divided by $(2\pi)^{n/2}$. Here p is called the **amplitude** of the operator P, $\langle x,\xi\rangle$ is its **phase function**, and

$$\widehat{u}(\xi) := \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} u\,(x)\; đx$$

denotes the Fourier transform of u (see the crash course in Appendix A).

There are a number of reasons why these *pseudo-differential operators* have commanded increasing attention since the appearance of the pioneering studies by Solomon Grigoryevich Mikhlin on *Singular Integral Equations* (1948). We mention the following overlapping aspects.

1. This class is large enough to contain in addition to the differential operators (Exercise 8.5 below, p.209) the Green operators (see also Chapter 2) and other *singular integral operators* which play a role in solving partial differential equations. In particular this class of pseudo-differential operators contains, with each elliptic operator, its *parametrix*, i.e., a quasi-inverse modulo an operator of *lower order*. In Theorem 9.7 (p.236) below we will incorporate this operator calculus into Hilbert space theory and in this fashion, we will be able to derive easily the classical results on elliptic operators (regularity theorems, finiteness of the index) using the elementary theory of Fredholm operators developed in Chapters 1-3. Thereby "some of the techniques used in the case of differential operators appear here as general properties of the class of integro-differential operators considered" (Seeley).

2. The class is small enough and close enough to the differential operators to allow convenient computations. This standpoint is important, particularly because the progress in functional analysis of the past decades permitted the definition of more and more general and involved operator classes and *phantom spaces* (Thom), while the exploration of their properties was too difficult and lagged behind. In contrast, turning to pseudo-differential operators, for which an exact calculus was developed, signalled "a trend in the theory of general partial differential equations towards essentially constructive methods" (Hörmander).

3. A special aspect is the attempt to deal with differential operators with variable coefficients, by means of pseudo-differential operators in *first* approximation, in the same way differential operators with constant coefficients are treated by means of the Fourier transform: For example, for $f \in C^\infty(\mathbb{R}^n)$ with compact support and $n > 2$, consider the inhomogeneous equation $\Delta u = f$, where $\Delta = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_n^2$ is the Laplace operator. With the Fourier transform (see the multiplication rule in Exercise A.5b, Appendix A), we obtain $-(\xi_1^2 + \cdots + \xi_n^2)\widehat{u}(\xi) = \widehat{\Delta u}(\xi) = \widehat{f}(\xi)$, i.e., $\widehat{u}(\xi) = -\widehat{f}(\xi)/|\xi|^2$, as an L^2 function (for $n > 2$), and further with the Fourier inversion formula,

$$(Qf)(x) = u(x) = -\int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} |\xi|^{-2} \widehat{f}(\xi) d\!\!\!{}^-\xi,$$

where Q is the inverse operator (*fundamental solution*) of Δ. In general, suppose P is a differential operator with *constant coefficients* which can be written as a polynomial $P = p(D)$ where $D = (-i\partial/\partial x_1, \dots, -i\partial/\partial x_n)$, and consider the inhomogeneous equation $p(D)u = f$, $f \in C^\infty(\mathbb{R}^n)$ of compact support. We obtain in the same way, at least formally, a solution $u = Qf$, where

$$(Qf)(x) := \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} q(\xi) \widehat{f}(\xi)\, d\!\!\!{}^-\xi,$$

and $q(\xi) := p(\xi)^{-1}$ is the amplitude. In the process, a number of difficulties arise. Indeed, Qf is in general not C^∞, and possibly only a distribution, and the integral must be interpreted, since the zeros of p can cause divergences. But these problems can be resolved almost completely; e.g., see the following references of quite different depth: [**Gru**, pp.108-110],[**Hö63**, Chs. III and IV], [**Hö2**], [**Tay1**, Chapter 3].

Now, if (as in Chapter 5) $U \subseteq \mathbb{R}^n$ is open and

$$P = p(x, D) = \sum\nolimits_{|\alpha| \le k} a_\alpha(x) D^\alpha, \text{ with } a_\alpha \in C^\infty(U),$$

is a differential operator with *variable coefficients*, then all these methods fail initially. But we can, "as a good physicist would" (Atiyah), formally invert the operator P by *freezing* the coefficients at a point $x_0 \in U$ and considering P as a perturbation of $p(x_0, D)$ which is a differential operator with constant coefficients. In this way we obtain as an approximate inverse of P a pseudo-differential operator with the amplitude $q(\xi) = p(x_0, \xi)^{-1}$. In order to get a better approximate inverse it is natural to slowly *thaw* the coefficients, i.e., to let the point x_0 vary in U. This yields the operator

$$(Qf)(x) := \int_{\mathbb{R}^n} e^{i\langle x, \xi\rangle} q(x, \xi) \widehat{f}(\xi)\, đ\xi$$

with the amplitude $q(x, \xi) = p(x, \xi)^{-1}$ ($x \in U, f \in C^\infty(\mathbb{R}^n)$ with compact support). This basic *perturbation argument*, which was supplied in the study of elliptic differential equations by the Italian mathematician Eugenio Elia Levi already in the year 1907, thus finds its theoretical framework within the class of pseudo-differential operators.

We remark (see Theorem 9.7, p. 236) that in the elliptic case an *equally good* approximation is obtained by choosing as amplitude the inverse of the principal part (symbol), i.e., the function $p_k(x, \xi)^{-1}$ which is homogeneous of degree $-k$ in ξ. There is a particularly simple calculus of such operators, since the asymptotic expansions of $p(x, \xi)$ and $q(x, \xi)$ and the underlying iteration (usually necessary) is avoidable here.

4. The theory of pseudo-differential operators allows a certain *relaxing* of customary precision, a precision which is senseless, or at least exaggerated, in a number of practical problems. Thus, in order to investigate regularity and solvability of the differential equation $Pu = f$, we do not need an actual inverse operator (fundamental solution), but (in the framework of Fredholm theory) it suffices to have a *parametrix*, i.e., a *quasi-inverse* modulo certain elementary operators (see Chapter 3). This has considerable computational advantages. In the case of differential operators $P = p(D)$ with constant coefficients (as in Item 3 above), we may take the amplitude to be

$$q(\xi) := \chi(\xi) p(\xi)^{-1}$$

where $\chi(\xi)$ is a fixed C^∞ function which is identically zero in a disk about the origin and identically 1 for large ξ. In this fashion we avoid the delicate convergence problems for the integral

$$(Qf)(x) := \int_{\mathbb{R}^n} e^{i\langle x, \xi\rangle} q(\xi) \widehat{f}(\xi)\, đ\xi$$

which are required for the amplitude $p(\xi)^{-1}$ because of singularity at the zeros of p. While $PQ = \mathrm{Id}$ and $QP = \mathrm{Id}$ are *not* valid, we still have

$$PQf = f + Rf \quad \text{for } f \in C^\infty(U), \text{ where}$$
$$(Rf)(x) := \int_{\mathbb{R}^n} r(x - \xi) f(\xi)\, đ\xi.$$

and $\widehat{r} = \chi - 1$; so $r \in C^\infty$, and R is a *smoothing operator*. Indeed,

$$(P(Qf))(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p(\xi) \widehat{Qf}(\xi)\, đ\xi \quad \text{and}$$

$$\widehat{Qf}(\xi) = \int_{\mathbb{R}^n} e^{-i\langle y,\xi\rangle} \left(\int_{\mathbb{R}^n} e^{i\langle y,\eta\rangle} q(\eta)\, \widehat{f}(\eta)\, đ\eta \right) đy = q(\xi)\widehat{f}(\xi) = \chi(\xi) p(\xi)^{-1} \widehat{f}(\xi).$$

Using the convolution formula of Exercise A.5e in Appendix A, we then have

$$\begin{aligned}
&(P(Qf))(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p(\xi)\chi(\xi)p(\xi)^{-1}\widehat{f}(\xi) đy\, đ\xi \\
&= \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \left(1 + (\chi(\xi) - 1)\right) \widehat{f}(\xi)\, đ\xi \\
&= f(x) + \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \left(\chi(\xi) - 1\right) \widehat{f}(\xi)\, đ\xi = f(x) + \int_{\mathbb{R}^n} r(x - \xi)\, f(\xi)\; đ\xi.
\end{aligned}$$

Since much is known about simple *correction* or *residue* operators such as R, a parametrix Q serves just as well as a true fundamental solution for which R vanishes. At any rate, fundamental solutions do not usually exist when passing to variable coefficients in the perturbation method (sketched in Item 3) and when replacing operators and their amplitudes by *principal symbols* (the terms of highest order in the amplitudes). However, the simple computations modulo *smoothing operators* and other *operators of lower order* can be used very efficiently and arise naturally in the theory of pseudo-differential operators.

5. In Section 6.7 (p. 186) we pointed out, in connection with the symbolic calculus, the basic significance of the change of levels in passing from *operators* to the *functions* which characterize them in approximation. The pseudo-differential operators form a class (and this is tied to the perturbation argument) whose operators can at least in approximation (actually, precise in the global approach described below in Section 8.5) be described by their *amplitudes* and *symbols*. The latter are functions satisfying simple rules of computation resulting in a particularly simple approximation theory for the corresponding operators; see for example the composition rules in Theorem 8.26, p. 226. Surely, mathematicians such as Vito Volterra, Erik Ivar Fredholm, David Hilbert and Friedrich (Frigyes) Riesz had this goal in mind when they developed the theory of integral equations as a means of dealing with differential operators. However, starting with the classical representation

$$(Qf)(x) = \int K(x, z) f(z)\, dx,$$

it turns out that the formulation of the correct conditions for the *weights* K are by far not as simple and natural as are those for the *amplitudes* and *symbols* in the representation via Fourier transform. The same is true for the transformation and composition rules (see Theorem 8.3, p. 208, and Theorem 8.13, p. 213).

6. From the topological standpoint, the following issues are particularly important: In the passage between the levels of consideration, we prefer to operate with symbols (they are *more accessible by topological means*) rather than operators. In Part III below, the larger class of pseudo-differential operators has a decisive advantage. Indeed, the associated extension of the symbol space beyond the polynomial maps to arbitrary C^∞ functions permits *lifting a homotopy* of the symbol of a differential operator to a homotopy of the operator itself in the space of pseudo-differential operators. This is generally impossible in the space of differential operators. Using heavier topological machinery, this difficulty can be dealt with without the use of pseudo-differential operators. However, the difficulties which occur are not to be underestimated. For example, not much is known about the *simplest* question of the existence an elliptic system (in $\mathbb{R}^n$) of N differential equations of order k with constant coefficients. For $k = 1$, this is the case exactly when the $(N-1)$-sphere S^{N-1} admits $n-1$ linearly independent vector fields; hence, for $N = n$, (according to a famous theorem of John Frank Adams) exactly for the values 2, 4, and 8. More about this is in [**Ati70a**]. In [**Fu99**], M. Furuta presented a full proof of the Atiyah-Singer Index Theorem without the use of pseudo-differential operators.

7. Finally, we point out that the class of pseudo-differential operators originally was developed only in connection with elliptic differential equations, and only there (and with the closely related *hypo-elliptic* differential equations) the beautiful properties listed above unfold fully. However, Lars Hörmander and other authors succeeded, in a series of papers and monographs, in generalizing the concept of a pseudo-differential operator in such a way that the theory *Fourier integral operators* so created leads to new results also in heat transfer and wave operators, for example. For this aspect, which we cannot pursue further, see [**Hö71b, Hö4**], and [**Tay74**, Chapters II, IV, VI and VII].

2. Canonical Pseudo-Differential Operators

We begin with the definition of the prototypes of our pseudo-differential operators in local form over the open subset $U \subseteq \mathbb{R}^n$ and acting on functions only:

$$(Pu)(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p\,(x,\xi)\,\widehat{u}(\xi)\,đ\xi =: (\mathrm{Op}(p)u)\,(x), \tag{8.1}$$

where $x \in U$, $u \in C_0^\infty(U)$ (i.e., $u \in C^\infty(U)$ and the support of u is compact).

Definition 8.1. a) The operator $P = \mathrm{Op}(p)$ is called a **canonical pseudo-differential operator of order** $k \in \mathbb{R}$, shortly $P \in \mathrm{L}^k(U)$, if the amplitude $p \in C^\infty(U \times \mathbb{R}^n)$ satisfies the following *asymptotic conditions* of growth as $|\xi| \to \infty$: For each compact subset $K \subset U$ and multi-indices $\alpha = (\alpha_1, ..., \alpha_n)$, $\beta = (\beta_1, ..., \beta_n) \in \mathbb{Z}_+^n$, there is a $C \in \mathbb{R}$ such that for all $x \in K$ and $\xi \in \mathbb{R}^n$, we have

$$\left|D_x^\beta D_\xi^\alpha p(x,\xi)\right| \le C\,(1+|\xi|)^{k-|\alpha|}\,. \tag{8.2}$$

Recall that $|\alpha| := |\alpha_1 + \cdots + \alpha_n|$ and that $D_x^{(0,\ldots,1,\ldots 0)} := -i\partial/\partial x^j$ where "1" stands in the j-th place of the multi-index $(0,\ldots,1,\ldots 0)$.
b) The set of amplitudes satisfying (8.2) will be denoted by $\mathrm{S}^k(U \times \mathbb{R}^n)$. We set $\mathrm{S}^\bullet(U \times \mathbb{R}^n) := \bigcup_{k\in\mathbb{R}} \mathrm{S}^k(U \times \mathbb{R}^n)$ and $\mathrm{L}^\bullet(U) := \bigcup_{k\in\mathbb{R}} \mathrm{L}^k(U)$.

Remark 8.2. a) Amplitudes satisfying (8.2) are often called symbols of Hörmander type $(1,0)$. The definition extends to matrix valued amplitudes, needed below for defining pseudo-differential operators acting on sections of vector bundles.
b) A global version of the estimates (8.2) is given below in (8.16).
c) The best constants in (8.2) provide a set of semi-norms which endow $\mathrm{S}^\bullet(U \times \mathbb{R}^n)$ with the structure of a Fréchet algebra.

The estimate (8.2) plays a key role in the derivation of many useful properties of pseudo-differential operators, as in the following

THEOREM 8.3. *Each canonical pseudo-differential operator is a linear map from $C_0^\infty(U)$ to $C^\infty(U)$.*

PROOF. For all $\xi \in \mathbb{R}$, the integrand $x \mapsto e^{i\langle x,\xi\rangle} p(x,\xi)\, \widehat{u}(\xi)$ is obviously a C^∞ function. To show that the function

$$x \mapsto (\mathrm{Op}(p)u)(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p(x,\xi)\, \widehat{u}(\xi)\, đ\xi$$

is also C^∞, we must show that the integral converges *sufficiently well* so that the order of integration and differentiation may be switched. More precisely, by the Dominated Convergence Theorem of Henri Lebesgue, a function which is the limit of a sequence of measurable functions, uniformly bounded by an (absolutely) integrable function, is itself integrable and the *limit* and *integral* may be interchanged. To apply this to our situation, we must show that, for each $x \in U$ and each multiindex β, the function

$$\xi \mapsto \left| D_x^\beta \left(e^{i\langle x,\xi\rangle} p(x,\xi) \right) \widehat{u}(\xi) \right| \quad (\xi \in \mathbb{R}^n)$$

can be estimated by an integrable function. Since the support of u is compact, we have (see Exercise A.5b, p. 675 of Appendix A) that

$$\xi^\alpha \widehat{u}(\xi) = \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} D^\alpha u(x)\; đx,$$

which goes to 0 as $|\xi| \to \infty$. Hence the function $\xi \mapsto |\xi^\alpha \widehat{u}(\xi)|$ is bounded for each multiindex α. Thus, $\widehat{u}(\xi)$ decreases faster than any power of $|\xi|^{-1}$ as $|\xi| \to \infty$; i.e., for each N there is a constant C_1 such that for all $\xi \in \mathbb{R}^n$,

$$|\widehat{u}(\xi)| \leq C_1 (1 + |\xi|)^{-N}.$$

By (8.2), we have $\left| D_x^{\beta'} p(x,\xi) \right| \leq C_2 (1 + |\xi|)^k$ for any $\beta' \leq \beta$. Hence

$$\left| D_x^\beta \left(e^{i\langle x,\xi\rangle} p(x,\xi) \right) \right| \leq C_3\, |\xi|^{|\beta|} (1 + |\xi|)^k, \text{ and so}$$

$$\left| D_x^\beta \left(e^{i\langle x,\xi\rangle} p(x,\xi) \right) \widehat{u}(\xi) \right| \leq C_3 (1 + |\xi|)^{k+|\beta|-N},$$

the right side being integrable for N sufficiently large. □

REMARK 8.4. a) Recall roughly that quantization in quantum mechanics attempts to convert functions of position and momentum (i.e., functions on T^*X) into operators. One may think of $\mathrm{Op}(p)$ as a quantization of p. See also [**GS**, Exercise 3.1, p.36f]. There you find a sketch of how, e.g., the common commutator relations of *Weyl quantization* can be derived from properties of Op. Readers, however, who have a geometric antenna and are truly interested in physics will be bothered by the fact that our $\mathrm{Op}(p)$ is naturally defined only in Euclidean space. A generalization of that quantization concept to manifolds depends on many choices. The common way depends on the choice of charts. That does not lead very far, see our Exercise 8.19, p.221 and our Chapter 9 or [**GS**, Exercise 3.4] and [**Gru**, Section 8.2]. Below in Section 8.5, we shall describe an alternative, thoroughly global way to define $\mathrm{Op}(p)$ for suitable p over a closed manifold. However, that approach will also depend on many choices (e.g., the choice of metric, connection and bump function), as we will see. Apart from these choices, there are other choices one can make. As [**GS**, Chapter 11] indicates, one probably has to return to the visionary notes by J.B. Keller, 1958, V.P. Maslov, 1972, and J. Leray, 1981 (precise references are

given in [**GS**, p.130]) to find ideas for a quantization concept which is physically realistic and geometrically meaningful.
b) The continuity of the operator $\mathrm{Op}(p)$ will be discussed later, see Theorem 9.2, p.234.
c) We also postpone the discussion, whether an amplitude is determined from a given pseudo-differential operator to Example 8.15, p.215.

Exercise 8.5. Show that the following standard operators define canonical pseudo-differential operators modulo smoothing operators (i.e., pseudo-differential operators, whose amplitudes have compact support in the second variable; see also Remark 8.8, p. 210).
a) $P = \sum_{|\alpha| \le k} a_\alpha D^\alpha$, where $a_\alpha \in C^\infty(U)$.
b) $(Pu)(x) = \int_{\mathbb{R}^n} K(x,y)u(y)dy$, where $K \in C^\infty(U \times U)$ and the support of $K(x,\cdot)$ is compact for all $x \in U$. For example, the convolution $u \mapsto u * \varphi$ with $\varphi \in C_0^\infty(U)$ (where $K(x,y) = \varphi(x-y)$) is a canonical pseudo-differential operator.
c) The *Riesz operator* $P = \sum a_\alpha R^\alpha$, where $a_\alpha \in C_0^\infty(U)$, with $a_\alpha = 0$ for all but finitely many multi-indices α, and

$$(R^\alpha u)(x) := \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \left(\tfrac{\xi}{|\xi|}\right)^\alpha \widehat{u}(\xi) đ\xi.$$

(See the footnote for Exercise 8.23, p. 223 below.)
[Hint: For a: One applies the differential operator P to

$$u(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \widehat{u}(\xi)\, đ\xi,$$

and then obtains a pseudo-differential operator with amplitude

$$p(x,\xi) = \sum\nolimits_{|\alpha| \le k} a_\alpha(x)\, \xi^\alpha, \ x \in U, \ \xi \in \mathbb{R}^n.$$

For b: Here also, we begin with the Fourier Inversion Formula. One obtains the amplitude

$$p(x,\xi) = \int_{\mathbb{R}^n} e^{i\langle y,\xi\rangle} K(x,y)\, dy.$$

For each fixed x, this is a multiple of the inverse Fourier transform of a function with compact support, and hence $p(x,\cdot) \in C_\downarrow^\infty(\mathbb{R}^n)$, the space of rapidly decreasing functions in $C^\infty(\mathbb{R}^n)$. (Argue as in the proof of Theorem 8.3, where partial integration interchanges multiplication and differentiation.) Thus, the required conditions on the amplitude hold for each $k \in \mathbb{Z}$.
For c: $p(x,\xi) = \chi(\xi) \sum a_\alpha(x) (\xi/|\xi|)^\alpha$ where $\chi(\xi)$ is a *cut-off function*, i.e., $\chi(\xi) = 0$ for $|\xi| < \rho$, $\rho > 0$, and $\chi(\xi) = 1$ for $|\xi| > \rho' > \rho$. The order of P is therefore $k = 0$, and different choices of χ lead, modulo smoothing functions, to the same pseudo-differential operator. One also calls Riesz operators *singular integral operators*, since they can be alternatively represented by singular convolutions. For example, if $\alpha = (1,0,...,0)$, then one can (up to a constant factor, which we will ignore) write

$$(a_\alpha R^\alpha u)(x) = \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} K(x,x-y)u(y)\, dy, \text{ where } K(x,z) := a_\alpha(x) z_1 |z|^{-n-1},$$

whence the weight function K has a singularity at the point $z = 0$. For the connection between Riesz operators, Hilbert transformations, and Wiener-Hopf operators in the case $n = 1$, see above our Chapter 4, [**Tay74**, p.36], and [**Prö72**], where an algebra of *pseudo-multiplication operators* in the half-space $\mathbb{R}^n_+$ is investigated. This algebra is formed with the help of pseudo-differential operators and contains the Wiener-Hopf operators.]

REMARK 8.6. While differential operators are *local* operators (see Exercise 5.2, p.136), a pseudo-differential operator can increase supports. For example, if P is defined as convolution with $\varphi \in C^\infty(U)$ as in Exercise 8.5b, we can have $\operatorname{supp} Pu = \operatorname{supp} u + \operatorname{supp} \varphi$. The *translation operator* with amplitude

$$p(x,\xi) := e^{i\langle x_0,\xi\rangle},\ x_0 \text{ fixed},$$

which sends $u(x)$ to $u(x+x_0)$ is *not* a canonical pseudo-differential operator. Actually, the asymptotic amplitude estimate guarantees **pseudo-locality**, a kind of locality modulo operators of *lower order*, whereby (rather than the support) the **singular support** (the closure of the set where a function is not C^∞) is not increased. Details are found in [**Gru**, p.177], [**Hö3**, Theorem 18.1.16], [**Ni**, p.151f] or [**Pal65**, p.260].

EXERCISE 8.7. Show that one can write every canonical pseudo-differential operator P as an integral operator (for some $\lambda \in \mathbb{R}$)

$$(Pu)(x) = \int_U K_\lambda(x, x-y)(1-\Delta)^\lambda u(y)\, đy,\ u \in C_0^\infty(U),$$

where the weight function $K_\lambda(x,z)$ is C^∞ for $z \neq 0$.
[Hint: Suppose the operator P has order $k \in \mathbb{Z}$ and amplitude $p \in C^\infty(U \times \mathbb{R})$. The case $k < -n$ is easily analyzed: Without loss of generality, suppose $z \neq 0$ and show as in the proof of Theorem 8.3 (repeated partial integration) that $K(x,z) := \int e^{i\langle z,\xi\rangle} p(x,\xi)\, đ\xi$ is actually C^∞ for $z \neq 0$, and $\lambda = 0$. $K(x,z)$ is continuous even at $z = 0$, and differentiable there for k sufficiently negative. In the case $k \geq -n$, formally write the integral for $K(x,z)$ as in the case $k < -n$ where $\lambda = 0$. The integral need not converge, since one no longer has the estimate

$$|p(x,\xi)| \leq (1+|\xi|)^{-n-1}.$$

Hence, insert the factor $(=1)$

$$(1+|\xi|^2)^\lambda (1+|\xi|^2)^{-\lambda}$$

in the integrand. Then note that for the usual Laplace operator $\Delta_z = -\sum_{j=1}^n D_j^2$ ($D_j = -i\partial/\partial z_j$), we have

$$e^{i\langle z,\xi\rangle}(1+|\xi|^2)^\lambda = (1-\Delta_z)^\lambda e^{i\langle z,\xi\rangle}.$$

If $\lambda > k+n$, the case $k \geq -n$ reduces to the case $\lambda = 0$. Details are found in [**Ni**, p.152].]

REMARK 8.8. By Definition 8.1, a canonical pseudo-differential operator, whose amplitude has compact support in the second variable, is of arbitrarily small order ("$k = -\infty$"), and so it can be presented as an integral operator with C^∞ weight function (i.e., a **smoothing operator**).

REMARK 8.9. Contrary to the classical notation for integral operators (where the singularities of the weight function lie on the diagonal of $U \times U$), we write $K(x, x-y)$ instead of $K(x,y)$ under the integral, and by this artifice obtain a weight function $K(x,z)$ that is singular only at $z = 0$.

3. Principally Classical Pseudo-Differential Operators

In order to adapt the theory of pseudo-differential operators to our problem of treating elliptic differential equations, first on closed manifolds and then on bordered domains, we must solve two problems.

Task 1 - Homogeneous Principal Symbol. Instead of the weight function K or the amplitude p, we require the notion of *principal symbol*, a sort of homogeneous main part of the amplitude. For a differential operator $P = \sum_{|\alpha|\le k} a_\alpha D^\alpha$, the amplitude was the polynomial function (Exercise 8.5a)

$$p(x,\xi) = \sum_{|\alpha|\le k} a_\alpha(x)\xi^\alpha, \quad x \in U, \xi \in \mathbb{R}^n.$$

From this, the principal symbol (or shortly "symbol") $\sigma(P)(x,\xi)$ of P was taken to be the homogeneous polynomial in ξ of order k obtained by taking the sum only over the terms of highest order ($|\alpha| = k$); see Chapter 5. This process of *separation* does not carry over to the amplitude of an arbitrary canonical pseudo-differential operator. Thus, we make the following four assumptions about the amplitude p of a pseudo-differential operator of order $k \in \mathbb{Z}$:

Assumptions 8.10. (i) For each compact subset $K \subset U$ and multi-indices $\alpha, \beta \in \mathbb{Z}_+^n$, there is a $C \in \mathbb{R}$ such that for all $x \in K$ and $\xi \in \mathbb{R}^n$, we have

$$\left|D_x^\beta D_\xi^\alpha p(x,\xi)\right| \le C\,(1+|\xi|)^{k-|\alpha|}.$$

(ii) The limit $\sigma_k(p)(x,\xi) := \lim_{\lambda\to\infty} \frac{p(x,\lambda\xi)}{\lambda^k}$ exists for all $x \in U$ and $\xi \in \mathbb{R}^n \setminus \{0\}$.

(iii) For some cut-off function $\chi \in C^\infty(\mathbb{R}^n)$ with $\chi(\xi) = \begin{cases} 0, & \text{for } |\xi| \text{ small,} \\ 1, & |\xi| \ge 1, \end{cases}$
$p(x,\xi) - \chi(\xi)\sigma_k(p)(x,\xi)$ is the amplitude of a canonical pseudo-differential operator of order $k-1$.

(iv) $p(x,\xi)$ has compact support in the variable x.

Note that (i) is just a repeat of (8.2). For us, conditions (iii) and (iv) serve only a technical purpose, since we then obtain convergence of integrals and estimates more easily (e.g., see the above hint to Exercise 8.7b). Actually, one can forgo these conditions and, as in Theorem 8.13 (p. 213), go over to a *Fourier integral operator* with a three-slot amplitude. For applications, we must drop these further assumptions, and we do so for the additional reason that we define our *global pseudo-differential operators* so that they possess amplitudes with compact support only in their *localized form* (see below).

In contrast to the *canonical pseudo-differential operators*, whose amplitudes only satisfy the estimate (i), we now say that P is a **pseudo-differential operator** (with compact support), if the amplitude p of P meets all four conditions (i) – (iv). We shall write $P \in \mathrm{L}^k_{\mathrm{pc}}(U)$, where the acronym "pc" stands for **principally classical** in accordance with one branch of modern literature, see [**BCLZ**, Section 2.3].

Remark 8.11. a) Main stream deals with **classical** pseudo-differential operators (written $P \in \mathrm{CL}^k(U)$). That are operators generated by elements of the subspace $\mathrm{CS}^k(U\times\mathbb{R}^n) \subset \mathrm{S}^k(U\times\mathbb{R}^n)$ consisting of **classical** (**polyhomogeneous**)

symbols. More precisely, an amplitude $p \in \mathrm{S}^k(U \times \mathbb{R}^n)$ belongs to $\mathrm{CS}^k(U \times \mathbb{R}^n)$, if it admits sequences $p_{k-j} \in \mathrm{C}^\infty(U \times \mathbb{R}^n), j \in \mathbb{Z}_+$ with

$$p_{k-j}(x, r\xi) = r^{k-j} p_{k-j}(x, \xi), \quad r \geq 1, |\xi| \geq 1, \tag{8.3}$$

such that

$$p - \sum_{j=0}^{N-1} p_{k-j} \in \mathrm{S}^{m-N}(U \times \mathbb{R}^n) \qquad \text{for all } N \in \mathbb{Z}_+. \tag{8.4}$$

The latter property is usually abbreviated $p \sim \sum_{j=0}^{\infty} p_{k-j}$.

b) Clearly we have $\mathrm{CL}^k(U) \subset \mathrm{L}^k_{\mathrm{pc}}(U) \subset \mathrm{L}^k(U)$, more precisely:

$$\mathrm{L}^k_{\mathrm{pc}}(U) = \mathrm{CL}^k(U) + \mathrm{L}^{k-1}(U). \tag{8.5}$$

For *index theory of elliptic operators*, it seems to us that the common restriction to *classical* pseudo-differential operators is not necessary. All we need can be done within the wider space $\mathrm{L}^\bullet_{\mathrm{pc}} = \bigcup \mathrm{L}^k_{\mathrm{pc}}$.

Task 2 - Manifolds and Coordinate Change. Our second task consists of defining pseudo-differential operators on a paracompact C^∞ *manifold* X. Thus, consider a linear map $P : C_0^\infty(X) \to C^\infty(X)$, where $C_0^\infty(X)$ again denotes the space of C^∞ functions with compact support. (We will consider operators on sections of vector bundles below in Exercise 8.20, p. 222). For each local coordinate system $\kappa : U \to \mathbb{R}^n$ with U open in X, P yields a *local operator*

$$P_\kappa u := P(\overline{u \circ \kappa}) \circ \kappa^{-1}, \quad u \in C_0^\infty(\kappa(U)), \qquad \text{where } \overline{u \circ \kappa} := \begin{cases} u \circ \kappa, & \text{on } U, \\ 0, & \text{on } X \setminus U. \end{cases}$$

DEFINITION 8.12. $P : C_0^\infty(X) \to C^\infty(X)$ is called a **pseudo-differential operator of order** k **on** X, if P_κ is a pseudo-differential operator (with compact support) for all C^∞ charts κ with relatively compact image. We write

$$P \in \mathrm{L}^k_{\mathrm{pc}}(X).$$

The definition seems to be analogous to the introduction of differential operators on manifolds. Actually, the situation here is different and more complex, since pseudo-differential operators do not need to be local (see the warning of Remark 8.6, p.210), while differential operators may actually be characterized by their locality (i.e., $\operatorname{supp} Pu \subseteq \operatorname{supp} u$); see Exercise 5.2, p. 136 and Remark 5.3. In particular, we have the following problems:

1. How *invariant* is the definition of $\mathrm{L}^k_{\mathrm{pc}}(X)$? Must one actually show that the induced *local operators* are pseudo-differential operators for all charts, or is it enough to check this for an atlas. This difficulty lies in the fact that the formation of the *local operators* is not *transitive*; i.e., in general, one obtains two different operators, if one first restricts a chart κ on $U \subseteq X$ to a open subset $U' \subseteq U$ obtaining $P_{(\kappa|_{U'})}$ and then considers the restriction P to $C_0^\infty(\kappa(U'))$. However, it turns out that the difference is a *smoothing operator* of the simple form treated in Exercise 8.5b.

2. With a differential operator P, the amplitude $p(x, \xi)$ and symbol $\sigma(P)(x, \xi)$ can be obtained *intrinsically* from the action of the operator, without explicitly

representing it in terms of local coordinates first. Namely, we have (see Exercise 6.34, p. 180 above) in a local chart

$$\sigma(P)(x,\xi)e = \frac{i^k}{k!} P\left((\varphi - \varphi(x))^k u\right)(x), \text{ with } d\varphi_x = \xi, u(x) = e,$$

and trivially

$$p(x,\xi) = e^{-i\langle x,\xi\rangle} P\left(\psi e^{i\langle x,\xi\rangle}\right)(x),$$

where $\psi \in C_0^\infty(X)$ with $\psi = 1$ in a neighborhood of x.

For pseudo-differential operators (in general, k is not positive) the first formula does not make sense, and there is no known simple fully *invariant* formula for the symbol of a pseudo-differential operator; the second formula holds only in an approximate sense (e.g., see [**Ni**, p.152f]); the amplitude of a pseudo-differential operator is not unique, but is only *asymptotically* determined by the operator. Thus, the task of defining a global symbol (for pseudo-differential operators defined on the whole manifold X) lies before us now. (Later, as announced before, we shall give a genuinely global definition of the principal symbol and the total symbol (amplitude), see Section 8.5 below.)

3. For this, we investigate the behavior of the *local operators* and their symbols under a coordinate change, and determine the transformation rule in order to obtain a global symbol. These calculations are somewhat lengthy, since under a coordinate change, the phase $\langle x,\xi\rangle$ and the amplitude $p(x,\xi)$ cannot be directly expressed in the form $\langle y,\eta\rangle$ and $q(y,\eta)$ in the new coordinates. By passing over to an apparently larger operator class (the so-called *Kuranishi Trick*, see also [**GS**, p.34f]), one can drastically simplify these computations (Theorem 8.18, p. 220), as well as the derivation of the composition rules, the formula for the symbol of the adjoint operator (Theorem 8.26, p. 226), and the multiplicative properties under *tensor product* (see [**Pal65**, p.206-209] and [**Hö71b**, p.96]).

The Kuranishi Trick. The following theorem is our entrance ticket to the micro-local analysis of pseudo-differential operators on manifolds. It is also of independent interest.

THEOREM 8.13 (M. Kuranishi, 1969). *Let $U \subseteq \mathbb{R}^n$ be open and $k \in \mathbb{Z}$. Let Q be an operator of the form*

(8.6)
$$(Qu)(x) = \int_{\mathbb{R}^n}\int_U e^{i\varphi(x,y,\xi)} q(x,y,\xi)\, u(y)\, đy\, đ\xi,\ x \in U,\ u \in C_0^\infty(U) =: \big(\mathrm{Op}(q)u\big)(x),$$

where the phase function φ is C^∞ and real-valued on $U \times U \times \mathbb{R}^n$, and linear in the variable ξ with

$$(8.7)\qquad \tfrac{\partial\varphi}{\partial\xi_1}(x,y,\xi) = \cdots = \tfrac{\partial\varphi}{\partial\xi_n}(x,y,\xi) = 0 \quad \text{for } \xi \neq 0 \qquad \Leftrightarrow \qquad x = y,$$

and for each fixed x (resp. y) φ is without critical points (y,ξ) (resp. (x,ξ)). In other words, for all $(x,y,\xi) \in U \times U \times (\mathbb{R}^n \setminus \{0\})$,

$$(d_\xi\varphi)_{(x,y,\xi)} = 0 \ \Leftrightarrow\ x = y, \quad \left(d_{(y,\xi)}\varphi\right)_{(x,y,\xi)} \neq 0 \quad \textit{and} \quad \left(d_{(x,\xi)}\varphi\right)_{(x,y,\xi)} \neq 0.$$

Moreover, we assume that the amplitude $q \in C^\infty(U \times U \times \mathbb{R}^n)$ meets the following conditions (analogous to the conditions (i)–(iv) on the amplitude of a pseudo-differential operator, p. 211):

(i′) *For each compact subset $K \subset U$ and multi-indices $\alpha, \beta, \gamma \in \mathbb{Z}_+^n$, there is a $C_{\alpha,\beta,\gamma} \in \mathbb{R}$ such that for all $x, y \in K$ and $\xi \in \mathbb{R}^n$, we have*
$$\left| D_\xi^\alpha D_x^\beta D_y^\gamma q\,(x, y, \xi) \right| \leq C_{\alpha,\beta,\gamma}(1 + |\xi|)^{k-|\alpha|}.$$
(ii′) $\sigma_k(q)(x, y, \xi) := \lim_{\lambda\to\infty} \frac{q(x,y,\lambda\xi)}{\lambda^k}$ *exists for $\xi \neq 0$ and $(x, y) \in U \times U$.*

(iii′) *For some cut-off function $\chi \in C^\infty(\mathbb{R}^n)$ with $\chi(\xi) = \begin{cases} 0, & \text{for } |\xi| \text{ small}, \\ 1, & |\xi| \geq 1, \end{cases}$*
$q\,(x, x, \xi) - \chi\,(\xi)\,\sigma_k(q)(x, x, \xi)$ is the amplitude of an element of $\mathrm{L}_{\mathrm{pc}}^{k-1}(U)$.

(iv′) *$q\,(x, y, \xi)$ has compact support in the x and y variables.*

Then Q can be written as a pseudo-differential operator (with compact support) of order k, i.e., $Q \in \mathrm{L}_{\mathrm{pc}}^k(U)$.

Note. Recall that we in this book deal mostly with the principal symbol of pseudo-differential operators and write shortly "symbol" and $\sigma(x, \xi)$ when we mean "principal symbol" and "$\sigma_k(x, \xi)$". In some places, however, we wish to mark the order of the operator in the notation for the symbol. That is the case in the preceding assumption (ii').

Remark 8.14. a) These operators are special types of *Fourier integral operators*. The term is due to L. Hörmander who in a series of papers, developed a precise theory for them, which can be applied to the general theory of partial differential equations. In doing so, he could resort to ideas of the Dutch mathematician and physicist Christian Huygens (1629-1695) and of the Russian mathematicians Vladimir Igorevich Arnold, Yuriy Vladimirovich Egorov, and Venyaminovich Clavdiy Maslov, who dealt with fundamentals of geometric optics and the formalization of its more or less intuitive methods (*aggregation principle*, *quantization*, etc., see also our Remark 8.4, p.208).
b) Below, in Step 1 of the proof of the preceding theorem, we shall address the delicate convergence questions related to the integral in (8.6). Such integrals are called **oscillatory integrals**. More precisely, let $U \subset \mathbb{R}^m$ and $\varphi = \varphi(x, \theta) \in C^\infty(U \times (\mathbb{R}^n \setminus \{0\}))$ real valued (or, at least of non-negative imaginary part) with $\varphi(x, \lambda\theta) = \lambda\varphi(x, \theta)$ for $\lambda > 0$ and $d\varphi \neq 0$ on all $U \times (\mathbb{R}^n \setminus \{0\})$. Let $p \in C^\infty(U \times \mathbb{R}^n)$ satisfy the asymptotic estimate (8.2) introduced on p.207 for fixed order $k \in \mathbb{R}$. Then the oscillatory integral
$$I(p, \varphi)(x) := \int_{\mathbb{R}^n \setminus \{0\}} e^{i\varphi(x,\theta)} p(x, \theta) d\theta$$
belongs to $C^j(U)$, if k is sufficiently negative, more precisely if $k + j < -n$. That is not very deep. The interesting aspect of oscillatory integrals is that they give also a meaning as distributions even if the order k of p is large. We shall explain that below in Step 1 for our special case where we replace U by $U \times U$, m by $2n$, $\varphi(x, \theta)$ by $\phi(x, y, \theta)$ and $p(x, \theta)$ by $q(x, y, \theta)$. A general and more systematic treatment of oscillatory integrals can be found in [**GS**, Chapter 1] and [**Gru**, p.168].
c) Obviously, every pseudo-differential operator with amplitude $p(x, \xi)$ can be written as a *Fourier integral operator* with $q(x, y, \xi) := p(x, \xi)$ and $\varphi(x, y, \xi) := \langle x - y, \xi \rangle$.

Before the proof of the preceding theorem, we shall emphasize that an amplitude is not determined from a given pseudo-differential operator.

EXAMPLE 8.15. Let $U \subset \mathbb{R}^n$, $\phi(x, y, \xi) := \langle x - y, \xi \rangle$, $a \in C_0^\infty(U) \setminus \{0\}$ and $1 \leq j \leq n$. Then the amplitude

$$q(x, y, \xi) := \xi_j a(y) - a(x)\xi_j - D_{x_j} a(x)$$

meets the conditions (i')-(iv') of Theorem 8.13 for $k = 1$, but $\mathrm{Op}(q)$ is just the zero operator.

REMARK 8.16. For pseudo-differential operators, the amplitude (also called the *total symbol* or the *dequantization*) $p(x, \xi)$ is neither uniquely determined from $\mathrm{Op}(p)$ in general, since a perturbation by an infinitely smoothing amplitude can generate the same operator. For the precise results we refer to [**Gru**, Proposition 7.8]. There is a vast literature on **properly supported** pseudo-differential operators, which have the nice (and somewhat misleading) property that they have uniquely determined amplitudes, see [**GS**, Chapter 3], [**Gru**, Section 7.2], [**Hö3**, Section 18.1], [**Shu**, Section 3.1], [**Tay81**, Section II.3]. The sad fact is the following: If you build on coordinates and coordinate shifts, it seems that only the symbol (i.e., what is also called the *principal* symbol) has a geometric meaning. That is well and easily defined for differential and pseudo-differential operators. It suppresses substantial parts of the underlying operator, but is sufficient for finding parametrices and calculating the index of elliptic operators, as we shall show below in Chapter 9. The good news is that there is a coordinate free definition of Fourier transformation and pseudo-differential operators leading to a one-to-one quantization $p \mapsto \mathrm{Op}(p)$, see below Section 8.5. For the full embedding proof of the Atiyah-Singer Index Theorem, mastering our coordinate free introduction of pseudo-differential operators and the one-to-one correspondence between operator and amplitude will be decisive. The non-geometric constructions of the analysis main-stream do not suffice. However, also our geometric construction below depends on choices (of metrics and connections, as mentioned already in Remark 8.4a, p. 208). So, it supports a powerful and transparent proof of the Index Theorem, but it does not offer a formal solution to the mysteries of dequantization.

PROOF OF THEOREM 8.13. **Step 0**: First we show the convergence of the integral defining $Qu(x)$. As the integral stands, it is only absolutely convergent when the order k of Q is very negative. However, the following integral is absolutely convergent for sufficiently large $r \in \mathbb{N}$:

$$\int_{\mathbb{R}^n} \int_U e^{i\varphi(x,y,\xi)} \left({}^tL\right)^r \left(q\left(x, y, \xi\right) u(y)\right) \,đy\, đ\xi,\ x \in U,$$

where ${}^tL$ denotes the formal adjoint of the operator

$$L := -i \frac{(d_y\varphi) \cdot d_y + (d_\xi\varphi) \cdot d_\xi}{|d_y\varphi|^2 + |\xi|^2 |d_\xi\varphi|^2} : C^\infty(U \times \mathbb{R}^n) \longrightarrow C^\infty(U \times \mathbb{R}^n),$$

which (under the assumptions) is a well-defined differential operator such that

$$Le^{i\varphi(x,\cdot,\cdot)} = e^{i\varphi(x,\cdot,\cdot)}.$$

In this way, the original integral can be replaced by an absolutely-convergent integral via repeated integration by parts. (Note that owing to the linearity of φ in ξ and the assumption $d_{(y,\xi)}\varphi \neq 0$, the term $|d_y\varphi|^2$ in the denominator of L grows like $|\xi|^2$ for fixed x and y, while $u(y)$ has compact support in y). Hence $Qu(x)$ is well-defined. Without difficulty it follows that Q is a linear map from $C_0^\infty(U)$ to $C^\infty(U)$.

Step 1: We now show that on a neighborhood Ω of the diagonal in $U \times U$, one can find a C^∞ map $\psi : \Omega \to \mathrm{GL}(n, \mathbb{R})$ (see Figure 8.1) such that for all $(x, y) \in \Omega$ and $\xi \in \mathbb{R}^n$, we have

$$\varphi(x, y, \psi(x, y)\xi) = \langle x - y, \xi \rangle .$$

By the assumption that φ is linear in ξ, we can write φ in the form

$$\varphi(x, y, \xi) = \sum\nolimits_{j=1}^{n} \varphi_j(x, y)\xi_j, \qquad \text{where } \varphi_j(x, y) := \tfrac{\partial \varphi}{\partial \xi_j}(x, y, \xi) .$$

We now show that the functional matrix

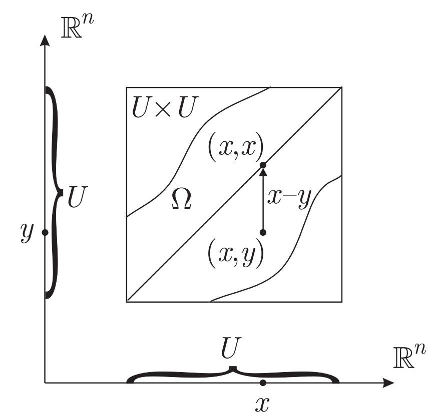


FIGURE 8.1. Finding ψ on a neighborhood Ω of the diagonal $U \times U$

$$F(x, y) := \begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1}(x, y) & \cdots & \frac{\partial \varphi_1}{\partial x_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(x, y) & \cdots & \frac{\partial \varphi_n}{\partial x_n}(x, y) \end{bmatrix}$$

is invertible for $x = y$. However, for all $y \in U$, $\varphi|_{U\times\{y\}\times(\mathbb{R}^n\setminus\{0\})}$ has no critical points, and so

$$\xi \neq 0 \quad \Rightarrow \quad \sum\nolimits_{j=1}^{n} \left| \tfrac{\partial \varphi}{\partial x_j}(x, y, \xi) \right| + \sum\nolimits_{j=1}^{n} \left| \tfrac{\partial \varphi}{\partial \xi_j}(x, y, \xi) \right| \neq 0.$$

By (8.7) $\sum_{j=1}^{n} \left| \frac{\partial \varphi}{\partial \xi_j}(x, y, \xi) \right| = 0 \Leftrightarrow x = y$. Thus, if $x = y$, then for *any* $\xi \neq 0$ there is some $j \in \{1, ..., n\}$ such that

$$0 \neq \tfrac{\partial \varphi}{\partial x_j}(x, y, \xi) = \tfrac{\partial}{\partial x_j} \left(\sum\nolimits_{k=1}^{n} \tfrac{\partial \varphi}{\partial \xi_k}(x, y, \xi)\, \xi_k \right) = \sum\nolimits_{j=1}^{n} \xi_k \tfrac{\partial \varphi_k}{\partial x_j}(x, y) .$$

Hence, the matrix $F(x, x)$ has a trivial kernel and must be invertible. By assumption $\varphi_j(x, x) = 0$. Thus we have the short Taylor expansion

$$\varphi_j(x, y) = \sum\nolimits_{\mu=1}^{n} \varphi_{\mu j}(x, y)(x_\mu - y_\mu),$$

where the functions $\varphi_{\mu j}(x,y)$ are C^∞ near the diagonal of $U \times V$. The matrix $\varphi(x,y) = [\varphi_{\mu j}(x,y)]$ is invertible in a neighborhood Ω of the diagonal, because $\varphi(x,x) = {}^tF(x,x)$. Since

$$\begin{aligned}\varphi(x,y,\xi) &= \sum\nolimits_{j=1}^n \varphi_j(x,y)\xi_j \\ &= \sum\nolimits_{\mu=1}^n (x_\mu - y_\mu) \sum\nolimits_{j=1}^n \varphi_{\mu j}(x,y)\xi_j = \langle x-y, \varphi(x,y)\xi\rangle\,,\end{aligned}$$

we have the desired property

$$\varphi(x,y,\psi\,(x,y)\,\xi) = \langle x-y,\xi\rangle\,,$$

where $\psi\,(x,y) := \varphi(x,y)^{-1}$ for $(x,y)\in\Omega$. For later, we note that

$$\varphi(x,x) = {}^tF(x,x) = \varphi''_{x\xi}(x,y,\xi)\big|_{y=x} := \left[\tfrac{\partial^2\varphi}{\partial x_i\partial\xi_j}(x,y,\xi)\Big|_{y=x}\right],$$

whence in particular,

$$\det\psi\,(x,x) = \frac{1}{\det\,\varphi''_{x\xi}(x,y,\xi)\Big|_{y=x}}.$$

Step 2: Now we *eliminate* the phase function φ. For this, we assume that for all ξ,

$$\operatorname{supp} q\,(\cdot,\cdot,\xi) \subseteq \Omega.$$

Then, for all $x\in U$, the integration domain in the formula for $(Qu)\,(x)$ (see above) is small enough so that the change of variable transformation $\xi = \psi(x,y)\theta$ can be applied to obtain

$$(Qu)\,(x) = \int_U\int_{\mathbb{R}^n} e^{i\langle x-y,\theta\rangle} q(x,y,\psi(x,y)\theta)\,|\det\psi(x,y)|\,u(y)\,đy\,đ\theta. \tag{8.8}$$

The new amplitude

$$(x,y,\theta)\mapsto a(x,y,\theta)\,|\det\psi(x,y)|$$

with $a(x,y,\theta) := q(x,y,\psi(x,y)\theta)$ then automatically satisfies the conditions (ii'), (iii'), and (iv'). To check (i'), we must calculate: By the chain rule, we obtain the formula (for $z := (x,y)\in\mathbb{R}^{2n}$)

$$(\partial_z a\,(z,\theta)\,,\partial_\theta a\,(z,\theta)) = (\partial_z q(z,\psi(z)\theta),\partial_\xi q(z,\psi(z)\theta))\begin{bmatrix}\mathrm{Id} & 0\\ \psi'(z)\theta & \psi(z)\end{bmatrix},$$

where ∂_z denotes the partial derivatives with respect to the first $2n$ variables and ∂_θ or ∂_ξ denote those with respect to the last n variables. Hence, we have

$$\begin{aligned}\left|\frac{\partial a}{\partial\theta_i}\,(z,\theta)\right| &= \left|\sum\nolimits_{j=1}^n \frac{\partial q}{\partial\xi_j}\,(z,\psi\,(z)\,\theta)\,\psi^{ij}\,(z)\right| \\ &\le nC_k\,(1+|\psi(z)\theta|)^{k-1}\max_{i,j,z}\left|\psi^{ij}\,(z)\right|,\end{aligned}$$

where C_k is a real number from assumption (i') for q when z varies within a compact domain $K\subseteq\Omega\subseteq U\times U\subseteq\mathbb{R}^{2n}$. Since we can find positive constants C_1 and C_2 with

$$C_1\,|\theta| \le |\psi\,(z)\,\theta| \le C_2\,|\theta|\,,\ \text{for all } z\in K \text{ and } \theta\in\mathbb{R}^n,$$

we finally have the estimate

$$\left|\frac{\partial a}{\partial\theta_i}\,(z,\theta)\right| \le \widetilde{C}_k\,(1+|\theta|)^{k-1}\,,\ \text{where } \widetilde{C}_k\in\mathbb{R}.$$

Similarly one can obtain estimates for the higher derivatives, wherein the factor $|\det \psi(x,y)|$ of the amplitude in (8.8) is irrelevant.

Step 3: Now, we consider the general case, where support $q(\cdot,\cdot,\xi)$ is not necessarily contained in Ω. In our applications of the theorem of Kuranishi (see Theorem 8.18, p. 220) we are only concerned with a local argument; i.e., we can manage with the case treated in step 2. We will therefore be brief in showing that in general we may assume the first case without loss of generality. We choose a non-negative C^∞ function χ on $U \times U$ having support in Ω and being equal to 1 in a neighborhood of the diagonal. Then Q can be written as the sum of two *Fourier integral operators*, where one has the amplitude χq of the form in step 2, and the other has the form

$$(Ru)(x) = \int_U \int_{\mathbb{R}^n} e^{i\theta(x,y,\xi)} r\,(x,y,\xi)\, u(y)\, đy\, đ\xi,$$

where $r = (1-\chi)q$ is a C^∞ function vanishing in a neighborhood of the diagonal in $U \times U$. Just as in Exercise 8.7, it follows that R can be written in the form

$$(Ru)(x) = \int_U K(x,y)\,(1-\Delta)^\lambda\, u(y)\, dy,$$

where the weight function K is C^∞ off the diagonal of $U \times U$ according to Exercise 8.7, and vanishes in a neighborhood of the diagonal by construction; K is then C^∞ everywhere. By Exercise 8.5b, R can then be written as a canonical pseudo-differential operator; the corresponding, conditions (*ii*), (*iii*), and (*iv*) are met without difficulty.

Step 4: Without loss of generality, we may now assume that the operator Q is given in the form

$$\begin{aligned}(Qu)(x) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle} q(x,y,\xi) u(y)\, đy\, đ\xi \\ &= \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \left\{ \int_{\mathbb{R}^n} e^{-i\langle y,\xi\rangle} q(x,y,\xi) u(y) đy \right\} đ\xi,\end{aligned}$$

where the braces enclose the Fourier transform of the product function $q(x,y,\xi)u(y)$ (extended by 0 values outside U) or equivalently, by Appendix A, the convolution of the Fourier transforms in the variable ξ

$$\int_{\mathbb{R}^n} e^{-i\langle y,\xi\rangle} q(x,y,\xi) u(y)\, đy = \int_{\mathbb{R}^n} \widehat{q}(x, \xi-\eta, \xi) \widehat{u}(\eta)\, đ\eta,$$

where $\widehat{q}(x,\cdot,\xi)$ is the Fourier transform of $y \mapsto q\,(x,y,\xi)$. Inserting the factor $e^{i\langle x,\eta\rangle} e^{-i\langle x,\eta\rangle}$ and reversing the order of integration, we obtain

$$\begin{aligned}(Qu)(x) &= \int_{\mathbb{R}^n} e^{i\langle x,\eta\rangle} \left(\int_{\mathbb{R}^n} e^{i\langle x,\xi-\eta\rangle} \widehat{q}(x,\xi-\eta,\xi) đ\xi \right) \widehat{u}(\eta)\, đ\eta \\ &= \int_{\mathbb{R}^n} e^{i\langle x,\eta\rangle} p\,(x,\eta)\, \widehat{u}(\eta)\, đ\eta,\end{aligned}$$

where (by a change of variables $\zeta = \xi - \eta$)

$$p(x,\eta) := \int_{\mathbb{R}^n} e^{i\langle x,\zeta\rangle} \widehat{q}(x,\zeta,\zeta+\eta)\, đ\zeta.$$

We now show that $p(x,\eta)$ is actually the amplitude of a pseudo-differential operator (the support is trivially compact by construction, whence (*iv*) already holds):

(i): Let α and β be multi-indices and let x range over a compact subset of $\mathbb{R}^n$. For the estimation of $\left|D_x^\beta D_\eta^\alpha p(x,\eta)\right|$, we first note that by the Fourier multiplication rule (Appendix A, Exercise A.5b, p. 675), we have

$$\left|D_x^\beta M_\theta^\gamma D_\eta^\alpha \widehat{q}(x,\theta,\eta)\right| = \left|\int_{\mathbb{R}^n} e^{-i\langle y,\zeta\rangle} D_x^\beta D_y^\gamma D_\eta^\alpha q(x,y,\eta)\, đy\right| \le c\,(1+|\eta|)^{k-\alpha},$$

where γ is a further multi-index and M_θ^γ is multiplication by $\theta^\gamma = \theta_1^{\gamma_1}\cdots\theta_n^{\gamma_n}$; the inequality follows from the assumptions (i') and (iv') for q. Hence, for each positive ν, we have

$$\left|D_x^\beta D_\eta^\alpha \widehat{q}(x,\theta,\eta)\right| \le c'\,(1+|\eta|)^{k-\alpha}\,(1+|\theta|)^{-\nu}. \tag{8.9}$$

Thus, by definition of p and by means of differentiation under the integral, we get

$$\left|D_x^\beta D_\eta^\alpha p\,(x,\eta)\right| \le c''\,(1+|\eta|)^{k-\alpha}$$

(ii): By the Mean-Value Theorem, we obtain for suitable ζ_0 between 0 and ζ

$$\begin{aligned} p(x,\eta) &= \int_{\mathbb{R}^n}\left(e^{i\langle x,\zeta\rangle}\widehat{q}(x,\zeta,\eta) + \sum\nolimits_{|\alpha|=1} e^{i\langle x,\zeta\rangle} D_\eta^\alpha \widehat{q}(x,\zeta,\eta+\zeta_0)\zeta^\alpha\right) đ\zeta \\ &= q(x,x,\eta) + \text{a correction term } E\,(x,\eta)\,. \end{aligned}$$

We have already seen (see (8.9)) that

$$\left|D_\eta^\alpha \widehat{q}(x,\zeta,\eta+\zeta_0)\right| \le C_\nu(1+|\eta+\zeta_0|)^{k-1}\,(1+|\zeta|)^{-\nu},$$

for arbitrarily large ν. Since $|\zeta_0| < |\zeta|$, we have

$$\left|D_\eta^\alpha \widehat{q}(x,\zeta,\eta+\zeta_0)\right| \le c'\,(1+|\eta|)^{k-1}\,(1+|\zeta|)^{-\nu+k-1}.$$

Integrating with respect to ξ, we find $|E\,(x,\eta)| \le C\,(1+|\eta|)^{k-1}$. Thus,

$$\lim_{\lambda\to\infty}\frac{E\,(x,\lambda\eta)}{\lambda^k} \le \lim_{\lambda\to\infty}\left|\frac{C\,(1+|\lambda\eta|)^{k-1}}{\lambda^k}\right| = 0, \text{ and so}$$

$$\sigma_k(p)(x,\eta) = \lim_{\lambda\to\infty}\frac{p\,(x,\lambda\eta)}{\lambda^k} = \lim_{\lambda\to\infty}\frac{q(x,x,\lambda\eta)}{\lambda^k} = \sigma_k(q)(x,x,\eta).$$

(iii): It follows easily from the estimate $E\,(x,\eta) \le C\,(1+|\eta|)^{k-1}$ and the corresponding assumption (iii') for $q(x,x,\eta)$, that $p(x,\eta)-\chi\,(\eta)\,\sigma_k(p)(x,\eta)$ is the amplitude of a canonical pseudo-differential operator of order $k-1$. □

Remark 8.17. We note that, from the preceding constructive proof of Masatake Kuranishi, a simple formula for the symbol follows from steps 2 and 4:

$$\sigma_k(p)(x,\eta) = \frac{\sigma_k(q)(x,x,\psi\,(x,x)\,\eta)}{\left|\det \varphi''_{x,\xi}\,(x,y,\xi)\big|_{y=x}\right|},$$

where $\psi\,(x,x)$ is the inverse of the functional matrix (also denoted $F(x,x)$) in step 1, namely

$$\varphi''_{x,\xi}\,(x,y,\xi)\big|_{y=x} = \left[\tfrac{\partial^2\varphi}{\partial x_i\partial\xi_j}(x,y,\xi)\Big|_{y=x}\right].$$

The term Ru in step 3 with amplitude r does not affect the symbol formula, since $r(x,x,\eta)=0$ for all x and η, whence $\sigma_k(r)(x,x,\eta)=0$.

Coordinate Change and Pseudo-Differential Operators on Manifold. Now we investigate the behavior of pseudo-differential operators under a coordinate change, exploiting the *Kuranishi Trick* of transgressing to Fourier integral operators, similarly in [**GS**, p.34f]:

THEOREM 8.18. *Let $\kappa : U \to V$ be a diffeomorphism between relatively compact open subsets of $\mathbb{R}^n$. If P is a pseudo-differential operator (with compact support) of order $k \in \mathbb{Z}$ on V, then the* transported *operator*

$$P_\kappa(u) := P(u \circ \kappa^{-1}) \circ \kappa,\ u \in C_0^\infty(U)$$

is a pseudo-differential operator (with compact support) of order k over U. If p and q are amplitudes for P and P_κ resp., then their symbols are related by

$$\sigma_k(q)(x,\xi) = \sigma_k(p)(\kappa(x), ({}^t\kappa'(x))^{-1}\xi),\ x \in U,\ \xi \in \mathbb{R}^n \setminus \{0\},$$

where

$${}^t\kappa'(x) = \begin{bmatrix} \frac{\partial \kappa_1}{\partial x_1} & \cdots & \frac{\partial \kappa_n}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial \kappa_1}{\partial x_n} & \cdots & \frac{\partial \kappa_n}{\partial x_n} \end{bmatrix}$$

is the transpose of the functional matrix of κ at x.

PROOF. Suppose that the operator P is of the form

$$\begin{aligned}(Pv)(y) &= \int_{\mathbb{R}^n} e^{i\langle y,\eta\rangle} p(y,\eta)\,\widehat{v}(\eta)\,đ\eta \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\langle y-\theta,\eta\rangle} p(y,\eta)\,v(\theta)\,đ\theta\,đ\eta,\ v \in C_0^\infty(V),\ y \in V,\end{aligned}$$

whence for $v = u \circ \kappa^{-1}$ and $y = \kappa(x)$, $u \in C_0^\infty(U)$ and $x \in U$:

$$\begin{aligned}P_\kappa(u)(x) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\langle \kappa(x)-\theta,\eta\rangle} p(\kappa(x),\eta)\,u\left(\kappa^{-1}(\theta)\right)\,đ\theta đ\eta \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\langle \kappa(x)-\kappa(\xi),\eta\rangle} p(\kappa(x),\eta)\,|\det \kappa'(\xi)|\,u(\xi)\,đ\xi đ\eta\end{aligned}$$

by means of the change of variable $\kappa(\xi) = \theta$. The *transported* operator is then a *Fourier integral operator* with the phase function $\varphi(x,\xi,\eta) := \langle \kappa(x) - \kappa(\xi), \eta\rangle$ and amplitude $q(x,\xi,\eta) := p(\kappa(x),\eta)\,|\det \kappa'(\xi)|$. Since φ and q meet the hypotheses of the Theorem of Kuranishi (Theorem 8.13, p. 213), we have

$$P_\kappa(u)(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\langle x-\xi,\eta\rangle} p(\kappa(x), \psi(x,\xi)\,\eta)\,D(x,\xi)\,u(\xi)\,đ\xi đ\eta,$$

where

$$D(x,\xi) := |\det \kappa'(\xi)|\,|\det \psi(x,\xi)|$$

and $\psi(x,\xi)$ is the matrix-valued function constructed in step 1 of the proof of Theorem 8.13; in particular,

$$\psi(x,x)^{-1} = \varphi''_{x,\eta}(x,\xi,\eta)\big|_{\xi=x} = \left[\frac{\partial \kappa_j}{\partial x_\nu}\right] = {}^t\kappa'(x)$$

and $D(x,x) = 1$.

By the Theorem of Kuranishi, P_κ is a pseudo-differential operator for which we derived an explicit formula for the amplitude in the above proof. For the symbol, we have

$$\sigma_k(q)(x,\eta) = \sigma_k(\widetilde{q})(x,x,\eta),$$

where

$$\widetilde{q}(x,\xi,\eta) = p\left(\kappa(x), \psi\left(x,\xi\right)\eta\right) D(x,\xi).$$

Since $\psi(x,x)^{-1} = {}^t\kappa'(x)$ and $D(x,x)=1$, we then obtain

$$\sigma_k(q)(x,\eta) = \sigma_k(p)(\kappa\left(x\right), \left({}^t\kappa'\left(x\right)\right)^{-1}\eta).$$

□

Exercise 8.19. Let X be a (paracompact) C^∞ n-manifold and $k \in \mathbb{Z}$.
a) Show that the space $\mathrm{L}^k_{\mathrm{pc}}(X)$, defined by *localization* at the beginning of this section just before Definition 8.12, coincides with the space of pseudo-differential operators of order k on X, when X is a bounded open subset of $\mathbb{R}^n$.
b) Define a canonical vector space structure on $\mathrm{L}^k_{\mathrm{pc}}(X)$.
c) Show $\mathrm{L}^k_{\mathrm{pc}}(X) \subset \mathrm{L}^{k+1}_{\mathrm{pc}}(X)$.
d) Let $\kappa : U \to \mathbb{R}^n$, U open in X, be a local coordinate system for X, $f \in C_0^\infty(U)$, and $P \in \mathrm{L}^k_{\mathrm{pc}}(X)$. Show that if $f \in C^\infty(X)$ is identically 1 in a neighborhood of $\kappa^{-1}(x)$, then the formula

$$(x,\xi) \mapsto q_f(x,\xi) := e^{-i\langle x,\xi\rangle}\left(P(f\left(\cdot\right)e^{i\langle\kappa(\cdot),\xi\rangle})\right)(\kappa^{-1}(x));\ x \in \kappa(U),\ \xi \in \mathbb{R}^n,$$

defines the amplitude of a pseudo-differential operator of order k on $\kappa(U)$, and that if p is the amplitude for the *localized* operator P_κ, then

$$\sigma_k\left(q_f\right)(x,\xi) = \sigma_k(p)(x,\xi).$$

e) Set $\mathrm{Smbl}_k(X) := \mathrm{Smbl}_k(\mathbb{C}_X, \mathbb{C}_X)$; see our definition in Equation (6.26), p. 181, in Chapter 6 above. Thus, $s \in \mathrm{Smbl}_k(X) \Leftrightarrow s : \mathring{T}^*X \to \mathbb{C}$ with $s(x,\lambda v) = \lambda^k s(x,v)$ for all $x \in X$ and $v \in T^*_xX$, $v \neq 0$. Show that the linear map

$$\sigma_k : \mathrm{L}^k_{\mathrm{pc}}(X) \to \mathrm{Smbl}_k(X)$$

is well defined and coincides with the earlier definition (see Exercise 6.34, p.180) on $\mathrm{Diff}_k(X) \subset \mathrm{L}^k_{\mathrm{pc}}(X)$.
[Hint: For a: Theorem 8.18.
For b: Proceed by using the vector space structure of $C^\infty(X)$.
For c: Use amplitude estimates.
For d: One can characterize $\mathrm{L}^k_{\mathrm{pc}}(X)$ within the space of linear operators from $C_0^\infty(x)$, to $C^\infty(X)$ as those such that, for all local coordinate systems *cut-off functions* f, the associated q_f is the amplitude of a pseudo-differential operator of order k on an open subset of $\mathbb{R}^n$. For details of the computation, see [**Hö71b**, p.112] and [**Ni**].
For e: Recall that $\mathring{T}^*X$ denotes the *symplectic cone* $T^*X \setminus X$ that consists of the punctured cotangent spaces. It remains only to show that the locally well defined symbol in d) transforms *correctly* under a coordinate change, so that it forms global homomorphism from $\mathring{T}^*X \times \mathbb{C}$ to $\mathring{T}^*X \times \mathbb{C}$ (which is homogeneous of degree k in the cotangent vectors. For this, check that the transformation rule in Theorem 8.18 can be written in the form

$$\sigma(q)(x,\widetilde{\kappa}\left(\eta\right)) = \sigma_k(p)(\kappa\left(x\right),\eta),$$

where η lies in $T^*(\mathbb{R}^n)_{\kappa(x)}$, which is the space of covectors at the point $\kappa(x)$ canonically identified with $\mathbb{R}^n$, and $\widetilde{\kappa}(\eta)$ is the *pull back* covector via κ; see (6.1), p.163, in Appendix B; further details are found in [**AB67**, p.404-407].]

EXERCISE 8.20. Define the space $\mathrm{L}^k_{\mathrm{pc}}(E,F)$ when E and F are complex vector bundles over the C^∞ manifold X, and show the existence of a canonical linear map $\sigma_k : \mathrm{L}^k_{\mathrm{pc}}(E,F) \to \mathrm{Smbl}_k(E,F)$.
[Hint: Represent an operator $P : C_0^\infty(E) \to C^\infty(F)$ locally; i.e., choose a chart $\kappa : U \to \mathbb{R}^n$, $U \subseteq X$ open, and $\kappa(U)$ relatively compact, and trivializations $E|_U \cong U \times \mathbb{C}^N$ and $F|_U \cong U \times \mathbb{C}^M$ as a $M \times N$ matrix of pseudo-differential operators (with compact support) of order k.]

REMARK 8.21. a) There is a slight ambiguity in our definition of the symbol space $\mathrm{Smbl}_k(E,F)$. As explained in our defining Equation 6.26, p.181, we require homogeneity

$$\sigma(x, r\xi) = r^k \sigma(x,\xi) \text{ for } x \in X, \xi \in T^*_x X \setminus \{0\} \text{ and } r > 0 \tag{8.10}$$

for $\sigma \in \mathrm{Smbl}_k(E,F)$. Our Assumption 8.10(ii),p.211 ensures (8.10). Clearly, homogeneity and smoothness at $\xi = 0$ contradict each other except for monomials. *Our* convention is that $\mathrm{Smbl}_k(E,F)$ denotes the space of homogeneous bundle homomorphisms of the lifted bundles π^*E, π^*F, where $\pi : \mathring{T}^*X \to X$ and $\mathring{T}^*X = T^*X \setminus X$, i.e., we exclude $\xi = 0$. In various applications, however, symbols should be smooth functions, thus the $\sigma(x,\xi)$ should be smooth everywhere but homogeneous only in the restricted sense:

$$\sigma(x, r\xi) = r^k \sigma(x,\xi) \text{ for } x \in X, |\xi| \geq 1 \text{ and } r \geq 1 \tag{8.11}$$

with a suitable Riemannian metric that yields the length of cotangent vectors.
b) In many places, we shall tacitly identify the homogeneous bundle mappings on $T^*X \setminus X$ by restriction with the smooth sections $C^\infty\big(S^*X, \mathrm{Hom}(\rho^*(E), \rho^*(F))\big)$. Here S^*X denotes the sphere bundle of cotangent vectors (relative to a fixed Riemannian metric), $\rho : S^*X \to X$ the natural projection, and $\mathrm{Hom}\big(\rho^*(E), \rho^*(F)\big)$ the bundle of smooth bundle homomorphisms.

REMARK 8.22. The definition of $\mathrm{L}^k_{\mathrm{pc}}(E,F)$ through *localizations* is unsatisfactory from a computational point of view. The choice of coordinates is awkward with its avalanche of subscripts which are frequently unavoidable even in fairly simple situations. The method is particularly unsatisfactory, when the operator (as in Exercise 8.23 below) can be written in closed, global form explicitly and much more clearly. [**Hö71b**, p.113f] contains the following idea for writing $\mathrm{L}^k_{\mathrm{pc}}(E,F)$ by means of the Kuranishi Theorem directly as a space of *Fourier Integral Operators* with phase function $\varphi : G \to \mathbb{R}$ and amplitude $q : G \to \mathrm{Hom}(E,F)$: Let G be a real vector bundle of fiber dimension n over a neighborhood of the diagonal in $X \times X$, e.g., $G = \pi^*(T^*X)$ where π is the projection $\pi(x,y) := y$, and

$$q(x,y,\xi) \in \mathrm{Hom}(E_y, F_x).$$

It turns out that one can formulate the necessary conditions on φ and q directly, globally and with little difficulty: For example, φ is linear in the fibers and the restriction of φ to a fiber has a critical point exactly when the fiber lies above a point of the diagonal of $X \times X$. Then (loc. cit.) $\mathrm{L}^k_{\mathrm{pc}}(E,F)$ consists of all operators that can be written as the sum of an operator with C^∞ kernel and one of the form

$$(Pe)(x) := (2\pi)^{-n} \int_{T^*X} e^{i\varphi(x,y,\eta)} q(x,y,\eta) e(y)\, dy d\eta,$$

where $e \in C_0^\infty(E)$, $dyd\eta$ is the invariant volume element on the cotangent bundle T^*X, and q an amplitude *of order* k which vanishes for (x,y) outside a small neighborhood of the diagonal of $X \times X$. By step 4 of the proof of Theorem 8.13, it follows that

$$\sigma_k\left(P\right)(x,\eta) = \lim_{\lambda\to\infty} \frac{q(x,x,\lambda\eta)}{\lambda^k},\ x \in X,\ \eta \in T_x^*X \setminus \{0\}.$$

We shall devote the whole Section 8.5, p.228ff to the details of a truly global construction of a normalized total symbol.

Singular Integral Operators. We show that the classical singular integral operators fit nicely under our heading of principally classical pseudo-differential operators.

EXERCISE 8.23. Show that the following *singular integral operators* are pseudo-differential operators of order 0 over $\mathbb{R}$ or $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ [1] and determine their symbols:
a) The Hilbert transform $Q : C_0^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, defined for $u \in C_0^\infty(\mathbb{R})$ by

$$(Qu)(x) := \frac{-1}{\pi i}\,(p.v.) \int_{-\infty}^{\infty} \frac{u\left(y\right)}{x-y}\, dy := \frac{-1}{\pi i} \lim_{\varepsilon\to 0^+} \int_{|x-y|>\varepsilon} \frac{u\left(y\right)}{x-y}\, dy.$$

b) The projection operator $P : C^\infty(S^1) \to C^\infty(S^1)$, defined by

$$Pe^{im\theta} := \begin{cases} e^{im\theta}, & \text{for } m \ge 0, \\ 0, & \text{for } m < 0. \end{cases}$$

c) The *Toeplitz operator*

$$gP + (\mathrm{Id} - P),\ \text{for } g \in C^\infty(S^1).$$

[Hint: For a: Show that

$$(Qu)(x) = \int_{-\infty}^{\infty} e^{ix\xi} \operatorname{sign}(\xi)\, \widehat{u}\left(\xi\right)\, d\!\!\!^{-}\xi,$$

as follows. We have

$$\begin{aligned}\int_{|x-y|>\varepsilon} \frac{u\left(y\right)}{x-y}\, dy &= \int_{|y|>\varepsilon} \frac{u\left(x-y\right)}{y}\, dy = (u * g_\varepsilon)(x) \\ &= \int_{-\infty}^{\infty} e^{ix\xi} \widehat{u}\left(\xi\right) \widehat{g}_\varepsilon\left(\xi\right)\, d\xi,\ \text{where}\end{aligned}$$

$$g_\varepsilon\left(x\right) := \begin{cases} \frac{1}{x}, & \text{for } |x| > \varepsilon, \\ 0, & \text{for } |x| \le \varepsilon, \end{cases}$$

[1] More precisely: Write them as a sum of a pseudo-differential operator of the kind treated so far and a *smoothing* operator. Many *classical* pseudo-differential operators Q are defined as here via an amplitude q which is homogeneous in the second variable, but has a singularity at the origin. Through multiplication by a C^∞ function χ which is identically 1 in a neighborhood of ∞, we obtain a *singularity-free* amplitude $\widetilde{q}(x,\xi) := \chi(\xi)q(x,\xi)$, which defines a pseudo-differential for $\widetilde{Q}$ in our (Hörmander's) sense. Then $\widetilde{Q} - Q$ has an amplitude with compact support and consequently (with reasoning as in Remark 8.8, p.210) can be represented as an integral operator with a C^∞ weight function.

and for the last equality the well-known convolution formula is used. Thus, it is natural to try to evaluate the improper integral

$$(p.v.)\int_{-\infty}^{\infty}\frac{e^{-i\xi t}}{t}\,dt := \sqrt{2\pi}\lim_{\varepsilon\to 0}\ \widehat{g}_\varepsilon(\xi).$$

Now distinguish cases according to the sign of ξ! One obtains

$$\sqrt{2\pi}\lim_{\varepsilon\to 0}\widehat{g}_\varepsilon(\xi) = -2i\ \operatorname{sign}(\xi)\lim_{\varepsilon\to 0}\int_{|\xi|\varepsilon}^{\infty}\frac{\sin t}{t}\,dt = -\pi i\ \operatorname{sign}(\xi)\,,$$

since $\int_0^\infty \frac{\sin t}{t}\,dt = \frac{\pi}{2}$ via contour integration of the function $f(z) := \frac{e^{iz}}{z}$, $z = t + is$, along the curve shown in Figure 8.2. Compare also [**DM**, p.93 and 150].

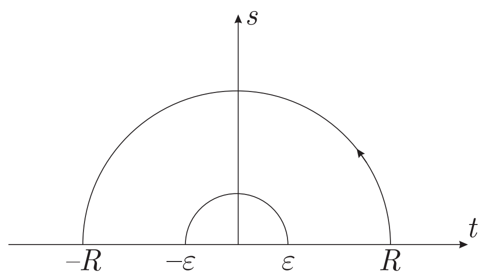


FIGURE 8.2. Contour for integrating the function f

For b: Reduce to a) by means of the formula $Pu = \frac{1}{2}(u + Hu)$, where

$$(Hu)(e^{i\theta}) := \frac{1}{\pi i}(p.v.)\int_{S^1}\frac{u(z)}{z - e^{i\theta}}\,dz,\ u\in C^\infty(S^1),$$

is the Cauchy-Hilbert transformation (on the circle) which is carried over to the Hilbert transformation Q (on the line $\mathbb{R}$) by means of the Cayley transformation; see p. 130 in Chapter 4. Details for this are in [**Tay74**, p.4-5, 36]. A more direct way may be found in [**AS68a**, p.525]. For this, as in Exercise 8.19d form the expression

$$\begin{aligned} q_f(x,\xi) &:= e^{-ix\xi}P(f(x)e^{ix\xi}) = \sum\nolimits_{n=0}^{\infty}\sqrt{2\pi}\widehat{f}(n-\xi)e^{ix(n-\xi)} \\ &= f(x) - \sqrt{2\pi}\sum\nolimits_{n=-1}^{-\infty}\widehat{f}(n-\xi)e^{ix(n-\xi)}, \end{aligned}$$

where $f\in C_0^\infty(\mathbb{R})$ with compact support in an interval of length $< 2\pi$, so that f may be regarded as a function on the circle with support in a *canonical* coordinate domain.

Trick: For $\xi < 0$, estimate $\sum_{n=0}^{\infty}\widehat{f}(n-\xi)e^{ix(n-\xi)}$ and its derivatives, showing that as $\xi\to-\infty$ they go to 0 faster than any power of $|\xi|^{-1}$. Show that for $\xi > 0$, the sum $\sum_{n=-1}^{-\infty}\widehat{f}(n-\xi)e^{ix(n-\xi)}$ has the corresponding property. By the hint for Exercise 8.19d, one is done and obtains

$$\sigma_0(P)(x,\xi) = \begin{cases} 1, & \text{for } \xi > 0,\\ 0, & \text{for } \xi < 0.\end{cases}$$

For c: Reduce to b). Note that in the notation of Chapter 4

$$gP + (\operatorname{Id} - P) = \begin{cases} T_g, & \text{on } C^\infty\left(S^1\right)\cap H_0,\\ \operatorname{Id}, & \text{on } C^\infty\left(S^1\right)\cap H_0^\perp \text{ (in } L^2(S^1)),\end{cases}$$

where T_g is the Wiener-Hopf operator induced by g with index $T_g = -W(g,0)$, if g is nowhere zero on S^1. In particular, by Exercise 1.5 (p. 5), $\operatorname{index}(gP + \mathrm{Id} - P) = -W(g,0)$.]

4. Algebraic Properties and Symbolic Calculus

Here we will show that all C^∞ symbols are obtained as symbols of pseudo-differential operators. Moreover, we explain how one can calculate with symbols instead of operators, using some simple rules. In a wording borrowed from algebraic topology, th symbolic calculus is a *functor* from the category of infinite-dimensional function spaces and systems of linear differential and pseudo-differential equations to parameter dependent linear algebra in finite dimensions.

THEOREM 8.24. *Let E and F be complex vector bundles over a C^∞ manifold X. Then there is an exact sequence*

$$0 \longrightarrow \mathrm{CL}^{k-1}(E,F) \hookrightarrow \mathrm{L}^k_{\mathrm{pc}}(E,F) \xrightarrow{\sigma_k} \mathrm{Smbl}_k(E,F) \longrightarrow 0\,, \tag{8.12}$$

where $\sigma_k(A)$ denotes the principal *(homogeneous leading) symbol of $A \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$.*

Recall that $\mathrm{L}^{k-1}(E,F)$ (respectively $\mathrm{CL}^k(E,F)$) denote the space of $(k-1)$th order canonical pseudo-differential operators from sections of E to sections of F (respectively kth order classical pseudo-differential operators) and that

$$\mathrm{L}^k_{\mathrm{pc}}(E,F) = \mathrm{CL}^k(E,F) + \mathrm{L}^{k-1}(E,F), \tag{8.13}$$

like in Remark 8.11b, p.211.

PROOF. Since $\mathrm{CL}^k(E,F) \cap \mathrm{L}^{k-1}(E,F) = \mathrm{CL}^{k-1}(E,F)$, the principal symbol map $\sigma_k : \mathrm{L}^k_{\mathrm{pc}}(E,F) \to \mathrm{Smbl}_k(E,F)$ is well defined. Because of the decomposition (8.13), it only remains to show that the symbol map is surjective.
Let $s \in \mathrm{Smbl}_k(E,F)$. If $\pi : X \times X \to X$ is the projection given by $\pi(x,y) = y$, then the pull-back bundle

$$G := \pi^*(T^*X) \to X \times X,$$

is a real vector bundle of fiber dimension n. In reference to Remark 8.22 after Exercise 8.20 it suffices to give a *phase function* $\varphi : G \to \mathbb{R}$ and an *amplitude* $a : G \to \mathrm{Hom}(E,F)$ with the properties required by definition (see the conditions (i')–(iv') in Theorem 8.13, p. 213) such that

$$a(x,y,\eta) = s(x,\eta),\ x \in X,\ \eta \in T^*_x X \setminus \{0\}\,.$$

For φ, we choose a real-valued function with $\varphi(x,x,\eta) = 0$ and

$$d\varphi_{(x,x,\eta)} = \eta \oplus -\eta \oplus 0 : T^*_{(x,x,\eta)}G \to \mathbb{R},\ \text{where}$$
$$T^*_{(x,y,\eta)}G \cong T^*_x X \oplus T^*_y X \oplus T^*_\eta\,(T^*X)\ \text{and}\ \eta \in T^*_x X.$$

Such a map φ, which is also linear on the fiber and only possesses critical points over the diagonal of $X \times X$, is locally easy to construct relative to a chart κ about the point x. One simply sets $\varphi(x,y,\eta) = \left\langle \kappa x - \kappa y, \left(\kappa^{-1}\right)^{\sim}\eta \right\rangle$ (see (6.1), p. 163 for the definition of "~"). A global construction may be carried out using a partition of unity (see Theorem 6.4, p. 157) in a neighborhood of the diagonal. For $a : G \to \mathrm{Hom}(E,F)$, we choose an arbitrary extension of $a(x,x,\eta) := s(x,\eta)$ to a neighborhood of the diagonal; then we can smoothly extend a by multiplying by a C^∞ function (with support in the neighborhood) which is identically 1 in a smaller neighborhood of the diagonal. An extension can be found, since the diagonal is closed in $X \times X$, and a can be regarded as a section of the lift of the bundle $\mathrm{Hom}(E,F)$ by means of the projection $(x,y,\eta) \mapsto (x,y)$. Compare with step 1 of the proof of Theorem B.5 (p. 680) in Appendix B, in connection with the Whitney Approximation Theorem; e.g., [**Na**, p.34f] or [**BJ**, 14.8]. With this, the proof (which strongly depends on the Theorem of Kuranishi, Theorem 8.13, p.213) is done. A direct proof can be found in [**Wells**, p.134f]. □

REMARK 8.25. We can fix a right inverse

$$\mathrm{Op} : \mathrm{Smbl}_k(E,F) \longrightarrow \mathrm{L}^k_{\mathrm{pc}}(E,F)$$

of σ_k, obtained by patching together the local Op-maps (8.1) via a fixed partition of unity.

THEOREM 8.26. *The direct sum* $\mathrm{L}^\bullet_{\mathrm{pc}}(E,F) := \bigoplus_k \mathrm{L}^k_{\mathrm{pc}}(E,F)$ *forms a graded algebra via composition, and is closed under the operation of taking formal adjoints. For the symbols, we have the following calculation rules:*
(*a*) *If* $E, F,$ *and* G *are complex vector bundles over the* C^∞ *manifold* X*, and* $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$ *and* $Q \in \mathrm{L}^j_{\mathrm{pc}}(F,G)$*, then* $Q \circ P \in \mathrm{L}^{j+k}_{\mathrm{pc}}(E,G)$ *and*

$$\sigma_{j+k}\left(Q \circ P\right)(x,\eta) = \sigma_j\left(Q\right)(x,\eta) \circ \sigma_k\left(P\right)(x,\eta),\ x \in X,\ \eta \in T^*_x X \setminus \{0\}\,.$$

(*b*) *Let* $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$*, where the bundles* E *and* F *are equipped with Hermitian metrics and the manifold* X *is Riemannian and oriented. Then, there is a unique operator* $P^* \in \mathrm{L}^k_{\mathrm{pc}}(F,E)$ *with*

$$\int_X \langle Pu, v\rangle_F = \int_X \langle u, P^*v\rangle_E\,,\ \textit{for all } u \in C^\infty_0(E), v \in C^\infty_0(F),\ \textit{and}$$
$$\sigma_k\left(P^*\right)(x,\eta) = \sigma_k\left(P\right)(x,\eta)^*,\ \textit{for } x \in X,\ \eta \in T^*_x X \setminus \{0\}\,.$$

PROOF. We begin with (b): The uniqueness of P^* is clear. For the proof of existence, we need only to show that for each $v \in C^\infty_0(F)$ and each open *coordinate domain* $U \subseteq X$ (with $E|_U$ and $F|_U$ trivial) there is $P^*_U v \in C^\infty(E|_U)$ such that

$$\int_X \langle Pu, v\rangle_F = \int_X \langle u, P^*_U v\rangle_E \text{ for all } u \in C^\infty_0(E|_U).$$

Indeed, when such a $P^*_U v$ exists, then it is uniquely determined, and so for a second *coordinate domain* U'

$$P^*_U v|_{U\cap U'} = P^*_{U'} v|_{U\cap U'}.$$

Then we have a global C^∞ section $P^*v \in C^\infty(E)$ with

$$\int_X \langle Pu, v\rangle_F = \int_X \langle u, P^*v\rangle_E \text{ for all } u \in C^\infty_0(E),$$

which is constructed by covering X with finitely many coordinate domains U and writing $u = \sum_j u_j$ with $u_j \in C^\infty_0\left(E|_{U_j}\right)$ by means of a C^∞ partition of unity.

Hence, let v and U be given. Without loss of generality, we, assume that there is a coordinate domain V which contains $\overline{U}$ as well as $\mathrm{supp}(v)$. (Otherwise one covers $\mathrm{supp}(v)$ with finitely many coordinate domains and pieces v together from a

C^∞ partition of unity.) In local coordinates, relative to a C^∞ N-framing for $E|_V$ and M-framing for $F|_V$, one can write P in the form

$$(Pu)(x) = \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} \langle p(x,\xi)\widehat{u}(\xi), v(x)\rangle \, đ\xi, \ u \in C_0^\infty(E|_V)$$

where p is an $M \times N$ matrix of amplitude functions with compact support. Hence,

$$\begin{aligned}\langle Pu, v\rangle &= \int\int e^{i\langle x,\xi\rangle} \langle p(x,\xi)\widehat{u}(\xi), v(x)\rangle_{\mathbb{C}^M} \, đ\xi\, dx \\ &= \int\int\int e^{i\langle x-y,\xi\rangle} \langle p(x,\xi)u(y), v(x)\rangle_{\mathbb{C}^M} \, đy\, đ\xi\, dx \\ &= \int \left\langle u(y), \int\int e^{i\langle x-y,\xi\rangle} p^*(x,\xi)v(x)\, d'x\, đ\xi \right\rangle_{\mathbb{C}^N} dy,\end{aligned}$$

where $p^*(x,\xi)$ is the adjoint of $p(x,\xi)$ (i.e., complex-conjugate transpose), an $N\times M$ matrix of complex numbers for each $x \in V$, $\xi \in \mathbb{R}^n$, and the integrals here and below are over $\mathbb{R}^n$. Thus, we get $\langle Pu, v\rangle = \langle u, P^*v\rangle$ for

$$(P^*v)(y) := \int e^{i\langle x-y,\xi\rangle} p^*(x,\xi)v(x)\, đxđ\xi = \int e^{i\varphi(x,y,\xi)} q(y,x,\xi)v(x)\, d'xđ\xi,$$

where $q(y,x,\xi) := p^*(x,\xi)$ and $\varphi(x,y,\xi) = \langle x-y,\xi\rangle$. By the Theorem of Kuranishi, there now exists an amplitude $\widetilde{p}$ so that

$$(P^*v)(y) = \int e^{i\langle y,\xi\rangle} \widetilde{p}(y,\xi)\widehat{v}(\xi)\, đ\xi.$$

Hence, we have found $P^*v \in P\operatorname{Diff}_k(\mathbb{C}^N, \mathbb{C}^M)$ with the desired property. Also, we have

$$\sigma(P^*)(x,\xi) = \lim_{\lambda\to\infty} \lambda^{-k} q(x,x,\lambda\xi) = \lim_{\lambda\to\infty} \lambda^{-k} p^*(x,\lambda\xi) = (\sigma(P)(x,\xi))^*.$$

Furthermore, we remark that with this we obtain (for all $u \in \mathbb{C}^\infty(\mathbb{C}^N_V)$ and $v \in \mathbb{C}^\infty(\mathbb{C}^M_V)$)

$$\begin{aligned}\langle Pu, v\rangle &= \int \left\langle u(y), \int e^{i\langle y,\xi\rangle}\widetilde{p}(y,\xi)\widehat{v}(\xi)\, đ\xi \right\rangle_{\mathbb{C}^N} dy \\ &= \int \left\langle \left\{ \int e^{-i\langle y,\xi\rangle} \widetilde{p}(y,\xi)^* u(y)\, đy \right\}, \widehat{v}(\xi) \right\rangle_{\mathbb{C}^M} đ\xi,\end{aligned}$$

where $\widetilde{p}(y,\xi)^*$ is the adjoint matrix of the amplitude $\widetilde{p}(y,\xi)$ of the operator P^*. By Parseval's Formula, $\widehat{Pu}(\xi)$ is exactly the expression in the curly braces. Hence, we have the *additional formula*

$$\widehat{Pu}(\xi) = \int e^{-i\langle y,\xi\rangle} \left(\widetilde{p}(y,\xi)\right)^* u(y)\, d'y. \tag{8.14}$$

For (a): This time we do not rely on the global representation of P and Q, but rather on their definition by localizations. Without loss of generality, let X be an open, relatively compact subset of $\mathbb{R}^n$. Let p and q be the amplitudes of orders k and j, belonging to P and Q, respectively. By (8.14), we obtain for

$$(QPu)(x) = \int e^{i\langle x-y,\xi\rangle} q(x,\xi) \left(\widetilde{p}(y,\xi)\right)^* u(y)\, đy\, đ\xi.$$

which is thus a pseudo-differential operator (of order $k+j$) by Theorem 8.13 (of Kuranishi), p. 213, and

$$\begin{aligned}\sigma_{k+j}(Q\circ P)(x,\eta) &= \lim_{\lambda\to\infty}\frac{q(x,\lambda\eta)\,(\widetilde{p}(y,\lambda\eta))^*}{\lambda^{k+j}}\\ &= \lim_{\lambda\to\infty}\frac{q(x,\lambda\eta)}{\lambda^j}\lim_{\lambda\to\infty}\frac{(\widetilde{p}(y,\lambda\eta))^*}{\lambda^k}\\ &= \sigma_j(Q)(x,\eta)\circ\sigma_k(P)(x,\eta),\end{aligned}$$

since $\sigma_k(\widetilde{p}) = \sigma_k(p)^*$ by (b). □

REMARK 8.27. The derivation of the formally adjoint operator seems trivial here in comparison to the lengthy calculation for differential operators; see Exercise 6.39, p.181. Actually, we have three entirely different problems: By definition, it is trivial that every Fourier-integral operator P possesses a formal adjoint P^*. To prove that P^* is a pseudo-differential operator if P is, we need more (namely, the Theorem of Kuranishi or somewhat long-winded direct computations). It is possible, by the way, to prove the *sharper* result $P\in \mathrm{Diff}_k \Rightarrow P^*\in \mathrm{Diff}_k$ in this fashion, by analyzing carefully the various transformations in the Theorem of Kuranishi with this goal in mind.

5. Normalized (Global) Amplitudes

Motivation. Our aim is to provide a framework for a proof of the Index Theorem, which, when compared to existing approaches, we believe is somewhat more streamlined and globally expressed (i.e., free of local coordinates). This method is based on defining pseudo-differential operators from sections of a vector bundle $E\to X$ to sections of a bundle $F\to X$ in terms of a *globally* defined symbol which is a section $p\in C^\infty(\mathrm{Hom}(\pi^*E,\pi^*F))$ of the bundle of homomorphisms between the lifts π^*E and π^*F to the cotangent bundle T^*X, where $\pi : T^*X\to X$. The definition of the operator, say $\mathrm{Op}(p)\in C^\infty(E,F)$, associated with p entails the introduction of a metric on X and connections on E and F. It could be argued that (without a considerable background in modern differential geometry) this is not easier than using local coordinates and framings, but there are advantages. This global approach to pseudo-differential operators is not new. It seems to have first appeared in the paper of Juliane Bokobza-Haggiag [**Bok**]. Many subsequent developments and applications have appeared steadily since then, in the work of Harold Widom [**Wid**], Ezra Getzler [**Get**], Marcus Pflaum [**Pf**] and Theodore Voronov [**Vor**], just to mention a few.

The application of the global approach to the index theorem, has been mostly in the context of the heat equation proof, rather than in the embedding proof. Nevertheless, in preliminaries leading up to their treatment of the embedding proof in their enlightening book, H. Blaine Lawson and Marie-Louise Michelsohn [**LaMi**, p. 188], point out the possible desirability of defining pseudo-differential operators with a global symbol. In essence, here we are exploring this possibility. We find that some of the difficulties are softened. In particular, the lifting a pseudo-differential operator to an invariant one in the proof of the twisted multiplication formula is made easier, since the global symbol can be lifted by means of a connection. Moreover the thorny problem of forming suitable products of individual pseudo-differential operators with identity operators over product manifolds (see [**LaMi**,

p. 250f] or [**AS68a**, p. 514f]) is alleviated by performing operations on the globally-defined total product symbols. This will be made clearer below. One fundamental challenge inspired by this program is the task of constructing a global symbol whose associated pseudo-differential operator is exactly the operator that one may want, as opposed to an approximate operator with essentially the same asymptotic principal symbol.

Normalized Fourier Transform. We work within the C^∞ category unless stated otherwise. Let ρ be the *injectivity radius* of the compact manifold X with Riemannian metric g and *Levi-Civita connection* ∇; i.e., for all $x \in X$, the exponential map $\exp_x : T_xX \to X$ relative to g is injective on the disk of radius ρ about $0_x \in T_xX$. Let $\pi_E : E \to X$ and $\pi_F : F \to X$ be complex Hermitian vector bundles equipped with *Hermitian connections* $\nabla^E : C^\infty(E) \to C^\infty(T^*X \otimes E)$ and $\nabla^F : C^\infty(F) \to C^\infty(T^*X \otimes F)$, where $C^\infty(E)$ denotes the space of (smooth) sections of $\pi_E : E \to X$. For $x, y \in X$, with $d(x, y) < \rho$, let $\tau^E_{x,y} : E_y \to E_x$ denote *parallel translation* relative to ∇^E along the unique *geodesic* from y to x with minimal length $d(x, y)$. Here We emphasized in italics the terms which were introduced more precisely in Sections 6.4 and 6.6.

Let $\psi : [0, \infty) \to [0, 1]$ be smooth, with $\psi(r) = 1$ for $r \in [0, \rho/3]$ and $\psi(r) = 0$ for $r \in [2\rho/3, \infty]$.

Definition 8.28. For $\pi : T^*X \to X$ and $u \in C^\infty(E)$, we define the **normalized Fourier transform** $u^\wedge \in C^\infty(\pi^*E)$ (where $x = \pi(\xi)$ and $\xi \in T^*X$) by

$$u^\wedge(\xi) := \int_{T_xX} e^{-i\xi(v)} \psi(|v|) \tau^E_{x, \exp_x v} \left[u(\exp_x v)\right] đv \in E_x \text{ for } \xi \in T^*_xX.$$

where $đv = (2\pi)^{-n/2} dv$ and dv is the volume element on T_xX associated with g_x.

For $x, y \in X$ with $d(x, y) < \rho$, we have $y = \exp_x v$ for a unique $v \in T_xX$ with $|v| = d(x, y)$, and we may define $\alpha \in C^\infty(X \times X, [0, 1])$ by

$$\alpha(x, y) := \begin{cases} \psi\left(d(x, y)\right) = \psi(|v|), & \text{for } d(x, y) < \rho\ , \\ 0, & \text{for } d(x, y) \geq \rho. \end{cases}$$

Note that we can think of the function (in $C^\infty(T_xX, E_x)$)

$$v \mapsto \psi(|v|) \tau^E_{x, \exp_x v} \left[u(\exp_x v)\right] = \tau^E_{x, \exp_x v} \left[\alpha(x, \exp_x v) u(\exp_x v)\right] \quad (v \in T_xX)$$

as a *pull-back* (of sorts), using τ^E and $\exp_x : T_xX \to X$, of the bump function $\alpha(x, \cdot)$ times $u(\cdot)$ in a neighborhood of x, and $u^\wedge|_{T^*_xX}$ is the Fourier transform of this "pull-back" of $\alpha(x, \cdot)u(\cdot)$. The **normalized inverse Fourier transform** $(u^\wedge)^\vee : TX \to E$ of $u^\wedge$ is given by

$$\left(u^\wedge\right)^\vee(v) := \int_{T^*_xX} e^{i\xi(v)} u^\wedge(\xi)\, đ\xi = \psi(|v|) \tau_{x, \exp_x v} \left[u(\exp_x v)\right],$$

where $đ\xi = (2\pi)^{-n/2} d\xi$. Since $\left(u^\wedge\right)^\vee(v) \in E_x$, $\left(u^\wedge\right)^\vee$ is a section of the pull-back of E to TX via $\pi : T^*X \to X$. Moreover, we can recover u locally about x from $u^\wedge{}_{|T^*_xX}$. In particular, for $v = 0_x \in T_xX$, we have $\left(u^\wedge\right)^\vee(0_x) = u(x)$.

Pseudo-Differential Operators and Normalized Amplitudes. For $\pi : T^*X \to X$ and a section $p \in C^\infty(\operatorname{Hom}(\pi^*E, \pi^*F))$, (of $\operatorname{Hom}(\pi^*E,\pi^*F) \to T^*X$) we define an operator $\operatorname{Op}(p) : C^\infty(E) \to C^\infty(F)$ via

$$\operatorname{Op}(p)(u)_x := \int_{T_x^*X} e^{i\xi(v)} p(\xi)\left(u^\wedge(\xi)\right)\, đ\xi \bigg|_{v=0} = \int_{T_x^*X} p(\xi)\left(u^\wedge(\xi)\right)\, đ\xi$$

$$= \int_{T_x^*X} p(\xi)\left(\int_{T_xX} e^{-i\xi(v)}\psi(|v|)\tau^E_{x,\exp_x v}\;\left[u(\exp_x v)\right] d'v\right) đ\xi$$

$$= \int_{T_xX\times T_x^*X} p(\xi)\left(e^{-i\xi(v)}\psi(|v|)\tau^E_{x,\exp_x v}\;\left[u(\exp_x v)\right]\right) đvđ\xi$$

$$= \int_{T_xX\times T_x^*X} e^{-i\xi(v)}p(\xi)\left(\tau^E_{x,\exp_x v}\;\left[\alpha(x,\exp_x v)u(\exp_x v)\right]\right) đvđ\xi. \tag{8.15}$$

Remark 8.29. a) In the spirit of Remark 8.4, p.208, one may think of $\operatorname{Op}(p)$ as a quantization of p, but $\operatorname{Op}(p)$ continues to depend on many choices also in our global setting (e.g., the choice of metric, connections, and $\alpha : X \times X \to [0,1]$).
b) Apart from these choices, there are other choices one can make, as is discussed in [**Vor**]. For example, if $s \in [0,1]$, let

$$T_{x,\exp_x sv} : T^*_{\exp_x sv}X \to T^*_xX$$

denote parallel translation (with respect to the Levi-Civita connection) for T^*X along the geodesic $t \mapsto \exp_x tv$ in the reverse direction from $\exp_x sv$ to x. In [**Vor**] (but with notation that differs from ours), an operator $\operatorname{Op}(p;s)$ (depending on s) is associated to p via

$$\operatorname{Op}(p;s)(u)_x = \int_{T_xX\times T_x^*X} đvđ\xi\; e^{-i\xi(v)}$$
$$\alpha(x,\exp_x v)\tau^F_{x,\exp_x sv}p(T_{x,\exp_x sv}(\xi))\tau^E_{\exp_x sv,\,\exp_x v}u(\exp_x v).$$

When $s = 0$, we get

$$\operatorname{Op}(p;0)(u)_x = \int_{T_xX\times T_x^*X} đvđ\xi\; e^{-i\xi(v)}\alpha(x,\exp_x v)p(\xi)\tau^E_{x,\,\exp_x v}\left(u\left(\exp_x v\right)\right)$$

which is precisely our $\operatorname{Op}(p)$. In cases of interest, the operators $\operatorname{Op}(p;s)$ for different s differ by "lower order" operators which do not affect the index (if defined). Hence, for simplicity, we only use $s = 0$. As stated in [**Vor**] the choice of s is related to the choice of operator ordering of monomials in position and momentum variables under quantization.

The connections ∇^E and ∇^F pull back via $\pi : T^*X \to X$ to connections on the bundles $\pi^*E \to T^*X$ and $\pi^*F \to T^*X$, which we continue to denote by ∇^E and ∇^F. The Levi-Civita connection for (X,g) determines a subbundle H of $T(T^*X)$ consisting of horizontal subspaces of $T(T^*X)$, which is complementary to the subbundle V of $T(T^*X)$ consisting of vectors which are tangent to the fibers of $T^*X \to X$. There is a natural Riemannian metric, say g^*, on T^*X such that V and H are orthogonal and g^* equals g on V and π^*g on H. Using ∇^E and ∇^F, along with the Levi-Civita connection for g^*, say ∇^*, we may construct a covariant derivative

$$\widetilde{\nabla} : C^\infty(\operatorname{Hom}(\pi^*E,\pi^*F)) \to C^\infty\left(T^*\left(T^*X\right)\otimes \operatorname{Hom}(\pi^*E,\pi^*F)\right).$$

Since ∇^* extends to $\otimes^k T^*(T^*X)$, we may "iterate" $\widetilde{\nabla}$ to obtain

$$\widetilde{\nabla}^k : C^\infty(\mathrm{Hom}(\pi^*E,\pi^*F)) \to C^\infty\left(\otimes^k T^*(T^*X)\otimes \mathrm{Hom}(\pi^*E,\pi^*F)\right).$$

Now we are ready to give a geometric and truly global definition of a variant of our spaces of principally classical pseudo-differential operators, introduced locally in Section 8.3.

DEFINITION 8.30. a) We say that $p \in C^\infty(\mathrm{Hom}(\pi^*E,\pi^*F)$ is a **normalized amplitude** (or **total symbol**) of order $k \in \mathbb{R}$ if for any $H_1,\dots,H_I \in C^\infty(H)$ with $|H_1|,\dots,|H_I| \le 1$ and $V_1,\dots,V_J \in C^\infty(V)$, there are constants C_{IJ} (depending only on I, J and p), such that

$$\left|\left(\widetilde{\nabla}^{I+J}p\right)(H_1,\dots,H_I,V_1,\dots,V_J)\right| \le C_{IJ}\left(1+\sum\nolimits_{j=1}^{J}|V_j|\right)^{k-J}. \tag{8.16}$$

Moreover, we require that the k-th order asymptotic **symbol** (or **principal symbol**) of p, namely

$$\sigma_k(p)(\xi) := \lim_{t\to\infty}\frac{p(t\xi)}{t^k} \quad (\text{for } \xi \ne 0) \tag{8.17}$$

exists, where the convergence is uniform on $S(T^*X)$.
b) Then we call $\mathrm{Op}(p)$ a **normalized pseudo-differential operator** of order k and write $\mathrm{Op}(p) \in \mathrm{L}^k_{\mathrm{norm}}(E,F)$. We set $\mathrm{L}^\bullet_{\mathrm{norm}}(E,F) := \bigcup \mathrm{L}^k_{\mathrm{norm}}(E,F)$.
c) We denote the set of amplitudes of order k by $\mathrm{Ampl}_k(E,F)$.

Clearly, for $k' > k$,

$$\mathrm{Ampl}_{k'}(E,F) \supset \mathrm{Ampl}_k(E,F) \supset \mathrm{Ampl}_{-\infty}(E,F) := \bigcap\nolimits_{k=0}^{-\infty}\mathrm{Ampl}_k(E,F).$$

For $p \in \mathrm{Ampl}_k(E,F)$, we then have the operator, say $\mathrm{Op}(p) : C^\infty(E) \to C^\infty(F)$, given by (8.15), which extends to a bounded operator $\mathrm{Op}_s(p) : W^s(E) \to W^{s-k}(F)$, where for any $s \in \mathbb{R}$, $W^s(E)$ denotes the s-th **Sobolev space** of sections of E, namely the completion of $C^\infty(E)$ with respect to the norm $\|\cdot\|_s$ defined by

$$\|u\|_s^2 := \int_{T^*X}\left(1+|\xi|^2\right)^s |u^\wedge(\xi)|^2\,d\xi.$$

Recall that for $k \in \mathbb{Z}^+$, and $s > n/2 + k$, there is a compact inclusion $W^s(E) \subset C^k(E)$. For each s, the linear map

$$\mathrm{Op}_s : \mathrm{Ampl}_k(E,F) \to \mathcal{B}(W^s(E), W^{s-k}(F)) \tag{8.18}$$

into the Banach space $\mathcal{B}(W^s(E), W^{s-k}(F))$ of bounded linear transformations is continuous (see [**LaMi**, p. 177f]). Moreover, for $\varphi \in \mathrm{Ampl}_{-\infty}(E,F)$, $\mathrm{Op}_s(\varphi)$ is a compact operator for any $s \in \mathbb{R}$, and $\mathrm{Op}_s(\varphi)(W^s(E)) \subset C^\infty(F)$; i.e., $\mathrm{Op}_s(\varphi)$ is a **smoothing operator**.

DEFINITION 8.31. We say that $p \in \mathrm{Ampl}_k(E,F)$, and the corresponding operator $\mathrm{Op}(p)$, are **elliptic** if for some constant $c > 0$, $p(\xi)^{-1}$ exists for $|\xi| > c$, and for some constant $K > 0$

$$\left|p(\xi)^{-1}\right| \le K(1+|\xi|)^{-k} \text{ for all } \xi \in T^*X \text{ with } |\xi| > c.$$

We set $\mathrm{Ell}_k(E,F) := \{p \in \mathrm{Ampl}_k(E,F) : p \text{ is elliptic}\}$.

For $p \in \mathrm{Ell}_k(E,F)$, there are $q \in \mathrm{Ampl}_{-k}(E,F)$, $\varphi_E \in \mathrm{Ampl}_{-\infty}(E,E)$ and $\varphi_F \in \mathrm{Ampl}_{-\infty}(F,F)$, such that

$$\mathrm{Op}_{s-k}(q) \circ \mathrm{Op}_s(p) = \mathrm{Id}_{W^s(E)} + \mathrm{Op}_s(\varphi_E) \text{ and}$$
$$\mathrm{Op}_s(p) \circ \mathrm{Op}_{s-k}(q) = \mathrm{Id}_{W^{s-k}(F)} + \mathrm{Op}_{s-k}(\varphi_F).$$

Since $\mathrm{Op}_s(\varphi_E)$ and $\mathrm{Op}_{s-k}(\varphi_F)$ are compact operators, it follows that $\mathrm{Op}_s(p)$ is Fredholm, and hence we may define

$$\mathrm{index}(\mathrm{Op}_s(p)) := \dim\ker(\mathrm{Op}_s(p)) - \dim\mathrm{coker}(\mathrm{Op}_s(p)).$$

Note also that if $\mathrm{Op}_s(p)u \in C^\infty(F)$, then

$$u = \mathrm{Op}_{s-k}(q)\,(\mathrm{Op}_s(p)u) - \mathrm{Op}_s(\varphi_E)u \in C^\infty(E).$$

Thus, $\dim\ker(\mathrm{Op}_s(p)) < \infty$, $\ker(\mathrm{Op}_s(p)) \subset C^\infty(E)$, and $\ker(\mathrm{Op}_s(p))$ is independent of s. As a consequence,

$$\mathrm{index}(\mathrm{Op}_s(p)) = \dim\ker(\mathrm{Op}(p)) - \dim\mathrm{coker}(\mathrm{Op}(p))$$

is independent of s.

Approximation of Differential Operators. The reader should be aware of a minor technical problem when dealing with our normalized amplitudes and normalized pseudo-differential operators: There can be differential operators which can not be generated by a normalized amplitude.

Let us have a closer look at the familiar case of an elliptic, linear differential operator $D : C^\infty(E) \to C^\infty(F)$ of given order k. As we have seen in the preceding chapter, associated with D is its principal symbol $\sigma_k(D) \in C^\infty(\mathrm{Hom}(\pi^*E, \pi^*F))$. We have checked that $\sigma_k(D)$ is independent of the choice of local coordinates and observed that this would not be the case if lower-order terms were included. If $\sigma_k(D)$ is invertible outside of the zero section of T^*X, then D is said to be elliptic, which we assume. If lower order terms were included and if we denoted this coordinate-dependent, locally-defined "full symbol" by $p_{\mathrm{loc}}(D)\,(\xi)$, then $\sigma_k(D)\,(\xi)$ at $\xi \in T^*_x X$ would be given by

$$\lim_{t\to\infty} \frac{p_{\mathrm{loc}}(D)\,(t\xi)}{t^k}$$

in comparison with (8.17). However, it is not clear that D is precisely $\mathrm{Op}(p)$ for some globally defined $p \in \mathrm{Ampl}_k(E,F)$. In the language of physicists, it is not clear that D can be precisely dequantized. If such p exists, it would clearly depend on choices of a Riemannian metric on X, connections for E and F and on the function $\alpha : X \times X \to [0,1]$ supported near the diagonal. However, in [**Bok**] and [**Wid**], it is shown that given such choices, p can be found so that $\mathrm{Op}(p)$ and D differ by an operator which is infinitely smoothing (and hence compact); i.e.,

$$D - \mathrm{Op}(p) = \mathrm{Op}(a) \text{ for } a \in \mathrm{Ampl}_{-\infty}(E,F).$$

By methods that are standard by now and to be summarized in the following chapter, it follows that D has Fredholm Sobolev extensions $D_s : W^s(E) \to W^{s-m}(F)$ for all s, with a common index, which is sometimes called the *analytic* index of D; it is just the usual operator-theoretic index. It is simply denoted by $\mathrm{index}(D)$ and if $D^* : C^\infty(F) \to C^\infty(E)$ denotes the formal L^2-adjoint of D, then

$$\begin{aligned}\dim\ker(D) - \dim\ker(D^*) &= \mathrm{index}(D)\\ &= \mathrm{index}(\mathrm{Op}(p) + \mathrm{Op}(a)) = \mathrm{index}(\mathrm{Op}(p)).\end{aligned}$$

Thus, readers (including the authors) who are bothered by the fact that differential operators may not be precisely dequantized, may take some solace in the fact that elliptic differential operators may be approximated by a pseudo-differential operator of the form $\mathrm{Op}(p)$, modulo smoothing operators which preserve the index.

We close this section with a few exercises.

Better hints? All correct?

EXERCISE 8.32. a) Show that $\mathrm{L}^{\bullet}_{\mathrm{norm}}(E,F)$ of Definition 8.30b is a graded *-algebra.
b) Show that the *quantization* $\mathrm{Op} : \mathrm{Ampl}_k(E,F) \to \mathrm{L}^k_{\mathrm{norm}}(E,F)$ is a bijection for all $k \in \mathbb{R}$.
c) Show that the spaces $\mathrm{Ampl}_k(E,F)$ and $\mathrm{L}^k_{\mathrm{norm}}(E,F)$ are independent of the choice of metrics and connections (contrary to the definition of Op).
d) Prove $\mathrm{L}^k_{\mathrm{norm}}(E,F) \subset \mathrm{L}^k_{\mathrm{pc}}(E,F)$.
e) Find a closed Riemannian manifold X, Hermitian bundles $E,F \to X$ with fixed connections and $k \in \mathbb{N}$ such that $\mathrm{Diff}_k(E,F) \not\subset \mathrm{L}^k_{\mathrm{norm}}(E,F)$. Conclude $\mathrm{L}^k_{\mathrm{norm}}(E,F) \neq \mathrm{L}^k_{\mathrm{pc}}(E,F)$. For that, can E,F be trivial bundles? Can you choose $X = S^1$ or, more generally, $X = S^n$?
[Hint: To a: Check composition and taking formal adjoints like in Section 8.4.
To b: By definition of $\mathrm{L}^k_{\mathrm{norm}}(E,F)$ and the linearity of Op it suffices to prove the injectivity of Op_s of (8.18) for $s = 0$. How can you exclude the existence of a not identically vanishing $p \in \mathrm{Ampl}_k(E,F)$ with $\mathrm{Op}(p) = 0$ in spite of the Example 8.15, p.215?
To c: Nasty!
To d: Fix the metric structures and connections. Then check the claim in coordinates.
To e: See the references given at the beginning of this section, in particular [**Bok**], [**Wid**] and [**Pf**].]

CHAPTER 9

Elliptic Operators over Closed Manifolds

In this chapter, we show that out of formal properties (e.g., invertibility) of elliptic symbols, a series of existence, regularity, and finiteness results for the associated differential and pseudo-differential operators (as Fredholm operators) can be obtained.

1. Continuity of Pseudo-Differential Operators

Convention: In what follows, the manifold X is *closed*, i.e. compact, without boundary. We make this convention, in part for convenience (in order to make some proofs go easier), but also because otherwise some of the following theorems would be meaningless or false; see Exercise 9.16 below. Moreover, X continues to be oriented and is furnished with a fixed Riemannian metric; E and F are Hermitian vector bundles. Without loss of generality, we will occasionally assume that the *Hilbertable* Sobolev spaces are already furnished with a fixed norm or scalar product.

Definition 9.1. A linear operator $P : C_0^\infty(E) \to C^\infty(F)$ is called an **operator of order** $k \in \mathbb{Z}$ if it extends to a continuous map $P_s : W^s(E) \to W^{s-k}(F)$ for all $s \in \mathbb{R}$ with $s, s-k > 0$. We denote by $\mathrm{OP}_k(E,F)$ the set of operators of order k.

Theorem 9.2. *For* $k \in \mathbb{Z}$, $\mathrm{L}^k_{\mathrm{pc}}(E,F) \subseteq \mathrm{OP}_k(E,F)$.

Proof. Formally, this theorem says that the *analytical order* of a pseudo-differential operator (determined by the asymptotic behavior its amplitude) coincides with its *order in the context of functional analysis* (which is expressed by its continuity relative to the norms of the Sobolev spaces). We need to prove the estimate $\|Pu\|_{s-k} \le C\,\|u\|_s$, $u \in C_0^\infty(E)$, where C only depends on $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$ and s, and not on u. Since $C_0^\infty(E)$ is dense in $W^s(E)$, the theorem follows, because then P can be extended to a continuous linear operator $P_s : W^s(E) \to W^{s-k}(F)$.

Since the norms in $W^s(E)$ and $W^{s-k}(F)$ are locally defined by Exercise 7.7 (p. 194), it suffices to show the inequality for $u \in C_0^\infty(\mathbb{R}^n)$. By Exercise 7.5 (p. 194) we can further assume without loss of generality that $k = 0$ and $s = 0$. Let

$$(Pu)(x) = \int e^{i\langle x,\xi\rangle} p(x,\xi)\widehat{u}(\xi)\, đ\xi,\ \ u \in C_0^\infty(\mathbb{R}^n),$$

be a pseudo-differential operator, whose amplitude $p(x,\xi)$ vanishes for sufficiently large x and satisfies the estimate (see 8.2, p. 207)

$$\left|D_x^\beta D_\xi^\alpha p(x,\xi)\right| \le C\,(1+|\xi|)^{k-|\alpha|}\,,\ (x,\xi) \in \mathbb{R}^n \times \mathbb{R}^n,$$

for all multi-indices α and β. Then the Fourier transform $\widehat{p}(\cdot,\xi)$ of the function $x \mapsto p(x,\xi)$ can be estimated by

$$|\widehat{p}(z,\xi)| \le C_N(1+|z|)^{-N} \text{ for all } N \in \mathbb{N}, \tag{9.1}$$

as was done for $\widehat{u}$ in the proof of Theorem 8.3, p. 208. We have

$$\begin{aligned}\widehat{Pu}(\eta) &= \int\int e^{-i\langle \eta, x\rangle} e^{i\langle x,\xi\rangle} p(x,\xi)\, đx\ \widehat{u}(\xi)\, đ\xi \\ &= \int \widehat{p}(\eta-\xi,\xi)\,\widehat{u}(\xi)\, đ\xi.\end{aligned}$$

By (9.1) and the Schwarz Inequality,

$$\begin{aligned}|\widehat{Pu}(\eta)|^2 &\le C_N^2 \left|\int (1+|\eta-\xi|)^{-N}\widehat{u}(\xi)\, đ\xi\right|^2 \\ &\le (2\pi)^{-n} C_N^2 \left(\int (1+|\eta-\xi|)^{-2N}\, đ\xi\right) \|\widehat{u}\|_0^2 .\end{aligned}$$

For N sufficiently large, integrating this with respect to η and using Parseval's Formula (Appendix A, Exercise A.5d), we obtain the desired result

$$\|Pu\|_0^2 = ||\widehat{Pu}||_0^2 \le C' \|\widehat{u}\|_0^2 = C' \|u\|_0^2 .$$

□

EXERCISE 9.3. Interpret and prove:

$$P \in \mathrm{L}^k_{\mathrm{pc}}(E,F) \Rightarrow (P_0)^* = (P^*)_0.$$

EXERCISE 9.4. Let $\sigma_k : \mathrm{L}^k_{\mathrm{pc}}(E,F) \to \mathrm{Smbl}_k(E,F)$ be the well-defined (Exercise 8.19, p. 221), surjective (Theorem 8.24, p. 225) symbol map. Show $\operatorname{Ker}\sigma_k \subseteq \mathrm{OP}_{k-1}(E,F)$. [Hint: This is trivial by axiom (iii) (see Section 8.3, p. 211) and the preceding Theorem 9.2.]

EXERCISE 9.5. Show that the short exact sequence

$$\mathrm{L}^k_{\mathrm{pc}}(E,F) \overset{\sigma_k}{\to} \mathrm{Smbl}_k(E,F) \to 0$$

splits; i.e., σ_k has a linear right inverse $\chi_k : \mathrm{Smbl}_k(E,F) \to \mathrm{L}^k_{\mathrm{pc}}(E,F)$ which satisfies the continuity condition

$$\begin{aligned}&\sup\left\{\|\chi_k(\rho)(u)\|_{s-k} : u \in C^\infty(X) \text{ and } \|u\|_s = 1\right\} \\ &\le C \sup\left\{|\rho(x,\xi)| : x \in X,\ |\xi| = 1\right\}\ (s, s-k > 0),\end{aligned}$$

where C does not depend on $\rho \in \mathrm{Smbl}_k(E,F)$, and $|\rho(x,\xi)|$ is the usual matrix norm which may be defined via the Hermitian inner products on E_x and F_x. [Hint: Go through the proof of Theorem 8.24 (p. 225) again.]

EXERCISE 9.6. Show conversely, that for all $s, s-k > 0$ and all $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$, we have the following inequality

$$\begin{aligned}&\sup\left\{|\sigma_k(P)(x,\xi)| : x \in X,\ |\xi| = 1\right\} \\ &\le \sup\left\{\|Pu\|_{s-k} : u \in C^\infty(X) \text{ and } \|u\|_s = 1\right\}.\end{aligned}$$

[Hint: By a theorem of Israil Gohberg (see also [**See65**, p.171]), one can find a sequence $\{\varphi_\nu\}$ of functions in $C_0^\infty(\mathbb{R}^n)$ such that

(*i*) $\varphi_\nu(x) = 0$ for $|x - x_0| > 1/\nu$
(*ii*) $\|\varphi_\nu\|_0 = 1$ for all ν, and
(*iii*) $\|P\varphi_\nu - \sigma(P)(x_0,\xi_0)\varphi_\nu\|_0 \to 0$ as $v \to \infty$, where $(x_0,\xi_0) \in \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})$ is any given point.

Details are found in [**See65**, p.179] or [**CaSc**, 22-05]. Caution: On $\mathrm{L}^k_{\mathrm{pc}}(E,F)$ itself there is a topology which is defined in a natural way by the condition that the map $P \mapsto P_s$ be continuous for all s. However, then σ_k is *not* continuous; see [**Pal65**, p.175].]

2. Elliptic Operators - Regularity and Fredholm Property

As a generalization of our earlier definition for differential operators (see Section 6.7 above), we call $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$ **elliptic**, if $\sigma_k(P)\,(x,\xi)$ is an isomorphism from E_x to F_x for all $x \in X$ and $\xi \in T^*_x X$, $\xi \neq 0$. We write $P \in \mathrm{Ell}_k(E,F)$.

THEOREM 9.7 (Main result). *For any $P \in \mathrm{Ell}_k(E,F)$, there is $Q \in \mathrm{Ell}_{-k}(E,F)$, such that $PQ - \mathrm{Id}_F \in \mathrm{OP}_{-1}(F,F)$ and $QP - \mathrm{Id}_E \in \mathrm{OP}_{-1}(E,E)$.*

REMARK 9.8. This existence theorem forms the foundation of our theory of elliptic operators. Using the terminology introduced by David Hilbert, one calls Q a parametrix for P, although a *crude* one: The classical parametrix (*Green's function*) inverts P, not only modulo operators of order -1, but also of order $-\infty$ (i.e., modulo so-called *smoothing operators*; see [**Hö71a**]).

PROOF. Theorem 8.24 (p. 225) guarantees the existence of a $Q \in \mathrm{L}^{-k}_{\mathrm{pc}}(F,E)$ with $\sigma_{-k}(Q)(x,\xi) := \sigma_k(P)(x,\xi)^{-1}$, whence $PQ \in \mathrm{L}^0_{\mathrm{pc}}(F,F)$ by Theorem 8.26a (p. 226) and $\sigma_0\,(PQ - \mathrm{Id}_F) = 0$, and $\sigma_0\,(PQ - \mathrm{Id}_F) \in \mathrm{OP}_{-1}(F,F)$ by Exercise 9.4. □

THEOREM 9.9. *Let $P \in \mathrm{Ell}_k(E,F)$ and $s, s-k \geq 0$. Then we have:*

a) ***Finiteness****: The extension $P_s : W^s(E) \to W^{s-k}(F)$ is a Fredholm operator with index independent of s.*
b) ***Existence****: P^* is elliptic and* $\mathrm{Coker}\, P_s \cong \mathrm{Ker}(P^*)_{s-k}$.
c) ***Regularity****:* $\mathrm{Ker}\, P_s = \mathrm{Ker}\, P$.
d) ***Homotopy-invariance****:* $\mathrm{index}\, P = \mathrm{index}\, P_s$ *depends only on the homotopy class of $\sigma(P)$ in $\mathrm{Iso}_{SX}(E,F)$. Here $\mathrm{Iso}_{SX}(E,F)$ is the space of C^∞ bundle isomorphisms $\pi^* E \cong \pi^* F$, where $\pi : SX \to X$ is the base-point map and $SX := \{\xi \in T^*_x X : x \in X \text{ and } |\xi|_x = 1)\}$ is the co-sphere bundle; $\mathrm{Iso}_{SX}(E,F)$ is equipped with a supremum norm as in Exercise 9.5.*

REMARK 9.10. In conjunction with c), existence says that the inhomogeneous equation $Pu = f$ has a solution exactly when $f \perp \mathrm{Ker}\, P^*$, and the solution is unique if constrained to be orthogonal to $\mathrm{Ker}\, P$ in $W^0(E)$. By regularity, all *classical solutions* (i.e., $u \in C^k(E)$) of homogeneous elliptic differential equations $Pu = 0$ (with C^∞ coefficients) lie in $C^\infty(E)$. In the context of distribution theory (see [**Hö63**]), one obtains the sharper result that every *weak solution* (in the distribution sense) is a *strong solution* (in the function sense); i.e., from the assumption $u \in W^0(E)$ and $\langle u, P^* f\rangle_0 = 0$ for all $f \in C^\infty(F)$, the conclusions $u \in C^\infty(E)$ and $Pu = 0$ follow. Such regularity results, which were first proved in 1940 by Herman Weyl in the case of the Laplace operator $P := \Delta$, are of special importance when one is solving partial differential equations by variational methods (i.e., solving through extremal conditions); see [**Hö63**, p.96] and [**LiMa**, p.214f].

PROOF. **For a**: If $Q \in \mathrm{L}^{-k}_{\mathrm{pc}}(F,E)$ is a parametrix (see Theorem 9.7) for P, then it follows from the Theorem of Franz Rellich (Theorem 7.15, p. 198) that the composition

$$W^s(E) \stackrel{Q_{s-k}P_s - \mathrm{Id}}{\longrightarrow} W^{s+1}(E) \hookrightarrow W^s(E)$$

is a compact operator on $W^s(E)$, and correspondingly, $P_sQ_{s-k} - \mathrm{Id}$ is a compact operator on $W^{s-k}(F)$. Thus, $P_s : W^s(E) \to W^{s-k}(F)$ is a Fredholm operator by Theorem 3.2, p. 64. (There, actually the proof was explicitly given only for endomorphisms, but this is no restriction for separable Hilbert spaces, since they are all isomorphic.) By continuity considerations (Theorem 3.11, p. 68 and the preceding Exercises 9.5 and 9.6, or easy norm comparison for P_s and P_t by means of Λ^{s-r} of Exercise 7.5, p. 194) it follows that $\operatorname{index} P_s = \operatorname{index} P_t$.

For b: Without loss of generality (Λ argument of Exercise 7.5, p. 194), let $k = s = 0$. Then b) follows directly from Exercise 9.3 and Theorem 2.7.

For c: By definition, we have $\operatorname{Ker} P_{s+1} \subseteq \operatorname{Ker} P_s$, since $W^{s+1}(E) \subseteq W^s(E)$. Conversely, by Theorem 9.7 there is a bounded operator $K : W^s(E) \to W^{s+1}(E)$, such that $Q_{s-k}P_su - \mathrm{Id}\, u = Ku$ for all $u \in W^s(E)$, where Q is a parametrix for P. Thus $u \in W^{s+1}(E)$, if $P_su = 0$. Hence, $\operatorname{Ker} P_s = \operatorname{Ker} P_{s+1} = \cdots = \operatorname{Ker} P$, since $C^\infty = \cap W^s$; see Exercise 7.3c (p. 192) and Theorem 7.13, p. 197).

For d: For $Q \in \mathrm{Ell}_k(E,F)$ and $\sigma_k(Q) = \sigma_k(P)$, we obtain $\operatorname{index} Q = \operatorname{index} P$ from Exercise 9.4 and the invariance of the index under perturbation by compact operators (Exercise 3.10, p. 65). In general, by Exercise 9.5, each continuous curve $\rho : I \to \mathrm{Smbl}_k(E,F)$ lifts to a corresponding continuous (in the operator norm) curve $r : I \to \mathrm{L}^k_{\mathrm{pc}}(E,F)$ with $\sigma_k \circ r = \rho$. Thus, if one can connect $\sigma_k(Q)$ with $\sigma_k(P)$ by a continuous curve in $\mathrm{Iso}^\infty_{SX}(E,F)$, then P and Q can be connected in $\mathrm{Ell}_k(E,F)$. Hence for all s, the Fredholm operators Q_s and P_s lie in the same component and have the same index by Theorem 3.11, p. 68. Finally, if $Q \in \mathrm{Ell}_j(E,F)$ is an operator whose symbol $\sigma_j(Q)$ coincides with $\sigma_k(P)$ on SX, then $\sigma_j(Q)^{-1} \circ \sigma_k(P)$ is the symbol of a self-adjoint operator $R \in \mathrm{Ell}_{k-j}(E,E)$. From $\sigma_k(P) = \sigma_j(Q) \circ \sigma_{k-j}(R) = \sigma_k(QR)$, it follows by preceding arguments that

$$\operatorname{index} P = \operatorname{index} QR = \operatorname{index} Q + \operatorname{index} R = \operatorname{index} Q,$$

since $\operatorname{index} R = 0$ by (c) (see also the composition rule in Exercise 1.10, p. 9, and the following Exercise 9.11a). □

Convention: In the following, we write for short $\sigma(P)$ for the restriction of $\sigma_k(P)$ to SX.

EXERCISE 9.11. Let E, F, G, H be Hermitian vector bundles over the closed, oriented Riemannian manifold X; $P \in \mathrm{Ell}_k(E,F), Q \in \mathrm{Ell}_j(F,G), R \in \mathrm{Ell}_{k'}(G,H)$. Show that the following expressions are defined, and prove the formulas:
a) $\operatorname{index} P^* = -\operatorname{index} P$
b) $\operatorname{index} QP = \operatorname{index} P + \operatorname{index} Q$
c) $\operatorname{index} P \oplus R = \operatorname{index} P + \operatorname{index} R$
d) $\operatorname{index} P = 0$, if $\sigma(P)(x,\xi)$ depends only on x and not on $\xi \in (SX)_x$
[Hint for d: A bundle isomorphism $E \to F$ and a multiplication operator $M_\psi \in \mathrm{Ell}_0(E,F)$ are defined via $\psi(x) := \sigma(P)(x,\xi)$ for $\xi \in (SX)_x$. Apply Theorem 9.9d.]

EXERCISE 9.12. a) Form the closure $\overline{\mathrm{Smbl}_k(E,F)}$ of $\mathrm{Smbl}_k(E,F)$ in the supremum norm, and show that one then obtains all the continuous symbols. In particular, the space $\mathrm{Iso}_{SX}(E,F)$ of all continuous isomorphisms from π^*E to π^*F, where $\pi : SX \to X$ is the projection, consists of the restrictions $p|_{SX}$ for $p \in \overline{\mathrm{Smbl}_k(E,F)}$
b) For each $s, s-k \geq 0$, form the closure (in the operator norm) of the set of all

operators P_s with $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$, and show that σ_k can be continuously extended to a surjective map on this space if the target of σ_k is enlarged to $\overline{\mathrm{Smbl}_k(E,F)}$.

We now write $P \in \overline{\mathrm{L}^k_{\mathrm{pc}}(E,F)}$, if P lies in the closure formed in Exercise 9.12, for all $s \geq 0$ (and $s-k \geq 0$). One easily sees that our results up to now (in particular on elliptic operators) remain valid in this larger class. The most important reason for passing to the closure arises from the multiplicative behavior of pseudo-differential operators:

Exercise 9.13. Consider two closed, oriented, Riemannian manifolds X and Y; and Hermitian vector bundles E and F over X, and G and H over Y. Moreover, let $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$ and $Q \in \mathrm{L}^k_{\mathrm{pc}}(G,H)$, $k \in \mathbb{N}$.

a) Show that an operator $P \otimes \mathrm{Id}_G \in \overline{\mathrm{L}^k_{\mathrm{pc}}(E \otimes G, F \otimes G)}$ is defined by

$$(P \otimes \mathrm{Id}_G)\,(u \otimes v) := Pu \otimes v,\ u \in C^\infty(E), v \in C^\infty(G),$$

which does not always lie in $\mathrm{L}^k_{\mathrm{pc}}(E \otimes G, F \otimes G)$.

b) Over the manifold $X \times Y$, define the operator

$$P\#Q : C^\infty(E \otimes G) \otimes C^\infty(F \otimes H) \to C^\infty(F \otimes G) \otimes C^\infty(E \otimes H)$$

by the matrix

$$P\#Q := \begin{bmatrix} P \otimes \mathrm{Id}_G & -\mathrm{Id}_F \otimes Q^* \\ \mathrm{Id}_E \otimes Q & P^* \otimes \mathrm{Id}_H \end{bmatrix}.$$

Prove the Multiplication Theorem: If P and Q are elliptic, then $P\#Q$ is elliptic, and $\mathrm{index}\,(P\#Q) = (\mathrm{index}\,P)(\mathrm{index}\,Q)$.

[Hint for a: For the sake of simplicity, assume that all bundles are trivial line bundles (i.e., the case of functions). Then $P \otimes \mathrm{Id}_{\mathbb{C}_Y} : C^\infty(X \times Y) \to C^\infty(X \times Y)$ is the operator obtained when P acts on the first variable while the second is held fixed. In local coordinates, for $(x,y) \in X \times Y$ and $u \in C^\infty(X \times Y)$, we have:

$$(P \otimes \mathrm{Id}_{\mathbb{C}_Y})\,u(x,y) = \int\!\!\int e^{i\langle x-\bar{x},\xi\rangle} p(x,\xi) u(\bar{x},y)\, d'\bar{x}\, d\!\!\!{}^{-}\xi.$$

The amplitude $\widetilde{p}$ of $P \otimes \mathrm{Id}_{\mathbb{C}_Y}$ is then given by $\widetilde{p}(x,y,\xi,\eta) = p(x,\xi)$ (up to a constant of integration). Show that the amplitude estimate

$$\left|D^\beta_x D^\alpha_\xi p(x,\xi)\right| \leq C\,(1 + |\xi| + |\eta|)^{k-|\alpha|}.$$

can only hold for large $|\alpha|$ when $D^\beta_x D^\alpha_\xi p(x,\xi)$ is identically zero (i.e., when p is a polynomial in ξ). For the proof that

$$P \otimes \mathrm{Id}_{\mathbb{C}_Y} \in \mathrm{L}^k_{\mathrm{pc}}(X \times Y) := \overline{\mathrm{L}^k_{\mathrm{pc}}(\mathbb{C}_{X\times Y}, \mathbb{C}_{X\times Y})},$$

explicitly construct a family $\{R^t : t \in (0,1]\}$ with

$$R^t \in \mathrm{L}^0_{\mathrm{pc}}(X \times Y) \text{ and } (P \otimes \mathrm{Id}_{\mathbb{C}_Y})\,R^t \in \mathrm{L}^k_{\mathrm{pc}}(X \times Y),$$

such that $(P \otimes \mathrm{Id}_{\mathbb{C}_Y})\,R^t$ converges in the operator norm to $P \otimes \mathrm{Id}_{\mathbb{C}_Y}$ as $t \downarrow 0$, as in [**AS68a**, p.513-516]. Another proof can be found in [**Hö71b**, p.96f], where the consideration of the difference variable z in x and y (and ζ in ξ and η) is carried out in the framework of the theory of Fourier integral operators, with its more flexible methods; see also the theorem of Kuranishi (Theorem 8.13, p. 213).

For b: For the origin of the somewhat strange form of $P\#Q$ compare with *golden rule* of tensoring chain complexes; see also Chapters 11 and 12. Details may be found in [**See65**, p.190-193], [**CaSc**, exp. 22], [**AS68a**, p.526-529].]

EXERCISE 9.14. Let $P \in \mathrm{L}^k_{\mathrm{pc}}(E,F)$ be an elliptic operator (i.e., $\sigma(P) \in \mathrm{Iso}_{SX}(E,F)$). Show that $\operatorname{index} P = 0$, if $\sigma(P)$ can be extended to an isomorphism over all of $BX := \{\xi \in T^*X : |\xi| \le 1\}$.
[Hint: Show that a homotopy between the symbol of P and the symbol of the multiplication operator $(Mu)(x) := \sigma(P)(x,0)(u(x))$ can be defined and apply Theorem 9.9d and Exercise 9.11d.]

EXERCISE 9.15. Now, let E and F be trivial line bundles over the closed manifold X. Show that $\operatorname{index} P = 0$, if $\dim X > 2$. [Hint: Reduce this to Exercise 9.14 by a suitable deformation of $\sigma(P)(x,\xi)$; see [**Ni**, p.160f]. Compare also with Section 14.3 below.]

EXERCISE 9.16. Carry the following theorems and exercises over to the case of manifolds with boundary: Theorem 9.2, Exercise 9.3 (if the formal adjoint operator is defined by $\langle Pu, v\rangle = \langle u, P^*v\rangle$ for all u and v with support contained in the interior of X), Exercises 9.4–9.6, Theorem 9.7 (Why not Theorem 9.9?), Exercise 9.12, and Exercise 9.13a.

CHAPTER 10

Local Elliptic Boundary-Value Systems

In this chapter, we shall develop a systematic theory of <u>*local*</u> *elliptic boundary-value systems. Here* <u>*systematic*</u> *means a generality and simplicity historically created to yield convenient proofs, and hence apparent arbitrariness. Put differently, as in the preceding chapters on Sobolev spaces, pseudo-differential operators and symbolic calculus, we mostly follow the course of history and the analysis mainstream, contrary to our own view and preferences: With hindsight, any rational discussion of elliptic boundary conditions should begin with R.T. Seeley's profound view of 1968 in* [**See68**]*, on what kind of boundary conditions yield* <u>*Fredholm operators*</u> *and* <u>*regularity*</u>*, the two essentials of the idea of* <u>*ellipticity*</u>*. Seeley's ideas found a spectacular application in the famous* <u>*Atiyah-Patodi-Singer Index Theorem*</u> *for boundary conditions defined by spectral projections, see our Section 14.8, p.323ff in Part III. There we shall give a short summary of Seeley's ideas and the main subsequent results. For a comprehensive presentation we refer to the monograph* [**BoWo**] *and the more recent review* [**BruLes**].

We postpone our principal, global view upon the index theory of elliptic boundary value problems for two reasons: firstly to <u>*compromise*</u> *with the preference of analysis mainstream for local theory, and secondly to show the* <u>*beautiful*</u> *relations between the local theory and global invariants.*

1. Differential Equations in Half-Space

Here we shall discuss the <u>*idea*</u> *of ellipticity of local boundary-value conditions, i.e., the connection between boundary-value problems of partial differential equations and initial conditions of ordinary differential equations and thus finally the* <u>*algebraic essence*</u> *of that idea of local ellipticity.*

We begin with the local considerations which arise from the classical theory of homogeneous differential equations with constant coefficients in the half-space. Thus, for the moment, let p be a homogeneous, complex-valued polynomial of degree k in n variables and $p(D) = p(-i\frac{\partial}{\partial x_1}, ..., -i\frac{\partial}{\partial x_n})$ be the associated homogeneous differential operator of order k with constant coefficients. In general, $\operatorname{Ker} p(D) = \{u \in C^\infty(\mathbb{R}^n) : p(D)u = 0\}$ will be an infinite dimensional function space. For example, for $n = 2$, let $p(\xi_1, \xi_2) = \xi_1^2 + \xi_2^2$, hence $p(D) = -\Delta$, whose kernel consists of all harmonic functions. From numerical mathematics, which is concerned with the approximation of arbitrary solutions by solutions of particularly simple form, we know the importance (beyond the realm of ordinary linear differential equations) of the *exponential solutions.* These are functions $u \in \operatorname{Ker} p(D)$ which can be written in the form $u(x) = f(x) \exp i(x_1\xi_1 + ... + x_n\xi_n)$ where f is a polynomial in n variables and $\xi_1, ..., \xi_n \in \mathbb{C}$. In [**Hö63**, p.76 f], we find the exact formulations and the proof that the exponential solutions are dense in $\operatorname{Ker} p(D)$ for a suitable topology. While systems of ordinary differential equations with constant coefficients can be solved completely by means of elementary functions, the same is true in approximation only, for the analogous partial differential equations.

The exponential solutions play a similarly significant role in the characterization of boundary value problems:

For each $(n-1)$-tuple $\eta = (\xi_1, ..., \xi_{n-1}) \in \mathbb{R}^{n-1}$, define in $\operatorname{Ker} p(D)$ the subspace

$$\mathcal{M}_\eta := \{u \in \operatorname{Ker} p(D) : u(x) = h(x_n) \exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1}),\ h \in C^\infty(\mathbb{R})\}$$

and show that

$$\mathcal{M}_\eta \cong \left\{h \in C^\infty(\mathbb{R}) : h \in \operatorname{Ker} p(\xi_1, ..., \xi_{n-1}, -i\tfrac{d}{dt})\right\}.$$

Trick: Compute that

$$\begin{aligned} &p(D)[(\exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1})h(x_n)] \\ &= [\exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1})]p(\xi_1, ..., \xi_{n-1}, -i\tfrac{d}{dx_n})h(x_n) \end{aligned}$$

and obtain the isomorphic representation of $\mathcal{M}_\eta$ by a mere separation of variables.

In this fashion we associated with the partial differential operator $p(D)$ a family of vector spaces parametrized by $\eta \in \mathbb{R}^{n-1}$ whose elements are solutions of ordinary differential equations with constant coefficients. A number of results about the structure of the spaces $\mathcal{M}_\eta$ follow painlessly by means of the most elementary stability considerations. From these, we will then develop the notion of ellipticity of boundary-value problems both in its analytic and its algebraic form.

THEOREM 10.1. *For a homogeneous polynomial p of degree k in n variables and for $\eta \in \mathbb{R}^{n-1}$, $\mathcal{M}_\eta$ consists of exponential solutions (i.e., coefficient functions h are polynomials). If p is elliptic (i.e., $\xi \in \mathbb{R}^n \setminus \{0\} \Rightarrow p(\xi) \neq 0$), then $\dim \mathcal{M}_\eta = k$, and $\mathcal{M}_\eta$ decomposes in a natural way into two subspaces and $\mathcal{M}_\eta^+$ and $\mathcal{M}_\eta^-$ which are complementary for $\eta \neq 0$.*

PROOF. It is well-known (see [**Po**, p.45 ff]) that the solutions of the differential equations $q(-i\frac{d}{dt})h(t) = 0$ (where q is a polynomial, in one variable, of degree k) are completely determined by the zeros of q. Namely, when $\lambda_1, ..., \lambda_m$ are the roots of q with respective multiplicities $r_1, ..., r_m$, then $\operatorname{Ker} q(-i\frac{d}{dt})$ is the linear span of the k exponential functions

$$t^p \exp i\lambda_\mu t,\ \mu = 1, ..., m \text{ and } p = 0, ..., r_\mu - 1.$$

For p elliptic, $q(t) = p(\xi_1, ..., \xi_{n-1}, t)$ remains degree k, and $\mathcal{M}_\eta$ is a k-dimensional subspace of the vector space of exponential solutions of the equation $p(D)u = 0$. The asymptotic behavior as $t \to \pm\infty$ (i.e., $x_n \to \pm\infty$) is determined by the imaginary parts of the zeros. If $\eta = (\xi_1, ..., \xi_{n-1}) \neq 0$, then $q(t)$ has no real zeros (at least for p elliptic), and $\mathcal{M}_\eta$ is the direct sum of $\mathcal{M}_\eta^+$ and $\mathcal{M}_\eta^-$, consisting of solutions which remain bounded (more precisely, go to zero) as $t \to +\infty$, resp. as $t \to -\infty$. □

REMARK 10.2. Conversely, the ellipticity of $p(D)$ can be defined by the condition

$$\mathcal{M}_\eta^+ \cap \mathcal{M}_\eta^- = \begin{cases} \{0\}, & \text{for } \eta \neq 0, \\ \{\text{constant functions}\}, & \text{for } \eta = 0. \end{cases}$$

This means that $p(D)$ possesses no bounded exponential solutions other than constant functions.

REMARK 10.3. In order to move on to boundary (or initial-value) problems, we consider the trajectories (*integral curves*) $h(t)$ of solutions together with their derivatives up to order $k-1$ in *phase space* $\mathbb{C}^k$. Each trajectory corresponds uniquely to its initial values $(h(0), Dh|_0, \ldots, D^{k-1}h|_0)$, where $D := -i\frac{d}{dt}$. The trajectories, corresponding to points in $\mathcal{M}_\eta^+$ $(\eta \neq 0)$ which tend to the origin, have initial values which lie in a subspace K_η^+ (the *plus-stable subspace*, see Figure 10.1) in phase space. The corresponding initial values for $\mathcal{M}_\eta^-$ lie in a subspace K_η^-

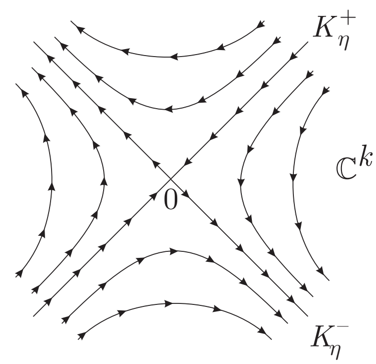


FIGURE 10.1. The plus- and minus-stable subspaces in phase space

(the *minus-stable subspace*) complementary to K_η^+. The remaining integral curves, whose initial values lie in the open set $\mathbb{C}^k - (K_\eta^+ \cup K_\eta^-)$, have a positive distance from the origin and are unbounded as $t \to \pm\infty$. The isomorphic images K_η^+ of $\mathcal{M}_\eta^+$ and K_η^- of $\mathcal{M}_\eta^-$ form the level set through a generalized saddle point at the origin of the phase space $\mathbb{C}^k$.[1] The necessary simple calculations may be found in [**NS**, p.187 f].

We come now to the definition of ellipticity of boundary-value problems in the half-space $x_n \geq 0$ of $\mathbb{R}^n$ for a single elliptic differential equation

$$p(D)u = 0 \tag{10.1}$$

with the boundary conditions

$$p_j(D)u\big|_{x_n=0} = 0,\ j = 1, ..., r. \tag{10.2}$$

Here p and $p_1, ..., p_r$ are homogeneous polynomials of degree k and $k_1, ..., k_r < k$, respectively.

DEFINITION 10.4. We call (10.1), (10.2) an **elliptic boundary-value problem**, if

$$k = 2r \tag{10.3}$$

and if

(10.4) the system of equations (10.1), (10.2) has no nontrivial solutions in $\mathcal{M}_\eta^+$ for each $\eta \neq 0$.

[1]Note that the spaces K_η^+ and $\mathcal{M}_\eta^+$ are canonically isomorphic (for given p and η); thus we can identify them with each other. Therefore, in the following, we will only speak of $\mathcal{M}_\eta^+$, and take $\mathcal{M}_\eta^+$ as a space of functions or initial values, depending on context.

Detailed motivation for this definition may be found in [**Hö63**, p.242-246]. According to this reference, the ellipticity of a boundary-value problem is important primarily for the proof of regularity theorems and finiteness of the index (see below). In addition, L. Hörmander proved that from a regularity assumption, saying that all solutions are still C^∞ on the boundary, it already follows that system (10.1), (10.2) can possess no bounded exponential solutions in the x_n-direction and that $k = 2r$. (The ellipticity of $p(D)$ is not needed here.) Elliptic boundary-value conditions, first formulated in this sense by Jaroslav Boresovich Lopatinsky and ???, are thus exactly those local conditions which insure the smoothness of solutions also on the boundary.

In order to isolate the algebraic kernel of the ellipticity conditions for the boundary-value system (10.1), (10.2), we again translate the conditions (10.3) and (10.4) into the language of ordinary differential equations and their initial conditions. Thus, let $p, p_1, ..., p_r$ be homogeneous polynomials as above, with p elliptic of degree k, and let $\eta \in \mathbb{R}^{n-1} \setminus \{0\}$. As before, we assign, to the pair (p, η), the vector space $\mathcal{M}_\eta^+$ which we can identify with the plus-stable subspace K_η^+. Now show:

Exercise 10.5. The data $p, p_1, ..., p_r$ of a boundary-value problem (p elliptic) defines a linear map $\beta_\eta^+ : \mathcal{M}_\eta^+ \to \mathbb{C}^r$, for each point $\eta \in \mathbb{R}^{n-1} \setminus \{0\}$.
Trick: Assign to each boundary operator $p_j(D)$, $j = 1, ..., r$, and each point $\eta = (\xi_1, ..., \xi_{n-1}) \in \mathbb{R}^{n-1} \setminus \{0\}$, the initial value problem

$$C^\infty(\mathbb{R}_+) \ni h \mapsto p_j(\xi_1, ..., \xi_{n-1}, -i\tfrac{d}{dt})h|_{t=0},$$

and thus a linear functional on the phase space $\mathbb{C}^k = \mathcal{M}_\eta$. The $p_1, ..., p_r$ together define (for each $\eta \neq 0$) a linear map β_η from $\mathbb{C}^k$ to $\mathbb{C}^r$; then consider the restriction to the plus-stable subspace $\mathcal{M}_\eta^+$.

Theorem 10.6. *A boundary-value system* (10.1), (10.2) *is elliptic, exactly when the map* $\beta_\eta^+ : \mathcal{M}_\eta^+ \to \mathbb{C}^r$ *is an isomorphism for each* $\eta \neq 0$. *Written out: For each* $\eta = (\xi_1, ..., \xi_{n-1}) \in \mathbb{R}^{n-1} \setminus \{0\}$ *and for each* r*-tuple of complex numbers* $(g_1, ..., g_r) \in \mathbb{C}^r$, *there is exactly one bounded (as* $t \to +\infty$*) function* $h \in \mathbb{C}^\infty(\mathbb{R})$ *such that* $p(\xi_1, ..., \xi_{n-1}, -i\frac{d}{dt})h = 0$ *and* $p_j(\xi_1, ..., \xi_{n-1}, -i\frac{d}{dt})h|_{t=0} = g_j$ *for* $j = 1, \ldots, r$.

Proof. First, suppose that (10.1), (10.2) is elliptic. The condition (10.4) says that $\mathcal{M}_\eta^+ \cap \{g \in \mathbb{C}^k : \beta_\eta(g) = 0\} = \{0\}$ for all $\eta \neq 0$; this means that β_η^+ is injective. To prove that β_η^+ is surjective, we use the condition (10.3). Recall (see Remark 10.2 after Theorem 10.1 above) that $\mathcal{M}_\eta^+$ and $\mathcal{M}_\eta^-$ together span the whole phase space $\mathbb{C}^k$ and are complementary to each other; in particular,

$$\dim \mathcal{M}_\eta^+ + \dim \mathcal{M}_\eta^- = k. \tag{10.5}$$

Since β_η^+ is injective, we have $\dim \mathcal{M}_\eta^+ \le r$. However, $\mathcal{M}_\eta^- \cong \mathcal{M}_{-\eta}^+$, since $u \in \operatorname{Ker} p(D) \Leftrightarrow \widetilde{u} \in \operatorname{Ker} p(D)$, where

$$u(x_1, ..., x_n) = (\exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1}))\, h(x_n) \text{ and}$$
$$\widetilde{u}(x_1, ..., x_n) = (\exp -i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1}))\, h(-x_n).$$

Thus, we also have $\dim \mathcal{M}_\eta^- \le r$, whence $\dim \mathcal{M}_\eta^+ = k - \dim \mathcal{M}_\eta^- \ge k - r = r$, since $k = 2r$ by (10.3). Thus, $\dim \mathcal{M}_\eta^+ = r$, and so β_η^+ is bijective.

Conversely, (10.4) follows easily from the bijectivity of β_η^+. Moreover, $\dim \mathcal{M}_\eta^+ = \dim \mathcal{M}_\eta^- = r$, and because of (10.5), we obtain condition (10.3). □

REMARK 10.7. Since the boundary conditions $p_1, ..., p_r$ are supposed to define a linear map from phase space $\mathbb{C}^k$ to $\mathbb{C}^r$, we had to first assume that the degrees of the polynomials $p_1, ..., p_r$ are smaller than the degree k of the elliptic polynomial p. Actually, we can algebraically reduce every polynomial of degree greater than p (e.g., see [**vdW**, Ch. 18]) and come to equivalent boundary (resp. initial) conditions.

REMARK 10.8. For an elliptic differential equation with constant coefficients (homogeneous of degree $2r$), we have first the natural (initial-value) isomorphism $\mathcal{M}_\eta \cong \mathbb{C}^{2r}$. Thus, elliptic boundary conditions cannot exist unless $\dim \mathcal{M}_\eta^+ = r$, i.e., $\mathcal{M}_\eta^+ \cong \mathbb{C}^r$, and this isomorphism is not canonically defined but depends on a choice of boundary conditions. In this sense, the algebraic-geometrical meaning of elliptic boundary-value problems lies in providing a fixed coordinate system (i.e., a fixed basis) for $\mathcal{M}_\eta^+$.

EXERCISE 10.9. a) For a boundary-value system of your choice, determine the families of vector spaces $\mathcal{M}_\eta$, $\mathcal{M}_\eta^+$, K_η^+ and the maps β_η^+ . For example show for the Laplace equation with $p(\xi_1, \xi_2) = \xi_1^2 + \xi_2^2$ that, at the point $\eta = \xi_1 > 0$,

$$\begin{aligned}\mathcal{M}_\eta &= \left\{c_1 e^{\eta(ix+y)} + c_2 e^{\eta(ix-y)} : c_1, c_2 \in \mathbb{C}\right\}\\ &\cong \left\{c_1 e^{\eta t} + c_2 e^{-\eta t} : c_1, c_2 \in \mathbb{C}\right\} \cong \mathbb{C}^2, \text{ and}\end{aligned}$$

$$\mathcal{M}_\eta^+ = \left\{c e^{\eta(ix-y)} : c \in \mathbb{C}\right\} \cong \left\{c e^{-\eta t} : c \in \mathbb{C}\right\} \cong \mathbb{C}.$$

Then determine the linear map β_η^+, e.g., for the Neumann boundary-value problem with $p_1(\xi_1, \xi_2) = \xi_2$, and show that it selects the basis vector $(0, -i/\eta)$ in $\mathcal{M}_\eta^+$.
b) Show that, for the Cauchy-Riemann equation with $p(\xi_1, \xi_2) := \xi_1 + i\xi_2$, the vector spaces $\mathcal{M}_\eta^+$ $(\eta \in \mathbb{R} \setminus \{0\})$ have different dimensions depending on whether $\eta > 0$ or $\eta < 0$.

2. Systems of Differential Equations with Constant Coefficients

In this section, p is a $N \times N$ matrix of homogeneous polynomials of degree k in n variables, and p is elliptic; i.e., $\det p(\xi) \neq 0$ for $\xi = (\xi_1, ..., \xi_n) \neq 0$.

EXERCISE 10.10. For $\eta = (\xi_1, ..., \xi_{n-1}) \in \mathbb{R}^{n-1} \setminus \{0\}$ define $\mathcal{M}_\eta$, $\mathcal{M}_\eta^+$, $\mathcal{M}_\eta^-$ as above.

THEOREM 10.11. *The following statements hold for* $\eta \neq 0$.

(*i*) $\dim \mathcal{M}_\eta = Nk$.
(*ii*) *Each solution is uniquely determined by its initial value in the phase space* $\mathbb{C}^{Nk}$.
(*iii*) *The vector spaces* $\mathcal{M}_\eta^+$ *and* $\mathcal{M}_\eta^-$, *interpreted as the plus-stable, resp. minus-stable subspaces, are complementary in* $\mathbb{C}^{Nk}$.

PROOF. For fixed $\eta = (\xi_1, ..., \xi_{n-1})$, the elements of $\mathcal{M}_\eta$ are of the form

$$u(x) = \begin{bmatrix} u_1(x) = (\exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1}))\, h_1(x_n) \\ \vdots \\ u_N(x) = (\exp i(x_1\xi_1 + ... + x_{n-1}\xi_{n-1}))\, h_N(x_n) \end{bmatrix}$$

where $x = (x_1, ..., x_n) \in \mathbb{R}^n$; or more simply of the form

$$h(t) = \begin{bmatrix} h_1(t) \\ \vdots \\ h_N(t) \end{bmatrix}$$

where $t \in \mathbb{R}$. For such C^∞ functions on $\mathbb{R}$ with values in $\mathbb{C}^N$, we write $h \in \mathbb{C}^\infty(\mathbb{C}^N_{\mathbb{R}})$; i.e., h is a C^∞ section of $\mathbb{C}^N_{\mathbb{R}}$, the N-dimensional trivial bundle over $\mathbb{R}$.

For the proof of (i), we go through the canonical transformation of the $N \times N$ system

$$p(\xi_1, ..., \xi_{n-1}, -i\tfrac{d}{dt})h = 0,\ h \in C^\infty(\mathbb{C}^N_{\mathbb{R}})$$

of order k to a $Nk \times Nk$ system

$$\frac{d\widetilde{h}}{dt} - A\widetilde{h} = 0,\ \widetilde{h} \in C^\infty(\mathbb{C}^{Nk}_{\mathbb{R}})$$

of order 1 (as in [**Po**, p.89f]). The solutions h can be written ([**Po**, p.97]) explicitly in the form $\overset{(\nu)}{q}(t) \exp i\lambda t$, where λ is an eigenvalue of A, $v \in \{1, ..., r\}$, r is the multiplicity of λ,

$$\overset{(\nu)}{q}(t) := \frac{t^{v-1}}{(v-1)!}e_1 + \frac{t^{v-2}}{(v-2)!}e_2 + \cdots + e_v,$$

and $e_1, e_2, ..., e_\nu$ is a system of vectors in $\mathbb{C}^{Nk}$ given through the Jordan normal form (see [**Po**, p.277-295]) so that

$$Ae_1 = \lambda e_1,\ Ae_2 = \lambda e_2 - ie_1, \dots,\ Ae_r = \lambda e_r - ie_{r-1}.$$

In this way, we establish (ii) and (iii) along with (i).

In [**Hö63**, p.269], we find a direct proof of (i) by directly working out the algebraic essence of the statement – without the conceptual, but perhaps somewhat diverting, discussion of solution curves: Set $P(t) := p(\xi_1, ..., \xi_{n-1}, t)$. One knows from the theory of elementary divisors (e.g., [**Alb**]) that invertible $N \times N$ matrices $A(t)$ and $B(t)$ exist, whose elements are polynomials as are the entries of the inverse matrices, such that $A(t)P(t)B(t) = Q(t)$, where $Q(t)$ is a diagonal matrix. For $D := -i\frac{d}{dt}$, the operators $A(D)$ and $B(D)$ are bijective on $C^\infty(\mathbb{C}^N_{\mathbb{R}})$, and we have $P(D) = A^{-1}(D)Q(D)B^{-1}(D)$. Thus,

$$\begin{aligned} \mathcal{M}_\eta &\cong \left\{h \in C^\infty(C^N_{\mathbb{R}}) : P(D)h = 0\right\} \\ &= \left\{B(D)g : g \in C^\infty(C^N_{\mathbb{R}}) \text{ and } A^{-1}(D)Q(D)g = 0\right\} \\ &= \{B(D)g : Q(D)g = 0\} \cong \left\{g \in C^\infty(C^N_{\mathbb{R}}) : Q(D)g = 0\right\}. \end{aligned}$$

Because of the polynomial form of $A(t)$ *and* $A(t)^{-1}$, $\det A(t)$ does not depend on t. The same holds for $\det B(t)$. Thus, there is a constant $C \neq 0$ such that $\det Q(t) = C \det P(t)$. The sum of the degrees of diagonal elements of $Q(t)$ (and hence, $\dim \mathcal{M}_\eta$) is the same as the degree of $\det P(t)$, namely Nk if p is elliptic of degree k.

Note that the transformation $B(D)$ does not alter the asymptotic behavior of g. Thus, the direct decomposition of $\mathcal{M}_\eta$ into $\mathcal{M}^+_\eta$ and $\mathcal{M}^-_\eta$ is directly given via $Q(D)$ (or through $\det P(t)$ and its zeros with positive imaginary part). □

For the differential operator $p(D_1, ..., D_n)$, a system of boundary conditions is given by $N \times M_j$ matrices p_j $(j = 1, ..., r)$, whose elements are homogeneous polynomials of degree k_j. The *boundary-value problem* is written exactly as in (10.1), (10.2) above. For *ellipticity*, the condition (10.3) must be correspondingly generalized to $Nk = 2\sum_{j=1}^{r} M_j$, while condition (10.4) needs no further modification.

EXERCISE 10.12. As above, define the initial-value map $\beta_\eta : \mathbb{C}^{Nk} \to \bigoplus_{j=1}^{r} \mathbb{C}^{M_j}$ and show that the boundary-value system is elliptic exactly when the restriction $\beta_\eta^+ : \mathcal{M}_\eta^+ \to \bigoplus_{j=1}^{r} \mathbb{C}^{M_j}$ is an isomorphism for all $\eta \neq 0$.

3. Variable Coefficients

The preceding developments easily carry over to the case where X is a C^∞ n-manifold (with boundary Y) equipped with a Riemannian metric, E and F are N-dimensional C^∞ complex vector bundles over X, and $P : C^\infty(E) \to C^\infty(F)$ is an elliptic differential operator of order k. Here, the geometrical character of our definitions (their invariance under coordinate changes) is manifest, so that we may pass to global concepts. For this, first show:

EXERCISE 10.13. For each $y \in Y$, the symbol $\sigma(P)(y, \cdot)$ of P defines a homogeneous partial differential operator with constant coefficients $P : \mathcal{E}_y \otimes E_y \to \mathcal{E}_y \otimes F_y$, where $\mathcal{E}_y$ is the space (isomorphic to $C^\infty(\mathbb{R}^n)$) of C^∞ functions on the tangent space T_yX; thus $\mathcal{E}_y \otimes E_y$ is the space of C^∞ functions on T_yX with values in E_y (isomorphic to the N-fold product $C^\infty(\mathbb{R}^n) \times \cdots \times C^\infty(\mathbb{R}^n)$).
[Warning: Here it does not matter whether P is actually elliptic or not, and this somewhat contrived construction of P applies at each $y \in X$, not only at boundary points.]

As usual, let $T'Y$ be the bundle of nonvanishing covectors for Y; i.e., T^*Y without the zero section. We obtain, as the first main result of the heuristic considerations of this paragraph, the following

THEOREM 10.14. *Each elliptic differential operator P over the manifold X with boundary Y defines C^∞ vector bundles $\mathcal{M}$, $\mathcal{M}^+$, and $\mathcal{M}^-$ (with fiber dimension not necessarily constant; see Exercise 10.9b) over $T'Y$.*

PROOF. By selecting an inner normal $\nu \in T_y^*X$ by means of the Riemannian metric, we obtain a half-space problem, which defines the spaces $\mathcal{M}_\eta$, $\mathcal{M}_\eta^+$, and $\mathcal{M}_\eta^-$ as above, where now $\eta \in T_y^*Y$, $\eta \neq 0$. Namely, set

$$\mathcal{M}_\eta := \{f \in \operatorname{Ker} P_y : f = \exp(\langle \cdot, \eta \rangle)\, h, \text{ where } h \in C^\infty(T_yX/T_yY, E_y)\}$$

(i.e., h is a C^∞ function on T_yX/T_yY with values in E_y), and correspondingly define, with the appropriate boundedness conditions, the spaces $\mathcal{M}_\eta^+$, and $\mathcal{M}_\eta^-$, where T_yX/T_yY is identified with $\mathbb{R}$ in an oriented way by means of the Riemannian metric so that we know what it means to go to $+\infty$ or $-\infty$. Passing to the equivalent system of ordinary differential equations, we obtain the polynomial $\sigma(P)(y, \eta + tv) = \sum_{\kappa=0}^{k} c_\kappa t^\kappa$ of degree k with coefficients in $\operatorname{Hom}(E_y, F_y)$. In a local coordinate system $u = (u_1, ..., u_n)$ about y (which respects the Riemannian metric at y and maps the neighborhood of the boundary Y near y to $\mathbb{R}^{n-1}$, the linear ordinary differential operator $\sigma(P)(y, \eta + Dv)$ takes the familiar form $p(\xi_1, ..., \xi_{n-1}, -i\frac{d}{dt})$, where $\eta = \xi_1 du_1 + ... + \xi_{n-1} du_{n-1}$ and $v = du_n$. In this way, the families $\mathcal{M}$, $\mathcal{M}^+$, and $\mathcal{M}^-$ of vector spaces are defined and parametrized by $T'Y$. Considerations

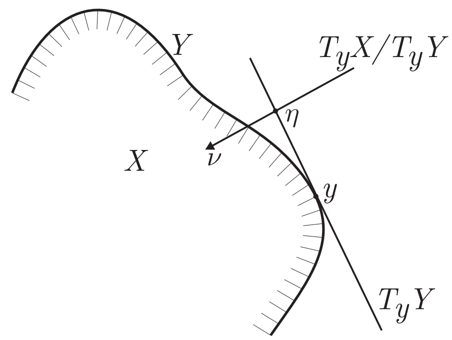


FIGURE 10.2. xxx1

in local coordinates show that $\mathcal{M}$, $\mathcal{M}^+$, and $\mathcal{M}^-$ are actually C^∞ vector bundles, since the respective solutions (resp., their initial values in phase space) depend smoothly on the coefficients $c_0, ..., c_k$, or more generally, on (y, η): This follows easily from the explicit form of the solutions given in the proof of Theorem 10.11 (see also the alternative proof there). □

Now let $G = \bigoplus_{j=1}^r G_j$ be a vector bundle on the boundary Y of X where $G_j \in \mathrm{Vect}(Y)$ has fiber dimension $M_j \in \mathbb{N}$. Let $R = (R_1, ..., R_r)$, for differential boundary operators $R_j : C^\infty(E) \to C^\infty(G)$ of order l_j. Our second main result introduces a further global invariant. For this, let $\pi_Y : T'Y \to Y$ be the natural projection (base-point map).

THEOREM 10.15. *For $P : C^\infty(E) \to C^\infty(F)$ elliptic, each boundary-value system*

$$(P, R) : C^\infty(E) \to C^\infty(F) \oplus C^\infty(G)$$

defines, in a natural way, a vector bundle homomorphism $\beta^+ : \mathcal{M}^+ \to \pi_Y^*(G)$.

PROOF. $\mathcal{M}^+$ is the bundle defined from $\sigma(P)$ in Theorem 10.14, and β^+ is the initial-value map given, at the point (y, η), by the composition

$$\beta_\eta^+ : \mathcal{M}_\eta^+ \overset{R_y|_{\mathcal{M}_\eta^+}}{\longrightarrow} \mathcal{E}_y \oplus G_y \overset{\delta}{\to} G_y.$$

Here, $R_y : \mathcal{E}_y \oplus G_y \to \mathcal{E}_y \oplus G_y$ is the differential operator with constant coefficients defined via the symbol $\sigma(R) = (\sigma(R_1), ..., \sigma(R_r))$ taken at the point y as in Exercise 10.13, and δ is the map defined by $\delta(f) := f(0)$. □

The boundary-value system (P, R) is defined to be **elliptic** when β^+ is an isomorphism.

EXERCISE 10.16. By introducing a coordinate system, show that this definition of ellipticity corresponds to the local ellipticity defined above (see Exercise 10.12 and before).

4. The Topology of Local Boundary Value Conditions (Case Study)

In Chapter 9 we saw that a topological object *$\sigma(P) \in \mathrm{Iso}_{SX}(E, F)$ is associated with each elliptic operator $P : \overline{C^\infty(E) \to C^\infty(F)}$, whereby* index P *depends only on the homotopy type of $\sigma(P)$. Here SX is the covariant sphere bundle for a*

Riemannian metric for X, and E and F are Hermitian vector bundles over the closed manifold X.

Now we consider, on a manifold X with boundary the local elliptic boundary-value system $(P,R) : C^\infty(E) \to C^\infty(F) \oplus C^\infty(G)$, where E, F are vector bundles over X, and G is a vector bundle over the boundary Y of X. The object $\sigma(P) \in \mathrm{Iso}_{SX}(E,F)$ is well defined, but it does not contain the necessary information on the index of the boundary value problem (P,R) which may depend on the specific choice of the boundary conditions R, by the Theorem of I. N. Vekua (Theorem 5.11, p.146). In the case of a differential operator P of first order (but see Exercise 10.20) we will show, roughly, how the given boundary conditions canonically determine a continuation of $\sigma(P)$ beyond SX to the closed manifold $SX \cup BX|_Y$. We obtain a topologically more significant object (conceptually: a closed line packs more topological information than an open one) whose homotopy type does in fact determine $\mathrm{index}(P,R)$, *as we will see in Section 14.8. Contrary to the technical explanations of the algebraic meaning of ellipticity of boundary-value problems in the preceding sections, we are concerned in the following case study with the geometric-topological interpretation.*

THEOREM 10.17. *An elliptic system (P,R) of N partial differential equations of first-order over the compact, oriented, Riemannian manifold X with $N/2$ boundary conditions over the boundary Y of X defines (uniquely, up to homotopy) a continuous map of the closed manifold $SX \cup (BX|_Y)$ into $\mathrm{GL}(2N,\mathbb{C})$ which coincides with $\sigma(P) \oplus \mathrm{Id}_N$ on SX. Here, $\mathrm{GL}(2N,\mathbb{C})$ denotes the group of complex, invertible $2N \times 2N$ matrices and $BX := \{\xi \in T^*X : |\xi| \le 1\}$.*

PROOF. (a verbal communication of I. M. Singer; see also [**AB64a**, p.180-184], [**CaSc**, 20/05-25/07] and [**Pal65**, p.346-350]): Let

$$(P,R) : C^\infty(E) \to C^\infty(F) \oplus C^\infty(G)$$

be an elliptic boundary-value system with $P \in \mathrm{Diff}_1(E,F)$ and $E = F = \mathbb{C}_X^N$ and $G = \mathbb{C}_Y^{N/2}$, N even. Let $\nu \in T_y^*X$ be the inner normal at the point $y \in Y$. Each covector $\xi \in T_y^*X$ can be written in the form $z\nu + \eta$ with $z \in \mathbb{R}$ and $\eta \in T_y^*Y$, where T_y^*Y can be taken to be a proper subspace of T_y^*X by means of the Riemannian metric (see Exercise 6.45, p. 188). We write $\sigma(\xi) := \sigma_1(P)(y,\xi)$ and obtain (since $\sigma(\xi)$ is a homogeneous polynomial in $\xi = (z\nu,\eta)$) :

$$\sigma(z\nu + \eta) = \sigma(z\nu) + \sigma(\eta) = z\sigma(\nu) + \sigma(\eta) : E_y \to F_y \text{ (linear).}$$

By a corresponding choice of basis for F_y, we may assume (without loss of generality) that $\sigma(\nu) = \mathrm{Id}$, whence

$$\sigma(z\nu + \eta) = z\,\mathrm{Id} + \sigma(\eta). \tag{10.6}$$

Consider the space $\mathcal{M}_\eta^+$, defined as in Theorem 10.14 (p. 246), consisting of the C^∞ functions $h : \mathbb{R} \to E$ with $\sigma(D\nu + \eta)h = Dh + \sigma(\eta)h = 0$ which remain bounded as $t \to +\infty$; here $D := -i\frac{d}{dt}$. $\mathcal{M}_\eta^+$ is spanned by the functions of the form $h(t) = h_0 e^{i\lambda t}$, where λ is an eigenvalue of the endomorphism $\sigma(\eta) : E_y \to E_y$ with $\mathrm{Im}\,\lambda > 0$ and $h_0 \in E_y$ is an element of the associated eigenspace. Corresponding remarks hold for $\mathcal{M}_\eta^-$. Thus, the spaces $\mathcal{M}_\eta^\pm$ are naturally isomorphic to the sum of the $(+)$ (resp. $(-)$) eigenspaces of $\sigma(\eta)$. For $\eta \neq 0$, the homomorphism $\lambda\,\mathrm{Id} + \sigma(\eta) = \sigma(\lambda\nu + \eta)$ is regular for $\lambda \in \mathbb{R}$ by the ellipticity of P; i.e., $\sigma(\eta)$ has no real eigenvalues and E_y can be represented as the direct sum

$$E_y \cong \mathcal{M}_\eta^+ \oplus \mathcal{M}_\eta^- . \tag{10.7}$$

Now let $G_1, ..., G_r$ be vector bundles over Y with $\bigoplus G_j = G$ and

$$R_j : C^\infty(E) \to C^\infty(G), j = 1, ..., r,$$

be boundary conditions given by differential expressions such that associated initial value map

$$\beta_\eta^+ : \mathcal{M}_\eta^+ \to G_y \tag{10.8}$$

is an isomorphism for all $y \in Y$ and $\eta \in T_y^*Y \setminus \{0\}$. By Theorem 10.15 (p. 247) and Exercise 10.16, this is the condition of ellipticity for boundary-value systems. We now show that $\sigma(p)(y, \cdot) : (SX)_y \to \mathrm{Iso}(E_y, E_y)$ is (stably) homotopic to a constant map, and that the homotopy is defined in a natural way by using $\left\{\beta_\eta^+ : \eta \in T_y^*Y \setminus \{0\}\right\}$. Thus, $\sigma(P)$ (more precisely $\sigma(P) \oplus \mathrm{Id}_{\mathbb{C}^N}$, see the Homo-

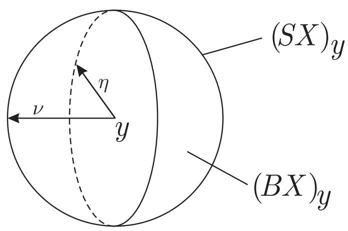


FIGURE 10.3. xxx2

topy 2 below) can be extended to $(BX)|_Y$.

Homotopy 1: For $\eta \in T_y^*Y \setminus \{0\}$, let $\pi_\eta^\pm : E_y \to \mathcal{M}_\eta^\pm \subset E_y$ be the projections defined by (10.7) . By means of

$$s\sigma(\eta) + (1 - s)(i\pi_\eta^+ - i\pi_\eta^-), \ s \in I, \tag{10.9}$$

we obtain a homotopy of $\sigma(\eta)$ to the map $i\pi_\eta^+ - i\pi_\eta^-$; the geometric meaning of this is that one may concentrate the eigenvalues of $\sigma(\eta)$ on the eigenvalues $+i$ and

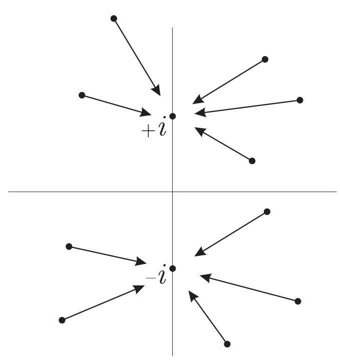


FIGURE 10.4. xxx3

$-i$ which are independent of η, while the eigenspaces still depend on η. By using $h_0 = h_0^+ + h_0^- \in E_y$ with $h_0^\pm \in \mathcal{M}_\eta^\pm$ one calculates that the eigenvalues of the

endomorphism defined in (10.9) always remain non-real; thus, $z\,\mathrm{Id} +$ (10.9) is non-singular for $z \in \mathbb{R}$. Hence, we have a homotopy in the space of elliptic symbols (i.e., in $(SX)|_Y \to \mathrm{GL}(N,\mathbb{C})$ here) from σ to σ_1 with $\sigma_1\,(z\nu+\eta) := z\,\mathrm{Id} + i\pi_\eta^+ - i\pi_\eta^-$.
Homotopy 2: By means of $e^{-i\varphi} i\pi_\eta^+ - e^{i\varphi} i\pi_\eta^-$, $\varphi \in [0,\pi/2]$, each $\sigma_1(\eta)$ in $\mathrm{Iso}(E_y,E_y)$ can be connected with the identity, but this deformation depends on the choice of η and does not go through uniformly for all $\eta \in T_y^*Y \setminus \{0\}$. $\mathrm{GL}(N,\mathbb{C})$ is too small to implement the further homotopy. Hence, we enlarge σ_1 by direct sum with $\mathrm{Id}_G \oplus \mathrm{Id}_G$ to a map $(SX)|_Y \to \mathrm{Iso}(E\oplus G\oplus G, E\oplus G\oplus G)$, which we also denote by σ_1. Because of the splitting $E_y \cong \mathcal{M}_\eta^- \oplus \mathcal{M}_\eta^+$, $\eta \in T_y^*Y\setminus\{0\}$, we have

$$\sigma_1\,(\eta) = \begin{bmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where the diagonal elements mean respectively the identities on $\mathcal{M}_\eta^+$ or $\mathcal{M}_\eta^-$ or G multiplied by the coefficients $-i$ or $+i$ or I. By the deformation $e^{-i\frac{1}{2}\pi s}\,\mathrm{Id}_G \oplus e^{i\frac{1}{2}\pi s}\,\mathrm{Id}_G$, $s \in [0,1]$, we can uniformly deform $\sigma_1\,(\eta)$ to

$$\sigma_2\,(\eta) = \begin{bmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \text{ on } \mathcal{M}_\eta^- \oplus \mathcal{M}_\eta^+ \oplus G_y \oplus G_y.$$

Homotopy 3: Now it remains to deform $\sigma_2(\eta)$ to a constant ($\sigma_2(\eta)$ still depends on the positions of the eigenspaces $\mathcal{M}_\eta^\pm$), in such a way that no real eigenvalues appear and $\sigma_2(z\nu+\eta) := z\,\mathrm{Id} + \sigma_2(\eta)$ does not become singular for $z \in \mathbb{R}$ under the deformation. This we achieve with the help of the boundary isomorphism $\beta_\eta^+ : \mathcal{M}_\eta^- \cong G_y$ given by the ellipticity of the boundary-value problem in (10.8). Indeed, there is a homotopy (which switches the second and third diagonal members in $\sigma_2\,(\eta)$) of $\sigma_2\,(\eta)$ to a constant map (on $(SY)_y$)

$$\eta \mapsto \sigma_3\,(\eta) := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \left(\beta_\eta^+\right)^{-1} & 0 \\ 0 & \beta_\eta^+ & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \circ \sigma_2\,(\eta) = \begin{bmatrix} -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \end{bmatrix}$$

which is multiplication by $-i$ on E_y and by i on $G_y \oplus G_y$; the homotopy

$$\begin{bmatrix} 0 & \left(\beta_\eta^+\right)^{-1} \\ \beta_\eta^+ & 0 \end{bmatrix} \sim \begin{bmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{bmatrix}$$

follows from the Homotopy Lemma which says that

$$\begin{bmatrix} 0 & B \\ A & 0 \end{bmatrix} \sim \begin{bmatrix} \mathrm{Id} & 0 \\ 0 & AB \end{bmatrix}$$

in the space of automorphisms of the vector space $V \times W$, if V and W are complex vector spaces and $A : V \to W$ and $B : W \to V$ are linear with $AB \in \mathrm{Iso}(W,W)$. One proves the Homotopy Lemma by composing two homotopies: First connect

$$\begin{bmatrix} 0 & B \\ A & 0 \end{bmatrix} \text{ and } \begin{bmatrix} i\,\mathrm{Id} & 0 \\ 0 & iAB \end{bmatrix} \text{ via } \begin{bmatrix} (i\sin\varphi)\,\mathrm{Id} & (\cos\varphi)\,B \\ (\cos\varphi)\,A & (i\sin\varphi)\,AB \end{bmatrix}$$

$\varphi \in [0, \pi/2]$, and then multiplication by $e^{-i\psi}$, $\psi \in [0, \pi/2]$, provides the final homotopy.

Homotopy 4: We have now deformed $\sigma \oplus \mathrm{Id}$ to the constant map σ_3 on $(SY)_y$; we extend this map to a map on all of $(SX)_y$ which is homotopic to $\sigma_1 \oplus \mathrm{Id}$, by means of the parametrization

$$(SX)_y = \{(\cos\theta)\,\nu + (\sin\theta)\,\eta : \eta \in (SY)_y \text{ and } \theta \in [0, \pi]\}\,.$$

Namely, define on $E_y \oplus (G_y \oplus G_y)$ the automorphism

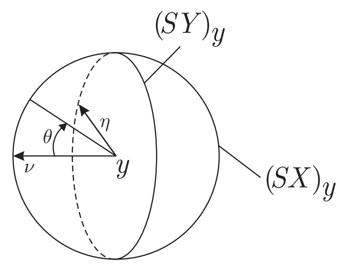


FIGURE 10.5. xxx4

$$\sigma_3\left((\cos\theta)\,\nu + (\sin\theta)\,\eta\right) := \begin{bmatrix} \cos\theta - i\sin\theta & 0 \\ 0 & \cos\theta + i\sin\theta \end{bmatrix}$$

which by definition is homotopic to the constant map $\mathrm{Id} \oplus \mathrm{Id}$. □

5. Generalizations (Heuristic)

REMARK 10.18. In the preceding proof, we have explicitly shown with a sequence of homotopies how one can continuously extend the map $\sigma \oplus \mathrm{Id} : SX \to \mathrm{GL}(2N, \mathbb{C})$ to a map $\widetilde{\sigma} : SX \cup (BX)|_Y \to \mathrm{GL}(2N, \mathbb{C})$ with $\widetilde{\sigma}(y, 0) = \mathrm{Id}_{2N}$ for all $y \in Y :=$boundary of X. In that proof, we have chosen formulations that make the generalization for arbitrary bundles clear. To be sure, we must make precise what we understand by *stable homotopy*; i.e., why we are content with an extension of $\sigma \oplus \mathrm{Id}_{2G} : (SX)|_Y \to \mathrm{Iso}\,(E \oplus G \oplus G, F \oplus G \oplus G)$ on $(BX)|_Y$, even though this object does not immediately extend over all of SX, since G is defined only over Y. See also under Section 11.3. To bring about further generalizations of the proof, we remark that the special form of the boundary conditions (which were given by differential operators) played no role, since only the isomorphism β^+_η was needed; β^+_η possibly could be defined through pseudo-differential boundary conditions. The definition of the spaces $\mathcal{M}^\pm_\eta$ and the construction of the symbol homotopies are made very easy by the polynomial form of $\sigma(P)$, i.e., its derivation from a differential operator. This is why we have devoted so much space to this case study. The construction of an extension of $\sigma(P) \oplus \mathrm{Id}_N$ as an isomorphism over $(BX)|_Y$ goes through more generally for elliptic pseudo-differential operators with *transmission properties* which allow elliptic boundary problems; see Chapter 10 and Section 14.8.

REMARK 10.19. To each elliptic operator $P \in \mathrm{L}^k_{\mathrm{pc}}(E, F)$ over the Riemannian n-manifold X, one can assign a **local index**, namely the homotopy class of

$$\begin{array}{cccc} \sigma(P)\,(x,\cdot): & (SX)_x & \to & \mathrm{Iso}\,(E_x, F_x) \\ & \updownarrow & & \updownarrow \\ & S^{n-1} & \to & \mathrm{GL}\,(N, \mathbb{C}) \end{array}$$

where N is the fiber dimension of E. By the Bott Periodicity Theorem, whereby (see Exercise 11.18, p. 282) for N sufficiently large (which we achieve here by adding the identity)

$$\pi_{n-1}(\mathrm{GL}(N,\mathbb{C})) = \begin{cases} \mathbb{Z}, & \text{for } n \text{ even,} \\ 0, & \text{for } n \text{ odd,} \end{cases}$$

we obtain an integer $\deg(P)$ for the local index; by continuity, it is independent of the choice of x. If X is a manifold with boundary Y, then (as in Chapter 10) we can determine the vector spaces $\mathcal{M}_\eta^\pm$ and the integer $\mu(P) := \dim \mathcal{M}_\eta^+ - \dim \mathcal{M}_\eta^-$, which is independent of the choice of $\eta \in T'Y$ and does not automatically vanish for $n = 2$.

If the operator P admits elliptic boundary conditions, then the condition $\mu(P) = 0$ follows from Theorem 10.6 and Theorem 10.15, and by the preceding Theorem 10.17, the condition $\deg(P) = 0$ holds. These two conditions are closely connected. By a communication from M. F. Atiyah, in the special case $N = 1, n = 2$ (where $\deg(P)$ is the classical winding number of $\sigma(P)(x,\cdot)$ in $\mathbb{C} \setminus \{0\}$ about the point 0), the equation

$$\deg(P) = \pm\mu(P) \tag{10.10}$$

holds. Also, in the general case, we have

$$\deg(P) = \pm \deg \mathcal{M}^+ \tag{10.11}$$

where $\deg \mathcal{M}^+$ is the integer degree of a map $f_y : S^{n-3} \to \mathrm{GL}(M,\mathbb{C})$, $y \in Y$, which is used to join together the trivial bundles over the upper and lower hemispheres of S^{n-2} along the equator S^{n-3} (see Appendix B, Exercise B.7, 681) to obtain the bundle $\mathcal{M}_y^+ := \left\{\mathcal{M}_\eta^+ : \eta \in T_y^*Y \setminus \{0\}\right\}$ over $(SY)_y \cong S^{n-2}$. Again, by continuity arguments, it is clear that degree $\deg \mathcal{M}^+$ does not depend on the choice of y in f_y. Equations (10.10) and (10.11) then represent a reformulation of the Bott Periodicity Theorem. For this, see also [**AB64a**, p.178], [**CaSc**, 25-05], and in particular [**Pal65**, p.351], where a K-theoretic formulation of (10.10) and (10.11) is given; see also the Hint for Exercise 10.20.

As a result of Theorem 10.17, we have obtained (in $\deg(P) \neq 0$) a topological obstruction to there being elliptic boundary conditions for the elliptic differential operator P. In the classical theory $E = F = \mathbb{C}_X$, the obstruction can arise only in the case $n = 2$, since for $n \geq 3$ every homogeneous elliptic polynomial is of even degree $2k$ and possesses an equal number k of zeros in the upper and lower half-planes [**Hö63**, p.246]. The situation is different for $n = 2$, where $\deg(\frac{\partial}{\partial \bar z}) = 1$. However, for all even $n \geq 4$ there are elliptic differential operators P with $\deg(P) \neq 0$; e.g., the Dirac operators (see [**Pal65**, p.91f] or [**Bo72**, p.26f]), defined using the Clifford module $\mathbb{R}^{2^{m-1}}$ over $X = \mathbb{R}^{2m}$, have local index 1.

The example is comparable to the peculiarity shown by the Cauchy-Riemann operator $\frac{\partial}{\partial \bar z}$ in the classical theory for $n = 2$. Actually, there are many elliptic operators which arise in Riemannian geometry, that have a nonvanishing local index, and, for closed manifolds, characterize important topological invariants such as the *Euler number* and *signature* by their global indices; see [**Bo72**, p.30] and [**APS75**, p.46]. In the last work, as an expedient for the calculation involving the corresponding topological invariants of manifolds with boundary, a nonlocal theory of boundary-value problems is applied, with which one obtains a Fredholm theory in which the above *obstructions* become irrelevant.

EXERCISE 10.20. For each elliptic system

$$(P, R) : C^\infty(E) \to C^\infty(E) \oplus C^\infty(G)$$

of partial differential equations of order $k > 1$ with boundary conditions on the bounded domain X of $\mathbb{R}$, where E, F, G are trivial bundles, construct an *equivalent* (in which sense?) elliptic system

$$(\widetilde{P}, \widetilde{R}) : C^\infty(\widetilde{E}) \to C^\infty(\widetilde{E}) \oplus C^\infty(\widetilde{G})$$

of differential equations of order 1 with boundary conditions.
[Hint: For the exact formulations, see [**ADN**]. The operator $\widetilde{P}$ which we define here is only elliptic in an extended sense, which one calls Douglis-Nirenberg elliptic. For reduction to the classical case (i.e., by splitting $\widetilde{E}$ and applying Λ-type operators componentwise), see also [**Hö66a**, p.134-136]. Here, we only give the rough idea (note also the Remark 10.22 below). Let $P = \sum_{|\alpha|\le k} a_\alpha D^\alpha$, where $a_\alpha \in C^\infty(\mathrm{Hom}(E, F))$ (i.e., $a_\alpha(x)$ is a linear map from $\mathbb{C}^N$ to $\mathbb{C}^N$, if $E = F = \mathbb{C}^N_X$ and $x \in X$. Then begin with the *order lowering* procedure for ordinary differential equations of higher order, constructing a differential operator $\widetilde{P} \in \mathrm{Diff}_1(\widetilde{E}, \widetilde{E})$, where we set $\widetilde{E} = \widetilde{F} = \sum_{|\alpha|\le k} E \times \{\alpha\}$ and $\widetilde{P}$ is given by the following matrix. (We consider only the case $N = 1$, $n = k = 2$, whence $P = a_{00} + a_{10}D_1 + a_{01}D_2 + a_{20}D_1^2 + a_{11}D_1D_2 + a_{02}D_2^2$, where $D_j = -i\partial/\partial x_j$.)

$$\begin{bmatrix} a_{00} & a_{10} & a_{01} & a_{20} & a_{11} & a_{02} \\ -D_1 & 1 & 0 & 0 & 0 & 0 \\ -D_2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -D_1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -D_1 & 0 & 1 & 0 \\ 0 & 0 & -D_2 & 0 & 0 & 1 \end{bmatrix}$$

Write the matrix in greater generality! How does the corresponding $\widetilde{R}$ appear, if $R = (R_1, ..., R_r)$ are differential boundary operators with $R_j : C^\infty(E) \to C^\infty(G_j)$ of order k_j, where $G_j = \partial X \times \mathbb{C}^{N_j}$ and $G = \bigoplus_{j=1}^r G_j$. Compare $\mathrm{Ker}(P, R)$ and $\mathrm{Ker}(\widetilde{P}, \widetilde{R})$ and correspondingly, $\mathrm{Coker}(P, R)$ and $\mathrm{Coker}(\widetilde{P}, \widetilde{R})$; show $\mathrm{index}(P, R) = \mathrm{index}(\widetilde{P}, \widetilde{R})$.]

EXERCISE 10.21. Consider the disk $X := \{z \in \mathbb{C} : |z| \le 1\}$ with the circle $Y := \{z \in \mathbb{C} : |z| = 1\}$ as boundary.
a) Go through the construction of Theorem 10.17 for the *transmission operator* (see Exercise 5.16)

$$T(u, v) := \left(\frac{\partial u}{\partial \bar{z}}, \frac{\partial v}{\partial z}, (u - v)\,|_Y\right)$$

b) Change the Laplace operator Δ on X as in Exercise 10.20 to a system $\widetilde{\Delta}$ of partial differential operators of order 1, and check that $\tilde{\Delta}$ is elliptic. What are the new boundary conditions that are obtained for the Dirichlet boundary-value problem $u|_Y = 0$? For $y \in Y$ and $\eta \in (SY)_y = \{\pm 1\}$, determine the spaces $\widetilde{\mathcal{M}}^\pm_\eta$ and the isomorphism $\widetilde{\beta}^+_\eta$. As in Theorem 10.17, extend $\sigma(\widetilde{\Delta})$ by means of $\widetilde{\beta}^+$ from $(SX)|_Y = S^1 \times S^1$ to a map $\sigma(\widetilde{\Delta}, \widetilde{\beta}^+)$ on the solid torus $(BX)|_Y$.
c) Go through the corresponding construction for the Neumann boundary-value problem $\frac{\partial u}{\partial \nu}|_Y = 0$, and compare $\sigma(\widetilde{\Delta}, \widetilde{\beta}^+)$ with $\sigma(\widetilde{\Delta}, \widetilde{\gamma}^+)$, where $\widetilde{\gamma}^+_\eta$ denotes the

boundary isomorphism obtained from the Neumann boundary-value problem relative to $\widetilde{\Delta}$.
d) Does the map $\sigma(\Delta)$, which is the constant $-i$ on SX, have different extensions to $SX \cup (BX)|_Y$? Can one apply the boundary isomorphisms β_η^+ or γ_η^+ here, and relate them to the results of a), b), and c)?

Remark 10.22. For the practical calculation of the index of an elliptic boundary-value system, the explicit determination of $\widetilde{P}$ is often lengthy and in fact superfluous, since the index is completely determined by the symbol and boundary isomorphisms; see Theorem 10.32 below. In particular, it suffices to reduce elliptic boundary-value systems of order k to elliptic systems of order 1 on the *symbol level*, just as in the previous Exercise where $\sigma(P) \oplus \mathrm{Id}_{\widetilde{E}}$ is homotopic in $\mathrm{Iso}_{SX}(E \oplus \widetilde{E}, F \oplus \widetilde{E})$ to $\sigma(\widetilde{P})$. This deformation procedure on the *symbol level* was introduced in [**AB64a**, p.180f], and is entirely analogous to the *linearization of polynomial clutching functions*, which plays a decisive role in one of the proofs of the Bott Periodicity Theorem (see [**AB64b**, p.241f]). It has the merit of also carrying over to the case of arbitrary manifolds with boundary and nontrivial bundles E, F, and G. This requires some standard tricks: One must identify the bundles E and F in a neighborhood of the boundary Y of X, by means of the isomorphism $\sigma(P)(y,\nu) : E_y \to F_y$, where ν is the inner normal at the boundary point y, etc.. In such general cases it is possibly advantageous to go yet a step higher: One does not specify a homotopy-theoretic extension of $\sigma(P)$ and Iso_{SX} to an isomorphism over $SX \cup BX|_Y$ by means of the boundary isomorphism β, but rather one directly constructs, from $\sigma(P)$ and β, a difference vector bundle in $K(BX, SX \cup BX|_Y)$, as sketched in [**Pal65**, p.346-351].

Beside these two ways of finding the *correct* topological object while preserving the information about the index of an elliptic boundary-value problem, viz.

- linearization and continuation of the symbol by means of homotopies, and
- K-theoretic axiomatic characterization of a difference vector bundle,

there is a third way oriented more strongly towards functional analysis, viz.

- assignment of families of Wiener-Hopf operators to elliptic boundary-value problems.

All three approaches depend not only on the Bott Periodicity Theorem (see Chapter 11), but on essential elements of the different proofs. Conversely, the Periodicity Theorem be derived via an investigation of a simple boundary-value problem on the disk (see Exercise 11.19, p. 282).

In Theorem 4.4, we already provided the Index Theorem for Wiener-Hopf operators. We will prove the Periodicity Theorem, following [**Ati69**, p.116-120], as a generalization of this theorem. Hence, the third approach via Hilbert space theory is most convenient for us, and offers some formal advantages for the demonstration of the analytical main theorems on the index of elliptic boundary-value problems. Admittedly, this approach may be less conceptual than the first, and less elegant than the second.

6. The Poisson Principle

Before coming to the definition and treatment of a class of boundary-value systems, chosen very large for our purposes, we elucidate by means of simple examples the technique of reducing a boundary-value problem to a problem on the boundary. This Poisson principle provides the main idea for the solution of elliptic boundary-value problems, particularly for the computation of the index.

Siméon-Denis Poisson and other mathematicians of the beginning of the 19-th century occupied themselves with the observation that in equilibrium the temperature and heat distribution in a three dimensional body is completely determined by the temperature distribution on its surface. Similar statements are true for static magnetic and electric fields (e.g., the potential of an electric field in a conductor is determined by the charge density on the surface). The mathematical content of such laws of nature can be expressed variously depending on the context:

Exercise 10.23. (Poisson Integral Formula): Show that every harmonic function on the disk $\{z \in \mathbb{C} : |z| < \rho\}$, which has a continuous extension to the boundary, can be expressed in terms of its restriction u_0 to the boundary curve $\{z \in \mathbb{C} : |z| = \rho\}$. [Hint: For $0 \le r < \rho$ and $0 < \psi < 2\pi$, prove the formula $u = Lu_0$, where L is the *Poisson operator* given by

$$Lu_0\left(re^{i\psi}\right) := \frac{1}{2\pi}\int_0^{2\pi} \frac{\rho^2 - r^2}{\rho^2 + r^2 - 2r\rho\cos(\psi - \varphi)} u_0\left(\rho e^{i\varphi}\right)\, d\varphi,$$

by going back to the Cauchy Integral Formula

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta,\ |z| < \rho,\ \text{where } \gamma(\varphi) = \rho e^{i\varphi},\ \varphi \in [0, 2\pi],$$

which holds when f is holomorphic for $|z| < \rho$ and continuous for $|z| \le \rho$. Recall that u is harmonic ($\Delta u = 0$) in a disk (or simply-connected region) exactly when u is the real part of a holomorphic function. Details are found in [**Ah**, p.175f], but see also [**DM**, p.164].]

Exercise 10.24. Let X be a codimension 0, bounded, submanifold of $\mathbb{R}^n$ with C^∞ boundary Y, and let Δ be the Laplace operator and R_0, R_1 be differential operators on Y. Set $u_0 := u|_Y$ and $u_1 := \frac{\partial u}{\partial \nu}|_Y$ for $u \in C^\infty(X)$, where $\frac{\partial}{\partial \nu}$ is the outward normal field on the boundary Y of X. Show: The solution of the boundary-value problem

$$(*) \qquad \Delta u = 0 \text{ in } X \text{ and } R_0u_0 + R_1u_1 = f \text{ on } Y$$

is *equivalent* to the solution of a system of pseudo-differential equations

$$(**) \qquad (\mathrm{Id} - Q_0)u_0 - Q_1u_1 = 0 \text{ and } R_0u_0 + R_1u_1 = f \ \text{ on } Y,$$

where $Q_0, Q_1 \in P\,\mathrm{Diff}_0(Y)$ can be given explicitly.
[Hint: Begin with Green's Formula,

$$\int_X (u\Delta v - v\Delta u)\, dx = \int_{\partial X} (u_0v_1 - v_0u_1)\, dy, \quad u, v \in C^\infty(X),$$

where dy is the volume element on $\partial X = Y$. (More generally, for every differential operator $P \in Diff_k(X)$, the difference $\langle u, Pv\rangle - \langle P^*u, v\rangle$ can be estimated, using Stokes' Theorem, by a form on ∂X in which only the derivatives on ∂X of u and v of order $\le k-1$ enter; see Exercise 6.17 (p. 169), and [**Pal65**, p.73-75]. Now determine (e.g., by elementary distribution theory) a *fundamental solution* $\mathcal{E}(x-x')$ of the Laplace operator (i.e., $(\Delta_x \mathcal{E})(x-x') = \delta(x-x')$); in classical terminology $\mathcal{E}$ is the *Newtonian kernel* of Δ. For $u \in \operatorname{Ker}\Delta$ and $v(x) := \mathcal{E}(x-x')$, derive the *Poisson Integral Formula* (for $x' \in X - \partial X$)

$$u(x') = L(u_0,u_1)(x') := \int_{\partial X} u_0(y)\,\tfrac{\partial \mathcal{E}}{\partial \nu}(x'-y)\,dy - \int_{\partial X} u_1(y)\mathcal{E}(x'-y)\,dy.$$

Now show (by letting $x' \to \partial X$) that pseudo-differential operators Q_0 and Q_1 are defined such that for each solution $u \in C^\infty(X)$ of the boundary-value problem $(*)$, we have the equation $u_0 = Q_0 u_0 + Q_1 u_1$.

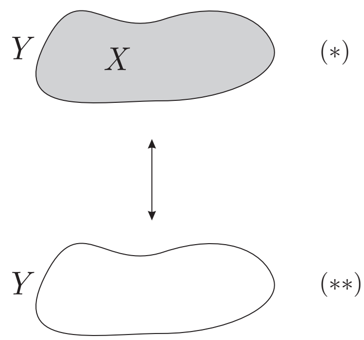


FIGURE 10.6. xxx5

Conversely, show that each solution pair (u_0,u_1) of the pseudo-differential system $(**)$ furnishes a solution $u := L(u_0,u_1)$ of $(*)$ by means of the Poisson integral, such that $u_1 = \frac{\partial u}{\partial \nu}$. Details are found in [**Hö66a**, p.187] or [**Jö**, 9.3-9.5].]

REMARK 10.25. Pseudo-differential operators thus turn up very naturally when one goes from a boundary-value system, which may only contain differential operators, to a problem on the boundary.

7. The Green Algebra

The treatment of boundary-value problems via functional analysis is often troublesome. One reason for this lies in the natural *asymmetry* of operators

$$(P,R) : C^\infty(E) \to C^\infty(F) \oplus C^\infty(G),\ u \mapsto (Pu, Ru),$$

where E and F are vector bundles over X, and G is a vector bundle over ∂X; so, a formal adjoint operator to (P,R) would be of an entirely different form and no longer defines a boundary-value problem. Also, one has no clear composition rules. Here, the following algebra which extends the class of boundary-value systems, is helpful; the construction goes back, in particular, to the work of Mark Josifovich Vishik and Grigory Ilyich Eskin. However, for the most part, we follow the specific approach of Louis Boutet de Monvel and employ his definition and method of proof; see [**Bou**] and [**GruG**], where there is a non-technical introduction to the Boutet-de-Monvelian calculus in the appendix.

We consider systems of equations of the form

$$\begin{aligned}(P+\mathcal{G})e + Lg &= f\\ Re + Qg &= h\end{aligned}$$

or (in other words) operators of the form

$$\begin{bmatrix} P+\mathcal{G} & L \\ R & Q \end{bmatrix} : \begin{matrix} C^\infty(E) \\ \oplus \\ C^\infty(G) \end{matrix} \to \begin{matrix} C^\infty(F) \\ \oplus \\ C^\infty(H) \end{matrix} ,$$

where E and F are Hermitian vector bundles over the compact, oriented, C^∞ Riemannian manifold X with boundary $Y := \partial X$, and G and H are Hermitian vector bundles over Y; e, f, g, h are C^∞ sections of the corresponding vector bundles. Here, $P \in \mathrm{L}^k_{\mathrm{pc}}(E, F)$ is a pseudo-differential operator over X ($\mathcal{G}$ is described in the next paragraph), $Q \in \mathrm{L}^{k'}_{\mathrm{pc}}(G, H)$ is a pseudo-differential operator over the boundary Y of X, $R := \sum Q_i(P_i(\cdot)|_Y)$ is a *trace operator* which is composed of pseudo-differential operators Q_i over Y and P_i over all of X. Moreover, $L := \{P_0((\cdot)\delta(Y))\}\,|_{X\setminus Y}$ is a *Poisson operator*, where $P_0 \in \mathrm{L}^k_{\mathrm{pc}}(\widetilde{G}, F)$ where $\widetilde{G}$ is a vector bundle over X with $\widetilde{G}|_Y = G$ and $\delta(Y)$ is the Lebesgue measure on the boundary Y of X, so that $g\delta(Y)$ can be regarded as a *generalized section* of the vector bundle $\widetilde{G}$ for $g \in C^\infty(G)$. Hence, the operator L is defined in the sense of distribution theory. For a direct definition of these *Poisson operators* (*potentials*) by means of Fourier representation (without distributions) see [**Bou**, p.26-28]. By a certain symmetry condition (the *transmission condition*) on the behavior of the amplitudes of the pseudo-differential operators on the boundary, it follows that the images of $C^\infty(E)$ under P and $C^\infty(G)$ under L lie in $C^\infty(F)$; this symmetry condition holds in particular for differential operators and more generally for rational symbols, and represents no real restriction for our considerations. More precisely, one shows that, for each $g \in C^\infty(G)$, Lg lies in $C^\infty(F|_{(X\setminus Y)})$ and extends to a C^∞ section of F (over all of X), which we also denote by Lg.

One has to add to P a *singular Green operator* $\mathcal{G}$ which is finite sum of our *Poisson operators* and *trace operators* if one wants to describe (for example) the changes in the solution of an elliptic boundary-value problem resulting from a modification of the boundary conditions, or, more generally, if one wishes to secure the existence of a *parametrix* of this form (see Theorem 10.32a, p. 260) for arbitrary or only differential boundary-value systems. We have the following theorems; we refer to [**Bou**] and [**VE**] for their proofs.

THEOREM 10.26. *Let $P, \mathcal{G}, L, R, Q$ be defined as above, and let m be order* (*the homogeneity of the symbol*) *of P, l the order of L, r order of R, and $-m+l+r+1$ the order of Q. Then the operator*

$$A = \begin{bmatrix} P+\mathcal{G} & L \\ R & Q \end{bmatrix} : \begin{matrix} C^\infty(E) \\ \oplus \\ C^\infty(G) \end{matrix} \to \begin{matrix} C^\infty(F) \\ \oplus \\ C^\infty(H) \end{matrix}$$

possesses (*for s sufficiently large*) *a continuous extension*

$$A_s = \begin{matrix} W^s(E) \\ \oplus \\ W^{t+l-\frac{1}{2}}(G) \end{matrix} \to \begin{matrix} W^t(F) \\ \oplus \\ W^{s-r-\frac{1}{2}}(H) \end{matrix} , \ t := s - m$$

to Sobolev spaces. (See Chapter 7 above.)

THEOREM 10.27. *The operators of the stated type form an algebra (called the Green algebra by Boutet de Monvel) which is closed under sums $A+B$, and composites $A \circ C$ (in so far as they are defined, and in particular their bundles or domains* *and codomains fit). In certain cases*[2]*, it is also closed under the formation of formal adjoints. Moreover, we have*

$$\sigma_X(A+B) = \sigma_X(A) + \sigma_X(B) \text{ and } \sigma_X(A \circ B) = \sigma_X(A) \circ \sigma_X(B),$$

where $\sigma_X(A) := \sigma(P) \in \mathrm{Hom}_{SX}(E,F)$ *is the* ***inner symbol*** *of* A*; hence,* $\sigma(P)(x,\xi) \in \mathrm{Hom}(E,F)$ *for* $x \in X$ *and* $\xi \in T_x^*X$ *with* $|\xi| = 1$*. (The notion of boundary symbol may be also introduced in greater generality, see the following discussion of the elliptic case.)*

8. The Elliptic Case

We first recall the approach chosen for deriving the main analytic theorems for elliptic operators on *closed* manifolds in Chapter 9 above: There, we were successful in deriving the various results (Fredholm properties) of the operators from formal properties of the associated symbols. Roughly speaking, the symbol map yields an *algebraic* version of the operator. The symbol is defined pointwise and for each cotangent vector individually so that this type of *snapshot* (pointwise *freezing*) produced, from an operator of functional analysis, a family of matrices which could be investigated with the tools of linear algebra. For example, ellipticity of the operator P was characterized by the condition that the matrices $\sigma(P)(x,\xi)$ are invertible for all non-vanishing covectors ξ.

For the analytical treatment of elliptic boundary-value systems, we need a refinement of this symbol calculus. For example, let us take an element

$$A = \begin{bmatrix} P+\mathcal{G} & L \\ R & Q \end{bmatrix} : \begin{matrix} C^\infty(E) \\ \oplus \\ C^\infty(G) \end{matrix} \rightarrow \begin{matrix} C^\infty(F) \\ \oplus \\ C^\infty(H) \end{matrix}$$

of the *Green algebra* over the manifold X with boundary Y. For elliptic P, by definition, $\sigma(P)(y,\eta+t\nu) : E_y \to F_y$ is an isomorphism of vector spaces, where $y \in Y$, $\eta \in (SY)_y$, $\nu \in (SX)_y$ is the inwardly directed normal relative to a Riemannian metric, and $t \in \mathbb{R}$. We denote by $p(y,\eta)$ the function defined on $\mathbb{R}$ by $t \to \sigma(P)(y,\eta+t\nu)$ with values in $\mathrm{Aut}(E_y)$, if we identify F with E by means of $\sigma(P)(y,\eta+0)$. If P is an elliptic operator of order 0 and $\lim_{t\to+\infty} p(y,\eta+t\nu) = \lim_{t\to-\infty} p(y,\eta+t\nu)$, then $p(y,\eta)$ defines (e.g., by means of the Cayley transformation; see p. 6) a continuous map $S^1 \to \mathrm{Aut}(E)$ and hence a *discrete Wiener-Hopf operator*

$$T_{p(y,\eta)} : H_0(S^1) \otimes E_y \to H_0(S^1) \otimes E_y,$$

which is a Fredholm operator with

$$\mathrm{index}\, T_{p(y,\eta)} = -W(\det(p(y,\eta)), 0).$$

Here (as in Chapter 4), $H_0(S^1) \otimes E_y$ is the Hilbert space of square integrable functions on the circle S^1, with values in the vector space E_y, which can be analytically continued to the disk $\{z \in \mathbb{C} : |z| < 1\}$ and T_f (for $f : S^1 \to \mathrm{Aut}(E_y)$) is the operator defined by pointwise multiplication by f and subsequent orthogonal projection onto $H_0(S^1) \otimes E_y$. See Chapter 4 for details, and in particular, Exercise 4.7, p. 128.

[2]While the adjoint L^* of a Poisson operator L always is a trace operator, the adjoint of a restriction operator $C^\infty(E) \to C^\infty(E|Y)$ (which arises in the Dirichlet problem, for example) is not a Poisson operator of the form $C^\infty(E|Y) \to C^\infty(E)$. One can place necessary and sufficient conditions on the corresponding symbol classes so that R^* is a Poisson operator (namely, if R is of *class 0*; see [**GruG**]). Similar restrictions hold for $\mathcal{G}$.

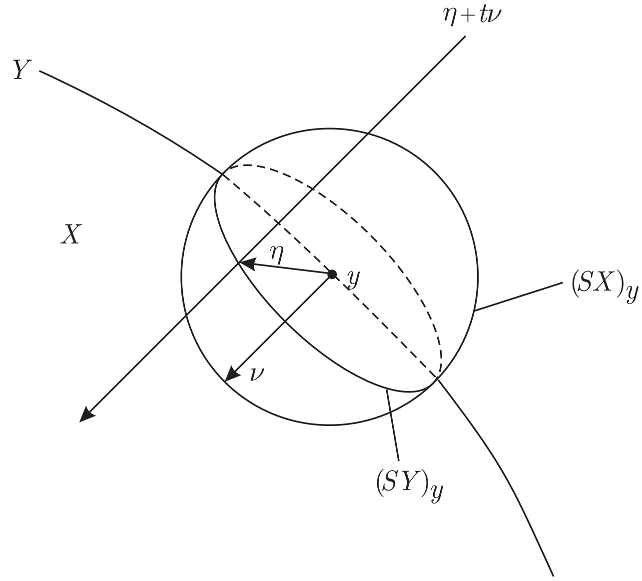


FIGURE 10.7. xxx6

If $p(y,\eta)$ does *not* define a continuous map $S^1 \to \mathrm{Aut}(E_y)$, these constructions go through with certain modifications: One then works, as in Exercise 4.10 and Theorem 4.11 (p.129), on the real line and forms the corresponding multiplication operators with projection, the *continuous Wiener-Hopf operators* $W_{p(y,\eta)}$ which are Fredholm operators on suitably chose *Hardy spaces*. If the order $k \neq 0$, we must develop the *Hardy spaces* $H_0(\mathbb{R})$ (introduced in (4.1), p.130) into a Sobolev chain, as we have done with the Hilbert space $L^2(X)$ in Chapter 7. Details may be found in [**VE**, p.304-311, esp. Lemma 3.1] and [**Bou**, p.14-20]; see also [**GruG**, Appendix]. In each case, from the ellipticity of P there follows the invertibility of $p(y,\eta)$, regarded as a homomorphism $\sigma(P)(y,\eta+t\nu) : E_y \to E_y$ for all $t \in E_y$ (but not the invertibility of $p(y,\eta)$, regarded as a Wiener-Hopf operator). To each elliptic operator P on a manifold X with boundary Y, we can assign a family of Fredholm operators $W_{p(y,\eta)}$ which depends continuously on the unit covector $\eta \in SY$ whose base point y lies in Y. By Theorem 3.30 (p. 84), an *index bundle* in $K(SY)$ can be constructed in a canonical way from such a family. We denote this bundle by $j(P)$, the **indicator bundle** of P.

EXERCISE 10.28. Compare the construction of $j(P)$ with the definition of the bundle $\mathcal{M}^+$ in Chapter 10 for an elliptic differential operator P. In particular, investigate $P \in \{\Delta, \partial/\partial\bar{z}, \partial/\partial\bar{z} \oplus \partial/\partial z\}$. [Hint: [**Bou**, p.35], [**VE**, p.328-331].]

EXERCISE 10.29. Show that each operator $A = \begin{bmatrix} P+\mathcal{G} & L \\ R & Q \end{bmatrix}$ of the Green algebra defines a family $\sigma_Y(A)$ of Wiener-Hopf operators $\sigma_Y(A)(y,\eta)$ parametrized by $\eta \in S(Y)$, and that the usual addition and composition rules for these *boundary symbols* hold. [Hint: [**GruG**, Appendix].]

EXERCISE 10.30. Show that, for $y \in Y$ and $\eta \in (SY)_y$, $\sigma_Y(A)(y,\eta)$ is a Fredholm operator, if P is elliptic. Then define the indicator bundle $j(A) \in K(SY)$, as the index bundle of this family of Fredholm operators, and show $j(A) = j(P) + [\pi^*(G)] - [\pi^*(H)]$, where $\pi : SY \to Y$ is the projection. [Hint: The boundary symbol $\sigma_Y(\widetilde{A})(y, \eta)$ with $\widetilde{A} = \begin{bmatrix} 0 & L \\ R & Q \end{bmatrix}$ is an operator of finite rank for each $y \in Y$ and $\eta \in (SY)_y$.]

We can now define $A = \begin{bmatrix} P + \mathcal{G} & L \\ R & Q \end{bmatrix}$ to be **elliptic** when P is elliptic and the Wiener-Hopf operator $\sigma_Y(A)(y, \eta)$ is invertible, for all $y \in Y$ and $\eta \in (SY)_y$. The purpose of the boundary conditions R and the potential L is, by this definition, exactly to *use up* the kernel and cokernel of the Wiener-Hopf operators $W_{p(y,\eta)}$. We then have $j(A) = 0$, and hence $j(P) = [\pi^*(H)] - [\pi^*(G)]$.

Exercise 10.31. For a differential operator P with (differential) boundary operator R, show that (P, R) is an elliptic boundary-value system in the sense of Chapter 10, exactly when $\begin{bmatrix} P & 0 \\ R & 0 \end{bmatrix}$ is an elliptic operator in the Green algebra. [Hint: Compare the condition of invertibility of the Wiener-Hopf operators $\sigma_Y(A)(y, \eta)$ with the Lopatinsky condition (10.4), p. 242. See also Exercise 10.28. How can β^+ be interpreted? [**Bou**, p.45f].]

Theorem 10.32. *If* $\mathrm{Ell}_k(X; Y)$ *is the class of elliptic operators in the Green algebra over the compact, oriented, Riemannian* C^∞ *manifold* X *with boundary* $\partial X = Y$*, then we have*
a) Each $A \in \mathrm{Ell}_k(X; Y)$ *possesses a parametrix* $B \in \mathrm{Ell}_{-k}(X; Y)$*; i.e.,* $AB - \mathrm{Id}$ *and* $BA - \mathrm{Id}$ *are operators of order* -1*.*
b) If $A \in \mathrm{Ell}_k(X; Y)$*, then the extensions* A_s*, defined above in* Theorem 10.26 *for* $s \gg 0$*, are Fredholm operators on the corresponding Sobolev spaces.*
c) If A, B*, and* C *are elliptic, then* $\mathrm{index}(A \oplus B) = \mathrm{index}\, A + \mathrm{index}\, B$*. Also, if* $A \circ C$ *is well defined, then* $\mathrm{index}(A \circ C) = \mathrm{index}\, A + \mathrm{index}\, B$*.*
d) Two elliptic operators A *and* B *are called* ***stably equivalent*** *(*$A \sim B$*), if the interior and boundary symbols of* $A \oplus \mathrm{Id}_N$ *and* $B \oplus \mathrm{Id}_M$ *can be continuously deformed into each other while maintaining ellipticity. Then, we have* $\mathrm{index}\, A = \mathrm{index}\, B$*.*
e) If R, R' *are boundary operators, and* L, L' *are Poisson operators, and* Q *is a pseudo-differential operator on the boundary, then*

$$\begin{bmatrix} \mathrm{Id} - L'R' & L \\ R & Q \end{bmatrix} \sim \begin{bmatrix} \mathrm{Id} - R'L' & -R'L \\ -RL' & Q - RL \end{bmatrix},$$

where the second operator is a pseudo-differential operator on the boundary.

Arguments: a) follows easily from the definition of ellipticity, if one knows that the inverse of the boundary symbol $\sigma_Y(A)$ is again the boundary symbol of an operator of the Green algebra; see [**Bou**, p.19f and 34f]. Incidentally, for an elliptic boundary system $A = \begin{bmatrix} P \\ Q \end{bmatrix}$ in the sense of Chapter 10, one can find a parametrix of the form $B = (\widetilde{P} + G, L)$. One obtains b), c), and d) from Theorem 10.26 and the general theory of Fredholm operators in Hilbert space as the corresponding statements in Chapter 9. e) is the consequential (see Exercise 10.33 below) result of a topological exercise [**Bou**, p.44f].

EXERCISE 10.33. (M. S. Agranovich and A. S. Dynkin, 1962): If $A_1 = \begin{bmatrix} P \\ R_1 \end{bmatrix}$ and $A_2 = \begin{bmatrix} P \\ R_2 \end{bmatrix}$ are two elliptic boundary-value systems for the same elliptic operator P on the manifold X with boundary Y, then $\operatorname{index} A_1 - \operatorname{index} A_2 = \operatorname{index} Q$, where Q is a pseudo-differential operator on Y defined in a canonical way by means of A_1 and A_2.
[Hint. If $B_1 := \begin{bmatrix} P' & L_1 \\ 0 & 0 \end{bmatrix}$ is a parametrix for A_1, set $Q := R_2 L_1$, and show that $A_2 B_1 \sim Q$ with Theorem 10.32e.]

Part III

The Atiyah-Singer Index Formula

"In perhaps most cases when we fail to find the answer to a question, the failure is caused by unsolved or insufficiently solved simpler and easier problems. Thus all depends on finding the easier problem and solving it with tools that are as perfect as possible and with notions that are capable of generalization." (D. Hilbert, 1900)

CHAPTER 11

Introduction to Algebraic Topology (K-Theory)

It is the goal of this part to develop a larger portion of algebraic topology by means of a theorem of Raoul Bott concerning the topology of the general linear group $\mathrm{GL}(N, \mathbb{C})$ on the basis of linear algebra, rather than the theory of "simplicial complexes" and their "homology" and "cohomology". There are several reasons for doing so. First of all, is of course a matter of taste and familiarity as to which approach "codifying qualitative information in algebraic form" (Atiyah) one prefers. In addition, there are objective criteria such as simplicity, accessibility and transparency, which speak for this path to algebraic topology. Finally, it turns out that this part of topology is most relevant for the investigation of the index problem.

Before developing the necessary machinery, it seems advisable to explain some basic facts on winding numbers and the topology of the general linear group $\mathrm{GL}(N, \mathbb{C})$. Note that the group $\mathrm{GL}(N, \mathbb{C})$ moved to fore in Part II already in connection with the symbol of an elliptic operator, and that the group $\mathbb{Z}$ of integers was in a certain sense the topic of Part I, the Fredholm theory. In the following, Part III, the concern is (roughly) the deeper connection between the previous parts. Thereby we will be guided by the search for the "correct" and "promising" generalizations of the theorem of Israil Gohberg and Mark Krein on the index of Wiener-Hopf operators. See Chapter 4, Theorem 4.4 (p.126), Exercise 4.9 (p.128), and Theorem 4.11 (p.129).

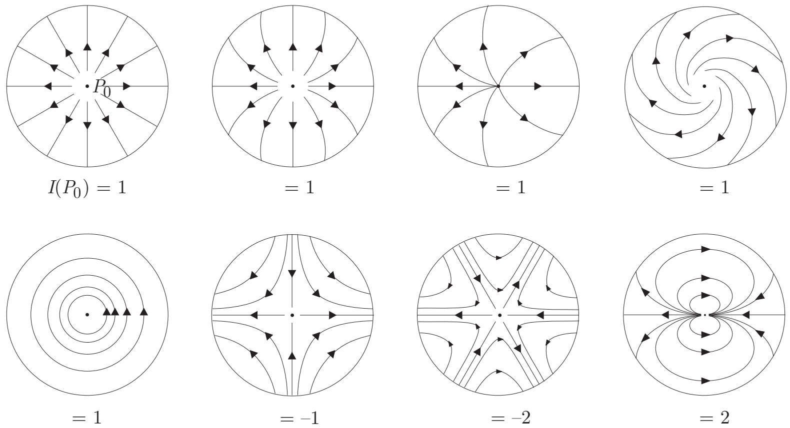
1. Winding Numbers

"How can numerical invariants be extracted from the raw material of geometry and analysis?" (Hirsch). A good example is the concept of winding number, surely the best known item of algebraic topology: In his studies of celestial mechanics the French physicist and mathematician Henri Poincaré turned to stability questions of planetary orbits. Many of the related problems are not completely solved even today (e.g., the "three body problem" of describing all possible motions of three points which are interact via gravitation. However, this problem is solved for practical purposes, as is shown by the successful landing of the lunar module Luna 1.)

As a tool for the qualitative investigation of nonlinear (ordinary) differential equations, Poincaré introduced in 1881 the notion of the "index" $I(P_0)$ of a "singular point" P_0 for a system of two ordinary differential equations

$$\dot{x} = F(x, y), \;\; \dot{y} = G(x, y).$$

To do this, surround P_0 by a closed curve C in the phase portrait (see the following examples) and measure on it the angle of the rotation



performed by the vector field $(F(x,y), G(x,y))$ when (x,y) traverses C once counterclockwise. The angle is an integral multiple of 2π, and this integer is $I(P_0)$. In case of a magnetic field one can actually see $I(P_0)$ in the rotation of the needle, when a compass is moved along C. Among other things, one has the theorem (see [**AM**, p.75-76]): If the equilibrium position P_0 is stable, then $I(P_0) = 1$. We will come back to this in Section 14.14.5.

Poincaré returned to this topological argument in 1895, when he considered all closed curves in an arbitrary "space" and classified them according to their deformation properties.[1] His simplest result can be expressed in today's terminology (we follow [**Ati67b**, p.237-241]) follows:

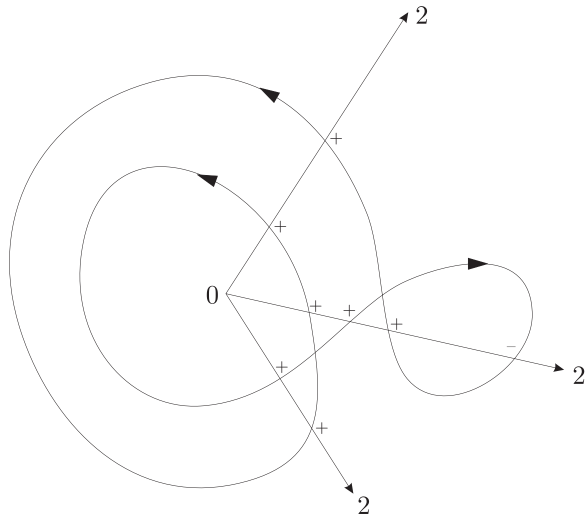
THEOREM 11.1. *Let $f : S^1 \to \mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ be a continuous mapping of the circle X to the punctured plane of non-zero complex numbers $\mathbb{C}^\times$. In other words, we have a closed path in the plane not passing through the origin. The following hold:*
(i) f possesses a "winding number" which states how many times path rounds the origin; we write $W(f,0)$ or $\deg(f)$.
(ii) This degree is invariant under continuous deformations.
(iii) $\deg(f)$ is the only such invariant, i.e., f can be deformed to g, if and only if $\deg(f) = \deg(g)$.
(iv) For each integer m, there is a mapping f with $\deg(f) = m$.

Arguments: Instead of a formal proof, we briefly assemble the different ways of defining or computing $\deg(f)$.

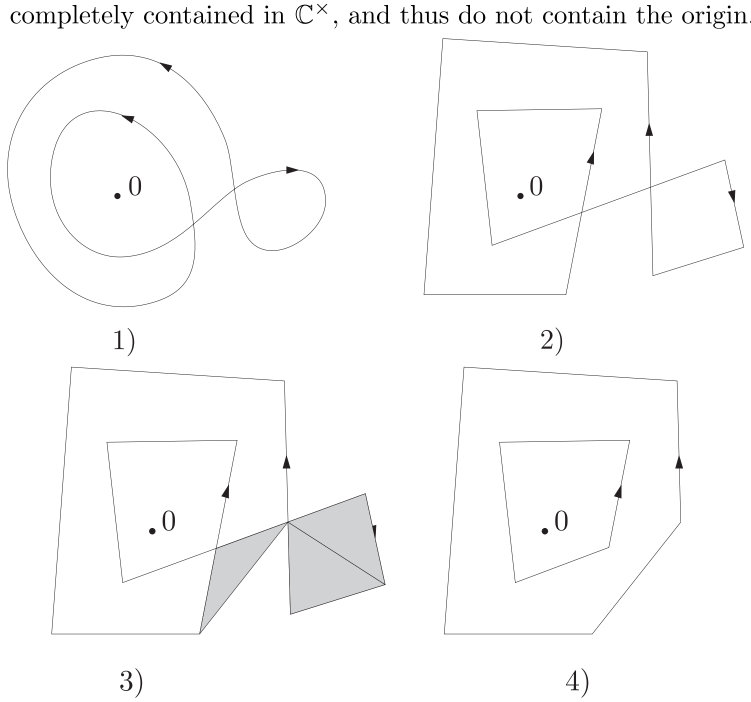
Geometrically: Replace f by $g := f/|f|$. This is a mapping from S^1 to S^1. Approximate g by a differentiable map h, and count (algebraically, i.e. with a sign

[1] The "fundamental group" was introduced in Poincaré's work Analysis Situs (Oeuvres 6, 193-288), whose theme is purely topological-geometric-algebraic: an "analysis situs in more than three dimensions." Poincaré expected the abstract formalism to "do in certain cases the service usually expected of the figures of geometry." He mentioned three areas of application: In addition to the Riemann-Picard problem of classifying algebraic curves and the Klein-Jordan problem of determining all subgroups of finite order in an arbitrary continuous group, he particularly stressed its relevance for analysis and physics: "one easily recognizes that the generalized analysis situs would allow treating the equations of higher order and specifically those of celestial mechanics (the same way as H. P. had done it before with simpler types of differential equations; B.B.)... I also believe that I did not produce a useless work, when I wrote this treatise." The complexity and limited understanding of the topological problems did not however permit Poincare to carry out his program completely: "Each time I tried to limit myself I slipped into darkness."

convention according to the derivative of the number of points in the preimage of a point which is in general position. This method can also be characterized as "counting of the intersection numbers": Draw an arbitrary ray emanating from the origin which does not pass through a self-intersection point of the path. Now count the intersections of the path with the ray according to the "traffic of the right of way" (H. Weyl) - thus with a plus sign if the path has the right of way, and a minus sign when the ray has the right of way.



Combinatorially: We approximate with a piecewise linear path g, and then use combinatorial methods; i.e., we permit deleting and adding of those edges of our polygonal path which are boundaries of 2-simplices (these are triangles whose interior is completely contained in $\mathbb{C}^\times$, and thus do not contain the origin.



Differential: We approximate by a differentiable g, and then set $\deg(f) := \frac{1}{2\pi i}\int \frac{dg}{g}$. Here, we have regarded g as a map $[0, 2\pi] \to \mathbb{C}^{\times}$, and then the integral is defined as $\int_0^{2\pi} \frac{g'(\tau)}{g(\tau)} d\tau$. From the Cauchy Integral Formula, it follows that the integral is a multiple of $2\pi i$, and hence $\deg(f)$ is an integer.

Algebraic: Approximate by a finite Fourier series

$$g(\phi) = \sum_{\nu=-k}^{k} a_\nu e^{i\nu\phi}$$

We regard g as a map $S^1 \to \mathbb{C}^{\times}$ with $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Consider now the extension of g to the disk $|z| < 1$, where it is a finite Laurent series. We then obtain a meromorphic function h, and set $\deg(f) := N(h) - P(h)$ where $N(h)$ and $P(h)$ denote the number of zeros and poles of h in $|z| < 1$.

Function Analytic: Set $\deg(f) := -\operatorname{index} T_f$, where T_f is the Wiener-Hopf operator, assigned to f, on the space $H_0(S^1) \subset L^2(S^1)$ spanned by the functions $z^0, z^1, z^2, \dots$. T_f is defined (for $u \in H_0(S^1)$) by

$$\left(\hat{T_f u}\right)(n) := \begin{cases} \sum_{k=0}^{\infty} \hat{f}(n-k)\,\hat{u}(k) & \text{for } n \geq 0 \\ 0 & \text{for } n < 0, \end{cases}$$

where $\hat{f}(m) := \langle f, z^m \rangle$ is the m-th Fourier coefficient of f. For the details of this, see Theorem 4.4 above and Theorem 4.11 for the analogous representation $\deg(f) = \operatorname{index}(I + K_\phi)$ via the (continuous) Wiener-Hopf operator

$$(K_\phi u)(x) := \int_0^\infty \phi(x-y)u(y)\,dy;\ x \in \mathbb{R}^+,\ u \in L^2(\mathbb{R}^+),$$

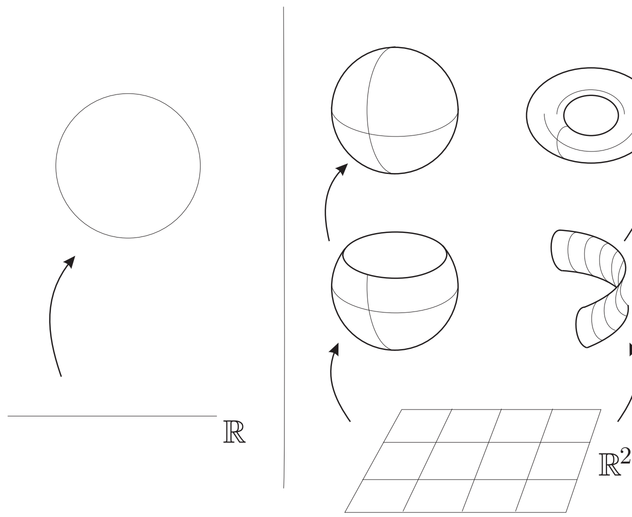
where $\phi \in L^1(\mathbb{R})$ with $\hat{\phi} = f \circ \kappa^{-1}$ and $\kappa t := \frac{t-i}{t+i}$ is the Cayley transformation.

As a first example, one may recall the standard map $a : S^1 \to \mathbb{C}^{\times}$ which is given by $a(z) = z$, $z \in \mathbb{C}$, $|z| = 1$. Here the equivalence of the various definitions is clear. We omit the proof for complicated cases and refer to [**Ah**, p.151]. Compare also Theorem 11.2 (p. 269) and Theorem 11.17 (p. 281) below and in another context [**Hir**, p.120-131] or [**BJ**, p.161 f].

In Theorem 11.1 all closed curves in the punctured plane are compared and the "essentially different" ones separated. The different definitions of winding number listed there reflect the main branches of topology with their different techniques, goals and connections. Accordingly, depending on the point of view chosen, very many generalizations of Theorem 11.1 to higher dimensions are possible (see also Exercise 11.22, p. 284):

If one sticks with the classification of systems of ordinary differential equations (which was the point of departure for Poincaré's topological papers) one would first try to distinguish the different possibilities of bending the real line into a closed curve in space or other higher dimensional spaces. In this fashion H. Poincaré (but also see the footnote above) conceived (among other things) the "fundamental group" $\pi_1(X, x_0)$ of a space X which arises from the homotopic classification of closed paths $S^1 \to X$ which pass through the point $x_0 \in X$. Here only the embedding question (which depends on the structure of X) is of interest, while the embedded images themselves of the compactified line are topologically identical. The reason is that there is just one way of bending the real line into a closed manifold, namely the form of a circle which may be traversed several times and may wind so many times around one or the other hole, but still remains topologically a circle.

The situation is different, when we pass from ordinary to partial differential equations. Here the classification essentially requires a differentiation between the various possibilities of bending the plane or higher dimensional Euclidean spaces into closed manifolds (see also [**Ati76b**].)



Now "genuine" global difficulties arise, since (as the above diagrams illustrate) there are already for the plane $\mathbb{R}^2$ different ways of "bending it together." For $\mathbb{R}^2$ it is still possible to survey completely the different forms which can be classified according to the "genus" of the surface, the number of its "handles" (see for instance [**Hir**, p.204 f]). The corresponding problem of the "bending of $\mathbb{R}^3$ has not been solved, although it is of special importance for the analysis of space-time processes of the real world by means of partial differential equations. For example, it took 100 years until the "Poincaré conjecture" was confirmed by Grigori Perelman in [**Per02, Per03a, Per03b**] (according to which every simply-connected, three-dimensional, closed manifold is homeomorphic to the 3-sphere S^3). We shall not comment on the proof in this monograph.

The following generalization of Theorem 11.1 is due to Raoul Bott. It is a true and fully understood achievement of topology, and in addition, touches on the heart, of the index prob1em for systems of elliptic differential equations.

2. The Topology of the General Linear Group

We consider continuous maps $f : S^{n-1} \to \mathrm{GL}(N, \mathbb{C})$, $2N \geq n$, where S^{n-1} is the unit sphere in $\mathbb{R}^n$ and $\mathrm{GL}(N, \mathbb{C})$ is the general linear group of all invertible linear maps from $\mathbb{C}^N$ to $\mathbb{C}^N$.

THEOREM 11.2. *(R. Bott, 1958). If n is odd, each such f can be deformed to a constant map. If n is even, one can define an integer* $\deg(f)$*, such that f can be deformed to another map g exactly when* $\deg(f) = \deg(g)$*. Moreover, there exist maps having arbitrarily prescribed integer degree.*

Arguments. (We will completely prove Theorem 11.2 in a different form below):

1. Theorem 11.1 is the special case of Theorem 11.2 for $n = 2$ and $N = l$. For $n = 1$ and N arbitrary, we recover the well-known fact that $GL(N, \mathbb{C})$ is pathwise connected; see the paragraph following Exercise 3.20, p. 74.

2. As we have formulated it here, the Bott Theorem is expressed in terms of deformations (i.e., "homotopies"), a central theme of topology. In the formalism of homotopy theory, the Bott Theorem says that for all n, N with $2N \geq n$ the homotopy groups $\pi_{n-1}(\mathrm{GL}(N, \mathbb{C}))$ (i.e., the group of homotopy classes of maps $S^{n-1} \to \mathrm{GL}(N, \mathbb{C})$) are given as follows:

$$\pi_{n-1}(\mathrm{GL}(N, \mathbb{C})) \cong \begin{cases} 0 & \text{if } n \text{ is odd} \\ \mathbb{Z} & \text{if } n \text{ is even.} \end{cases}$$

In the second case, the isomorphism is given by

$$\deg : \pi_{n-1}(\mathrm{GL}(N, \mathbb{C})) \stackrel{\cong}{\to} \mathbb{Z}.$$

Using these concepts the classically expressed result of our Theorem 11.1 above means that the first homotopy group (the "fundamental group") of $\mathbb{C}^\times$ is isomorphic to $\mathbb{Z}$.

Since Theorem 11.2 yields an isomorphism of $\pi_{n+1}(\mathrm{GL}(N, \mathbb{C}))$ with $\pi_{n-1}(\mathrm{GL}(N, \mathbb{C}))$, it is also known as a "periodicity theorem." Incidentally, there is a corresponding theorem (with period 8) for $\mathrm{GL}(N, \mathbb{R})$. It has close connections with the theory of real elliptic skew-adjoint operators; see Section 14.9 below.

3. As with Theorem 11.1 above, there are various ways in which the degree (for even n) can be defined: First, a differential definition of $\deg(f)$ is possible with the help of a known, explicitly defined, invariant differential form ω on the C^∞ manifold $\mathrm{GL}(N, \mathbb{C})$. For the not entirely simple definition of this "Weltkonstante" (F. Hirzebruch), we refer to [**Hi66b**, p.587 f]. One then sets

$$\deg(f) = \int_{S^{n-1}} f^*(\omega),$$

where $f^*(\omega)$ is the pull-back form over S^{n-1}, and shows (!) that the invariant, so defined, is an integer. As an alternative to this direct, somewhat computationally cumbersome definition, one can also define $\deg(f)$ *geometrically*, by means of a stepwise reduction to the more intuitive, but topologically no less demanding, notion of "mapping degree" of a continuous mapping of the $(n-1)$-sphere into itself. (In Theorem 11.1 the two concepts still coincided.) First one shows that, without loss of generality, one can take $2N = n$, since in the case $2N > 2$, f can be deformed into a map of the form

$$g(x) = \begin{bmatrix} h(x) & 0 \\ 0 & \mathrm{Id} \end{bmatrix},$$

where $h : S^{n-1} \to \mathrm{GL}(n/2, \mathbb{C})$. All further constructions, do not depend on the choice of g, since this gives, more precisely,

$$\pi_{n-1}(\mathrm{GL}(N, \mathbb{C})) \cong \pi_{n-1}(\mathrm{GL}(n/2, \mathbb{C})) \text{ for } N \geq n/2.$$

For those with enough background, we offer the following explanation of this:

From the exact homotopy sequence

$$\pi_{i+1}(\mathrm{U}(m), \mathrm{U}(n)) \to \pi_i(\mathrm{U}(n)) \to \pi_i(\mathrm{U}(m)) \to \pi_i(\mathrm{U}(m), \mathrm{U}(n))$$

induced by $\mathrm{U}(n) \hookrightarrow \mathrm{U}(m)$ for $n \leq m$ (where $\mathrm{U}(n) := \{A \in GL(n, \mathbb{C}) : AA^* = I\}$, and the exact sequence

$$\pi_{i+1}\left(S^{2n+1}\right) \to \pi_i(\mathrm{U}\,(n)) \to \pi_i(\mathrm{U}\,(n+1)) \to \pi_i\left(S^{2n+1}\right)$$

from the "fibration" $\mathrm{U}(n) \to \mathrm{U}(n+1) \to S^{2n+1}$. For $i < 2n$, we have $\pi_{i+1}\left(S^{2n+1}\right) \cong \pi_i\left(S^{2n+1}\right) = 0$; namely, show with the Sard Theorem (after differentiable approximation) that on account of different dimensions, not every map in such homotopy classes can be surjective, whence one always has a point in the complement of the image in S^{2n+1} from which' one can contract. Thus, it follows that $\pi_{i-1}(\mathrm{U}(m)) \to \pi_{i-1}(\mathrm{U}(n))$ for $1 \leq i \leq 2n \leq 2m$. Since the unitary group $\mathrm{U}(n)$ is a deformation ret of $\mathrm{GL}(n, \mathbb{C})$ (prove!), the statement follows; see [**Ste**, 5.6 and 19.5].

Hence, let $2N = n$. Then the first column of the matrix f define a map $f_1 : S^{n-1} \to \mathbb{C}^N \setminus \{0\}$. Since $\mathbb{C}^N \setminus \{0\} = \mathbb{R}^n \setminus \{0\}$, we have a map $g := f_1/\left|f_1\right|$, for which a degree (the "mapping degree", the "most" natural generalization of Theorem 11.1) may be easily defined: One approximates g by a differentiable map h, chooses a point $y \in S^{n-1}$ in "general position". This means that the differential h_{*x} (see Section 6.3 above) has maximal rank for all $x \in h^{-1}(y)$; i.e., that h_{*x} is an isomorphism of tangent spaces. Then one counts the algebraic number of points in $h^{-1}(y)$ (i.e., the points x where h_{*x} reverses orientation are counted negatively). Details for this are found in [**Hir**, p.121-131]. For other definitions of the "mapping degree" of g, see [**BJ**, 14.9.6-10] (on "intersection numbers" – thus again a geometrical definition, but in a more general setting), and [**ES**, p304 ff] (on the "homology" of S^{n-1}; i.e., fundamentally a combinatorial definition); furthermore, see Exercise 11.22, p. 284 ("K-theoretical" definition).

It turns out that the mapping degree of g

(a) contains the essential information on the qualitative behavior of f and

(b) is always divisible by $(N-1)!$, so that we then define

$$\deg(f) := \frac{(-1)^{N-1} \deg{(g)}}{(N-1)!},$$

where the sign $(-1)^{N-1}$ is there for secondary technical reasons.

One can perhaps best visualize (a) (concerning the term "visualize" see 4 below. Strictly speaking it is more a comprehension aid.) as follows: Topologically, the linear independence of the column vectors of a matrix means "approximately" the same as their being perpendicular. Thus, if such is the case, each function $S^{n-1} \to \mathrm{GL}(N, \mathbb{C})$ is so "rigid" that it can be completely classified topologically by means of a single column.

In fact (and this now concerns (b)) the maps $S^{n-1} \to \mathrm{GL}(N, \mathbb{C})$ are so "rigid" that in general, by far not every function $S^{n-1} \to \mathbb{C}^N \setminus \{0\}$ can appear as the first column. What is the reason? Isn't it possible to make an invertible matrix out of any non-zero vector by adding orthogonal vectors? Yes and no: It is possible to do so at each single point, but not if the additions are to be made uniformly in continuous dependence on the points of S^{n-1}. One may think of a sphere which as the 2-sphere has no unit tangential vector field (see Exercise 11.23, p. 284), and thus eliminates the identity map as a possible first column.

For an evaluation of the "difficulty" of the divisibility theorem (b), see [**?**, p.182], where the theorem was proven in somewhat different form "pretty much at the end of the study (on the Theorem of Riemann-Roch, B.B.) as a corollary of cobordism theory, but which on the other hand, belongs "not at the end, but at the beginning", namely with "the Bott periodicity theory which is the basis of the newer proofs of the Riemann-Roch Theorem".[2]

4. What can be said so far on the *substance* of the Bott Periodicity Theorem? We stay in the case n even, $2N \geq n$. Then, we have three statements:
(i) degree : $\pi_{n-1}(\mathrm{GL}(N,\mathbb{C})) \to \mathbb{Z}$ is well defined,
(ii) the map is surjective, and
(iii) the map is injective.
As we have seen, for (i) two approaches are available, a "differential" one and a "geometrical" one. In this way, the degree was defined one time as an integral of a differential form (hence it is, a priori, a real number), and the other time as a quotient of two integers (a priori, as a rational number). Either way, statement (i) causes no difficulties except for the integrability and divisibility theorems needed, since the homotopy invariance of degree is rather clear from the definition. One might simply say: (i) concerns the definition of an integer homotopy invariant - this is homology and comparatively simple.

Also, (ii) is comparatively simple: Namely, one can form (as when tensoring elliptic operators in Exercise 9.13, p. 238) the tensor product of $f : S^{n-1} \to \mathrm{GL}(N,\mathbb{C})$ and $g : S^{m-1} \to \mathrm{GL}(M,\mathbb{C})$,

$$f\#g : S^{m+n-1} \to \mathrm{GL}(2MN,\mathbb{C}), \text{ defined by}$$

$$(f\#g)(x,y) := \begin{bmatrix} f(x)\otimes \mathrm{Id}_M & -\mathrm{Id}_N \otimes g^*(y) \\ \mathrm{Id}_N \otimes g(y) & f^*(x)\otimes \mathrm{Id}_M \end{bmatrix},$$

where f and g are extended homogeneously to all of $\mathbb{R}^n$ resp. $\mathbb{R}^m$. From the simple multiplication formula $\deg(f\#g) = (\deg f)(\deg g)$ it then follows that $\deg(a_k) = 1$, where $a : S^1 \to \mathrm{GL}(1,\mathbb{C})$ is the standard map ($a(z) := z$) of degree 1 and the map a_k is its k-fold power:

$$a_k := a\#\overset{k \text{ times}}{\cdots}\#a : S^{2k-1} \to \mathrm{GL}(2^{k-1},\mathbb{C}).$$

Thus, product theory furnishes a generating map meaning a generating element of the infinite cyclic group of homotopy classes of maps $S^{n-1} \to \mathrm{GL}(N,\mathbb{C})$ for $n = 2k$, whence the surjectivity of "degree" follows.

The statement (iii) is in contrast highly nontrivial: While in (i) and (ii) one had to construct particular objects (a homotopy-invariant "number" and a suitable homotopically nontrivial map), one must now show that mappings of the same degree may be deformed into each other, in particular, that each map of degree 0 is homotopic to a constant map. Having defined the invariants, one wants to verify their relevance and determine their value. This is exactly "homotopy" and not at all intuitive. It can be seen from the fact that, although the homotopy classes of maps $S^{n-1} \to \mathrm{GL}(N,\mathbb{C})$ (for $2N < n$) as well as those from spheres to spheres are in a fashion "closer" to our space-time comprehension, they are largely unknown, because of their extreme complexity.

---

[2]In the language of Section 14.1, the facts are the following: The Chern character $ch_N : \widetilde{K}(S^{2N}) \to H^{2N}(S^{2N};\mathbb{Q})$ is given by $(-1)^{N-1}c_N/(N-1)!$, since all other terms in the formula for ch_N cancel because of Bott periodicity which here takes the form $\widetilde{K}(S^{2N}) \cong H^{2N}(S^{2N};\mathbb{Q})$. If E is the vector bundle over S^{2N} of complex dimension N, which is constructed by gluing with $f : S^{2N-1} \to \mathrm{GL}(N,\mathbb{C})$, then (regarding the sign see above) $\deg f = ch_N([E] - [\mathbb{C}^N_{S^{2N}}])$, where $[E]$ is the class of E in $K(S^{2N})$ furthermore the N-th Chern class $C_N(E)$ is mapped to $\deg(g)$ under the isomorphism $H^{2N}(S^{2N};\mathbb{Z}) = \mathbb{Z}$, where $g : S^{2N-1} \to S^{2N-1}$ is defined by means of f as before.

5. Without commenting on the very different proofs of the Periodicity Theorem which exist to date, we would like to point out that *all* proofs proceed by induction on n, more precisely by induction from n to $n+2$. In the language of the product theory presented above, it must be shown that $f \mapsto f\#a$ is an isomorphism of the homotopy group of dimension $n-1$ with the homotopy group of dimension $n+1$.

6. Interestingly, the "*algebraic*" definition of winding number (see Theorem 1) was initially not susceptible to generalization to the present situation. However, the discovery of the topological significance of elliptic boundary-value problems (see above Chapters 10-**??** and below, Section 14.14.8) lead Raoul Bott and Michael F. Atiyah to a new and elementary[3] proof of the Periodicity Theorem. In a very deep sense, which we will explain below in a particularly suitable form when presenting the proof, this proof generalizes and unifies the algebraic and function analytic definitions and furthermore "gets to the heart of the problem" (Atiyah).

The "Lie group" $GL(N, \mathbb{C})$ and its real analogue play an important role everywhere in mathematics. Accordingly the Periodicity Theorem and its off-spring "K-theory" (see below) not only have immediate applications to the index problem of elliptic oscillations (see particularly Chapters 12-14 below) but also proved to be a useful tool for a number of deep geometric problems. An example is the computation of the number of linearly independent vector fields on a sphere which the British mathematician John Frank Adams carried out in 1962 with just these tools. A detailed exposition of this and some other applications can be found in [**Hu**, Ch.15].

The deep relation between the topology of the general linear group and the geometry of differentiable manifolds, which becomes manifest in these successes, can be described intuitively as follows: Among the simplest global topological invariants of a compact oriented n-dimensional manifold X is the "Euler characteristic" $e(X)$. A famous theorem, proven in 1895 by Henri Poincaré for $n = 2$, and in general by Heinz Hopf in 1925, says that $e(X)$ can be found by means of a differentiable structure on X as the number of "singularities". Here a singularity is an isolated zero x of a tangent vector field v on X, and the counting must be done with the proper multiplicity, namely the (local) "index" of v at x (i.e., the degree of the mapping $S^{n-1} \to S^{n-1}$) which is given by v on the surface of a ball about x. For a conceptually very plausible proof for $n = 2$ see [**Bri76**, p.166-171], see also Sections 14.4 and 14.5 below.

Today very many global topological invariants are known which are .) defined on X by means of a classical (Riemannian or complex) structures see [**Hi66a**]. These "characteristic classes" are without exception generalizations of the Euler characteristic since "roughly, considers the cycles where a given number of vector fields become dependent", as [**Ati68b**, p.59] remarks. The question, of which way a system of linearly independent vectors can become dependent, forms the link between topology and $GL(N, \mathbb{C})$ (and $GL(N, \mathbb{R})$).

[3] in comparison with the original proof which employed essentially the tools of the modern calculus of variations due to Marston Morse; see, e.g. [**Mil**, p.124-132].

3. The Ring of Vector Bundles

Let X be a compact topological space. We consider the abelian semigroup $\mathrm{Vect}(X)$ (defined in the Appendix) of isomorphism classes of complex vector bundles over X. If X consists of a single point, then $\mathrm{Vect}(X) \cong \mathbb{Z}$. Now we generalize the construction which one uses to go from the semi-group $\mathbb{Z}_+$ to the group $\mathbb{Z}$, in such a way that we can assign a group $K(X)$ to the semigroup $\mathrm{Vect}(X)$.

THEOREM 11.3. *Each abelian semi-group A (with zero element) yields in canonical way an abelian group $B := A \times A/ \sim$ and a semi-group homomorphism $g : A \to B$ induced by $a \to (a, 0)$. Here $\sim$ is the equivalence relation on $A \times A$ defined by*

$$\begin{aligned} &(a_1, a_2) \sim (a_1', a_2') \\ &\Leftrightarrow \exists a, a' \in A \textit{ such that } (a_1 \oplus a, a_2 \oplus a) = (a_1' \oplus a', a_2' \oplus a')\,. \end{aligned}$$

PROOF. Let $\Delta : A \to A \times A$ be the diagonal homomorphism $a \to (a, a)$ of semi-groups. Then B consists of the cosets of $\Delta(A)$ in $A \times A$; i.e., $B = \{(a_1, a_2) + \Delta(A) : a_i \in A,\ i = 1, 2\}$, where

$$(a_1, a_2) + \Delta(A) := \{(a_1 \oplus a, a_2 \oplus a) : a \in A\}\,.$$

B is a quotient semi-group in which there is an inverse for each element given by

$$-(a_1, a_2) + \Delta(A) = (a_2, a_1) + \Delta(A).$$

Thus, B is a group. In this notation, the semi-group homomorphism is given by

$$\phi(a) := (a, 0) + \Delta(A).$$

□

REMARK 11.4. It is advisable to go through the definition of B carefully as we did in the proof, since the intuition one gains in the transition from $\mathbb{Z}_+$ to $\mathbb{Z}$ is partly deceptive: Namely, a semi-group is not always embedded in a group. (The cancellation rule must already hold in the semi-group). The natural homomorphism $\phi : A \to B$ is not necessarily injective; see Exercise 11.7 below.

REMARK 11.5. In a certain sense, B is the "best possible" group that can be made from the semi-group A. More precisely, the following universal property of ϕ holds: Every semi-group homomorphism $h : A \to G$ from A to an arbitrary group G can be "factored" through ϕ in exactly one way; i.e., there is exactly one group homomorphism $h' : B \to G$ such that the adjacent diagram

$$\begin{array}{ccc} A & \stackrel{\phi}{\to} & B \\ & h \searrow & \downarrow h' \\ & & G \end{array}$$

is commutative. The universal property is a generalization of the trivial observation that ϕ becomes an isomorphism, if A is already a group, and, conversely, the latter follows from the functorial property of the assignment $A \mapsto (B, \phi)$ on the category of semi-groups; see [**Ati67a**, p.42f]. Alternatively, B can be defined as a quotient group FA/RA, where FA is the "free abelian group" on A (which consists of all finite linear combinations of elements of A with coefficients in $\mathbb{Z}$), and RA is the subgroup of FA generated by the subset

$$\{1\,(a_1 \oplus a_2) + (-1)\,a_1 + (-1)\,a_2\}\,.$$

The homomorphism $\phi : A \to FA/RA$ is defined in the natural way and fulfills the homomorphism condition $\phi\,(a_1 \oplus a_2) = \phi\,(a_1) + \phi\,(a_2)$, since $\phi\,(a_1 \oplus a_2) - \phi\,(a_1) - \phi\,(a_2)$ is represented by $1\,(a_1 \oplus a_2) + (-1)\,a_1 + (-1)\,a_2$. That ϕ is the "universal solution of the factorization problem" is shown as follows: The uniqueness of h' is clear, since "$h'(\phi(a)) = h(a)$ for $a \in A$" implies that h' is already given on a set of generators of FA/RA, whence h' itself is uniquely defined. From $h'\,(\phi\,(a_1 \oplus a_2)) - h'\,(\phi\,(a_1)) - h'\,(\phi\,(a_2)) = 0$, it follows that h' is a group homomorphism. From the universal property, it follows easily that the two methods of group construction are equivalent, and in particular that B and FA/RA are isomorphic.

Now let $A := \mathrm{Vect}(X)$, X compact. We denote the associated abelian group B by $K(X)$. Then, for each vector bundle E over X, we obtain (by means of ϕ) an element $[E] \in K(X)$, and every element of $K(X)$ be written as a linear combination of such elements; see also Exercise 11.9a below. For $\left[\mathbb{C}_X^N\right]$ we also simply write N.

Exercise 11.6. In the formalism developed here, describe the canonical extension of subtraction $\delta : \mathbb{Z}_+ \times \mathbb{Z}_+ \to \mathbb{Z}$ to the "difference-bundle construction" $\mathrm{Vect}(X) \times \mathrm{Vect}(X) \to K(X)$. First show $K(X) \cong \mathbb{Z}$, if X is a point.

Exercise 11.7. Show that the cancellation rule does not always hold in $\mathrm{Vect}(X)$. [Hint: First illustrate with the two real bundles TS^2 and $\mathbb{R}^2_{S^2}$ over the 2-sphere. These are not isomorphic (see Exercise 11.23, p. 284), but forming the direct sum of each with the trivial line bundle $\mathbb{R}_{S^2}$, we arrive at isomorphic bundles (with the tangent bundle TS^2 one thinks of the direct sum with the normal bundle NS^2 of the canonical embedding of S^2 in $\mathbb{R}^3$). In general search for a nontrivial vector bundle that becomes trivial when a trivial bundle is added to it. A detailed discussion of special "cancellation type" rules can be found in [**Hu**, Ch.8]; for example, the "Uniqueness Theorem for Vector Bundles" says that trivial bundles over manifolds of dimension n may be cancelled when the other summand has fiber dimension $\geq n/2$.]

Exercise 11.8. a) Show that each element of $K(X)$ can be written in form $[E] - N$, where $E \in \mathrm{Vect}(X)$ and $N \in \mathbb{N}$.
b) Show that two vector bundles E and F define the same element of $K(X)$ (i.e., $[E] = [F]$ exactly when $E \oplus \mathbb{C}_X^N = F \oplus \mathbb{C}_X^N$, for some N.
c) One says that the bundles E and F are **stably-equivalent** when there are natural numbers M and N such that

$$E \oplus \mathbb{C}_X^N \cong F \oplus \mathbb{C}_X^M.$$

Show that the set $I(X)$ of stable equivalence classes form a group relative to the operation of direct sum. [Hint for c): See Appendix, Exercise 11.8, p. 275.]

Exercise 11.9. a) Show that, by means of the tensor product $\otimes$ for vector bundles (see Appendix, Exercise B.4, p. 679), a multiplicative structure for $K(X)$ is furnished, making it a commutative ring with unit $[\mathbb{C}_X]$.
b) Show that each continuous map $f : Y \to X$ induces a ring homomorphism $f^* : K\,(X) \to K(Y)$ which only depends on the homotopy class of f.
c) Let $i : Y \to X$ be the inclusion of a closed subset Y of X. Set $K(X, Y) :=$

$\mathrm{Ker}(K(X/Y) \to K(Y/Y))$, where X/Y is the space obtained from X when Y is "collapsed" to a point $\{Y/Y\}$.
(i) Let Y consist only of a single point x_0. Show that the group $K(X,Y)$ forms an ideal in $K(X)$, and that $K(X)$ splits into a direct sum

$$K(X) \cong K(X,x_0) \oplus K(x_0) \cong K(X,x_0) \oplus \mathbb{Z} = \widetilde{K}(X) \oplus \mathbb{Z},$$

where $\widetilde{K}(X) := K(X,x_0)$, the "essential" part of $K(X)$, is isomorphic to $I(X)$.
(ii) In general, define a natural map $j^* : K(X,Y) \to K(X)$, and show that the short sequence $K(X,Y) \xrightarrow{j^*} K(X) \xrightarrow{i^*} K(Y)$ is exact.
[Hint for a): Use the "universal property" (Remark 2) to factorize $\mathrm{Vect}(X) \times \mathrm{Vect}(X) \to K(X)$ through $K(X) \times K(X)$. For b): See Appendix, Theorem B.8, p. 682; in particular, $K(X) \cong K(Y)$ when X and Y are homotopy equivalent. For c): Work with the retraction $r : X \to x_0$, for the splitting in (i). See Exercise 11.8 for $I(X)$. First define j^* in (ii) more generally, when $f : (X',Y') \to (X,Y)$ is a map of pairs of spaces (i.e., $j : X' \to X$ continuous with $j(Y') \subseteq Y$). Then set $X' = X$ and $Y' = Y$. To check that $\mathrm{Im}(j^*)c = \mathrm{Ker}(i^*)$, factor

$$\begin{array}{ccc} (Y,\phi) & \overset{j\circ i}{\to} & (Y,X) \\ & \searrow & \downarrow \\ & & (Y,Y) \end{array}$$

and note that $K(Y,Y) = 0$. To prove the other direction, work with the representation as in Exercise 11.8a. See also [**Ati67a**, p.69 f].]

THEOREM 11.10. *Let X be a compact space and $[X,\mathcal{F}]$ the set of homotopy classes of continuous maps $T : X \to \mathcal{F}$, where $\mathcal{F}$ is the space of Fredholm operators in a Hilbert space H. The construction of index bundles (see Section 3.3.7 above) induces a bijective map* $\mathrm{index} : [X,\mathcal{F}] \to K(X)$ *under which composition in $\mathcal{F}$ and addition in $K(X)$ correspond, as do adjoints in $\mathcal{F}$ and negatives in $K(X)$.*

PROOF. See Theorem 3.40, p. 89. □

1. **K-Theory and Functional Analysis**. In the proof of Theorem 11.3 the subsequent remark, we learned two different constructions of the group $K(X)$ and of these the first is probably the most natural in connection with functional analysis (as in Theorem 11.10) . On the other hand, the second construction immediately yields the "universal property" which historically motivated this formal group construction in the papers of Claude Chevalley and Alexandre Grothendieck concerned with algebraic geometry. It arose as a tool for the study of problems in which functions are involved which are additive on a semigroup with integral values. This also explain the relevance of $K(X)$ for our index problem of elliptic operators:

In Part II, we associated with each elliptic pseudo-differential operator $P : C^\infty(E) \to C^\infty(F)$ (E and F complex C^∞ bundles over the closed, C^∞ Riemannian n-manifold X) its symbol $\sigma(P) \in \mathrm{Iso}_{SX}(E,F)$, and we proved that $\mathrm{index}\, P$ only depends on the homotopy type of $\sigma(P)$. Now let $S'X := B^+X \cup_{SX} B^-X$ be the n-sphere bundle over X which arises by gluing two copies B^+X and B^-X of the covariant unit-ball bundle $BX := \{(x,\xi) : x \in X \text{ and } \xi \in T_x^*X \text{ with } |\xi| \le 1\}$ along their common boundary SX. We lift E over B^+X and F over B^-X and glue them (see Appendix, Exercise B.7, p. 681) over SX by means of $\sigma(P)$. This way we obtain a vector bundle on $S'X$ whose isomorphism class $[E \overset{\sigma(P)}{\to} F]$ only depends on the homotopy type of $\sigma(P)$. Conversely, the space of pseudo-differential operators is so rich that $\sigma(P)$ has any desired homotopy type for suitable P. Because of the special form of $S'X$, we therefore obtain all isomorphism classes of vector bundles over $S'X$ in this fashion. Thus the theory of elliptic equations yields a semigroup homomorphism index : $\mathrm{Vect}(S'X) \to \mathbb{Z}$ which fits into the following diagram

$$\begin{array}{ccccc} \mathrm{Ell}\,(X) & \to & \mathrm{Vect}(S'X) & \to & K\,(S'X) \\ & \underset{\searrow}{\text{index}} & \downarrow^{\text{index}} & \underset{\swarrow}{K\text{- index}} & \\ & & \mathbb{Z} & & \end{array}$$

Here Ell(X) is the class of elliptic pseudo-differential operators on the closed manifold X. Thus the universal property of $K\,(S'X)$ guarantees the existence and uniqueness of a K-index, which makes the diagram commutative and which being a group homomorphism is easier to analyze especially after more is known about the group $K\,(S'X)$.[4]

2. K-theory and cohomology. Exercise 11.8 says that K is a contravariant functor on the category of compact topological spaces and continuous maps into the category of commutative rings (with identity) and ring homomorphisms, and this functor bears great resemblance (also in aspects not explained here) with the cohomology functor H^*; see [**ES**, p.13 f]. The "characteristic classes" of vector bundles (see [**Hi66a**]) yield a variety of interesting operations $K \to H^*$, and one can show that in fact $K(X) \otimes_{\mathbb{Z}} \mathbb{R} \cong H^{\text{even}}(X;\mathbb{R})$; see [**Ati67a**].

4. K-Theory with Compact Support

Until now, we have only defined $K(X)$ for compact X. For locally compact X, we now set $K(X) := K(X^+,+) = \mathrm{Ker}(K(X) \overset{i^*}{\to} K(+))$, where $X^+ = X \cup \{+\}$ is the 1-point compactification of X (by the addition of the point $+$, where $i : \{+\} \to X$ is the inclusion of this point). For compact X, this definition brings nothing new. Alternatively, $K(X)$ can be expressed in terms of "complexes" of vector bundles - see [**Ati68a**, p.489 ff]; e.g., the elements of $K(Y - Z) = K(Y, Z)$ can be taken to be equivalence classes of isomorphisms $\sigma : E|_Z \cong F|_Z$, where E and F are (complex) vector bundles over the compact set Y with closed subset Z. (One thinks of the

[4] Presently, three different ways are known for associating with an elliptic operator P a K-theoretic object via its symbol. The construction presented here which ends up in $K(B^+X \cup_{SX} B^-X)$ is the most conceptual, since $[E \overset{\sigma(P)}{\to} F] \in K(B^+X \cup_{SX} B^-X)$ can be represented directly by a vector bundle on $B^+X \cup_{SX} B^-X$. The other constructions yield objects in $K(BX, SX)$ or (equivalently, see the following section) in $K(TX)$ and basically amount to forming the difference class $[E \overset{\sigma(P)}{\to} F] - [F]$. The advantage of this "non-conceptual" construction of a difference bundle (for details see Chapters 12 and 13 below) rests on the fact that the object $[E \overset{\sigma(P)}{\to} F]$ contains too much useless information on the specific form of the vector bundles E and F which is completely irrelevant to the index problem. For example, one can make E (or F) trivial by adding a vector bundle V (Appendix, Exercise B.10, p.683) while the index of $P \oplus \mathrm{Id}_V : C^\infty(E \oplus V) \to C^\infty(E \oplus V)$ does not change. More precisely: By evaluating the row-exact commutative diagram

$$\begin{array}{ccccccc} \leftarrow & K\left(B^+X\right) & \leftarrow & K\left(B^+X \cup_{SX} B^-X\right) & \leftarrow & K\left(B^+X \cup_{SX} B^-X, B^-X\right) & \leftarrow \\ & \downarrow\cong & \swarrow\nearrow & & & \downarrow\cong & \\ & K\,(X) & & & & K\left(B^+X, SX\right), & \end{array}$$

where $\swarrow \circ \nearrow$ is the identity on $K(X)$, one finds that $K(B^+X \cup_{SX} B^-X) \cong K(X) \oplus K(BX, SX)$, whereby the first summand is irrelevant for the index problem and the second one is best dealt with in the non-relative form $K(TX)$ in the framework of "K-theory with compact support."

symbol of an elliptic operator over the manifold X and sets $Y := BX$ and $Z := SX$, where $Y - Z$ is then diffeomorphic to the full cotangent bundle T^*X.)

EXERCISE 11.11. a) Verify that $K(X)$ is a ring (without unit element) when X is non-compact. b) Show functoriality for proper maps $f : Y \to X$ these are the maps which can be continuously extended to Y^+.
[Hint for b): One may also define f to be proper exactly when $f^{-1}(K)$ is compact for all compact subsets $K \subseteq X$. From this comes the notion of "K-Theory with compact support" - see also [**ES**, p.5 and 269 ff]. In particular, each homeomorphism is proper, and we have $K(X) \cong K(Y)$ for homeomorphic X and Y, and $f^* = \mathrm{Id}$, if $Y = X$ and f is homotopic (within the class of homeomorphisms) to the identity. On the other hand, the mere homotopy type of X does not ' determine $K(X)$. (Example: $K(\mathbb{R}) \ncong K(point)$).]

THEOREM 11.12. *If X and Y are locally compact spaces, then (in addition to the ring structures of $K(X)$ and $K(Y)$) there is an* ***outer product*** $\boxtimes : K(X) \otimes K(Y) \to K(X \times Y)$.

REMARK 11.13. The outer product admits a particularly simple and natural definition, if one adopts the above introduction of $K(X)$ via complexes, and forms the tensor product of complexes; see [**AS68a**, p.490] and also the outer tensor product for elliptic operators in Exercise 9.13b, (p. 238) and for matrix-valued functions in Section 11.2 above.

PROOF. 1) If X and Y are compact, then $\boxtimes$ is defined by forming the vector bundle $E \boxtimes F$ over $X \times Y$, where E is a vector bundle over X, F is over Y, and $E \boxtimes F$ has fiber $E_x \otimes F_y$ over (x, y).
2) In order to carry this definition over to locally compact X and Y, we prove the exactness of the short sequence

$$0 \to K(X \times Y) \to K(X^+ \times Y^+) \to K(X^+) \oplus K(Y^+), \tag{11.1}$$

whereby $K(X \times Y)$ is identified with the subgroup of $K(X^+ \times Y^+)$ which vanishes on the "axes" X^+ and Y^+. For this, we begin with the short exact sequence

$$K(A, B) \xrightarrow{j^*} K(A) \xrightarrow{i^*} K(B) \tag{11.2}$$

from Exercise 11.9c (above) for compact topological spaces A and B with $i : B \to A$ and for the case where B is a "retract" of A; i.e., there is a continuous map $r : A \to B$ which is the identity on B. Then $ri = \mathrm{Id}$ on B and $i^*r^* = \mathrm{Id}$ on $K(B)$, where $r^* : K(B) \to K(A)$ is naturally defined; thus, we see that i^* is surjective and r^* is injective. Furthermore, we obtain (prove!) a $g^* : K(A) \to K(A, B)$ with $g^*j^* = \mathrm{Id}$, whence j^* is injective. One says: The sequence "splits" (see [**ES**, p.229 f]), and we obtain a decomposition

$$K(A) \cong K(A, B) \oplus K(B). \tag{11.3}$$

From (11.3), we get (11.1): one time we set

$$A := X^+ \times Y^+ \text{ and } B := X^+ \times \{+\} \tag{11.4a}$$

and then

$$A := \left(X^+ \times Y^+\right)/X^+ \text{ and } B := Y^+. \tag{11.5}$$

Since B is a retract of A in both cases, (11.2) splits and we obtain the formulas

$$K(X^+ \times Y^+) \cong K(X^+) \oplus K(X^+ \times Y^+, X^+) \tag{11.6}$$

and

$$K((X^+ \times Y^+)/X^+) \cong K(Y^+) \oplus K((X^+ \times Y^+)/X^+, Y^+), \tag{11.7}$$

from which the desired splitting

$$K(X^+ \times Y^+) \cong K(X^+) \oplus K(Y^+) \oplus K(X \times Y)$$

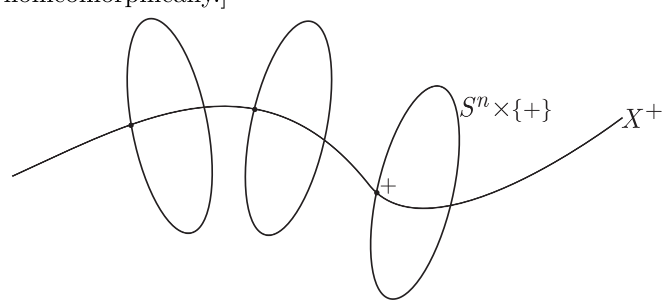
follows because $K((X^+ \times Y^+)/X^+, Y^+) \cong K(X^+ \times Y^+, +) = K(X \times Y)$. One can check that the splitting is compatible with the naturally defined arrows in (11.1).
3) Now let $x \in K(X) \subseteq K(X^+)$ and $y \in K(Y) \subseteq K(Y^+)$. Then $x \boxtimes y \in K(X^+ \times Y^+)$ is well defined by 1). Actually, $x \boxtimes y$ can be regarded also as an element of $K(X \times Y)$ by (11.1), since $i^*(x \boxtimes y) = 0$, where $i : X^+ \hookrightarrow X^+ \times Y^+$ is the canonical inclusion (and correspondingly for $Y^+ \hookrightarrow X^+ \times Y^+$). For a proof of this, we write (Exercise 11.8a, p. 275) $x = [E] - N$ and $y = [F] - M$, where $E \in \mathrm{Vect}(X)$ and $F \in \mathrm{Vect}(Y)$, $N, M \in \mathbb{Z}_+$; we then have

$$x \boxtimes y = [E \boxtimes F] - [N \boxtimes F] - [E \boxtimes M] + [N \boxtimes M], \text{ and so}$$
$$i^*(x \boxtimes y) = [E \otimes F_+] - [E \otimes M] - [N \otimes F_+] + [N \otimes M] = 0,$$

since the fiber $F_+ \cong \mathbb{C}^M$. (Beware: M denotes the trivial M-dimensional bundle over Y in the first formula, but in the second, it is only vector space $\mathbb{C}^M$.) □

An important example of a locally compact space is furnished by Euclidean space $\mathbb{R}^n$, whose 1-point compactification is the n-sphere S^n; by definition $K(S^n) \cong K(\mathbb{R}^n) \oplus \mathbb{Z}$ holds, where the second summand is $K(\{+\}) := \mathbb{Z}$, given by the dimension of the vector bundle. This shows that $K(\mathbb{R}^n)$ is actually the interesting part of $K(S^n)$.

Exercise 11.14. For an arbitrary paracompact X, go through the splitting $K(S^n \times X) \cong K(\mathbb{R}^n \times X) \oplus K(X)$.
[Hint: Work with the sequence (11.2) from step 2 of the preceding proof, where $B := X^+$ is a retract of $A := S^n \times X^+/S^n$. Note that $A = (S \times X)^+$ and $A/B = (\mathbb{R}^n \times X)^+$, homeomorphically.]



The importance of K-theory with compact support stems for one thing from the fact that applications frequently involve non-compact, but locally compact, spaces such as Euclidean space or tangent spaces. Of course it is possible, without undue difficulties, to avoid non-compact spaces altogether (as with the passage from the tangent bundle, TX to the double ball bundle $B^+X \cup B^-X$ in the Section 11.3 above) - as artificial as this construction may appear. However, the splitting of Exercise 11.14 (also see the footnote p. 277) makes the locally compact formalism genuinely simpler and perhaps conceptually clearer. In the following proof of the Bott Periodicity Theorem which we adopt from [**Ati69**], we will therefore stay in the category of locally compact spaces: We already know that $K(\mathbb{R}^0) = K(point) \cong \mathbb{Z}$ and we obtain $K(\mathbb{R}^1) = 0$, since all complex vector bundles on the circle are trivial ($\mathrm{GL}(N, \mathbb{C})$ is connected..., see the gluing classification in the Appendix, Theorem B.8, p. 682). With some pains (as well as some projective geometry, see Appendix, Exercise B.2a, p. 678, and [**Ati69**, p.46 f]), we could still compute $K(\mathbb{R}^2) \cong \mathbb{Z}$. How does it go?

The theorem which we will prove says that the sequence of these K-groups continues. Thus, $K(\mathbb{R}^3) = 0$, $K(\mathbb{R}^4) \cong \mathbb{Z}$, $K(\mathbb{R}^5) \cong 0$, etc., and that, more generally, for each locally compact X, there is a natural isomorphism $K(\mathbb{R}^2 \times X) \cong K(X)$.

5. Proof of the Periodicity Theorem of R. Bott

In Chapter 4, we became acquainted with Wiener-Hopf operators, and as a generalization of the discrete (and there rather trivial) index theorem of Israil Gohberg and Mark Krein (Theorem 4.4, p. 126), we gave a construction $(V, f) \mapsto F \mapsto \operatorname{index} F$. In the pair (V, f), V is a vector bundle over a compact parameter space X, and f is an automorphism of π^*V, where $\pi : S^1 \times X \to X$ is the projection. For $z \in S^1$ and $x \in X$, $f(z, x)$ is then an automorphism of the fiber V_x, and it depends continuously on x and z. Now, $F : X \to \mathcal{F}$ is the associated family of Fredholm operators (after the Wiener-Hopf recipe, formed on certain Hilbert spaces of "half-space functions") and $\operatorname{index} F \in K(X)$ is the "index bundle" of F, which is defined (in the special case that the kernels of all of the Fredholm operators F_x have constant dimension) as $[\operatorname{Ker} F] - [\operatorname{Coker} F]$. Moreover, we have seen that $\operatorname{index} F$ only depends on the homotopy class of f.

Exercise 11.15. a) From the pair (V, f), how can one construct a vector bundle over $S^2 \times X$ that only depends on V and the homotopy class of f?
b) Show that every $E \in \operatorname{Vect}(S^2 \times X)$ can be obtained in this way.
[Hint for a): Decompose S^2 into the two hemispheres B^+ and B^- with $B^+ \cap B^- = S^1$, and form the bundle $(\pi^+)^*V \cup_f (\pi^-)^*$ by means of the clutching construction (see Appendix, Exercise B.7, p. 681), where $\pi^\pm : B \times X \to X$ is the projection.
For b): Argue as in the proof of Theorem B.8 (p. 682) of the Appendix, where X consists only of a single point. The parameter space plays only a subordinate role, and so the proof actually carries over. It is convenient to normalize the map f that one obtains so that $f(1, x)$ is the identity on V_x.]

Theorem 11.16. *Let X be locally compact. Then a homomorphism $\alpha : K(\mathbb{R}^2 \times X) \to K(X)$ can be defined (here, by the index of a Wiener-Hopf family of Fredholm operators; see also Exercise 11.19) with the following properties:*
(i) α is functorial in X.
(ii) If Y is another locally compact space, then we have the following multiplication rule, expressed by the commutative diagram:

$$\begin{array}{ccc} K(\mathbb{R}^2 \times X) \otimes K(Y) & \xrightarrow{t'} & K(\mathbb{R}^2 \times X \times Y) \\ \downarrow \alpha_X \otimes \mathrm{Id} & & \downarrow \alpha_{X \times Y} \\ K(X) \otimes K(Y) & \xrightarrow{t} & K(X \times Y), \end{array}$$

where t and t' are the outer tensor products $\boxtimes$ defined in Theorem 11.12
(iii) $\alpha(b) = 1$, where b is the Bott class, a kind of basis in $K(\mathbb{R}^2)$ that we define

by $b : [E_{-1}] - [E_0] \in K(S^2)$*, where the line bundle defined over* S^2 *by means of the clutching function,* $f(z) = z$*; see Appendix, Theorem B.8 (p. 682) or the preceding Exercise 11.15a. Since* E_{-1} *and* E_0 *have the same dimension (one),* b *lies in* $K(S^2, \{+\}) = K(\mathbb{R}^2)$*. Thus, here we set* $X = \{point\}$*, and identify* $K(point)$ *with* $\mathbb{Z}$ *as in Exercise 11.6.*

Proof. 1) We start with X compact. By Theorem 4.4 (p. 126) and Exercise 4.9 (p. 128), in conjunction with the preceding Exercise 11.15, we can (for each vector bundle E over $S^2 \times X$) go through a construction $E \mapsto (V, f) \mapsto F \mapsto \operatorname{Index} F$. In this way, a semi-group homomorphism $\operatorname{Vect}(S \times X) \to K(X)$ is defined, which can be extended (see 11.5, p. 274) uniquely to a group homomorphism $\alpha' : K(S^2 \times X) \to K(X)$. The restriction of α' to $K(\mathbb{R}^2 \times V)$ (which may be regarded as a subgroup of $K(S^2 \times X)$ by Exercise 11.14) then provides a homomorphism $\alpha : K(\mathbb{R}^2 \times X) \to K(X)$. The "functoriality" of α means that, for each element $u \in K(\mathbb{R}^2 \times X)$ and each continuous map $g : X' \to X$ (where X' is another compact space), we have

$$\alpha_{X'}\left((g \times \mathrm{Id}_{\mathbb{R}^2})^* u\right) = g^* \alpha_X(u),$$

which is clear from the functorial nature of index bundles (Exercise 3.37, p. 88).

2) We apply this definition for $X = \{point\}$ and calculate $\alpha(b) = \alpha'[E_{-1}] - \alpha'[E_0]$. By construction, we have $\alpha'[E_m] = -m$ ($m \in \mathbb{Z}$), since by Theorem 4.4 (p. 126) we have the assignments

$$E_m \overset{\text{clutch}}{\longmapsto} (\mathbb{C}, z^m) \overset{\text{Wiener-Hopf}}{\longmapsto} T_{z^m} \overset{\text{index}}{\longmapsto} -m,$$

Hence $\alpha(b) = 1$, and so (iii) is fulfilled.

3) To prove the multiplication rule (ii) – first for X, Y compact – we must consider the difference $t(\alpha_X \otimes \mathrm{Id})(u \otimes v) - \alpha_{X \times Y}(t(u \otimes v))$ for $u \in K(\mathbb{R}^2 \times X)$ and $v \in K(Y)$; without loss of generality, we may assume $v = 1$ (the class of the trivial line bundle $\mathbb{C}_Y$ over Y), since all of the maps arising here are $K(Y)$-module homomorphisms. By the functoriality of α, the difference $\pi^* \alpha_X(u) - \alpha_{X \times Y}(\pi^* u)$ vanishes, as needed. (Here $\pi : X \times Y \to X$ is the projection.)

4) The definition of $u : K(E \times X) \to K(X)$ for locally compact X carries over just as in the proof of (ii) in the compact case via the one-point compactification with the decomposition above in the proof of Theorem 11.12 and in Exercise 11.14. □

Theorem 11.17. ***(Periodicity Theorem)****. For each locally compact space* X*, we have that* $\alpha : K(\mathbb{R}^2 \times X) \to K(X)$ *is an isomorphism, whose inverse* $\beta : K(X) \to K(\mathbb{R}^2 \times X)$ *is given via outer multiplication* $x \mapsto \beta(x) := b \boxtimes x$ *by the Bott class* b*.*

Proof. The proof follows by repeated use of the multiplicative property of α expressed in Theorem 11.16(ii). In the following, for short we write only xy for the outer product $x \boxtimes y$ or $t(x \otimes y)$.

$\alpha\beta = \mathrm{Id}$: For the proof here, in (ii) substitute (point) for X and X for Y; then we have $\alpha\beta(x) = (\alpha(b))(x)$ for each $x \in K(X)$ by (ii), whence $\alpha\beta = \mathrm{Id}$, since $\alpha(b) = 1$ by (iii).

$\beta\alpha = \mathrm{Id}$: Let $u \in K(\mathbb{R}^2 \times X)$. We want to show that $\beta\alpha u := b(\alpha(u)) = u$, or equivalently (if multiplying by b from the right) $(a(u))b = \tilde{u}$, where $\tilde{u} := \rho^* u$ and $\rho : X \times \mathbb{R}^2 \to \mathbb{R}^2 \times X$ switches factors. From (ii) with $\mathbb{R}^2$ for Y, it follows that

$(\alpha(u))(b)$ equals $\alpha(ub)$. Now comes the trick, through which the proof of $\beta\alpha = \mathrm{Id}$ can be reduced to the rather banal fact $\alpha\beta = \mathrm{Id}$ already proven: On $K(\mathbb{R}^2 \times X \times \mathbb{R}^2)$, where the element ub lies, the map τ^*, which is the "lifting" along the switching map

$$\tau : \mathbb{R}^2 \times X \times \mathbb{R}^2 \to \mathbb{R}^2 \times X \times \mathbb{R}^2, \text{ where } \tau\,(a, b, c) = (c, b, a)\,,$$

is the identity, since τ is homotopic to the identity on $\mathbb{R}^2 \times X \times \mathbb{R}^2$ through homeomorphisms. (See the hint for Exercise 11.11b.) Namely, on $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$, τ is given by the matrix

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

which has determinant $+1$ and hence (one thinks of the transition to Jordan normal form) lies in the same connected component of $\mathrm{GL}(4, \mathbb{R})$ as the identity. Hence, we have

$$(\alpha(u))b = \alpha(ub) = \alpha(\tau^*(ub)) = \alpha(b\tilde{u}) = \alpha\beta\tilde{u} = \tilde{u}.$$

□

Exercise 11.18. Show the following consequences of Theorem 11.17:
a) $K(X \times S^2) \cong K(X) \otimes K(S^2)$ for X compact.
b) $K(X \times \mathbb{R}^n) = \begin{cases} \mathbb{Z} & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd.} \end{cases}$
c) $K(S^n) = \begin{cases} \mathbb{Z} \oplus \mathbb{Z} & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd.} \end{cases}$
d) $N \geq n/2 \Rightarrow \pi_{n-1}(\mathrm{GL}(N, \mathbb{C})) = \begin{cases} \mathbb{Z} & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd.} \end{cases}$.
[Hint: While a), b) and c) follow directly from Theorem 11.17 with Exercise 11.14, the derivation of d) requires two further considerations. First, for a compact manifold X of dimension $n-1$, we have that for each $E \in Vect_N(X)$ with $N \geq m := [n/2 - 1]$, there is an $F \in \mathrm{Vect}(X)$ such that $E = F \oplus \mathbb{C}_X^{N-m}$, i.e., each vector bundle over X is stably equivalent (see Exercise 11.8c, p. 275) to a vector bundle of fiber dimension m. This is the "basis theorem" for vector bundles. The "uniqueness theorem" says that (under the same assumptions) stably-equivalent vector bundle fiber dimension $N \geq m+1$ are isomorphic. For the proofs of these two lemmas (e.g., see [**Hu**, Ch.8]) one needs some homotopy theory. The rest is trivial, since we can now represent the group $I\,(X)$ of stable-equivalence classes of bundles for $N \geq n/2$ by $\mathrm{Vect}(X)$. In particular (see Exercise 11.9c, p. 275), $\mathrm{Vect}_N(S^n) \cong I(S^n) \cong K(\mathbb{R}^n)$. The "classical" form of the periodicity theorem now follows, since $\pi_{n-1}(\mathrm{GL}(N, \mathbb{C})) \cong \mathrm{Vect}_N(S^n)$ by Theorem B.8 (p. 682) of the Appendix.]

Exercise 11.19. As an alternative to Theorem 11.16, give a construction of the isomorphism $\alpha : K(\mathbb{R}^2 \times X) \to K(X)$ by means of a family of elliptic boundary-value problems.
[Hint: For $X = \{point\}$ and $f : S^1 \to \mathrm{GL}(N, \mathbb{C})$, consider (over the disk $|z| < 1$) the "transmission operator" (see Exercise 5.16, p. 151)

$$A_f : (u, v) \mapsto \left(\frac{\partial u}{\partial \bar{z}}, \frac{\partial v}{\partial z}, fu|_{S^1} - v|_{S^1} \right),$$

where u, v are N-tuples of complex-valued functions. By the same recipe, one can also treat families of such boundary-value problems which are parametrized over a space X; see [**Ati68a**, p.118-122].]

REMARK 11.20. The connection with the construction of α via Wiener-Hopf operators lies, roughly speaking, in the "Poisson principle" (i.e., in the Agranovich-Dynkin formula in Exercise 10.33, p. 260), by which boundary-value problems can be translated into problems on the boundary. Namely, extend (as in Exercise 8.23, p. 223) the discrete Wiener-Hopf operator T_f to a pseudo-differential operator of order 0 on the circle S^1 (via the identity on the basis elements of the form z^m with $m < 0$), which we still denote by T_f. Then, we have that

$$S_f : (u, v) \mapsto \left(\frac{\partial u}{\partial \bar{z}}, \frac{\partial v}{\partial z}, T_f \left(zu|_{S^1} - v|_{S^1} \right) \right)$$

is an elliptic problem with pseudo-differential boundary conditions; it has the same kernel as the transmission problem A_{zf} and isomorphic cokernel. On the other hand, the operator S_f is formed directly by the composition of $A_{z\,\mathrm{Id}}$ with the "primitive boundary-value problem" $(u, v, w) \to (u, v, T_f w)$. Here Id is the constant function that assigns the identity in $\mathrm{GL}(N, \mathbb{C})$ to each $z \in S^1$. While Exercise 5.16 says that $\operatorname{index} A_{\mathrm{Id}} = 1$, one finds $\operatorname{index} A_{z\,\mathrm{Id}} = 0$, whence $\operatorname{index} A_{zf} = \operatorname{index} S_f = \operatorname{index} T_f$. Within the Green algebra (see Chapter **??**), then induces the connection

$$\left[\begin{array}{cc|c} \frac{\partial}{\partial \bar{z}} & 0 & 0 \\ 0 & \frac{\partial}{\partial z} & 0 \\ \hline M_{zf} \circ r & \mathrm{Id} \circ r & 0 \end{array}\right] \overset{S_f}{\leftrightarrow} \left[\begin{array}{cc|c} \mathrm{Id} & 0 & 0 \\ 0 & \mathrm{Id} & 0 \\ \hline 0 & 0 & T_f \end{array}\right]$$

between a conventional elliptic system of partial differential equations of the first order over the disk B^2 with "ideally simple" boundary-value conditions which are formed via the restriction $r : C^\infty(B^2) \to C^\infty(S^1)$ and a trivial multiplication operator, and a "primitive boundary-value problem" that consists of a (somewhat complex) elliptic pseudo-differential operator only on the boundary. Actually, the "change of planes" from differential boundary value problems in the plane to pseudo-differential operators on the boundary line can be pushed further, and using a polynomial approximation to f, a purely algebraic definition of the homomorphism α can be given. In doing so, no Hilbert space theory and Fredholm operators are needed, just as in the wholly elementary, but in parts quite tedious, proof [**AB64b**]. A detailed discussion of the advantages and disadvantages of the different methods for the construction of α is in [**Ati68a**, p.131-136].

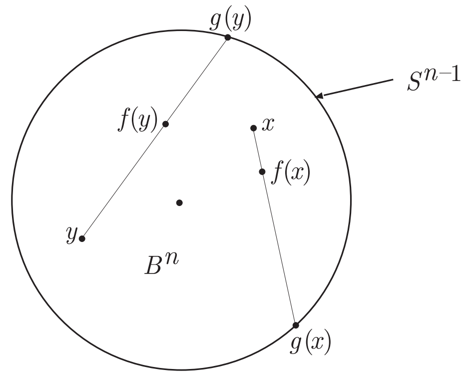
As a simple corollary of the periodicity theorem, we prove the classical fixed point theorem of topology.

THEOREM 11.21. *(L. E. J. Brouwer, 1911): Each continuous map f of the closed n-dimensional ball B^n into itself has a fixed point.*

PROOF. If $f(x) \neq x$ for all $x \in B^n$, a continuous map $g : B^n \to S^{n-1}$ is defined by

$$g(x) := (1 - \alpha(x))f(x) + \alpha(x)x,$$

where $\alpha(x) \geq 0$ is chosen such that $\|g(x)\| = 1$; g is the identity on S^{n-1}, and therefore a retraction of B^n to S^{n-1} (i.e., $g \circ i = \mathrm{Id}$, where $i : S^{n-1} \to B^n$ is inclusion). For odd n, we note that the composition



$$K(S^{n-1}) \xrightarrow{g^*} K(B^n) \xrightarrow{i^*} K(S^{n-1})$$

is the identity since $g \circ i = \mathrm{Id}$. But it cannot be the identity, since $K(S^{n-1}) \cong \mathbb{Z} \oplus \mathbb{Z}$ is not cyclic, whereas $i^*(K(B^n))$ is cyclic because $K(B^n) \cong \mathbb{Z}$. For even n, note that the corresponding composition of "suspensions" (see Appendix, Exercise B.7, p. 681):

$$K(S^n) \overset{(Sg)^*}{\to} K(SB^n) \overset{(Si)^*}{\to} K(S^n)$$

is the identity, and yet it cannot the identity since $K(S^{n-1}) \cong \mathbb{Z} \oplus \mathbb{Z}$ is not the image of $K(SB^n) \cong \mathbb{Z}$. □

Exercise 11.22. Define the **mapping degree**, $\deg f \in \mathbb{Z}$, of an arbitrary continuous map $f : S^n \to S^n$, and show that homotopic maps have the same mapping degree.
[Hint: Each homomorphism $h : \mathbb{Z} \to \mathbb{Z}$ clearly has a "degree", namely d, such that $h(m) = dm$ for all $m \in \mathbb{Z}$. Thus, for even n work with the subgroup $K(\mathbb{R}^n) \cong \mathbb{Z}$ of $K(S^n)$, which is carried into itself by f^*. For odd n, consider the suspension $Sf : S^{n+1} \to S^{n+1}$.]

Exercise 11.23. Show that the tangent bundle $T(S^n)$ is nontrivial, for even $n \geq 2$.
[Hint: More generally, each nowhere vanishing vector field v on S^n ($v \in C^\infty(TS^n)$) provides a homotopy between the identity and the antipodal map $A : S^n \to S^n$, while $\deg \mathrm{Id} \neq \deg A$ for even $n \geq 2$, contradicting Exercise 11.22.]

The preceding examples show that important notions and results of classical algebraic topology can be developed as well, and perhaps more quickly and more easily, on the basis of linear algebra via K-theory than they can by establishing, say, homology theory by means of simplicial theory.

A number of sharper results, such as the well-known converse by Hopf of Exercise 11.22 (equality of mapping degrees implies homotopy) in case $n \geq 2$, cannot be obtained by K-theoretic means but require deeper geometric considerations. At least according to Atiyah, the Periodicity Theorem is not only simpler than most major theorems of classical algebraic topology, but also more relevant for the index problem and, more generally, for many investigations of manifolds. Hereby, the "philosophy" is that the usual algebraic topology destroys the structures too much while K-theory "comparable to molecular biology" (Atiyah) searches for the essential macromolecules which make up the manifold. It is then clear that for a "comparison" of the topology of the intricate manifold with the "building blocks" one needs to first know the topology of the general linear group which comprises the transition functions and, more generally, the Periodicity Theorem in its K-theoretic form.

In the following chapters we will apply the Periodicity Theorem in various ways for computing the index of elliptic problems. Conversely, the index of special elliptic differential equations served to prove the Periodicity Theorem; see especially Theorem 11.16 and Exercise 11.19. This is no contradiction but an indication of how closely K-theory and index theory of linear elliptic oscillation equations are related both being linked by the catchwords "linear", "finite dimensional" and "deformation invariant".

CHAPTER 12

The Index Formula in the Euclidean Case

1. Index Formula and Bott Periodicity

In the preceding chapter, we used analytic tools (the Gohberg-Krein Index Theorem for Wiener-Hopf operators) to prove the Bott Periodicity Theorem. We will now use it to derive an index theorem for elliptic integral operators in $\mathbb{R}^n$. The basic idea is perhaps best described via homotopy theory:

An elliptic pseudo-differential operator of order k in $\mathbb{R}^n$ given in the form

$$(Pu)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle} p(x,\xi)\hat{u}(\xi)\, d\xi,$$

where u is a C^∞ function on $\mathbb{R}^n$ with compact support and values in $\mathbb{C}^N$, and the "amplitude" (see above Chapter 8) p is an $N \times N$ matrix-valued function with

$$\sigma(P)(x,\xi) := \lim_{\lambda\to\infty} \frac{p(x,\lambda\xi)}{\lambda^k} \in \mathrm{GL}(N,\mathbb{C}),\ \xi \neq 0.$$

Thus for fixed x, we have a continuous map

$$\sigma(P)(x,\cdot) : S^{n-1} \to \mathrm{GL}(N,\mathbb{C})$$

that has a well-defined degree for n even and N sufficiently large. This degree does not depend on x, because of continuity ($\mathbb{R}^n$ is connected) and it was denoted by $\deg(P)$ in Remark 10.19, p. 251.

In order to get interesting global problems, P is usually combined with $Nk/2$ boundary conditions to form an elliptic system in the sense of Chapter 10. However, this is only possible in the stated fashion if the "local index" $\deg(P)$ vanishes. A sort of extremely simple boundary condition lacking the topological and analytical difficult discussed in Sections 10.4 and **??** arises when we put $k = 0$ and $p(x,\xi) = \mathrm{Id}$ for x outside a compact subset K of $\mathbb{R}^n$. The class of elliptic pseudo-differential operators of order 0 in $\mathbb{R}^n$ with this property of "being equal to identity at infinity" was first investigated by Robert T. Seeley and will be denoted by $\mathrm{Ell}_c(\mathbb{R}^n)$. Obviously, see Exercise 12.1 below, every $P \in \mathrm{Ell}_c(\mathbb{R}^n)$ has a finite dimensional kernel and cokernel; therefore $\operatorname{index} P$ is well-defined. If furthermore $p(x,\xi) = \mathrm{Id}$ for $|x| \geq r$, r real, then $\sigma(P)(x,\xi) \in \mathrm{GL}(N,\mathbb{C})$ for $|x| + |\xi| \geq r$. In this way P defines a continuous mapping $S^{2n-1} \to \mathrm{GL}(N,\mathbb{C})$ where S^{2n-1} is the $(2n-1)$-sphere of radius r in $\mathbb{R}^{2n}$ (the (x,ξ) -space). Since $\operatorname{index} P = \operatorname{index}(P + \mathrm{Id})$ (Id the identity operator on functions), we may assume without loss of generality that $N \geq n$. By the homotopy theoretic form of the Bott Periodicity Theorem (Theorem 11.2, p. 269, or Exercise 11.18d, p. 282), we have $\pi_{2n-1}(\mathrm{GL}(N,\mathbb{C})) \cong \mathbb{Z}$. Thus we have three integral invariants:

index P	$\{$ the analytic index (defined in the sense of analytic function theory,
$\deg(\sigma(P)(\cdot,\cdot))$	$\{$ the topological index (defined via homotopy theory by the global behavior of $\sigma(P)$,
$\deg P$	$\{$ the local index (defined via homotopy for even n by the pointwise behavior of $\sigma(P)(x,\cdot)$.

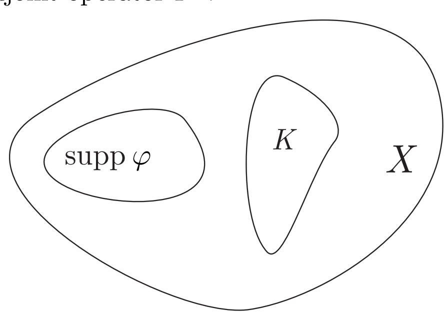
We have $\deg(P) = 0$ (which is trivial) and index $P = \pm\deg(\sigma(P)(\cdot,\cdot))$ (see Theorem 12.3, p. 291; be careful with the sign). The second formula is not trivial. Just as with the Gohberg-Krein Index Formula for Wiener-Hopf operators on the circle and the straight line [Theorem 4.4 (p. 126), Exercise 4.7 (p. 128), and Theorem 4.11 (p. 129)], its significance derives from the fact that on the left side the analytic index defined globally by the operator P is an object of the analysis of infinite-dimensional function spaces, while on the right side the topological index is given by the symbol, i.e., by locally defined data of the linear algebra of finite dimensional vector spaces (which are suitably "integrated").

The proof of this index formula (see Theorem 12.3, p. 291) roughly rests on the fact that $\mathrm{Ell}_c(\mathbb{R}^n)$ is so rich that the analytic index (which as in Theorem 9.9d, p. 236, only depends on the symbol and does not change under "small" deformations of the symbol) can be considered an additive function on $\pi_{2n-1}(\mathrm{GL}(N,\mathbb{C}))$, and thus as a multiple of the topological index. Comparing the topological and the analytic index for the generators of the homotopy groups, we obtain equality.

2. The Difference Bundle of an Elliptic Operator

We will no longer pursue these homotopy-theoretic arguments, but rather we carry out the details of the proofs in the more convenient formalism of K-theory, which also permits an easier transition to the more general situation of the following section.

EXERCISE 12.1. Let X be a (not necessarily compact) oriented C^∞ Riemannian manifold, and let $\mathrm{Ell}_c(X)$ be the class of elliptic pseudo-differential operators of order 0 on X which are "equal to the identity at infinity"; i.e., for each $P \in \mathrm{Ell}_c(X)$, there is a compact subset $K \subseteq X$, such that $P\varphi = \varphi$ for all C^∞ sections φ (in the domain of definition of P) with support $\varphi \cap K = \phi$. The same condition should hold for the formal adjoint operator P^*.



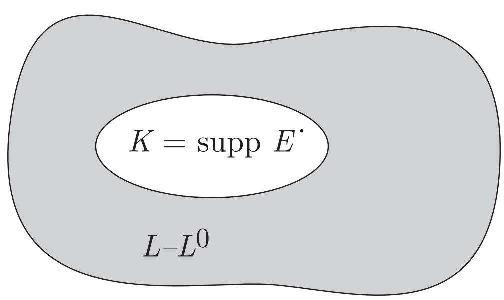
a) Show that this definition coincides with the definition given in Section 12.1 for $X = \mathbb{R}^n$.
b) Show that $\operatorname{index} P$ is well defined, depends only on $\sigma(P)$, remains constant under a C^∞ homotopy of the symbol within the space elliptic symbols which are the identity at infinity.
[Hint for b): Instead of repeating the proofs of Chapter 9, one can also reduce the present case to the results of Chapter 9 directly. Indeed, one can embed K in a bounded, compact, codimension-zero submanifold Y of X and then investigate the "doubled" operator $\tilde{P}$ on the closed manifold $\tilde{X} := Y \cup_{\partial Y} Y$. Show that $\operatorname{index} \tilde{P} = 2 \operatorname{index} P$.]

Much more generally one can define, for locally compact Y, the group $K(Y)$ through "**complexes of vector bundles with compact support**", these are short sequences $0 \to E^0 \xrightarrow{\alpha} E^1 \to 0$, where E^0 and E^1 are complex vector bundles over Y and α is a vector bundle isomorphism outside a compact subset of Y. Two complexes $E^\cdot = 0 \to E^0 \xrightarrow{\alpha} E^1 \to 0$ and $F^\cdot = 0 \to F^0 \xrightarrow{\beta} F^1 \to 0$ are called equivalent, if there is a complex $G^\cdot = 0 \to G^0 \xrightarrow{\gamma} G^1 \to 0$ with compact support over $Y \times I$ such that $E^\cdot = G^\cdot|_{Y\times\{0\}}$ and $F^\cdot = G^\cdot|_{Y\times\{1\}}$. The equivalence classes form a semigroup $C(Y)$ with sub-semigroup $C_\phi(Y)$ of *elementary complexes* with empty "support" (i.e., the bundle maps over Y are isomorphisms). Then the sequence

$$0 \to C(Y)/C_\phi(Y) \xrightarrow{d} K(Y^+) \xrightarrow{i^*} K(+) \to 0$$

is exact and splits; hence, $C(Y)/C_\phi(Y)$ proves to be isomorphic to $K(Y)$.

The construction (which goes back to Michael Atiyah and Friedrich Hirzebruch) of "difference bundles" $d(E^\cdot)$ for complexes $E^\cdot = 0 \to E^0 \xrightarrow{\alpha} E^1 \to 0$ over Y with compact support K goes roughly as follows: We choose a compact neighborhood L of K, such that K is contained in the interior L^0 of L. In order to extend the complex $E^\cdot$ to all of Y^+,



we replace it by an equivalent complex whose bundles are trivial over $L - L^0$:

$$0 \to E^0|_L \oplus F \xrightarrow{\alpha\oplus\mathrm{Id}} E^1|_L \oplus F \to 0$$

where $F \in \operatorname{Vect}(L)$ is chosen (by means of Appendix, Exercise B.10, p. 683) so that $E^1|_L \oplus F$ is trivial. Since α is an isomorphism on $L - L^0$, we must have that $E^0|_L \oplus F$ is trivial at least on $L - L^0$. Let

$$\tau_i : \left(E^i|_L - L^0\right) \oplus \left(F|_L - L^0\right) \to \left(L - L^0\right) \times \mathbb{C}^N,\ i = 0, 1,$$

be trivializations with τ_1 arbitrary and $\tau_0 := \tau_1 \circ (\alpha \oplus \mathrm{Id})$. Then the clutched bundle (see Appendix, Theorem B.8, p. 682)

$$G^i := \left(E^i|_L \oplus F\right) \cup_{\tau_i} \left((Y^+ - L^0) \times \mathbb{C}^N\right) \in \mathrm{Vect}(Y^+),$$

and we set $d(E^\cdot) := [G^0] - [G^1] \in K(Y^+)$. Since the fiber dimensions of G^0 and G^1 coincide, $d(E^\cdot)$ lies in $K(Y)$. Incidentally, one calculates easily that $d(E^\cdot) = d(E^\cdot \oplus H^\cdot)$, if $H^\cdot$ is an elementary complex, and that $d(E^\cdot)$ does not depend on the choice of F.

Because of the universal property (see Remark 11.5, p. 11.5) of K, it suffices to define the *splitting homomorphism*

$$b : K(Y^+) \to C(Y)/C_\phi(Y) \tag{12.1}$$

additively on $\mathrm{Vect}(Y^+)$. For $E \in \mathrm{Vect}(Y^+)$, one sets $b(E) := E^\cdot|_Y$, where $E^\cdot$ is the complex $0 \to E \xrightarrow{\beta} p^* i^* E \to 0$, $p : Y^+ \to \{+\}$ is the retraction, and β is an arbitrary extension of $\beta_+ := \mathrm{Id}$ to an isomorphism on a neighborhood of $+$. The support of $E^\cdot$ is then compact and contained in Y, and hence $b(E) \in C(Y)$. As an element of $C(X)/C_\phi(X)$, $b(E)$ is independent of the choice of the extension β. One sees immediately that $db \oplus p^* i^* = \mathrm{Id}$, whence in particular $K(+)$ is the cokernel of d; and with suitable homotopies for the vector bundle homomorphisms, we have $bd = \mathrm{Id}$, whence d is injective.

Further details of this construction are found in [**AS68a**, p.489 ff] and in [**Seg**, p.139-151], where complexes

$$0 \to E^0 \xrightarrow{\alpha_1} E^1 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_n} E^n \to 0$$

of length n with $a_{i+1} \circ a_i = 0$ are considered, which are exact outside a compact subset of Y. Incidentally, by means of tensor products of complexes of arbitrary length (e.g., see [**ES**, p.140 ff] a **ring structure** on $K(Y)$ may be introduced in a natural way[1].

If the locally compact space Y can be represented in the form $Z - A$ (where Z is compact and A is closed in Z), then often in the literature for this special ease one sees

$$K(Y) = K\left(Z - A\right) = K(Z, A) = C_{Z-A}(Z)/C_\phi(Z).$$

Hence, an element of $K(Y)$ is written as an equivalence class of an isomorphism $\sigma : E^0|_A \to E^1|_A$, where E^0 and E^1 are complex vector bundles over Z. In our applications (where Y is the tangent bundle TX) one prefers $K(BX, SX)$ to $K(TX)$, since $B(X)$ and $S(X)$ are compact for X compact. In fact, the construction

[1]For complexes $E^\cdot = 0 \to E^0 \xrightarrow{\alpha} E^1 \to 0$ and $F^\cdot = 0 \to F^0 \xrightarrow{\beta} F^1 \to 0$ of length 1, one obtains as an outer product the complex

$$E^\cdot \boxtimes F^\cdot := 0 \to E^0 \boxtimes F^0 \xrightarrow{\phi} \left(E^1 \boxtimes F^0\right) \oplus \left(E^0 \boxtimes F^1\right) \xrightarrow{\psi} E^1 \boxtimes F^1 \to 0$$

of length 2 where ("golden rule" of multilinear algebra)

$$\phi := \alpha \boxtimes \mathrm{Id} + \mathrm{Id} \boxtimes \beta \text{ and } \psi := -\mathrm{Id} \boxtimes \beta + \alpha \boxtimes \mathrm{Id}$$

With the help of Hermitian metrics on the vector bundles one can rewrite $E^\cdot \boxtimes F^\cdot$ as a complex of length 1

$$0 \to \left(E^0 \boxtimes F^0\right) \oplus \left(E^1 \boxtimes F^1\right) \xrightarrow{\theta} \left(E^1 \boxtimes F^0\right) \oplus \left(E^0 \boxtimes F^1\right) \to 0, \text{ where}$$

$$\theta := \begin{bmatrix} \alpha \boxtimes \mathrm{Id} & -\mathrm{Id} \boxtimes \beta^* \\ \mathrm{Id} \boxtimes \beta & \alpha^* \boxtimes \mathrm{Id} \end{bmatrix}$$

and α^* and β^* sre the adjoint homomorphisms. Details are in [**Ati67a**, p.93 f].

of the difference bundles in K-theory began with compact base space, although the proof (see e.g., [**Ati67a**, p.88-94]) in its basic idea is not as simple as that for the more general locally compact space.

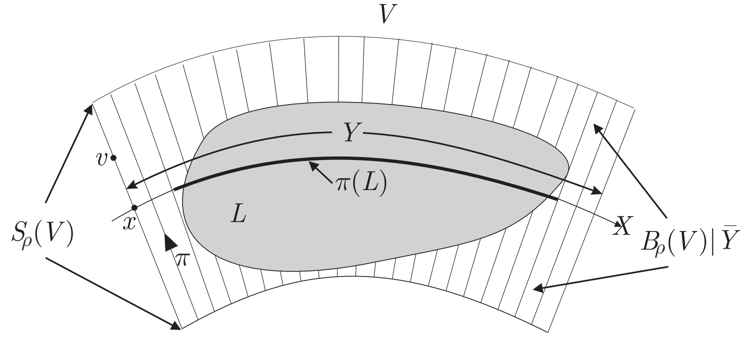
EXERCISE 12.2. Show that the symbol $\sigma(P)$ of an elliptic operator $P \in \mathrm{Ell}_c(X)$ defines an element $[\sigma(P)] \in K(TX)$ in a natural way, and that each $a \in K(TX)$ can be represented in this way; here, X is as in Exercise 12.1 and TX is the tangent bundle of X, which can be identified with the cotangent bundle T^*X by means of the Riemannian metric on X.
[Hint: Construct the "difference bundle" $[\sigma(P)] \in K(TX)$ as in the preceding. In addition, show

$$[\sigma(P)] = \left[\mathbb{C}^N_{B_0} \cup_{\sigma(P)} \mathbb{C}^N_{B_\infty}\right] - [N]$$

in the special case $X = \mathbb{R}^n$ where $TX^+ = (\mathbb{R}^{2n})^+ = S^{2n} = B_0 \cup B_\infty$ with $B_0 \cap B_\infty = S^{2n-1}$, and

$$\sigma(P)(\cdot,\cdot) : S^{2n-1} = \{(x,\xi) : |x| + |\xi| = r\} \to \mathrm{GL}(N, \mathbb{C}^N),$$

where r is so large that $P\varphi = \varphi$ for all N-tuples φ of complex-valued functions such that $\mathrm{support}(\varphi) \cap \{x : |x| \le r\} = \phi$.
For the reverse direction set $V := TX$ and represent $a \in K(V)$ (as with the "splitting homomorphism" (12.1)) by a complex $0 \to F^0 \xrightarrow{\phi} F^1 \to 0$, where the bundles F^i are restrictions to V of bundles of the same fiber dimension N over V; i.e., outside a compact subset $L \subseteq V$, we have isomorphisms $\tau_i : F^i|_{(V-L)} \overset{\cong}{\to} (V - L) \times \mathbb{C}^N$ such that $\phi := (\tau_1)^{-1}\, \tau_0$ is a bundle isomorphism over $V - L$.
If $\pi : V \to X$ is the base point map, then replace the bundles F^i on $V - L$ (where they are trivial) by $\pi^* E^i$, where E^i is the restriction of F^i to the zero section of V. On L this cannot be done in general. However, the following artifice (after [**AS68a**, p.492 f]) is helpful: Choose an open, relatively compact subset Y of X that includes $\pi(L)$ and a real number $\rho > 0$, such that the compact set L is

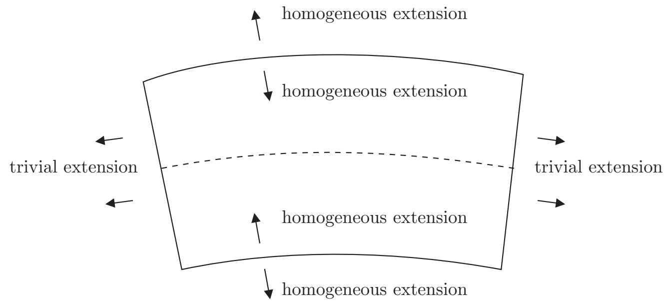


contained in the "ball bundle" $B_\rho(V)|_{\bar{Y}}$. Now show that $\bar{Y}$ is a deformation retract of $B_\rho(V)|_{\bar{Y}}$, and conclude (with Appendix, Theorem B.5, p. 680) that there are isomorphisms, $\theta_i : F^i|(B_\rho(V)|_{\bar{Y}}) \to \pi^* E^i|(B_\rho(V)|_{\bar{Y}})$, which are extensions of the trivialization isomorphisms given above over $\bar{Y} - Y$. Thus one must show that θ_i can be chosen so that the homomorphisms

$$\theta_i\,(v) : F^i_v \to \left(\pi^* E^i\right)_v = F^i_x,\ x \in \bar{Y} - Y \text{ and } \pi\,(x) = v$$

coincide with the composition $(\tau_i\,(x))^{-1}\, \tau_0\,(v)$. (Furthermore, if we require that θ_i is the identity on the zero section then it is uniquely determined up to homotopy.)

Now define $\alpha := \theta_1 \circ \phi \circ \theta_0^{-1}$ over $\partial\left(B_\rho(V)|_{\bar{Y}}\right) = (S_\rho(V)|_{\bar{Y}}) \cup \Big(B_\rho(V)|_{(\bar{Y}-Y)}\Big)$, and on $V|_{\bar{Y}}$ (modulo the zero section) extend it to be homogeneous of degree 0, and the given trivialization on $\pi^{-1}(Y - \bar{Y})$.



Finally, approximate the complex $0 \to \pi^* E^0 \overset{\alpha}{\to} \pi^* E^1 \to 0$ so obtained (where α is homogeneous of degree 0, and is induced by an isomorphism $E^0 \to E^1$ outside a compact subset of the base X) by a C^∞ mapping with the same properties, where (without loss of generality – see Appendix, Exercise B.11b, 683) the E^i can be taken to be C^∞ vector bundles. Incidentally, what simplifications can be made for $X = \mathbb{R}^N$?]

3. The Index Formula

Theorem 12.3. *For all $P \in \mathrm{Ell}_c(\mathbb{R}^n)$, we have the index formula*

$$\operatorname{index} P = (-1)^n \alpha^n([\sigma(P)]).$$

Here $a^n : K(\mathbb{R}^{2n}) \overset{\cong}{\to} K(\mathbb{R}^0) = \mathbb{Z}$ is the "periodicity homomorphism" produced by iteration of

$$\alpha_X : K(\mathbb{R}^2 \times X) \to K(X) \text{ for } X = \mathbb{R}^{2(n-1)},\ \mathbb{R}^{2(n-2)}, \ldots;$$

(see Theorem 11.17, p. 281).

Proof. 1) For locally compact X, we have the following commutative diagram:

$$\begin{array}{ccc} \mathrm{Ell}_c(X) & \overset{[\sigma(\cdot)]}{\longrightarrow} & K(TX) \\ {\scriptstyle \text{index}} \searrow & & \swarrow {\scriptstyle \text{index}} \\ & \mathbb{Z} & \end{array}$$

Here the "analytical index" is defined on $\mathrm{Ell}_c(X)$ by Exercise 12.1b, and by Exercise 12.2 (surjectivity of the difference bundle construction $[\sigma(\cdot)]$) it is well defined on $K(TX)$ and trivially additive (Exercise 1.5, p. 5). Hence, for $X = \mathbb{R}^n$ (where $K(TX)$ is isomorphic to $\mathbb{Z}$ by the Bott Periodicity Theorem above), the index is a multiple of this isomorphism, whence

$$\operatorname{index} P = C_n \alpha^n[\sigma(P)],$$

where the constant does not depend on P, but indeed may depend on n.

2) We now want to show that $C_n = (-1)^n$. For this we must find a $P \in \mathrm{Ell}_c(\mathbb{R})$ with $[\sigma(P)] = b \boxtimes \cdots \boxtimes b \in K(\mathbb{R}^{2n})$ ($b \in K(\mathbb{R}^2)$ the Bott class of Theorem 11.16, p. 280) and index $P = (-1)^n$. The main problem consists in finding a sufficiently simple operator P, so that one can compute its analytical index. We already know that P cannot have constant coefficients, since P is the identity at infinity; also,

P is not a differential operator, since it has vanishing order. These days, there are various ways to solve this problem; see [**Ati67b**, p.243 f], [**Ati70a**, p.110 ff], and [**Hö71**, p.141-146]. For us, it is most convenient to first show that we can restrict ourselves to the case $n = l$. As in Exercise 9.13b (p. 238), we have (with analogous proof) the following multiplicative properties: $P = Q\#R$, $P \in \mathrm{Ell}_c(\mathbb{R}^n), Q \in \mathrm{Ell}_c(\mathbb{R}^m), R \in \mathrm{Ell}_c(\mathbb{R}^k)$, and $m + k = n$ imply $\sigma(P) = \sigma(Q)\#\sigma(R)$ and $[\sigma(P)] = [\sigma(Q)] \boxtimes [\sigma(R)]$ and $\operatorname{index} P = (\operatorname{index} Q)(\operatorname{index} R)$. Since α^n is multiplicative by construction, we have $C^n = C_1^n$.

3) Thus, let $n = 1$ (i.e., $X = \mathbb{R}$ and $TX = \mathbb{R}^2 = \mathbb{C} = \{x + i\xi\}$. By definition, b is represented by the complex

$$\begin{array}{ccccccc} 0 & \to & \mathbb{C}_{\mathbb{C}} & \overset{\cdot (x+i\xi)^{-1}}{\to} & \mathbb{C}_{\mathbb{C}} & \to & 0 \\ & & \downarrow & & \downarrow & & \\ & & TX & & TX, & & \end{array}$$

which, as it stands, still does not represent any pseudo-differential operator. As in Exercise 12.2, we can deform the bundle map ϕ, which at the point (x, ξ) is defined on the fiber $\mathbb{C}$ by

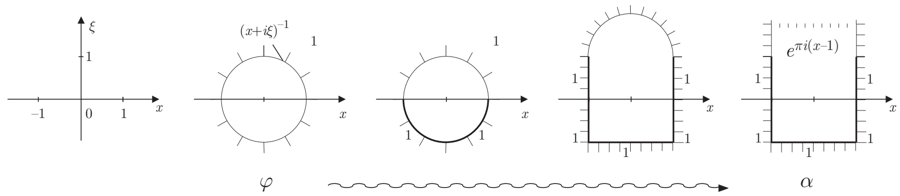
$$\phi(x, \xi) : z \mapsto z(x + i\xi)^{-1}$$

to a map α with

$$\begin{array}{cl} \alpha(x,\xi) = 1 & \text{for } |x| \geq 1 \\ \alpha(x, \lambda\xi) = \alpha(x,\xi) & \text{for } \lambda > 0 \\ \alpha(x,\xi) \neq 0 & \text{for } \xi \neq 0. \end{array}$$

For $|x| \leq 1$, we explicitly set

$$\alpha(x,\xi) = \begin{cases} e^{i\pi(x-1)} & \text{for } \xi > 0 \\ 1 & \text{for } \xi < 0 \end{cases}$$



Deformation: dilation of one

Thus, after smoothing, we can represent α as the symbol of an elliptic pseudo-differential operator T of order 0 on $\mathbb{R}$, which is the identity outside of the interval $[-1, 1]$ and inside it equals the "Toeplitz operator" $\tilde{T} := e^{i\pi(x-1)}P + (\mathrm{Id} - P)$ on the circle $S^1 \cong [-1, 1]$, where $P : \sum_{-\infty}^{\infty} a_\nu z^\nu \mapsto \sum_0^\infty a_\nu z^\nu$ is the projection operator. By construction, $\operatorname{index} T = \operatorname{index} \tilde{T}$, and according to Exercise 8.23c, we obtain $\operatorname{index} \tilde{T} = W(e^{i\pi(x-1)}, 0) = -1$. Thus, $\operatorname{index}(b) = -1$, and $C_1 = -1$ then follows. $\square$

Exercise 12.4. How can one directly prove $C_2 = 1$, without using induction from step 2 of the preceding proof?
[Hint: Consider, on the disk $X := B^2$, the "transmission operator"

$$A : (u, v) \mapsto \left(\frac{\partial u}{\partial \bar{z}}, \frac{\partial v}{\partial z}, (u - v)\,|_{S^1} \right)$$

with index $A = 1$ (Exercise 5.16, p. 151), and construct an operator A' (in the Green algebra $\mathrm{Ell}(X, \partial X)$) which is stably equivalent to A, and is equal to the identity in a neighborhood of ∂X; use the deformation procedure of Exercise 10.21a (p. 253) or Theorem 10.17, p. 248. Show index $A' =$ index A with Theorem 10.32d, whence index $A'' = 1$ if $A'' \in \mathrm{Ell}_c(K)$ is the extension of A' to all of $\mathbb{R}^2$ with $A'' = \mathrm{Id}$ outside B^2.

Then it only remains to show that $[\sigma(A'')]$ is actually $b \boxtimes b$. For this, represent b as in Theorem 12.3 by the complex $0 \to \mathbb{C}_{\mathbb{C}} \overset{\cdot\zeta^{-1}}{\to} \mathbb{C}_{\mathbb{C}} \to 0$ and derive (using the recipe given in the above footnote, p. 289) the representation of $b \boxtimes b$ by the complex $0 \to \mathbb{C}_{TZ} \oplus \mathbb{C}_{TZ} \overset{\theta}{\to} \mathbb{C}_{TZ} \oplus \mathbb{C}_{TZ} \to 0$ over the tangent bundle $TZ = \mathbb{C}^2 = \{(z, \zeta)\}$ of the space $Z = \mathbb{R}^2 = \mathbb{C} = \{z\}$. Then

$$\theta\,(z, \zeta) := \begin{bmatrix} z^{-1} & -\left(\bar{\zeta}\right)^{-1} \\ \zeta^{-1} & \left(\bar{z}\right)^{-1} \end{bmatrix}$$

can be deformed into $\sigma(A')\,(z, \zeta)$.]

Theorem 12.3 is a beautiful result of a purposeful application of modern topological methods to questions of analysis. Its theoretical ramifications are manifold, and we mention briefly:

(i) Generalizations of the Gohberg-Krein Index Formula for Wiener-Hopf operators on the circle or half-line (see above Chapter 4) to elliptic pseudo-differential operators of order 0 on n-dimensional Euclidean space. See also [**Prö72**] for a systematic comparison of these two interesting operator classes.

(ii) Analytic definition of the degree of a $(2n-1)$-dimensional homotopy class of $\mathrm{GL}(N, \mathbb{C})$: In its homotopy theoretic form (see Section 12.1), Theorem 12.3 supplies an explicit formula for the index, if one uses either one of the definitions of "degree" given explicitly in Section 11.2 above. Conversely, one can use Theorem 12.3 to define the degree (by analytic means) and generalize in this fashion the "algebraic" definition, which in Section 11.1 was available only for $n = 1$, to the case $n > 1$. Although "the index of an operator is usually a less computable quantity than an integral, say, the actual computation is for many theoretical goals not important, while the analytic definition entails numerous theoretical advantages." ([**Ati67b**, p.244], where a number of "advantages" – a priori integrality, connections with analytic function theory and Lie groups are discussed in detail.)

(iii) Applications to Boundary-Value Problems. In Section 10.4, we learned methods for "trivializing" the symbol of an elliptic boundary-value problem along the boundary. Thereby (more precise discussion in [**Bou**, p.40] one usually obtains an operator of order 0 which equals the identity near the boundary and therefore can be continued to an operator of $\mathrm{Ell}_c(\mathbb{R}^n)$, if the manifold with boundary is a region with boundary in $\mathbb{R}^n$. As was sketched in the hint to Exercise 12.4, the index does not change under these manipulations. Therefore, the index of a boundary-value problem can be computed as the case may be via Theorem 12.3; however, the derivation of a closed formula (see Chapter 14 below) requires a reduction to the Agranovich-Dynkin Formula (Theorem 10.32e, p. 260 or Exercise 10.33, p. 260) and thus to the study of pseudo-differential operators also on the boundary which is a closed manifold.

(iv) In the following chapter, finally, we will derive from Theorem 12.3, an index formula for elliptic operators on closed (= compact without boundary) manifolds, provided the latter can be imbedded in Euclidean space "trivially" (i.e. with trivial normal bundle).

CHAPTER 13

The Index Theorem for Closed Manifolds

1. The Index Formula

Let X be a closed (i.e., compact, without boundary), oriented, Riemannian manifold of dimension n, which is "trivially" embedded (i.e., with trivial normal bundle) in the Euclidean space $\mathbb{R}^{n+m}$. Let E and F be Hermitian vector bundles over X and $P \in \mathrm{Ell}_k(E,F)$, $k \in \mathbb{Z}$. Then we have the following formula:

THEOREM 13.1. *(M. F. Atiyah, I. M. Singer 1963).*

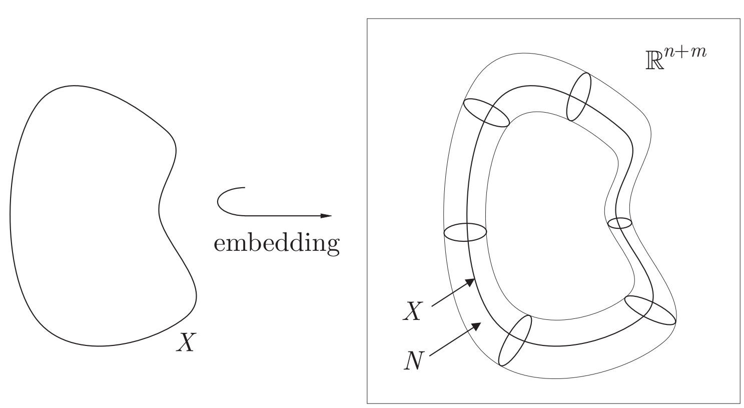
$$\operatorname{index} P = (-1)^n \alpha^{n+m}([\sigma(P)] \boxtimes b^m),$$

where $[\sigma(P)] \in K(TX)$ *is the "symbol class" of* P*,* $b \in K(\mathbb{R}^2)$ *is the "Bott class", and* $\alpha^{n+m} : K(\mathbb{R}^{2(n+m)}) \to \mathbb{Z}$ *is the iteration of the Bott isomorphism.*

REMARK 13.2. The preceding situation of "trivial" embedding arises in applications; e.g., when X is a hypersurface, in particular the boundary of a bounded domain in $\mathbb{R}^{n+1}$. For the more general case of "non-trivial" embedding, for different modes of expressing the "topological index" (the right side of the formula), and for a comparison of the various proofs, see the commentary below.

PROOF. 1) First, we want to visualize the contents of the formula. The left side is a well-defined (by Chapter 9) integer which depends only on the homotopy type of the symbol $\sigma[P] \in \mathrm{Iso}_{SX}(E,F)$. However, how is the right side defined? The construction of $\sigma[P] \in K(TX)$ was carried out in Exercise 12.2 (p.290) for $k=0$; the case $k \neq 0$ adds nothing new. (One can reduce it to the case $k=0$ directly via composition with Λ^{-k}.)

Now, consider the following figure:



Here N is a "tubular neighborhood" of X in $\mathbb{R}^{n+m}$; i.e., a neighborhood of X which locally (and also globally, because of the "triviality" of the embedding) has the form $X \times \mathbb{R}^m$.

The index formula then says that the following diagram is commutative:[1]

$$\begin{array}{ccccccc}
\mathrm{Ell}_k\,(E,F) & \overset{[\sigma(\cdot)]}{\rightarrow} & K(TX) & \overset{\boxtimes(-b)^m}{\rightarrow} & K(TX\times\mathbb{R}^{2m}) & = K(TN) \\
 & \searrow^{\text{index}} & \downdownarrows & & \downarrow \mathrm{ext} & \\
 & & \mathbb{Z} & \overset{(-1)^{n+m}\alpha^{n+m}}{\leftarrow} & K(T\mathbb{R}^{n+m}) & = K(\mathbb{R}^{2(n+m)})
\end{array}$$

Here, ext is induced by the map $K(T\mathbb{R}^{n+m})^+ \to (TN)^+$, which maps the complement of the open set TN in $T\mathbb{R}^{n+m}$ to the point at infinity $+ \in (TN)^+$.

2) We choose a very direct way for the proof of the index formula by systematically replacing the horizontal homomorphisms (K-theoretic operations) by operations associated with corresponding elliptic operators: Thus, for each elliptic operator P over X, we construct an elliptic operator P' over N, which is the identity at infinity and satisfies $\operatorname{index} P' = (-1)^m \operatorname{index} P$ and $[\sigma(P')] = [\sigma(P)] \boxtimes b^m$.

We take P' to be $P\#T^m$, where $T \in \mathrm{Ell}_c(\mathbb{R})$ is the standard operator with $\operatorname{index} T = -1$ given in the proof of Theorem 12.3, p. 291; further note that in Exercise 9.13 (p.238) the tensor product $\#$ was only defined. operators of order $k > 0$. Hence, more precisely, we take $P' \in \mathrm{Ell}_c\,(N)$ to be an operator whose symbol on the unit cosphere bundle coincides with $\sigma(P)\#\sigma(T)\#\overset{m}{\cdots}\#\sigma(T)$; e.g., for $m = 1$

$$(\sigma(P)\#\sigma(T))\,(x,t;\xi,\tau) = \begin{bmatrix} \sigma(P)(x,\xi) & -\operatorname{Id} F_x \otimes \sigma(T^*)(t,\tau) \\ \operatorname{Id} E_x \otimes \sigma(T)(t,\tau) & \sigma(P^*)(x,\xi) \end{bmatrix},$$

where $(x,t;\xi,\tau) \in (X\times\mathbb{R})\times S\,(X\times\mathbb{R})$. As in Exercise 12.2 (p.290) (since $\sigma(T)(t,\tau) = 1$ for t sufficiently large and for τ negative), this symbol can be conveniently deformed to a symbol σ' with $\sigma'(x,t;\xi,\tau) = \mathrm{Id}$ for t sufficiently large, if we identify $E' \oplus F'$ and $F' \oplus E'$ by means of switching the summands. Here $E' := p^*E$ where $p : N = X \times E \to X$ is the projection, whence $E'_{x,t} = E_x$, and F' is defined similarly.

5) Without loss of generality, we can take $E' \oplus F'$ to be a trivial bundle, since otherwise we can form $P' \oplus \mathrm{Id}_G$, where the bundle G over N is chosen so that $E' \oplus F' \oplus G$ is trivial. Hence, $E' \oplus F'$ is extended to a trivial bundle over all of $\mathbb{R}^{n+m}$. Then extend the operator P' to all of $\mathbb{R}^{n+m}$, by the identity outside N, to an operator $P'' \in \mathrm{Ell}_c(\mathbb{R}^{n+m})$. Since P' is the identity near the "boundary" $\bar{N} - N$ of N it follows that if $P''u = 0$ (on $\mathbb{R}^{n+m}$), then the support of u lies entirely in the interior of N, whence $\operatorname{Ker} P'' = \operatorname{Ker} P'$. By the same argument for the formal adjoint operators, it follows that $\operatorname{index} P'' = \operatorname{index} P'$; $[\sigma(P'')] = \mathrm{ext}[\sigma(P')]$ by construction.

4) The formula given for index P then follows from the index formula in the Euclidean case (Theorem 12.3, p.291). □

[1] We prefer working with the finitely generated abelian groups $K(TX)$, rather than with single elliptic operators or classes of such: It is precisely the advantage of topological methods that complicated objects of analysis, whose structure is only partially explored, can be replaced purposely by simple "quantities" (in the case before us, by the rank r of the group $K(TX)$). Incidentally, it is known that $r = \sum_{k=0}^{n} \operatorname{rank} H^{2k}(TX)$. Therefore, justified by Exercise 12.2, we will henceforth consider the "analytic index" not on $\mathrm{Ell}_k\,(E,F)$ but directly on $K(TX)$ as indicated by the double arrow in the diagram.

REMARK 13.3. Instead of tensoring with the standard operator $T \in \mathrm{Ell}_c(\mathbb{R})$, we can also (for $m = 2$) tensor with the standard transmission operator $T \in \mathrm{Ell}(B^2, S^1)$, thereby constructing an elliptic boundary-value problem over the bounded manifold $\bar{N}$, whose inner symbol can be deformed by the well-known procedure (using the boundary symbols) such that it becomes the identity near the boundary $\bar{N} - N$. The desired operators P' and P'' are then provided. Correspondingly, for arbitrary m, one can find a boundary-value problem whose index is 1 and whose symbol induces the bundle $b \boxtimes \cdots \boxtimes b$ (m-times): For this, one takes the differential operator $d + d^*$ (in the exterior calculus of differential forms; see Exercise 6.17, p. 169) from the forms of even order to those with odd order, with a suitable elliptic boundary-value problem in the sense of Chapter 10.

REMARK 13.4. The advantage of the construction of P' via boundary-value problems is best exhibited, when the embedding of X in $\mathbb{R}^{n+m}$ is not "trivial" i.e., when N is no longer $X \times \mathbb{R}^m$. In [**Ati68a**] and [**Ati70a**] devices of "equivariant K-theory" (with transformation groups) are employed for explicitly stating or axiomatically characterizing the desired operator P' with the help of symmetry properties of standard operators over the sphere. The use of "equivariant K-theory" is avoided (also in the case $N \neq X \times \mathbb{R}^m$ not treated by us) by passing to boundary-value problems and by another method proposed by [**Hö71c**] using "hypo-elliptic" operators and stronger analytical tools.

REMARK 13.5. Just as we learned (in Sections 11.1 and 11.2) different ways for defining "degree", the index, computed here in K-theoretic terms, can be determined in cohomological or integral form; see Section 14.14.1. This is possible without new or modified proofs, but simply by routine exercises in algebraic topology, the transition from K-theory to cohomology, whereby simply "one set of topological invariants is translated into another...." Which formula provides the "best answer" is largely a matter of taste. It depends on which invariants are most familiar or can be computed most easily (M. F. Atiyah, I. M. Singer).

2. Comparison of the Proofs: The Cobordism Proof

Michael Atiyah and Isadore Singer gave two more proofs of the Index Formula, in addition to the "embedding proof" given above. These are the original "cobordism proof" and the newer "heat equation proof". We cannot summarize them here, but we will comment briefly (see the tabulated crude survey at the end of this section). Since all three of these proofs appear to be somewhat complicated, several authors (among others, [**Boj**], [**Cald67**], [**See65**] and [**See67**]) tried to give simpler or more elementary proofs for the Euclidean case. In the judgement of [**Ati67b**, p.245] "these different proofs differ only in the use and presentation of algebraic topology" (instead of, and at times together with, the Bott Periodicity Theorem "older but not at all elementary parts of topology" are employed) −"the analysis is essentially the same in origin".

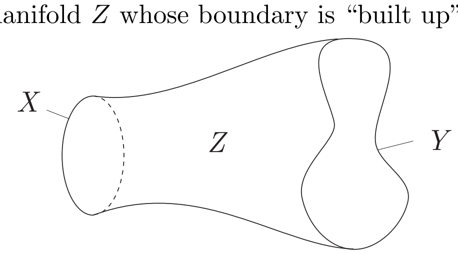
The "Cobordism proof" is sketched in [**AS63**] and worked out in detail in [**Bri63**], [**CaSc**] and [**Pal65**]. It was the first proof: It begins with a compact, oriented Riemannian manifold (without boundary) of dimension $2l$ and defines $d : \Omega^j \to \Omega^{j+1}$ and $d^* : \Omega^{j+1} \to \Omega^j$ as the exterior ("Cartan"-) derivative of forms and its adjoint. (These are considered in Exercise 6.17, p. 169. More precisely, we have here $\Omega^j := C^\infty\left(\Lambda^j\left(T^*X\right)\otimes\mathbb{C}\right)$.) For $\Omega := \sum_{j=0}^{2l}\Omega^j$, $d+d^* : \Omega\to\Omega$ is a self-adjoint, elliptic differential operator of first or whose square is the Laplace operator Δ of Hodge theory. If $* : \Omega^p \to \Omega^{2l-p}$ is the Hodge star (duality) operator, the formula

$$\tau\left(v\right) := i^{p(p-1)+l} * v,\ v\in\Omega^p \tag{13.1}$$

defines on Ω an "involution" (i.e., $\tau\circ\tau = \mathrm{Id}$). If $\Omega^\pm$ denote the ± 1 eigenspaces of τ, we define the signature operator (see Section 14.4 below and [**AS68b**, p.575]) $\left(d+d^*\right)^+ : \Omega^+ \to \Omega^-$ to be the restriction of $d+d^*$ to Ω^+. One can show that $\left(d+d^*\right)^+$ is an elliptic operator and its index is the "signature" of the manifold X, and is often named after Friedrich Hirzebruch.

For sufficiently many special manifolds (specifically for $X = S^{2l}$ and $X = \mathbb{P}^l\left(\mathbb{C}\right) :=$ complex projective space of complex dimension l) one can now compute the signature (index $\left(d+d^*\right)^+$) using cohomology theory and derive an index formula for manifolds of even dimension. "Sufficiently many" here means four things:

(i) By a deep result of cobordism theory by René Thom, every even-dimensional manifold Y is in a certain sense "cobordant" to the special manifolds; in other words, there is a bounded manifold Z whose boundary is "built up" from X and Y.



(ii) Furthermore, René Thom proved the vanishing of the signature for bounded manifolds. Hence the index of the signature operator on an arbitrary $2l$-dimensional manifold X can be computed from the indices of the special signature operators [**Hi66a**, p.58].

(iii) For a Hermitian C^∞-vector bundle E over X, let

$$\Omega^j_E := C^\infty\left(E\otimes\Lambda^j\left(T^*X\right)\right)$$

be the space of j-forms "with coefficients in E". By means of a "covariant derivative" ∇^E (inducing parallel translation along paths), one can define the operator

$$d^E : \Omega^j_E \to \Omega^{j+1}_E \text{ via}$$
$$d^E\left(v\otimes u\right) := \nabla^E v\wedge u + v\otimes du, \text{ for } u\in\Omega^j_E \text{ and } v\in C^\infty(E).$$

and its adjoint d^{E*}, and one can establish the index formula for the generalized signature operator $\left(d^E + d^{E*}\right)^+$ for the vector bundle E which is defined by means of an involution on $\Omega_E := \sum_{j=0}^{2l}\Omega^j_E$.

This concludes the proof since every elliptic operator on a closed even-dimensional oriented manifold X is "equivalent" in the sense of K-theory to a generalized signature operator. More precisely: $K(TX)$ is a ring over $K(X)$ and the subgroup $K(X)\cdot[\sigma((d+d^*)^+)]$ generated by generalized signature operators via $\left[\sigma\left(\left(d^E+d^{E*}\right)^+\right)\right] = [E]\cdot[\sigma((d+d^*)^+)]$ is so large, namely a "subgroup of finite index", that "practically" all of $K(TX)$ is generated. Here "practically" means up to the image of

$K(X)$ in $K(TX)$ and up to 2-torsion, where the index must vanish as an additive function with values in $\mathbb{Z}$.

At this place, the theory of pseudo-differential operators enters in order to achieve that the symbols are arbitrary bundle isomorphisms over SX, and to allow reduction of the index computation from $\mathrm{Ell}(X)$ to $K(TX)$. Further, the Bott Periodicity Theorem is used in somewhat generalized form in representing $K(TX)$ approximately by $K(X)$; see [**ABP**, p.321 f].

(iv) The general index formula can be extended to an odd-dimensional manifold X, using the multiplicative property of the index by tensoring with the standard operator T with index 1 on S^1 and by passing to the even-dimensional manifold $X \times S^1$. If one is not interested in the sign in the index formula, one can avoid the explicit definition of T and simply pass to the "squared" (relative to the tensor product) operator on the even-dimensional manifold $X \times X$.

3. Comparison of the Proofs: The Imbedding Proof

The "Imbedding Proof", which was given in Section 13.1, following [**Ati67b**], [**Ati68a**] and [**Ati70a**]. In this proof, the methods remain topological with the consideration of $K(TX)$ instead of the operator space $\mathrm{Ell}(X)$. The idea goes back to the proof of the Riemann-Roch Theorem by Alexander Grothendieck, Section 14.7 below. One shows first using the Bott Periodicity Theorem that every elliptic operator on the sphere or Euclidean space is equivalent, in sense of K-theory, to one of infinite many (more precisely ($|\mathbb{Z}|$-many), standard operators, and then the case of an arbitrary elliptic operator on arbitrary closed manifold is reduced to the standard case by imbedding.

The advantage, as well as the weakness, of this proof lies in its perhaps somewhat forced directness. It succeeds on the one hand in eliminating cohomology and cobordism theory completely, bringing out the function analytic and topological pillars (Theorems of Gohberg-Krein and Bott) plainly and in their most elementary form, and in achieving through this simplicity of tools the greatest susceptibility to generalization (see Section 14.11). On the other hand, under the imbedding (except for particularly smooth ones, e.g. holomorphic embeddings of algebraic manifolds in a complex projective space) the special structure of "classical" operators is completely destroyed. For example, the signature operator does not become another signature operator under the imbedding, and to prove the Riemann-Roch Theorem for arbitrary compact complex manifolds (to mention another problem defined by "classical" operators; see also Section 14.7 below), one has to leave this category, whereby many of the interesting and sometimes open problems of modern differential topology become less transparent.

4. Comparison of the Proofs: The Heat Equation Proof

In comparison with the first two proofs, which argue more topologically, the "heat equation proof" offers a completely different and initially, purely analytic approach to the index problem. The germinal idea goes back to papers of Marcel Riesz on spectral theory of positive self-adjoint operators, and was presented by M. F. Atiyah as early as 1966 at the International Congress of Mathematicians in

Moscow, and then published, also in connection with applications of the Index Formula to fixed point problems in [**AB67**], [**Ati68b**] and in related form in [**Cald67**] and [**See67**]:

1. One starts with an operator $P \in \mathrm{Ell}_k(E, F)$, $k > 0$, where E and F are Hermitian C^∞ vector bundles on the n-dimensional, closed, oriented, Riemannian manifold X. Then the operator P^*P is a non-negative self-adjoint operator of order $2k$ with a discrete spectrum (see Chapter 3 above) of non-negative eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \cdots$ (the multiplicity may be larger than 1, hence "$\le$"); and the series

$$\theta_{P^*P}(t) := \sum\nolimits_{m=1}^{\infty} e^{-t\lambda_m} \tag{13.2}$$

converges for all $t > 0$. By the way, for $X = S^1$ and $P = -i\frac{d}{dx}$, we obtain the "theta function" $\theta(t) = \sum_{m=0}^{\infty} e^{-tm^2}$ of analytic number theory since the square integers are exactly the eigenvalues of $P^*P = \Delta = -d^2/dx^2$.

Correspondingly, one forms the function θ_{PP^*}. The operators P^*P and PP^* have the same non-zero eigenvalues, and only the eigenvalue 0 has in general different multiplicities, namely $\dim \mathrm{Ker}\, P$ and $\dim \mathrm{Ker}\, P^*$ (see the Remark 2.11, p. 17). This way, one has a new index formula

$$\operatorname{index} P = \theta_{P^*P}(t) - \theta_{PP^*}(t),\ t > 0. \tag{13.3}$$

Trivially, one may choose an arbitrary function ϕ on E with $\phi(0) = 1$ instead of the function $m \mapsto e^{-tm}$ and thus obtain for each ϕ a further index formula

$$\operatorname{index} P = \sum_{\lambda \in \mathrm{Spec}(P^*P)} \phi(\lambda) - \sum_{\lambda \in \mathrm{Spec}(PP^*)} \phi(\lambda)$$

Now, the theta function is distinguished by permitting near $t = 0$ an asymptotic development

$$\theta_{P^*P}(t) \sim \sum_{m \ge -n} t^{m/2k} \int_X \mu_m\left(P^*P\right),\ (\text{as } t \to 0^+) \tag{13.4}$$

where $\mu_m\left(P^*P\right)$ is, for each $m \in \mathbb{Z}$, a certain density on X which can be formed canonically with the coefficients of the operator P^*P. Using (13.3) we get from (13.4) the explicit integral representation

$$\operatorname{index} P = \int_X \mu_0\left(P^*P\right) - \mu_0\left(PP^*\right). \tag{13.5}$$

2. The convergence of the series in (13.2) has implications for the construction of solutions of the heat conduction equation (where $\square := P^*P$ is a generalized Laplace operator) [**Gi95**, p.64-65].

$$\frac{\partial u}{\partial t}(x,t) + \square u(x,t) = 0, \quad x \in X,\ t \in [0, \infty)$$

with the initial condition $u(\cdot, 0) = u_0 \in L^2(E)$. Here u is the unknown function on $X \times [0, \infty)$ with values in the bundle E (the "heat distribution"). Then

$$H_t := e^{-t\square} = \mathrm{Id} - t\square + \frac{t^2}{2!}\square^2 - \frac{t^3}{3!}\square^3 + \cdots,\ t \ge 0$$

is a well-defined family of bounded operators on the Hilbert space $L^2(E)$ which satisfies the heat equation

$$\frac{dH_t}{dt} + \Box H_t = 0$$

with initial value $H_0 = \mathrm{Id}$. Thus H yields for each initial distribution u_0 the heat distribution at time t via the formula $u(\cdot, t) = H_t u_0$.

Since the eigenfunctions $\{v_m : m \in \mathbb{Z}^+\}$ of a form a complete orthogonal system for $L^2(E)$, the formula

$$\operatorname{trace} e^{-t\Box} = \theta_\Box(t) = \sum\nolimits_{m=1}^{\infty} e^{-t\lambda_m}$$

is meaningful. The convergence of the series in (13.2) means that the evolution operators H_t of the parabolic heat conduction equation belongs to the "trace class" for $t > 0$. By means of the theory of pseudo-differential operators it follows more precisely that H_t is a "smoothing operator", i.e., an operator of order $-\infty$ which is representable as an integral operator

$$(H_t v)(x) = \int_X K_t(x, y)\, v(y)\, \omega_y, \ \ v \in L^2(E), \ \ x \in X$$

with C^∞ weight function $(x, y) \mapsto K_t(x, y) \in L(E_y, E_x)$ and volume element ω. Then $\theta_\Box(t) = \operatorname{trace} H_t = \int_X \mu_t$, $t > 0$, where

$$\mu_t(x) := \operatorname{trace}(K_t(x, x))\, \omega_x = \sum\nolimits_{m=1}^{\infty} e^{-t\lambda_m} \left|v_m(x)\right|^2 \omega_x$$

defines a density on X which at each $x \in X$ is the pointwise trace of the operator K_t, times ω_x. While μ_t can be expressed in terms of the coefficients of the operator P only very indirectly (via the eigenvalues and eigenfunctions of $\Box = P^*P$), one has at each point $x \in X$ an asymptotic expansion

$$\mu_t(x) \sim \sum_{m \geq -n} t^{m/2k} \mu_m(x)$$

where the μ_m are purely local invariants of P^*P which then implies (13.4).

4. A proof of (13.4) with a recipe for the computation of μ_m extracted from the theory of pseudo-differential operators is due to [**See67**]. It shows that the μ_m depend rationally on the coefficients of P and their derivatives of orders $\leq n$. A more intuitive and heuristic description of the μ_m for the special case $X = T^n := \mathbb{R}^n/(2\pi\mathbb{Z}^n)$, to which we paid special attention in our Sobolev case studies (Chapter 7), can be found in [**APS73**, p.300 f]. The idea of the proof goes back to the mathematicians Subbaramiah Minakshisundaram and Åke Pleijel, who in 1949 (long before an effective machinery for pseudo-differential operators was established) computed the μ_m for the case $P^*P = \Delta$, where Δ is the invariantly defined Laplace-Beltrami operator which depends only on the Riemannian metric on X, i.e., $k = 1$ and $E = \mathbb{C}_X$. Precisely, S. Minakshisundaram and Å. Pleijel (and later R. T. Seeley, when generalizing their results) studied in place of P^*P the positive self-adjoint operator $\Box = \mathrm{Id} + P^*P$ and, in place of the theta function, the zeta function $\zeta(z) := \sum_{m=1}^{\infty} (\lambda_m)^{-z}$ summed over all (discrete positive) eigenvalues of $\Box$. As is well-known, the zeta function is well-defined for $\mathrm{Re}(z) > \dim X$ and can be continued to a meromorphic function in the z-plane with finitely many real poles of order 1 with behavior at poles known in principle. Among other things it

is found that $z = 0$ is not a pole and that the value $\zeta(0)$ can be expressed explicitly in terms of $\square$; in fact,

$$\zeta(0) = \int_X \rho_0\left(\square\right), \tag{13.6}$$

where the right hand side is fairly complicated but can be computed in principle. On the other hand, $\zeta(z)$ can be interpreted within spectral theory as trace($\square^{-z}$), and finally $\zeta(0)$ appears as the constant term in the asymptotic expansion of $\theta(t)$ as $t \to +\infty$. This establishes the connection with the heat conduction approach. In particular, the measure $\mu_0(\square)$ sought there is identical with the measure $\rho_0\left(\square\right)$ in equation (13.6), and (13.4) follows from (13.6) and similar computations of the residues of ζ at its poles.

5. Thus, a general algorithm is available that is capable of producing the right hand side of the index formula (13.5) in finitely many steps by means of a computer for example. In contrast to the index formulas of the cobordism and imbedding proofs (into which enter the derivatives of the coefficient of P up to order 2 only) this formula is in the general case complicated numerically and algebraically mainly by the appearance of derivatives up to order n $(= \dim X)$. While for algebraic curves of complex dimension 1 (= Riemannian surfaces, $n = 2$), the formula can be handled well computationally, the general situation requires so much effort that M. F. Atiyah and R. Bott by their own admission had initially "little hope for interpreting these integrals directly in terms of the characteristic classes of E and X", and therefore "the beautiful formula appeared, be useless in this context." Only a series of more recent papers on curvature tensors revealed that in the special case when X is even-dimemsional and P is the signature operator $(d + d^*)^+$, all higher order derivatives "cancel out" in Seeley's formula for the measure $\mu_0(P^*P)$, and only the derivatives up to order 3 remain. Details of such computations appear first in [**MS**], in the case (not all that fortunate for aspect) that P is the operator $d + d^* : \Omega^{\mathrm{even}} \to \Omega^{\mathrm{odd}}$ with the Euler characteristic as index (see Section 14.4). This was generalized in 1971 by V. K. Patodi – again by means of symmetry considerations – and extend to Riemann-Roch operators (see Section 14.7) in particular. P. Gilkey succeeded shortly thereafter in replacing Patodi's complicated group theoretic cancellation procedure for the higher derivatives by an axiomatic argument which was drastically simplified in [**ABP**] through the use of stronger tools of Riemannian geometry. It says roughly that in each integrand with the general qualitative properties of the annoying higher derivatives can be disregarded and $\mu_0(P^*P)$ can identified with the (normalized) Gaussian curvature.

In this fashion, a new purely analytic proof of the Hirzebruch Signature Theorem (the index formula for "classical" operators) is achieved which implies the general index formula, as in the cobordism proof (see above, items (iii) and (iv) in the Section 13.2), and with the same topological arguments.

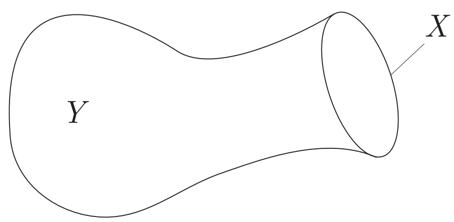
6. The significance of the heat equation proof, which cannot be extended (just like the cobordism proof) to families of elliptic operators and operators with group action, is at present difficult to estimate. Its authors, who (as an aside) acknowledge that their "whole thinking on these questions has been stimulated and influenced very strongly by the recent paper of Gelfand on Lie algebra cohomology" [**Ge**], point out that their proof is hardly shorter than the imbedding proof since it "uses more analysis, more differential geometry and no less topology. On the other hand, it is more direct and explicit for the classical operators associated with

Riemannian structures: In particular, the local form of the Signature Theorem and its generalizations are of considerable interest in itself and should lead to further developments" [**ABP**, p.281].

This prediction appears to materialize even beyond the realm of differential geometry: The approach via the zeta function of the Laplace-Beltrami operator Δ on X (whose values yield real-valued invariants of the Riemannian metric ρ of X – *spectral invariants* – at any point where the zeta function does not have a pole) has been extended to the systems case, where the Laplace equation is replaced by the system of partial differential equations of the *total* Laplace operator of Hodge theory which can be represented as the square of a formally self-adjoint operator A (the *Dirac operator*). In analogy to the zeta function, one considers for an operator A, which is not positive, the function $\eta(z) := \sum_{\lambda\neq 0}(\operatorname{sign}\lambda)\,|\lambda|^{-z}$ where summation is over the eigenvalues of A with proper multiplicity. Again $\eta\,(z)$ is a holomorphic function for large $|z|$, which can be continued meromorphically to the whole z-plane. Corresponding to the asymptotic expansion above in equation (13.4) for the theta and zeta functions, one can look for an integral formula for $\eta\,(0)$. Such a formula

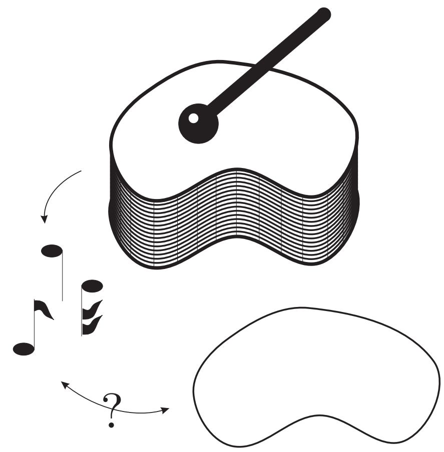
$$\eta\,(0) = \int_\gamma \alpha\,(\tilde\rho) - \text{integer} \tag{13.7}$$

is proven in [**APS75**, Thm. 4.14], when X can be obtained as the boundary of a 4-dimensional manifold Y with Riemannian metric $\tilde\rho$.



Here $\tilde\rho$ is assumed to induce on X the metric ρ and to render a neighborhood of X in Y isometric to $X \times [0, 1)$. The integrand $\alpha\,(\tilde\rho)$ is explicitly known (the "l-th Hirzebruch L-polynomial in Pontryagin forms of the Riemannian metric $\tilde\rho$"), as well as the integral correction term (the "signature of Y" defined by the topology of Y). In this way, a formula is obtained which relates the spectral invariant $\eta(0)$, measuring the asymmetry of the spectrum of A, with a differential geometric and a purely topological invariant.

This formula lies deeper than (13.4) or (13.5), where the integrand was of local type, and has numerous relations to the cobordism as well as the imbedding (particularly via boundary-value problems) proof. It is not as esoteric as it may appear to someone not so much interested in differential-topological problems, since it yields (roughly) a direct geometric interpretation for the peculiarities of the distribution of the eigenvalues of A: Newer results in this direction by Victor W. Guillemin and others, say for example, that a Riemannian manifold X is isometric with S^n if the spectra of the Laplace operators coincide. One cannot always expect that two manifolds with equal spectra of their Laplace operators are isometric (the 16-dimensional tori yield a counterexample, see), but it appears that at least "extreme" distributions of the eigenvalues (when they are not randomly distributed, but lumped together near integers for example) carry with them "extreme" geometric situations (in our example the closedness of the geodesics). According to an announcement in [**Si75**], this answers in principle the classical question "Can you hear the shape of the drum?" (Mark Kac) which had already motivated [**MS**].



Further, a gate opened to the "inverse problem"[2] of "mathematical modelling" of real phenomena: Theoreticians frequently and perforce only "apply" theory; i.e., they try, similar to the axiomatic method within mathematics, to draw conclusions as far-reaching as possible about the concrete behavior from relatively modest assumptions about the existence of certain laws. Conversely, the practitioner needs in general the "inverse" of the theory, namely the exposure of regularities in the observations at hand. (Somewhat overstated, the practitioner desires to fit a curve to given measurements, while the theoretician sees his strength in a detailed discussion of the properties of a given curve.) In this sense, the novelty consists in the attempt to estimate the parameters of a differential equation, when information about special solutions (eigenfunctions and eigenvalues, for example) is available. The following last chapter contains a survey on some reformulations, applications and generalizations of the Atiyah-Singer Index Formula.

[2]According to a communication by Richard Bellmann, the "inverse problem" in its most general formulation goes back to Carl Gustav Jacob Jacobi (1804-1851). Today, the Inverse Problem is studied frequently in very different contexts, reaching from algebraic problems [**Ul**, p.32] "structure identification and parameter estimation" in control theory. A special version of the Inverse Problem is the inversion of spectral analysis, "spectral synthesis", which goes far beyond the newer results of Riemannian geometry presented here [**Ben**].

CHAPTER 14

Applications (Survey)

With the Atiyah-Singer Index Formula, we proved "one of the deepest and hardest results of mathematics" which "is probably enmeshed more widely with topology and analysis than any other single result" [**HiZa**, VIII]. *In this book, we are mainly interested in the varied sources and parts which flow together or are put together in the Index Formula. But the formula itself is of great interest also. It can express important and far-reaching ideas in various areas of application, through numerous corollaries, specializations, reformulations and generalizations.*

The appraisals of the role of the Atiyah-Singer Index Formula in these applications are contradictory: On the one hand, it permits us to attack complicated topological questions with relatively simple analytic methods; on the other hand, this formula frequently serves only "to derive a number of wholly elementary identities, which could have been proved much more easily by direct means." This is the judgement of Friedrich Hirzebruch and Don Zagier about the relationship between the Index Formula and topics from elementary number theory. Some of the following principal theorems, which at first could be proved only in the framework of the Atiyah-Singer theory, have been proved directly in the meantime. This is true for the general Riemann-Roch Theorem (see Section 14.7 below) in [**TT**], *its consequences for the classification of certain algebraic surfaces drawn by Kunihiko Kodaira in* [**Hi73**], *and for some of the vector field computations of Atiyah-Dupont (see Section 14.5) in* [**Kos**]. *There is much that is unclear in the relationship between the Atiyah-Singer Index Theorem and its applications. It is important and interesting that these many relationships exist, although considerable future efforts will be required to bare the real reasons for these relationships, and, in the end, to better understand the unity of mathematics and the specifics, and interrelationships of its parts. Once again, we quote* [**HiZa**, VII]: *"That a connection exists, a number of people realized essentially at the same time.... And since neither we nor anybody knows why there must be such a connection, this seemed like an ideal topic for a book in order to confront other mathematicians with a puzzle for their embarrassment or their entertainment – as the case may be." While some of the following areas of applications will be presented in more detail in Part IV, this chapter is more of an overview and a literature survey.*

1. Cohomological Formulation of the Index Formula

In Theorem 13.1 (p. 295), the Index Formula is phrased in the language of K-theory: Its right side only involves vector bundles and operations vector bundles. The conversion to cohomological form is carried out in [**Ati68b**, p.546-559]. While the individual calculations are somewhat complicated (if more or less routine topological exercises days), the underlying method (namely, the construction of a "functor" that assigns to each vector bundle a "characteristic" cohomology class of the base X of E with coefficients in $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$) is rather clear. Here, we follow [**MilS**].

Let E be a complex vector bundle of fiber dimension N over paracompact space X, with projection $\pi : E \to X$, and let E_0 be the subspace of E obtained by removing the zero section. As a real vector bundle of fiber dimension $2N$, E is oriented, since all complex bases $e_1, ..., e_N$ of the fiber E_x, $x \in X$, yield real bases e_1, ie_1, e_2, ie_2, $\ldots$,e_N, ie_N of the same orientation. In the language of cohomology, an orientation for E_x is the choice of a generating element $\mathcal{O}_x$ for $H^{2N}(E_x, E_x \setminus \{0\}\,; \mathbb{Z})$. Now, the following observations go back to R. Thom:

(i) $H^i(E, E_0; \mathbb{Z}) = 0$ for all $i < 2N$.

(ii) The orientation of E defines a "total orientation class" $U \in H^{2N}(E, E_0; \mathbb{Z})$ by the condition $(j_x)^* U = \mathcal{O}_x$ for all $x \in X$, where $j_x : (E_x, E_x \setminus \{0\}) \to (E, E_0)$ is the embedding.

(iii) Via the cup product $u \mapsto \pi^*(u) \cup U$, $u \in H^i(X; \mathbb{Z})$, with the orientation class U, an isomorphism $\Phi : H^*(X; \mathbb{Z}) \cong H^*(E, E_0, 2E)$ is defined which raises the dimension of the cohomology classes by $2N$. This is the *Thom isomorphism* that can be given more generally for all real, oriented vector bundles of arbitrary fiber dimension. For a comparison with the Bott isomorphism of K-theory, see [**Ati68b**, p.546-559].

The N-th Chern class $c_N(E)$ of E is the class $\Phi^{-1}(U \cup U) \in H^{2N}(X; \mathbb{Z})$, the total Chern class

$$c(E) = 1 + c_1(E) + ... + c_N(E);\ c_i(E) \in H^{2i}(X, \mathbb{Z})$$

is obtained (for example) by the axiomatic conditions of functoriality (i.e., $f^* c(E) = c(f^* E)$ for $f : Y \to X$) and homomorphy (i.e., $c(E \oplus F) = c(E) \cup c(F)$). Because of this homomorphism property, the Chern classes are not only defined on $\mathrm{Vect}(X)$, but also on $K(X)$, since $c(E) = 1$ if E is a trivial bundle.

By means of the formal factorization $c(E) = (1 + y_1) \cup \cdots \cup (1 + y_N)$ with $y_i \in H^2(X, Z)$, where the $-y_i$. are the hypothetical "zeros" of the polynomial $1 + c_1(E)t + ... + c_N(E)t^N$, one obtains the Chern character

$$ch(E) := \sum_{i=1}^N e^{y_i} = N + \sum_{i=1}^N y_i + \frac{1}{2!} \sum_{i=1}^N y_i^2 + \cdots$$

which extends to a ring homomorphism $ch : K(X) \to H^*(X; \mathbb{Z})$, thereby providing a natural transformation from K-theory to singular cohomology theory with rational coefficients and compact support.

Results. a) If P is an elliptic operator on a closed oriented manifold X of dimension n, then we have

$$\operatorname{index} P = (-1)^{n(n+1)} \{\Phi^{-1} ch[\sigma(P)] \cup \tau(TX \otimes \mathbb{C})\} \, [X] \,.$$

Here $[\sigma(P)] \in K(X)$ is the difference bundle of the symbol $\sigma(P)$ of P, $ch[\sigma(P)] \in H^*(TX; \mathbb{Q})$ is its Chern character, $\Phi : H^*(X; \mathbb{Q}) \to H^*(TX, (TX)_0, \mathbb{Q}) = H_c(TX; \mathbb{Q})$ is the Thom isomorphism, $[X] \in H_n(X; \mathbb{Q})$ is the fundamental cycle of the orientation of X, and

$$\tau(E) := \frac{y_1}{1 - e^{-y_1}} \cdot \cdots \cdot \frac{y_N}{1 - e^{-y_N}} \in H^*(X; \mathbb{Q})$$

is the Todd class of the complex vector bundle E of fiber dimension N, whose Chern class is factored as above. (Here, take E to be the complexification $TX \otimes \mathbb{C}$ with fiber dimension n.) Incidentally, by means of Riemannian geometry one can express $ch(E)$ and $\tau(E)$ by the "curvature matrix" of the vector bundle E which one equips with a Hermitian metric; e.g., see [**AS68b**, p.551] and [**APS73**, p.310], or Section 16.7.

b) More generally, one can drop the orientability of X and obtain the formula

$$\operatorname{index} P = (-1)^n \left\{ch[\sigma(P)] \cup \pi^* \tau(TX \otimes_{\mathbb{R}} \mathbb{C})\right\}[TX].$$

Here $[TX]$ is the fundamental cycle of the tangent bundle TX which admits an orientation as an "almost-complex manifold" (namely, divide the tangent space of TX into a "horizontal" = "real" and a "vertical" = "imaginary" part); and $\pi : TX \to X$ is the projection. The calculation of the right sides in a) and b), naturally involves only the evaluation the highest dimensional components of the cup product of the respective fundamental cycles.

2. The Case of Systems (Trivial Bundles)

In [**Ati68b**, p.600-602], a drastic simplification of the Index Formula is proved for the case of trivial vector bundles (i.e., the elliptic operator P is applied to a system of N complex-valued functions). The symbol of such an operator is then a continuous map $\sigma(P) : SX \to \mathrm{GL}(N, \mathbb{C})$, where SX is the unit sphere bundle of X. Look at the induced cohomology homomorphism $(\sigma(P))^* : H^*(\mathrm{GL}(N, \mathbb{C})) \to H^*(SX)$ (coefficients arbitrary), it becomes evident that in order to obtain a useful formula in this case, one must know something about the cohomology of the Lie group $\mathrm{GL}(N, \mathbb{C})$. Since the unitary group $\mathrm{U}(N)$ is a deformation retract of $\mathrm{GL}(N, \mathbb{C})$ we have $H^*(\mathrm{GL}(N, \mathbb{C})) = H(\mathrm{U}(N))$. Moreover, since we have a natural map $\rho_N : \mathrm{U}(N) \to \mathrm{U}(N)/\mathrm{U}(N-1) = S^{2N-1}$, we may obtain all essential information on $H^*(\mathrm{GL}(N, \mathbb{C}))$ from the well-known cohomology of S^{2N-1}. More precisely: Let $u_i \in H^{2i-1}(S^{2i-1})$ be the natural generating element, where S^{2i-1} is oriented as the boundary of the ball in $\mathbb{C}^i$. Then set $h_i^i : (\rho_i)^* (u_i) \in H^{2i-1}(\mathrm{U}(i))$; e.g., $h_N^N \in H(\mathrm{U}(N))$. For $i \le N$, we obtain additional elements $h_i^N \in H^{2i-1}(\mathrm{U}(N))$ by means of the the normalization condition $j^* h_i^N = h_i^i$, where $j : \mathrm{U}(i) \to \mathrm{U}(N)$ is the canonical embedding. The h_i^N $(i \le N)$ which form a system of generators for the algebra $H^*(\mathrm{U}(N))$.

Results. For an elliptic system P of N pseudo-differential equations for N complex-valued functions on a closed manifold X of dimension n, we have

a)

$$\operatorname{index} P = (-1)^n \left\{ \sum\nolimits_{i=1}^{N} (-1)^{i-1} \frac{(\sigma(P))^* h_i^N}{(i-1)!} \cup \tau(X) \right\} [SX].$$

Here $[SX]$ is the fundamental cycle of the canonical orientation of SX and $\tau(X)$ is the lift to SX of the Todd class of the complexification of TX; see Section 14.1 above.

b) If $n \leq 3$ or if X is a hypersurface in $\mathbb{R}^{n+1}$, then $\tau(X) = 1$, and hence

$$\operatorname{index}(P) = \begin{cases} (-1)^{N+n-1} \frac{(\sigma(P))^* h_n^N}{(n-1)!} & \text{for } N > n \\ -\frac{\text{mapping degree } (\rho \circ \sigma(P))}{(n-1)!} & \text{for } N = n \\ 0 & \text{for } N < n. \end{cases}$$

For the determination of the mapping degree of the composition $\rho \circ \sigma(P) : SX \to S^{2n-1}$, see [**BJ**, 14.9.6-10].

c) In the framework of Hodge theory (which provides a canonical isomorphism between the "harmonic differential forms" of degree p on a manifold and the p-th cohomology of the manifold with coefficients in $\mathbb{C}$ – see also Section 14.4 below), there are explicit differential forms $\omega_i \in \Omega^{2i-1}(\mathrm{U}(N))$ (so-called "bi-invariant forms"), which represent the generators h_i^N. If $\pi : SX \to X$ is the projection, $\tau \in \Omega^*(X)$ is the total differential form corresponding to the Todd class (involving the curvature of the Riemannian manifold X) and $\omega := \sum_{i=1}^N \frac{(-1)^{i-1}\omega_i}{(i-1)!} \in \Omega^{2i-1}(\mathrm{U}(N))$ is the "total bi-invariant form", then one obtains the integral formula

$$\operatorname{index} P = (-1)^n \int_{SX} \sigma(P)^* \omega \wedge \pi^* \tau.$$

3. Examples of Vanishing Index

In a series of special cases one can conclude that the index of an operator vanishes by using the index formula in Theorem 13.1 (p. 295) or its alternative formulations in Sections 14.1 and 14.2 above, without having to go through all of the somewhat complicated topological computations. For some of these results [e.g., for a) and the special case $N = 1$ and $n > 2$ in b)], one does not need the full index formula, but rather only the simpler theorem (see Exercise 12.2, p. 290) that the index is a homomorphism $K(TX) \to \mathbb{Z}$.

Results. Let X be a closed manifold of dimension n, $E, F \in \mathrm{Vect}_N(X)$ and $P \in \mathrm{Ell}_k(E, F)$. Then we have $\operatorname{index}(P) = 0$ in the following cases

a) n odd and P a differential operator.

b) $N < n$ and (X a hypersurface in $\mathbb{R}^{n+1}$or $n \leq 3$)
c) $N = n/2$ and Euler number $e(X) \neq 0$
d) $N = n/2$ and n not divisible by four
e) $N < n/2$.

(b)–e): and E and F trivial)

Arguments. We went over b) in Result b) . The derivation c)–e) from the Result a) of Section 14.2 can be found in [**AS68b**, p.602 f]We show that a) follows very nicely from Result b) of Section 14.1: Let $\alpha : \xi \to -\xi$ be the antipodal map on the tangent bundle TX. Since $\sigma(P)$ at the point x written in terms of a matrix of homogeneous polynomials of the k-th degree with coefficients in $\mathbb{C}$ and coordinates in T_x^*X as variables, we have the symmetry condition

$$\sigma(P)\,(\alpha(\xi)) = (-1)^k \sigma(P)(\xi), \quad \xi \in T_x X. \tag{14.1}$$

Here we have identified TX and T^*X by means of a Riemannian metric on X. Via multiplication by $e^{it\pi}$, $t \in [0,1]$, one obtains a homotopy in $\mathrm{Iso}_{SX}(E, F)$ from $\sigma(P)$

to $-\sigma(P)$, hence $[\sigma(P)]$ and $[-\sigma(P)]$ are equal in $K(TX)$. We can then neglect the sign in (14.1) and obtain

$$\alpha^* [\sigma(P)] = [\sigma(P)] \tag{14.2}$$

if P is a differential operator. We now apply Result b) of Section 14.1:

$$\begin{aligned} \operatorname{index} P &= (-1)^n \{ch[\sigma(P)] \cup \tau(X)\} [Tx] \\ &= (-1)^n \{\alpha^* ch[\sigma(P)] \cup \tau(X)\} (\alpha_*[TX]) \\ &= (-1)^n \{ch[\sigma(P)] \cup \tau(X)\} ((-1)^n [TX]) \\ &= -\text{index } P, \text{ whence } \operatorname{index} P = 0. \end{aligned}$$

Here we have used (14.2) in the third equality, to obtain $\alpha^* ch[\sigma(P)] = ch[\sigma(P)]$ in $H^*(TX, \mathbb{Q})$. Note also that α inverts only the vertical part of the tangent space TX, leaving the horizontal part unchanged: in local coordinates $(x_1, ..., x_n)$, with ξ represented by $(x_1, ..., x_n, \xi_1, ..., \xi_n)$ where $\xi = \sum \xi_i dx^i$, we have $\alpha(\xi)$ represented by $(x_1, ..., x_n, -\xi_1, ..., -\xi_n)$. Thus, the orientation of TX is reversed by α, precisely when n is odd.

Incidentally, with somewhat more topology, (see [**AS68b**, p.600] one can show directly for odd n and P a differential operator that $\sigma(P)(\alpha(\cdot))$ and $(\sigma(P)(\cdot))^{-1}$ are stably homotopic, whence $[\sigma(P)] + [\sigma(P)] = 0$ and $[\sigma(P)]$ is of finite order. Then $\operatorname{index} P \in \mathbb{Z}$ must also be of finite order and hence zero, since index : $K(TX) \to Z$ is a homomorphism. In this way, one obtains a) without recourse to the explicit index formula.

One can also directly prove e) for the special case $N = 1$ and $n > 2$ without the full index theorem (see also Exercise 9.15, p. 239, and the literature given there, where the same result is derived topologically in a "pedestrian" way). For trivial line bundles the space of elliptic symbols can be expressed very simply: Since $\mathrm{GL}(1, \mathbb{C}) = \mathbb{C}^\times$ can be contracted to the circle S^1, the index is defined on the set of homotopy classes $[SX, S^1] = H^1(SX; \mathbb{Z})$. Since (by Exercise 9.14, p. 239) the index of an elliptic operator P is zero when its symbol $\sigma(P)$ depends only on x (and not on $\xi \in (SX)_x$), it follows that the index vanishes on the image of π^* in the following long exact cohomology sequence:

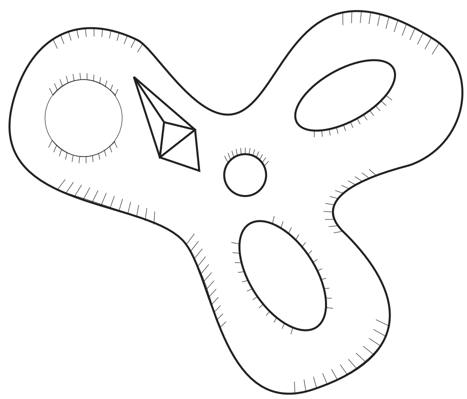
$$\begin{array}{ccccccc} \cdots \to & H^1(BX) & \to & H^1(SX) & \to H^2(BX, SX) \to H^2(BX) \to \cdots \\ & & \nearrow_{\pi^*} & \downarrow \text{index} & \\ & H^1(X) & & \mathbb{Z} & \end{array}$$

(We have omitted the coefficient ring $\mathbb{Z}$ from the cohomology groups here.) As we already reported in (i) of Section 14.1 above, by a classical result of R. Thom, $H^2(BX, SX; \mathbb{Z}) = 0$ for $n > 2$, whence π^* is surjective. Thus, we have proved that $\operatorname{index} P = 0$ for each elliptic operator P defined on the space of complex-valued functions on a manifold of dimension > 3. Compare with [**Ati70a**, p.103 f], where similar topological arguments are needed in certain cases for $n = 2$ and for nontrivial bundles.

4. Euler Number and Signature

We have seen already how (e.g., in our proof of the Bott Periodicity Theorem) analytic methods are utilized in topology and, conversely, how the Index Formula expresses the analytic index by topological means. This deep inner relationship between the topology of manifolds and the analysis of linear elliptic operators is further revealed by the fact that certain invariants of manifolds can be realized as indices of "classical" elliptic operators which can be defined quite naturally on these manifolds. A detailed presentation is contained in [**May**]; for the case bordered manifolds, which remained obscure for a long time (since the "classical" operators do not all admit elliptic boundary value systems in the sense of 10-**??**; see [**Bo72**]), we refer to [**APS75**].

The invariant which we consider here is the Euler number $e(X)$. It is defined by the observation (a matter of solid geometry and probably known long ago to Greek mathematicians) that, for every "triangulation" of a closed oriented surface X, the alternating sum $e(X) = \alpha_0 - \alpha_1 + \alpha_2$ of the number α_0 of vertices (α_0), edges (α_1) and faces (α_2) is the same and depends only on the number of "handles", the "genus" the surface: $e(X) = 2 - 2g$. (The picture shows the beginning of a triangulation and the surface has genus 4.)



The Euler number can be interpreted as the alternating sum $e(X) = \beta_0 - \beta_1 + \beta_2$ of the 0-, 1- and 2-dimensional "holes" whereby the interior of X consists of a 0-dimensional and g 1-dimensional holes and the exterior of further g 1-dimensional holes and the 2-dimensional total space, thus $\beta_0 = \beta_2 = 1$ and $\beta_1 = 2g$.

In the language of singular homology g is the rank of the i-th group of homology $H_i(X;\mathbb{Z})$, the i-th "Betti number", and in this form the definition of the Euler number can be extended to a topological manifold X of dimension $n > 2$. If the α_i are again defined by a triangulation of X, then we obtain

$$e(X) = \alpha_0 - \alpha_1 + \cdots + (-1)^n \alpha_2 = \beta_0 - \beta_1 + \cdots + (-1)^n \beta_n$$

The Euler number is among the best understood topological invariants. For example one knows (see e.g., [**AH**, p.309 and 358], [**ST**, p.246] and [**Gr**, p.99-103]):

$$\begin{aligned} &\text{(i)} \quad e\left(X \times Y\right) = e\left(X\right) \cdot e\left(Y\right), \\ &\text{(ii)} \quad \dim X \text{ odd} \Rightarrow e\left(X\right) = 0, \\ &\text{(iii)} \quad e\left(S^{2m}\right) = 2 \text{ and } e\left(\mathbb{P}^m\left(\mathbb{C}\right)\right) = m+1 \end{aligned}$$

An invariant which is sharper in several aspects (see Section 14.5) is the signature of an oriented topological manifold X of dimension $4q$ defined as follows. Let the real-valued symmetric bilinear form

$$Q : H^{2q}\left(X;\mathbb{R}\right) \times H^{2q}\left(X;\mathbb{R}\right) \to \mathbb{R} \text{ be defined by}$$
$$Q\left(a,b\right) := \left(a \smile b\right)\left[X\right],$$

where $a \smile b$ is the cup product and $[X]$ is the fundamental cycle of the orientation of X. Then

$$\operatorname{sig}(X) := \operatorname{sig}(Q) := p^+ - p^-$$

where p^+ (resp. p^-) denotes the maximal dimension of subspaces of $H^{2q}(X;\mathbb{R})$ on which Q is positive (resp. negative)-definite.

Q is non-degenerate, whence $p^+ + p^- = \beta_{2q}$ (recall $\beta_i = \dim H_i(X;\mathbb{R}) = \dim H^i(X;\mathbb{R})$). Furthermore, the Poincaré duality $H_i(X;\mathbb{R}) \cong H^{4q-i}(X;\mathbb{R})$ implies $e(X) \equiv \beta_{2q} \bmod 2$, and we obtain the formula

$$\text{(iv)}\ \ \dim X = 4q \Rightarrow e(X) \equiv \operatorname{sig}(X) \bmod 2.$$

We now want to describe these two invariants analytically, when the oriented, closed manifold X, with $\dim X = n$ even, is equipped with a differentiable structure and a Riemannian metric. Referring to Exercise 6.17 (p. 169) and Section 13.2 above, let $\Omega := \bigoplus_{j=1}^n \Omega^j$ be the space of (complexified) exterior differential forms, with exterior derivative $d : \Omega \to \Omega$ and its adjoint $d^* : \Omega \to \Omega$.

Results. a) The operator $d + d^* : \Omega \to \Omega$ is an elliptic, self-adjoint differential operator of order 1, whence $\operatorname{index}(d + d^*) = 0$. To show the ellipticity, one first checks that

$$\sigma(d)(x,\xi)\nu = i\xi \wedge \nu \text{ and } \sigma(d^*)(x,\xi)\nu = -i * (\xi \wedge *\nu), \tag{14.3}$$

For this, note that

$$\nu \in \Omega^k,\ \ \alpha \in \Omega^{k+1} \Rightarrow *\alpha \in \Omega^{n-(k+1)} \text{ and } \xi \wedge *\alpha \in \Omega^{n-k}.$$

Thus, if μ is the volume element, we have

$$\begin{aligned}
&\langle \xi \wedge \nu, \alpha\rangle\, \mu = (\xi \wedge \nu) \wedge *\alpha = (-1)^k\, \nu \wedge (\xi \wedge *\alpha) \\
&= (-1)^k\, \nu \wedge \left((-1)^{(n-k)k} **\right)(\xi \wedge *\alpha) = (-1)^{nk} \langle \nu, *(\xi \wedge *\alpha)\rangle\, \mu \\
&= \langle \nu, *(\xi \wedge *\alpha)\rangle\, \mu\ .
\end{aligned}$$

since n is even. The square of $d + d^*$ is the (Hodge) Laplace operator

$$\Delta := (d + d^*)^2 = dd^* + d^*d : \Omega \to \Omega$$

which is of order 2 and homogeneous (i.e., $\Delta\left(\Omega^j\right) \subseteq \Omega^j$). The ellipticity $d + d^*$, (as well as that of Δ) follows, once it is shown that

$$(\sigma(d + d^*)(x,\xi) \circ \sigma(d + d^*)(x,\xi))\, \nu = \|\xi\|^2\, v,$$

since then $(\sigma(d + d^*)(x,\xi))^{-1} = \|\xi\|^{-2}\, \sigma(d + d^*)(x,\xi)$. We have

$$\begin{aligned}
&(\sigma(d + d^*)(x,\xi) \circ \sigma(d + d^*)(x,\xi))\, \nu \\
&= i\xi \wedge (i\xi \wedge \nu - i * (\xi \wedge *\nu)) - i * (\xi \wedge * (i\xi \wedge \nu - i * (\xi \wedge *\nu))) \\
&= \xi \wedge (*(\xi \wedge *\nu)) + *(\xi \wedge (*\xi \wedge \nu - ** (\xi \wedge *\nu))) \\
&= \xi \wedge (*(\xi \wedge *\nu)) + *((\xi \wedge *\xi \wedge \nu - \xi \wedge ** (\xi \wedge *\nu))) \\
&= \xi \wedge (*(\xi \wedge *\nu)) + *((\xi \wedge * (\xi \wedge \nu) \pm \xi \wedge \xi \wedge *\nu)) \\
&= \xi \wedge (*(\xi \wedge *\nu)) + *((\xi \wedge * (\xi \wedge \nu)) = \|\xi\|^2\, \nu,
\end{aligned}$$

where suffices to check the last equality for $\|\xi\| = 1$ in the case $\nu \in \Lambda^k\left(\xi^\perp\right)$, and in the case $\nu = \xi \wedge \eta$ where $\eta \in \Lambda^{k-1}\left(\xi^\perp\right)$, both of which are straightforward (extend ξ to an oriented, orthonormal basis).

For $\Delta_j := \Delta|_{\Omega_j}$, we have the "main theorem of Hodge theory"

$$\operatorname{Ker}\left((d+d^*)\,|_{\Omega_j}\right) = \operatorname{Ker}\Delta_j \cong H^j(X;\mathbb{C}).$$

Moreover, there is a splitting of Ω^j into an L^2 orthogonal direct sums

$$\Omega^j = \operatorname{Ker}\Delta \oplus \operatorname{Im}\Delta = \operatorname{Ker}\Delta \oplus \operatorname{Im} d_{j-1} \oplus \operatorname{Im} d_j^*.$$

For the details, see Theorem 18.57, p. 571, and Corollary 18.58.

b) By restricting $d+d^*$ to the space $\Omega^{\text{even}} := \bigoplus_{j \text{ even}} \Omega^j$ of even forms, we obtain an elliptic differential operator

$$(d+d^*)^{\text{even}} : \Omega^{\text{even}} \to \Omega^{\text{odd}} := \bigoplus_{j \text{ odd}} \Omega^j$$

of order 1. By b) we have

$$\begin{aligned}
\operatorname{index}\left((d+d^*)^{\text{even}}\right) &= \sum_{j \text{ even}} \dim \operatorname{Ker}\Delta_j - \sum_{j \text{ odd}} \dim \operatorname{Ker}\Delta_j \\
&= \sum_{j \text{ even}} \dim H^j(X;\mathbb{C}) - \sum_{j \text{ odd}} \dim H^j(X;\mathbb{C}) = e(X).
\end{aligned}$$

c) If $n \equiv 0 \bmod 4$, say $n = 4q$, and $\tau : \Omega \to \Omega$ is the involution (defined by (13.1), p. 298) with the ± 1-eigenspaces $\Omega^{\pm}$, then $(d+d^*)(\Omega^+) = \Omega^-$. We show that the index of the elliptic operator $(d+d^*)^+ : \Omega^+ \to \Omega^-$ is given by

$$\operatorname{index}(d+d^*)^+ = \operatorname{sig}(X). \tag{14.4}$$

First note that since Δ and τ commute, it follows that

$$\operatorname{Ker}(d+d^*)^+ = (1+\tau)\operatorname{Ker}\Delta = \bigoplus_{j=0}^{2q} (1+\tau)\operatorname{Ker}\Delta_j \text{ and}$$

$$\operatorname{Ker}\left((d+d^*)^+\right)^* = \operatorname{Ker}(d+d^*)^- = (1-\tau)\operatorname{Ker}\Delta = \bigoplus_{j=0}^{2q} (1-\tau)\operatorname{Ker}\Delta_j.$$

For $j < 2q$ (strict), $(1-\tau)\operatorname{Ker}\Delta_j \cong \operatorname{Ker}\Delta_j \cong (1+\tau)\operatorname{Ker}\Delta_j$, and so

$$\begin{aligned}
\operatorname{index}(d+d^*)^+ &= \dim \operatorname{Ker}(d+d^*)^+ - \dim \operatorname{Ker}(d+d^*)^- \\
&= \dim\left((1+\tau)\operatorname{Ker}\Delta_{2q}\right) - \dim\left((1-\tau)\operatorname{Ker}\Delta_{2q}\right) \\
&= \dim\left((1+*)\operatorname{Ker}\Delta_{2q}\right) - \dim\left((1-*)\operatorname{Ker}\Delta_{2q}\right),
\end{aligned}$$

where we have used the fact that for $v \in \Omega^{2q}$,

$$\tau(v) = i^{2q(2q-1)+2q} * v = i^{4q^2} * v = *v.$$

For $n = 4q$ and $k = 2q$, we have $*^2|_{\Omega^k} = (-1)^{k(n-k)} = \mathrm{Id}$. Hence, for $v \in \operatorname{Ker}\Delta_{2q} \cong H^j(X;\mathbb{C})$, we get

$$\int_X v \wedge v = \int_X v \wedge (* * v) = \int_X \langle v, *v\rangle = \begin{cases} \geq 0 & \text{for } * v = v \\ \leq 0 & \text{for } * v = -v, \end{cases}$$

and so $\operatorname{index}(d+d^*)^+ = p^+ - p^-$; i.e., (14.4) holds.

d) As defined in Section 14.1, let $U \in H^n\left(TX, (TX)_0\right)$ be the orientation class and let $\Phi : H^n(X) \to H^{2n}\left(TX, (TX)_0\right)$ be the Thom isomorphism. There is a characteristic class $\chi(TX) = \Phi^{-1}(U \smile U) \in H^n(X;\mathbb{Z})$, namely the *Euler class* of the tangent bundle TX, for which $\chi(TX)[X] = e(X)$. The class $\chi(TX)$ is represented by the a certain n-form $\mathrm{GB}\left(\Omega^\theta\right)$, the *Gauss-Bonnet form*, which is defined in terms of the curvature Ω^θ of X; i.e., $\mathrm{GB}\left(\Omega^\theta\right)$ is a multiple of the Pfaffian of Ω^θ. Thus,

$$e(X) = \int_X \mathrm{GB}\left(\Omega^\theta\right), \tag{14.5}$$

and this is known as the Gauss-Bonnet-Chern formula. The formula (14.5) is proved directly (i.e., without reference to U, Φ, or the fact that $\Phi^{-1}(U \smile U)[X] = \chi(TX)[X] = e(X)$) using the Local Index Theorem in Part IV (see Theorem 18.63, p. 580) applied to certain components of the operator $(d+d^*)^{\text{even}}$. When $n=2$, the formula reduces to the classical result $e(X) = \frac{1}{2\pi}\int_X K$ of C. F. Gauss and O. Bonnet, where K is the Gaussian curvature of the closed surface X.

e) If $n = 4q$, the signature $\operatorname{sig}(X)$ can also be expressed as

$$\operatorname{sig}(X) = L_q(p_1,\ldots,p_q)[X] = \int_X L_q(\widetilde{p}_1,\ldots,\widetilde{p}_q), \tag{14.6}$$

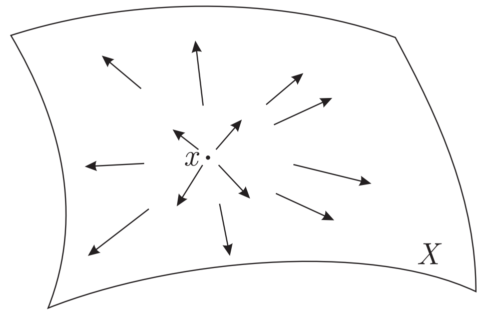
which is known as the Hirzebruch Signature Formula. Here the p_j are the Pontryagin characteristic classes which are defined in terms of Chern classes via $p_j := (-1)^j c_{2j}(TX\otimes\mathbb{C})$ and the polynomial $L_q(p_1,\ldots,p_q)$ (in which the p_j are multiplied via cup product) is described as follows. We have an expansion

$$\frac{x_1}{\tanh x_1}\cdots\frac{x_q}{\tanh x_q} = \sum_{k=1}^{\infty} L_k(\sigma_1,\ldots,\sigma_q)$$

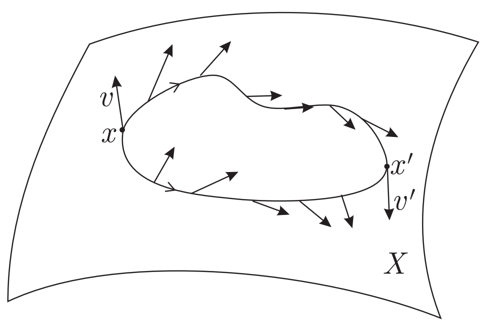
where $\sigma_1,\ldots,\sigma_q$ denote the elementary symmetric polynomials in $x_1^2,\ldots,x_q^2$ and $L_k(\sigma_1,\ldots,\sigma_q)$ is ultimately homogeneous of degree k in $x_1^2,\ldots,x_q^2$. We then replace $\sigma_1,\ldots,\sigma_q$ in $L_q(\sigma_1,\ldots,\sigma_q)$ by $p_1,\ldots,p_q$ to obtain $L_q(p_1,\ldots,p_q)$. In particular, $L_1(p_1) = \frac{1}{3}p_1$, $L_2(p_1,p_2) = \frac{1}{45}\left(7p_2 - p_1^2\right)$, etc.; see (16.109), p. 421. One can represent p_k by a $4k$-form $\widetilde{p}_k$ which involves the curvature tensor of X. The formula for $\widetilde{p}_k$ in terms of the curvature tensor of X is found in Section 16.7, specifically (16.100), p. 418, where $\widetilde{p}_j$ is denoted there by $p_k\left(\Omega^\theta\right)$ to indicate its dependence on the curvature form Ω^θ of the Levi-Civita connection θ. In $L_q(\widetilde{p}_1,\ldots,\widetilde{p}_q)$ the forms $\widetilde{p}_k$ are multiplied via wedge product. For a statement and proof (using the Local Index Theorem) of the twisted generalization of the Hirzebruch Signature Formula, see Theorem 18.60, p. 574.

5. Vector Fields on Manifolds

Among the basic concepts of the analysis of a dynamical system, as well as of the geometry of a differential manifold X, is the notion of a "vector field". It is a C^∞ section v of the tangent bundle (in classical terminology a "variety of line elements"), thus $v \in C^\infty(TX)$.



Each v defines an ordinary differential equation $c'(t) = v(c(t))$ for differentiable paths ("trajectories") $c : \mathbb{R} \to X$. According to the classical existence and uniqueness theorems, the equation has a unique solution (at least defined on some open interval about 0) for any given "initial value" $c(0) = x \in X$; for details see [**BJ**, p.74-87]. In every theory of flows the "singularities" in $\Sigma := \{x \in X : v(x) = 0\}$ are of special interest, as they are the "stagnation points" of dynamics, the equilibrium points, or the so-called "stationary solutions" (the constant paths $c(t) = x$). For applications, one may think intuitively of examples from oceanography, of a magnetic field, or of dynamic laws of economics. We saw above in Section 11.1 that every isolated singularity x of a vector field v defines locally a map $S^{n-1} \to S^{n-1}$, $n = \dim X$, whose mapping degree we denote by $I_v(x)$. The geometric interest in vector fields usually stems from the question of "parallelizability" of the manifold: Is it possible to assign to a tangent vector $v \in T_xX$ at a point x a tangent vector $v' \in T_{x'}X$ "parallel" to v in a way which is independent of the path from x to x' used in the process?



This question is important for the physical notion of space and finally for the analysis of motion (since the concept of acceleration depends on parallel displacement). It is equivalent to the question of whether an n-dimensional manifold X possesses n vector fields that are linearly independent at each point $x \in X$; i.e., whether the tangent bundle TX is trivial. (For example, by a famous theorem of J. F. Adams, S^{n-1} is parallelizable if and only if $\mathbb{R}^n$ is a "division algebra", i.e., for $n = 2m$ and $m = 1, 2, 4$.)

The geometrical idea and also the historical origin (with Eduard Stiefel, 1935, and almost simultaneously and in a similar connection with Hassler Whitney) of the topological invariants of Riemannian or complex manifolds (nowadays called "characteristic classes"– see Section 14.1 above) lies in the general investigation of r vector fields $v_1, ..., v_r$ on a manifold and their singular set Σ consisting of those points x where, $v_1(x), ..., v_r(x)$ are linearly dependent. One knows, for example, that set Σ generically has dimension $r-1$, and that the "cycle" Σ defines a homology class (with suitable coefficient group) which is a "characteristic class" of X and is independent of $v_1, ..., v_r$; see also the large survey [**Th**].

The index formula (see Section 14.1 above) says that the symbol of an elliptic operator is a kind of characteristic class, and its index is a "characteristic" integer. From this arose the program to illuminate the connection between elliptic operators and vector fields on manifolds. Specifically, the existence of a certain number of vector fields implies certain symmetry properties for "classical operators" and corresponding results for their indices such as Euler number, signature and "characteristic numbers".

Results. Let X be a closed, oriented, Riemannian manifold.

a) Each vector field v "generically" has only finitely many zeros (i.e., by an arbitrarily small perturbation a vector field can be put in this form), and we have $e(X) = \sum_{x\in\Sigma} I_v(x)$, where $I_v(x)$ is the local index of v at x defined above. Thus, the right side of the formula consists of the finite weighted sum of zeros of v.

b) If $\dim X = 4q$ and X has a tangent field of 2-dimensional planes (i.e., an oriented 2-dimensional subbundle of the tangent bundle), then $e(X) \equiv 0 \bmod 2$ and $\operatorname{sig}(X) = e(X) \bmod 4$.

c) If the $4q$-dimensional manifold X has at least r vector fields that are independent everywhere, then $\operatorname{sig}(X) = 0 \bmod b_r$, where the values of b_r are given in the following table ($b_{r+8} = 16 b_r$):

r	1	2	3	4	5	6	7	8
b_r	2	4	8	16	16	16	16	32

Arguments: For $\dim X = 2$, a) is a classical result of H. Poincaré, for which one can find a very clear sketch of the proof in [**Bri76**, p.166-171]. The generalization for $\dim X > 2$ originated from H. Hopf, who also showed that $e(X) = 0$, if and only if, there is a nowhere zero vector field on X. For b) and c), we refer to [**May**], [**Ati70c**] and [**AD**], where a series of related results are proved by combining results of analysis, topology, and algebra, and in some cases using the theory of real elliptic operators (see Section 14.9 below). To illustrate the methods, we will only treat the theorem

$$\exists v \in C^\infty(TX) \text{ with } v(x) \neq 0 \ \forall x \in X \Rightarrow e(X) = 0.$$

One can obtain this from the explicit Atiyah-Singer Index Formula or the Gauss-Bonnet Formula which says that $e(X) = \chi(TX)[X]$ (see Section 14.4d). By definition, $\chi(TX) = \Phi^{-1}(U \smile U) = \tilde{v}^* i^* U$, where $U \in H^n(BX, SX; \mathbb{Z})$ is the orientation class of TX, $\Phi : H^j(X;\mathbb{Z}) \to H^{j+n}(BX, SX; \mathbb{Z})$ is the Thom isomorphism, $i : (BX, \phi) \to (BX, SX)$ is the trivial embedding, and $\tilde{v} : X \to BX$ is the normalized vector field $\tilde{v} = v/|v|$, where $|v| \neq 0$ by assumption. Since $i\tilde{v}(X) \subset SX$, we have $\tilde{v}^* i^* = 0$.

To get the same result, without recourse to the explicit index formula, one can apply the general theory of elliptic operators (see Chapter 9 above) and the results of Section 14.4 b and c of Hodge theory. The key issue for this is the fact that $\sigma(d + d^*)(x, \xi) w = i\xi \times w$ for $x \in X$, $\xi \in T_x^* X$, and $w \in \Lambda^*(T_x^* X)$. Here $\times : \Lambda^*(V) \times \Lambda^*(V) \to \Lambda^*(V)$, $V := T_x^* X$ is the Clifford product (in contrast to the "exterior" multiplication) given (in particular) for $\xi \in T_x^* X$ and $w \in \Lambda^*(T_x^* X)$ by

$$\xi \times w = \xi \wedge w - \xi \llcorner w \text{ and } w \times \xi = \xi \wedge w - w \lrcorner \xi$$

where $\llcorner$ and $\lrcorner$ are left and right interior multiplication. [Note $\xi \llcorner w = *(\xi \wedge *\nu)$ when n is even, so that indeed $\sigma(d+d^*)(x,\xi)w = i\xi \times w$ by (14.3), p. 311]. A general treatment of Clifford algebras is provided in Section 18.1. By means of the Riemannian metric on X, a vector field v can be regarded as a 1-form which yields a 0-th order differential operator $R_v : \Omega^*(X) \to \Omega^*(X)$ given by $R_v(u) := u \times v$, $u \in \Omega^*(X)$ (i.e., via pointwise right Clifford multiplication by v). Since $(u \times v) \times v = u \times (v \times v) = -|v|^2 u$, $|v| > 0$ implies that R_v is a automorphism of $\Omega^*(X)$ which maps $\Omega^{\text{even/odd}}(X)$ bijectively to $\Omega^{\text{odd/even}}(X)$. Since

$$\begin{aligned} R_v \circ \sigma(d+d^*)(x,\xi) &= w(i\xi \times w) \times v(x) \\ &= i\xi \times (w \times v(x)) = \sigma(d+d^*) \circ R_v(x,\xi), \end{aligned}$$

$\sigma(d+d^*)$ commutes with $\sigma(R_v) = R_v$. Let $R_v^{\text{odd}} := R_v|_{\Omega^{\text{odd}}(X)}$. Then

$$\left(R_v^{\text{odd}}\right)^{-1} \circ (d+d^*)^{\text{even}} \circ R_v^{\text{odd}} - (d+d^*)^{\text{odd}} \in \mathrm{OP}_0 \,.$$

Using $((d+d^*)^{\text{even}})^* = (d+d^*)^{\text{odd}}$, we then have

$$\begin{aligned} \operatorname{index}(d+d^*)^{\text{even}} &= \operatorname{index}\left(\left(R_v^{\text{odd}}\right)^{-1} (d+d^*)^{\text{even}} \circ \circ R_v^{\text{odd}}\right) \\ &= \operatorname{index}(d+d^*)^{\text{odd}} = -\operatorname{index}(d+d^*)^{\text{even}}, \end{aligned}$$

whence $e(X) = \operatorname{index}(d+d^*)^{\text{even}} = 0$.

6. Abelian Integrals and Riemann Surfaces

Perhaps one of the first (in our topological sense) "quantitative" results of analysis is contained in the major work of the Norwegian mathematician Niels Henrik Abel entitled "Mémoire sur une propriété générale d'une classe trés étendue de fonctions transcendantes". It was written in 1826, but published only posthumously in 1841. In it Abel takes up a dispute of the numerical analysis of the 18th century about "rectifiability", the possibility of solving integrals by means of elementary functions (algebraic functions, circular functions, logarithm and exponential functions). Since Jakob Bernoulli and Gottfried Wilhelm Leibniz[1] there was an interest particularly in the integration of irrational functions, turn up in many problems of science and technology. The efforts (already of the 17th century) to rectify the ellipse, whose arc length important for astronomy, leads to the computation of the integral

$$I(x) = a \int_0^x \frac{1 - k^2 t^2}{\sqrt{(1-t^2)(1-k^2t^2)}} dt, \quad k = 1 - \frac{b^2}{a^2}.$$

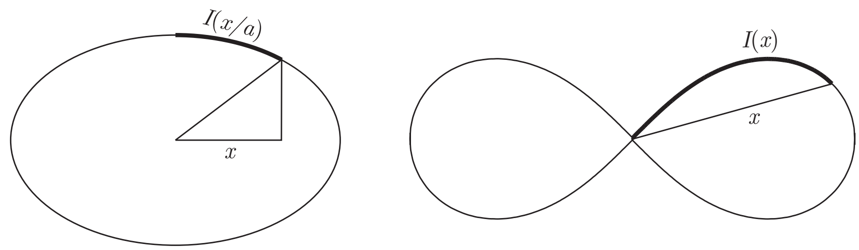
While investigating the deformation of an elastic rod under the influence of forces acting at its extremities, Jakob Bernoulli (1694) came across further irrational integrands. In this connection he also introduced the "lemniscate"

$$\left\{ (\pm\sqrt{(x^2+x^4)/2}, \pm\sqrt{(x^2-x^4)/2}) : 0 \le x \le 1 \right\},$$

whose arclength is given by

$$I(x) = \int_0^x \frac{dt}{\sqrt{1-t^4}}.$$

[1]Who drew the attention of the mathematicians to constructions and curves "quas natura ipsa simplici et expedito motu producere potest" (which nature itself can produce by simple and complete motions), quoted after [**BN**, p.124]. See also [**Kli**, p.411 f].



(For this and the following, see [**Sie**, 1.1-1.2].) For the "lemniscate integral" Leonhard Euler proved (1753) the addition theorem

$$I(u) + I(v) = I(w) \text{ where } w = \frac{u\sqrt{1-v^4} + v\sqrt{1-u^4}}{1+u^2v^2}$$

It generalized a result of the Italian mathematician Giulio Carlo de'Toschi di Fagnano (1714) on the doubling of the circumference of the lemniscate with compass and ruler alone:

$$2I(u) = I(w) \text{ for } w^2 = \frac{4u^2\left(1-u^4\right)}{\left(1+u^4\right)^2}$$

Prior to this, Johann Bernoulli (1698) "accidentally" discovered that the difference of two arcs of the cubic parabola ($y = x^3$) is integrable by elementary functions and added to,the problem of rectifying curves[2] the new problem of finding arcs of parabolas, ellipses, hyperbolas etc. whose sum or difference is an elementary quantity – "just as arcs of a circle could be compared with one another through the expressions for $\sin(\alpha+\beta)$, $\sin 2\alpha$, etc." [**BN**, p.206].

A little later L. Euler succeeded in extending his addition theorem for the leminiscatic integral to "elliptic integrals of the first kind" i.e., in proving that

$$I(u) + I(v) = I(w) \text{ with } w = \frac{u\sqrt{P(v)} + v\sqrt{P(u)}}{1+u^2v^2}, \text{ where} \tag{14.7}$$

$$I(u) = \int_0^u \frac{dt}{\sqrt{P(t)}} \text{ and } P(t) := 1 + at^2 - t^4$$

Based on a comparison of elliptic arcs also due to Fagnano, Euler finally found another generalization to "elliptic integrals of higher kind". These are integrals of the form

$$I(u) = \int_0^u \frac{r(t)}{\sqrt{P(t)}} dt,$$

where r is a rational function in one variable and P is a polynomial of third or fourth degree with simple zeros. Here the addition theorem takes the form

$$I(u) + I(v) = I(w) + W(u,v) \tag{14.8}$$

where w is, as above, an algebraic function of the arbitrary upper integration limits u and v and $W(u,v) = S_1(u,v) + \log S_2(u,v)$ with rational functions S_1, S_2.

[2]It was suspected already that this was impossible. But only in 1835, did J. Liouville prove rigorously that the "elliptic integrals" could not be solved "elementarily", i.e., expressed by a finite combination of algebraic, circular, logarithmic and exponential functions. (As an aside, it is now known that the problem of the elementary integration of arbitrary functions is recursively undecidable.)

Euler already noticed that his methods cannot be used for the treatment of "hyperelliptic integrals" (with a polynomial P of fifth or higher degree). But only N. H. Abel found the explanation for "the difficulties which Euler's formulation necessarily encountered when dealing with hyperelliptic integrals: The constant of integration ($I(w)$ in (14.7) or (14.8)), appearing in the transcendental equation, could not be replaced, as in the elliptic case, by a single integral but only by two or more hyperelliptic integrals – a remarkable circumstance which was in no way predictable... The question about the minimal number of integrals which a given sum of integrals could be reduced, remained as the cardinal question; it caused Abel to produce the elaborate and laborious counts which constitute the main results of his great Paris paper and which brought him into the possession of the notion of "genus" of an algebraic structure long before Riemann" [**BN**, p.211 f]

Abelian Addition Theorem. Let R be a rational function and F a polynomial in two variables. For $a, x \in \mathbb{R}$ (a fixed), consider the "Abelian integral" $I(x) := \int_a^x R(t,y)dt$, where y satisfies the equation $F(t,y) = 0$. (For $F(t,y) = y^2 - P(t)$ with P as above and $R(t,y) = 1/y$, we obtain an elliptic integral of the first kind; and for $R(t,y) = r(t)/y$ with r a rational function, we get an elliptic integral of a higher kind.) For a given value of t, there may be several corresponding solutions ("roots") of the equation $F(t,y) = 0$. Thus, one must specify which root will be substituted for y in $R(t,y)$. Hence, one selects an integration path $\gamma : I \to \mathbb{C}$ in the plane (with $\operatorname{Re}\gamma(0) = a$, $\operatorname{Re}\gamma(1) = x$, and $F(\operatorname{Re}\gamma, \operatorname{Im}\gamma) = 0$), and regards $I(x)$ as a line integral. Then the sum of m (m sufficiently large) arbitrary Abelian integrals $I(x_i)$ with respective fixed integration paths ($x_i \in \mathbb{R}$, $i = 1, ..., m$) can be written as the sum of only g Abelian integrals $I(\hat{x}_j)$ ($j = 1, ..., g$) and a remainder term $W(x_1, ..., x_m)$ which is the sum of a rational function S_1 and the log of another rational function S_2 of the limits of integration $x_1, ..., x_m$:

$$\sum\nolimits_{i=1}^{m} I(x_i) = \sum\nolimits_{j=1}^{g} I(\hat{x}_j) + W(x_1, ..., x_m), \text{ where}$$

where the $\hat{x}_j = \hat{x}_j(x_1, ..., x_m)$ are algebraic functions and the number g only depends on the specific form of the polynomial F.

According to C. G. J. Jacobi, Abel attempted to solve two problems with his addition theorem, "the representability of an integral by closed expressions", and "the investigation of general properties of integrals of algebraic functions" [**BN**, p.205]. In fact, for one thing, the addition theorem says something about the rectification question: If $g = 0$ then $I(x) = W(x)$ for $m = 1$ where W is constructed from rational functions and the logarithm. Thus, in this case, the integral $I(x)$ can be solved by elementary functions. If $g > 0$ then in general at least g additional higher transcendental functions are needed, namely the $I(\hat{x}_j)$. Most of all, the theorem is an addition theorem just like the sum formulas for trigonometric functions, and can be used in the back-up files of computers as interpolation formulas for realizing standard functions of science and technology.[3]

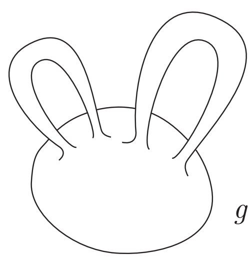
We quoted Abel's theorem, in order to point out one of the earliest occurrences of the fundamental invariant g which, as the index, possess the dual character of being both analytic and algebraic. Without entering a discussion of the deep function theoretic aspects and geometric interpretations of Abel's addition theorem, we will

[3]See Hart, J. F. and E. W. Cheney, *Computer Approximations*, SIAM Series in Applied Mathematics, Wiley, New York, 1968, §§1.5, 4.2.

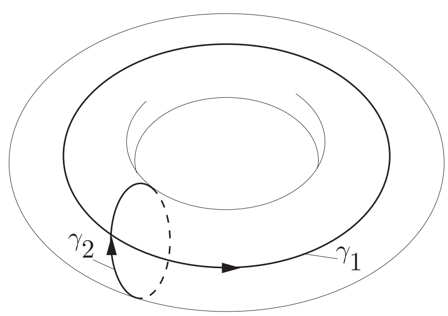
note some results exhibiting the significance of the quantity g. These essentially are due to Bernhard Riemann (except c).

Results. a) Each polynomial $F(t, y)$ defines a compact *Riemann surface*, an oriented surface with a complex-analytic structure (i.e., a topological manifold with a distinguished atlas whose coordinate changes are holomorphic functions). Conversely, one can define a complex-analytic structure on each compact, oriented topological surface X such that Xcan be regarded as the Riemann surface of a polynomial $F(t, y)$.

b) Topologically, a compact Riemann surface X is characterized by its "genus" g, the number of handles that must be fastened to the sphere in order to obtain X. Twice g is the number of closed curves needed to generate the first homology of X; i.e., there are $\gamma_1, \dots, \gamma_{2g}$ closed curves (namely, along the lengths and girths of the handles) such that every closed curve γ in X is "homologous" to a unique integral linear combination $\sum n_i \gamma_i$.



Analytically (see the Abelian Addition Theorem) X has another invariant, the maximal number g_1 of linearly independent holomorphic differential forms $\alpha_1, ..., \alpha_{g_1}$ of degree 1 on X. Actually, $g = g_1$; i.e., the "numerical complexity"[4] of the Abelian integral $\int R(t, y)dt$ with $F(t, y) = 0$ is equal to the genus of the Riemann surface F. In particular, for the elliptic integrals, one obtains an "*elliptic curve*", namely the torus of genus 1 with two generating cycles γ_1 and γ_2 (For the connection with the classical notion of "double periodicity" in elliptic integrals, see [**Mu**, p.149-155], for example.)



More generally, one can define a "periodicity matrix" $\omega_{ij} := \int_{\gamma_i} \alpha_j$ for $i = 1, ..., 2g$, which determines the complex-analytic structure of X by a theorem of R. Torelli.

c) On each Riemann surface X of genus g there is a "*Cauchy-Riemann operator*" $\bar{\partial} : f \mapsto \frac{\partial f}{\partial \bar{z}} d\bar{z}$ which assigns to each complex-valued C^∞ function f on X a complex deferential form of degree 1. $\bar{\partial}$ is an elliptic operator and index $\bar{\partial} = 1 - g$.

[4]Riemann (1857) calls this quanity the "Klassenzahl" (class number). The term "genus" originated with Alfred Clebsch (1864).

Arguments: a) and b) follow from the classical theory of Riemann surfaces. For c), note that $\dim \operatorname{Ker} \bar{\partial} = 1$, since $\operatorname{Ker} \bar{\partial}$ consists of the global holomorphic functions and these must be constant by the maximum principle. $\operatorname{Coker} \bar{\partial} = \operatorname{Ker} \bar{\partial}^*$ consists of the "(anti-) holomorphic differential forms" which constitute a vector space isomorphic to the space of holomorphic 1-forms.

One can also directly obtain c) from the Atiyah-Singer Index Formula, since the Euler class $\chi(TX)$ is the only Chern class of a Riemann surface, whence for some constant C

$$\operatorname{index} \bar{\partial} = C\chi(TX)[X] = Ce(X) = C(1 - 2g + 1) = C(2 - 2g).$$

It is then not difficult to derive that $C = 1/2$; see also the following Section 14.7 (Results a), c) and d)), which include c) as a special case; and [**Mu**, p.132-141] where c) is proven, with these generalizations in mind, in the form "arithmetic genus = geometrical genus" by elementary geometrical and algebraic tools of classical projective geometry.

7. The Theorem of Riemann-Roch-Hirzebruch

We deal next with a class of theorems for which the Atiyah-Singer Index Formula yields new proofs or generalizations (see d)). According to [**BN**, p.280 f], who introduced the term "Riemann-Roch Theorem", it deals with the "counting of the constants of an algebraic function", and more generally, with establishing relations between the "constants" (number of singularities, degree, order, genus etc.) of an algebraic curve, algebraic surface or complex manifold. Thus the history and motivation of "Riemann-Roch" – detailed in [**BN**][5] – are closely tied to the unsolved problem of "completely" classifying algebraic varieties, see [**Hi73**]. We cannot convey the abundance of results in this subject. They are still too scattered and the diversity of approaches too uncertain[6]. We therefore restrict ourselves to a few stages which are essential for us, and where, out of the complexity of the problems, developed little by little, some unifying, very rich, and consequential aspects:

– Bernhard Riemann's "transcendental" idea of the analysis on Riemannian manifolds, i.e., his attempt to consider the totality of integrals of a fixed algebraic function field (just the Abelian integrals).

– The function theoretic treatment of the problems by Karl Weierstrass.

– The interpretation from the point of view of differential geometry in the language of Hodge theory due to Kunihiko Kodaira which was the basis for Friedrich Hirzebruch's generalization of Riemann-Roch to higher dimensional algebraic varieties.

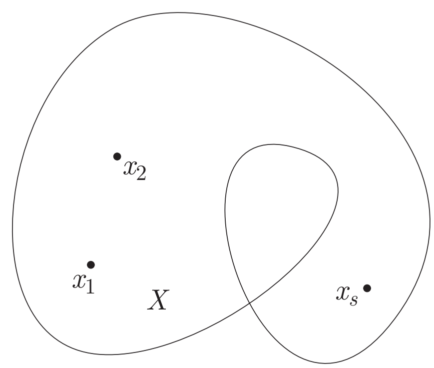
Results: a) On a compact Riemann surface X of genus g, we consider meromorphic functions $w : X \to \mathbb{C}$ which have poles at the points $x_i \in X$ $(i = 1, ..., r)$ of order at most $m_i \in \mathbb{N}$ and zeros at the points $x_j \in X$ $(j = r + 1, ..., s)$ of order

[5]For a function theoretic interpretation of Riemann-Roch, see also [**Sie**, 4.6-4.7], where Riemann-Roch is used to prove Abel's theorem and is expressly called "algebraic" in contrast to the "transcendental nature" of Abel's Theorem.

[6]Alfred Clebsch, "who surely did not lack knowledge and versatility" wrote very openly in a letter of August 1864 to Gustav Roch that he himself "under- stood very little of Riemann's treatise even after greatest efforts, and that Roch's dissertation remained for the most part incomprehensible to him." (Quoted after [**BN**, p.320].) The historical process of "correctly understanding" the theorem apparently has not reached its conclusion at least as far as algebraic functions are concerned.

at least $-m_j \in \mathbb{N}$. These form a complex vector space $L(\mathcal{D})$ whose dimension $l(\mathcal{D})$ is given by the following formula

$$l(\mathcal{D}) - l'(\mathcal{D}) = \deg(\mathcal{D}) - g + 1,$$



where $\mathcal{D} := \sum_{j=1}^{s} m_j x_j$ (formal sum), $\deg(\mathcal{D}) := \sum_{j=1}^{s} m_j$ and $l'(\mathcal{D})$ is the dimension of vector space $L'(\mathcal{D})$ of meromorphic differential 1-forms on X with the corresponding behavior on the zeros and poles.

b) For $\deg(\mathcal{D}) \geq 2g-1$, $l'(\mathcal{D})$ vanishes, whence one obtains a proper formula for $l(\mathcal{D})$ in this case.

c) For each formal integral linear combination $\mathcal{D}$ of points of X called a *divisor*), there is a holomorphic vector bundle $\{\mathcal{D}\}$ of complex fiber dimension 1 such that $L(\mathcal{D})$ is isomorphic to the vector space $\operatorname{Ker} \bar\partial_{\{\mathcal{D}\}}$ of holomorphic sections of $\{\mathcal{D}\}$, and $L'(\mathcal{D})$ is isomorphic to the vector space $\operatorname{Coker} \bar\partial_{\{D\}} \cong \operatorname{Ker} \bar\partial^*_{\{D\}}$ of (anti-) holomorphic 1-forms "with coefficients in $\{\mathcal{D}\}$". Here, $\bar\partial_{\{\mathcal{D}\}} : \Omega^0(\{\mathcal{D}\}) \to \Omega^{0,1}(\{\mathcal{D}\})$ is the elliptic differential operator obtained from $\bar\partial$ by "tensoring (see d) below). With this construction, we can write result a) in the form $\operatorname{index} \bar\partial_{\{\mathcal{D}\}} = \deg(\mathcal{D}) - g + 1$.

d) More generally, one can consider the Dolbeault complex

$$0 \to \Omega^0 \xrightarrow{\bar\partial} \Omega^{0,1} \xrightarrow{\bar\partial} \cdots \xrightarrow{\bar\partial} \Omega^{0,n} \to 0$$

for a Kähler manifold X of (complex) dimension n (i.e., a complex manifold with Hermitian metric $\sum g_{ik} dz_i d\bar z_k$, whose associated 2-form $\alpha = \sum g_{ik} dz_i \wedge d\bar z_k$ is closed, i.e., $d\alpha = 0$). Here $\Omega^{0,p}$ denotes the space of complex exterior differential forms of degree p, which can be written in the form $\sum_{(i)} a_{i_1 \cdots i_p} d\bar z_{i_1} \wedge \cdots \wedge d\bar z_{i_p}$ relative to local coordinates $z_1, \ldots, z_n$. As in Exercise 6.17b (p. 169), $\bar\partial$ is the exterior derivative. If V is a holomorphic vector bundle over X, then via tensoring (as above with the signature operator) one can construct the generalized Dolbeault complex

$$0 \to \Omega^0(V) \xrightarrow{\bar\partial_V} \Omega^{0,1}(V) \xrightarrow{\bar\partial_V} \cdots \xrightarrow{\bar\partial_V} \Omega^{0,n}(V) \to 0,$$

where $\Omega^0(V) = C^\infty(V)$. Using the fact that V is holomorphic, one verifies that $\bar\partial_V^2 = 0$, and thus one can define the Dolbeault cohomology groups

$$H^q(\mathcal{O}_V) := \frac{\operatorname{Ker}\left(\bar\partial_V|_{\Omega^{0,q}(V)}\right)}{\bar\partial_V\left(\Omega^{0,q-1}(V)\right)}.$$

In analogy with the above definition of the Euler number, one defines the *Euler characteristic of* V as

$$\chi(X,V) := \bigoplus_{q=0}^{n} (-1)^q \dim H^q(\mathcal{O}_V).$$

As with Hodge theory in the real category

$$H^q(\mathcal{O}_V) \cong \mathcal{H}^q(V) := \operatorname{Ker}\left(\left(\bar\partial_V + \bar\partial_V^*\right)|_{\Omega^{0,q}(V)}\right) = \operatorname{Ker}\left(\bar\Box_V|_{\Omega^{0,q}(V)}\right),$$

where $\bar\Box_V := \left(\bar\partial_V + \bar\partial_V^*\right)^2 = \bar\partial_V\bar\partial_V^* + \bar\partial_V^*\bar\partial_V$, so that

$$\chi(X,V) = \operatorname{index}\left(\left(\bar\partial_V + \bar\partial_V^*\right)^{\text{even}} : \bigoplus_{j \text{ even}} \Omega^{0,j}(V) \to \bigoplus_{j \text{ odd}} \Omega^{0,j}(V)\right)$$

The "Riemann-Roch-Hirzebruch Theorem" then states that

$$\chi(X,V) = (ch(V) \smile \tau(TX))[X]. \tag{14.9}$$

Here $ch(V)$ is the Chern character of V and $\tau(TX)$ is the Todd class of X; see Section 14.1 above.

If V is the trivial line bundle $\mathbb{C}_X$, then $\chi(X) := \chi(X, \mathbb{C}_X)$ is called the *arithmetic genus* of X and (14.9) states that this is the same as the *Todd genus* $\tau(TX)[X]$. We denote by c_i^p the evaluation $c_i(TX)^p[X]$ of the p-th power (with $i{\cdot}p = \dim_{\mathbb{C}} X$) of the Chern class $c_i(TX)$ on the fundamental class $[X] \in H_{2n}(X;\mathbb{Z})$ (c_i^p is called a "Chern number"). One then computes the table

$\dim_{\mathbb{C}} X$	1	2	3	4
$\chi(X)$	$\frac{1}{2}c_1$	$\frac{1}{12}\left(c_2 + c_1^2\right)$	$\frac{1}{24}c_1c_2$	$\frac{1}{720}\left(-c_4 + c_3c_1 + 3c_2^2 + 4c_2c_1^2 - c_1^4\right).$

Arguments: **a**) follows using c) and d) from the Atiyah-Singer Index Formula; a direct proof can be found in [**We64**, p.117-119]. If $\mathcal{D} = 0$ (in particular, $\deg(\mathcal{D}) = 0$), then one recovers results b) and c) of Section 14.6. Notice that $l'(\mathcal{D}) = l(K_X - \mathcal{D})$, where K_X like $\mathcal{D}$ is a divisor which canonically (up to a certain "linear equivalence") defined by the zeros and poles of an arbitrary nontrivial form on X. One has $\deg(K_X) = 2 - 2g$, For details of the definition of "canonical divisors", and for an elementary proof of $l(\mathcal{D}) - l(K_X - \mathcal{D}) = \deg(\mathcal{D}) - g + 1$, see [**Mu**, p.104-107 and 145-147]; the duality $l'(\mathcal{D}) = l(K_X - \mathcal{D})$ originated with Richard Dedekind and Heinrich Weber, and is a special case of a very general, duality theorem of Jean-Pierre Serre (cf. [**Hi66a**, 15.4]).

For b): A non-trivial meromorphic function on a compact Riemann surface has as many poles as zeros (counting orders). Thus, $l(\mathcal{D}) = 0$ if $\deg(0) < 0$. Since $\deg(K_X) = 2g - 2$, it follows that $\deg(K_X - \mathcal{D}) < 0$, for $\deg(\mathcal{D}) \geq 2g - 1$.

For c): For the construction of the line bundle $\{\mathcal{D}\}$, we cover X with a finite collection of open sets $\{U_j\}_{j\in J}$ such that on each U_j there is defined a meromorphic function f which has the zero and pole behavior that is the opposite of the portion of $\mathcal{D}$ involving points of U_j; so that f_i/f_j is a nowhere-zero holomorphic function on $U_i \cap U_j$. For example, one can choose U_j so small that at most one point (of the finitely many points $x_1, ..., x_s$ found in $\mathcal{D}$) lies in U_j, and then take f_j to be a locally defined rational monomial which has a zero of order m_k if $m_k > 0$, and a pole of order $-m_k$ if $m_k < 0$. The clutching construction (Exercise B.7, p. 681, of the Appendix) then yields the bundle $\{\mathcal{D}\}$ for which $\{f_i\}$ defines a global meromorphic section f with zero behavior opposite $\mathcal{D}$. Via $g \mapsto g/f$, an isomorphism is defined from the global holomorphic sections of $\{\mathcal{D}\}$ to the vector space of meromorphic functions on X with the zero and pole behavior prescribed by the divisor $\mathcal{D}$. Details of this construction, which is rather typical for the topological approach to analytic problems of function theory and algebraic geometry, can be found in [**Hi66a**, 15.2].

For d): The derivation from the Atiyah-Singer Index Formula can found in Section 18.6 (beginning p. 583); see also [**Pal65**, p.324 ff] and [**Schw66**], and in a somewhat more general context (see Section 14.11 below) also in [**AS68b**, p.563-565]. This transition from the classical Riemann-Roch Theorem to complex manifolds of arbitrary dimension was initiated by Friedrich Hirzebruch (1953). His proof depended essentially on the additional condition that X can be holomorphically embedded in a complex projective space of a suitable dimension. By going back to the index formula for elliptic operators, one can drop this restriction. Moreover, recently a rather differential geometric proof using the heat equation has been given for d). Also, our proof in Section 18.6 ultimately depends on the heat equation approach to the Local Index Theorem for Dirac operators. For the rather easily calculable ease of an elliptic curve ($n = 1$, $g = 1$) refer to [**ABP**, p.311 f], and for a sketch of the general case see [loc. cit., p. 317 f].

The Riemann-Roch-Hirzebruch Theorem appears here as a special application of the Atiyah-Singer Index Formula for elliptic operators of which one, namely $\bar{\partial}_V + \bar{\partial}_V^*$, was constructed taking advantage of the complex structure. Even so, d) and already a) are the quintessential models for the structure of the Index Formula: On the left side we have the difference of globally defined quantities each of which can change even under small variations of the initial data (as the dimension $l(\mathcal{D})$ in a)), while on the right we have an expression in terms of topological invariants of the problem [in a) it is the genus g of the Riemannian surface X and the degree of the divisor $\mathcal{D}$].

In the end, for complex manifolds, the general index formula and the Riemann-Roch-Hirzebruch Theorem (the special index formula for elliptic operators $\bar{D}_V := \left(\bar{\partial}_V + \bar{\partial}_V^*\right)^{even}$) are equivalent: Let

$$\bar{D} = \left(\bar{\partial} + \bar{\partial}^*\right)^{even} : \Omega^{0,even}\left(\mathbb{C}^n\right) \to \Omega^{0,odd}\left(\mathbb{C}^n\right)$$

be the "Riemann-Roch operator" for $X = \mathbb{C}^n$. Then $\sigma\left(\bar{D}\right)$ produces a generator of the homotopy group $\pi_{2n-1}(\mathrm{GL}(N, \mathbb{C}))$, $N = 2^{n-1}$, and all of $K(TX)$ is generated by the $\left[\sigma\left(\bar{D}_V\right)\right]$ modulo the image of $K(X)$; see [**ABP**, p.321 f]

A survey of function theoretic and geometric applications (e.g., estimates for the dimensions of systems of curves or differential forms and determination of Betti numbers of complicated manifolds) can be found for the case of curves, $\dim_{\mathbb{C}} X = 1$) in [**Mu**, p.147 ff], for the classification of surfaces ($\dim_{\mathbb{C}} X = 2$) in [**Kod**] including a multitude of very concrete geometric facts which are derived directly from the general Riemann-Roch-Hirzebruch Theorem (d) for line bundles), and for the general case in [**Schw66**].

8. The Index of Elliptic Boundary-Value Problems

We will now have a look at the index problem for boundary-value problems which was solved in the forties and fifties for a number of special cases in [**Ve62**, p.316-330], [**Boj**, p.15-18] and the sources stated there (see also Theorem 5.11, p. 146). In the programmatic article of [**Ge**], it was called "description of linear elliptic equations and their boundary problems in topological terms", and this initiated the search for the index formula for elliptic operators on closed manifolds [**AS63**]. In the center of the rapid development of "elliptic topology" (1963-1975), with alternate forms and proofs of the index formula and their widespread applications, remained manifolds without boundary, although [**AB64a**] demonstrated the topological significance of elliptic boundary conditions, and showed how, in principle, an index formula for elliptic boundary value problems may be obtained by reducing it to the unbordered case ("Poisson principle", see Sections 10.4 and 10.6). In connection with Section 14.3, one can in this fashion very quick proofs (and supplements) for the boundary-value problems with vanishing index listed in [**Ag**].

In working out the details of the Atiyah-Bott concept, one encounters two main difficulties:

1. Let $\sigma \in \mathrm{Iso}_{SX}(E,F)$ be the symbol of an elliptic (differential operator $P \in \mathrm{Ell}(E,F)$ over the bordered Riemannian manifold X with boundary Y, where E and F are vector bundles over X. If N is a sufficiently large natural number, it is relatively easy to extend (as exercised in Section 10.4) $\sigma \oplus \mathrm{Id}_{\mathbb{C}^N}$ on the "symbol level" to an elliptic symbol $\sigma' \in \mathrm{Iso}_{SX\cup BX|_Y}(E',F')$, using elements of the proof of the Bott Periodicity Theorem in such a way that $\sigma'(y,\xi) = \mathrm{Id}_{E'_y}$ for all $y \in Y$, and $\xi \in T^*_y X$ with $|\xi| \le 1$. There E_y was identified with F_y by means of $\sigma(y,\nu)$, where ν was an inner normal, and $E' := E \oplus \mathbb{C}^N_X$. The only hypothesis needed is that P admits elliptic boundary-value problems. The particular choice of boundary-value problem determines more exactly the way in which $\sigma \oplus \mathrm{Id}_{\mathbb{C}^N}$ is extended.

Doing the corresponding deformation on the "operator level", i.e., deforming $P \oplus \mathrm{Id}^N$ to P' "stably equivalent" to P and $P' = \mathrm{Id}'$ near Y, poses some essential difficulties. It is, for example, necessary to pass from differential operators to the class of pseudo-differential operators, and for the transition, the boundary behavior must be suitably restricted (see above before Theorem 10.26, p. 257); specifically, in [**Bou**, p.39 f] it is shown that if the "transmission condition" $\sigma(P)(y,-\nu) = (-1)^k\sigma(P)(y,\nu)$ is postulated, then the desired operator deformation is possible if and only if the indicator bundle $j(P) \in K(SY)$ (see before Exercise 10.28, p. 259) vanishes. On the other hand, the "free" extension of $\sigma \oplus \mathrm{Id}_{\mathbb{C}^N}$ characterized above exists (see result b) below), if only the difference bundle $j(P)$ lies in the image of the "lifting" $\pi^*_y : K(Y) \to K(SY)$.

It was mainly L. Boutet de Monvel, M. I. Vishik and G. I. Eskin who dealt in long series of papers with this predominantly analytic problem of finding the "correct" operator class for the formation of the "parametrix" and the necessary operator deformations. On the positive side, these difficulties demonstrate the effectiveness of topological methods in analysis: Once the index formula (see result d) below) is established, the computation of the index of a boundary-value problem with, say, a partial differential operator, does not require the deformation of the operators with all its difficulties, but only that of the symbol according to the simple rules given in Section 10.4.

2. Another basic difficulty is the following. Operators (such as the signature operator $D^+ := (d+d^*)^+)$ which play such an important role for closed manifolds in applications of the index formula as well as in the "cobordism" and "heat equation" proofs (see above Sections 14.4, 13.2 and 13.4) do not admit elliptic boundary value problems in the sense of Chapters 10-**??** on a bordered manifold X. See result b) below and [**APS75**, p.46]. Accordingly $\sigma((d+d^*)^+)$ cannot be deformed over the boundary Y of X to the identity (or any map which only depends on the base point $y \in Y$ but not on the covectors $\xi \in T^*_y X$). However, M. F. Atiyah, V. K. Patodi and I. M. Singer (loc. cit.) discovered that D^+ admits in an extended sense certain "globally elliptic" boundary conditions R^+ which canonically belong to D^+ and for which index $(D^+ \oplus R^+)$ is well-defined and equal to $\mathrm{sig}(X)$. Using the Gauss-Bonnet Formula $e(X) = (2\pi)^{-1}(\int_X K + \int_Y s)$ for the Euler number $e(X)$ as a model (where X is a bordered surface with Gaussian curvature K on X and with "geodesic curvature" s of Y in X), they (loc. cit., p. 54-57) could derive an index formula for a certain class of boundary value problems (P, R). The right hand side of this formula is the sum of two terms which are given explicitly in the "heat equation proof" and of which one only depends on the behavior of P in the interior (or the doubling of X), and the other only depends on the boundary behavior of (P, R). In particular, they obtained for $P := D^+$ the new formula

$$\operatorname{sig}(X) = \int_X L_q(p_1, \dots, p_q) - \eta(0)$$

(recall 13.7, p. 303) for the signature of a $4q$-dimensional compact, oriented, bordered Riemannian manifold. See also [**Gi75**] Efforts to obtain similar explicit "analytic" formulas for arbitrary elliptic boundary-value problems apparently did not achieve full success; see especially [**Cald67**], [**See69**] and [**Fe**]. Thus for the time being, we depend on our crude topological methods of symbol formation [see below results a) and d) below], in spite of the associated computationally unfortunate destruction of the special problem structure which may be simpler initially, and despite the delicate technical questions of lifting homotopies to the "operator levels" in the proofs. (The the topology chosen for the symbol deformation is not clear. While we deformed in the C^0-topology (see p. 235), there is also the C^1-topology which (according to a recent theorem of F. Waldhausen) has possibly less room for movement but a richer homotopy type.)

Results: Let X be an n-dimensional, compact, oriented Riemannian manifold with boundary Y.

a) There is a unique construction, whereby one can assign to elliptic element A of the Green algebra over X (see Chapter **??** a difference bundle $[A] \in K(T(X-Y))$ such that the following conditions hold:

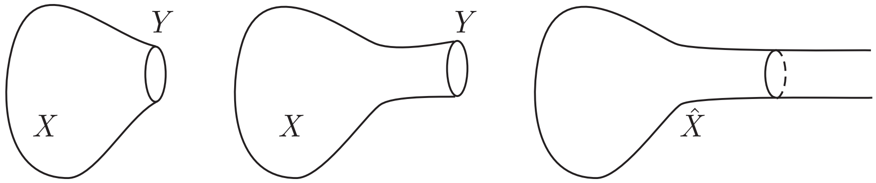
(i) $[A] = [B]$, if A and B are stably equivalent,

(ii) $[A \oplus B] = [A] + [B]$ and $[A \circ B] = [A] + [B]$, when the composition is defined,

(iii) $[A] = [\sigma(P)] + t\,[\sigma(Q)]$, for $A = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$ and $P = \mathrm{Id}$ near Y so that $[\sigma(P)] \in K(T(X-Y))$ is well defined. Here t is the composition already investigated above in Chapter 13, namely

$$K(TY) \overset{\boxtimes b}{\cong} K(TY \times \mathbb{R}^2) \cong K(TN) \overset{ext}{\to} K(T\hat{X}) \cong K(T(X-Y))$$

for the "trivial" embedding of Y into the open manifold $\hat{X} = X \cup (Y \times [0, \infty))$ with trivial normal bundle N; $\hat{X}$ and $X - Y$ are diffeomorphic,



whence we have the isomorphism on the right.

(iv) $[A] = [\bar{\sigma}(P, \beta^+)]$, if $A = (P, R)$ is an elliptic boundary-value system of differential operators and $\bar{\sigma}(P, \beta^+)$ is the extension of $\sigma(P) \oplus \mathrm{Id}_{\mathbb{C}^N} \in \mathrm{Iso}_{SX}$ (defined by the boundary isomorphism β^+) to an object in $\mathrm{Iso}_{SX\cup BX|_Y}$, see Chapter 10 and, in particular, Section 10.4.

b) For a differential operator P, at each point $y \in Y$, there is a vector bundle $\mathcal{M}_y^+$ over the base $(SY)_y = S^{n-2}$, and a difference bundle $[\sigma_y(P)] \in K(\mathbb{R}^n)$ arising from $\sigma_y(P) := \sigma(P)(y, \cdot)$. Now, $\mathcal{M}_y^+ - \dim \mathcal{M}_y^+$ defines an element of $K(\mathbb{R}^{n-2})$, and we have (because of an unfortunate sign choice, we use P^* instead of P) $[\sigma_y(P^*)] = -b \boxtimes (\mathcal{M}_y^+ - \dim \mathcal{M}_y^+)$. The tensoring with the Bott element on the right side extends to a homomorphism $\gamma : K(SY) \to K(BX|_Y, SX|_Y)$, and we have the global formula $r^*[\sigma(P^*)] = -\gamma(\mathcal{M}^+)$, which can be written in the form $r^*[\sigma(P)] = \gamma(j(P))$ in the case that P is a pseudo-differential operator. Here $r^* : K(BX, SX) \to K(BX|_Y, SX|_Y)$ is the restriction homomorphism, and $\mathcal{M}^+$ resp. $j(P)$ are the "indicator bundles" defined via ordinary differential equations (Chapter 10) or Wiener-Hopf operators (Chapter **??**); these indicator bundles coincide (up to a sign convention) for differential operators.

c) An elliptic pseudo-differential operator P over X can then be completed to produce an elliptic element $\begin{bmatrix} P & * \\ * & * \end{bmatrix}$ of the Green algebra over X, if and only if the following two equivalent conditions are met (for suitable N):

(i) $j(P) \in \operatorname{Ker} \gamma$

(ii) $\sigma(P) \oplus \mathrm{Id}_{\mathbb{C}^N}$ extends to an isomorphism over $SX \cup BX|_Y$.

If P is a differential operator, then $\deg(P) = \nu(P)$, where $\deg(P)$ is defined as the "degree" of $\sigma(P)(y, \cdot) : (SY)_y \to \mathrm{GL}(N, \mathbb{C}), y \in Y$ (i.e., the local degree of P), and $\nu(P)$ is the difference $\dim \mathcal{M}_\eta^+ - \dim \mathcal{M}_\eta^-$, $\eta \in (SY)_y$.

d) Each embedding of the bordered manifold X in a Euclidean space $\mathbb{R}^m$ defines (as in Chapter 13 and also in a(iii) above) a homomorphism $t' : K(T(X - Y)) \to K(\mathbb{R}^{2m})$, and we then have the formula $\operatorname{index} A = \alpha^m t'[A]$ for each elliptic element A of the Green algebra. Here $[A]$ is the difference bundle of A (well defined by a), and $\alpha^m : K(\mathbb{R}^{2m}) \to K(\{0\})$ is the iteration of the Bott isomorphism.

The cohomological version (see Section 14.1 above) of the index formula is

$$\operatorname{index} A = (-1)^{n(n+1)/2} \left\{ \Phi^{-1} ch\,[A] \smile \tau\,(TX \otimes \mathbb{C}) \right\} [X].$$

In this formula, $[X] \in H_n(X, Y; \mathbb{Q})$ denotes the fundamental cycle of the orientation of X, $\tau\,(TX \otimes \mathbb{C}) \in H^*(X; \mathbb{Q})$ is the Todd class of the complexification of TX,

$$ch : K(T(X - Y)) \to H^*(BX, SX \cup BX|_Y; \mathbb{Q}) = H_c^*(T(X - Y)); \mathbb{Q})$$

is the "Chern character ring homomorphism", $\Phi : H^*(X, Y; \mathbb{Q}) \to H^*(BX, SX \cup BX|_Y; \mathbb{Q})$ is the "relative Thom isomorphism", and $\smile : H^*(X, Y; \mathbb{Q}) \times H^*(X; \mathbb{Q}) \to H^*(X, Y; \mathbb{Q})$ is the "relative-cup product".

In the case of systems (see Section 14.2 above), where A is given by a system of equations, $[A]$ is defined by a continuous matrix-valued map $\sigma(A) : SX \cup BX|_Y \to \mathrm{GL}$, and we have the integral formula

$$\operatorname{index} A = \int_{\partial(BX)} \sigma(A)^* \omega \wedge \pi^* \tau,$$

where ω and τ are the explicit differential operators (given in Section 14.2 above) on U resp. GL and on X, and $\pi : \partial\,(BX) \to X$ is the projection.

In the classical case, where X is a bounded domain in $\mathbb{R}^n$, one finally obtains the simple formula $\operatorname{index} A = \sigma^n[A]$, which coincides, in the special case $A = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$ with $P = \mathrm{Id}$ near Y, with Theorem 12.3 (p. 291) and Theorem 13.1 (p.295).

Arguments: a) was announced in [**Pal65**, p.349 f], and proved in [**Bou**, p.47-50]. There, (iv) is replaced by the weaker requirement that $[A] = 0$, if A belongs to a certain list of canonical operators of "Dirichlet type".

For b) and c), note that r^* and γ fit into the short exact sequences which are horizontal and vertical in the following diagram

$$\begin{array}{ccccccc} & & & & K(BX, SX \cup BX|_Y) & \cong & K(T(X-Y)) \\ & & & & \downarrow q^* & & \\ & & & & K(BX, SX) & \cong & K(TX) \\ & & & & \downarrow r^* & & \\ K(Y) & \overset{\pi_Y^*}{\to} & K(SY) & \overset{\gamma}{\to} & K(BX|_Y, SX|_Y) & \cong & K(\mathbb{R} \times TY) \end{array}$$

Here, $\pi_Y : SY \to Y$ is the projection, and r and q are both embeddings; γ is the composition $K(SY) \overset{\cong}{\to} K(\mathbb{R} \times TY) \cong K(BX|_Y, SX|_Y)$. See [**Bou**, p.39 f], and also, for the local version, [**Pal65**, p.351].

c) is a trivial consequence of Exercise 10.30, p. 259: By definition, $j(A) = j(P) + [\pi^* G] - [\pi^* H]$ vanishes for each elliptic element $A : C^\infty(E) \oplus C^\infty(G) \to C^\infty(F) \oplus C^\infty(H)$ of the Green algebra (G and H vector bundles over Y). Hence, $j(P) \in K(SY)$ lies in the image of π_Y^*, and hence in the kernel of γ. Because of b) and the exactness of the horizontal sequence, this means that it is possible to "lift" $[\sigma(P)] \in K(BX, SX)$ to $K(BX, SX \cup BX|_Y)$, obtaining the difference bundle $[A] \in K(T(X - Y))$. Incidentally, a) goes much further than c), since it is shown in a) how a special completion of the elliptic operator P to an element A of the Green algebra, generates the "lifting". In [**Bou**, p.35 f] it is shown in addition that each bundle $\pi^*(G)$ with arbitrary $G \in K(Y)$ is the "indicator bundle" $j(P)$ of an elliptic pseudo-differential operator P on X.

With d), concrete computations are made easy for some classical boundary-value problems. Indeed, one recovers without difficulty the Theorem of Vekua (Theorem 5.11, p. 146) using Exercise 10.21, p. 253. In view of the Atiyah-Patodi-Singer Signature Formula (see Difficulty 2 above), it seems probable (in order to remain in the language of integral formulas) that the index integral over the "integration domain" $\partial(BX) = SX \cup BX|_Y$ should be replaced by the sum of two integrals with different integration domains. In fact, the proof of d) (in which one lifts the K-theoretic resp. homotopy-theoretic construction in a) to the level of operators) already runs in this direction. Roughly speaking, the proof of d) goes as follows: We start with $A \in \mathrm{Ell}_k(X;Y)$. Via composition with certain standard operators (construction modeled on operators like the Riesz Λ-operators; see Exercise 7.5, p. 194) of vanishing index, we obtain an operator $A' \in \mathrm{Ell}_0(X;Y)$ with $\operatorname{index} A' = \operatorname{index} A$. To A', we can now add another elementary boundary-value system B with $\operatorname{index} B = 0$, such that the indicator bundle $j(A' \oplus B)$ vanishes. With a topological argument, it follows that $A' \oplus B$ is stably equivalent to a $A'' = \begin{bmatrix} P'' & L'' \\ R'' & Q'' \end{bmatrix} \in \mathrm{Ell}_0(X;Y)$, where P'' equals the identity near Y. With Theorem 10.32e (p. 260), this allows a deformation of A'' to $A''' := \begin{bmatrix} P'' & 0 \\ 0 & Q'' \end{bmatrix}$, whence index A = index A''' = index P'' + index Q''. Thus, in principle, the index problem for boundary-value systems is reduced to the index problem over closed manifolds, since index P'' = index $\bar{P}''$, where $\bar{P}''$ is the operator extended to $X \cup X$ by means of the identity. Details are in [**Bou**, p.47 f], where a) and d) are proved simultaneously. For a heuristic treatment of the situation, when all the operators arising are differential operators at the outset, see also [**CaSc**, 25-08 f], [**Schu76a**] and [**Schu76b**].

9. Real Operators

Up to now, we have considered operators between spaces of sections of *complex* vector bundles. One can also consider operators with *real* coefficients which operate only on sections of real bundles. If such a real elliptic and skew-adjoint operator, then trivially index $P = 0$. This is uninteresting, but now $\dim(\operatorname{Ker} P)$ is a homotopy-invariant mod 2. The reason for this stems from the fact that the nonzero eigenvalues of P all come in complex conjugate pairs $(\lambda, \bar{\lambda})$; if one deforms the the operator P so that λ goes to zero, then $\bar{\lambda}$ goes to zero and consequently $\dim(\operatorname{Ker} P)$ increases by two. Already in 1959, R. Bott had discovered a real analogue to his "periodicity theorem" (see Chapter 11 above) and proved that the homotopy groups $\pi_i(\mathrm{GL}(N, \mathbb{R}))$, for large N, periodic in i with period 8, and for $i \equiv 0$ or $i \equiv 1 \bmod 8$ are isomorphic to $\mathbb{Z}$.

In [**AS69**] the connection between these two analytic and topological mod 2 invariants was determined, and the index theorem was carried over to the real case, see also the elaboration in [**Fu99**]. (Actually, linear differential operators on real vector bundles are decisive for M. Furuta's geometric proof of the Index Theorem, avoiding the use of pseudo-differential operators.) Together, these considerations provide a new and topologically much simpler proof of the real Bott Periodicity Theorem. Two details are particularly noteworthy: First, in order to connect the two invariants of the real theory, one must go outside of the real theory, since the amplitude p of a real skew self-adjoint differential operator P is defined via the Fourier transform, and is thus not real in general, but rather complex with the condition $p\,(x, -\xi) = \overline{p\,(x, \xi)}$. Thus, the symbol of a real operator does not immediately yield a suitable element of $\pi_i(\mathrm{GL}(N, \mathbb{R}))$, but rather, must be interpreted as a mapping $f : S^{2n-1} \to \mathrm{GL}(N, \mathbb{C})$ (as in Chapter 12) with the condition $f(-\xi) = \overline{f(\xi)}$. Another peculiarity lies in the fact that these mod 2 invariants (although having only the values 0 or 1) in a certain sense are more complicated topologically, or in any case, of a different type than the usual homology or cohomology classes (also if one takes $\mathbb{Z}_2$ coefficients). Thus, in concrete situations they can provide decisive additional information. This program was carried out for vector fields (see Section 14.5 above) in [**Ati70c**] and [**AD**]. For a different approach see also the cited [**Fu99**].

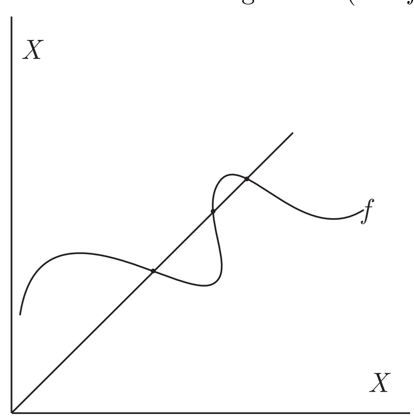
10. The Lefschetz Fixed-Point Formula

Let $f : X \to X$ be a continuous map with X compact, and let $\operatorname{Fix}(f) = \{x \in X : f(x) = x\}$ be the fixed-point set of f. Salomon Lefschetz (1926) introduced the formula $L(f) = \sum \nu(x)$, where the sum is over all fixed-points of f. For details of the definition of the integer $\nu(x)$ (which is 1 for an isolated fixed point and equals 0 for a point where $f = \mathrm{Id}$ in a neighborhood) For the proof, see [**AH**, p.531-542] or [**Gr**, p.222-224]. The Lefschetz number" $L(f)$ is defined as the alternating sum $\sum(-1)^i \operatorname{trace}(H^i f)$, where $H^i f : H^i(X, \mathbb{C}) \to H^i(X, \mathbb{C})$ is the cohomology endomorphism of the complex vector space $H^i(X, \mathbb{C})$ induced by f. M. F. Atiyah and R. Bott refined this beautiful formula in the mid 60's. Furthermore, the formula is simplicially defined and in general hardly computable or in other words, only the topology of X enters into the formula, while additional structures are ignored. Atiyah and Bott removed this weakness by bringing the additional structures into play.

Results. Let X be a compact C^∞ manifold without boundary, and let P be an elliptic differential operator. Let $f : X \to X$ be differentiable and commute with P; where we need to assume that f "lifts" to a bundle mapping, say $\tilde{f}$, so that f acts on sections via $(f \cdot s)(x) = \tilde{f}(s(f^{-1}(x)))$. Then f yields a well defined endomorphism of the finite-dimensional vector spaces $\operatorname{Ker} P$ and $\operatorname{Coker} P$. We define the *Atiyah-Bott-Lefschetz number* as

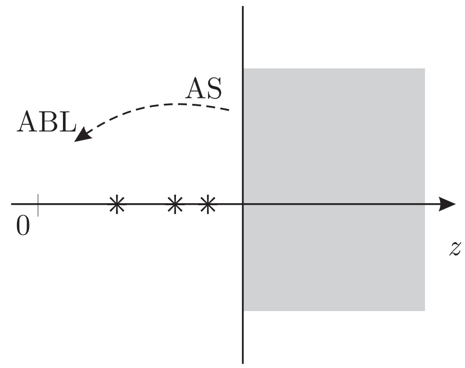
$$L(f, P) := \operatorname{trace}(f|_{\operatorname{Ker} P}) - \operatorname{trace}(f|_{\operatorname{Coker} P}).$$

If f is the identity, then by definition $L(f, P) = \operatorname{index} P$, and one can apply the Atiyah-Singer Index Formula. In the other "extreme case", where f has only isolated fixed points with multiplicity $+1$, one obtains the formula $L(f, P) = \sum v(x)$ where x runs over the fixed-points of f and the complex number $v(x)$ depends only on the differential $f_*(x) : T_x X \to T_x X$. (The "simplicity" or "transversality" of the fixed-point x means that the endomorphism $\mathrm{Id} - f_*(x)$ is invertible, and the "multiplicity" ± 1 is understood to be the sign of $\det(\mathrm{Id} - f_*(x))$.)



In the following list, there are some applications of the Atiyah-Bott-Lefschetz formula (ABL) with the expressions fog the respective values of $v(x)$:

For further details and the proof of (ABL), we refer to [**AB65**], [**AB66**], and [**AB67**]. Incidentally, the proof is essentially simpler than the proof of the index formula (AS), and represents a "weak" version of the "heat equation proof" discussed above in 13.4: One considers the zeta function $\zeta(z) : \operatorname{trace}(f^* \circ \Delta^{-z})$ whose value at $z = 0$ is easier to calculate, since it turns out that this zeta function is holomorphic not only for $\operatorname{Re}(z) > \dim X$, but also in all of $\mathbb{C}$ because of the assumptions on the fixed points; see also the following section.



11. Analysis on Symmetric Spaces

Let X be a closed C^∞ manifold and G a compact group of diffeomorphisms of X. One may think of $G = \{\mathrm{Id}\}$ or $G = \left\{\mathrm{Id}, g, g^2, ..., g^{r-1}\right\}$, where g is a diffeomorphism of order r (i.e., $g^r = \mathrm{Id}$). G can also be a more general finite group, or a compact Lie group. Let P be a "G-invariant" elliptic operator over X, which commutes with the operations of the group elements (via suitable bundle automorphisms), so that $\operatorname{Ker} P$ and $\operatorname{Coker} P$ are not only finite-dimensional complex vector spaces, but can be regarded as G-modules. The homomorphism $g \mapsto g|_{\operatorname{Ker} P}$, $g \in G$, is then a finite-dimensional representation of G, and $g \mapsto \operatorname{trace}(g|_{\operatorname{Ker} P})$ is its character. Corresponding considerations apply to $\operatorname{Coker} P$. As in the preceding section, we have

$$\operatorname{index}_G(P) : g \mapsto L(g, P) := \operatorname{trace}(g|_{\operatorname{Ker} P}) - \operatorname{trace}(g|_{\operatorname{Coker} P}),\ \ g \in G,$$

which is a "virtual" character (i.e., an element of the "representation ring" $R(G)$). If $G = \{\mathrm{Id}\}$, then $R(G) = K(\mathit{point}) = \mathbb{Z}$ and $\operatorname{index}_G(P) = \operatorname{index} P$. In the general case, $R(G) \cong K(*) \cong *$ is a complicated object and correspondingly, $\operatorname{index}(P)$ is a sharper (also homotopy-) invariant than the integer $\operatorname{index}(P)$.

Results. In the framework of "equivariant" K-theory with groups $K_G(X)$ (esp. $K_G(TX)$) for the "symmetric space" X, one obtains an analogous index formula for $\operatorname{index}_G(P) \in R(G)$, and for each $g \in G$ a Lefschetz formula

$$L(g, P) = \sum\nolimits_{i=1}^{l} \int_{X_i} \alpha_i,$$

where the sum on the right is over the connected components $X_1, ..., X_l$ of the fixed point set $\operatorname{Fix}(g)$. Incidentally, $\operatorname{Fix}(g)$ is a mixed-dimensional submanifold of X, since one can introduce local coordinates in a suitable neighborhood of any fixed-point such that g operates linearly. Each individual term in the formula has the form of our index formula (see Section 14.1 or Chapter 13 above). For the precise definitions and the various formulations of these formulas, we refer to **??** and **??**, as well as **??**, **??** and **??**. An abundance of applications mainly for the case of a complex manifold X with $P := \left(\partial + \bar{\partial}\right)^{\text{even}}$ and holomorphic g (and more generally, applications in the domain of elementary number theory) can be found in [**HiZa**]. Here P is held fixed and attention is paid to the correct choice of G or g. Conversely, in [**Bott**] or [**Wal**, p.114-138] one finds an investigation of "homogeneous differential operators" on a fixed homogeneous space $X = G/H$, where H is a closed subgroup of G. Here the relevance lies in the fact that these formulas connect the fixed points of the action by G with global invariants of the symmetric space X. Thus, deep formulas are obtained for the way in which (in the language of cohomology) individual characteristic classes of invariants can be synthesized to yield invariants of the entire cohomology ring of the manifold.

The proof follows essentially as in the argument carried out here in Chapters 12 and 13. For the construction of the "topological" Index $K_G(TX) \to R(G)$, one needs an equivariant sharpening of the Bott Periodicity Theorem which can be obtained by the same function analytic tools used in 11 for the "classical" Bott Periodicity Theorem; incidentally, it is noteworthy that these tools are *necessary*, since (roughly) only this approach can be freed of the inductive argument which cannot succeed in general for the equivariant case. See [**Ati68a**], and for a non-technical introduction [**Ati67b**, p.246 f].

12. Further Applications

With the preceding overview, we have in no way encompassed all of the connections, alternative formulations, generalizations and special applications of the Atiyah-Singer Index Formula, but we have only touched those which have reached a certain definitive form in their development. For an insight into further perspectives and open problems, we refer to [**Si69**], **??**, and [**Ati76b**] and the literature given there.

Part IV

Index Theory in Physics and the Local Index Theorem

CHAPTER 15

Physical Motivation and Overview

Synopsis. Mode of Reasoning in Physics. String Theory and Quantum Gravity. The Experimental Side. Classical Field Theory: Newton-Maxwell-Lorentz, Faraday 2-Form, Abstract Flat Minkowski Space-Time, Relativistic Mass, Relativistic Kinetic Energy, Inertial System, Lorentz Transformations and Poincaré Group, Relativistic Deviation from Flatness, *Twin Paradox*, Variational Principles. Kaluza-Klein Theory: Simultaneous Geometrization of Electro-Magnetism and Gravity, Other *Grand Unified Theories*, String Theory. Quantum Theory: Photo-Electric Effect, Atomic Spectra, Quantizing Energy, State Spaces of Systems of Particles, Basic Interpretive Assumptions. Heisenberg Uncertainty Principle. Evolution with Time - The Schrödinger Picture. Nonrelativistic Schrödinger Equation and Atomic Phenomena. *Minimal Replacement* and Covariant Differentiation. Anti-Particles and Negative-Energy States. Unreasaonable Success of Standard Model. Dirac Operator vs. Klein-Gordon Equation. Feynman Diagrams

Both the goals and methods of physics are different from those of mathematics. Mathematicians have the rather nebulous goal of exploring and establishing that which is logically possible and interesting, depending on the fashions of the time and the tastes of the individual. Physicists have the sharper goals of discovering, explaining and predicting actual phenomena in the physical world. The mode of reasoning in physics is rather fuzzy by mathematical standards, as it is often partly based on conventional wisdom and folklore rather than clear axioms. However, this reasoning is of great value if it provides a satisfying explanation of experimental data and makes promising, testable predictions. If physicists were forced to be mathematically rigorous every step of the way, physics would not have advanced toward its goals nearly as much as it has. Although the creative process in mathematics is generally fuzzy in its initial phases, a result is not usually publishable until it has been proven within a quite definite framework of commonly accepted logical standards, which goes far beyond the notion of "reasonable doubt" in a court of law. Mathematicians are uneasy with many of the heuristic arguments used by physicists sometimes involving manipulations of expressions which have not been shown to exist (e.g., path integrals, infinite renormalizations, nonconvergent series, etc.). On the other hand, physicists cannot be expected to have interest in mathematics that seems unrelated to physical phenomena.

Since index theory was developed as a mathematical achievement, there has emerged a prominent group of theoretical physicists who appear to be somewhat unconventional, namely the string theorists (and we may include supersymmetrists and quantum gravitists as well). It is clear that most string theorists believe that what they are doing is of physical relevance, but as yet no direct experimental confirmation has emerged. What they have certainly uncovered is a truly awesome body of mathematics that has had a big positive impact on purely mathematical research in neighboring areas. Many mathematicians envy the mathematical *insights that string theorists have had. Indeed, Seiberg-Witten Theory is but a small portion of the mathematics inspired largely by the insights of Edward Witten, one of the leading string theorists. However, one characteristic of a physical theory that conventional physicists deem essential is that the theory be testable by experiment. Currently, the physical refutation of string theory seems just as remote as its confirmation. Based on this, the conventional physicist can argue that string theory (or quantum gravity) is not really a physical theory at all. This is not because it is false, but because it is not falsifiable, in the sense that it seems unlikely that it can be proven or disproved experimentally in the foreseeable future. José Gracia-Bondía* [**Gra**, p.6], *e.g., emphasises that masses and energies on our planet are much too small to make a difference for possible falsifications of common ideas of string theory and quantum gravity. However,when Giampiero Esposito in* [**Esp11**, Section 8.1] *addresses the experimental side of quantum gravity, one of his points is the immense capacity of modern computer supported and partly space based astronomy, which gives access to data involving previously unimaginable large masses and energies. See also the recent Nobel citations in physics for a non-technical view on the new observational capacities. Moreover, as noted by Bryce DeWitt in* [**DeWi**, p.417], *string theory provides (in some cases) substantially simplified schemes and diagrams for basic calculations. His example is the replacement of four different Feynman diagrams by a single one in string theory: a thing "that, from a nonspecialists point of view, make it look rather pretty".*

If it were suddenly found that string theory has no physical relevance, most likely only a handful of string theorists would remain, namely those who really consider themselves to be primarily mathematicians.

There are such mathematicians (misnamed mathematical physicists) who are interested in strict mathematics that seems to have physical relevance or is motivated by physical considerations. In the overview that follows, it is hoped that the reader may gain some understanding of why many concepts in this book (e.g., elliptic operators, complex vector bundles, pseudo-differential operators, Hilbert spaces, distributions, etc.) may have great physical relevance and how in large part they were initially motivated by physical considerations. Of course, mathematics being motivated by physics is not a new phenomena, but rather an old one. There was hardly any distinction between mathematics and theoretical physics before the 1900s. The dubious mid-twentieth century goal of attaining ivory purity in mathematics, devoid of any hint of lowly physical application, seems to have been largely temporary, although many practitioners remain.

The reader is not expected to understand every detail in the following lengthy (yet necessarily incomplete and historically vague) overview of quantum field theories in modern physics. However, she or he may take whatever is digestible, realizing that this material is neither a prerequisite nor a substitute for the more precise (if drier) mathematics of the chapters that follow. Those who are unfamiliar with relativity or quantum physics are likely to discover that the logical possibilities of the world of physics can be every bit as beautiful and strange as those encountered in far-reaching mathematical diversions.

1. Classical Field Theory

In classical (as opposed to quantum) physics, particles are viewed as point-like objects that move along paths which are solution curves of systems of ordinary differential equations determined by a force field. For example, there is **Newton's equation** $m\mathbf{r}''(t) = \mathbf{F}(\mathbf{r}(t))$, where $\mathbf{F}:\mathbb{R}^3 \to \mathbb{R}^3$ is a given force field. More generally, the force may also depend on the velocity $\mathbf{r}'(t)$ as well as $\mathbf{r}(t)$. Indeed, an electromagnetic (E-M) field consists of a pair of vector fields $\mathbf{E}$ and $\mathbf{B}$ (which may be time-dependent). A *test* particle of charge e moves according to the **Lorentz force law**

$$\frac{d}{dt}(m\mathbf{r}'(t)) = e\mathbf{E}(\mathbf{r}(t), t) + \frac{e}{c}\mathbf{r}'(t) \times \mathbf{B}(\mathbf{r}(t), t). \tag{15.1}$$

The situation is complicated not only by the fact that a real (not test) particle contributes to $\mathbf{E}$ and $\mathbf{B}$, but also by the fact that $\mathbf{E}$ and $\mathbf{B}$ satisfy a system of partial differential equations, namely **Maxwell's equations**

$$\begin{array}{ll} (1)\ \nabla \times \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t} = \mathbf{0} & (2)\ \nabla \cdot \mathbf{B} = \mathbf{0} \\ (3)\ \nabla \cdot \mathbf{E} = \rho & (4)\ \nabla \times \mathbf{B} - \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t} = \dfrac{1}{c}\mathbf{J}. \end{array} \tag{15.2}$$

Here c denotes the *speed of light*, $\rho : \mathbb{R}^3 \to \mathbb{R}$ is proportional to the *charge density* of a continuous medium of charged particles, and $\mathbf{J}$ is essentially the *current density* of the medium (i.e., $\mathbf{J} = \rho\mathbf{v}$, where $\mathbf{v}$ is the *velocity vector field* of the medium). Thus, $\mathbf{E}$ and $\mathbf{B}$ are influenced by each other, as well as the by the motions of the charged medium that they are supposed to influence via the Lorentz force law (15.1).

Maxwell's equations can be simplified conceptually by considering the following 2-form, called the **E-M field strength** or **Faraday 2-form**

$$F := cE_1 dx \wedge dt + cE_2 dy \wedge dt + cE_3 dz \wedge dt + B_1 dy \wedge dz + B_2 dz \wedge dx + B_3 dx \wedge dy. \tag{15.3}$$

EXERCISE 15.1. (a) Check that Maxwell's equations (1) and (2) are equivalent to $dF = 0$.
(b) Defining the **source 1-form** j by

$$j = \rho dt - c^{-2}(J_1 dx + J_2 dy + J_3 dz), \tag{15.4}$$

verify that the Maxwell equations (3) and (4) say that $\delta F = j$, where δ is the codifferential (the formal adjoint of d) relative to the Lorentz-Minkowski metric $c^2dt^2 - dx^2 - dy^2 - dz^2$.

This exercise implies that Maxwell's equations can be immediately generalized to arbitrary 4-manifolds with Lorentz metric tensors (i.e., *space-times*). Hence, Maxwell's equations fit quite naturally into general relativity, although historically, relativity was built around Maxwell's equations. The Lorentz force law (15.1) also can be written invariantly on a space-time M. Indeed, the "world line" of a test particle of rest mass m_0 and charge e is a curve $s \mapsto \gamma(s) \in M$ which obeys the equation

$$m_0 \frac{D}{ds}\gamma'(s) = \frac{e}{c}F(\gamma'(s), \cdot)^{\#} \tag{15.5}$$

where γ' denotes the tangent vector field of γ, $\frac{D}{ds}$ denotes covariant differentiation along γ (i.e., $\nabla_{\gamma'}$) and the sharp "#" on the right side indicates that the covector $F(\gamma'(s), \cdot)$ has been converted to a vector by "raising indices" using the metric. In flat Minkowski space, $\frac{D}{ds}(\gamma'(s))$ is simply $\gamma''(s)$, but s is not necessarily the time coordinate. Rather think of s as the arc length as an inherent parametrisation. Equation (15.5) implies that the length of $\gamma'(s)$ is constant. Indeed,

$$\begin{aligned}\tfrac{1}{2}m_0\tfrac{d}{ds}\left|\gamma'(s)\right|^2 &= m_0\left\langle \frac{D}{ds}\gamma'(s), \gamma'(s)\right\rangle = \frac{e}{c}\left\langle F(\gamma'(s),\cdot)^{\#}, \gamma'(s)\right\rangle \\ &= \frac{e}{c}F(\gamma'(s),\gamma'(s)) = 0,\end{aligned} \tag{15.6}$$

since F is anti-symmetric. Note that (15.5) is not scale invariant; i.e., if γ is replaced by γ_a where $\gamma_a(s) := \gamma(as)$, then γ_a is not necessarily a solution of (15.5) for $a \neq 1$. In Minkowski space with the metric $c^2dt^2 - dx^2 - dy^2 - dz^2$, equation (15.5) splits into spatial and temporal components which are empirically correct only when $\left|\gamma'(s)\right|^2 = c^2$. Indeed,

$$\begin{aligned}\gamma(s) &= (t(s), x(s), y(s), z(s)) =: (t(s), \mathbf{r}(s)) \\ &\Rightarrow \left|\gamma'(s)\right|^2 = c^2t'(s)^2 - \left|\mathbf{r}'(s)\right|^2 = c^2t'(s)^2\left(1 - c^{-2}\left|\frac{\mathbf{r}'(s)}{t'(s)}\right|^2\right) \\ &= c^2t'(s)^2\left(1 - \frac{\left|\mathbf{v}(t(s))\right|^2}{c^2}\right),\end{aligned} \tag{15.7}$$

where $\mathbf{v}(t) = \frac{d}{dt}\mathbf{r}(s(t))$ assuming that $t'(s) > 0$, so that $t = t(s)$ can be inverted. Then

$$\left|\gamma'(s)\right|^2 = c^2 \quad \Leftrightarrow \quad t'(s) = \left(1 - \frac{\left|\mathbf{v}(t(s))\right|^2}{c^2}\right)^{-\frac{1}{2}} =: \beta(s), \tag{15.8}$$

and $\gamma'(s) = (\beta(s), \beta(s)\mathbf{v}(t(s)))$.

Exercise 15.2. Check that (15.5) splits into the pair of equations

$$\text{(a) } \tfrac{d}{dt}(m_0\beta\mathbf{v}) = e\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right) \qquad \text{(b) } \tfrac{d}{dt}\left(m_0\beta c^2\right) = e\mathbf{E}\cdot\mathbf{v} \tag{15.9}$$

Note that (a) is the Lorentz force law (15.1), where $m = m_0\beta$ is the so-called **relativistic mass** ($m \approx m_0$ for $|\mathbf{v}| \ll c$, and $m \to +\infty$ as $|\mathbf{v}| \uparrow c$). The right side of (b) is the rate at which the E-M field does work on the particle; note that $\mathbf{B}$ does no work since $(\mathbf{v}\times\mathbf{B})\cdot\mathbf{v} = 0$. Thus, $mc^2 = m_0\beta c^2$ on the left side of (b) must be the energy E of the particle (i.e., $E = mc^2$). Note that m_0c^2 is the **rest energy** and the **relativistic kinetic energy** is

$$\begin{aligned}mc^2 - m_0c^2 &= m_0c^2(\beta - 1) \\ &= \tfrac{1}{2}m_0\left|\mathbf{v}\right|^2 + c^2\mathcal{O}\left(\left(|\mathbf{v}|/c\right)^4\right) \quad \text{as} \quad |\mathbf{v}|/c \to 0.\end{aligned} \tag{15.10}$$

Abstract Minkowski space-time consists of a four-dimensional vector space (or more precisely, affine space) M with scalar product $\langle\cdot,\cdot\rangle$ of signature $(+,-,-,-)$. By translation and the usual indentifications, $\langle\cdot,\cdot\rangle$ determines a scalar product on the tangent space at each point of M. A coordinate system $(t,\mathbf{r}) := (t,x,y,z)$ on M is called an **inertial system** if $\langle\cdot,\cdot\rangle = c^2dt^2 - d\mathbf{r}^2 := c^2dt^2 - dx^2 - dy^2 - dz^2$. If $(\bar{t},\bar{\mathbf{r}}) := (\bar{t},\bar{x},\bar{y},\bar{z})$ is another inertial system, then there is a linear transformation $L : \mathbb{R}^4 \to \mathbb{R}^4$ and a point $(t_0,\mathbf{r}_0) \in \mathbb{R}^4$, such that

$$(\bar{t},\bar{\mathbf{r}}) = L((t,\mathbf{r})) + (t_0,\mathbf{r}_0). \tag{15.11}$$

The fact that $c^2 d\bar{t}^2 - d\bar{\mathbf{r}}^2 = \langle \cdot, \cdot \rangle = c^2 dt^2 - d\mathbf{r}^2$ places a restriction on L. It preserves the scalar product with diagonal matrix $I_{1,3}$ having diagonal entries $1, -1, -1, -1$, in the sense that $[L]^T I_{1,3} [L] = I_{1,3}$, where $[L]$ is the matrix of L relative to the standard basis of $\mathbb{R}^4$. Such $L \in \mathrm{GL}\left(\mathbb{R}^4\right)$ are known as **Lorentz transformations** and comprise the Lorentz group $\mathrm{O}(1,3)$. The Lorentz transformations together with the translations of $\mathbb{R}^4$, generate the **Poincaré group**. Even if inertial systems $(\bar{t}, \bar{\mathbf{r}})$ and $(t, \mathbf{r})$ are based at the same point $O \in M$, we do *not* necessarily have the equality $\bar{t} = t$ of time coordinate functions (i.e., coordinate time has no absolute meaning). On the other hand, if $\gamma : (a,b) \to M$ is a smooth curve, the condition $|\gamma'(s)|^2 := \langle \gamma'(s), \gamma'(s) \rangle = c^2$ does have invariant meaning. The correct interpretation is that s represents the time on a clock carried by the particle with world line γ. Equation (15.8) tells us that $t'(s) = \beta \geq 1$ meaning that coordinate time in the inertial system $(t, \mathbf{r})$ generally runs faster than the proper time of a particle which is moving relative to this inertial system. Assuming that the earth does not deviate from the t-axis of some inertial system in a nearly Minkowskian space-time, a high velocity space traveler with world line γ will find his earth-bound twin is older when he returns. The earth-bound twin ages according to his proper time which coincides with coordinate time t, while the space traveler ages according to his proper time, namely s, where $|\gamma'(s)|^2 = c^2$. Thus, assuming that their clocks are synchronized just before departure at $t = s = 0$, we then have

$$t'(s) = \left(1 - |\mathbf{v}|^2 / c^2\right)^{-\frac{1}{2}} > 1 \Rightarrow t > s. \tag{15.12}$$

One might argue that by symmetry, we should also have that $s > t$, but the situation is not symmetric, since the world line of the space traveler is not close to the time axis of an inertial system, because his acceleration is considerable (i.e., his world line is not nearly straight). Thus, while we have the so-called twin paradox, there is no contradiction. The result has been confirmed by experiment using particles instead of humans.

Things are complicated by the fact the metric tensor $g_{\mu\nu}$ on a realistic space-time is not flat like that of Minkowski space. The deviation from flatness is due to the presence of E-M fields (i.e., radiation), particles (neutral and charged) and gravity waves that can propagate in space-time even in the absence of radiation and matter. One measure of curvature is the symmetric Ricci curvature tensor $R_{\mu\nu}$ (where $\mu, \nu = 0, 1, 2, 3$) to be defined later (see 16.61). The vanishing of the Ricci tensor is necessary (but not sufficient) in order that a space-time be locally isometric to Minkowski space. The *scalar curvature* is the trace $S = g^{\mu\nu} R_{\mu\nu}$, where we automatically sum over repeated indices on different levels (the Einstein convention). The **Einstein field equation** (10 scalar equations) is

$$R_{\mu\nu} - \tfrac{1}{2} S g_{\mu\nu} = \frac{-8\pi K}{c^2} T_{\mu\nu}, \tag{15.13}$$

where K is the universal gravitational constant and $T_{\mu\nu}$ is the symmetric stress-energy-momentum tensor which is formed in a canonical way from the E-M field and a continuous approximation of the energy-momentum density of particle-like matter (cosmologists sometimes take these particles to be entire galaxies). In essence, the Einstein field equation (15.13) tells us how the nongravitational stress-energy-momentum $T_{\mu\nu}$ of radiation and matter influences the curvature of space-time. Neutral particles "move" along geodesics $\gamma(s)$ of space-time such that $|\gamma'(s)|^2 = c^2$.

The apparent curvature of such geodesics when projected onto what we perceive as space, is due to gravity which is just the geometry of space-time.

The vanishing of the Ricci tensor (and hence the scalar curvature) does not imply that space-time is locally flat. Indeed, the full curvature tensor has ten additional components that constitute the *Weyl conformal curvature tensor* (defined in Section 16.5). Thus, it is quite possible to have a curved space-time which is devoid of matter and radiation (i.e., $T_{\mu\nu} = 0$) which satisfies the so-called empty space equation $R_{\mu\nu} - \frac{1}{2} S g_{\mu\nu} = 0$. This equation can be formulated in terms of a variational principle. Indeed, let D be a compact domain in a space-time M with metric tensor g. Let L be the functional that assigns to each metric tensor g', the quantity

$$L_D(g) := \int_D S(g)\, \mu_g, \tag{15.14}$$

where $S(g)$ is the scalar curvature of g and μ_g is its volume element. It is found (e.g., see [**Bl81**, p. 125] for a coordinate-free proof) that g is a critical point of L within the space of those g' agreeing with g on the boundary of D, if and only if g satisfies the empty space equation $R_{\mu\nu} - \frac{1}{2} S g_{\mu\nu} = 0$ in D. In order to obtain the full equation (15.13) including $T_{\mu\nu}$, it is necessary to add additional terms to L for each type of nongravitational particle or field that resides in space-time. These terms are known as "actions" or Lagrangians (L itself is the purely gravitational Lagrangian). The action over D for an E-M field $F = F_{\mu\nu} dx^\mu \wedge dx^\nu$ (see (15.3)) for a fixed metric g is proportional to $-\frac{1}{2}\int_D |F|_g^2\, \mu_g$, where

$$|F|_g^2 := \tfrac{1}{2} g^{\mu\nu} g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma} \tag{15.15}$$

(i.e., the standard "Lorentz-invariant" norm-square relative to the metric tensor g). In the special case of Minkowski space with flat metric g, writing F in terms of inertial coordinates as in (15.3), we have

$$-\tfrac{1}{2} |F|_g^2 = \tfrac{1}{2}\left(|\mathbf{E}|^2 - |\mathbf{B}|^2\right). \tag{15.16}$$

Under a change of inertial coordinate system on Minkowski space $|\mathbf{E}|^2$ and $|\mathbf{B}|^2$ can change, but $|\mathbf{E}|^2 - |\mathbf{B}|^2$ is invariant. The 2-form F is a "covariant object", but (as with coordinate time) $\mathbf{E}$ and $\mathbf{B}$ separately have no absolute significance. The combined Lagrangian over D is

$$L_D(g,F) := \int_D S(g)\, \mu_g - \tfrac{k}{2} \int_D |F|_g^2\, \mu_g, \tag{15.17}$$

For an arbitrary covariant symmetric 2-tensor h, we have the following for the partial directional derivative of $L(g,F)$ at g in the direction h

$$\begin{aligned} &\tfrac{d}{dt} L(g+th,F)\big|_{t=0} \\ &= \int_D \left(-R_{\mu\nu} + \tfrac{1}{2} S g_{\mu\nu} + k F_\mu{}^\sigma F_{\nu\sigma} - \tfrac{k}{4} |F|_g^2\, g_{\mu\nu}\right) h^{\mu\nu}\, \mu_g \end{aligned} \tag{15.18}$$

Thus, g is a critical point for $L(g,F)$ in the sense that this directional derivative is 0 for all h, when

$$R_{\mu\nu} - \tfrac{1}{2} S g_{\mu\nu} = k\left(F_\mu{}^\sigma F_{\nu\sigma} - \tfrac{1}{4} |F|_g^2\, g_{\mu\nu}\right). \tag{15.19}$$

For suitable k, depending on the choice of units, this is the Einstein field equation (15.13) for a space-time with E-M radiation. The right side is indeed proportional to the accepted stress-energy-momentum tensor for the E-M field F.

We can also get Maxwell's equation $\delta F = 0$ from a variational principle; note that $j = 0$ in the absence of sources which we assume here. Indeed, consider $\frac{1}{2}\int_D |F|_g^2 \; \mu_g$ as a functional of F, instead of g. We assume that F and its variations satisfy the other Maxwell equation $dF = 0$. Assuming that D is simply-connected, we can write any variation of F' of F as dA' for some 1-form A'. The associated variation of $\frac{1}{2}\int_D |F|_g^2 \; \mu_g$ is the directional derivative

$$\begin{aligned} &\tfrac{d}{dt}\left(\tfrac{1}{2}\int_D |F + tF'|_g^2 \; \mu_g\right)\bigg|_{t=0} = \int_D \langle F', F\rangle \; \mu_g \\ (15.20) \qquad &= \int_D \langle dA', F\rangle \; \mu_g = \int_D \langle A', \delta F\rangle \; \mu_g, \end{aligned}$$

assuming that A' vanishes on the boundary of D. This variation is 0 for all such A' exactly when $\delta F = 0$ in D. In summary, the vanishing of the first variations of $L(g, F)$ with respect to g and F are Einstein's equation (15.19) and Maxwell's equation $\delta F = 0$, respectively.

The Maxwell equation $dF = 0$ implies that F can be written locally as $F = -dA$, where A is a 1-form known to physicists as the "4-vector potential", and the minus sign stems from the fact that in mechanics forces generally act in the direction opposite the gradient of the potential energy. Such an A (satisfying $F = -dA$) exists on any simply-connected domain where $dF = 0$ and is called a *gauge potential* for F. However, A is not unique, since for any function $\varphi \in C^\infty(M)$, $F = -dA = -d(A + d\varphi)$, whence $A + d\varphi$ also serves as a gauge potential for F. The transformation $A \mapsto A + d\varphi$ is called a "gauge transformation". In terms of A, the equation $\delta F = 0$ becomes the wave equation $-\delta dA = 0$; in Minkowski space, $-\delta d = c^{-2}\partial_t^2 - \partial_x^2 - \partial_y^2 - \partial_z^2$. It is convenient to regard A as more fundamental than F, since $dF = 0$ follows immediately from $F = -dA$, and from the wave equation $\delta dA = 0$, we see that singularities of A propagate with speed c. However, the fact that A is not uniquely determined by F is somewhat of a drawback. To alleviate this, the so-called *Lorentz condition* $\delta A = 0$ is sometimes imposed on A, but there are generally plenty of functions φ for which $\delta(A + d\varphi) = 0$, namely solutions φ of the scalar wave equation $\delta d\varphi = 0$.

The E-M field strength F is not built into the metric $g_{\mu\nu}$, and Einstein spent many years trying to incorporate F into the geometry of space-time, thereby obtaining a "unified field theory". Actually, in [**Kal**] and [**Kle26**], it was shown that E-M and gravity could be geometrized simultaneously by forming a 5-dimensional manifold by attaching circles to the points of space-time. In this remarkable Kaluza-Klein theory, the illusion that the universe has only four space-time dimensions is not necessarily due to the smallness of the circles (although the theory predicts that they are very small), but rather it is due to the perfect homogeneity in the circular direction. Off hand, it is difficult to believe that anything useful can come from adding an unobservable dimension, but indeed all "grand unified theories" (GUTs) add at least 24 dimensions in order to unify all known non-gravitational forces. The strings in string theory live in a 10 or 26 dimensional space-time (depending on whether supersymmetry is incorporated or not), not even counting the gauge dimensions. For the remainder of this section we describe the geometrical unification that occurs for the simplest Kaluza-Klein theory.

In modern terminology (made precise in Section 16.1), the original Kaluza-Klein theory introduces a fiber bundle $\pi : P \to M$ over space-time M, whose fibers $\pi^{-1}(x)$ are circles. The pull-back $\pi^* g$ of the metric g on M is degenerate on the fibers. Thus, we need additional structures to complete $\pi^* g$ to a Lorentzian metric on P. While we do not assume that each fiber $\pi^{-1}(x)$ is explicitly identified with a standard circle, we do suppose that there is a notion of what it means for a point $p \in P$ to be rotated through an angle θ along its fiber, say $p \mapsto R_\theta(p)$. We let ∂_θ be the vector field given by

$$(\partial_\theta)_p = \left.\tfrac{d}{d\theta} R_\theta(p)\right|_{\theta=0} \tag{15.21}$$

Of course, there is a standard metric on the fibers which gives ∂_θ fixed length throughout P, but we need to specify what vectors in TP should be orthogonal to the fibers. This is conveniently accomplished by introducing a 1-form $\tilde{A}$ on P, such that $\tilde{A}(\partial_\theta) = 1$ and $R_\theta^* \tilde{A} = \tilde{A}$, so that the distribution of kernels of $\tilde{A}$ is preserved by R_θ. A non-degenerate metric $\tilde{g}$ on P is then given (for $X, Y \in T_p P$) by

$$\tilde{g}(X,Y) = (\pi^* g)(X,Y) + k\tilde{A}(X)\,\tilde{A}(Y) = g(\pi_* X, \pi_* Y) + k\tilde{A}(X)\,\tilde{A}(Y), \tag{15.22}$$

where $k > 0$ is some constant to be determined (i.e., $\tilde{g} = \pi^* g + k\tilde{A} \otimes \tilde{A}$). The subspace of $T_p P$ which is orthogonal to the fiber through p (i.e., orthogonal to $(\partial_\theta)_p$) is then the kernel of $\tilde{A}_p$. Note that the rotation $R_\theta : P \to P$ is an isometry of $(P, \tilde{g})$ since

$$\begin{aligned} R_\theta^* \tilde{g} &= R_\theta^* \pi^* g + k R_\theta^* \left(\tilde{A} \otimes \tilde{A}\right) = (\pi \circ R_\theta)^* g + k \left(R_\theta^* \tilde{A}\right) \otimes \left(R_\theta^* \tilde{A}\right) \\ &= \pi^* g + k\tilde{A} \otimes \tilde{A} \end{aligned} \tag{15.23}$$

where we have used the fact that $\pi \circ R_\theta = \pi$ (i.e., R_θ preserves fibers setwise), and $R_\theta^* \tilde{A} = \tilde{A}$. Note that $\tilde{A}$ can be recovered from $\tilde{g}$ by taking $\operatorname{Ker}\left(\tilde{A}_p\right)$ to be the subspace of $T_p P$ which is $\tilde{g}$-orthogonal to ∂_θ, and $\tilde{A}_p(\partial_\theta) = 1$.

We now come to the main reasons for introducing this 5-dimensional cylindrical universe $(P, \tilde{g})$ with its associated 1-form $\tilde{A}$ derived from $\tilde{g}$ and the circle action R_θ. The first very striking fact (e.g., see [**Bl81**]) is that if γ is a geodesic in P relative to $\tilde{g}$, with $\tilde{A}(\gamma') \neq 0$, then the projection $\bar{\gamma} := \pi \circ \gamma$ of γ onto M is the path of a charged particle subject to an E-M field. The Faraday 2-form F of this E-M field is the unique 2-form on M such that $\pi^* F = -d\tilde{A}$. The existence of such F is in part a consequence of the invariance of A under pull back by R_θ. The charge/mass ratio of the particle is proportional to $\tilde{A}(\gamma')$, or equivalently $\tilde{g}(\gamma', \partial_\theta)$, i.e., essentially the vertical component of γ' relative to $\tilde{g}$. If $\tilde{A}(\gamma') = 0$, then $\bar{\gamma}$ is the path of neutral particle in (M, g), i.e., a space-time geodesic. The fact that ∂_θ is a vector field generated by a 1-parameter group of isometries R_θ implies that the charge/mass ratio $\tilde{g}(\gamma'(s), \partial_\theta)$ is constant, independent of s, as it should be. The equation $\pi^* F = -d\tilde{A}$ suggests that $\tilde{A}$ is related to potential 1-forms of F. Indeed, suppose that there is an open set $U \subseteq M$ and a map $\sigma : U \to P$, such that $\pi \circ \sigma = I$ (i.e., σ is a local section of the bundle $\pi : P \to M$). Then $A_\sigma := \sigma^* \tilde{A}$ is locally a potential 1-form of F, since on U we have

$$-dA_\sigma = -d\left(\sigma^* \tilde{A}\right) = -\sigma^* d\tilde{A} = \sigma^* \pi^* F = (\pi \circ \sigma)^* F = F. \tag{15.24}$$

EXERCISE 15.3. Suppose that $\sigma' : U \to P$ is another local section, say $\sigma'(x) = R_{\varphi(x)}(\sigma(x))$, for some function $\varphi \in C^{\infty}(U)$. Verify that

$$A_{\sigma'} := \sigma'^{*}\tilde{A} = \sigma^{*}\tilde{A} + d\varphi = A_{\sigma} + d\varphi. \tag{15.25}$$

Thus, $A_{\sigma'}$ is related to A_{σ} by a gauge transformation.

REMARK 15.4. Actually, nowadays many refer to the transformation $p \mapsto R_{\varphi(\pi(p))}(p)$ of $\pi^{-1}(U)$ as a gauge transformation, and the transformation $A_{\sigma} \mapsto A_{\sigma'} = A_{\sigma} + d\varphi$ is induced by it. Moreover, differential geometers refer to the invariant 1-form $\tilde{A}$ on P as a *connection 1-form*. We will develop these notions much more systematically in the next chapter.

The crucial point here is that E-M forces and gravity are simultaneously encoded in the metric $\tilde{g}$ on P, thereby achieving a geometrical unification of these forces. It is interesting to observe that the addition of an extra dimension to space-time to geometrize E-M is very much in the same spirit that time was adjoined to space in order to geometrize gravity in general relativity.

Even by itself, the fact that the geodesics of $(P, \tilde{g})$ project to paths of charged particles would be sufficient to take the 5-dimensional Kaluza-Klein theory seriously, but there is yet another surprise. The scalar curvature $S(\tilde{g})$ of the metric $\tilde{g}$ is constant on each fiber and thus projects to a well-defined function on M, still denoted by $S(\tilde{g})$. However, $S(\tilde{g})$ is not just the scalar curvature $S(g)$ of g, but rather it is given (e.g., see [**Bl81**]) by

$$S(\tilde{g}) = S(g) - \tfrac{1}{2}k\,|F|_g^2\,. \tag{15.26}$$

Consequently, the combined Lagrangian of gravity and E-M (see (15.17)) can then be written simply as

$$L_D(g,F) = \int_D S(g)\ \mu_g - \tfrac{k}{2}\int_D |F|_g^2\ \mu_g = \int_D S(\tilde{g})\ \mu_g. \tag{15.27}$$

Thus, the scalar curvature $S(\tilde{g})$ of P yields the combined Lagrangian. Roughly put, the Einstein field equation in a non-empty universe with an E-M field (but no matter) is obtained from an empty bundle universe, in the sense that the E-M stress-energy-momentum source is encoded in the geometry of the metric $\tilde{g}$. All of this admits suitable generalization to the case where the fibers are not just circles, but rather general Lie groups (typically $\mathrm{SU}(N)$, $\mathrm{SO}(N)$ or products of these in physical applications). The 1-forms on these higher dimensional bundles are Lie-algebra-valued connection 1-forms which physicists call "gauge potentials" when they are pulled down to M via a local section. The corresponding field strengths (known as "curvatures" to differential geometers) are no longer $\mathbb{R}$-valued 2-forms such as the Faraday F, but rather they have values in certain vector bundles over M. There is the rather obvious hope that these field strengths describe the other forces. For example there are the weak forces that cause, among other events, the decay of the neutron; and the strong forces that are indirectly responsible for holding the nucleus together and directly responsible for binding quarks together inside individual hadrons such as the proton, neutron, pions, etc.. However, one must be wary about extrapolating classical field theory (which is all we have discussed up to this point) to such small systems which are governed by quantum theory.

2. Quantum Theory

Just as Newtonian mechanics breaks down for systems moving at high speeds near that of light, classical field theory does not describe systems of atomic dimensions or smaller very well. The classical picture told us that there are diffuse and wave-like background fields such as the electromagnetic (E-M) field F and the metric tensor g of general relativity, and in sharp distinction to these there were point-like particles that move in trajectories determined by these fields, as well as influencing them. However, the reader has no doubt heard that under certain conditions, light (E-M radiation) produces results that are better understood by assuming that it is made of a stream of particles, known as photons. Notably, when light falls on certain metallic surfaces in a vacuum electrons are emitted from the atoms at a rate which can typically be billions of times larger than the rate that is calculated under the assumption that each atom absorbs all of the energy of the energy it receives from the continuous E-M wave that contacts it. The most natural explanation of this photoelectric effect, is that the E-M wave is *not* continuous, but rather it is made of chunks (quanta) that have sufficient energy (depending on the wave length) to immediately dislodge the electrons from the atoms they come in contact with. It was Einstein who was awarded a Nobel Prize in 1922 in part for his explanation of the photo-electric effect in terms of the quantum theory of light. However, he veered away from the dramatic developments in quantum mechanics, preferring to work on unifying the classical field theories of E-M and gravity without adding an extra dimension as in the Kaluza-Klein theory. He did not succeed. Just as electro-magnetic fields exhibit particle-like properties, it was also discovered that particles (e.g., electrons) exhibit wave-like properties. In an experiment where electrons are fired at a double-slit they collectively make a diffraction pattern of impacts on a screen behind the slit. Thus, the sharp "particle versus wave" dichotomy in classical physics must admit some fuzziness.

Quantum mechanics and quantum field theory grew out of the attempt to describe this state of affairs and to make predictions as accurately as possible. One of the most perplexing phenomena confronting the founders of quantum mechanics was that of atomic spectra. A spectrograph reveals that atoms emit and absorb light at fairly discrete wave-lengths or energies. The classical "planetary model" for the hydrogen atom has the electron circling the proton under the inverse square Coulomb law. It predicts that the electron will radiate E-M energy at a continuously increasing rate and will actually spiral into the proton as it gives up its energy in a short time, rendering the atom unstable. Although it first seems speculative, one might hypothesize that the various energy levels of the atom are actually eigenvalues of some differential operator, just as the frequencies of a vibrating string are the eigenvalues of a constant multiple of $\frac{d^2}{dx^2}$ acting on the space of functions vanishing at the ends. If the eigenvalues of the operator are to represent energies, the operator should have the physical dimensions of energy. The Coulomb potential energy (due to the charge of the nucleus) of an electron at distance r to the nucleus of a hydrogenic atom (or ion) with Z protons is $-Ze^2/r$, where e is the proportional to the charge of the electron, depending on the system of units. Moreover, the simplest rotationally invariant differential operator on $\mathbb{R}^3$ is the Laplacian $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$. The most obvious operator, having dimensions of energy, formed from $-Ze^2/r$ and Δ is $\alpha\Delta - Ze^2/r$ where α is a constant with dimensions of energy times length^2. The point spectrum (assuming $\alpha < 0$) of this operator (densely defined on

$L^2(\mathbb{R}^3))$ is found to be $Z^2e^4/(4\alpha n^2)$, $n = 1, 2, 3, \ldots$. For hydrogenic atoms, there is a choice of α such that this spectrum is consistent with the observed energy spectrum. (Note that the observed spectral lines are at energies which are actually *differences* of the above eigenvalues, as the electron jumps between the possible energy levels.) In order to describe α, let m be the mass of the electron and let M be the much larger mass of the nucleus. We define the reduced mass of the system to be $\mu = m(1 + m/M) \approx m$. The experimentally suitable value for α is found to be $-\frac{\hbar^2}{2\mu}$, where $\hbar = \frac{h}{2\pi} \approx 6.6256 \times 10^{-27}$ erg. sec., and h is Planck's constant introduced by Max Planck around 1900 in connection with black body radiation. One can easily check that $\frac{\hbar^2}{2\mu}$ has units of energy times length^2 so that $-\frac{\hbar^2}{2\mu}\Delta$ has units of energy. Hence we have the striking coincidence that the point spectrum of the operator

$$\hat{E} := -\frac{\hbar^2}{2\mu}\Delta - Ze^2/r \tag{15.28}$$

coincides to good approximation with the energy levels

$$\frac{Z^2e^4}{4\alpha n^2} = -\frac{Z^2e^4}{4\frac{\hbar^2}{2\mu}n^2} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} \tag{15.29}$$

of a hydrogenic atom with Z protons.

A procedure, by which one replaces a classical observable A (e.g., energy, momentum, position) by an operator $\hat{A}$ whose spectrum ranges over all possible experimentally observed values of the observable, is known as a "quantization." Thus, we have roughly succeeded in quantizing the energy, say E, of an electron in a Coulomb potential, by replacing this energy by the operator $\hat{E}$ (15.28). Note that

$$-\frac{\hbar^2}{2\mu}\Delta = -\frac{\hbar^2}{2\mu}\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right) = \frac{1}{2\mu}\left((\pm i\hbar\partial_x)^2 + (\pm i\hbar\partial_y)^2 + (\pm i\hbar\partial_z)^2\right) \tag{15.30}$$

resembles the classical expression $\frac{1}{2\mu}\left(p_x^2 + p_y^2 + p_z^2\right)$ for the kinetic energy of an object of mass m and momentum $\mathbf{p} = p_x\mathbf{i} + p_y\mathbf{j} + p_z\mathbf{k}$. This suggests that the quantization of the classical observable $\mathbf{p}$ should be (where the minus sign is conventional)

$$\hat{\mathbf{p}} = \hat{p}_x\mathbf{i} + \hat{p}_y\mathbf{j} + \hat{p}_z\mathbf{k} := (-i\hbar\partial_x)\,\mathbf{i} + (-i\hbar\partial_y)\,\mathbf{j} + (-i\hbar\partial_z)\,\mathbf{k}, \tag{15.31}$$

and $-\frac{\hbar^2}{2\mu}\Delta$ is the quantization of the kinetic energy of an object of mass μ. Our atomic example then suggests that the quantization of a classical potential energy function $V(\mathbf{r})$ would be the multiplication operator should be the multiplication operator $\hat{V}$ given by $\hat{V}(\psi)(\mathbf{r}) := V(\mathbf{r})\,\psi(\mathbf{r})$, for $\psi : \mathbb{R}^3 \to \mathbb{C}$. In particular, for the coordinate function x, we should have $\hat{x}(\psi)(\mathbf{r}) = x\psi(\mathbf{r})$, etc.. We have been very vague about the domains (spaces of functions $\psi : \mathbb{R}^3 \to \mathbb{C}$) of these operators. In practice, physicists feel comfortable with this vagueness, as long as they can make physical sense of their results. For example, if $a \in \mathbb{R}^3$, the function $\psi(\mathbf{r}) = e^{i\mathbf{a}\cdot\mathbf{r}/\hbar}$ is a simultaneous eigenfunction of the momentum operators $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ in the sense that $\hat{p}_x(\psi) = a_1\psi$, etc., but this ψ is not in $L^2(\mathbb{R}^3, \mathbb{C})$. Also, the Dirac delta distribution $\delta(\mathbf{r} - a)$ might be regarded as an eigenfunction for the position operators $\hat{x}$, $\hat{y}$, $\hat{z}$. Perhaps the most appropriate domain would be the space of tempered distributions (i.e., continuous linear functionals on the Schwartz space of rapidly decreasing functions); at least this would be big enough to encompass the above examples. At any rate, the "functions" in the domains of these operators are known as *states* of the particle (e.g., the electron, in the above atomic example). We mention that the states describing *systems* of particles are essentially tensor products (sometimes symmetric and sometimes skew-symmetric) of the individual particles. States which differ by a constant, complex factor are identified (considered physically indistinguishable), making the space of states an infinite-dimensional complex projective space. The following is a basic interpretive assumption of quantum mechanics, which gives it some physical sense.

Postulate. Let $[\psi]$ be a state with representative ψ of norm $\|\psi\| = 1$ in a Hilbert space H (typically $L^2\left(\mathbb{R}^3, \mathbb{C}\right)$ in the single particle case). Suppose that a quantized observable (i.e., a self-adjoint operator) A has a eigenvalue λ. Then the probability that the observable is measured to be λ when the particle (or system) is in the state $[\psi]$ is the norm-square of the projection of ψ onto the eigenspace of λ. More generally, if the self-adjoint operator A has a spectral resolution (i.e., a projection-valued measure P on $\mathbb{R}$ with $A = \int_{-\infty}^{\infty} \lambda \, dP_\lambda$), then the probability that the observable is measured to be in some interval I when the particle (or system) is in state $[\psi]$ is $\left\langle \left(\int_I dP_\lambda\right)(\psi), \psi\right\rangle$.

If the self-adjoint, quantized observable A has a complete set of eigenvectors u_n, $n = 1, 2, 3, \ldots$, with $Au_n = \lambda_n u_n$, then according to the Postulate, the expectation of measurements of this observable for the state $[\psi]$ ($\|\psi\| = 1$) is simply

$$\begin{aligned} \sum_{n=1}^{\infty} \lambda_n \left|\langle \psi, u_n\rangle\right|^2 &= \sum_{n=1}^{\infty} \lambda_n \langle \psi, u_n\rangle \overline{\langle \psi, u_n\rangle} = \sum_{n=1}^{\infty} \langle \psi, Au_n\rangle \overline{\langle \psi, u_n\rangle} \\ &= \left\langle A\psi, \sum_{n=1}^{\infty} \langle \psi, u_n\rangle u_n \right\rangle = \langle A\psi, \psi\rangle . \end{aligned} \tag{15.32}$$

We get the same end result in the general case where A has a spectral resolution.

As a consequence of (15.32) in the single particle case, we show that $\mathbf{r} \mapsto |\psi(\mathbf{r})|^2$ (where $\|\psi\|^2 = \int_{\mathbb{R}^3} |\psi(\mathbf{r})|^2 \, d^3r = 1$) is the probability density for the position of the particle in state $[\psi]$. Indeed, for a domain $D \subseteq \mathbb{R}^3$, let $\chi_D : \mathbb{R}^3 \to \{0, 1\}$ be the characteristic function of D. Classically, this observable is 1 if the particle is in D and 0 otherwise. As for functions on $\mathbb{R}^3$ in general, the quantization of χ_D is the multiplication operator $\hat{\chi}_D$ on $L^2\left(\mathbb{R}^3, \mathbb{C}\right)$ given by $\hat{\chi}_D(u)(\mathbf{r}) = \chi_D(\mathbf{r})\, u(\mathbf{r})$. According to (15.32), the quantum mechanical expectation of this observable for the state $[\psi]$ is

$$\langle \hat{\chi}_D \psi, \psi\rangle = \int_{\mathbb{R}^3} \langle \chi_D(\mathbf{r})\, \psi(\mathbf{r}), \psi(\mathbf{r})\rangle \, d^3r = \int_D |\psi(\mathbf{r})|^2 \, d^3r. \tag{15.33}$$

This shows that $|\psi|^2$ is the probability density for the position of the particle. Our aim now is to show that the probability density for the momentum of the particle is the function $\mathbf{p} \mapsto \left|\tilde{\psi}(\mathbf{p})\right|^2$, where

$$\tilde{\psi}(\mathbf{p}) := (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} \psi(\mathbf{r})\, e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} \, d^3r, \tag{15.34}$$

which is essentially the Fourier transform of ψ. By the Fourier Inversion Theorem,

$$\psi(\mathbf{r}) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} \tilde{\psi}(\mathbf{p})\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \, d^3p, \tag{15.35}$$

and formally

$$\hat{p}_x(\psi)(\mathbf{r}) = -i\hbar\partial_x\psi(\mathbf{r}) = (2\pi\hbar)^{-3/2}\int_{\mathbb{R}^3} p_x\tilde{\psi}(\mathbf{p})\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\, d^3p. \tag{15.36}$$

More generally, for suitable functions $f(\mathbf{p})$ (momentum-dependent observables), it is natural to let

$$\hat{f}(\psi)(\mathbf{r}) := (2\pi\hbar)^{-3/2}\int_{\mathbb{R}^3} f(\mathbf{p})\,\tilde{\psi}(\mathbf{p})\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\, d^3p. \tag{15.37}$$

For a domain D in *momentum* space and associated characteristic function $\chi_D(\mathbf{p})$, we then have

$$\hat{\chi}_D(\psi)(\mathbf{r}) = (2\pi\hbar)^{-3/2}\int_{\mathbb{R}^3} \chi_D(\mathbf{p})\,\tilde{\psi}(\mathbf{p})\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\, d^3p. \tag{15.38}$$

The quantum mechanical expectation that the particle has momentum in D is then

$$\langle\hat{\chi}_D(\psi),\psi\rangle = \left\langle \chi_D\tilde{\psi},\tilde{\psi}\right\rangle = \int_D \left|\tilde{\psi}(\mathbf{p})\right|^2 d^3p, \tag{15.39}$$

where we have used Parseval's equality. This identifies $\left|\tilde{\psi}\right|^2$ as the probability density for the momentum of the particle in the state $\left[\tilde{\psi}\right]$. Note that pseudo-differential operators are essentially quantizations of functions of momentum.

In classical physics, particles move along trajectories and have well-defined positions and momenta at all times. In quantum mechanics, position and momentum cannot both be simultaneously determined with arbitrarily high precision. This is a consequence of the (Heisenberg) Uncertainty Principle which can be deduced as follows. The uncertainty of an observable A for a particle in state $[\psi]$ (with $\|\psi\| = 1$) is the standard deviation of the measurements of the observable, or equivalently, the root of the expectation of measurements of the observable $(A - \langle A\psi,\psi\rangle)^2$, namely

$$\Delta_\psi A := \left\langle (A - \langle A\psi,\psi\rangle)^2\psi,\psi\right\rangle^{1/2} = \|(A - \langle A\psi,\psi\rangle)\psi\| \tag{15.40}$$

As shown below, the Cauchy-Schwarz inequality and a little algebra reveal that for observables A and B we have the Uncertainty Principle

$$\Delta_\psi A\cdot\Delta_\psi B \geq \tfrac{1}{2}\left|([A,B]\psi,\psi)\right|, \tag{15.41}$$

where $[A,B] = AB - BA$. Indeed,

$$\begin{aligned}
\Delta_\psi A\cdot\Delta_\psi B &= \left((A-(A\psi,\psi))^2\psi,\psi\right)^{1/2}\left((B-(B\psi,\psi))^2\psi,\psi\right)^{1/2} \\
&= \|(A-(A\psi,\psi))\psi\|\,\|(B-(B\psi,\psi))\psi\| \\
&\geq |((A-(A\psi,\psi))\psi,(B-(B\psi,\psi))\psi)| \\
&\geq |\mathrm{Im}\,((A-(A\psi,\psi))\psi,(B-(B\psi,\psi))\psi)| \\
&= \tfrac{1}{2}\left|\begin{array}{c} ((A-(A\psi,\psi))\psi,(B-(B\psi,\psi))\psi) \\ -((B-(B\psi,\psi))\psi,(A-(A\psi,\psi))\psi) \end{array}\right| \\
&= \tfrac{1}{2}\left|\left(\begin{array}{c} (A-(A\psi,\psi))(B-(B\psi,\psi)) - \\ ((B-(B\psi,\psi))(A-(A\psi,\psi)))\psi,\psi \end{array}\right)\right| \\
&= \tfrac{1}{2}\left|((AB-BA)\psi,\psi)\right| = \tfrac{1}{2}\left|([A,B]\psi,\psi)\right|. \qquad (15.42)
\end{aligned}$$

Consequently, for two noncommuting quantized observables, there is a possible obstruction to obtaining arbitrarily low uncertainties in the measurements of both.

For example, $[\hat{x}, \hat{p}_x](\psi) = x(-i\hbar\partial_x\psi) - (-i\hbar\partial_x(x\psi)) = i\hbar\psi$ (i.e., $[\hat{x}, \hat{p}_x] = i\hbar I$), and so

$$\Delta_\psi \hat{x} \cdot \Delta_\psi \hat{p}_x \geq \tfrac{1}{2}|(i\hbar\psi, \psi)| = \tfrac{\hbar}{2}|(\psi, \psi)| = \tfrac{\hbar}{2}. \tag{15.43}$$

Thus in quantum mechanics, the position and momentum of a particle in the same direction cannot both be determined with arbitrarily high precision. For example, initial conditions for Newton's equation (or its relativistic analogs) for a particle cannot be exactly specified, and the philosophy of determinism loses its grip on reality. However, classical mechanics works well in ordinary circumstances due to the smallness of $\hbar$.

EXERCISE 15.5. Show that if a function $\psi \in C^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ is a state for which $\Delta_\psi \hat{x} \cdot \Delta_\psi \hat{p}_x$, $\Delta_\psi \hat{y} \cdot \Delta_\psi \hat{p}_y$ and $\Delta_\psi \hat{z} \cdot \Delta_\psi \hat{p}_z$ all have the minimal value $\frac{\hbar}{2}$, then $|\psi|^2$ is a normal (Gaussian) distribution in each variable. For simplicity, you may wish to assume (at first) that $(\hat{x}\psi, \psi) = (\hat{p}_x\psi, \psi) = 0$ (and similarly for y and z). To proceed, consider when the inequalities in 15.42 are equalities.

So far we have said nothing about how states and/or observables evolve with time in quantum mechanics. While there are a number of standard ways of introducing time evolution, here we will proceed in a somewhat unusual manner, by drawing upon special relativity for motivation. Suppose that

$$\gamma(s) = \left(x^0(s), x^1(s), x^2(s), x^3(s)\right) = (ct(s), \mathbf{r}(s)) \tag{15.44}$$

is the trajectory of a particle of rest mass m_0 in Minkowski space-time, parametrized in the standard way so that $\langle \gamma'(s), \gamma'(s)\rangle = c^2 t'(s)^2 - \|\mathbf{r}'(s)\|^2 = c^2$. The energy-momentum 4-vector of the particle is

$$\begin{aligned} p(s) &:= m_0\gamma'(s) = m_0(ct'(s), \mathbf{r}'(s)) \\ &= m_0(c\beta(s), \beta(s)\mathbf{v}(t(s))) = (mc, m\mathbf{v}) = (E/c, \mathbf{p}). \end{aligned} \tag{15.45}$$

Since the quantization of $\mathbf{p}$ is $\hat{\mathbf{p}} = -i\hbar\nabla$, it is fitting that (on the basis of covariance) more generally we should have $\hat{p}^\mu = i\hbar g^{\mu\nu}\partial_\nu$. As $g^{ii} = -1$ for $i = 1, 2, 3$, we then have $\hat{\mathbf{p}} = -i\hbar\nabla$, whereas for $\mu = 0$, we obtain $\hat{E}/c = \hat{p}^0 = i\hbar g^{00}\partial_0 = +i\hbar\partial_0 = +i\hbar\frac{1}{c}\partial_t$ (plus!). In other words, one ought to have $\hat{E} = i\hbar\partial_t$. However, most physicists do not think of $i\hbar\partial_t$ as being the quantization $\hat{E}$ of the classical energy E, since $\hat{E}$ is generally determined by other means. For example, in the nonrelativistic setting where $E = \frac{1}{2m}\|\mathbf{p}\|^2 + V(\mathbf{r})$ for some potential V, we have seen that $\hat{E} = -\frac{\hbar^2}{2m}\Delta + \hat{V}$. In the so-called Schrödinger picture, the relation $\hat{E} = i\hbar\partial_t$ is regarded as defining the time evolution of states $[\psi]$, in the sense that

$$i\hbar\partial_t\psi = \hat{E}\psi. \tag{15.46}$$

Actually, we should be more precise here. The E in $p = (E/c, \mathbf{p})$ is the total energy of the particle, which is close to the sum of kinetic energy, potential energy, *and* rest energy

$$E \approx \frac{1}{2m_0}\|\mathbf{p}\|^2 + V(\mathbf{r}) + m_0c^2, \tag{15.47}$$

where we assume that $V(\mathbf{r})$ is shifted to zero by adding a constant when $\mathbf{p} = \mathbf{0}$, so that $E = m_0c^2$ when the particle is at rest. Thus, according to (15.46)

$$i\hbar\partial_t\psi = \hat{E}\psi = -\frac{\hbar^2}{2m_0}\Delta\psi + V(\mathbf{r})\psi + m_0c^2\psi. \tag{15.48}$$

This is not quite the usual Schrödinger equation because of the term $m_0c^2\psi$. However, if in (15.48) we make the replacement $\psi(\mathbf{r},t) \to e^{-im_0c^2t/\hbar}\psi(\mathbf{r},t)$ (which at each time replaces ψ by an equivalent state), we then obtain the official time-dependent Schrödinger equation

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m_0}\Delta\psi + V(\mathbf{r})\psi \tag{15.49}$$

An obvious defect of (15.49) is its noninvariance under Lorentz transformations which arises from the approximation (15.47). However, it does have the virtue that $\psi(\cdot,0) \mapsto \psi(\cdot,t)$ is a unitary transformation, so that for each time t, $|\psi(\cdot,t)|^2$ may be regarded as a probability density if $\|\psi(\cdot,0)\| = 1$. This unitary property can be seen from the fact that the infinitesimal generator $\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m_0}\Delta + \hat{V}\right)$ is a skew-hermitian operator because of the factor of i.

Exercise 15.6. Show that $\int_{\mathbb{R}^3} |\psi(\mathbf{r},t)|^2 \, d^3r$ is constant by formally differentiating under the integral, using (15.49). You may assume that $\psi(\cdot,t)$ and its spacial derivatives decay rapidly enough as $|\mathbf{r}| \to \infty$ to neglect boundary terms when integrating by parts.

Another nice consequence of (15.49) is that (under suitable decay assumptions on $\psi(\mathbf{r},t)$ and its derivatives as $\|\mathbf{r}\| \to \infty$) the expectation of the position vector of the particle in state $[\psi]$, namely

$$\mathbf{R}(t) := \int_{\mathbb{R}^3} |\psi(\mathbf{r},t)|^2 \, \mathbf{r} \, d^3r, \tag{15.50}$$

obeys not only

$$\mathbf{R}'(t) = \frac{1}{m_0}\mathbf{P}(t) := \frac{1}{m_0}\langle\hat{\mathbf{p}}\psi,\psi\rangle = \frac{1}{m_0}\int_{\mathbb{R}^3} -i\hbar\nabla\psi(\mathbf{r},t)\,\psi(\mathbf{r},t)^* \; d^3r \tag{15.51}$$

where $\mathbf{P}(t)$ is the expectation of the momentum, but also "Newton's equation"

$$m_0\mathbf{R}''(t) = -\int_{\mathbb{R}^3} |\psi(\mathbf{r},t)|^2 \, \nabla V(\mathbf{r}) \; d^3r, \tag{15.52}$$

where we note that the right side is the expectation of the force on the particle in state $[\psi]$.

Exercise 15.7. Formally derive equations (15.51) and (15.52). Again, assume that ψ and its derivatives suitably decay so that the boundary terms produced when integrating by parts can be discarded.

One reason why the nonrelativistic (15.49) and its many-particle generalizations are so successful in dealing with atomic phenomena is that electrons in atoms "travel" at speeds of only around $c/100$, according to a simple approximate classical calculation.

Perhaps it would have been better to insist on Lorentz invariance from the beginning, thereby replacing the relation $(E/c)^2 - \|\mathbf{p}\|^2 = m_0^2c^2$ by its "quantized" analog, namely the Klein-Gordon equation for the $\mathbb{C}$-valued function ψ on Minkowski space

$$-\hbar^2\left(c^{-2}\partial_t^2 - \Delta\right)\psi = m_0^2c^2\psi. \tag{15.53}$$

However, note that there is no vestige of a potential in this equation. If we are to introduce electromagnetism in some way, we should do it in a Lorentz invariant way. If we write (15.53) in the covariant form

$$-\hbar^2 g^{\mu\nu}\partial_\mu\partial_\nu\psi = m_0^2c^2\psi, \tag{15.54}$$

then the most obvious way of introducing E-M is simply to add a multiple of the gauge potential $A = (A_\mu)$ to the operator $\partial = (\partial_\mu)$. However, in order that the resulting equation be invariant under gauge transformations $A \to A' := A + d\varphi$, where $\varphi : M \to \mathbb{R}$, we need to subject ψ to a gauge transformation. In order to keep $|\psi|^2$ gauge-invariant, we might try $\psi \to \psi' := e^{i\varphi}\psi$. Indeed, we have the identity

$$\begin{gathered}\left(\partial_\mu - i\left(A_\mu + \partial_\mu\varphi\right)\right)e^{i\varphi}\psi = e^{i\varphi}\left(\partial_\mu - iA_\mu\right)\psi \\ \text{or} \quad \left(\partial_\mu - iA'_\mu\right)\psi' = \left(\left(\partial_\mu - iA'_\mu\right)\psi\right)'\end{gathered} \tag{15.55}$$

Thus, we obtain the desired invariance

$$\begin{gathered}-\hbar^2 g^{\mu\nu}\left(\partial_\mu - iA_\mu\right)\left(\partial_\nu - iA_\nu\right)\psi = m_0^2c^2\psi \Leftrightarrow \\ -\hbar^2 g^{\mu\nu}\left(\partial_\mu - iA'_\mu\right)\left(\partial_\nu - iA'_\nu\right)\psi' = m_0^2c^2\psi'.\end{gathered} \tag{15.56}$$

In order for the units to work out, A_μ must be replaced by (const.) $\cdot\, A_\mu$ having the same dimensions as ∂_μ, namely $(\text{length})^{-1}$. The natural choice is $\frac{e}{c\hbar}A_\mu$ where e is the charge. Thus, one may incorporate E-M into the Klein-Gordon equation by replacing ∂_μ by $\partial_\mu - \frac{ie}{c\hbar}A_\mu$, obtaining

$$-\hbar^2 g^{\mu\nu}\left(\partial_\mu - \frac{ie}{c\hbar}A_\mu\right)\left(\partial_\nu - \frac{ie}{c\hbar}A_\nu\right)\psi = m_0^2c^2\psi. \tag{15.57}$$

Physicists call this "minimal replacement," while differential geometers recognize that this amounts to replacing ordinary derivatives by covariant derivatives. Note that ψ changes under a change of gauge. Thus, rather than taking ψ to be a $\mathbb{C}$-valued function on space-time, this wave function is more properly regarded as an equivariant $\mathbb{C}$-valued function on the Kaluza-Klein circle bundle P (or equivalently as a section of the associated complex line bundle). Covariant differentiation is then forced upon us, since ordinary differentiation of sections of a vector bundle makes no invariant sense. Observe that $\left|e^{i\varphi}\psi\right|^2 = |\psi|^2$, whence $|\psi|^2$ is gauge invariant. However there is a problem with interpreting $|\psi(t,\cdot)|^2$ as a probability, because, even with $A = 0$, it does not follow from (15.57) that $\int_{\mathbb{R}^3}|\psi(t,\mathbf{r})|^2\, d^3r$ is constant, as with solutions of Schrödinger's equation (15.49). Instead, one finds that the real quantity

$$
\begin{aligned}
&\int_{\mathbb{R}^3} 2\,\mathrm{Im}\left(\psi(t,\mathbf{r})\left(\partial_0 + \frac{ie}{c\hbar}A_0(t,\mathbf{r})\right)\psi(t,\mathbf{r})^*\right)\,d^3r \\
&= \int_{\mathbb{R}^3} i\left(\psi^*\left(\partial_0 - \frac{ie}{c\hbar}A_0\right)\psi - \psi\left(\partial_0 + \frac{ie}{c\hbar}A_0\right)\psi^*\right)\,d^3r \\
&= \int_{\mathbb{R}^3} i\left(\psi^*\partial_0\psi - \psi\partial_0\psi^*\right) - \frac{2e}{c\hbar}A_0\,|\psi|^2\ d^3r \qquad (15.58)
\end{aligned}
$$

is conserved (i.e., independent of t). However, the integrand, say $\rho(t,\mathbf{r})$, is not necessarily of fixed sign everywhere even if $A_0 = 0$, and thus does not represent a probability density. The usual interpretation is that $\rho(t,\mathbf{r})$ is proportional to a charge probability density, but this is odd because ψ is supposedly the state of a single particle of a definite charge e. This difficulty foreshadows the fact that relativistic quantum theories are generally multi-particle theories in which the number of particles is not fixed in the presence of an external potential, and anti-particles of charge $-e$ are naturally built in at the outset. The presence of anti-particles manifests itself in "negative-energy states" ψ for which $i\hbar\partial_t\psi = E\psi$ with $E < 0$.

The full story is known as *quantum field theory,* as opposed to relativistic quantum mechanics. For physicists, quantum field theory (particularly, quantum electrodynamics (QED)) is enormously successful, since very accurate, verifiable predictions are made. However, for mathematicians, conventional quantum field theory leads to an unsatisfactory state of affairs. Indeed, no mathematical quantum field theory (which satisfies a reasonable set of axioms) with realistic interacting particles has ever been constructed in four space-time dimensions. From a mathematical perspective, the great tragedy is that conventional physicists are getting fantastic answers, with little concern that a solid theoretical foundation has yet to be found.

Returning to the Klein-Gordon equation (15.57), it is possible to make contact with nonrelativistic quantum mechanics in a limiting sense, as follows. In order to relate (15.57) to Schrödinger's equation (15.49), for a solution ψ of (15.57), we define $\psi_s := e^{im_0c^2t/\hbar}\psi$. Note that

$$
i\hbar\partial_t\psi_s = i\hbar\partial_t\left(e^{im_0c^2t/\hbar}\psi\right) = e^{im_0c^2t/\hbar}\left(i\hbar\partial_t\psi - m_0c^2\psi\right). \qquad (15.59)
$$

Thus, the transformation $\psi \to \psi_s$ has the effect of removing the rest energy m_0c^2 from ψ. For a solution ψ of (15.57), we expect that ψ_s will approximately satisfy (15.49) for some choice of V. We will show this under the assumptions that $A_i = 0$ for $i = 1,2,3$, A_0 is time-independent ($\partial_t A_0 = 0$) and terms without factors of c^2 are negligible in comparison with those that do. Using (15.59), we then have

$$
\begin{aligned}
i\hbar\partial_t\psi &= e^{-im_0c^2t/\hbar}\left(i\hbar\partial_t\psi_s + m_0c^2\psi_s\right) \approx e^{-im_0c^2t/\hbar}m_0c^2\psi_s \\
-\hbar^2\partial_t^2\psi &= i\hbar\partial_t\left(i\hbar\partial_t\psi\right) = i\hbar\partial_t\left(e^{-im_0c^2t/\hbar}\left(i\hbar\partial_t\psi_s + m_0c^2\psi_s\right)\right) \\
&= e^{-im_0c^2t/\hbar}\left(-\hbar^2\partial_t^2\psi_s + 2m_0c^2 i\hbar\partial_t\psi_s + \left(m_0c^2\right)^2\psi_s\right) \\
&\approx e^{-im_0c^2t/\hbar}\left(2m_0c^2 i\hbar\partial_t\psi_s + \left(m_0c^2\right)^2\psi_s\right). \qquad (15.60)
\end{aligned}
$$

Then under the above assumptions, (15.57) yields

$$\begin{aligned} m_0^2c^2\psi &= -\hbar^2\left(c^{-1}\partial_t - \frac{ie}{c\hbar}A_0\right)\left(c^{-1}\partial_t - \frac{ie}{c\hbar}A_0\right)\psi + \hbar^2\Delta\psi \\ &\approx -c^{-2}\hbar^2\partial_t^2\psi + 2c^{-2}eA_0 i\hbar\partial_t\psi + \frac{e^2}{c^2}A_0^2\psi + \hbar^2\Delta\psi \\ &\approx c^{-2}\left(e^{-im_0c^2t/\hbar}\left(2m_0c^2 i\hbar\partial_t\psi_s + \left(m_0c^2\right)^2\psi_s\right)\right) \\ &+ 2c^{-2}eA_0\left(e^{-im_0c^2t/\hbar}m_0c^2\psi_s\right) + \frac{e^2}{c^2}A_0^2\psi + \hbar^2\Delta\psi \end{aligned} \tag{15.61}$$

or

$$\begin{aligned} i\hbar\partial_t\psi_s &\approx \frac{-\hbar^2}{2m_0}\Delta\psi_s - eA_0\psi_s + \frac{e^2}{2m_0c^2}A_0^2\psi_s \\ &= \frac{-\hbar^2}{2m_0}\Delta\psi_s - eA_0\left(1 - \frac{eA_0}{2m_0c^2}\right)\psi_s \end{aligned} \tag{15.62}$$

Thus, assuming that the electrostatic potential energy eA_0 is small compared with the rest energy m_0c^2 so that $\frac{eA_0}{2m_0c^2}$ is negligible, we approximately have Schrödinger's equation (15.49) with potential $V = -eA_0$. Incidentally, we mention that expressing the density in the conserved quantity (15.58) in terms of $\psi_s = e^{im_0c^2t/\hbar}\psi$ (using the first equation in (15.60)) yields

$$\begin{aligned} & i\left(\psi^*\partial_0\psi - \psi\partial_0\psi^*\right) - \frac{2e}{c\hbar}A_0\left|\psi\right|^2 \\ &= \frac{1}{c\hbar}\left(i\hbar\left(\psi_s^*\partial_t\psi_s - \psi_s\partial_t\psi_s^*\right) + 2\left(m_0c^2 - eA_0\right)\left|\psi_s\right|^2\right) \\ &= \frac{2m_0c}{\hbar}\left(\frac{i\hbar}{2m_0c^2}\left(\psi_s^*\partial_t\psi_s - \psi_s\partial_t\psi_s^*\right) + \left(1 - \frac{eA_0}{m_0c^2}\right)\left|\psi_s\right|^2\right). \end{aligned} \tag{15.63}$$

Thus, the density in (15.58) is for $eA_0 \ll m_0c^2$, etc., is approximately $2m_0c/\hbar$ times the usual Schrödinger probability density $\left|\psi_s\right|^2$.

For the hydrogenic atom where $A_0 = Ze/r$ and $\mathbf{A} = \mathbf{0}$, the relevant exact product solutions $\psi(t,\mathbf{r}) = e^{-iEt/\hbar}R(r)Y_{l,m}(\theta,\varphi)$ of (15.57) can be found (see [**Schi**, p. 470]) and the corresponding energies levels to order 4 in the parameter $\gamma := Ze^2/(\hbar c)$ are given by

$$E_{n,l} = m_0c^2\left(1 - \frac{\gamma^2}{2n^2} - \frac{\gamma^4}{2n^4}\left(\frac{n}{l+\frac{1}{2}} - \frac{3}{4}\right) + \mathrm{O}\left(\gamma^6\right)\right) \tag{15.64}$$

where $n = 1,2,3,\dots$ is the total quantum number and $l = 0,1,\dots,n-1$ is the azimuthal quantum number of the state. The first term m_0c^2 is the rest energy and the second term is

$$-m_0c^2\frac{\gamma^2}{2n^4} = -m_0c^2\frac{\left(Ze^2/(\hbar c)\right)^2}{2n^4} = -\frac{m_0e^4Z^2}{2\hbar^2n^4} \tag{15.65}$$

which coincides with (15.29) with $m_0 = \mu$. The third term is a relativistic correction that predicts that there is a small spread (fine structure) in the energy levels for a fixed n, since the different values $l = 0,1,\dots,n-1$ yield different energies (i.e., the degeneracy is broken). However, the predicted spread is larger than the observed spread. The problem is that, while the Klein-Gordon equation 15.57 is the most obvious relativistic wave equation, it is not the correct one for electrons. Indeed, there is a first-order relativistic equation for a multi-component (spinorial) wave function that works much better, namely the Dirac equation which we consider next.

The search for a first-order relativistic wave equation was partly motivated by the fact that Schrödinger's equation only involves a first derivative with respect to t, and the evolution of states is simply given by a one-parameter group of unitary transformations on $L^2\left(\mathbb{R}^3\right)$ generated by the skew-Hermitian operator $-i\hat{E}/\hbar$ formed from the quantized energy $\hat{E}$. As was eventually discovered, the problem with the Klein-Gordon equation really is not with the second-order time derivative per se, but Dirac's search for a first-order relativistic equation led to the correct equation for electron wave functions. What follows is a rough outline of his reasoning. Consider a first-order differential operator with constant (but possibly complex matrix) coefficients γ^μ, say

$$A = \gamma^\mu \partial_\mu = \gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^2\partial_2 + \gamma^3\partial_3 \quad \left(\partial_0 := c^{-1}\partial_t\right). \tag{15.66}$$

If A is relativistically (or Lorentz) invariant, then so is

$$A^2 = \left(\gamma^\mu\partial_\mu\right)^2 = \tfrac{1}{2}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right)\partial_\mu\partial_\nu. \tag{15.67}$$

For $g_{\mu\nu}dx^\mu dx^\nu = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2$, we have $g^{\mu\nu} = g_{\mu\nu}$, and constant multiples of the operator $g^{\mu\nu}\partial_\mu\partial_\nu$ are Lorentz invariant. Thus, it is reasonable to impose the condition (for some scalar or matrix $K \neq 0$) that

$$\tfrac{1}{2}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right) = Kg^{\mu\nu}. \tag{15.68}$$

If the γ^μ and K are assumed to be complex scalars then there are no such γ^μ, since these scalars would be nonzero and would anticommute. With some perseverance, one can prove that if the γ^μ are $n \times n$ matrices and $K = I_n$ (the $n \times n$ identity), then the least n for which there are solutions to (15.68) is $n = 4$. One standard solution is

$$\gamma^0 = \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}, \quad \gamma^j = \begin{bmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{bmatrix}, \tag{15.69}$$

where

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{15.70}$$

are the so-called Pauli matrices. A source of headaches is the fact that the γ^μ are not unique, since we can replace γ^μ by $B\gamma^\mu B^{-1}$ for any invertible 4×4 matrix B. At any rate, Dirac's equation for a $\mathbb{C}^4$-valued (4-component) wave function ψ is

$$i\hbar\gamma^\mu\partial_\mu\psi = m_0 c\psi \tag{15.71}$$

where m_0 is the rest mass of the particle associated with ψ. The electromagnetic gauge potential 1-form $A = A_\mu dx^\mu$ is again naturally introduced via minimal replacement:

$$i\hbar\gamma^\mu\left(\partial_\mu - \frac{ie}{c\hbar}A_\mu\right)\psi = m_0 c\psi. \tag{15.72}$$

For $(A_\mu) = (Ze/r, \mathbf{0})$, this equation can be separated (see [**Schi**, p. 486]) and one finds that there are solutions of the form $\psi(t, \mathbf{r}) = e^{-iEt/\hbar}\Psi(\mathbf{r})$, where $\Psi(\mathbf{r})$

decays suitably as $\mathbf{r} \to \infty$ the energy levels. They are indexed by $n = 1, 2, 3, \ldots$ and $|k| = 1, 2, \ldots, n$ and are given (to fourth order in $\gamma := Ze^2/(\hbar c)$) by

$$E_{n,|k|} = m_0 c^2 \left(1 - \frac{\gamma^2}{2n^2} + \frac{\gamma^4}{2n^4} \left(\frac{n}{|k|} - \frac{3}{4} \right) + \mathrm{O}\left(\gamma^6\right) \right). \tag{15.73}$$

The fine structure exhibited by the third term agrees much better with observations than that predicted by the Klein-Gordon equation (see (15.64)). Although discrepancies with the observed spectrum still exist (e.g., the Lamb shift), they are accounted for within the more accurate context of quantum electrodynamics (QED) which is a quantum *field* theory. In this theory, the wave function ψ and the E-M gauge potential A are replaced by distributions with values that are operators in a multi-particle Hilbert space of states. Although the mathematics of QED is shady, the formalities involved give rise to recipes for computing physical quantities in terms of formal power series in the dimensionless fine structure constant $\alpha := e^2/(c\hbar) \approx 1/137$. Such series are known are called renormalized perturbations series. The coefficients of such series are computed by summing up integrals associated with so-called Feynman diagrams. One trouble is that integrals associated with Feynman diagrams that have loops are infinite. By various procedures known as renormalization techniques, finite values for the coefficients of the perturbations series are extracted. Field theories for which this is the case (e.g., QED) are known as renormalizable. The various renormalization techniques all lead to the same values for the coefficients, which is reassuring. For QED, the terms of these series become smaller at least initially. Eventually, the coefficients become incalculable due to the huge number and complexity of the Feynman diagrams. The consensus among those who have studied these series in some detail is that the coefficients eventually increase rapidly enough so that the series do not converge. Contrary to popular misconceptions (even held by good physicists) a formal power series in α does not necessarily converge, even if $\alpha \approx 1/137$. However, as with asymptotic series, before divergence sets in, one obtains amazing accuracy compared with experimental results. For example, we have the following values for the magnetic moment of the electron:

$$\begin{array}{rl} \text{Experiment:} & \frac{e\hbar}{2m_e c}\left(1.00115965241 \pm 20\right) \\ \text{QED:} & \frac{e\hbar}{2m_e c}\left(1.00115965238 \pm 26\right) \end{array} \tag{15.74}$$

Thus, in spite of the profound mathematical problems with QED (e.g., its very existence as a mathematical theory, beyond computational recipes), QED is hailed as one of the most successful physical theories from the perspective of most physicists.

Returning to the Dirac equation (15.71), we have not yet indicated the sense in which it is Lorentz invariant. For the Klein-Gordon equation (15.54), Lorentz-invariance means that if ψ is a solution and $L \in \mathrm{O}(1,3)$ is a Lorentz transformation, then $\psi \circ L$ is also a solution. For the Dirac equation, one considers the universal double cover $C : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{O}_0(1,3)$ of the identity component $\mathrm{O}_0(1,3)$ of $\mathrm{O}(1,3)$. There is a representation $r : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{GL}(4,\mathbb{C})$, such that if ψ is a solution of Dirac's equation $i\hbar\gamma^\mu\partial_\mu\psi = m_0 c\psi$, and $A \in \mathrm{SL}(2,\mathbb{C})$ then $A^{-1}\psi \circ r(A)$ is also a solution. This is the meaning of Lorentz-invariance for the Dirac equation. Since the representation r is the sum of *two* irreducible spin-$\frac{1}{2}$ representations, the Dirac equation not only takes into account the spin of the electron, but it forecasts the existence of the positron. In the terminology of modern differential geometry (introduced in the next Chapter), the Dirac wave function is a section of a complex

4-dimensional vector bundle (the Dirac bispinor bundle) which is associated to a double cover (by spinor frames) of the bundle of space-time oriented orthonormal Lorentz frames. If the E-M gauge potential is included, then gauge invariance dictates that the Dirac bispinor bundle is associated with the fibered product of the spinor frame bundle and the Kaluza-Klein U(1) (circle) bundle.

Of course, electrons and photons are just part of the total picture. In place of the U(1) circle bundle used in the original Kaluza-Klein theory to introduce E-M, one uses a principal bundles P with a larger Lie group to incorporate other non-gravitational forces. The wave functions of the fundamental particles of matter are sections of various vector bundles that are associated (via group representations) to the so-called fibered product of the bundle P with the bundle of spinor frames over space-time. The known fundamental particles of matter include the 6 "flavors" of quarks (u (up), d (down), s (strange), c (charm), t (top), b (bottom)) each in three "colors" (R (red), G (green), B (blue)) together with the leptons (e^- (electron), μ^- (muon), τ^- (tau)) together with their associated neutrinos (ν_e (electron neutrino), ν_μ (muon neutrino), τ^- (tau neutrino)). These are organized into three generations (see the table below) which are essentially identical except in the masses of the corresponding particles, higher generation particles generally being heavier than their lower generation counterparts; e.g., the masses of e^-, μ^-, τ^- are approximately .511, 105.66, and 1784 MeV, respectively.

Generation		1	2	3
quarks	charge $+\frac{2}{3}$	$u_{(R,G,B)}$	$c_{(R,G,B)}$	$t_{(R,G,B)}$
	charge $-\frac{2}{3}$	$d_{(R,G,B)}$	$s_{(R,G,B)}$	$b_{(R,G,B)}$
leptons	charge -1	e^-	μ^-	τ^-
	charge 0	ν_e	ν_μ	ν_τ

We mention that for each quark $q_{(R,G,B)}$, there is an anti-quark $\bar{q}_{(C,M,Y)}$ of the opposite complementary color $(C, M, Y) =$ (cyan, magenta, yellow). In addition to gravity, the known forces are as follows. There is the strong force of QCD (quantum chromodynamics) acting between the quarks inside strongly interacting particles (hadrons) such as neutrons, protons and pions, which is mediated by 8 gluons (one for each vector in a basis for $\mathfrak{su}(3)$:=Lie algebra of SU(3)). The colors of the quarks can be regarded as charges that respond to the strong force in the sense that electrically charged particles respond to E-M fields. The fact that quarks are confined within hadrons seems to be related to the fact that SU(3) is nonabelian, which causes gluons to interact with each other. Unlike the coulomb force, gluon forces between quarks weaken as the separation distance decreases to zero, but gluon forces strengthen dramatically when the distance increases to the diameter of a hadron. Individual leptons (e.g., electrons, neutrinos) appear in the open (unconfined) since, being colorless, they are unaffected by gluons. In mathematical terms, SU(3) acts trivially on the lepton sector of the relevant representation. Corresponding to 4 generators of the Lie algebra of SU(2)× U(1), there is the weak force mediated by the Z and $\pm W$ vector bosons and the E-M force due to photons. Incidentally, U (1) of E-M is not simply the $U\,(1)$ factor in SU(2)× U(1). In terms of the Pauli matrices σ_k of (15.70), a set of standard generators for the complexified Lie algebra $\mathfrak{su}(2) \otimes \mathbb{C}$ are the matrices $\frac{1}{2}\sigma_3$ and $\sigma^{\pm} := \frac{1}{2}\,(\sigma_1 \pm i\sigma_2)$. The generator for electric charge in the Lie algebra $\mathfrak{su}(2) \oplus \mathfrak{u}\,(1)$ is a linear combination of $\frac{i}{2}\sigma_3 \in \mathfrak{su}(2)$ and the generator $i \in \mathfrak{u}\,(1)$, while the Z boson is associated with an independent linear combination of these generators. The $\pm W$ bosons are associated with $\sigma^{\pm}$. The weak force acts on leptons as well as quarks. Among other things, it is responsible for the decay of an isolated neutron into a proton, electron and antineutrino in about 15 minutes on average. It was primarily for their work in exhibiting that the electro-weak unification was feasible in the context of a spontaneously broken SU(2)× U(1) gauge theory (the GWS-theory) that Sheldon Glashow, Steven Weinberg and Abdus Salam were awarded the 1979 Nobel Prize for Physics. The Z and $\pm W$ vector bosons were detected by experimentalists in 1983. Unlike true gauge bosons such as the photon and gluons, the Z and $\pm W$ are massive, due to the fact that the $\mathrm{SU}(2) \times \mathrm{U}(1)$ gauge symmetry is broken, leaving the U(1) of E-M as the surviving gauge group. The standard explanation for how the symmetry was broken is known as the "Higgs mechanism," but the Higgs particles in the theory have yet to be found. There is another peculiarity of the weak force in that it acts only the so-called left halves of quarks and leptons. To understand this a little better, recall that a bispinor field for a particle is locally $\mathbb{C}^4$-valued. Two of these components correspond to the particle and two correspond to the antiparticle. Of the two components for the particle, one is the left-handed component and one is the right-handed component, and these are called the *chiral halves* of the particle; the antiparticle part also has chiral halves. Mathematically, the weak force associated with "$\mathfrak{su}(2)$-like" broken generators in the complexified $\mathfrak{su}(2) \oplus \mathfrak{u}(1)$ act in the usual way (as SU(2) acts on $\mathbb{C}^2$) on certain chiral doublets, such as $\left(e_L^-, v_{eL}\right)$ and (u_L, d_L'). In a less complicated world, d_L' might simply be the left-handed chiral half d_L of the down quark, but in *this* world, d_L' is a linear combination of d_L, s_L, and b_L, where the coefficients form the first row of the so-called Cabibbo-Kobayashi-Maskawa 3×3 matrix. The other rows of the CKM matrix are determined by the weak SU(2) doublets (c_L, s_L') and (t_L, b_L'), when s_L' and b_L' are written as linear combinations of d_L, s_L, and b_L. Incidentally, in a less complicated world where $d_L' = d_L$, etc. or where there are fewer than 3 generations, we might not exist. Indeed, then certain time-asymmetric weak reactions (e.g., K^0 meson decay) would not occur (see [**?**, p. 725]. It has been speculated that in the absence of such reactions, certain quark-nonconservation processes in grand unified theories might cause so much matter and anti-matter annihilation that there would be too few quarks left to make enough nucleons (see [**CD**, p. 176-7]). At any rate, to account for all of these known nongravitational forces, the gauge group of the principal bundle, say P, over space time M (before symmetry-breaking) must include $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$, a far cry from the U(1) group for the circle bundle of the original Kaluza-Klein theory. As a point of historical interest, Oscar Klein was the first to introduce SU(2) gauge fields (commonly known as Yang-Mills fields) and he even anticipated their use in modeling weak interactions in [**Kle38**], 30 years before the GWS model was developed.

In the search for symmetry, many have attempted to incorporate the group $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$ into a larger simple group (e.g., SU(5), SO(10), etc.), thereby obtaining a *grand unified theory* (GUT). The SO(10) (or more precisely, Spin(10)) GUT is particularly tidy, since the two fundamental spinor representation of Spin(10) are 16-dimensional and the left chiral halves of an individual generation of fundamental particles of matter fits perfectly (and correctly) into one of these 16-dimensional representations, while the right chiral halves fit into the other. For the first generation of (say, left) chiral halves, 12 of the 16 dimensions are accounted for by the 4 left chiral halves of the up and down quarks and antiquarks replicated in three colors, 2 dimensions come from the left chiral halves for the electron and positron, and the remaining 2 dimensions are occupied by the left chiral halves for the electron neutrino and antineutrino. Incidentally, for years many believed that neutrinos were massless and not bispinorial (i.e., having just 2-component wave functions), but recent experiments strongly suggest (if not prove) that neutrinos have a small mass which necessitates the existence of left-handed and right-handed neutrinos and antineutrinos, instead of just left-handed neutrinos and right-handed antineutrinos. It should be emphasized that not only do all of the left chiral halves of the fundamental particles fit by virtue of dimension count into a fundamental representation of Spin (10), but under the usual inclusions

$$\mathfrak{su}(3)\times\mathfrak{su}(2)\times\mathfrak{u}(1)\subset\mathfrak{su}(5)\subset\mathfrak{so}(10)\cong\mathfrak{spin}(10) \tag{15.75}$$

of Lie algebras the left chiral halves all respond correctly to the various forces under the spinor representation. In the Lie algebras of possible grand unification groups, there are generators which are not in $\mathfrak{su}(3)\times\mathfrak{su}(2)\times\mathfrak{u}(1)$ and hence do not correspond to standard known forces. The forces associated with these generators are thought to be involved in processes that convert quarks to leptons, which for example can lead to proton decay, a process which has yet to be detected. The force of gravity is not encoded in a grand unification group, but rather it is "gauged" in a different sense by the Lorentz group (or its cover $\mathrm{SL}(2,\mathbb{C})$) for the bundle of frames (or spinor frames) for space-time itself. In this way, gravity seems to resist attempts to unify it with other forces, and no universally convincing method of quantizing it has been forthcoming. The best hope seems to reside in string theory.

It must be stressed that unlike the Schrödinger wave function ψ_s which specifies the quantum state of a single particle in nonrelativistic quantum mechanics and whose modulus square $|\psi_s|^2$ is the position probability density, the Dirac bispinor wave functions ψ (or sections) for leptons and quarks, do not admit such an easy interpretation. One source of confusion is that these wave functions are not regarded as quantum fields even though in a certain nonrelativistic limit, two of the components of the Dirac wave function ψ can be identified with the two components of the Pauli wave function that satisfies the so-called Pauli equation which is a Schrödinger equation for single particles with spin $\frac{1}{2}$. In view of this, Dirac bispinor wave functions ψ are referred to as "first quantized" wave functions, while the process that converts such ψ to operator-valued distributions is known as "second quantization." However, as some correctly point out, an object should only be quantized once. Thus, one should regard Dirac bispinor wave functions ψ as classical states which have yet to be quantized. However, there are difficulties with interpreting such ψ as classical states. For example, there are two pointwise scalar products for Dirac bispinors ψ, one is simply $\psi^*\psi$, where ψ^* is conjugate transpose of the $\mathbb{C}^4$-valued, while the other is $\bar{\psi}\psi := \psi^*\gamma^0\psi$ where $\bar{\psi} := \psi^*\gamma^0$ is the so-called dual bispinor. Since $\psi^*\psi \geq 0$ and its integral over $\mathbb{R}^3$ is time-independent (as a consequence of Dirac's equation $\gamma^\mu\partial_\mu\psi = mc\psi$), one might be tempted to think of it as a probability density. However, when ψ is quantized (i.e., turned into a suitable operator-valued distribution), the expectation values of $\psi^*\psi$ have the interpretation as the charge density of a collection of positively and negatively charged particles and antiparticles, and so the positivity of $\psi^*\psi$ does not survive quantization. Moreover, the prequantum indefinite scalar product $\bar{\psi}\psi$ has positive expectation values after quantization and is interpreted as an energy operator. Generally, many classical fields cannot be given a reasonable physical significance until they are quantized. For example, although the forces of QCD are mediated by mass 0 gluons (corresponding to photons in QED) and thus might be expected to have a long range (decaying as r^{-2}, rather than exponentially), no unconfined, long range effects of gluons are evident, unlike the case of E-M fields of photons. Since gluons seem to be confined to very small regions inside hadrons, it would appear very speculative to treat them as classical wave-like fields. By the photoelectric effect, we know that E-M does not behave much like a wave even at the vastly greater dimensions of an atom. It is nevertheless believed that the classical solutions of the field equations for nonabelian gauge fields (particularly, in 4-dimensional *Euclidean* space, with positive-definite metric) do yield at least a first-order approximation to certain quantum effects for such fields, especially with regard to "tunneling phenomena." Although we cannot go into the details of how this works in quantum field theory, there is a similar situation in quantum mechanics, which can be understood through the following discussion.

Consider the Schrödinger operator $H := H_0 + V$ where $H_0 := -\frac{\hbar^2}{2m}\Delta$ and $V : \mathbb{R}^3 \to \mathbb{R}$ is a suitable potential (e.g., $V_- := \min(V,0) \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ and $V_+ := \max(V,0) \in L^2_{\mathrm{loc}}(\mathbb{R}^3)$). Then $H_0 + V$ is an unbounded essentially self-adjoint on the dense domain $C_0^\infty(\mathbb{R}^3)$ of $L^2(\mathbb{R}^3)$, see [**ReSi75**, p. 185]. One then obtains a strongly continuous 1-parameter group of unitary transformations $\exp(-itH/\hbar)$ of $L^2(\mathbb{R}^3)$, see [**ReSi72**, p. 265]. For f in the Schwartz space of rapidly decreasing functions on $\mathbb{R}^3$, the solution of the problem

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi + V\psi, \quad \psi(x,0) = f(x) \tag{15.76}$$

is given by $\psi(x,t) = [\exp(-itH/\hbar) f](x)$. Even though H_0 and V do not commute in general, there is a formula due to T. Kato and H. F. Trotter (see [**Kato**] and [**Tro**]) that yields

$$\exp\left(-\frac{i}{\hbar}tH\right) = \lim_{k\to\infty}\left(\left[\exp\left(\frac{-it}{\hbar k}V\right)\exp\left(\frac{-it}{\hbar k}H_0\right)\right]^k\right). \tag{15.77}$$

It is well-known (and not hard to prove) that

$$\exp\left(-\frac{i}{\hbar}tH_0\right)[f](x) = \left(\frac{2\pi i\hbar t}{m}\right)^{-3/2}\int_{\mathbb{R}^3}\exp\left(\frac{im\,|x-y|^2}{2\hbar t}\right)f(y)\;d^3y, \tag{15.78}$$

whence

$$\begin{aligned}&\exp\left(\frac{-it}{\hbar k}V\right)\exp\left(\frac{-it}{\hbar k}H_0\right)[f](x)\\ &= \left[\frac{2\pi i\hbar t}{mk}\right]^{-3/2}\int_{\mathbb{R}^3} e^{\frac{it}{\hbar k}\left(\frac{m}{2}\frac{|x-x_0|^2}{(t/k)^2}-V(x)\right)}f(x_0)\;d^3x_0.\end{aligned} \tag{15.79}$$

For $x_0, x_1, \ldots, x_k \in \mathbb{R}^3$, let

$$A_t(x_0,x_1,\ldots,x_k) := \sum_{j=1}^{k}\left(\frac{m}{2}\frac{|x_j-x_{j-1}|^2}{(t/k)^2} - V(x_j)\right)\frac{t}{k}. \tag{15.80}$$

Taking $x_k = x$, the Kato-Trotter formula then yields

$$\begin{aligned}\psi(x,t) &= [\exp(-itH/\hbar) f](x) \\ &= \lim_{k\to\infty}\left[\exp\left(\frac{-it}{\hbar k}V\right)\exp\left(\frac{-it}{\hbar k}H_0\right) f\right]^k (x) \\ &= \lim_{k\to\infty}\left[\frac{2\pi i\hbar t}{mk}\right]^{-3k/2}\int_{\mathbb{R}^{3k}} e^{\frac{i}{\hbar}A_t(x_0,\ldots,x_k)} f(x_0)\, d^3x_0\cdots d^3x_{k-1}. \qquad (15.81)\end{aligned}$$

Let $\gamma : [0,t] \to \mathbb{R}^3$ be a path with $\gamma(jt/k) = x_j$ for $j = 0,\ldots,k-1$ and

$$\gamma(s) = (x_{j+1} - x_j)\left(\frac{k}{t}s - j\right) + x_j \quad \text{for} \quad s \in [jt/k,(j+1)t/k]. \qquad (15.82)$$

Then the "classical action" for the path γ is

$$\begin{aligned}A(\gamma) &:= \int_0^t \frac{m}{2}\left|\gamma'(s)\right|^2 - V(\gamma(s))\, ds \\ &= \sum_{j=0}^{k-1}\int_{jt/k}^{(j+1)t/k} \frac{m}{2}\left(\frac{|x_j - x_{j-1}|}{t/k}\right)^2 - V(\gamma(s))\, ds \\ &\approx \sum_{j=1}^{k}\left(\frac{m}{2}\left(\frac{|x_j - x_{j-1}|}{t/k}\right)^2 - V(x_j)\right)\frac{t}{k} \\ &= A_t(x_0,x_1,\ldots,x_k). \qquad (15.83)\end{aligned}$$

Now, integrating with respect to $x_1,\ldots,x_{k-1}$ is like integrating over all polygonal paths have $k-1$ segments, starting at x_0 at time 0 and ending at some arbitrary point x at time t. As $k \to \infty$, the variety of such paths is sufficiently great so that we (as R. P. Feynman) are tempted to write

$$\psi(x,t) = \int_{\mathbb{R}^3}\left(\int_{P_t(x_0,x)} e^{\frac{i}{\hbar}A(\gamma)} d\gamma\right) f(x_0)\, d^3x_0, \qquad (15.84)$$

where $P_t(x_0,x)$ is the space of continuous paths $\gamma : [0,t] \to \mathbb{R}^3$ with $\gamma(0) = x_0$ and $\gamma(t) = x$, $A(\gamma)$ is the action defined in (**??**), and $d\gamma$ is some kind of "measure" on $P_t(x_0,x)$. In other words, the kernel for the operator $\exp\left(-\frac{i}{\hbar}tH\right)$ is formally

$$K(x_0,x,t) = \int_{P_t(x_0,x)} e^{\frac{i}{\hbar}A(\gamma)} d\gamma. \qquad (15.85)$$

The integrand $e^{\frac{i}{\hbar}A_t(\gamma)}$ oscillates the least about paths for which $A(\gamma)$ is stationary among paths in $P_t(x_0,x)$, namely classical paths that are solutions of Newton's equation $m\gamma'' = -\nabla V$. Hence we expect $K(x_0,x,t)$ to be most greatly influenced by classical paths (typically only one) with $\gamma(0) = x_0$ and $\gamma(t) = x$. As $\hbar \to 0$, this effect becomes more pronounced, and presumably we obtain classical mechanics in the limit. Note that formally

$$|K(x_0,x,t)|^2 = \left|\left\langle \exp\left(-\frac{i}{\hbar}tH\right)\delta(x_0), \delta(x)\right\rangle\right|^2 \qquad (15.86)$$

is the probability density that at time t the particle will be found at x, given that it was at x_0 at time t. While the path integral is a suggestive formalism for the rigorous Kato-Trotter limit, M. Kac (see [**Kac**]) noticed that if one replaces the time variable t by a pure imaginary parameter $-i\tau$, then we obtain

$$\frac{i}{\hbar}A_{-i\tau}(x_0,x_1,\dots,x_k) = -\frac{1}{\hbar}\sum_{j=1}^{k}\left(\frac{m}{2}\frac{|x_j-x_{j-1}|^2}{(\tau/k)^2}+V(x_j)\right)\frac{\tau}{k}$$

$$= -\frac{1}{\hbar}\int_0^\tau \frac{m}{2}|\gamma'(s)|^2 + V(\gamma(s))\,ds =: -\frac{1}{\hbar}A_E(\gamma), \tag{15.87}$$

where $\gamma : [0,\tau] \to \mathbb{R}^3$ is a polygonal path through $x_0, x_1, \dots, x_k$ and $A_E(\gamma)$ is the so-called "Euclidean action." Then M. Kac was able to express $K(x_0,x,-i\tau)$ rigorously as a path integral in terms of conditional Wiener measure $W^\tau_{x_0,x}$ on the set $P_\tau(x_0,x)$ of continuous paths $\gamma:[0,\tau]\to\mathbb{R}^3$ with $\gamma(0)=x_0$ and $\gamma(\tau)=x$. The Feynman-Kac Formula (see [**GJ**, Theorem 3.2.3]) is then

$$K(x_0,x,-i\tau) = \int \exp\left(-\frac{1}{\hbar}\int_{-\tau/2}^{\tau/2} V(\gamma(s))\,ds\right)dW^\tau_{x_0,x}(\gamma). \tag{15.88}$$

Note that the kinetic part $\int_0^\tau \frac{m}{2}|\gamma'(s)|^2\,ds$ of the Euclidean action $A_E(\gamma)$ has been absorbed into the measure $W^\tau_{x_0,x}$, and (since it is convenient for some purposes) the paths γ are reparametrized symmetrically using $[-\tau/2,\tau/2]$ instead of $[0,\tau]$. It is also of interest that the subset of paths γ with Hölder exponent larger than $\frac{1}{2}$ (which includes the piecewise C^1 paths) has Wiener measure 0 and hence this subset does not contribute to the integral. However, one still expects that the greatest contributions to $K(x_0,x,-i\tau)$ come from fluctuations about a path which minimizes the Euclidean action $A_E(\gamma)$. Such paths are solutions of the Euler-Lagrange equation (with conditions $\gamma(0)=x_0$ and $\gamma(\tau)=x$)

$$m\gamma''(s) = \nabla V(\gamma(s)) \tag{15.89}$$

which differs from Newton's equation $m\gamma''(s) = -\nabla V(\gamma(s))$ by a minus sign. We now interpret $K(x_0,x,-i\tau)$, at least formally. For simplicity, suppose that $H = -\frac{\hbar^2}{2m}\Delta + V$ where $V(\mathbf{r})$ increases rapidly enough as $\|\mathbf{r}\|\to\infty$ so that there is a complete orthonormal set of eigenfunctions $u_0,u_1,u_2,\dots$ of H with eigenvalues (energies) arranged in increasing order, say $E_0 \le E_1 \le E_2 \le \dots$ (degeneracy allowed). The kernel of $\exp\left(-\frac{i}{\hbar}tH\right)$ is given by

$$K(x_0,x,t) = \sum_{n=0}^{\infty} e^{-\frac{i}{\hbar}tE_n}u_n(x)\overline{u_n(x_0)}. \tag{15.90}$$

Indeed,

$$\psi(x,t) := \int_{\mathbb{R}^3}\left(\sum_{n=0}^{\infty} e^{-\frac{i}{\hbar}tE_n}u_n(x)\overline{u_n(x_0)}\right)f(x_0)\,d^3x_0 \tag{15.91}$$

solves Schrödinger's equation at least formally and $\psi(x,0)$ is the eigenfunction expansion of $f(x)$. Replacing t by $-i\tau$, we obtain

$$K(x_0,x,-i\tau) = \sum_{n=0}^{\infty} e^{-\frac{\tau}{\hbar}E_n}u_n(x)\overline{u_n(x_0)}. \tag{15.92}$$

Letting $\tau \to \infty$, we formally obtain (for some $k > 0$)

$$e^{\frac{\tau}{\hbar}E_0} K\left(x_0, x, -i\tau\right) = \sum_{E_n = E_0} u_n\left(x\right) \overline{u_n\left(x_0\right)} + \mathrm{O}\left(e^{-kt}\right) \tag{15.93}$$

where the sum is only over lowest energy states. If x_0 is a minimum for V, then the constant path $\gamma : [0, \tau] \to \{x_0\}$ clearly minimizes the Euclidean action $\int_0^\tau \frac{m}{2}\left|\gamma'\left(s\right)\right|^2 + V\left(\gamma\left(s\right)\right) ds$ for loops at x_0. Hence, such paths are likely to make $K\left(x_0, x_0, -i\tau\right) \approx e^{-\frac{\tau}{\hbar}E_0} \sum_{E_n=E_0}^{\infty} \left|u_n\left(x_0\right)\right|^2$ larger at minima for V than at other x_0. Indeed, we expect the position probability densities for ground energy states to be concentrated about the minima for V. Classical intuition leads us to suspect that if there are N absolute minima for V, then there ought to be N independent eigenfunctions, each peaked at a different minimum; i.e., that there is an N-fold degeneracy in the lowest energy level (i.e., $E_n = E_0$ for $n = 0, 1, \ldots, N-1$). However, there is a very general result stating that if V is continuous and bounded below and $H = -\frac{\hbar^2}{2m}\Delta + V$ is essentially self-adjoint, then the ground state is nondegenerate and is represented by a real, positive function. Indeed, there is an elegant proof of this in [**GJ**, Corollary 3.3.4] based in part on the Feynman-Kac Formula (15.88). We will examine a simple example in dimension 1 (i.e., $x \in \mathbb{R}$) to illustrate this and to introduce the concept of an instanton.

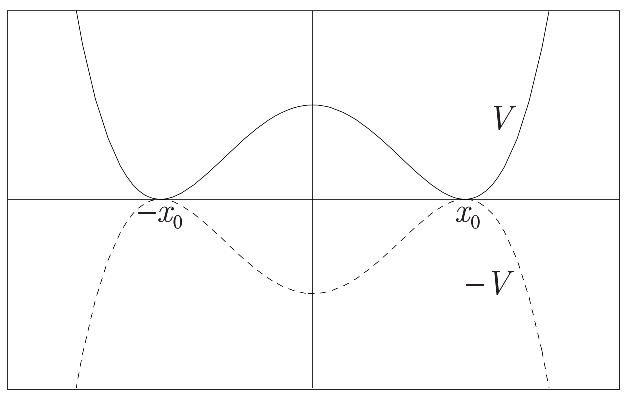


FIGURE 15.1. Quartic potential and its negative (dashed)

Let $V\left(x\right) = \frac{1}{2}\left(x^2 - x_0^2\right)^2$, as shown in Figure 15.1. Note that the Schrödinger operator $H = -\frac{\hbar^2}{2m}\Delta + V$ commutes with the parity operator P given by $\left(Pf\right)\left(x\right) := f\left(-x\right)$, and hence the eigenspaces of H split into even functions and odd functions ($+1$ and -1 eigenspaces of P). Classical intuition falsely suggests that there are two normalized E_0-energy eigenfunctions, say ψ peaked at x_0 and $P\psi$ peaked at $-x_0$ ($\left\|\psi\right\|^2 = \left\|P\psi\right\|^2 = 1$). Suppose that this is the case, and let $\psi_+ = \left(\psi + P\psi\right)/\sqrt{2}$ and $\psi_- = \left(\psi - P\psi\right)/\sqrt{2}$ be the associated even and odd states. Then for any $x \in \mathbb{R}$, as $\tau \to \infty$,

$$\begin{aligned} e^{\frac{\tau}{\hbar}E_0} K\left(x, -x, -i\tau\right) &\sim \psi_+\left(-x\right)\overline{\psi_+\left(x\right)} + \psi_-\left(-x\right)\overline{\psi_-\left(x\right)} \\ &= \left|\psi_+\left(x\right)\right|^2 - \left|\psi_-\left(x\right)\right|^2 . \end{aligned} \tag{15.94}$$

From (15.88), $K(x_0, x, -i\tau)$ is the path integral of a positive function, and hence $e^{\frac{\tau}{\hbar}E_0} K(x, -x, -i\tau) \geq 0$ so that

$$|\psi_+(x)|^2 - |\psi_-(x)|^2 \geq 0. \tag{15.95}$$

If this inequality were strict even at a single point, then we would get the contradiction

$$0 < \|\psi_+(x)\|^2 - \|\psi_-(x)\|^2 = 1 - 1 = 0. \tag{15.96}$$

We now argue (as physicists might) that the inequality (15.95) is strict near $x = x_0$. Consider the Euclidean potential $-V$ whose graph is shown dashed in Figure 15.1. There is a classical solution $\gamma_\infty : (-\infty, \infty) \to (-x_0, x_0)$ of the Euclidean equation of motion $m\gamma''(s) = \nabla V(\gamma(s))$ for a particle, with total energy $\frac{1}{2}m\gamma'_\infty(s)^2 - V(\gamma_\infty(s)) = 0$, that moves from the top of the left hill to the top of the right hill. Physicists call such trajectories (and their analogs in quantum field theory) "instantons." The Euclidean action $A_E(\gamma) = \int_0^\tau \frac{m}{2}|\gamma'(s)|^2 + V(\gamma(s))\, ds$ of the instanton is finite, since making the change of variable $x = \gamma_\infty(s)$, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} \tfrac{1}{2}m\gamma'_\infty(s)^2 + V(\gamma_\infty(s))\, ds &= \int_{-\infty}^{\infty} 2V(\gamma_\infty(s))\, ds \\
&= \int_{-x_0}^{x_0} 2V(x)\left(\tfrac{dx}{ds}\right)^{-1} dx = \int_{-x_0}^{x_0} 2V(x)\frac{1}{\sqrt{2V(x)/m}}dx \\
&= \int_{-x_0}^{x_0} \sqrt{2mV(x)}\, dx < \infty.
\end{aligned} \tag{15.97}$$

The instanton γ_∞ minimizes the action functional $A_E(\gamma)$ among suitable competing paths from $-x_0$ to $-x_0$. Let $\gamma_\tau := \gamma_\infty|_{[-\tau/2,\tau/2]}$. We have for all $\tau > 0$,

$$A_E(\gamma_\tau) = \int_{-\tau/2}^{\tau/2} 2V(\gamma_\tau(s))\, ds < \int_{-x_0}^{x_0} \sqrt{2mV(x)}\, dx < \infty. \tag{15.98}$$

Since the kinetic part of $A_E(\gamma_\tau)$ is implicit in $dW^\tau_{-x_0,x_0}$, by a formal application of Laplace's method, we expect that there is a finite, nonzero contribution to $K(-x_0, x_0, -i\tau)$ proportional to $\exp\left(-\frac{1}{\hbar}\int_{-\tau/2}^{\tau/2} A_E(\gamma_\tau)\, ds\right)$; i.e., for some constant $C > 0$,

$$\begin{aligned}
K(x_0, -x_0, -i\tau) &= \int \exp\left(-\frac{1}{\hbar}\int_{-\tau/2}^{\tau/2} V(\gamma(s))\, ds\right) dW^\tau_{x_0,-x_0}(\gamma) \\
&\geq C\exp\left(-\frac{1}{\hbar}\int_{-\tau/2}^{\tau/2} A_E(\gamma_\tau)\, ds\right) \geq C\exp\left(-\frac{1}{\hbar}\int_{-\infty}^{\infty} A_E(\gamma_\infty)\, ds\right) \\
&= C\exp\left(-\frac{1}{\hbar}\int_{-x_0}^{x_0} \sqrt{2mV(x)}\, dx\right) > 0.
\end{aligned} \tag{15.99}$$

Since $E_0 \geq 0$, as $\tau \to \infty$,

$$\begin{aligned}
|\psi_+(x_0)|^2 - |\psi_-(x_0)|^2 &\sim e^{\frac{\tau}{\hbar}E_0} K(-x_0, x_0, -i\tau) \\
&\geq C\exp\left(-\frac{1}{\hbar}A_E(\gamma_\infty)\right) \\
&\geq C\exp\left(-\frac{1}{\hbar}\int_{-x_0}^{x_0} \sqrt{2mV(x)}\, dx\right) > 0.
\end{aligned} \tag{15.100}$$

Hence, we arrive at the contradiction (15.96). In the above heuristic argument, the contradiction was produced by the instanton γ_∞, and hence physicists are led to attribute the nondegeneracy of the ground state to the presence of this instanton. From our previous discussion, we know that in fact there is a unique ground state represented by a positive wave function ψ_0. Since H commutes with the parity operator P, we know that $P\psi_0$ also is a positive representative of the ground state and hence $P\psi_0 = \psi_0$ (i.e., ψ_0 is even). The evenness of the ground state is also a consequence of (**??**). Regardless of whether ψ_0 is even or odd, we have that $|\psi_0|^2$ is even and hence a particle in the state ψ_0 has the same probability of appearing in an interval about x_0 as it does in the reflected interval about $-x_0$, even for measurements made in rapid succession. This is so even though the height of the potential barrier between $-x_0$ and x_0 forbids travel of a classical particle of energy E_0 between these two points. In other words, we have the phenomenon of "quantum tunneling." There is also a quantitative link between the instanton and quantum tunneling. It turns out that by WKB methods, the transmission amplitude for a particle to penetrate the potential barrier from x_0 to $-x_0$ is proportional to $\exp\left(-\frac{1}{\hbar}A_E(\gamma_\infty)\right)$ (see [**Kak**, p. 545-554]).

Much of this discussion on the role of instantons in removing degeneracy and tunneling caries over, at least metaphorically, to quantum field theory. Since we will be primarily concerned with nonabelian gauge fields (e.g., where the gauge group is SU(2), isomorphic to the unit quaternions S^3), we confine ourselves to a very brief account of what instantons are in context of the quantum field theory of pure Yang-Mills fields and how they correspond to the ones we have discussed in relation to quantum mechanics. The configuration space in the quantum mechanics of a single particle is simply $\mathbb{R}^3$, but for pure Yang-Mills fields, the configuration space is essentially the space $\Omega^1\left(\mathbb{R}^3, \mathfrak{su}(2)\right)$ of all smooth 1-forms on $\mathbb{R}^3$ (with values in the Lie algebra $\mathfrak{su}(2)$ of traceless skew-Hermitian matrices), modulo the action by the group of gauge transformations. Of course, any SU(2)-bundle over $\mathbb{R}^3$ is just a product, $\mathbb{R}^3 \times \mathrm{SU}(2)$. Hence, a gauge transformation amounts to a map $\phi : \mathbb{R}^3 \to \mathrm{SU}(2)$;i.e., $\phi \in C^\infty\left(\mathbb{R}^3, \mathrm{SU}(2)\right)$. The right action of ϕ on $A \in \Omega^1\left(\mathbb{R}^3, \mathfrak{su}(2)\right)$ is given by

$$A \cdot \phi := \phi^{-1}(A\phi + d\phi). \tag{15.101}$$

Note that

$$\begin{aligned} A \cdot (\phi\eta) &= (\phi\eta)^{-1}(A(\phi\eta) + d(\phi\eta)) = \eta^{-1}\phi^{-1}(A(\phi\eta) + (d\phi)\eta + \phi d\eta) \\ &= \eta^{-1}\left(\left(\phi^{-1}A\phi + \phi^{-1}d\phi\right)\eta + d\eta\right) \\ &= \eta^{-1}\left(\left(\phi^{-1}(A\phi + d\phi)\right)\eta + d\eta\right) \\ &= (A \cdot \phi) \cdot \eta. \end{aligned} \tag{15.102}$$

We require that $\phi(\mathbf{r}) \to \mathrm{Id} \in \mathrm{SU}(2)$ as $\|\mathbf{r}\| \to \infty$, and ϕ yields a map $\phi' : S^3 \to \mathrm{SU}(2) \cong S^3$ which is classified up to homotopy by its degree. A gauge transformation ϕ is called homotopically trivial if ϕ' has degree 0. Moreover, it is required of $A \in \Omega^1\left(\mathbb{R}^3, \mathfrak{su}(2)\right)$ that the field strength

$$F_A := dA + A \wedge A \in \Omega^2\left(\mathbb{R}^3, \mathfrak{su}(2)\right) \tag{15.103}$$

be square integrable; i.e., $\int_{\mathbb{R}^3} |F_A|^2 < \infty$, where $|b|^2 := \frac{1}{2} Tr\,(b^* b)$ for $b \in \mathfrak{su}\,(2)$. More precisely, for

$$\Omega^1\left(\mathbb{R}^3, \mathfrak{su}\,(2)\right)_{\text{finite}} := \left\{ A \in \Omega^1\left(\mathbb{R}^3, \mathfrak{su}\,(2)\right) : \int_{\mathbb{R}^3} |F_A|^2 < \infty \right\} \tag{15.104}$$

the configuration space is

$$\mathcal{C} := \frac{\Omega^1\left(\mathbb{R}^3, \mathfrak{su}\,(2)\right)_{\text{finite}}}{C^\infty\left(\mathbb{R}^3, \mathrm{SU}(2)\right)_0} \tag{15.105}$$

where $C^\infty\left(\mathbb{R}^3, \mathrm{SU}(2)\right)_0$ is the group of homotopically trivial gauge transformations. The nonlinear "Yang-Mills" functional $\mathcal{V} : \mathcal{C} \to [0, \infty)$, defined by

$$\mathcal{V}\left([A]\right) := \tfrac{1}{2} \int_{\mathbb{R}^3} |F_A|^2, \tag{15.106}$$

plays the role of the potential energy function V in quantum mechanical setting of a single particle. Note that $\mathcal{V}$ is well defined, since one can verify that $F_{A\cdot\phi} = \phi^{-1} F_A \phi$ and $|F_{A\cdot\phi}|^2 = \left|\phi^{-1} F_A \phi\right|^2 = |F_A|^2$ using the fact that for $b \in \mathfrak{su}\,(2)$ and $C \in \mathrm{SU}(2)$

$$\begin{aligned} \left|C^{-1} b C\right|^2 &= \tfrac{1}{2} Tr\left(\left(C^{-1} b C\right)^* \left(C^{-1} b C\right)\right) \\ &= \tfrac{1}{2} Tr\left(C^{-1} b^* \left(C C^{-1}\right) b C\right) = \tfrac{1}{2} Tr\left(C^{-1} b^* b C\right) = |b|^2 . \end{aligned} \tag{15.107}$$

Suppose that $\phi_i : \mathbb{R}^3 \to \mathrm{SU}\,(2)$ $(i = 1, 2)$ are inequivalent (i.e., $\deg\left(\phi_1'\right) \neq \deg\left(\phi_2'\right)$). For $A_i := \phi_i^{-1} d\phi_i \in \Omega^1\left(\mathbb{R}^3, \mathfrak{su}\,(2)\right)$, we have $A_i = 0 \cdot \phi_i$ and so $F_{A_i} = F_0 = 0$. There is no ϕ with $\deg\left(\phi'\right) = 0$ such that $A_1 \cdot \phi = A_2$. Indeed,

$$\begin{aligned} A_1 \cdot \phi = A_2 &\Rightarrow \phi_1^{-1} d\phi_1 \cdot \phi = \phi_2^{-1} d\phi_2 \Rightarrow \left(0 \cdot \phi_1\right) \cdot \phi = 0 \cdot \phi_2 \\ &\Rightarrow 0 \cdot \left(\phi_1 \phi \phi_2^{-1}\right) = 0 \Rightarrow d\left(\phi_1 \phi \phi_2^{-1}\right) = 0 \\ &\Rightarrow \phi_1 \phi \phi_2^{-1} = \mathrm{Id} \Rightarrow \phi_1 \phi = \phi_2 \Rightarrow \deg\left(\phi_1'\right) = \deg\left(\phi_2'\right). \end{aligned} \tag{15.108}$$

Thus, $[A_1]$ and $[A_2]$ are distinct, absolute minima for the "potential" $\mathcal{V}$, and there are infinitely many distinct minima of the form $\left[\phi^{-1} d\phi\right]$, one for each possible value of $\deg\left(\phi'\right)$. In analogy with the single particle setting, an instanton is a certain curve connecting minimum $[A_1]$ to $[A_2]$ in the configuration space $\mathcal{C}$ parametrized by $\tau \in (-\infty, \infty)$. Such a curve can be regarded as the class of point, say $A \in \Omega^1\left(\mathbb{R}^4, \mathfrak{su}\,(2)\right)$ where we mod out by suitable gauge transformations in $C^\infty\left(\mathbb{R}^4, \mathrm{SU}(2)\right)_0$. Again by analogy, we want to minimize the Euclidean action of this curve among competitors running from $[A_1]$ to $[A_2]$. This Euclidean action is naturally taken to be $\int_{\mathbb{R}^4} |F_A|^2$, where $F_A := dA + A \wedge A$ as before, and $|F_A|^2$ is computed using the Euclidean metric on $\mathbb{R}^4$, as opposed to the Minkowski metric. Note that this is analogous to replacing t by $-i\tau$ in the single particle setting. One major goal of the following chapters is to use the Atiyah-Singer Index Theorem to prove that the set of instantons connecting $[A_1] = \left[\phi_1^{-1} d\phi_1\right]$ to $[A_2] = \left[\phi_2^{-1} d\phi_2\right]$ form an $(8k - 3)$-dimensional manifold, where $k = |\deg \phi_1 - \deg \phi_2|$. It is important to know this dimension in order to estimate its full effect with regard to vacuum tunneling between $[A_1]$ and $[A_2]$. The interested reader may find some insight into this very tricky business in [**Ber**] and [**CG**]. It should also be noted that the Lagrangians (actions) of other fields (e.g., Dirac bispinor fields for fundamental particles) must be added to the self-actions of pure gauge potentials. The Index Theorem is also essential for estimating the effects of these other fields (e.g., Euclidean fermionic lowest energy modes) on Green's functions of quantum field theory (see [**Schw79**]).

CHAPTER 16

Geometric Preliminaries

Here we provide fundamental definitions and results concerning the geometry and topology of fiber bundles that is essential to understanding gauge theories. In most journal articles on the subject, it is assumed that the reader knows this material or can dig it out from various sources. To cut down on the frustration, we develop the following topics, assuming a more modest background:

16.1. Principal G-bundles
16.2. Connections and Curvature
16.3. Equivariant Forms and Associated Bundles
16.4. Gauge Transformations
16.5. Curvature in Riemannian Geometry
16.6. Bochner-Weitzenböck Formulas
16.7. Characteristic Classes and Curvature Forms
16.8. Holonomy

1. Principal G-Bundles

A *Lie group* is simply a group which is a smooth (C^∞) manifold for which the map $(g_1, g_2) \mapsto g_1 g_2^{-1}$ is a C^∞ map from $G \times G$ to G. Let P be a manifold on which a Lie group G acts freely and smoothly on the right. Thus, there is a smooth map $P \times G \to P$ which we denote by $(p, g) \mapsto pg$ such that $(pg_1)\, g_2 = pg_1 g_2$, and if $pg = p$ for some $(p, g) \in P \times G$, then g is the group identity, say e. For our purposes, one may assume that G is a matrix group, such as $\mathrm{SU}(N)$, $\mathrm{SO}(N)$, etc.. We assume that the quotient space $M := P/G$ can be made into a manifold such that the projection $\pi : P \to M$ is smooth. Also, we assume that P is locally trivial. This means that each $x \in M$ has a neighborhood U, such that there is a diffeomorphism $T : \pi^{-1}(U) \to U \times G$ with $T(pg) = (\pi(p), s(p))$ where $s : \pi^{-1}(U) \to G$ satisfies $s(pg) = s(p)\, g$ for all $p \in \pi^{-1}(U)$ and $g \in G$. In other words, P is locally equivalent to a product with the standard action.

Definition 16.1. If the conditions of the preceding paragraph hold, then we say that $\pi : P \to M$ is a **principal G-bundle** with **total space** P, **base space** M, and **projection** π.

The diffeomorphism $T : \pi^{-1}(U) \to U \times G$ is known as a *local trivialization.* In the case where U can be taken to be all of M (i.e., $T : P = \pi^{-1}(M) \to M \times G$), we say that principal G-bundle $\pi : P \to M$ is (globally) *trivial.* Let U be an open subset of M, and let $\sigma : U \to P$ be a map such that $\pi \circ \sigma = 0$. Then σ is called a *local section.* There is a one-to-one correspondence between local sections and local trivializations. Indeed, given σ, define $T : \pi^{-1}(U) \to U \times G$ by $T(\sigma(x)\, g) = (x, g)$.

Note that T is well-defined since any $p \in \pi^{-1}(U)$ can be written uniquely as $\sigma(x)\, g$ since G acts freely and transitively on the fiber $\pi^{-1}(\pi(p))$. Conversely, given a local trivialization T, the equation $T(\sigma(x)\, g) = (x, g)$ serves to define σ (i.e., $\sigma(x) = T^{-1}(x, g)\, g^{-1} = T^{-1}(x, 1)$). It follows that a principal G-bundle is trivial precisely when it has a *global* section (i.e., local section $\sigma : U \to P$ with $U = M$). The corresponding statement for sphere bundles is false. For example, the unit tangent bundle $S(K)$ of a Klein bottle K is a circle bundle which is not globally a product, since otherwise one could define a frame field on K even though K is nonorientable. Nevertheless, K does have a unit tangent vector field. This also shows that $S(K)$ cannot be made into a principal S^1 bundle, where S^1 is regarded as the group $U(1) := \{e^{i\theta} : \theta \in \mathbb{R}\}$. Indeed, an orientation is precisely what is necessary in order to define a free S^1 action on the unit tangent bundle $S(M)$ of a surface M (with metric). Thus, for *orientable* surfaces M, $S(M)$ is trivial exactly when there is a unit tangent vector field.

Exercise 16.2. Recall that

$$\begin{aligned} \mathrm{SU}(2) &:= \{A \in \mathrm{GL}(2, \mathbb{C}) : A^* A = I_2,\ \det A = 1\} \\ &= \left\{ \begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix} : a, b \in \mathbb{C},\ |a|^2 + |b|^2 = 1 \right\} \cong S^3. \end{aligned}$$

Let $G \cong S^1$ be the subgroup of elements with $b = 0$ (i.e., $a = e^{i\theta}$). Show that the left coset space $\mathrm{SU}(2)/G = \{AS^1 : A \in \mathrm{SU}(2)\}$ can be identified with S^2 so that the quotient map $Q : \mathrm{SU}(2) \to \mathrm{SU}(2)/G$ may be regarded as a principal S^1-bundle $\pi : \mathrm{SU}(2) \to S^2$ (which is known as a **Hopf bundle**). Show that this bundle is nontrivial. [Hint. Identify $\mathbb{R}^3$ with the space of traceless Hermitian matrices via the "Pauli map"

$$\sigma : \mathbf{r} = (x, y, z) \mapsto \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix}.$$

For $A \in \mathrm{SU}(2)$, show that $c(A)(\mathbf{r}) := \sigma^{-1}(A\sigma(\mathbf{r})\, A^*)$ defines a homomorphism $c : \mathrm{SU}(2) \to \mathrm{SO}(3)$. Let $\pi : \mathrm{SU}(2) \to S^2$ be given by $\pi(A) := c(A)\, \mathbf{e}_3$ where $\mathbf{e}_3 := (0, 0, 1)$ and check that $\pi(Ae^{i\theta}) = \pi(A)$.]

2. Connections and Curvature

Connections (or gauge fields) on principal G-bundles can be defined in various equivalent ways. From a conceptual standpoint, the following definition is perhaps best:

Definition 16.3. A **connection** on a principal G-bundle $\pi : P \to M$ smoothly assigns to each $p \in P$ a subspace H_p of the tangent space $T_p P$, such that $\pi_{*p} : H_p \to T_{\pi(p)} M$ is an isomorphism and $R_{g*}(H_p) = H_{pg}$, where $R_g : P \to P$ is defined by $R_g(p) = pg$.

The so-called *horizontal subspace* H_p is complementary to the *vertical subspace* $V_p := \mathrm{Ker}(\pi_{*p})$ which is the tangent space of the fiber $\pi^{-1}(\pi(p))$ at p. Thus, a connection serves to select a "horizontal" complement to each vertical subspace in a smooth G-invariant fashion, see Figure 16.1. One way of defining H_p would be

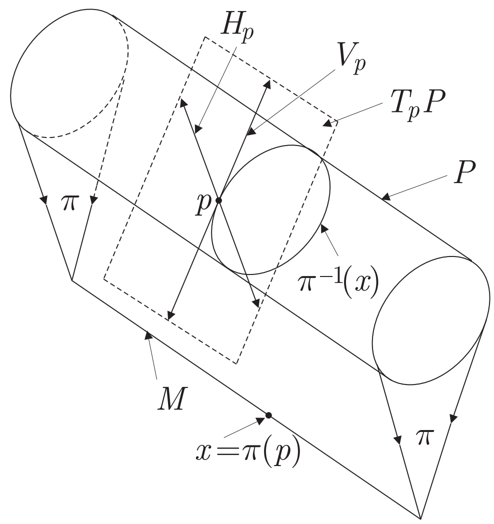


FIGURE 16.1. Selecting a *horizontal* complement H_p to each *vertical* subspace V_p of the tangent space T_pP at $p \in P$

to let it be the subspaces of T_pP annihilated by a differential 1-form of maximal rank with values in a fixed vector space having the same dimension as V_p, namely $\dim P - \dim M = \dim G$. The natural choice for this vector space is the Lie algebra of G which we denote by $\mathfrak{g}$. The Lie algebra of G is the tangent space T_eG of G at the identity $e \in G$, and for $A, B \in \mathfrak{g}$, there is a Lie bracket $[A, B] \in \mathfrak{g}$. While we will not go into the definition of $[A, B]$ for general Lie groups, in the case matrix Lie groups $G \subseteq \mathrm{GL}\,(N, \mathbb{C})$ it is easy to describe. Indeed, as $\mathrm{GL}\,(N, \mathbb{C})$ is an open subset of the linear space $\mathfrak{gl}\,(N, \mathbb{C})$ of all $N \times N$ complex matrices, the tangent space of T_IG at the identity matrix I may be identified with a subspace of $\mathfrak{gl}\,(N, \mathbb{C})$, and for the $A, B \in \mathfrak{g} = T_IG \subseteq \mathfrak{gl}\,(N, \mathbb{C})$, the Lie bracket $[A, B]$ is just the commutator $AB - BA$. In all of what follows, we will assume that G is a matrix Lie group. For a matrix $A \in \mathfrak{gl}\,(N, \mathbb{C})$, we define

$$\exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Then we show that $\mathfrak{g}$ is the set, say $\mathfrak{s}$, of all $A \in \mathfrak{gl}\,(N, \mathbb{C})$ such that $\exp(tA) \in G$ for all $t \in \mathbb{R}$. Since $\mathfrak{g} := T_IG$, it is clear that $\mathfrak{s} \subseteq \mathfrak{g}$. To show $\mathfrak{g} \subseteq \mathfrak{s}$, suppose that $A \in \mathfrak{g}$, and let $\widetilde{A}$ be the vector field on G defined by $\widetilde{A}_g := L_{g*}(A)$, where $L_g : G \to G$ is given by $L_g(g') = gg'$. Since G is a matrix group where tangent vectors are considered to reside in $\mathfrak{gl}\,(N, \mathbb{C})$ and L_g is a linear transformation of $\mathfrak{gl}\,(N, \mathbb{C})$, $\widetilde{A}_g = L_{g*}(A)$ is simply gA. Now, $t \mapsto \exp(tA)$ is the solution curve at I of the vector field $\widetilde{A}$ on G, since

$$\begin{aligned}
\tfrac{d}{dt}\left(\exp(tA)\right) &= \tfrac{d}{dt}\left(\sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k\right) = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^k \\
&= \left(\sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^{k-1}\right) A = \exp(tA)\, A = \widetilde{A}_{\exp(tA)}.
\end{aligned}$$

In particular, $\exp(tA) \in G$ for all $t \in \mathbb{R}$ and hence $\mathfrak{g} \subseteq \mathfrak{s}$. Note that for $g \in G$, the map $Ad_g : G \to G$ given by $Ad_g(h) := ghg^{-1}$ fixes the identity I. Thus its derivative $(Ad_{g*})_I$ at I is a linear transformation of $T_I G$ or $\mathfrak{g}$. For $A \in \mathfrak{g}$, we have (at $t = 0$)

$$\begin{aligned}(Ad_{g*})_I(A) &= \tfrac{d}{dt} Ad_g(\exp(tA)) \\ &= \tfrac{d}{dt}\left(g \exp(tA) g^{-1}\right) = \tfrac{d}{dt}\exp\left(tgAg^{-1}\right) = gAg^{-1}.\end{aligned}$$

We denote $(Ad_{g*})_I(A)$ by $ad_g(A)$ and the homomorphism $ad : G \to \mathrm{GL}(\mathfrak{g})$ given by $g \mapsto ad_g$ is known as the *adjoint representation* of G. Observe that for $A, B \in \mathfrak{g}$,

$$ad_{\exp(tA)}(B) = \exp(tA) B \exp(-tA) = B + t(AB - BA) + \mathrm{O}\left(t^2\right),$$

and so

$$\tfrac{d}{dt}\left(ad_{\exp(tA)}(B)\right)\Big|_{t=0} = AB - BA = [A, B].$$

Thus, the derivative $ad_{*I} : \mathfrak{g} \to End(\mathfrak{g})$ of the map $ad : G \to \mathrm{GL}(\mathfrak{g})$ at I is given by $(ad_{*I}(A))(B) = [A, B]$. It is convenient to denote ad_{*I} by $\mathfrak{ad}$, and so $\mathfrak{ad} : \mathfrak{g} \to End(\mathfrak{g})$ is given by

$$(\mathfrak{ad}(A))(B) = (ad_{*I}(A))(B) = [A, B].$$

For a principal G-bundle $\pi : P \to M$ and any $A \in \mathfrak{g}$, there is a vector field A^* (known as a *fundamental vertical vector field*) defined at $p \in P$ by

$$A_p^* := \tfrac{d}{dt}\left(p \exp(tA)\right)\Big|_{t=0}.$$

Definition 16.4. A **connection** 1-**form** for the principal G-bundle $\pi : P \to M$ is a $\mathfrak{g}$-valued 1-form ω on P, such that for each $A \in \mathfrak{g}$, $g \in G$, $p \in P$ and $X \in T_p M$, we have

$$\begin{aligned}&(\mathrm{C}_1)\ \omega(A^*) = A \text{ and} \\ &(\mathrm{C}_2)\ \omega_{pg}(R_{g*}X) = ad_{g^{-1}}(\omega_p(X)) = g^{-1}\omega_p(X)\, g.\end{aligned} \tag{16.1}$$

The Definitions 16.3 and 16.4 are related as follows. Given ω as in definition 16.4, define

$$H_p := \{X \in T_p M : \omega_p(X) = 0\}.$$

Then condition C_1 in (16.1) insures that $\pi_{*p} : H_p \to T_{\pi(p)}M$ is an isomorphism, and C_2 guarantees that $R_{g*}(H_p) = H_{pg}$. Note that $ad_{g^{-1}}$ in C_2 is needed so that it be consistent with C_1, since

$$\begin{aligned}R_{g*}(A^*) &= \tfrac{d}{dt}(p \exp(tA) g) = \tfrac{d}{dt}\left(pgg^{-1}\exp(tA) g\right) \\ &= \tfrac{d}{dt}\left(pgAd_{g^{-1}}(\exp(tA))\right) = \left(ad_{g^{-1}}(A)\right)^*_{pg}\end{aligned}$$

and C_1 imply that

$$\omega_{pg}(R_{g*}(A^*)) = \omega_{pg}\left(\left(ad_{g^{-1}}(A)\right)^*_{pg}\right) = ad_{g^{-1}}(A) = ad_{g^{-1}}(\omega_p(A^*)).$$

Of course, when G is abelian, $ad_{g^{-1}}$ is the identity and C_2 says that ω_{pg} is invariant under R_g (i.e., $R_g^*\omega = \omega$ when G is abelian).

Exercise 16.5. Let G be a (closed) Lie subgroup of a matrix Lie Group $\overline{G}$. It can be verified that $\overline{G}/G$ naturally has the structure of a manifold and that $\pi : \overline{G} \to \overline{G}/G$ is a principal G-bundle. Let $\overline{\mathfrak{g}}$ denote the Lie algebra of $\overline{G}$. The **Maurer-Cartan form** for $\overline{G}$ is the $\overline{\mathfrak{g}}$-valued 1-form $\overline{\omega} \in \Omega^1\left(\overline{G}, \overline{\mathfrak{g}}\right)$ on $\widetilde{G}$ given at $\overline{g}$ by $\overline{\omega}_{\overline{g}}(L_{\overline{g}*}A) = \overline{\omega}_{\overline{g}}(\overline{g}A) = A$. Suppose that $\overline{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{m}$ where $ad_g(\mathfrak{m}) = \mathfrak{m}$ for all

$g \in G$ (i.e., $\mathfrak{m}$ is an ad_G-invariant subspace of $\overline{\mathfrak{g}}$). Let $\pi_{\mathfrak{g}} : \overline{\mathfrak{g}} \to \mathfrak{g}$ be the projection onto $\mathfrak{g}$ along $\mathfrak{m}$.
(a) Check that the form $\omega := \pi_{\mathfrak{g}} \circ \overline{\omega} \in \Omega^1\left(\overline{G}, \mathfrak{g}\right)$ is a connection 1-form for $\pi : \overline{G} \to \overline{G}/G$. Why do we need $ad_g(\mathfrak{m}) = \mathfrak{m}$?
(b) Use this construction to explicitly find a natural connection 1-form $\omega \in \Omega^1(\mathrm{SU}(2), \mathfrak{g})$ in the case of the Hopf bundle $\mathrm{SU}(2) \to \mathrm{SU}(2)/G$ in Exercise 16.2.

Given a connection 1-form ω, we can decompose any $X \in T_pP$ into *horizontal and vertical parts*, X^H and X^V respectively, where $\omega_p\left(X^H\right) = 0$, $\pi_{*p}\left(X^V\right) = 0$, and $X = X^H + X^V$. If ϕ is a k-form on P with values in a vector space W, then we may define a new k-form ϕ^H on P by

$$\phi^H(X_1, \ldots, X_k) := \phi\left(X_1^H, \ldots, X_k^H\right).$$

We define the *covariant derivative* $D^\omega\phi$ of ϕ relative to ω to be the W-valued $(k+1)$-form

$$D^\omega\phi := (d\phi)^H. \tag{16.2}$$

The *curvature* Ω^ω *of* ω is simply the covariant derivative of ω relative to itself

$$\Omega^\omega := D^\omega\omega = (d\omega)^H. \tag{16.3}$$

PROPOSITION 16.6. *We have*

$$\Omega^\omega = D^\omega\omega = d\omega + \omega \wedge \omega, \tag{16.4}$$

where

$$(\omega \wedge \omega)(X, Y) := \omega(X)\,\omega(Y) - \omega(Y)\,\omega(X).$$

PROOF. Both sides of (16.4) agree on any pair (X, Y) of vectors where one of X^V or Y^V is 0, since then $\omega(X) = 0$ or $\omega(Y) = 0$. Thus, it suffices to show that both sides agree on a pair (A^*, B^*) of fundamental vertical fields. In this case, $(D^\omega\omega)(A^*, B^*) = 0$, and

$$\begin{aligned}
& d\omega(A^*, B^*) + (\omega \wedge \omega)(A^*, B^*) \\
&= A^*[\omega(B^*)] - B^*[\omega(A^*)] - \omega([A^*, B^*]) + (\omega \wedge \omega)(A^*, B^*) \\
&= -\omega([A^*, B^*]) + \omega(A^*)\,\omega(B^*) - \omega(B^*)\,\omega(A^*) \\
&= -\omega([A^*, B^*]) + [A, B].
\end{aligned}$$

Thus, it suffices to check that

$$[A^*, B^*] = [A, B]^*. \tag{16.5}$$

We have (evaluating derivatives with respect to s and t at 0)

$$\begin{aligned}
[A^*, B^*]_p &= \tfrac{d}{dt} R_{\exp(-tA)*}\left(B^*_{p\exp(tA)}\right) \\
&= \tfrac{d}{dt} R_{\exp(-tA)*}\left(\tfrac{d}{ds} p\exp(tA)\exp(sB)\right) \\
&= \tfrac{d}{dt}\tfrac{d}{ds}\left(p\exp(tA)\exp(sB)\exp(-tA)\right) \\
&= \tfrac{d}{ds}\tfrac{d}{dt} p\exp(tA)\exp(sB)\exp(-tA) \\
&= \tfrac{d}{ds} pAd_{I*}(A)(sB) = \tfrac{d}{ds}\left(p\exp(s[A, B])\right) \\
&= [A, B]^*_p,
\end{aligned}$$

verifying (16.5). □

We mention that for general Lie groups (16.4) is written as

$$D^{\omega}\omega = d\omega + \tfrac{1}{2}[\omega,\omega], \tag{16.6}$$

where

$$\tfrac{1}{2}[\omega,\omega](X,Y) := \tfrac{1}{2}\left([\omega(X),\omega(Y)] - [\omega(Y),\omega(X)]\right)$$

is defined purely in terms of Lie brackets instead of commutators of matrices.

EXERCISE 16.7. We use the notation in Exercise 16.5. For $A \in \overline{\mathfrak{g}}$, let $\widetilde{A} \in C^{\infty}\left(T\overline{G}\right)$ be the vector field on $\overline{G}$ given by $\widetilde{A}_{\overline{g}} := L_{\overline{g}*}A = \overline{g}A$. Similarly, for $B \in \overline{\mathfrak{g}}$, we define $\widetilde{B}$.
(a) Show that $\left[\widetilde{A},\widetilde{B}\right] = \widetilde{[A,B]}$.
(b) Use formula 6.7 (p. 172) to show that

$$d\overline{\omega}\left(\widetilde{A},\widetilde{B}\right) = -\overline{\omega}\left(\left[\widetilde{A},\widetilde{B}\right]\right) = -[A,B].$$

(c) Conclude that the curvature $\Omega^{\omega} \in \Omega^2\left(\overline{G},\mathfrak{g}\right)$ of the connection ω in Exercise 16.5, is given by $\Omega^{\omega}\left(\widetilde{A},\widetilde{B}\right) = \pi_{\mathfrak{g}}([A,B])$.
(d) Consider the special case $\overline{G} = \mathrm{SU}(2)$ and $G = \{\exp(it\sigma(\mathbf{e}_3)) : t \in \mathbb{R}\}$, as in Example 16.2. Check that $\left\{-\frac{i}{2}\sigma(\mathbf{e}_1), -\frac{i}{2}\sigma(\mathbf{e}_2), -\frac{i}{2}\sigma(\mathbf{e}_3)\right\}$ is a basis for $\mathfrak{su}(2)$ and $\left[-\frac{i}{2}\sigma(\mathbf{v}), -\frac{i}{2}\sigma(\mathbf{w})\right] = -\frac{i}{2}\sigma(\mathbf{v}\times\mathbf{w})$ for $\mathbf{v},\mathbf{w} \in \mathbb{R}^3$. Conclude that if $\mathfrak{m} \subseteq \mathfrak{su}(2)$ in Exercise 16.7 (b) is chosen to be span $\left\{-\frac{i}{2}\sigma(\mathbf{e}_1), -\frac{i}{2}\sigma(\mathbf{e}_2)\right\}$, then at any $\overline{g} \in \mathrm{SU}(2)$ we have

$$\Omega^{\omega}_{\overline{g}}\left(\overline{g}\left(-\tfrac{i}{2}\sigma(\mathbf{v})\right), \overline{g}\left(-\tfrac{i}{2}\sigma(\mathbf{w})\right)\right) = (\mathbf{e}_3\cdot(\mathbf{v}\times\mathbf{w}))\,\tfrac{i}{2}\sigma(\mathbf{e}_3).$$

The following notion of horizontal lift will be crucial in many key computations.

DEFINITION 16.8. For principal G-bundle $\pi : P \to M$ with connection ω and a vector field Y on M the vector field X on P, such that $\omega(X) = 0$ and $\pi_*(X) = Y$ is called the **horizontal lift of** Y.

REMARK 16.9. Note that X is unique since $\pi_* : H_p \to T_{\pi(p)}M$ is an isomorphism. Moreover, for any $g \in G$, note that $R_{g*}(X)$ satisfies $\pi_*(R_{g*}(X)) = (\pi \circ R_g)_*(X) = \pi_*(X) = Y$ and $\omega(R_{g*}(X)) = 0$. Thus, $R_{g*}(X)$ is also a horizontal lift of Y, and by uniqueness $R_{g*}(X) = X$ (i.e., horizontal lifts are R_{g*}-invariant). In particular, for a fundamental vertical vector field A^* ($A \in \mathfrak{g}$) and a horizontal lift X, we have (at any $p \in P$)

$$[A^*,X]_p = \tfrac{d}{dt}\left(R_{\exp(-tA)*}\left(X_{p\exp(tA)}\right)\right)\Big|_{t=0} = \tfrac{d}{dt}(X_p)\Big|_{t=0} = 0. \tag{16.7}$$

3. Equivariant Forms and Associated Bundles

Let $r : G \to \mathrm{GL}(W)$ be a representation (i.e., a homomorphism), where $\mathrm{GL}(W)$ is the general linear group of a vector space W. For a principal G-bundle $\pi : P \to M$, there is a right action of G on $P \times W$ given by $(p,w)g = \left(pg, r\left(g^{-1}\right)w\right)$. Let $[p,w]$ denote the orbit $\{(p,w)g : g \in G\}$ of (p,w), and let $P \times_G W$ be the quotient space

$$P \times_G W := \frac{P \times W}{G} = \{[p,w] : (p,w) \in P \times W\}.$$

It is not difficult to verify that $\pi_W : P \times_G W \to M$, where $\pi_W([p,w]) := \pi(p)$, is a vector bundle which is known as an *associated vector bundle of P via r*. Note that for $p \in P$ and $x = \pi(p)$, any two points in the fiber $\pi_W^{-1}(x)$ have unique representatives of the form (p, w_1) and (p, w_2), and it is easily verified that $[p,w_1] + [p,w_2] := [p, w_1 + w_2]$ is a well-defined addition. The fibers $\pi_W^{-1}(x)$ are isomorphic to W, but not canonically so, since the isomorphism $[p,w] \mapsto w$ depends on the choice of $p \in \pi^{-1}(x)$.

Let $s : M \to P \times_G W$ be a section, and define $f : P \to W$ by the equation $s(\pi(p)) = [p, f(p)]$. Since

$$[p, f(p)] = s(\pi(p)) = s(\pi(pg)) = [pg, f(pg)] = [p, r(g) f(pg)],$$

f has the equivariance property

$$f(pg) = r(g)^{-1} f(p).$$

Thus, we have an isomorphism $(s \mapsto f)$

$$C^\infty(P \times_G W) \cong \overline{\Omega}^0(P,W) := \left\{ f \in C^\infty(P,W) : f(pg) = r(g)^{-1} f(p) \right\}.$$

$\overline{\Omega}^0(P,W)$ is called the space of *W-valued equivariant functions* (0-forms) on P. Whether one works with sections or equivariant functions, is largely a matter of taste or convenience. We will use equivariant forms as well.

Definition 16.10. Let $\pi : P \to M$ be a principal G-bundle and let $r : G \to \mathrm{GL}(W)$ be a representation. We denote the space of C^∞, W-valued k-forms on P by $\Omega^k(P,W)$. For $k > 0$, the space of **horizontal**, **equivariant**, **W-valued k-forms** on P, denoted by $\overline{\Omega}^k(P,W)$, consists of all $\alpha \in \Omega^k(P,W)$, such that for all vector fields $X_1, \ldots, X_k$ and $g \in G$, we have

(H) $\alpha(X_1, \ldots, X_k) = 0$ if $\pi_*(X_i) = 0$ for some $i \in \{1, \ldots, k\}$ and
(E) $\alpha(R_{g*}X_1, \ldots, R_{g*}X_k) = r(g)^{-1} \alpha(X_1, \ldots, X_k)$ (i.e., $R_g^* \alpha = r(g)^{-1} \alpha$).

Remark 16.11. Note for the representation $ad : G \to \mathrm{GL}(\mathfrak{g})$, we may consider $\overline{\Omega}^k(P,\mathfrak{g})$. However, a connection 1-form ω is not in $\overline{\Omega}^1(P,\mathfrak{g})$, since $\omega(A^*) = A$ in violation of Condition H in Definition 16.10. Nevertheless, if ω' is another connection, then $\omega - \omega' \in \overline{\Omega}^1(P,\mathfrak{g})$, since $(\omega - \omega')(A^*) = 0$. In other words, $\mathcal{C}(P)$ is an affine space based on $\overline{\Omega}^1(P,\mathfrak{g})$.

Exercise 16.12. (a) In the notation of the preceeding Remark, while $\omega \notin \overline{\Omega}^1(P,\mathfrak{g})$, show that the curvature 2-form $\Omega^\omega \in \overline{\Omega}^2(P,\mathfrak{g})$, relative to the representation $ad : G \to \mathrm{GL}(\mathfrak{g})$. [Hint. One may use (16.3) and (16.4) in conjuction with (C_2) in (16.1).]
(b) If G is abelian, deduce that $\Omega^\omega = \pi^* \Omega_0^\omega$ for a unique form $\Omega_0^\omega \in \Omega^2(M,\mathfrak{g})$, where π^* is pull-back on forms induced by $\pi : P \to M$. (c) If ω is as in Exercise 16.7 (d) and $\pi : \mathrm{SU}(2) \to S^2$ (see also Exercise 16.2, p. 366), show that the form Ω_0^ω in (b) is $\frac{i}{2}\sigma(\mathbf{e}_3)\nu$, where ν is the area 2-form of S^2. [Hint. Show that $\pi_{\overline{g}*}\left(\overline{g}\left(-\frac{i}{2}\sigma(\mathbf{v})\right)\right) = c(\overline{g})(\mathbf{v} \times \mathbf{e}_3)$ and note that $\nu_{\mathbf{x}}(\mathbf{a},\mathbf{b}) = \mathbf{x} \cdot (\mathbf{a} \times \mathbf{b})$ for $\mathbf{x} \in S^2$ and $\mathbf{a}, \mathbf{b} \in T_x S^2 \subset \mathbb{R}^3$.]

REMARK 16.13 (basic forms). Let $s \in \overline{\Omega}^0(P,W)$, $\eta \in \Omega^k(M,\mathbb{R})$ and $\beta = \pi^*\eta$. Define $s \otimes \beta \in \Omega^k(P,W)$ by

$$(s\otimes\beta)_p(X_1,\dots,X_k) := \beta(X_1,\dots,X_k)\,s(p) = \eta(\pi_*X_1,\dots,\pi_*X_k)\,s(p).$$

Then $s\otimes\beta$ clearly meets Condition H, and it meets Condition E, since

$$\begin{aligned}&(s\otimes\beta)_{pg}(R_{g*}X_1,\dots,R_{g*}X_k) = \eta(\pi_*(R_{g*}X_1),\dots,\pi_*(R_{g*}X_k))\,s(pg)\\ &= \eta(\pi_*X_1,\dots,\pi_*X_k)\,r(g)^{-1}s(p) = r(g)^{-1}(\eta(\pi_*X_1,\dots,\pi_*X_k)\,s(p))\\ &= r(g)^{-1}(s\otimes\beta)_p(X_1,\dots,X_k).\end{aligned}$$

Thus, $s\otimes\beta \in \overline{\Omega}^k(P,W)$, and we call such forms **basic**. Although not every $\alpha \in \overline{\Omega}^k(P,W)$ is basic, any $\alpha \in \overline{\Omega}^k(P,W)$ can be written as a finite sum $\sum_i s_i \otimes \beta_i$ of basic forms. Hence many facts concerning forms in $\overline{\Omega}^k(P,W)$ can be verified first for basic forms, and then extended by linearity.

Equivariance is preserved by covariant differentiation, namely

$$D^\omega : \overline{\Omega}^k(P,W) \to \overline{\Omega}^{k+1}(P,W).$$

For this, note that $D^\omega\alpha = d\alpha^H$ satisfies condition H in Definition 16.10, and since the distribution of horizontal subspaces is R_{g*} invariant, $R_{g*}(X^H) = R_{g*}(X)^H$ and so $(R_g^*\beta)^H = R_g^*(\beta^H)$ for any form β on P. Then $D^\omega\alpha$ meets condition E, since

$$\begin{aligned}R_g^*(D^\omega\alpha) &= R_g^*\left((d\alpha)^H\right) = \left(R_g^* d\alpha\right)^H = \left(d\left(R_g^*\alpha\right)\right)^H = \left(d\left(r(g)^{-1}\alpha\right)\right)^H\\ &= \left(r(g)^{-1}d\alpha\right)^H = r(g)^{-1}(d\alpha)^H = r(g)^{-1}D^\omega\alpha.\end{aligned}$$

Moreover, there is a very convenient formula given in the following

PROPOSITION 16.14. *For a representation $r : G \to \mathrm{GL}(W)$ and $\alpha \in \overline{\Omega}^k(P,W)$, we have*

$$D^\omega\alpha = d\alpha + r'(\omega)\wedge\alpha, \tag{16.8}$$

where r' is the Lie algebra representation (i.e., the derivative of $r : G \to \mathrm{GL}(W)$ at I) and (where σ runs over all permutations of $\{1,\dots,k\}$)

$$(r'(\omega)\wedge\alpha)(X_1,\dots,X_{k+1}) := \frac{1}{k!}\sum\nolimits_\sigma (-1)^\sigma\, r'(\omega(X_{\sigma_1}))\,\alpha\left(X_{\sigma_2},\dots,X_{\sigma_{k+1}}\right).$$

PROOF. To verify (16.8), we need to show that

$$\begin{aligned}&D^\omega\alpha(X_1,\dots,X_{k+1})\\ &= d\alpha(X_1,\dots,X_{k+1}) + (r'(\omega)\wedge\alpha)(X_1,\dots,X_{k+1})\end{aligned} \tag{16.9}$$

when each X_i is a fundamental vertical field or horizontal. If all of the X_i are horizontal, then both sides of (16.9) agree, since $\omega(X_i) = 0$. By (6.8), p. 172,

$$d\alpha(X_1,\dots,X_{k+1}) = \sum_{i=1}^{k+1}(-1)^{i+1}X_i\left[\alpha\left(X_1,\dots,\widehat{X_i},\dots,X_{k+1}\right)\right]$$

$$+\sum_{1\le i<j\le k+1}^{k+1}(-1)^{i+j}\alpha\left([X_i,X_j],X_1,\dots,\widehat{X_i},\dots,\widehat{X_j},\dots,X_{k+1}\right) \tag{16.10}$$

Note that for fundamental vertical fields A^* and B^*, we have $[A^*, B^*] = [A, B]^*$. Thus, both sides of (16.9) are zero when two or more of the X_i are vertical. In the remaining case, one of the X_i is vertical (say $X_1 = A^*$) and the rest are horizontal. Then, the left side of (16.9) is 0 and it remains to verify the right side is 0. For this we may assume that the horizontal $X_2, \ldots, X_{k+1}$ are horizontal lifts of vector fields $Y_2, \ldots, Y_{k+1}$ on M. By (16.7), $[A^*, X_2] = \cdots = [A^*, X_{k+1}] = 0$ and the right side of (16.9) is (at $p \in P$)

$$\begin{aligned}
&A^* \left[\alpha\left(X_2, \ldots, X_{k+1}\right)\right] + r'\left(\omega\left(A^*\right)\right) \alpha\left(X_2, \ldots, X_k\right) \\
&= \tfrac{d}{dt} \alpha_{p \exp tA}\left(R_{(\exp tA)*} X_2, \ldots, R_{(\exp tA)*} X_{k+1}\right)\big|_{t=0} \\
&+ r'(A)\, \alpha\left(X_2, \ldots, X_k\right) \\
&= \tfrac{d}{dt}\left(R^*_{\exp tA} \alpha\left(X_2, \ldots, X_{k+1}\right)\right)\big|_{t=0} + r'(A)\, \alpha\left(X_2, \ldots, X_k\right) \\
&= \tfrac{d}{dt} r\left(\exp(-tA)\right) \alpha_p\left(X_2, \ldots, X_{k+1}\right)\big|_{t=0} + r'(A)\, \alpha\left(X_2, \ldots, X_k\right) \\
&= r'(-A)\, \alpha\left(X_2, \ldots, X_k\right) + r'(A)\, \alpha\left(X_2, \ldots, X_k\right) = 0.
\end{aligned}$$

□

The space $\overline{\Omega}^k(P, W)$ can be identified with the space $\Omega^k(P \times_G W)$ of k-forms with values in the vector bundle $P \times_G W$ as follows. If $Y_1, \ldots, Y_k$ are vector fields on M, with horizontal lifts $\widetilde{Y}_1, \ldots, \widetilde{Y}_k$, then it is easy to check that for $\alpha \in \overline{\Omega}^k(P, W)$, we have

$$\alpha\left(\widetilde{Y}_1, \ldots, \widetilde{Y}_k\right) \in \overline{\Omega}^0(P, W) \cong C^\infty(P \times_G W).$$

Then we define a form $\alpha_M \in \Omega^k(P \times_G W)$ at $x = \pi(p)$ by

$$\alpha_M\left(Y_1, \ldots, Y_k\right) = \left[p, \alpha(p)\left(\widetilde{Y}_1, \ldots, \widetilde{Y}_k\right)\right].$$

Conversely, the same equation can be used to define $\alpha \in \overline{\Omega}^k(P, W)$ for a given $\alpha_M \in \Omega^k(P \times_G W)$. It is easy to see that the correspondence

$$\overline{\Omega}^k(P, W) \cong \Omega^k(P \times_G W) \quad (\alpha \leftrightarrow \alpha_M) \tag{16.11}$$

is actually independent of the choice of connection ω. Moreover, via this correspondence $D^\omega : \overline{\Omega}^k(P, W) \to \overline{\Omega}^{k+1}(P, W)$ provides us with a corresponding operator (also denoted D^ω), say

$$D^\omega : \Omega^k(P \times_G W) \to \Omega^{k+1}(P \times_G W).$$

Remark 16.15. Since it would be cumbersome to adhere to the notation α_M, we shall simply use the same symbol α, whether we regard α as in $\overline{\Omega}^k(P, W)$ or as in $\Omega^k(P \times_G W)$. The context will either be clear, irrelevant, or made explicit.

Of fundamental importance is

Proposition 16.16 (Bianchi Identity). *If ω is any connection on a principal G-bundle $\pi : P \to M$ and $\Omega^\omega \in \overline{\Omega}^2(P, \mathfrak{g})$ is the curvature of ω, then*

$$D^\omega\left(\Omega^\omega\right) = 0. \tag{16.12}$$

PROOF. We compute

$$\begin{aligned}
D^\omega(\Omega^\omega) &= d\Omega^\omega + \mathfrak{ad}(\omega)\wedge\Omega^\omega \\
&= d\Omega^\omega + \omega\wedge\Omega^\omega - \Omega^\omega\wedge\omega \\
&= d(d\omega+\omega\wedge\omega) + \omega\wedge(d\omega+\omega\wedge\omega) - (d\omega+\omega\wedge\omega)\wedge\omega \\
&= d(\omega\wedge\omega) + \omega\wedge d\omega - (d\omega)\wedge\omega \\
&= ((d\omega)\wedge\omega - \omega\wedge d\omega) + \omega\wedge d\omega - d\omega\wedge\omega = 0.
\end{aligned}$$

□

Note that $D^\omega(D^\omega\omega) = D^\omega(\Omega^\omega) = 0$. However, unlike ordinary exterior differentiation $d:\Omega^k(M,\mathbb{R})\to\Omega^{k+1}(M,\mathbb{R})$ which satisfies $d^2=0$, the composition $D^\omega\circ D^\omega : \overline{\Omega}^k(P,W)\to\overline{\Omega}^{k+2}(P,W)$ is *not* zero in general, and the curvature Ω^ω is the obstruction in the following sense.

PROPOSITION 16.17. *If ω is any connection on a principal G-bundle $\pi:P\to M$, $\Omega^\omega\in\overline{\Omega}^2(P,\mathfrak{g})$ is the curvature of ω, and $\alpha\in\overline{\Omega}^k(P,W)$, we have*

$$D^\omega(D^\omega\alpha) = r'(\Omega^\omega)\wedge\alpha. \tag{16.13}$$

PROOF. We compute

$$\begin{aligned}
D^\omega&(D^\omega\alpha) = D^\omega(d\alpha + r'(\omega)\wedge\alpha) \\
&= d(d\alpha + r'(\omega)\wedge\alpha) + r'(\omega)\wedge(d\alpha + r'(\omega)\wedge\alpha) \\
&= d(r'(\omega)\wedge\alpha) + r'(\omega)\wedge d\alpha + r'(\omega)\wedge(r'(\omega)\wedge\alpha) \\
&= (d(r'(\omega)))\wedge\alpha - r'(\omega)\wedge d\alpha + r'(\omega)\wedge d\alpha + (r'(\omega)\wedge r'(\omega))\wedge\alpha \\
&= (r'(d\omega) + r'(\omega)\wedge r'(\omega))\wedge\alpha \\
&= r'(d\omega+\omega\wedge\omega)\wedge\alpha = r'(\Omega^\omega)\wedge\alpha.
\end{aligned}$$

□

In order to define a formal adjoint to D^ω, we need to introduce some inner products. Let h be a Riemannian metric on M, and suppose that K is an inner product on W for which $r:G\to\mathrm{GL}(W)$ is *orthogonal*; i.e., for all $g\in G$,

$$r(g)\subseteq \mathrm{O}(W) := \{A\in\mathrm{GL}(W) : K(Aw_1,Aw_2) = K(w_1,w_2)\}.$$

Such a K always exists if G is compact, by an averaging argument. Then we define an inner product on each fiber $\pi_W^{-1}(x)$ of the associated vector bundle $\pi_W : P\times_G W\to M$ by

$$\langle [p,w_1],[p,w_2]\rangle_x := K(w_1,w_2),$$

which is independent of the choice of representatives $(p,w_i)\in[p,w_1]$ by the orthogonality of $r:G\to\mathrm{GL}(W)$. This pointwise inner product gives us a pairing

$$\langle\cdot,\cdot\rangle : C^\infty(P\times_G W)\times C^\infty(P\times_G W)\to C^\infty(M,\mathbb{R})$$

by simply defining $\langle s,t\rangle(x) := \langle s(x),t(x)\rangle_x$ for $s,t\in C^\infty(P\times_G W)$. We can show that

$$d\langle s,t\rangle = \langle D^\omega s,t\rangle + \langle s,D^\omega t\rangle\in\Omega^1(M,\mathbb{R}), \tag{16.14}$$

where the inner product on the right is just between the values in $P \times_G W$. First note that for $A \in \mathfrak{g}$, the product rule for differentiation yields (at $u = 0$)

$$0 = \tfrac{d}{du} K\left(r\left(\exp uA\right) w_1, r\left(\exp uA\right) w_2\right) = K\left(r'\left(A\right) w_1, w_2\right) + K\left(w_1, r'\left(A\right) w_2\right)$$

Then regarding $s, t \in \overline{\Omega}^0\left(P, W\right)$ and using (16.8), we get

$$\begin{aligned} d\left(K\left(s,t\right)\right) &= K\left(ds, t\right) + K\left(s, dt\right) \\ &= K\left(ds, t\right) + K\left(s, dt\right) + K\left(r'\left(\omega\right) s, t\right) + K\left(s, r'\left(\omega\right) t\right) \\ &= K\left(ds + r'\left(\omega\right) s, t\right) + K\left(s, dt + r'\left(\omega\right) t\right) \\ &= K\left(D^\omega s, t\right) + K\left(s, D^\omega t\right), \end{aligned}$$

an equality of right-invariant $\mathbb{R}$-valued 1-forms on P, which yields (16.14). Using the Riemannian metric h, there is a pairing

$$\langle \cdot, \cdot \rangle : \Omega^k\left(P \times_G W\right) \times \Omega^k\left(P \times_G W\right) \to C^\infty\left(M, \mathbb{R}\right), \tag{16.15}$$

such that for $s, t \in C^\infty\left(P \times_G W\right)$ and $\beta, \gamma \in \Omega^k\left(M, \mathbb{R}\right)$

$$\langle s \otimes \beta, t \otimes \gamma \rangle = \langle s, t \rangle\, h\left(\beta, \gamma\right),$$

where $h\left(\beta, \gamma\right)$ is given locally by

$$h\left(\beta, \gamma\right) := \frac{1}{k!} \beta^{i_1 \cdots i_k} \gamma_{i_1 \cdots i_k} := \frac{1}{k!} h^{i_1 j_1} \cdots h^{i_k j_k} \beta_{j_1 \cdots j_k} \gamma_{i_1 \cdots i_k}. \tag{16.16}$$

Here, $\beta_{j_1 \cdots j_k} = \beta\left(\partial_{j_1}, \ldots, \partial_{j_k}\right)$ for local coordinate vector fields $\partial_1, \ldots, \partial_n$ and the h^{ij} are the entries of the inverse of the matrix $[h_{ij}]$, where $h_{ij} := h\left(\partial_i, \partial_j\right)$. In (16.16) and elsewhere we adopt the Einstein summation convention where repeated indices on different levels are assumed to be summed from 1 to $n = \dim M$. Rather than introducing sections with compact support, let us assume that M is compact. Then we have the inner product

$$\left(\cdot, \cdot\right) : \Omega^k\left(P \times_G W\right) \times \Omega^k\left(P \times_G W\right) \to \mathbb{R} \tag{16.17}$$

given by

$$\left(\alpha_1, \alpha_2\right) := \int_M \langle \alpha_1, \alpha_2 \rangle\ \left|\nu_h\right|,$$

where $\left|\nu_h\right|$ is the density on M relative to h, given locally in coordinates $\left(x^1, \cdots, x^n\right)$ by

$$\left|\nu_h\right| = \left|\det\left(h_{ij}\right)\right|^{1/2} dx^1 \cdots dx^n.$$

For $\alpha \in \Omega^k\left(P \times_G W\right)$, we define

$$\left\|\alpha\right\|^2 := \left(\alpha, \alpha\right) \in \mathbb{R} \ \text{ and } \ \left|\alpha\right|^2 := \langle \alpha, \alpha \rangle \in C^\infty\left(M, \mathbb{R}\right).$$

Suitable modifications can be made to handle the case where W is complex, with Hermitian scalar product K and $r : G \to \mathrm{GL}\left(W\right)$ is a unitary representation.

To explicitly construct a formal adjoint of D^ω on $\Omega^k\left(P \times_G W\right)$, we introduce the (Hodge) *star operator* (for $k \in \{1, \ldots, n = \dim M\}$

$$* : \Omega^m\left(M, \mathbb{R}\right) \to \Omega^{n-m}\left(M, \mathbb{R}\right) \ \text{ for } \ m \in \{1, \ldots, n = \dim M\}.$$

In order to define $*$, we need to assume that M is oriented with volume form given locally in an oriented coordinate system $\left(x^1, \cdots, x^n\right)$ by

$$\nu_h := \left|\det\left[h_{ij}\right]\right|^{1/2} dx^1 \wedge \cdots \wedge dx^n$$

Then $*$ is defined to be the unique linear map, such that for all $\alpha, \beta \in \Omega^m(M, \mathbb{R})$

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \nu_h. \tag{16.18}$$

Note that $*$ can be defined pointwise. There also is a local formula (proven in [**Bl81**, p. 5])

$$(*\beta)_{j_{m+1}\cdots j_n} = \frac{1}{m!} \left|\det [h_{ij}]\right|^{1/2} \beta^{j_1\cdots j_m} \varepsilon_{j_1\cdots j_m j_{m+1}\cdots j_n}, \tag{16.19}$$

where ε is antisymmetric in its indices with $\varepsilon_{12\cdots n} = 1$ and $\nu_h(\partial_1, \ldots, \partial_n) > 0$. In [**Bl81**, p. 5-6] it is also shown that for $\beta \in \Omega^m(M, \mathbb{R})$,

$$*^2\beta := *(*\beta) = \operatorname{sign}(\det [h_{ij}]) (-1)^{m(n-m)} \beta.$$

For a Riemannian (positive definite) metric h, sign($\det [h_{ij}]$) $= 1$, but for h Lorentzian sign($\det [h_{ij}]$) $= -1$. In particular, for $n = \dim M = 4$, note that $*^2 = \mathrm{Id}$ on $\Omega^2(M, \mathbb{R})$ for Riemannian h, while $*^2 = -\mathrm{Id}$ for Lorentzian h. Thus, for Riemannian h, one has a decomposition

$$\Omega^2(M, \mathbb{R}) = \Omega^2_+(M, \mathbb{R}) \oplus \Omega^2_-(M, \mathbb{R})$$

where

$$\Omega^2_\pm(M, \mathbb{R}) := \left\{\beta \in \Omega^2(M, \mathbb{R}) : *\beta = \pm\beta\right\}.$$

Forms in $\Omega^2_+(M, \mathbb{R})$ are called *self-dual*, while forms in $\Omega^2_-(M, \mathbb{R})$ are called *anti-self-dual*. For a Lorentzian 4-manifold, there is a similar notion, but only after complexification where we can decompose $\Omega^2(M, \mathbb{C})$ into the $\pm i$ eigenspaces of $*$. Unless otherwise stated, we assume that h is Riemannian. Of course, we can extend the notion of star operator to the spaces $\Omega^m(P \times_G W) \cong C^\infty(P \times_G W) \otimes \Omega^m(M, \mathbb{R})$ via $*(s \otimes \beta) = s \otimes *\beta$. Moreover, for $\dim M = 4$, we still have a decomposition

$$\Omega^2(P \times_G W) = \Omega^2_+(P \times_G W) \oplus \Omega^2_-(P \times_G W)$$

into self-dual and anti-self-dual 2-forms. Note that if the orientation of M is reversed, then according to (16.18), $*$ changes sign, and $\Omega^2_+(\cdot)$ and $\Omega^2_-(\cdot)$ are interchanged. Also, for $\dim M = 4$, $* : \Omega^2(\cdot) \to \Omega^2(\cdot)$ is invariant under a conformal change of metric. Indeed, if h is replaced by λh for a positive $\lambda \in C^\infty(M, \mathbb{R})$, in the local formula (16.19) $\left|\det (h_{ij})\right|^{1/2}$ gains a factor of λ^2, while $\beta^{j_1 j_2}$ gains a factor of λ^{-2} from the raising of the two indices (since h^{ij} becomes $\lambda^{-1}h^{ij}$).

For $s, s' \in C^\infty(P \times_G W)$, $\beta \in \Omega^m(M, \mathbb{R})$ and $\beta' \in \Omega^{m'}(M, \mathbb{R})$, the following definition is convenient

$$K((s \otimes \beta) \wedge (s' \otimes \beta')) := K(s, s')\, \beta \wedge \beta'.$$

Then for $\alpha \in \Omega^m(P \times_G W)$ and $\alpha' \in \Omega^{m'}(P \times_G W)$,

$$K(\alpha \wedge \alpha') \in \Omega^{m+m'}(M, \mathbb{R})$$

is defined by linearity. We have

$$d(K(\alpha \wedge *\alpha')) = K(D^\omega \alpha \wedge \alpha') + (-1)^m K(\alpha \wedge D^\omega \alpha'), \tag{16.20}$$

since (using (16.14))

$$\begin{aligned} dK((s\otimes\beta)\wedge(s'\otimes\beta')) &= d(K(s,s'))\wedge\beta\wedge\beta' \\ &\quad + K(s,s')\,d\beta\wedge\beta' + K(s,s')(-1)^m\beta\wedge d\beta' \\ &= K(D^\omega s,s')\wedge\beta\wedge\beta' + K(s,D^\omega s')\wedge\beta\wedge\beta' \\ &\quad + K(s,s')\,d\beta\wedge\beta' + K(s,s')(-1)^m\beta\wedge d\beta' \\ &= K((D^\omega s\wedge\beta + s\otimes d\beta)\wedge(s'\otimes\beta')) \\ &\quad + (-1)^m K((s\otimes\beta)\wedge(D^\omega s'\wedge\beta' + s\otimes d\beta')) \\ &= K(D^\omega(s\otimes\beta)\wedge(s'\otimes\beta')) \\ &\quad + (-1)^m K((s\otimes\beta)\wedge D^\omega(s'\otimes\beta')). \end{aligned}$$

Moreover, if $m'=m$, we have

$$K(\alpha\wedge *\alpha') = \langle\alpha,\alpha'\rangle\, v_h.$$

PROPOSITION 16.18. *The formal adjoint of $D^\omega : \Omega^m(P\times_G W)\to\Omega^{m+1}(P\times_G W)$ on a compact, oriented, Riemannian n-manifold M is the covariant codifferential $\delta^\omega : \Omega^{m+1}(P\times_G W)\to\Omega^m(P\times_G W)$ given by*

$$\delta^\omega := -(-1)^{nm} * D^\omega * \tag{16.21}$$

In other words, for $\alpha\in\Omega^m(P\times_G W)$ and $\alpha'\in\Omega^{m+1}(P\times_G W)$, we have

$$(D^\omega\alpha,\alpha') = (\alpha,\delta^\omega\alpha').$$

PROOF. If we show that

$$d\gamma = (\langle D^\omega\alpha,\alpha'\rangle - \langle\alpha,\delta^\omega\alpha'\rangle)\, v_h,$$

then $(D^\omega\alpha,\alpha')-(\alpha,\delta^\omega\alpha') = \int_M d\gamma = 0$ by Stoke's Theorem. Using (16.20), we compute

$$\begin{aligned} d\gamma &= d(K(\alpha\wedge *\alpha')) \\ &= K(D^\omega\alpha\wedge *\alpha') + (-1)^m K(\alpha\wedge D^\omega(*\alpha')) \\ &= K(D^\omega\alpha\wedge *\alpha') + (-1)^m K\left(\alpha\wedge(-1)^{(n-m)m} *^2 D^\omega(*\alpha')\right) \\ &= K(D^\omega\alpha\wedge *\alpha') + K(\alpha\wedge *((-1)^{nm} * D^\omega(*\alpha'))) \\ &= K(D^\omega\alpha\wedge *\alpha') - K(\alpha\wedge *(\delta^\omega\alpha')) \\ &= (\langle D^\omega\alpha,\alpha'\rangle - \langle\alpha,\delta^\omega\alpha'\rangle)\, v_h, \end{aligned}$$

as required. □

To obtain formulas for D^ω and δ^ω in local coordinates, let $\sigma : U\to P$ be a local section on a coordinate neighborhood U, and let $\alpha\in\Omega^m(P\times_G W)$. For each $x\in U$, we have an isomorphism $(P\times_G W)_x = \pi_W^{-1}(x)\to W$ given by $[\sigma(x),w]\mapsto w$. This yields an isomorphism $\Omega^m((P\times_G W)|U)\cong\Omega^m(U,W)$ which we denote by $\alpha\mapsto\widetilde{\alpha}$. One can easily check that if $\overline{\alpha}\in\overline{\Omega}^k(P,W)$ is the equivariant form corresponding to α, then $\widetilde{\alpha}=\sigma^*\overline{\alpha}$. By (16.8), we have $D^\omega\overline{\alpha} = d\overline{\alpha} + r'(\omega)\wedge\overline{\alpha}$ and

$$\begin{aligned} \widetilde{D^\omega\alpha} &= \sigma^*(D^\omega\overline{\alpha}) = \sigma^*(d\overline{\alpha} + r'(\omega)\wedge\overline{\alpha}) \\ &= d(\sigma^*\overline{\alpha}) + r'(\sigma^*\omega)\wedge\sigma^*\overline{\alpha} \\ &= d\widetilde{\alpha} + r'(\sigma^*\omega)\wedge\widetilde{\alpha}. \end{aligned}$$

In the local coordinates $\left(x^1,\ldots,x^m\right)$ on U we may write

$$\widetilde{\alpha}=\frac{1}{m!}\sum\widetilde{\alpha}_{i_1\ldots i_m}dx^{i_1}\wedge\cdots\wedge dx^{i_m},$$

where it is assumed that $\widetilde{\alpha}_{i_1\ldots i_m}$ is antisymmetric in $i_1,\ldots,i_m$. Then

$$\left(\widetilde{D^\omega\alpha}\right)_{j_1\cdots j_{m+1}}=\sum_{k=1}^{n}\left(-1\right)^{k+1}\left(\partial_{j_k}\left(\widetilde{\alpha}_{j_1\ldots\widehat{j_k}\ldots j_{m+1}}\right)+r'\left(\sigma^*\omega\left(\partial_{j_k}\right)\right)\widetilde{\alpha}_{j_1\ldots\widehat{j_k}\ldots j_{m+1}}\right), \tag{16.22}$$

where $\widehat{j_k}$ means that j_k is omitted. Using $\widetilde{*\alpha}=*\widetilde{\alpha}$, we have

$$\begin{aligned}-\left(-1\right)^{n(m-1)}\widetilde{\delta^\omega\alpha}&=*\widetilde{D^\omega*\alpha}=*\widetilde{D^\omega*\alpha}\\&=*\left(d\left(\widetilde{*\alpha}\right)+r'\left(\sigma^*\omega\right)\wedge\widetilde{*\alpha}\right)\\&=*\left(d\left(*\widetilde{\alpha}\right)+r'\left(\sigma^*\omega\right)\wedge*\widetilde{\alpha}\right).\end{aligned} \tag{16.23}$$

While it is possible to get this formula with (16.19), in order to find the components $\left(\delta^\omega\alpha\right)_{i_1\ldots i_{m-1}}$, it is easier to compute the formal adjoint of the operator $\widetilde{\alpha}\mapsto\widetilde{D^\omega\alpha}$, using (16.22) and integration by parts assuming that α is compactly supported in U. The final result is

$$\begin{aligned}\left(\widetilde{\delta^\omega\alpha}\right)_{i_1\ldots i_{m-1}}&=-\left|h\right|^{-1/2}h_{i_1j_1}\cdots h_{i_{m-1}j_{m-1}}\partial_i\left(\left|h\right|^{1/2}\widetilde{\alpha}^{ij_1\ldots j_{m-1}}\right)\\&\quad-r'\left(\left(\sigma^*\omega\right)\left(\partial_i\right)\right)\widetilde{\alpha}^i_{\ i_1\ldots i_{m-1}}.\end{aligned} \tag{16.24}$$

Associated principal bundles induced by Lie group homomorphisms

The construction of vector bundles associated to a given principal G-bundle $\pi:P\to M$ via a representation $r:G\to\mathrm{GL}\left(W\right)$ can be generalized to the case where r is replaced by left action of G on any manifold F, say $r:G\to C^\infty\left(F,F\right)$. Of particular use to us in applications will be the case where G acts on a second group G' via $g\cdot g':=\gamma\left(g\right)g'$ where $\gamma:G\to G'$ is a homomorphism. The following proposition shows that in this case the associated bundle is a principal G'-bundle $P'\to M$. Also, a connection for $\pi:P\to M$ gives rise to a connection for $\pi:P'\to M$. Moreover, an equivarant map between vector representation spaces V and V' of G and G' gives rise to a vector bundle morphism between the associated vector bundles $P\times_G V$ and $P'\times_{G'}V'$.

PROPOSITION 16.19. *For Lie groups G and G', let $\pi:P\to M$ be a principal G-bundle and let $\gamma:G\to G'$ be a homomorphism. Then there is a canonically constructed principal G'-bundle $\pi':P'\to M$. Moreover, there is a canonical map $\Gamma:P\to P'$ which is γ-equivariant, in the sense that $\Gamma\left(pg\right)=\Gamma\left(p\right)\gamma\left(g\right)$ (i.e., $\Gamma\circ R_g=R_{\gamma(g)}\circ\Gamma$). For any $p'\in P'$, $\Gamma^{-1}\left(p'\right)$ is an orbit of the action of* $\mathrm{Ker}\,\gamma$ *on P, so that Γ is an embedding only if* $\mathrm{Ker}\,\gamma=0$.

PROOF. We define the principal G'-bundle $\pi':P'\to M$ as follows. There is a right action $R:\left(P\times G'\right)\times G\to P\times G'$ of G on $P_0\times G'$ given by

$$R\left(g\right)\left(p,g'\right)=\left(p,g'\right)\cdot g:=\left(pg,\gamma\left(g^{-1}\right)g'\right).$$

Let the orbit of $\left(p,g'\right)$ be $\left[p,g'\right]:=\left\{\left(p,g'\right)\cdot g:g\in G\right\}$ and let

$$P':=P\times_G G':=\frac{P\times G'}{G}=\left\{\left[p,g'\right]:p\in P,g\in G'\right\}.$$

Define an action of G' on P by $[p,g']\cdot h' := [p,g'h']$ for $h'\in G'$. Since

$$\left[pg, h\left(g^{-1}\right)g'h'\right] = [p,g'h'],$$

this action is well defined. The action is free, since

$$\begin{aligned}[p,g'h'] = [p,g'] &\Rightarrow (p,g'h') = \left(pg,\gamma\left(g^{-1}\right)g'\right) \text{ for some } g\in G\\ &\Rightarrow g = e \text{ and } g'h' = \gamma\left(g^{-1}\right)g' = g' \Rightarrow h' = e',\end{aligned}$$

where $e\in G$ and $e'\in G'$ are the identities. Then $\pi' : P' \to M$ (with $\pi'([p,g']) = \pi(p)$) is a principal G'-bundle. Let $\Gamma : P\to P'$ be given by $\Gamma(p) := [p,e']$. We check that Γ is γ-equivariant:

$$\Gamma(pg) = [pg,e'] = \left[pgg^{-1},\gamma(g)e'\right] = [p,e']\gamma(g) = \Gamma(p)\gamma(g).$$

Note that $\Gamma^{-1}(p')$ is an orbit of the action of $\operatorname{Ker}\gamma$ on P: For $p_1,p_2\in P$,

$$\begin{aligned}\Gamma(p_1) = \Gamma(p_2) = p &\Leftrightarrow [p_1,e'] = [p_2,e']\\ &\Leftrightarrow \exists g\in G, \text{ s.t. } \left(p_1g,\gamma\left(g^{-1}\right)e'\right) = (p_2,e')\\ &\Leftrightarrow p_2 = p_1g \text{ and } \gamma\left(g^{-1}\right) = e' \text{ (i.e. } g\in\operatorname{Ker}\gamma).\end{aligned}$$

□

REMARK 16.20. If $\widetilde{P}'$ is a another principal G'-bundle and $\widetilde{\Gamma} : P\to\widetilde{P}'$ is γ-equivariant, then $\widetilde{P}'$ is isomorphic to P' via the bijection $\Gamma(p)g' \longleftrightarrow \widetilde{\Gamma}(p)g'$. This bijection is well defined (and then clearly G'-equivariant). Indeed, for $\pi(p_1) = \pi(p_2)$, we have $p_2 = p_1g$ for some $g\in G$, and

$$\begin{aligned}\Gamma(p_1)g_1' = \Gamma(p_2)g_2' &\Leftrightarrow \Gamma(p_1)g_1' = \Gamma(p_1g)g_2' = \Gamma(p_1)\gamma(g)g_2'\\ &\Leftrightarrow g_1' = \gamma(g)g_2' \Leftrightarrow \widetilde{\Gamma}(p_1)g_1' = \widetilde{\Gamma}(p_1)\gamma(g)g_2' = \widetilde{\Gamma}(p_2)g_2'.\end{aligned}$$

With regard to induced connections, we have

PROPOSITION 16.21. *Let $\gamma' : \mathfrak{g}\to\mathfrak{g}'$ denote the Lie algebra map for $\gamma : G\to G'$. For any connection 1-form $\omega\in\Omega^1(P,\mathfrak{g})$ on P, there is a unique connection 1-form $\omega'\in\Omega^1(P',\mathfrak{g}')$, such that $\Gamma^*\omega' = \gamma'\circ\omega$. Moreover, we have $\Gamma^*\left(\Omega^{\omega'}\right) = \gamma'\circ\Omega^{\omega}$.*

PROOF. Since $\Gamma\circ R_g = R_{\gamma(g)}\circ\Gamma$ for all $g\in G$, the G-invariant distribution $\mathcal{H}$ of ω-horizontal subspaces is mapped by $\Gamma_* : P\to P'$ to a well-defined $\gamma(G)$-invariant distribution of horizontal subspaces, say $\Gamma(\mathcal{H})$ on $\Gamma(P)$. Via the various $R_{g'*}$ for $g'\in G'$, $\Gamma(\mathcal{H})$ uniquely extends to a G'-invariant horizontal distribution, say $\mathcal{H}'$, on all of P'. Then $\mathcal{H}'$ determines a connection 1-form ω' on P'. By definition of ω', we have $\Gamma^*\omega' = 0$ on the horizontal subspaces of ω. Thus, $\Gamma^*\omega' = \gamma'\circ\omega$ on $\mathcal{H}$. As for vertical vectors, if $A\in\mathfrak{g}$, we have

$$\begin{aligned}\Gamma_*\left(A_p^*\right) &= \tfrac{d}{dt}\Gamma(p\exp(tA)) = \tfrac{d}{dt}\Gamma(p)\gamma(\exp(tA))\\ &= \tfrac{d}{dt}\left(\Gamma(p)\exp(t\gamma'(A))\right) = \gamma'(A)^*_{\Gamma(p)}, \text{ and so}\end{aligned}$$

$$(\Gamma^*\omega')(A^*) = \omega'(\Gamma_*(A^*)) = \omega'\left(\gamma'(A)^*\right) = \gamma'(A) = (\gamma'\circ\omega)(A^*).$$

Hence, $\Gamma^*\omega' = \gamma' \circ \omega$, and uniqueness is clear. Moreover,

$$\begin{aligned}\Gamma^*\left(\Omega^{\omega'}\right) &= \Gamma^*\left(d\omega' + \tfrac{1}{2}[\omega',\omega']\right) = d\Gamma^*\omega' + \tfrac{1}{2}[\Gamma^*\omega',\Gamma^*\omega'] \\ &= d\left(\gamma'\circ\omega\right) + \tfrac{1}{2}[\gamma'\circ\omega,\gamma'\circ\omega] = \gamma'\circ\left(d\omega + \tfrac{1}{2}[\omega,\omega]\right) \\ &= \gamma'\circ\Omega^{\omega}.\end{aligned}$$

□

Another useful fact concerns associated bundles.

PROPOSITION 16.22. *Suppose that we have representations*

$$r : G \to \mathrm{GL}\left(V\right) \text{ and } r' : G' \to \mathrm{GL}\left(V'\right)$$

and a linear map $\phi : V \to V'$ which is equivariant, in the sense that

$$\phi\left(r\left(g\right)\left(v\right)\right) = r'\left(\gamma\left(g\right)\right)\left(\phi\left(v\right)\right).$$

Then there is a vector bundle morphism

$$\Phi : P \times_G V \to P' \times_{G'} V' \text{ given by } \Phi\left([p,v]\right) = \left[\Gamma\left(p\right),\phi\left(v\right)\right].$$

PROOF. Note that Φ is well-defined, since

$$\begin{aligned}&[p_1,v_1] = [p_2,v_2] \\ &\Rightarrow (p_1,v_1) = \left(p_2 g, \left(r\left(g^{-1}\right)\right)(v_2)\right) \text{ for some } g\in G \\ &\Rightarrow \left([p_1,e'],\phi\left(v_1\right)\right) = \left([p_2 g,e'],\phi\left(r\left(g^{-1}\right)v_2\right)\right) \text{ for some } g\in G \\ &\Rightarrow \left([p_1,e'],\phi\left(v_1\right)\right) = \left([p_2,e']\gamma\left(g\right), r'\left(\gamma\left(g\right)^{-1}\right)\left(\phi\left(v_2\right)\right)\right) \text{ for some } g\in G \\ &\Rightarrow \left[[p_1,e'],\phi\left(v_1\right)\right] = \left[[p_2,e'],\phi\left(v_2\right)\right] \\ &\Rightarrow \left[\Gamma\left(p_1\right),\phi\left(v_1\right)\right] = \left[\Gamma\left(p_2\right),\phi\left(v_2\right)\right].\end{aligned}$$

□

REMARK 16.23. It is easy to see that Φ is injective (or surjective) if and only if ϕ is injective (or surjective). In particular, if ϕ is an an isomorphism, then so is $\Phi : P \times_G V \to P \times_{G'} V'$.

4. Gauge Transformations

While physicists speak of gauge transformations of particle fields and gauge potentials, each of these is induced by gauge transformations of a principal bundle defined as follows.

DEFINITION 16.24. A **gauge transformation** of a principal G-bundle $\pi : P \to M$ is a diffeomorphism $F : P \to P$, such that for all $p \in P$ and $g \in G$,

$$\begin{aligned}&(G_1)\ F\left(pg\right) = F\left(p\right)g \text{ and} \\ &(G_2)\ \pi\left(F\left(p\right)\right) = \pi\left(p\right).\end{aligned}$$

We denote the group of gauge transformations by $\mathrm{GA}\left(P\right)$.

REMARK 16.25. Condition G_2 implies that the fibers are mapped into themselves. If G_2 were dropped, then G_1 and the fact that F is a diffeomorphism imply that there is a diffeomorphism $f : M \to M$, such that $\pi\left(F\left(p\right)\right) = f\left(\pi\left(p\right)\right)$. In this more general case (i.e., if G_2 is dropped), F is called an **automorphism** of P. We denote the group of automorphisms of P by $\mathrm{Aut}\left(P\right)$.

$\mathrm{Aut}(P)$ acts to the left on the space $\mathcal{C}(P)$ of connection 1-forms on P, as well as the spaces $\overline{\Omega}^k(P,W)$ of horizontal, equivariant k-forms on P (relative to a representation $r: G \to \mathrm{GL}(W)$) via pull-back:

$$F \cdot \alpha := \left(F^{-1}\right)^* \alpha, \quad \alpha \in \begin{cases} \mathcal{C}(P) \\ \overline{\Omega}^k(P,W). \end{cases}$$

To prove that $F\cdot\alpha \in \mathcal{C}(P)$ for $\alpha \in \mathcal{C}(P)$, note that F^{-1} preserves the fundamental vertical fields A^* for $A \in \mathfrak{g}$, since $F^{-1}(p \exp tA) = F^{-1}(p) \exp tA$. Thus,

$$\left(\left(F^{-1}\right)^* \alpha\right)(A^*) = \alpha\left(F_*^{-1}(A^*)\right) = \alpha(A^*) = A.$$

(i.e., Condition C_1 of (16.1) is met by $\left(F^{-1}\right)^*\alpha$). Also, since $F^{-1} \circ R_g = R_g \circ F^{-1}$,

$$R_g^*\left(\left(F^{-1}\right)^* \alpha\right) = \left(F^{-1}\right)^* R_g^*(\alpha) = \left(F^{-1}\right)^*\left(g^{-1}\alpha g\right) = g^{-1}\left(\left(F^{-1}\right)^*\alpha\right) g, \tag{16.25}$$

so that C_2 of (16.1) is met. It is also easy to prove that $F \cdot \alpha \in \overline{\Omega}^k(P,W)$ for $\alpha \in \overline{\Omega}^k(P,W)$. Indeed, for any fundamental vertical field A^*,

$$F \cdot \alpha(A^*, \cdots) = \left(F^{-1}\right)^* \alpha(A^*, \cdots) = \alpha\left(\left(F^{-1}\right)_*(A^*), \cdots\right) = 0,$$

since we have observed that $\left(F^{-1}\right)_*(A^*)$ is vertical. Also $R_g^*(F\cdot\alpha) = r\left(g^{-1}\right)(F\cdot\alpha)$ by the same sort of computation as (16.25). Since $\overline{\Omega}^k(P,W) \cong \Omega^k(P\times_G W)$ (see (16.11)), $\mathrm{GA}(P)$ also acts on $\Omega^k(P\times_G W)$. Recall (see (16.15)) that if the representation $r: G \to \mathrm{GL}(W)$ is orthogonal relative to some inner product k on W and there is a Riemannian metric h on M, then there is a pairing $\langle\cdot,\cdot\rangle : \Omega^m(P\times_G W) \times \Omega^m(P\times_G W) \to C^\infty(M,\mathbb{R})$. It is straightforward to check that this pairing is invariant under $\mathrm{GA}(P)$, and it follows that $\mathrm{GA}(P)$ acts by isometries on the pre-Hilbert spaces $\Omega^m(P\times_G W)$; for this, it may be easier to work with the related pairing on $\overline{\Omega}^k(P,W)$.

Proposition 16.26. *For $F \in \mathrm{GA}(P)$ and $\alpha \in \overline{\Omega}^k(P,W)$, we have*

$$F\cdot(D^\omega\alpha) = D^{F\cdot\omega}(F\cdot\alpha).$$

Proof. Since wedge and d commute with pull-back,

$$\begin{aligned} F\cdot(D^\omega\alpha) &= \left(F^{-1}\right)^*(D^\omega\alpha) = \left(F^{-1}\right)^*(d\alpha + r(\omega)\wedge\alpha) \\ &= \left(F^{-1}\right)^*(d\alpha) + \left(F^{-1}\right)^*(r(\omega)\wedge\alpha) \\ &= d\left(\left(F^{-1}\right)^*\alpha\right) + \left(r\left(\left(F^{-1}\right)^*\omega\right)\wedge\left(F^{-1}\right)^*\alpha\right) \\ &= D^{F\cdot\omega}(F\cdot\alpha). \end{aligned}$$

□

There is another way of looking at $\mathrm{GA}(P)$. Let

$$C(P,G) := \left\{f \in C^\infty(P,G) : f(pg) = g^{-1}f(p)\,g = Ad_{g^{-1}}f(p)\right\}. \tag{16.26}$$

Since we have assumed that G is a matrix Lie group (i.e., $G \subseteq \mathrm{GL}(N,\mathbb{C})$), the adjoint action of G on itself (i.e., $g\cdot g_0 = Ad_g(g_0) = g g_0 g^{-1}$) can be regarded as a representation $Ad: G \to \mathrm{GL}(\mathfrak{gl}(N,\mathbb{C}))$. Hence

$$C(P,G) \subseteq \overline{\Omega}^0(P,\mathfrak{gl}(N,\mathbb{C})) \cong \Omega^0(P\times_G \mathfrak{gl}(N,\mathbb{C})).$$

Note that $C(P,G)$ has a group operation given, for $f_1, f_2 \in C(P,G)$, by $(f_1 f_2)(p) = f_1(p) f_2(p)$.

PROPOSITION 16.27. *There is an isomorphism $\Phi : C(P,G) \to \mathrm{GA}(P,G)$ of groups given by*

$$\Phi(f)(p) := pf(p)^{-1} \quad \text{for } f \in C(P,G).$$

PROOF. Note that $\Phi(f) \in \mathrm{GA}(P,G)$, since

$$\Phi(f)(pg) = pgf(pg)^{-1} = pg\left(g^{-1}f(p)g\right)^{-1} = pf(p)^{-1}g = \Phi(f)(p)g,$$

and Φ is a homomorphism, since

$$\begin{aligned}\Phi(f_1 f_2)(p) &= p(f_1(p) f_2(p))^{-1} \\ &= p\left(f_2(p)^{-1} f_1(p)^{-1}\right) = (\Phi(f_1) \circ \Phi(f_2))(p).\end{aligned}$$

For $F \in \mathrm{GA}(P,G)$, define $\Psi(F) \in C^\infty(P,G)$ by $F(p) = p\Psi(F)(p)^{-1}$. Then $\Psi(F) \in C(P,G)$, since

$$\begin{aligned}&pg\Psi(F)(pg)^{-1} = F(pg) = F(p)g = p\Psi(F)(p)^{-1}g \\ &\Rightarrow g\Psi(F)(pg)^{-1} = \Psi(F)(p)^{-1}g \\ &\Rightarrow \Psi(F)(pg) = g^{-1}\Psi(F)(p)g.\end{aligned}$$

Note that Ψ is the inverse of Φ, since $p\Psi(\Phi(f))(p)^{-1} = \Phi(f)(p) = pf(p)^{-1}$ and $\Phi(\Psi(F))(p) = p\Psi(F)(p)^{-1} = F(p)$. □

PROPOSITION 16.28. *Let $f \in C(P,G)$. For $\omega \in \mathcal{C}(P)$ and $X \in T_pP$, we have*

$$(\Phi(f) \cdot \omega)(X) = f(p)\left(f^{-1}\right)_{*p}(X) + f(p)\,\omega(X)\,f(p)^{-1}. \tag{16.27}$$

For $\alpha \in \overline{\Omega}^m(P,W)$ and $X_1, \ldots, X_k \in T_pP$, we have

$$(\Phi(f) \cdot \alpha)_p(X_1, \ldots, X_m) = r(f(p))(\alpha_p(X_1, \ldots, X_m)). \tag{16.28}$$

PROOF. Let $\gamma : \mathbb{R} \to P$ be a curve with $\gamma'(0) = X \in T_pP$, and let $f \in C(P,G)$. At $t = 0$, we have

$$\begin{aligned}\Phi(f)_*(X) &= \tfrac{d}{dt}\Phi(f)(\gamma(t)) = \tfrac{d}{dt}\gamma(t)f(\gamma(t)) \\ &= \tfrac{d}{dt}pf(\gamma(t)) + \tfrac{d}{dt}\gamma(t)f(p) \\ &= \tfrac{d}{dt}pf(p)f(p)^{-1}f(\gamma(t)) + \tfrac{d}{dt}R_{f(p)}(\gamma(t)) \\ &= \left(f(p)^{-1}f_{*p}(X)\right)^*_{pf(p)} + R_{f(p)*}(X).\end{aligned}$$

We have used the fact that $t \mapsto f(p)^{-1}f(\gamma(t))$ is a curve through $I \in G$ with tangent vector $f(p)^{-1}f_{*p}(X) \in T_IG = \mathfrak{g}$, and hence $\frac{d}{dt}pf(p)f(p)^{-1}f(\gamma(t))$ coincides with the fundamental vertical field $f(p)^{-1}df_p(X)^*$ at $pf(p)$. Since $\Phi(f)^{-1} = \Phi\left(f^{-1}\right)$, we also have

$$\left(\Phi(f)^{-1}\right)_*(X) = \left(f(p)\left(f^{-1}\right)_{*p}(X)\right)^*_{pf(p)^{-1}} + R_{f(p)^{-1}*}(X)$$

Thus,

$$\begin{aligned}(\Phi(f)\cdot\omega)(X) &= \left(\left(\Phi(f)^{-1}\right)^{*}\omega\right)(X) = \omega\left(\left(\Phi(f)^{-1}\right)_{*}(X)\right)\\ &= \omega\left(\left(f(p)\left(f^{-1}\right)_{*p}(X)\right)^{*}_{pf(p)^{-1}} + R_{f(p)^{-1}*}(X)\right)\\ &= f(p)\left(f^{-1}\right)_{*p}(X) + R^{*}_{f(p)^{-1}}\omega(X)\\ &= f(p)\left(f^{-1}\right)_{*p}(X) + f(p)\,\omega(X)\,f(p)^{-1}.\end{aligned}$$

Using the fact that α vanishes on vertical vectors, we also have

$$\begin{aligned}&(\Phi(f)\cdot\alpha)_p(X_1,\dots,X_m) = \left(\Phi(f)^{-1*}\alpha\right)_p(X_1,\dots,X_m)\\ &= \alpha\left(\left(\Phi(f)^{-1}\right)_{*}X_1,\dots,\left(\Phi(f)^{-1}\right)_{*}X_m\right)\\ &= \alpha\left(R_{f(p)^{-1}*}(X_1),\dots,R_{f(p)^{-1}*}X_m\right)\\ &= \left(R^{*}_{f(p)^{-1}}\alpha\right)(X_1,\dots,X_m) = r\left(f(p)^{-1}\right)^{-1}(\alpha_p(X_1,\dots,X_m))\\ &= r(f(p))(\alpha_p(X_1,\dots,X_m)),\end{aligned}$$

yielding (16.28). □

Corollary 16.29. *For $\omega \in \mathcal{C}(P)$ and $f \in C(P,G)$,*

$$\Omega^{\Phi(f)\cdot\omega} = f\Omega^{\omega}f^{-1}.$$

Proof. Using Proposition 16.6 (p. 369) and (16.28) where the representation r is $ad : G \to \mathrm{GL}(\mathfrak{g})$,

$$\begin{aligned}\Omega^{\Phi(f)\cdot\omega} &= \Omega^{\Phi(f)^{-1*}\omega} = d\Phi(f)^{-1*}\omega + \Phi(f)^{-1*}\omega\wedge\Phi(f)^{-1*}\omega\\ &= \Phi(f)^{-1*}d\omega + \Phi(f)^{-1*}(\omega\wedge\omega) = \Phi(f)^{-1*}\Omega^{\omega}\\ &= \Phi(f)\cdot\Omega^{\omega} = ad(f)\,\Omega^{\omega} = f\Omega^{\omega}f^{-1}.\end{aligned}$$

□

Corollary 16.30. *The pairing*

$$\langle\cdot,\cdot\rangle : \Omega^k(P\times_G W)\times\Omega^k(P\times_G W)\to C^{\infty}(M,\mathbb{R}),$$

of (16.15) is preserved under the action of $\mathrm{GA}(P)$ on $\Omega^k(P\times_G W)\cong\overline{\Omega}^k(P,W)$ in the sense that for $\alpha,\alpha'\in\Omega^k(P\times_G W)$,

$$\langle F\cdot\alpha, F\cdot\alpha'\rangle = \langle\alpha,\alpha'\rangle. \tag{16.29}$$

Proof. In the special case $k=0$, with $s,s'\in\overline{\Omega}^0(P,W)\cong\Omega^0(P\times_G W)$, and letting $F=\Phi(f)$, we have

$$\begin{aligned}\langle F\cdot s, F\cdot s'\rangle &= \langle\Phi(f)\cdot s,\Phi(f)\cdot s'\rangle\\ &= K(r(f)(s), r(f)(s')) = K(s,s') = \langle s,s'\rangle,\end{aligned}$$

since $r : G\to\mathrm{GL}(W)$ is orthogonal relative to K. For $s,s'\in\Omega^0(P\times_G W)$ and $\beta,\beta'\in\Omega^k(P\times_G W)$, for the basic forms $s\otimes\beta$ and $s'\otimes\beta'$, we have

$$\begin{aligned}&\langle F\cdot(s\otimes\beta), F\cdot(s'\otimes\beta')\rangle = \langle(F\cdot s)\otimes\beta,(F\cdot s')\otimes\beta'\rangle\\ &= \langle F\cdot s, F\cdot s'\rangle\, h(\beta,\beta') = \langle s,s'\rangle\, h(\beta,\beta') = \langle s\otimes\beta, s'\otimes\beta'\rangle.\end{aligned}$$

For arbitrary $\alpha, \alpha' \in \Omega^k(P \times_G W)$, (16.29) follows by linearity. □

COROLLARY 16.31. *For* $F \in \mathrm{GA}(P)$ *and* $\beta \in \overline{\Omega}^{k+1}(P, W)$, *we have*

$$F \cdot (\delta^\omega \beta) = \delta^{F \cdot \omega}(F \cdot \beta).$$

PROOF. In view of Corollary 16.30, the global inner product $(\cdot, \cdot)$ of (16.17) is also preserved by the action of $\mathrm{GA}(P)$. Thus, for all $\alpha \in \overline{\Omega}^k(P, W)$, we have

$$\begin{aligned}&\left(F \cdot \alpha, \delta^{F \cdot \omega}(F \cdot \beta)\right) = \left(D^{F \cdot \omega}(F \cdot \alpha), F \cdot \beta\right)\\&= (F \cdot D^\omega \alpha, F \cdot \beta) = (D^\omega \alpha, \beta) = (\alpha, \delta^\omega \beta)\\&= (F \cdot \alpha, F \cdot \delta^\omega \beta),\end{aligned}$$

and it follows that $F \cdot (\delta^\omega \beta) = \delta^{F \cdot \omega}(F \cdot \beta)$. □

Let

$$C(P, \mathfrak{g}) := \overline{\Omega}^0(P, \mathfrak{g}) = \left\{ s \in C^\infty(P, \mathfrak{g}) : s(pg) = ad_{g^{-1}} s(p) \right\}.$$

There is a map Exp: $C(P, \mathfrak{g}) \to C(P, G)$, defined for $s \in C(P, \mathfrak{g})$ by

$$\mathrm{Exp}(s)(p) := \exp(s(p)) = \sum_{m=0}^{\infty} \frac{1}{m!} s(p)^m. \tag{16.30}$$

Note that $s(p) \in \mathfrak{g} \subseteq \mathfrak{gl}(N, \mathbb{C})$ so that $s(p)^m$ makes sense, and $\mathrm{Exp}(s) \in C(P, G)$ since

$$\begin{aligned}\mathrm{Exp}(s)(pg) &= pg \exp(s(pg)) = pg \exp\left(g^{-1} s(p) g\right)\\&= pgg^{-1} \exp(s(p)) g = p \exp(s(p)) g = \mathrm{Exp}(s)(p) g.\end{aligned}$$

Thus, we also have

$$\Phi \circ \mathrm{Exp} : C(P, \mathfrak{g}) \to \mathrm{GA}(P, G).$$

Just as $\exp : \mathfrak{g} \to G$ is a local diffeomorphism on a neighborhood of $0 \in \mathfrak{g}$ to a neighborhood of $I \in G$, relative to the C^k topology ($k \geq 0$), $\Phi \circ \mathrm{Exp}$ is a continuous bijection of a neighborhood of $0 \in C(P, \mathfrak{g})$ to a neighborhood of $\mathrm{Id} \in \mathrm{GA}(P, G)$. One might think of $C(P, \mathfrak{g})$ as the "Lie algebra" of the "Lie group" $\mathrm{GA}(P, G)$. Then corresponding to the group representation of $\mathrm{GA}(P, G)$ on $\overline{\Omega}^k(P, W)$, we have a Lie algebra representation given, for $s \in C(P, \mathfrak{g})$ and $\alpha \in \overline{\Omega}^k(P, W)$, by

$$\begin{aligned}&(s \cdot \alpha)_p (X_1, \ldots, X_m) := \tfrac{d}{dt} \left((\Phi(\mathrm{Exp}(ts)) \cdot \alpha)_p (X_1, \ldots, X_m) \right)\Big|_{t=0}\\&= \tfrac{d}{dt} \left(r(\mathrm{Exp}(ts)(p)) \alpha(X_1, \ldots, X_m) \right)\Big|_{t=0}\\&= \tfrac{d}{dt} \left(r(\exp(ts(p))) \alpha(X_1, \ldots, X_m) \right)\Big|_{t=0}\\&= r'(s(p)) (\alpha(X_1, \ldots, X_m)).\end{aligned}$$

There is also an infinitesimal version of the action of $C(P, \mathfrak{g})$ on $\mathcal{C}(P)$ defined by

$$s \cdot \omega := \tfrac{d}{dt} (\Phi(\mathrm{Exp}(ts)) \cdot \omega)\Big|_{t=0}.$$

PROPOSITION 16.32. *For* $s \in C(P, \mathfrak{g})$ *and* $\omega \in \mathcal{C}(P)$,

$$s \cdot \omega = -(ds + [\omega, s]) = -D^\omega s \in \overline{\Omega}^1(P, \mathfrak{g}).$$

PROOF. For $X \in T_pP$,

$$\begin{aligned} s \cdot \omega &= \tfrac{d}{dt} \left(\Phi \left(\operatorname{Exp}(ts) \right) \cdot \omega \right)(X) \Big|_{t=0} \\ &= \tfrac{d}{dt} \left(\operatorname{Exp}(ts)(p) \left(\operatorname{Exp}(ts)^{-1} \right)_* (X) + \operatorname{Exp}(ts)\, \omega(X) \operatorname{Exp}(ts)^{-1} \right) \Big|_{t=0} \\ &= -ds(X) - [\omega(X), s]. \end{aligned}$$

We computed the derivative of the first term as follows. For $\gamma : \mathbb{R} \to P$ a curve with $\gamma'(0) = X \in T_pP$, we have (at $t = u = 0$)

$$\begin{aligned} &\tfrac{d}{dt} \left(\operatorname{Exp}(ts)(p) \left(\operatorname{Exp}(ts)^{-1} \right)_* (X) \right) \\ &= \tfrac{d}{dt} \left(\exp(ts(p)) \tfrac{d}{du} \exp(-ts(\gamma(u))) \right) \\ &= \tfrac{d}{dt} \left(\exp(ts(p)) \left(-tds(\gamma'(0)) \right) \right) \\ &= \tfrac{d}{dt} \left(\exp(ts(p)) \left(-tds(X) \right) \right) \\ &= \tfrac{d}{dt} \left(\exp(ts(p)) \right)(0) + I \tfrac{d}{dt} \left(-tds(X) \right) = -ds(X). \end{aligned}$$

□

EXERCISE 16.33. Let $\pi : P \to M$ be a principal G-bundle. Suppose that U_α is open in M and $T_\alpha : \pi_{-1}(U_\alpha) \to U_\alpha \times G$ is a local trivialization of P, so that

$$T_\alpha(p) = (\pi(p), s_\alpha(p)), \text{ where } s_\alpha(pg) = s_\alpha(p)\, g.$$

Similarly let $T_\beta = \pi \times s_\beta : \pi_{-1}(U_\beta) \to U_\beta \times G$ be another local trivialization with $U_\alpha \cap U_\alpha \neq \phi$. Define a local section $\sigma_\alpha : U_\alpha \to \pi^{-1}(U_\alpha)$ by $\sigma_\alpha(x) = T_\alpha^{-1}(x, e)$, where e is the identity of G, and similarly define $\sigma_\beta : U_\beta \to \pi^{-1}(U_\beta)$.
(a) Show that $T_\alpha(\sigma_\alpha(x)\, g) = (x, g)$ and $\left(T_\alpha \circ T_\beta^{-1} \right) : (U_\alpha \cap U_\beta) \times G \hookleftarrow$ is given by

$$\left(T_\alpha \circ T_\beta^{-1} \right)(x, g') = (x, g_{\alpha\beta}(x)\, g'),$$

where $g_{\alpha\beta} : U_\alpha \cap U_\beta \to G$ is a well-defined function (known as a **transition function for** P) given by

$$g_{\alpha\beta}(\pi(p)) := s_\alpha(p)\, s_\beta(p)^{-1}.$$

Conclude that $T_\alpha \circ T_\beta^{-1}$ is a gauge transformation of the trivial principal G-bundle $(U_\alpha \cap U_\beta) \times G \to G$.
(b) Let $\omega \in \Lambda^1(P, \mathfrak{g})$ be a connection 1-form on P. Show that

$$\begin{aligned} \sigma_\beta(x) &= \sigma_\alpha(x)\, g_{\alpha\beta}(x), \\ {\sigma_\beta}^* \omega &= g_{\alpha\beta}^{-1} (\sigma_\alpha^* \omega)\, g_{\alpha\beta} + g_{\alpha\beta}^{-1} dg_{\alpha\beta}, \text{ and} \\ {\sigma_\beta}^* \Omega^\omega &= g_{\alpha\beta}^{-1} (\sigma_\alpha^* \Omega^\omega)\, g_{\alpha\beta}. \end{aligned}$$

5. Curvature in Riemannian Geometry

Let M be a C^∞ n-manifold. A *linear frame* of M at $x \in M$ is an isomorphism $u : \mathbb{R}^n \to T_x M$. Note that if $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$, then $u(e_1), \ldots, u(e_n)$ is a basis of $T_x M$. Let LM_x denote the set of all linear frames at x, let $LM = \cup_{x \in M} LM_x$, and let $\pi : LM \to M$ be the map $u \mapsto x$. For $A \in \mathrm{GL}(n, \mathbb{R})$ and $u \in LM_x$, we have $u \circ A \in LM_x$, with $u \circ A = u$ only if $A = I$. Thus, $\mathrm{GL}(n, \mathbb{R})$ acts freely to the right on LM. Indeed, LM can by made into a C^∞ $(n + n^2)$-manifold, such that $\pi : LM \to M$ is a principal $\mathrm{GL}(n, \mathbb{R})$-bundle, known as the *bundle of linear frames for M*. The reader is probably familiar with tensor representations of $\mathrm{GL}(n, \mathbb{R})$. For example, let $S^2(\mathbb{R}^n) \subset (\mathbb{R}^n)^* \otimes (\mathbb{R}^n)^*$ be the space of symmetric, bilinear forms on $\mathbb{R}^n$. We have the tensor representation $r : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(S^2(\mathbb{R}^n))$ given, for $s \in S^2(\mathbb{R}^n)$ and $A \in \mathrm{GL}(n, \mathbb{R})$, by

$$(r(A)(s))(v_1, v_2) := s\left(A^{-1}v_1, A^{-1}v_2\right) \quad \text{for} \quad v_1, v_2 \in \mathbb{R}^n. \tag{16.31}$$

Using the abbreviation $G_n = \mathrm{GL}(n, \mathbb{R})$, the associated bundle $LM \times_{G_n} S^2(\mathbb{R}^n) \to M$ is the bundle of symmetric, bilinear forms on the tangent spaces of M. For $[u, s] \in LM \times_{G_n} S^2(\mathbb{R}^n)$ and $X, Y \in T_x M$, note that $[u, s](X, Y) := s(u(X), u(Y))$ defines a bilinear form on $T_x M$. In the case of the "defining representation" $r : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$, namely the identity map $A \mapsto A$, we have the isomorphism

$$LM \times_{G_n} \mathbb{R}^n \cong TM \text{ given by } [u, v] \mapsto u(v), \tag{16.32}$$

for $u \in LM$ and $v \in \mathbb{R}^n$.

Definition 16.34. The **canonical** 1-**form** on LM is the element $\varphi \in \overline{\Omega}^1(LM, \mathbb{R}^n)$ defined for $u \in LM$ and $X \in T_u LM$ by

$$\varphi_u(X) := u^{-1}(\pi_*(X)).$$

We know (see (16.11)) that $\overline{\Omega}^1(LM, \mathbb{R}^n)$ is isomorphic to $\Omega^1(LM \times_{G_n} \mathbb{R}^n) = \Omega^1(TM)$, the space of 1-forms on M with values in TM, or in other words, the space of endomorphisms of TM. The canonical 1-form φ corresponds to the identity endomorphism. Indeed, for $u \in LM$ and $X \in T_u LM$, we have

$$\varphi_M(\pi_* X) := [u, \varphi_u(X)] = \left[u, u^{-1}(\pi_*(X))\right] = u\left(u^{-1}(\pi_*(X))\right) = \pi_*(X),$$

where we have used the identification (16.32).

A connection 1-form ω for the principal $\mathrm{GL}(n, \mathbb{R})$-bundle $\pi : LM \to M$ is called a *linear connection for M*. The *torsion Φ of ω* is the covariant derivative of the canonical 1-form with respect to ω, namely

$$\Phi := D^\omega \varphi \in \overline{\Omega}^2(LM, \mathbb{R}^n) \cong \Omega^2(M, TM), \tag{16.33}$$

which (as indicated) can be regarded as a 2-form on M with values in TM. A linear connection gives us a way of differentiating a vector field, say Y on M, with respect to another vector field X, as follows. Let $\widetilde{Y}$ be the ω-horizontal lift of Y to LM. Then $\varphi\left(\widetilde{Y}\right) \in \overline{\Omega}^0(LM, \mathbb{R}^n)$, and $D^\omega\left(\varphi\left(\widetilde{Y}\right)\right) \in \overline{\Omega}^1(LM, \mathbb{R}^n) \cong \Omega^1(M, TM)$. Thus, we may regard $D^\omega\left(\varphi\left(\widetilde{Y}\right)\right)$ as a 1-form with values in TM. Evaluating this

1-form on X gives us a vector field, commonly denoted by $\nabla_X Y$. Explicitly, for any $u \in LM$, we have

$$\begin{aligned}(\nabla_X Y)_{\pi(u)} &= u\left(D^\omega\left(\varphi\left(\widetilde{Y}\right)\right)_u\left(\widetilde{X}\right)\right) = u\left(d\left(\varphi\left(\widetilde{Y}\right)\right)_u\left(\widetilde{X}\right)\right) \\ &= u\left(\widetilde{X}_u[\varphi\left(\widetilde{Y}\right)]\right) \in T_{\pi(u)}M. \end{aligned} \tag{16.34}$$

EXERCISE 16.35. (a) Show that for vector fields $X, Y \in C^\infty(TM)$ and $f \in C^\infty(M)$,

K_1. $\nabla_{fX} Y = f\nabla_X Y$ and
K_2. $\nabla_X(fY) = df(X)Y + f\nabla_X Y$.

(b) A Kozul connection ∇ for M is defined to be a map $C^\infty(TM) \times C^\infty(TM) \to C^\infty(TM)$ written as $(X, Y) \mapsto \nabla_X Y$ such that K_1 and K_2 hold. Show that a Kozul connection arises from a unique *linear connection for* M.
(c) Let M be a submanifold of $\mathbb{R}^{n+k}$. For $X, Y \in C^\infty(TM)$, say $X = (X_1, \ldots, X_{n+k})$ and $Y = (Y_1, \ldots, Y_{n+k})$, where $X_i, Y_i \in C^\infty(M)$, define

$$X[Y] := (X[Y_1], \ldots, X[Y_{n+k}]),$$

where $X[Y_i]_x := \frac{d}{dt} Y_i(\gamma(t))\big|_{t=0}$ and γ is a curve in M with $\gamma'(0) = X_x$. Verify that the following defines a Kozul connection for M:

$$(\nabla_X Y)_x := P_x(X[Y]) \quad \text{for any } x \in M,$$

where P_x is orthogonal projection of $\mathbb{R}^{n+k}$ onto $T_x M$.

The torsion Φ of ω is 0, if and only if $\nabla_X Y - \nabla_Y X - [X, Y] = 0$ for all vector fields X and Y. Indeed, for $u \in LM$, we have

$$\begin{aligned} u\left(\Phi\left(\widetilde{X}, \widetilde{Y}\right)\right) &= u\left((D^\omega\varphi)\left(\widetilde{X}, \widetilde{Y}\right)\right) \\ &= u\left(\widetilde{X}\left(\varphi\left(\widetilde{Y}\right)\right)\right) - u\left(\widetilde{Y}\left(\varphi\left(\widetilde{X}\right)\right)\right) - u\left(\varphi\left(\left[\widetilde{X}, \widetilde{Y}\right]\right)\right) \\ &= \nabla_X Y - \nabla_Y X - \pi_*\left(\left[\widetilde{X}, \widetilde{Y}\right]\right) = \nabla_X Y - \nabla_Y X - [X, Y]. \end{aligned} \tag{16.35}$$

We can similarly express the curvature Ω^ω in terms of ∇. First note that

$$\begin{aligned} \Omega^\omega\left(\widetilde{X}, \widetilde{Y}\right) &= d\omega\left(\widetilde{X}, \widetilde{Y}\right) = \widetilde{X}\left[\omega\left(\widetilde{Y}\right)\right] - \widetilde{Y}\left[\omega\left(\widetilde{X}\right)\right] - \omega\left(\left[\widetilde{X}, \widetilde{Y}\right]\right) = -\omega\left(\left[\widetilde{X}, \widetilde{Y}\right]\right) \\ &\Rightarrow \left[\widetilde{X}, \widetilde{Y}\right]_u^V = -\Omega_u^\omega\left(\widetilde{X}, \widetilde{Y}\right)^*. \end{aligned}$$

Then recall that for $A \in \mathfrak{gl}(n, \mathbb{R}^n)$ and A^* the fundamental vertical vector field on LM, we have

$$A^*\left[\varphi\left(\widetilde{Z}\right)\right] = \frac{d}{dt}\left[\varphi_{u\exp(tA)}\left(\widetilde{Z}\right)\right] = \frac{d}{dt}\left(\exp(tA)^{-1}\right)\varphi_u\left(\widetilde{Z}\right) = -A\varphi_u\left(\widetilde{Z}\right).$$

Thus,

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$
$$= u\left(\widetilde{X}_u\left[\widetilde{Y}\left[\varphi\left(\widetilde{Z}\right)\right]\right]\right) - u\left(\widetilde{Y}_u\left[\widetilde{X}\left[\varphi\left(\widetilde{Z}\right)\right]\right]\right) - u\left(\left[\widetilde{X},\widetilde{Y}\right]^H\left[\varphi\left(\widetilde{Z}\right)\right]\right)$$
$$= u\left(\left(\left[\widetilde{X},\widetilde{Y}\right] - \left[\widetilde{X},\widetilde{Y}\right]^H\right)\left[\varphi\left(\widetilde{Z}\right)\right]\right) = u\left(\left[\widetilde{X},\widetilde{Y}\right]^V\left[\varphi\left(\widetilde{Z}\right)\right]\right)$$

(16.36)
$$= u\left(-\Omega_u^\omega\left(\widetilde{X},\widetilde{Y}\right)^*\left[\varphi\left(\widetilde{Z}\right)\right]\right) = u\left(\Omega^\omega\left(\widetilde{X}_u,\widetilde{Y}_u\right)\varphi_u\left(\widetilde{Z}\right)\right) =: \Omega^\omega(X,Y)(Z),$$

where the final equality defines what it means to regard the curvature form $\Omega^\omega \in \overline{\Omega}^2(LM, \mathfrak{gl}(n,\mathbb{R}^n))$ as being in $\Omega^2(\operatorname{End} TM)$.

Now, suppose that h is a Riemannian metric on M. Then $u \in LM_x$ is called an *orthonormal frame* if $u : \mathbb{R}^n \to T_xM$ is an isometry where $\mathbb{R}^n$ has its standard inner product (i.e., $h(u(v), u(w)) = v \cdot w = v^1w^1 + \cdots + v^nw^n$). The set FM of all orthonormal frames at all points of M is the total space of a principal $\mathrm{O}(n)$-bundle $\pi : FM \to M$ called the *orthonormal frame bundle of* M relative to h. Note that FM is a submanifold of LM.

Definition 16.36. If ω is a linear connection for M whose horizontal subspaces (those subspaces annihilated by ω) at points of FM are contained in tangent spaces of FM, then ω is called a **metric connection** (relative to h).

Note that $\omega|_{FM}$ is automatically a connection 1-form for the principal $\mathrm{O}(n)$-bundle $FM \to M$, but the horizontal subspace of ω at $u \in FM$ is not necessarily $\operatorname{Ker}(\omega|_{T_uFM})$; i.e., ω is not necessarily metric. There is another useful characterization of metric connections described as follows. The metric h on M corresponds to some $H \in \overline{\Omega}^0\left(LM, S^2(\mathbb{R}^n)\right)$, where $r : \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}\left(S^2(\mathbb{R}^n)\right)$ is the tensor representation for the space $S^2(\mathbb{R}^n)$ of symmetric bilinear forms, defined in (16.31). Indeed for $u \in LM$ and $v_1, v_2 \in \mathbb{R}^n$,

$$H(u)(v_1,v_2) := h(u(v_1), u(v_2)). \tag{16.37}$$

For $u \in FM$, $H(u)(v_1,v_2) = v_1 \cdot v_2$, so that $H|_{FM}$ is constant, namely the usual dot product. In fact, if ι denotes the usual dot product, $FM = H^{-1}(\iota)$. Let ω be a connection on 1-form on LM. According to (16.8),

$$D^\omega H = dH + r'(\omega)(H).$$

Proposition 16.37. *The following are equivalent*

1. *ω is a metric connection*
2. *$D^\omega H = 0$ on LM*
3. *$X[h(Y,Z)] = h(\nabla_X Y, Z) + h(Y, \nabla_X Z)$, for all vector fields X, Y and Z.*

Proof. Since $H|_{FM}$ is constant, for $X \in T_uFM$, we have $dH(X) = 0$ and

$$((D^\omega H)(X))(v_1,v_2) = dH(X)(v_1,v_2) + (r'(\omega(X))(H))(v_1,v_2)$$
$$= -(\omega(X)v_1)\cdot v_2 - v_1\cdot(\omega(X)v_2)$$

Since the left side is 0 for X vertical and the right side is 0 for X horizontal, both sides are zero for all $X \in T_uFM$, namely

$$(D^\omega H)(X) = 0 \text{ and } (\omega(X)v_1)\cdot v_2 + v_1\cdot(\omega(X)v_2) = 0 \tag{16.38}$$

Suppose that ω is metric. Then for any $u \in FM$, all horizontal $Y \in T_uLM$ are in T_uFM, and hence $(D^\omega H)(Y) = 0$ by (16.38). Since $\mathrm{GL}(n, \mathbb{R})$ acts transitively via R_{g*} on the set of all horizontal subspaces of LM at points of a given fiber, we then have that $(D^\omega H)(Y) = 0$ for all horizontal $Y \in T_uLM$ for all $u \in LM$. Since $(D^\omega H)(Y) = 0$ for all vertical Y, we then have ω metric $\Rightarrow D^\omega H = 0$ on LM. Now, suppose that $D^\omega H = 0$ on LM, and $\omega(Y) = 0$ for $Y \in T_uLM$. To show that ω is metric, we need to prove that $Y \in T_uFM$. However,

$$0 = (D^\omega H)(Y) = dH(Y) + (r'(\omega(Y))(H)) = dH(Y).$$

Thus, as dH has maximal rank at $u \in FM$ (Exercise), $Y \in T_u\left(H^{-1}(\iota)\right) = T_uFM$. For (3) note that for $x = \pi(u)$,

$$\begin{aligned}
X_x[h(Y,Z)] &= \widetilde{X}_u\left[H\left(\varphi\left(\widetilde{Y}\right), \varphi\left(\widetilde{Z}\right)\right)\right] = \widetilde{X}_u[H]\left(\varphi\left(\widetilde{Y}\right), \varphi\left(\widetilde{Z}\right)\right) \\
&+ H_u\left(\widetilde{X}_u\left[\varphi\left(\widetilde{Y}\right)\right], \varphi\left(\widetilde{Z}\right)\right) + H_u\left(\varphi\left(\widetilde{Y}\right), \widetilde{X}_u\left[\varphi\left(\widetilde{Z}\right)\right]\right) \\
&= (D^\omega H)_u\left(\varphi\left(\widetilde{Y}\right), \varphi\left(\widetilde{Z}\right)\right) + h_x(\nabla_X Y, Z) + h_x(Y, \nabla_X Z).
\end{aligned}$$

□

Proposition 16.38 (Fundamental Lemma of Riemannian Geometry). *For each Riemannian metric h on a manifold M, there is a unique metric linear connection 1-form ω on LM relative to h with torsion 0 (i.e., $\Phi = D^\omega\varphi = 0$).*

Remark 16.39. The connection ω in this proposition restricts to an $\mathfrak{so}(n)$-valued connection $\theta := \omega|_{FM}$ which we call the **Levi Civita connection for** h. The existence and uniqueness proof below is also valid for nondegenerate, indefinite metrics h. One just replaces the standard dot product on $\mathbb{R}^n$ by a standard, nondegenerate, indefinite scalar product.

Proof. By (16.8), for the canonical 1-form φ, we have

$$D^\omega\varphi = d\varphi + \omega \wedge \varphi,$$

or more precisely, for all $u \in LM$ and $W_1, W_2 \in T_uLM$,

$$D^\omega\varphi(W_1, W_2) = d\varphi(W_1, W_2) + \omega(W_1)\varphi(W_2) - \omega(W_2)\varphi(W_1)$$

where $\omega(W_1) \in \mathfrak{gl}(n, \mathbb{R})$ is operating on $\varphi(W_2) \in \mathbb{R}^n$. A vector field Z on M corresponds to some $Z' \in \overline{\Omega}^0(LM, \mathbb{R}^n)$ via $Z'(u) = u^{-1}(Z)$. Let $\widetilde{Z}$ be the unique vector field on LM such that $\pi_*\widetilde{Z} = Z$ and $\widetilde{Z}$ is the horizontal lift of Z relative to some arbitrary, fixed connection ω_0 on LM (i.e., $\omega_0\left(\widetilde{Z}\right) = 0$). From the fact that horizontal subspaces are sent to horizontal subspaces by R_{g*}, it follows that $\widetilde{Z}$ is invariant under R_{g*}. Moreover, we have $\varphi\left(\widetilde{Z}\right) = Z'$. Indeed, at each $u \in LM$,

$$Z'(u) = u^{-1}\left(Z_{\pi(u)}\right) = u^{-1}(\pi_*(\widetilde{Z}_u)) = \varphi_u(\widetilde{Z}_u).$$

Let X and Y be vector fields on M. Then for an arbitrary connection 1-form ω on LM

$$\begin{aligned}
D^\omega\varphi(\widetilde{X}, \widetilde{Y}) &= d\varphi(\widetilde{X}, \widetilde{Y}) + \omega(\widetilde{X})\varphi(\widetilde{Y}) - \omega(\widetilde{Y})\varphi(\widetilde{X}) \\
&= d\varphi(\widetilde{X}, \widetilde{Y}) + \omega(\widetilde{X})Y' - \omega(\widetilde{Y})Y',
\end{aligned}$$

or for all vector fields Z on M with corresponding $Z' \in \overline{\Omega}^0(LM, \mathbb{R}^n)$,

$$\begin{aligned} & H(D^\omega \varphi(\widetilde{X}, \widetilde{Y}), Z') \\ & = H(d\varphi(\widetilde{X}, \widetilde{Y}), Z') + H(\omega(\widetilde{X})Y', Z') - H(\omega(\widetilde{Y})X', Z'). \end{aligned} \tag{16.39}$$

We also have

$$\begin{aligned} (D^\omega H)(\widetilde{Z})(X', Y') &= dH(\widetilde{Z})(X', Y') + r'(\omega(\widetilde{Z}))(H)(X', Y') \\ &= dH(\widetilde{Z})(X', Y') - H(\omega(\widetilde{Z})X', Y') - H(X', \omega(\widetilde{Z})Y') \end{aligned} \tag{16.40}$$

From (16.39) and (16.40), we see that if $D^\omega \varphi = 0$ and $D^\omega H = 0$, then

$$H(d\varphi(\widetilde{X}, \widetilde{Y}), Z') = H(\omega(\widetilde{Y})X', Z') - H(\omega(\widetilde{X})Y', Z')$$

and

$$dH(\widetilde{Z})(X', Y') = H(\omega(\widetilde{Z})X', Y') + H(X', \omega(\widetilde{Z})Y').$$

We want to solve for $H(\omega(\widetilde{X})Y', Z')$. This is accomplished by forming the following mysterious combination which one can derive from Young diagrams for 3-tensors (see [**We46**]), or (with a little luck) by trial and error.

$$\begin{aligned} & dH(\widetilde{X})(Y', Z') + dH(\widetilde{Y})(X', Z') - dH(\widetilde{Z})(X', Y') \\ & - \Big(H(d\varphi(\widetilde{X}, \widetilde{Y}), Z') + H(d\varphi(\widetilde{Z}, \widetilde{X}), Y') + H(d\varphi(\widetilde{Z}, \widetilde{Y}), X')\Big) \\ & = \left(\begin{array}{c} \left(\begin{array}{c} H(\omega(\widetilde{X})Y', Z') + H(Y', \omega(\widetilde{X})Z') \\ +H(\omega(\widetilde{Y})X', Z') + H(X', \omega(\widetilde{Y})Z') \\ -H(\omega(\widetilde{Z})X', Y') - H(X', \omega(\widetilde{Z})Y') \end{array} \right) \\ - \left(\begin{array}{c} H(\omega(\widetilde{Y})X', Z') - H(\omega(\widetilde{X})Y', Z') \\ +H(\omega(\widetilde{X})Z', Y') - H(\omega(\widetilde{Z})X', Y') \\ +H(\omega(\widetilde{Y})Z', X') - H(\omega(\widetilde{Z})Y', X') \end{array} \right) \end{array} \right) \\ & = 2H(\omega(\widetilde{X})Y', Z'). \end{aligned}$$

Thus, we have

$$\begin{aligned} & 2H(\omega(\widetilde{X})Y', Z') \\ & = dH(\widetilde{X})(Y', Z') + dH(\widetilde{Y})(X', Z') - dH(\widetilde{Z})(X', Y') \\ & - \Big(H(d\varphi(\widetilde{X}, \widetilde{Y}), Z') + H(d\varphi(\widetilde{Z}, \widetilde{X}), Y') + H(d\varphi(\widetilde{Z}, \widetilde{Y}), X')\Big). \end{aligned} \tag{16.41}$$

Hence, if $D^\omega \varphi = 0$ and $D^\omega H = 0$, then ω is uniquely determined by (16.41). Conversely, suppose that we *define* ω by (16.41) and by the requirement that $\omega(A^*) = A$. Then using (16.41), it is straightforward to check that not only do we have that ω is torsion-free, namely

$$\begin{aligned} & H(D^\omega \varphi(\widetilde{X}, \widetilde{Y}), Z') \\ & = H(d\varphi(\widetilde{X}, \widetilde{Y}), Z') + H(\omega(\widetilde{X})Y', Z') - H(\omega(\widetilde{Y})X', Z') = 0, \end{aligned}$$

but also that ω is a metric connection, namely

$$\begin{aligned} & (D^\omega H)(\widetilde{Z})(X', Y') \\ & = dH(\widetilde{Z})(X', Y') - H\left(\omega(\widetilde{Z})X', Y'\right) - H\left(X', \omega(\widetilde{Z})Y'\right) = 0. \end{aligned}$$

Of course one should also check that ω defined by (16.41) and $\omega(A^*) = A$, also has the property $R_g^*\omega = g^{-1}\omega g$. This is automatic on vertical vectors, but we need

$$\left(R_g^*\omega\right)(\widetilde{X}_u) = g^{-1}\omega(\widetilde{X}_u)g.$$

for all vector fields X on M and $g \in \mathrm{GL}(n,\mathbb{R})$, or equivalently,

$$H(ug)\left(\left(R_g^*\omega\right)(\widetilde{X}_u)Y'(ug), Z'(ug)\right) = H(ug)\left(g^{-1}\omega(\widetilde{X}_u)gY'(ug), Z'(ug)\right) \tag{16.42}$$

for all vector fields X, Y, Z on M and $g \in \mathrm{GL}(n,\mathbb{R})$. For this we use

$$\begin{aligned} &H(ug)\left(\left(R_g^*\omega\right)(\widetilde{X}_u)Y'(ug), Z'(ug)\right) \\ &= H(ug)\left(\omega(R_{g*}\widetilde{X}_u)Y'(ug), Z'(ug)\right) \\ &= H(ug)\left(\omega(\widetilde{X}_{ug})Y'(ug), Z'(ug)\right) \end{aligned} \tag{16.43}$$

and

$$\begin{aligned} &H(ug)\left(g^{-1}\omega(\widetilde{X}_u)gY'(ug), Z'(ug)\right) \\ &= \left(r\left(g^{-1}\right)\cdot H(u)\right)\left(g^{-1}\omega(\widetilde{X}_u)gY'(ug), Z'(ug)\right) \\ &= H(u)\left(\omega(\widetilde{X}_u)gY'(ug), gZ'(ug)\right) \\ &= H(u)\left(\omega(\widetilde{X}_u)Y'(u), Z'(u)\right), \end{aligned} \tag{16.44}$$

and then apply (16.41) to the final expressions in (16.43) and (16.44). The result (16.42) then follows by showing that the $\mathbb{R}$-valued functions of the form $dH(\widetilde{X})(Y', Z')$ and $H\left(d\varphi(\widetilde{X}, \widetilde{Y}), Z'\right)$ are R_g invariant. As can be easily checked, this follows from the fact that H, φ and X' are equivariant, and $R_{g*}(\widetilde{X}) = \widetilde{X}$ (and similarly for Y and Z). Thus, the 1-form ω defined by (16.41) is a connection, and is the unique torsionless metric connection on LM relative to h. □

Let $\left(x^1, \ldots, x^n\right)$ be a system of local coordinates in a neighborhood U of M. Then the coordinate vector fields $\partial_k := \frac{\partial}{\partial x^k}$ yield a section $\sigma : U \to LM$ of the frame bundle given, for $v \in \mathbb{R}^n$ and $x \in U$, by $\sigma(x)(v) = \sum v^k (\partial_k)_x$. The image $\sigma_*(T_x M)$ is a subspace of $T_{\sigma(x)}LM$ which is a complement of the vertical subspace of $T_{\sigma(x)}LM$. As $A \in \mathrm{GL}(n,\mathbb{R})$ varies, the subspaces $H^0_{\sigma(x)A} := R_{A*}\left(T_{\sigma(x)}LM\right)$ then define a connection on $LM|_{\pi^{-1}(U)}$. This may locally serve to define the fixed connection ω_0 in the above proof. For a vector field $V(x) = \sum_{k=1}^n v(x)^k \partial_k$, note that $\widetilde{V}_{\sigma(x)A} := R_{A*}(\sigma_{*x}(V_x))$ is the ω_0-horizontal lift of V, and

$$\begin{aligned} V'_{\sigma(x)A} &:= \varphi\left(\widetilde{V}\right) = (\sigma(x)A)^{-1}(\pi_*(R_{A*}(\sigma_{*x}(V_x)))) = (\sigma(x)A)^{-1}(V_x) \\ &= A^{-1}\sigma(x)^{-1}(V_x) = A^{-1}v(x). \end{aligned}$$

We have $H(\sigma(x)A)(v,w) = h(\sigma(x)A(v), \sigma(x)A(v))$, and in particular

$$\begin{aligned} &h_{ij}(x) := h_x(\partial_i, \partial_j) = h(\sigma(x)(e_i), \sigma(x)(e_j)) = H(\sigma(x))(e_i, e_j) \text{ or} \\ &(\sigma^* H)(e_i, e_j) = \sigma^*(H(e_i, e_j)) = h_{ij}. \end{aligned}$$

We have $\widetilde{\partial_k}_{\sigma(x)} = \sigma_{*x}(\partial_k)$ and $(\partial'_k)_{\sigma(x)} = \varphi(\sigma_{*x}(\partial_k)) = \sigma(x)^{-1}(\partial_k) = e_k$. Thus, the result (see (16.41))

$$\begin{aligned}
&2H(\omega(\widetilde{e_i})e'_j, e'_k)\\
&= dH(\widetilde{e_i})\left(e'_j, e'_k\right) + dH(\widetilde{e_j})\left(e'_i, e'_k\right) - dH(\widetilde{e_k})\left(e'_i, e'_j\right)\\
&- \left(H(d\varphi(\widetilde{e_i}, \widetilde{e_j}), e'_k) + H(d\varphi(\widetilde{e_k}, \widetilde{e_i}), Y') + H(d\varphi(\widetilde{e_k}, \widetilde{e_j}), e'_i)\right)
\end{aligned} \tag{16.45}$$

becomes (on the image $\sigma(U)$)

$$\begin{aligned}
&2H((\omega(\sigma_{*x}(\partial_i))e_j, e_k)\\
&= dH(\widetilde{\partial}_i)(e_j, e_k) + dH(\widetilde{\partial}_j)(e_i, e_k) - dH(\widetilde{\partial_k})(e_i, e_j)\\
&- \Big(H_{\sigma(x)}(d\varphi(\widetilde{\partial}_i, \widetilde{\partial}_j), e_k) + H(d\varphi(\widetilde{\partial_k}, \widetilde{\partial}_i), Y') + H(d\varphi(\widetilde{\partial_k}, \widetilde{\partial}_j), e_i)\Big).
\end{aligned} \tag{16.46}$$

On the image $\sigma(U)$ again,

$$\begin{aligned}
H(d\varphi(\widetilde{\partial}_i, \widetilde{\partial}_j), e_k) &= H(\widetilde{\partial}_i\left[\varphi(\widetilde{\partial}_j)\right] - \widetilde{\partial}_j\left[\varphi(\widetilde{\partial}_i)\right] - \varphi\left(\left[\widetilde{\partial}_i, \widetilde{\partial}_j\right]\right), e_k)\\
&= H(\widetilde{\partial}_i[e_j] - \widetilde{\partial}_j[e_i] - \varphi(0), e_k) = 0,
\end{aligned}$$

and

$$\begin{aligned}
(dH)_{\sigma(x)}(\widetilde{\partial}_i)(e_j, e_k) &= d\left(H|_{\sigma(U)}\right)(e_j, e_k)(\sigma_{*x}(\partial_i))\\
&= \left(\sigma^* d\left(H|_{\sigma(U)}\right)\right)(\partial_i)(e_j, e_k)\\
&= \left(d\left(\sigma^*\left(H|_{\sigma(U)}\right)(e_j, e_k)\right)\right)(\partial_i)\\
&= \partial_i(h_{jk}).
\end{aligned}$$

Thus, letting $(\omega_i)_j^k$ be defined by

$$(\sigma^*\omega)(\partial_i)(e_j) = \omega(\sigma_*(\partial_i))(e_j) = \omega(\widetilde{e_i})(e_j) = \sum_k (\omega_i)_j^k e_k,$$

we have

$$\begin{aligned}
2H_{\sigma(x)}(\omega(\widetilde{\partial}_i)e_j, e_k) &= 2H_{\sigma(x)}(\sum_k (\omega_i)_j^l e_l, e_k) = 2H_{\sigma(x)}(\sum_k (\omega_i)_j^l e_l, e_k)\\
&= 2\sum_l h_{kl}(\omega_i)_j^l.
\end{aligned}$$

Hence, using the classical notation for the *Christoffel symbols* $\Gamma_{ij}^k := (\omega_i)_j^k$, we get

$$2\sum_l h_{kl}\Gamma_{ij}^l = 2\sum_l h_{kl}(\omega_i)_j^l = \partial_i[h_{jk}] + \partial_j[h_{ik}] - \partial_k[h_{ij}] \quad \text{or}$$

$$\Gamma_{ij}^l = (\omega_i)_j^l = \tfrac{1}{2}h^{lk}\left(\partial_i[h_{jk}] + \partial_j[h_{ik}] - \partial_k[h_{ij}]\right), \tag{16.47}$$

which is the classical formula. Note that Γ_{ij}^l is symmetric in i and j, and hence the Levi Civita connection is often called a symmetric connection. We now show that

$$\nabla_{\partial_i}\partial_j = \sum_{l=1}^n \Gamma_{ij}^l \partial_l. \tag{16.48}$$

Since $\varphi_{\sigma(x)}\left(\widetilde{\partial}_i\right) = e_i$ is constant, we have $d\left(\varphi\left(\widetilde{\partial}_i\right)\right) = 0$, and so

$$H_{\sigma(x)}\left(D^{\omega}\left(\varphi\left(\widetilde{\partial}_i\right)\right)\left(\widetilde{\partial}_j\right), e_k\right) = H_{\sigma(x)}\left(d\left(\varphi\left(\widetilde{\partial}_i\right)\right) + \omega\left(\widetilde{\partial}_j\right)\varphi\left(\widetilde{\partial}_i\right), e_k\right)$$
$$= H_{\sigma(x)}\left(\omega\left(\widetilde{\partial}_j\right)e_i, e_k\right) = \sum_l h_{kl}\left(\omega_j\right)_i^l = \sum_l h_{kl}\Gamma_{ji}^l.$$

Then for each k,

$$h_x\left(\sum_l \Gamma_{ji}^l \partial_l, \partial_k\right) = \sum_l \Gamma_{ji}^l h_x\left(\partial_l, \partial_k\right) = \sum_l h_{kl}\Gamma_{ji}^l$$
$$= H_{\sigma(x)}\left(D^{\omega}\left(\varphi\left(\widetilde{\partial}_i\right)\right)\left(\widetilde{\partial}_j\right), e_k\right)$$
$$= h_x\left(\sigma(x)\left(D^{\omega}\left(\varphi\left(\widetilde{\partial}_i\right)\right)\left(\widetilde{\partial}_j\right)\right), \sigma(x)(e_k)\right) = h_x\left(\nabla_{\partial_i}\partial_j, \partial_k\right),$$

from which (16.48) follows. Let $\mathbf{h}$ denote the matrix whose entries are h_{ij}. There is a key identity that we will use later, namely (using the summation convention)

$$h^{ij}\nabla_{\partial_i}\partial_j = -\sum_{l=1}^{n}\frac{1}{\sqrt{\det\mathbf{h}}}\partial_i\left(h^{li}\sqrt{\det\mathbf{h}}\right)\partial_l \quad \text{or}$$
$$h^{ij}\Gamma_{ij}^l = -\frac{1}{\sqrt{\det\mathbf{h}}}\partial_i\left[h^{li}\sqrt{\det\mathbf{h}}\right]. \tag{16.49}$$

This is based on the identity

$$\partial_k\left[\det\mathbf{h}\right] = \left(\det\mathbf{h}\right)\sum_{i,j=1}^{n} h^{ij}\partial_k\left[h_{ij}\right] \quad \text{or} \quad h^{ij}\partial_k\left[h_{ij}\right] = \partial_k\left[\log\det\mathbf{h}\right],$$

which is shown as follows. Let $\mathbf{h}_j$ be the j-th column of $\mathbf{h}$ and write the determinant as a multilinear function of its columns, say $\det\mathbf{h} = \det\left(\mathbf{h}_1, \dots, \mathbf{h}_n\right)$. Then (where we use Cramer's rule for the third equality)

$$\partial_k\left[\det\mathbf{h}\right] = \sum_{j=1}^{n}\det\left(\mathbf{h}_1, \dots, \mathbf{h}_{j-1}, \partial_k\left[\mathbf{h}_j\right], \mathbf{h}_{j+1}, \dots, \mathbf{h}_n\right)$$
$$= \left(\det\mathbf{h}\right)\sum_{j=1}^{n}\frac{1}{\det\mathbf{h}}\det\left(\mathbf{h}_1, \dots, \mathbf{h}_{j-1}, \partial_k\left(\mathbf{h}_j\right), \mathbf{h}_{j+1}, \dots, \mathbf{h}_n\right)$$
$$= \left(\det\mathbf{h}\right)\sum_{j=1}^{n}\left(\mathbf{h}^{-1}\partial_k\left(\mathbf{h}_j\right)\right)^j$$
$$= \left(\det\mathbf{h}\right)\sum_{i,j=1}^{n} h^{ji}\partial_k\left(h_{ij}\right) = \left(\det\mathbf{h}\right)\sum_{i,j=1}^{n} h^{ij}\partial_k\left(h_{ij}\right).$$

To obtain (16.49), we then compute (where a sum over i and j is implicit)

$$\begin{aligned} h^{ij}\Gamma^l_{ij} &= h^{ij}\Gamma^l_{ij} = \tfrac{1}{2}h^{ij}h^{lk}\left(\partial_i\left[h_{jk}\right] + \partial_j\left[h_{ik}\right] - \partial_k\left[h_{ij}\right]\right) \\ &= h^{ij}\left(h^{lk}\partial_i\left[h_{jk}\right] - \tfrac{1}{2}h^{lk}\partial_k\left[h_{ij}\right]\right) \\ &= h^{ij}\left(\partial_i\left[h^{lk}h_{jk}\right] - \partial_i\left[h^{lk}\right]h_{jk} - \tfrac{1}{2}h^{lk}\partial_k\left[h_{ij}\right]\right) \\ &= -\partial_i\left[h^{lk}\right]h^{ij}h_{jk} - \tfrac{1}{2}h^{lk}h^{ij}\partial_k\left[h_{ij}\right] = -\partial_i\left[h^{li}\right] - \tfrac{1}{2}h^{lk}h^{ij}\partial_k\left[h_{ij}\right] \\ &= -\partial_i\left[h^{li}\right] - \tfrac{1}{2}h^{lk}\partial_k\left(\log\det\mathbf{h}\right) = -\partial_i\left[h^{li}\right] - h^{lk}\partial_k\left(\log\sqrt{\det\mathbf{h}}\right) \\ &= \frac{-1}{\sqrt{\det\mathbf{h}}}\left(\sqrt{\det\mathbf{h}}\partial_i\left[h^{li}\right] + h^{lk}\partial_k\left[\sqrt{\det\mathbf{h}}\right]\right) \\ &= \frac{-1}{\sqrt{\det\mathbf{h}}}\partial_i\left[h^{li}\sqrt{\det\mathbf{h}}\right]. \end{aligned}$$

As in (16.4) and (16.6), the curvature of the Levi-Civita connection $\theta = \omega|_{FM}$ is

$$\Omega^\theta := D^\theta\theta = d\theta + \theta\wedge\theta = d\theta + \tfrac{1}{2}\left[\theta,\theta\right] \in \overline{\Omega}^2\left(FM, \mathfrak{so}\left(n\right)\right),$$

where $\mathfrak{so}\left(n\right) = \mathfrak{o}\left(n\right) := \left\{A \in \mathfrak{gl}\left(n,\mathbb{R}\right) : A^T = -A\right\}$ is the Lie algebra of $\mathrm{O}(n)$ (or $\mathrm{SO}\left(n\right)$). Let us denote the bundle of skew-symmetric (relative to h) endomorphisms of the tangent spaces T_xM by $\mathfrak{so}\left(TM\right)$. Then $\mathfrak{so}\left(TM\right) = FM \times_{O(n)} \mathfrak{so}\left(n\right)$ and

$$\overline{\Omega}^2\left(FM,\mathfrak{so}\left(n\right)\right) \cong \Omega^2\left(M, FM\times_{O(n)}\mathfrak{so}\left(n\right)\right) = \Omega^2\left(M,\mathfrak{so}\left(TM\right)\right)$$

Thus, we may regard Ω^θ as in $\Omega^2\left(M,\mathfrak{so}\left(TM\right)\right)$, and for vectors $X, Y \in T_xM$, $\Omega^\theta\left(X,Y\right) \in \mathfrak{so}\left(T_xM\right)$, we define $R \in C^\infty\left(M, \otimes^4\left(TM\right)^*\right)$ by

$$R\left(W,Z,X,Y\right) := h\left(\Omega^\theta\left(X,Y\right)\left(Z\right), W\right). \tag{16.50}$$

Note that $R\left(W,Z,X,Y\right)$ is antisymmetric in (X,Y) and in (W,Z). We can relate the curvature tensor to the Gaussian curvature of surfaces as follows. Let $X, Y \in T_xM$ be and let S be the surface composed of geodesic segments issuing from x with initial tangent vectors in the 2-plane $\Pi = span\left(X,Y\right) \subseteq T_xM$. Then the Gaussian curvature of S (with the induced metric) at x is given by

$$K\left(\Pi\right) := \frac{R\left(X,Y,X,Y\right)}{h\left(X,X\right)h\left(Y,Y\right) - h\left(X,Y\right)^2}, \tag{16.51}$$

which is independent of the choice of the basis $\{X,Y\}$ and is known as the *sectional curvature* of Π.

EXERCISE 16.40. Here we show that the curvature tensor of the unit n-sphere S^n (with metric tensor induced from $\mathbb{R}^{n+1}$) at the point $x \in S^n$ is given by

$$R\left(W,Z,X,Y\right) = \left\langle X,Z\right\rangle\left\langle Y,W\right\rangle - \left\langle Y,Z\right\rangle\left\langle X,W\right\rangle \tag{16.52}$$

where $W,Z,X,Y \in T_xS^n = x^\perp \subset \mathbb{R}^{n+1}$, and $\langle\cdot,\cdot\rangle$ is the usual dot product. This implies that all the sectional curvatures $K\left(\Pi\right)$ of S^n are equal to 1.
(a) Let $e_{n+1} = (0,\dots,0,1) \in S^n$. Show that the subgroup $\{g \in \mathrm{O}\left(n+1\right) : ge_{n+1} = e_{n+1}\}$ can be identified with $\mathrm{O}\left(n\right)$ and that the map $\pi : \mathrm{O}\left(n+1\right) \to S^n$ given by $\pi\left(\overline{g}\right) = \overline{g}e_{n+1}$ induces a diffeomorphism $\frac{\mathrm{O}(n+1)}{\mathrm{O}(n)} \cong S^n$. Hence, $\pi : \mathrm{O}\left(n+1\right) \to S^n$ is the special case of the bundle $\pi : \overline{G} \to \overline{G}/G$ in Exercise 16.5, where where $\overline{G} = \mathrm{O}\left(n+1\right)$ and $G = \mathrm{O}\left(n\right)$.
(b) Show that for $x \in S^n$ and $\overline{g} \in \pi^{-1}\left(x\right)$, we have $\overline{g}|_{\mathbb{R}^n} : \mathbb{R}^n \cong x^\perp = T_xS^n$ is an isometry. This shows that $\pi : \mathrm{O}\left(n+1\right) \to S^n$ may be regarded as the *orthonormal*

frame bundle $\pi : F(S^n) \to S^n$.
(c) Noting that any vector in $T_{\overline{g}}(F(S^n)) = T_{\overline{g}}(\mathrm{O}(n+1))$ is of the form $\overline{g}A$ for some $A \in \mathfrak{o}(n+1) = \{A \in \mathfrak{gl}(n+1, \mathbb{R}) : A^T = -A\}$, show that the canonical 1-form $\varphi \in \overline{\Omega}^1(F(S^n), \mathbb{R}^n)$ is given by

$$\varphi(\overline{g}A) = Ae_{n+1}$$

Note that $A \in \mathfrak{o}(n+1) \Rightarrow Ae_{n+1} \cdot e_{n+1} = 0 \Rightarrow Ae_{n+1} \in \mathbb{R}^n$.
(d) Show that

$$\begin{aligned} &\mathfrak{o}(n+1) = \mathfrak{o}(n) \oplus \mathfrak{m}, \text{ where} \\ &\mathfrak{m} := \{A \in \mathfrak{o}(n+1) : Ae_{n+1} \cdot e_{n+1} = 0\} \cong \mathbb{R}^n, \end{aligned}$$

and where the isomorphism $\mathfrak{m} \cong \mathbb{R}^n$ is defined by $A \mapsto Ae_{n+1}$, with inverse $\alpha : \mathbb{R}^n \to \mathfrak{m}$ given by

$$\alpha(x)(y) := \langle e_{n+1}, y\rangle x - \langle x, y\rangle e_{n+1} \text{ for } x, y \in \mathbb{R}^n.$$

Also verify that $ad_g\mathfrak{m} = g\mathfrak{m}g^{-1} = \mathfrak{m}$, and for all $x, y \in \mathbb{R}^n$

$$\begin{gathered} [\alpha(x), \alpha(y)](e_{n+1}) = 0 \text{ so that } [\alpha(x), \alpha(y)] \in \mathfrak{o}(n), \text{ while} \\ [\alpha(x), \alpha(y)](z) = \langle x, z\rangle y - \langle y, z\rangle x \text{ for } x, y, z \in \mathbb{R}^n. \end{gathered}$$

(e) According to Exercise 16.5 (p. 368) with $\overline{G} = \mathrm{O}(n+1)$ and $G = \mathrm{O}(n)$, the $\mathfrak{o}(n)$-valued 1-form

$$\omega := \pi_{\mathfrak{o}(n)} \circ \overline{\omega} \in \Omega^1(\mathrm{O}(n+1), \mathfrak{o}(n)) \quad (\text{where } \overline{\omega}(\overline{g}A) := A)$$

is a connection 1-form for $\pi : \mathrm{O}(n+1) \to \mathrm{O}(n+1)/\mathrm{O}(n)$ or $\pi : F(S^n) \to S^n$. Show that ω is the Levi-Civita connection for S^n (i.e., $D^\omega\varphi = 0$). For this it is best to evaluate $D^\omega\varphi$ on the pair horizontal vector fields $\overline{g} \mapsto \overline{g}\alpha(v)$ and $\overline{g} \mapsto \overline{g}\alpha(w)$ for some $v, w \in \mathbb{R}^n$.
(f) For $X \in T_xS^n = x^\perp \subseteq \mathbb{R}^{n+1}$, show that the horizontal lift $\widetilde{X}_{\overline{g}}$ of X at $\overline{g} \in \mathrm{O}(n+1) = FS^n$ is $\overline{g}\alpha(\overline{g}^{-1}X)$. Then, using the formula $\Omega^\omega(\widetilde{A}, \widetilde{B}) = \pi_{\mathfrak{g}}([A, B])$ of Exercise 16.7 (p. 370) and Part (d) above, show that for $X, Y, Z \in T_xS^n$,

$$R(X, Y)(Z) := \overline{g}\Omega^\omega_{\overline{g}}(\widetilde{X}, \widetilde{Y})(\varphi_{\overline{g}}(\widetilde{Z})) = \langle X, Z\rangle Y - \langle Y, Z\rangle X,$$

so that (16.52) holds.

REMARK 16.41. Note that under a dialation of the metric, say $h \mapsto ch$ for some $c > 0$, the Levi-Civita connection ∇ does not change, as is apparent from (16.48) and (16.47). Hence by virtue of (16.36), the curvature $\Omega^\omega \in \Omega^2(\mathrm{End}\, TM)$ does not change However, the curvature tensor $R(W, Z, X, Y) := h(\Omega^\theta(X, Y)(Z), W)$ changes by a factor of c, because of the involvement of the metric. Moreover, due to the factor of c^2 in the denominator of (16.51), the sectional curvatures then change by a factor of c^{-1}. Thus, a sphere of radius r (whose metric tensor is $c = r^2$ times that of the unit sphere) sectional curvatures all equal to r^{-2}, which is the Gaussian curvature of all its great 2-spheres.

Proposition 16.43 below implies (by linearity) that R is uniquely determined at a point x by its sectional curvatures at x, but first we need to derive the so-called First Bianchi Identity. Using the fact that θ is torsion-free and (16.13), we obtain

$$0 = D^\theta \Phi = D^\theta \left(D^\theta \varphi\right) = \Omega^\theta \wedge \varphi,$$

Regarding $\Omega^\theta \in \Omega^2 \left(M, \mathfrak{so}\left(TM\right)\right)$ and $\varphi = \mathrm{Id} \in \Omega^1 \left(M, \mathrm{End}\left(TM\right)\right)$, we then have

$$0 = \left(\Omega^\theta \wedge \varphi\right)(X, Y, Z) = \tfrac{1}{3}\left(\Omega^\theta (X, Y)(Z) + \Omega^\theta (Z, X)(Y) + \Omega^\theta (Y, Z)(X)\right)$$

Taking the h inner product with W yields the *First Bianchi Identity*, namely

$$R(W, Z, X, Y) + R(W, Y, Z, X) + R(W, X, Y, Z) = 0. \tag{16.53}$$

Exercise 16.42. From the antisymmetry of $R(W, Z, X, Y)$ in (W, Z) and in (X, Y) and the First Bianchi Identity identity (16.53), obtain the identity

$$R(W, Z, X, Y) = R(X, Y, W, Z). \tag{16.54}$$

[Hint. Add the equations (each a First Bianchi Identity)

$$\begin{array}{r} R(W,Z,X,Y) + R(W,Y,Z,X) + R(W,X,Y,Z) = 0 \\ R(Y,W,Z,X) + R(Y,X,W,Z) + R(Y,Z,X,W) = 0 \\ -R(X,Y,W,Z) - R(X,Z,Y,W) - R(X,W,Z,Y) = 0 \\ -R(Z,X,Y,W) - R(Z,W,X,Y) - R(Z,Y,W,X) = 0. \end{array}$$

and use the above antisymmetry.]

Proposition 16.43. *If $R(X, Y, X, Y) = 0$ for all $X, Y \in T_x M$, then $R = 0$ at x.*

Proof. We know that $R(W, X, Y, Z)$ is antisymmetric in (W, X) and in (Y, Z). Hence, if we knew that $R(W, X, Y, Z)$ is antisymmetric in any other pair, say (W, Y), then it would be antisymmetric in all pairs; e.g., for (X, Y),

$$R(W, X, Y, Z) = -R(X, W, Y, Z) = R(Y, W, X, Z) = -R(W, Y, X, Z).$$

The First Bianchi Identity would then yield the desired result

$$\begin{aligned} 0 &= R(W,X,Y,Z) + R(W,Y,Z,X) + R(W,Z,X,Y) \\ &= R(W,X,Y,Z) - R(W,X,Z,Y) - R(W,X,Z,Y) \\ &= R(W,X,Y,Z) + R(W,X,Y,Z) + R(W,X,Y,Z) \\ &= 3R(W,X,Y,Z). \end{aligned}$$

Thus, it remains to prove $R(Y, X, W, Z) = -R(W, X, Y, Z)$. By assumption,

$$\begin{aligned} 0 &= R(W, X+Z, W, X+Z) \\ &= R(W,X,W,X) + R(W,X,W,Z) + R(W,Z,W,X) + R(W,Z,W,Z) \\ &= R(W,X,W,Z) + R(W,Z,W,X) \\ &= 2R(W,X,W,Z) \text{ by (16.54).} \end{aligned}$$

Then, as required,

$$\begin{aligned} 0 &= R(Y+W, X, Y+W, Z) \\ &= R(Y,X,Y,Z) + R(W,X,Y,Z) + R(Y,X,W,Z) + R(W,X,W,Z) \\ &= R(W,X,Y,Z) + R(Y,X,W,Z). \end{aligned}$$

$\square$

The general Bianchi Identity of (16.12), yields $D^{\theta}\left(\Omega^{\theta}\right)=0$ which in the context of Levi-Civita connections is called the *Second Bianchi Identity.* In order to write this identity in terms of R, it is convenient to introduce the notion of the standard horizontal vector field $\overline{w}$ on FM associated to $w\in\mathbb{R}^n$

DEFINITION 16.44. Given $w\in\mathbb{R}^n$, the **standard horizontal vector field** $\overline{w}$ on FM associated to w assigns to each $u\in FM$ the unique vector $\overline{w}_u\in T_uFM$, such that $\theta\left(\overline{w}_u\right)=0$ and $\varphi\left(\overline{w}_u\right)=w$.

A standard horizontal vector field $\overline{w}$ is not R_{g*} invariant, since

$$\varphi\left(R_{g*}\overline{w}_u\right)=g^{-1}\varphi\left(\overline{w}_u\right)=g^{-1}w\Rightarrow R_{g*}\overline{w}_u=\left(\overline{g^{-1}w}\right)_{ug},\tag{16.55}$$

and so $\overline{w}$ is not a horizontal lift. One can also define $\overline{w}$ on LM relative to a given linear connection ω on LM.

For $x,y,z\in\mathbb{R}^n$, we let $\overline{x},\overline{y},\overline{z}$ be the associated standard horizontal vector fields. Now $D^{\theta}\varphi=0$ implies (using the fact that $\varphi\left(\overline{y}\right)=y$ is constant) that

$$0=\left(D^{\theta}\varphi\right)\left(\overline{x},\overline{y}\right)=\overline{x}\left[\varphi\left(\overline{y}\right)\right]-\overline{y}\left[\varphi\left(\overline{x}\right)\right]-\varphi\left(\left[\overline{x},\overline{y}\right]\right)=\varphi\left(\left[\overline{x},\overline{y}\right]\right),\tag{16.56}$$

In other words, the Lie bracket $\left[\overline{x},\overline{y}\right]$ is a vertical vector field. Indeed, from

$$\begin{aligned}\Omega^{\theta}\left(\overline{x},\overline{y}\right)&=\left(D^{\theta}\theta\right)\left(\overline{x},\overline{y}\right)=d\theta\left(\overline{x},\overline{y}\right)\\&=\overline{x}\left[\theta\left(\overline{y}\right)\right]-\overline{y}\left[\theta\left(\overline{x}\right)\right]-\theta\left(\left[\overline{x},\overline{y}\right]\right)=-\theta\left(\left[\overline{x},\overline{y}\right]\right),\end{aligned}\tag{16.57}$$

we see that the vertical part of $\left[\overline{x},\overline{y}\right]$ is $-\Omega^{\theta}\left(\overline{x},\overline{y}\right)$. Using (16.56), we have

$$\begin{aligned}0&=\tfrac{1}{2}D^{\theta}\left(\Omega^{\theta}\right)\left(\overline{x},\overline{y},\overline{z}\right)=\tfrac{1}{2}d\Omega^{\theta}\left(\overline{x},\overline{y},\overline{z}\right)\\&=\overline{x}\left[\Omega^{\theta}\left(\overline{y},\overline{z}\right)\right]+\overline{y}\left[\Omega^{\theta}\left(\overline{z},\overline{x}\right)\right]+\overline{z}\left[\Omega^{\theta}\left(\overline{x},\overline{y}\right)\right]\\&\quad+\Omega^{\theta}\left(\left[\overline{x},\overline{y}\right],\overline{z}\right)+\Omega^{\theta}\left(\left[\overline{z},\overline{x}\right],\overline{y}\right)+\Omega^{\theta}\left(\left[\overline{y},\overline{z}\right],\overline{x}\right)\\&=\overline{x}\left[\Omega^{\theta}\left(\overline{y},\overline{z}\right)\right]+\overline{y}\left[\Omega^{\theta}\left(\overline{z},\overline{x}\right)\right]+\overline{z}\left[\Omega^{\theta}\left(\overline{x},\overline{y}\right)\right]\\&=d\left(\Omega^{\theta}\left(\overline{y},\overline{z}\right)\right)\left(\overline{x}\right)+d\left(\Omega^{\theta}\left(\overline{z},\overline{x}\right)\right)\left(\overline{y}\right)+d\left(\Omega^{\theta}\left(\overline{x},\overline{y}\right)\right)\left(\overline{z}\right).\end{aligned}\tag{16.58}$$

We can identify $R\in C^{\infty}\left(M,\otimes^4\left(TM\right)^*\right)$ with $\mathfrak{R}\in\Omega^0\left(FM,\otimes^4\mathbb{R}^{n*}\right)$ defined, for $u\in FM$, by

$$\mathfrak{R}\left(u\right)\left(w,z,x,y\right)=\left(\Omega_u^{\theta}\left(\overline{x},\overline{y}\right)\left(z\right)\right)\cdot w,$$

Since

$$\left(D^{\theta}\mathfrak{R}\right)\left(\overline{v}\right)\left(w,z,x,y\right)=\left(\left(d\mathfrak{R}\right)\left(\overline{v}\right)\right)\left(w,z,x,y\right)=\left(d\left(\Omega^{\theta}\left(\overline{x},\overline{y}\right)\right)\left(\overline{v}\right)\left(z\right)\right)\cdot w,$$

we see from (16.58) that

$$\left(D^{\theta}\mathfrak{R}\right)\left(\overline{x}\right)\left(w,v,y,z\right)+\left(D^{\theta}\mathfrak{R}\right)\left(\overline{y}\right)\left(w,v,z,x\right)+\left(D^{\theta}\mathfrak{R}\right)\left(\overline{z}\right)\left(w,v,x,y\right)=0.\tag{16.59}$$

In terms of R this identity (the Second Bianchi Identity) is typically written as

$$\left(\nabla_XR\right)\left(W,V,Y,Z\right)+\left(\nabla_YR\right)\left(W,V,Z,X\right)+\left(\nabla_ZR\right)\left(W,V,X,Y\right)=0.\tag{16.60}$$

EXERCISE 16.45. A C^{∞} Riemannian manifold M with metric h has constant sectional curvaure at a point x if all of the 2-planes Π in T_xM have the same sectional curvature (automatic if $\dim M=2$.) Use (16.60) to show that if $\dim M\geq 3$, M is connected, and M has constant sectional curvature $K\left(x\right)$ at each $x\in M$, then in fact $K\left(x\right)$ is independent of x. This is known as Schur's Theorem. [Hint. First use Proposition 16.43 to deduce that at $x\in M$ $R\left(W,Z,X,Y\right)=$

$K(x)(h(X,Z)h(Y,W) - h(Y,Z)h(X,W))$. Then apply (16.60), and choose $W = Z$ and $V = Y$ with X, Y, Z orthogonal.]

DEFINITION 16.46. The **Ricci curvature** of h is the contraction (trace) of R in the first and third slots, namely for an orthonormal basis $E_1, \ldots, E_n$ of $T_x M$ and $X, Y \in T_x M$,

$$Ric(X,Y) := \sum_{i=1}^{n} R(E_i, X, E_i, Y) = \sum_{i=1}^{n} R(E_i, Y, E_i, X) = Ric(Y,X), \tag{16.61}$$

where we have used (16.54) to deduce that Ric is a symmetric 2-tensor. The **scalar curvature** S is the contraction of Ric, namely

$$S = \sum_{i=1}^{n} Ric(E_i, E_i).$$

EXERCISE 16.47. The Einstein tensor of h is defined to be $Ric - \frac{1}{2}Sh$. By contracting (16.60) in the pairs (W,Y) and (V,Z). Show that

$$0 = \operatorname{div}\left(Ric - \tfrac{1}{2}Sh\right) := \nabla_{E_i}\left(Ric - \tfrac{1}{2}Sh\right)(E_i, \cdot) \in \Omega^1(M).$$

When h has signature $(3,1)$ (e.g, when (M,h) is a space-time), the Einstein equation of general relativity is $Ric - \frac{1}{2}h = \frac{-8\pi K}{c^2}T$ (see (15.13)), where T is the symmetric stress-energy-momentum tensor which is known to be divergence-free by conservation of energy and momentum. The left side $Ric - \frac{1}{2}h$ of Einstein's equation is the most obvious geometric candidate for a divergence-free symmetric tensor.

We will need to study the abstract space $\mathcal{R}(\mathbb{R}^n)$ of all possible curvature tensors on $\mathbb{R}^n$, defined as follows. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. For $R \in \otimes^4(\mathbb{R}^n)^*$, let $R_{hijk} := R(e_h, e_i, e_j, e_k)$. We define $\mathcal{R}(\mathbb{R}^n)$ to be the set of all $R \in \otimes^4(\mathbb{R}^n)^*$, such that

$$\begin{array}{ll}\text{(A)} & R_{hijk} = -R_{ihjk} = R_{ihkj} \\ \text{(B)} & R_{hijk} + R_{hkij} + R_{hjki} = 0.\end{array} \tag{16.62}$$

Thus, $\mathcal{R}(\mathbb{R}^n)$ consists of those $R \in \otimes^4(\mathbb{R}^n)^*$ antisymmetric in the first pair and the last pair of indices, and which satisfy the First Bianchi Identity. We have seen that $R_{hijk} = R_{jkhi}$ then follows. Hence, $\mathcal{R}(\mathbb{R}^n)$ can be regarded as the subspace of the vector space $\mathcal{S}\left(\Lambda^2(\mathbb{R}^n)\right)$ of symmetric linear endomorphisms of $\Lambda^2(\mathbb{R}^n)$ which satisfy Condition B in (16.62). Since Condition B is automatic in $\mathcal{S}\left(\Lambda^2(\mathbb{R}^n)\right)$ except when h, i, j,and k are distinct,

$$\begin{aligned}&\dim(\mathcal{R}(\mathbb{R}^n)) = \dim\left(\mathcal{S}\left(\Lambda^2(\mathbb{R}^n)\right)\right) - \tbinom{n}{4} \\ &= \tfrac{1}{2}\tfrac{n(n-1)}{2}\left(\tfrac{n(n-1)}{2} + 1\right) - \tbinom{n}{4} = \tfrac{1}{12}n^2\left(n^2 - 1\right).\end{aligned}$$

Let $\mathcal{S}(\mathbb{R}^n)$ be the space of symmetric bilinear forms on $\mathbb{R}^n$. There is a linear map, which we call the *Ricci map*,

$$r : \mathcal{R}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \text{ given by } r(R) := \sum_{h=1}^{n} R_{hihk},$$

and a *scalar map*

$$s:\mathcal{R}\left(\mathbb{R}^n\right)\to\mathbb{R} \text{ given by } s\left(R\right):=Tr\left(R\right):=\sum_{i=1}^{n} r\left(R\right)_{ii}=\sum_{h,i=1}^{n} R_{hihi}\,.$$

Note that $\operatorname{Ker} r\subseteq \operatorname{Ker} s$, so that $\operatorname{Ker} s=\left(\operatorname{Ker} s\cap\left(\operatorname{Ker} r\right)^{\perp}\right)\oplus\operatorname{Ker} r$. Thus, we have a decomposition

$$\mathcal{R}\left(\mathbb{R}^n\right)=\mathcal{R}_1\oplus\mathcal{R}_2\oplus\mathcal{R}_3:=\left(\operatorname{Ker} s\right)^{\perp}\oplus\left(\operatorname{Ker} s\cap\left(\operatorname{Ker} r\right)^{\perp}\right)\oplus\operatorname{Ker} r. \tag{16.63}$$

The tensor representation $\mathrm{O}(n)\to\mathrm{GL}\left(\mathcal{R}\left(\mathbb{R}^n\right)\right)$ is given for $g\in\mathrm{O}(n)$ by

$$\left(g\cdot R\right)\left(v_1,\ldots,v_4\right):=R\left(g^{-1}v_1,\ldots,g^{-1}v_4\right).$$

and the representation $\mathrm{O}(n)\to\mathrm{GL}\left(\mathcal{S}\left(\mathbb{R}^n\right)\right)$ is defined similarly. Since the map r is $\mathrm{O}(n)$-invariant and s is $\mathrm{O}(n)$-equivariant (i.e., $g\cdot s\left(R\right)=s\left(g\cdot R\right)$, for $g\in\mathrm{O}(n)$), (16.63) is a decomposition of $\mathcal{R}\left(\mathbb{R}^n\right)$ into subspaces which are $\mathrm{O}(n)$-invariant. The subspace $\mathcal{R}_1,\mathcal{R}_2$ and $\mathcal{R}_3$ are actually irreducible, since $\mathcal{R}\left(\mathbb{R}^n\right)$ is irreducible as a $\mathrm{GL}\left(n,\mathbb{R}^n\right)$-module corresponding to the Young symmetrizer diagram

$$\begin{array}{|c|c|}\hline 1 & 3\\\hline 2 & 4\\\hline\end{array} \tag{16.64}$$

and r and s are the only independent contractions in $\mathcal{R}\left(\mathbb{R}^n\right)$ (see [**We46**, 153ff]). This symmetrizer takes a tensor in $\otimes^4\left(\mathbb{R}^n\right)^*$ and symmetrizes it in the indices in positions 1 and 3 (the top row of the diagram), and in positions 2 and 4 (the bottom row of the diagram). Then the result is antisymmetrized in the indices in positions 1 and 2 (the first column of the diagram) and in positions 3 and 4 (the second column of the diagram). In order to determine the $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$ components of $R\in\mathcal{R}\left(\mathbb{R}^n\right)$, we introduce a bilinear, symmetric map (the *Kulkarni-Nomizu product*, up to a constant factor)

$$\vee:S^2\left(\mathbb{R}^n\right)\times S^2\left(\mathbb{R}^n\right)\to\mathcal{R}\left(\mathbb{R}^n\right)$$

say $\vee\left(P,Q\right):=P\vee Q$, where

$$\left(P\vee Q\right)_{hijk}:=\tfrac{1}{2}\left(P_{hj}Q_{ik}-P_{ij}Q_{hk}+P_{ik}Q_{hj}-P_{hk}Q_{ij}\right). \tag{16.65}$$

Note that this is simply twice the result of applying the diagram (16.64) to $P\otimes Q$.

EXERCISE 16.48. Check that $P\vee Q\in\mathcal{R}\left(\mathbb{R}^n\right)$;i.e., verify conditions (A) and (B) in (16.62).

Moreover, if I denotes the usual dot product on $\mathbb{R}^n$ (i.e., $I_{ij}=\delta_{ij}$ =Kronecker delta), then

$$\begin{aligned}\left(I\vee Q\right)_{hijk}&=\tfrac{1}{2}\left(\delta_{hj}Q_{ik}-\delta_{ij}Q_{hk}+\delta_{ik}Q_{hj}-\delta_{hk}Q_{ij}\right),\\ r\left(I\vee Q\right)_{ik}&=\tfrac{1}{2}\left(nQ_{ik}-Q_{ik}+\delta_{ik}Tr\left(Q\right)-Q_{ik}\right)\\ &=\tfrac{1}{2}\left(\left(n-2\right)Q_{ik}+Tr\left(Q\right)\delta_{ik}\right).\end{aligned}$$

Hence,

$$r\left(I\vee Q\right)=\tfrac{1}{2}\left(\left(n-2\right)Q+Tr\left(Q\right)I\right), \tag{16.66}$$

$$s\left(I\vee Q\right)=\left(n-1\right)Tr\left(Q\right), \tag{16.67}$$

and in particular,

$$r\left(I \vee I\right) = \tfrac{1}{2}\left(\left(n-2\right)I + nI\right) = \left(n-1\right)I, \text{ and} \tag{16.68}$$

$$s\left(I \vee I\right) = \left(n-1\right)n. \tag{16.69}$$

Proposition 16.49. *The adjoint r^* of $r : \mathcal{R}\left(\mathbb{R}^n\right) \to \mathcal{S}\left(\mathbb{R}^n\right)$ is $\frac{1}{2}I\vee : \mathcal{S}\left(\mathbb{R}^n\right) \to \mathcal{R}\left(\mathbb{R}^n\right)$; i.e., $r^*\left(Q\right) = 2I \vee Q$, in the sense that for $R \in \mathcal{R}\left(\mathbb{R}^n\right)$ and $Q \in \mathcal{S}\left(\mathbb{R}^n\right)$,*

$$\left\langle r\left(R\right), Q\right\rangle = \left\langle R, \tfrac{1}{2}I \vee Q\right\rangle.$$

For $n > 2$, $I\vee : \mathcal{S}\left(\mathbb{R}^n\right) \to \mathcal{R}_1 \oplus \mathcal{R}_2$ is an isomorphism of vector spaces. More precisely, if

$$\mathcal{S}_0\left(\mathbb{R}^n\right) := \left\{P \in \mathcal{S}\left(\mathbb{R}^n\right) : Tr\left(P\right) = 0\right\},$$

then

$$I\vee : \mathcal{S}_0\left(\mathbb{R}^n\right)^{\perp} \oplus \mathcal{S}_0\left(\mathbb{R}^n\right) \to \mathcal{R}_1 \oplus \mathcal{R}_2$$

respects the summands, and for $P, Q \in \mathcal{S}_0\left(\mathbb{R}^n\right)$,

$$\left\langle I \vee P, I \vee Q\right\rangle = \left(n-2\right)\left\langle P, Q\right\rangle. \tag{16.70}$$

Proof. For $R \in \mathcal{R}\left(\mathbb{R}^n\right)$ and $Q \in \mathcal{S}\left(\mathbb{R}^n\right)$, we have (summing over repeated indices),

$$\begin{aligned}\left\langle R, I \vee Q\right\rangle &= R_{hijk}\left(I \vee Q\right)_{hijk} \\ &= \tfrac{1}{2}R_{hijk}\left(\delta_{hj}Q_{ik} - \delta_{ij}Q_{hk} + \delta_{ik}Q_{hj} - \delta_{hk}Q_{ij}\right) \\ &= \tfrac{1}{2}\Big(r\left(R\right)_{ik}Q_{ik} + r\left(R\right)_{hk}Q_{hk} + r\left(R\right)_{hj}Q_{hj} + r\left(R\right)_{ij}Q_{ij}\Big) \\ &= 2r\left(R\right)_{ik}Q_{ik} = \left\langle 2r\left(R\right), Q\right\rangle.\end{aligned}$$

This shows that $r^*\left(Q\right) = \frac{1}{2}I \vee Q$. Also, $\left\langle R, I \vee I\right\rangle = \left\langle 2r\left(R\right), I\right\rangle = 2s\left(R\right)$, which shows that $\operatorname{Ker}\left(s\right)^{\perp}$ is spanned by $I \vee I$. Since the image of r^* is $\left(\operatorname{Ker} r\right)^{\perp}$, the mapping $I\vee : \mathcal{S}\left(\mathbb{R}^n\right) \to \mathcal{R}\left(\mathbb{R}^n\right)$ is onto $\left(\operatorname{Ker} r\right)^{\perp} = \mathcal{R}_1 \oplus \mathcal{R}_2$. Moreover, for $P, Q \in \mathcal{S}\left(\mathbb{R}^n\right)$,

$$\begin{aligned}\left\langle I \vee P, I \vee Q\right\rangle &= \left\langle 2r\left(I \vee P\right), Q\right\rangle = \left\langle \left(\left(n-2\right)P + Tr\left(P\right)I\right), Q\right\rangle \\ &= \left(n-2\right)\left\langle P, Q\right\rangle + Tr\left(P\right)Tr\left(Q\right).\end{aligned} \tag{16.71}$$

In particular, (16.70) holds for $P, Q \in \mathcal{S}_0\left(\mathbb{R}^n\right)$, and

$$\left|I \vee Q\right|^2 = \left(n-2\right)\left|Q\right|^2 + Tr\left(Q\right)^2, \tag{16.72}$$

which shows that $I\vee : \mathcal{S}\left(\mathbb{R}^n\right) \to \mathcal{R}\left(\mathbb{R}^n\right)$ is injective for $n \neq 2$. Note that $\left(I\vee\right)\left(\mathcal{S}_0\left(\mathbb{R}^n\right)\right) = \mathcal{R}_2$, since

$$I \vee Q \in \mathcal{R}_2 \Leftrightarrow 0 = s\left(I \vee Q\right) = Tr\left(r\left(I \vee Q\right)\right) = \left(n-1\right)Tr\left(Q\right).$$

Also, as $\mathcal{S}_0\left(\mathbb{R}^n\right)^{\perp}$ is spanned by I, $\left(I\vee\right)\left(\mathcal{S}_0\left(\mathbb{R}^n\right)^{\perp}\right) = \mathcal{R}_1$ by (16.68). □

For $R \in \mathcal{R}\left(\mathbb{R}^n\right)$, suppose that according to the decomposition (16.63)

$$R = R_1 + R_2 + R_3.$$

Then R_1 is the projection of R onto $\operatorname{Ker}\left(s\right)$, namely

$$R_1 := \tfrac{\left\langle R, I \vee I\right\rangle}{\left\|I \vee I\right\|^2}\left(I \vee I\right) = \tfrac{s\left(R\right)}{n\left(n-1\right)}\left(I \vee I\right),$$

where (16.71) is used to compute $\|I \vee I\|^2 = (n-2)\,n + n^2 = 2n\,(n-1)$. For $n = 2$, $\dim\left(\mathcal{R}\left(\mathbb{R}^2\right)\right) = 1$, and hence $R = R_1 = \frac{1}{2} s\,(R)\,(I \vee I)$, and $R_2 = R_3 = 0$. Thus, we now assume $n > 2$. A candidate for R_2 is obtained as follows. Note that $I \vee r\,(R) = 2r^*\,(r\,(R)) \in (\operatorname{Ker} r)^{\perp}$, and subtracting off the $(\operatorname{Ker} s)^{\perp}$ component, we get an element of $\operatorname{Ker} s \cap (\operatorname{Ker} r)^{\perp} = \mathcal{R}_2$, namely $(n-2)\langle P, Q\rangle + Tr\,(P)\,Tr\,(Q)$

$$\begin{aligned} I \vee r\,(R) - \tfrac{\langle I \vee r(R), I\vee I\rangle}{|I \vee I|^2}\,(I \vee I) &= I \vee \left(r\,(R) - \tfrac{(n-2)s(R)+ns(R)}{n(n-1)} I\right) \\ &= I \vee \left(r\,(R) - \tfrac{1}{n} s\,(R)\,I\right). \end{aligned}$$

The operator $A : \mathcal{R}\,(\mathbb{R}^n) \to \mathcal{R}_2$, given by

$$A\,(R) := I \vee \left(r\,(R) - \tfrac{1}{n} s\,(R)\,I\right)$$

is not quite a projection onto $\operatorname{Ker} s \cap (\operatorname{Ker} r)^{\perp}$, but rather note that for $R \in \mathcal{R}_2$, say $R = I \vee Q$ *with* $Tr\,(Q) = 0$, we have (using (16.66) and (16.67))

$$A\,(R) = A\,(I \vee Q) = I \vee \left(r\,(I \vee Q) - \tfrac{1}{n} s\,(I \vee Q)\,I\right) = \tfrac{1}{2}\,(n-2)\,(I \vee Q)\,.$$

Hence for $n > 2$, $\frac{2}{n-2} A|_{\mathcal{R}_2}$ is the identity and $A|_{(\mathcal{R}_1 \oplus \mathcal{R}_3)} = 0$, since clearly $A|_{\mathcal{R}_3} = A|_{\operatorname{Ker} r} = 0$ and

$$\begin{aligned} A\,(I \vee I) &= I \vee \left(r\,(I \vee I) - \tfrac{1}{n} s\,(I \vee I)\,I\right) \\ &= I \vee \left((n-1)\,I - \tfrac{1}{n}\,(n-1)\,nI\right) = 0. \end{aligned}$$

Thus, $\frac{2}{n-2} A : \mathcal{R}\,(\mathbb{R}^n) \to \mathcal{R}_2$ is orthogonal projection and

$$R_2 = \tfrac{2}{n-2} A\,(R) = \tfrac{2}{n-2} I \vee \left(r\,(R) - \tfrac{1}{n} s\,(R)\,I\right).$$

In summary, we have

PROPOSITION 16.50. *For* $n = 2$, $R = R_1 = s\,(R)\,(I \vee I)$, *while for* $n > 2$, $R = R_1 + R_2 + R_3$, *where*

$$\begin{aligned} R_1 &= \tfrac{s(R)}{n(n-1)}\,(I \vee I) \in \mathcal{R}_1 = (\operatorname{Ker} s)^{\perp} \\ R_2 &= \tfrac{2}{n-2} I \vee \left(r\,(R) - \tfrac{1}{n} s\,(R)\,I\right) \in \mathcal{R}_2 = \operatorname{Ker} s \cap (\operatorname{Ker} r)^{\perp} \\ R_3 &= R - R_1 - R_2 \in \operatorname{Ker} r. \end{aligned}$$

The parts R_1, R_2 and R_3 have names:

$$\begin{aligned} &R_1 \text{ is the } \textit{constant curvature part of } R \\ &R_2 \text{ is the } \textit{traceless Ricci part of } R \\ &R_3 \text{ (usually denoted } W\text{) is the } \textit{Weyl part of } R. \end{aligned} \tag{16.73}$$

R_1 gets its name as follows. For R of the form $\frac{s(R)}{n(n-1)}\,(I \vee I)$, we have (from (16.65)) that for independent vectors $X, Y \in \mathbb{R}^n$ with $\Pi := span\,(X, Y)$,

$$\begin{aligned} R\,(X, Y, X, Y) &= \tfrac{s(R)}{n(n-1)}\,(I \vee I)\,(X, Y, X, Y) \\ &= \tfrac{s(R)}{n(n-1)}\left((X \cdot X)\,(Y \cdot Y) - (X \cdot Y)^2\right), \end{aligned}$$

so that the sectional curvature (see (16.51, p. 394)) of Π, namely

$$K\,(\Pi) = \frac{R\,(X, Y, X, Y)}{\|X\|^2\,\|Y\|^2 - (X \cdot Y)^2} = \frac{s\,(R)}{n\,(n-1)},$$

is independent of the plane Π (i.e., K is a *constant* function on the set of planes). Note $R_2 = \frac{2}{n-2} I \vee \left(r(R) - \frac{1}{n} s(R) I\right)$, is determined by the *traceless Ricci tensor* $r(R) - \frac{1}{n} s(R) I$, called so since $Tr\left(r(R) - \frac{1}{n} s(R) I\right) = s(R) - s(R) = 0$. The Weyl part of the curvature tensor of a Riemannian manifold M with metric h is known as the *Weyl conformal curvature tensor* (or simply *Weyl tensor*) of (M, h), and is denoted by W. We should mention that each $T_x M$ can be isometrically identified with $\mathbb{R}^n$. Such an identification is unique up to $\mathrm{O}(n)$ and the various spaces $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$ are $\mathrm{O}(n)$-invariant, so that the split $R = R_1 + R_2 + R_3$ can be made invariantly. Indeed, for the curvature tensor field R, we can replace I in (16.73) by the metric h to define R_1, R_2 and R_3. Thus,

$$\begin{aligned} W = R_3 &= R - R_1 - R_2 \\ &= R - \tfrac{s(R)}{n(n-1)} (h \vee h) - \tfrac{2}{n-2} h \vee \left(r(R) - \tfrac{1}{n} s(R) h\right). \end{aligned}$$

Let $W^\#$ be the $(1,3)$-tensor obtained from W by raising the first index, namely $\left(W^\#\right)^i{}_{jkl} = h^{im} W_{mjkl}$. Then $W^\#$ has the property that it is invariant under a conformal change of metric, say $h \mapsto e^{2\sigma} h$ for some $\sigma \in C^\infty(M)$. If $(R_\sigma)^{\#_\sigma}$ is the $(1,3)$ version of the Riemann curvature tensor of $e^{2\sigma} h$, then a somewhat lengthy computation in [**Ei**] yields

$$R^\# - (R_\sigma)^{\#_\sigma} = |d\sigma|^2 (h \vee h)^\# + 2 (\widetilde{\sigma} \vee h)^\#, \tag{16.74}$$

where $\widetilde{\sigma} := \nabla(d\sigma) - d\sigma \otimes d\sigma$ is a symmetric 2-tensor. The 2-tensor $\nabla(d\sigma)$ is the covariant derivative of $d\sigma$ with respect to the Levi-Civita connection θ of h, and it is known as the *Hessian* of σ. The symmetry of $\nabla(d\sigma)$ is due to fact that θ is torsion-free. Since the right side of (16.74) has Weyl part 0, the $(1,3)$-version of the Weyl tensor is unchanged. Also in [**Ei**] it is shown that if $W = 0$ and $\dim(M) \geq 4$, then about each point $x \in M$, there is a neighborhood U and a function $\sigma \in C^\infty(U)$, such that the curvature tensor R_σ of $e^{2\sigma} h$ is 0 (i.e., (M, h) is conformally flat). The result (16.74) shows that $W = 0$ is necessary in order that (M, h) be conformally flat. If $\dim(M) = 3$, then $W = 0$ is automatic, since

$$6 = \frac{1}{12}\left(3^2\right)\left(3^2 - 1\right) = \dim\left(\mathcal{R}\left(\mathbb{R}^3\right)\right) = \dim\left(\mathcal{R}_1\left(\mathbb{R}^3\right) \oplus \mathcal{R}_2\left(\mathbb{R}^3\right)\right)$$

Thus, for $\dim(M) = 3$, R is determined by the Ricci tensor. However for $\dim(M) = 3$, $W = 0$ does not imply that (M, h) is conformally flat. For $\dim(M) = 2$, $W = 0$ is again automatic, but conformal flatness does not follow from $W = 0$. Instead, one proves conformal flatness (i.e., the existence of isothermal parameters) by other means.

We will focus on oriented Riemannian 4-manifolds (M, h). First however, in any dimension, due of the antisymmetry of $R(W, Z, X, Y)$ in (X, Y) and in (W, Z), we can view R as a section of $\mathrm{End}\left(\Lambda^2(TM^*)\right)$, or an operator sending 2-forms to 2-forms. This operator is called the *curvature operator*. In terms of a local frame field, the curvature operator applied to a 2-form α yields the 2-form $\widehat{R}(\alpha)$ defined by

$$\widehat{R}(\alpha)_{ij} := \tfrac{1}{2} R_{ijkl} \alpha^{kl}, \tag{16.75}$$

where $\alpha^{kl} = h^{kp} h^{lq} \alpha_{pq}$. Note that for a metric of constant sectional curvature 1, $\widehat{R}$ is the identity. The symmetry $R(W, Z, X, Y) = R(X, Y, W, Z)$ implies that $\widehat{R}$ is a symmetric endomorphisms of $\Lambda^2(TM^*)$, since

$$h(\widehat{R}(\alpha),\beta) = \tfrac{1}{4}R_{ijkl}\alpha^{kl}\beta^{ij} = \tfrac{1}{4}R_{klij}\alpha^{kl}\beta^{ij} = h(\alpha,\widehat{R}(\beta)).$$

For oriented Riemannian 4-manifolds, we have another automorphism of $\Lambda^2(TM^*)$, namely the Hodge star $*:\Lambda^2(TM^*)\to\Lambda^{4-2}(TM^*)$. While the decomposition, $\mathcal{R}(\mathbb{R}^n) = \mathcal{R}_1\oplus\mathcal{R}_2\oplus\mathcal{R}_3$ of (16.63) consists of $\mathrm{O}(n)$-irreducible subspaces, we now show that, for $n=4$, $\mathcal{R}_3$ is *not* $\mathrm{SO}(4)$-irreducible. Recall that the Hodge star $*$ satisfies $*^2 = \mathrm{Id}$ on $\Lambda^2(\mathbb{R}^{4*})$, and we have $\Lambda^2(\mathbb{R}^{4*}) = \Lambda^2_+\oplus\Lambda^2_-$, where

$$\Lambda^2_+ := (1+*)\left(\Lambda^2\left(\mathbb{R}^{4*}\right)\right) \text{ and } \Lambda^2_- = (1-*)\left(\Lambda^2\left(\mathbb{R}^{4*}\right)\right)$$

are the self-dual and anti-self-dual subspaces (±1 eigenspaces of $*$). Relative to the standard basis e^1,e^2,e^3,e^4 of $\mathbb{R}^{4*}$, a basis of $\Lambda^2_\pm$ is

$$e^2\wedge e^3 \pm e^1\wedge e^4,\ e^3\wedge e^1\pm e^2\wedge e^4,\ e^1\wedge e^2\pm e^3\wedge e^4.$$

We can write $\widehat{R}$ and $*$ in block form relative to the decomposition $\Lambda^2(\mathbb{R}^{4*}) = \Lambda^2_+\oplus\Lambda^2_-$, say

$$\widehat{R} = \begin{bmatrix} A & B \\ B^T & C\end{bmatrix},\quad * = \begin{bmatrix} I & 0 \\ 0 & -I\end{bmatrix}, \tag{16.76}$$

where $A = A^T$ and $C = C^T$ since $\widehat{R}$ is symmetric. We know that $\dim\mathcal{R}(\mathbb{R}^4) = \frac{1}{12}4^2(4^2-1) = 20$, but the dimension of the space of all 6×6 symmetric matrices is $6(6+1)/2 = 21$. The discrepancy is due to the fact that $*$ is orthogonal to $\mathcal{R}(\mathbb{R}^4)$ in the space $\mathcal{S}(\Lambda^2)$ of *all* symmetric endomorphisms of $\Lambda^2(\mathbb{R}^{4*})$ due to the Bianchi Identity, namely for $R\in\mathcal{R}(\mathbb{R}^4)$

$$0 = \varepsilon^{ijkl}(R_{ijkl}+R_{iljk}+R_{iklj}) = 3\varepsilon^{ijkl}R_{ijkl} = 3\langle *,R\rangle, \tag{16.77}$$

where we recall from (16.19) that $(*\alpha)_{ij} = \frac{1}{2}\varepsilon^{ijkl}\alpha_{kl}$. Thus,

$$\mathcal{R}(\mathbb{R}^4)\cong\mathcal{S}'(\Lambda^2) := *^\perp := \left\{S\in\mathcal{S}(\Lambda^2) : \langle S,*\rangle = 0\right\}. \tag{16.78}$$

By (16.76) and (16.77), we have $0 = \langle *,R\rangle = Tr(A) - Tr(C)$, while $Tr(A) + Tr(C) = Tr(R) = \frac{1}{2}R^{ij}{}_{ij} = \frac{1}{2}s(R)$. Hence, $Tr(A) = Tr(C) = \frac{1}{4}s(R)$, and defining $\widetilde{A} := A - \frac{1}{3}Tr(A)I$ and $\widetilde{C} := C - \frac{1}{3}Tr(C)I$, we have

$$R = \frac{s(R)}{12}\begin{bmatrix} I & 0\\ 0 & I\end{bmatrix} + \begin{bmatrix} 0 & B\\ B^T & 0\end{bmatrix} + \begin{bmatrix}\widetilde{A} & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix} 0 & 0\\ 0 & \widetilde{C}\end{bmatrix}. \tag{16.79}$$

As $s(R)$, B, and (traceless symmetric) $\widetilde{A}$ and $\widetilde{C}$ vary independently, each of the summands in (16.79) varies over a subspace of $\mathcal{R}(\mathbb{R}^4)$, say $\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3$, and $\mathcal{C}_4$, from left to right. Note that

$$\begin{aligned}
\mathcal{C}_1 &= \left\{\alpha\,\mathrm{Id}\in\mathcal{S}(\Lambda^2) : \alpha\in\mathbb{R}\right\},\\
\mathcal{C}_2 &= \left\{R\in\mathcal{S}(\Lambda^2) : R\circ * = -*\circ R\right\},\\
\mathcal{C}_3 &= \left\{R\in\mathcal{S}(\Lambda^2) : \langle *,R\rangle = 0,\ R\circ * = *\circ R = R\right\} \text{ and}\\
\mathcal{C}_4 &= \left\{R\in\mathcal{S}(\Lambda^2) : \langle *,R\rangle = 0,\ R\circ * = *\circ R = -R\right\}.
\end{aligned}$$

Thus, as $*$ is $\mathrm{SO}(4)$-invariant, each of $\mathcal{C}_i$ is an $\mathrm{SO}(4)$-invariant subspace for the tensor representation $\mathrm{SO}(4)\to\mathrm{O}\left(\mathcal{S}(\Lambda^2)\right)$. For Proposition 16.52 below, we will use the following

LEMMA 16.51. *Let $r : \mathrm{O}(n) \to \mathrm{O}(V)$ be an irreducible representation. Then the restriction $r|_{\mathrm{SO}(n)}$ is either irreducible, or V is the direct sum of two irreducible $r|_{\mathrm{SO}(n)}$-invariant subspaces of equal dimension.*

PROOF. Suppose that $r|_{\mathrm{SO}(n)}$ is not irreducible, and let V' be a proper, irreducible $r|_{\mathrm{SO}(n)}$-invariant subspace of V. Let $A \in \mathrm{O}(n)$ with $\det A = -1$. Then $r(A)(V')$ is $r|_{\mathrm{SO}(n)}$-invariant, since for any $B \in \mathrm{SO}(n)$, we have $C := A^{-1}BA \in \mathrm{SO}(n)$, and

$$r(B)(r(A)(V')) = r(BA)(V') = r(AC)(V') = r(A)(r(C)V') \subseteq r(A)(V').$$

Since V' is irreducible, either $r(A)(V') \cap V' = V'$ or $r(A)(V') \cap V' = \{0\}$. If $r(A)(V') \cap V' = V'$, then V' is a proper r-invariant subspace of V, contrary to assumption. If $r(A)(V') \cap V' = \{0\}$, then $V' + r(A)(V')$ is a direct sum. Moreover, $V' + r(A)(V')$ is also r-invariant, since any $C \in \mathrm{O}(n)$ is of the form BA^{-1} and AB' for some $B, B' \in \mathrm{SO}(n)$, and we have

$$\begin{aligned} r(C)(V' + r(A)(V')) &= r(C)(V') + r(C)\,r(A)(V') \\ &= r(AB')(V') + r\left(BA^{-1}\right) r(A)(V') \\ &= r(A)(V') + r(B)(V') = V' + r(A)(V'). \end{aligned}$$

Thus, $V = V' + r(A)(V')$ is a direct sum of two $r|_{\mathrm{SO}(n)}$-invariant subspaces of equal dimension. It remains to show that $r(A)(V')$ is $r|_{\mathrm{SO}(n)}$-irreducible. If W is a proper, irreducible $r|_{\mathrm{SO}(n)}$-invariant subspace of $r(A)(V')$, then (contrary to the $r|_{\mathrm{SO}(n)}$-irreducibility of V') $r(A)^{-1}(W)$ is a proper, $r|_{\mathrm{SO}(n)}$-invariant subspace of V', since for any $B \in \mathrm{SO}(n)$, $BA^{-1} = A^{-1}B'$ for some $B' \in \mathrm{SO}(n)$, and so

$$r(B)\left(r(A)^{-1}(W)\right) = r(A)^{-1}(r(B')(W)) = r(A)^{-1}(W).$$

□

PROPOSITION 16.52. *With respect to the decomposition* (*see* (*16.63*))

$$\mathcal{R}\left(\mathbb{R}^4\right) = \mathcal{R}_1 \oplus \mathcal{R}_2 \oplus \mathcal{R}_3 := (\operatorname{Ker} s)^{\perp} \oplus \left(\operatorname{Ker} s \cap (\operatorname{Ker} r)^{\perp}\right) \oplus \operatorname{Ker} r, \tag{16.80}$$

under the identification $\mathcal{R}\left(\mathbb{R}^4\right) \cong \mathcal{S}'\left(\Lambda^2\right)$ of (16.78), we have

$$\begin{aligned} &\mathcal{C}_1 \cong \mathcal{R}_1 = \{\textit{constant curvature parts}\}, \\ &\mathcal{C}_2 \cong \mathcal{R}_2 = \{\textit{traceless Ricci parts}\} \ \textit{and} \\ &\mathcal{C}_3 + \mathcal{C}_4 \cong \mathcal{R}_3 = \{\textit{Weyl parts}\}, \end{aligned}$$

in the terminology of (*16.73*).

PROOF. We know that the $\mathcal{R}_i$ are irreducible, $\mathrm{O}(4)$-invariant subspaces of $\mathcal{R}\left(\mathbb{R}^4\right)$, with $\dim(\mathcal{R}_1) = 1$, $\dim(\mathcal{R}_2) = \dim\left(\mathcal{S}_0\left(\mathbb{R}^4\right)\right) = 4(4+1)/2 - 1 = 9$, and

$$\dim(\mathcal{R}_3) = \dim\left(\mathcal{R}\left(\mathbb{R}^4\right)\right) - 10 = 4^2\left(4^2 - 1\right)/12 - 10 = 10.$$

By Lemma 16.51, the odd-dimensional $\mathrm{O}(4)$-irreducible subspaces $\mathcal{R}_1$ and $\mathcal{R}_2$ remain irreducible under $\mathrm{SO}(4)$. Since the isomorphism $\mathcal{R}\left(\mathbb{R}^4\right) \cong \mathcal{S}'\left(\Lambda^2\right)$ is $\mathrm{SO}(4)$-equivariant and there are four $\mathrm{SO}(4)$-invariant summands $\mathcal{C}_i$, $\mathcal{R}_3$ must split into two irreducible, $\mathrm{SO}(4)$-invariant summands, each of dimension 5. These are necessarily $\mathcal{C}_3$ and $\mathcal{C}_4$, and then clearly $\mathcal{C}_1 \cong \mathcal{R}_1$ and $\mathcal{C}_2 \cong \mathcal{R}_2$. □

We can decompose the Riemann curvature tensor R of an oriented, Riemannian 4-manifold (M,h) as

$$R = R_1 + R_2 + W = R_1 + R_2 + W^+ + W^-,$$

where W^+ and W^- correspond respectively to the $\mathcal{C}_3$ and $\mathcal{C}_4$ components of the Weyl tensor W. If $W^- = 0$, then (M,h) is called *self-dual,* and if $W^- = 0$, then (M,h) is called *anti-self-dual.* The terminology is fitting, since considering $*$ and $W^\pm \in C^\infty(End(\Lambda(TM^*)))$, we have $* \circ W^\pm = W^\pm \circ * = \pm W^\pm$.

6. Bochner-Weitzenböck Formulas

Let $\pi : P \to M$ be a principal G-bundle and let ω be a connection1-form on P. Suppose that $\rho : G \to \mathrm{GL}(W)$ is a representation, and let $\overline{\Omega}^k(P,W)$ be the space of horizontal equivariant forms (see Definition 16.10, p. 371). In this section, we assume that M is compact, oriented Riemannian n-manifold with metric h. Let θ be the Levi-Civita connection (see Remark 16.39, p. 389) on the principal $\mathrm{SO}(n)$-bundle $\pi_F : FM \to M$ of oriented, orthonormal frames. Let

$$P \times_f FM := \{(p,u) \in P \times FM : \pi(p) = \pi_F(u)\}.$$

The group $G \times \mathrm{SO}(n)$ acts freely on $P \times_f FM$ via $(p,u)(g_1,g_2) = (pg_1, ug_2)$, and

$$\pi \times_f \pi_F : P \times_f FM \to M$$

is readily verified to be a principal $G \times \mathrm{SO}(n)$-bundle, called the *fibered product* of P and FM. The subscript f in $\times_f$ stands for "fibered" (note that $P \times_f FM \neq P \times FM$). Observe that

$$\pi_1 : P \times_f FM \to P \text{ and } \pi_2 : P \times_f FM \to FM$$

given by $\pi_1(p,u) := p$ and $\pi_2(p,u) := u$ are principal bundles with groups $\mathrm{SO}(n)$ and G respectively. Note that $\pi_1^*\omega$ is a $\mathfrak{g}$-valued 1-form on $P \times_f FM$, while $\pi_2^*\theta$ is a $\mathfrak{so}(n)$-valued 1-form on $P \times_f FM$. The direct sum $\pi_1^*\omega \oplus \pi_2^*\theta$ is a $\mathfrak{g} \oplus \mathfrak{so}(n)$-valued 1-form on $P \times_f FM$. It is not hard to verify that $\pi_1^*\omega \oplus \pi_2^*\theta$ is a connection 1-form for $\pi \times_f \pi_F : P \times_f FM \to M$. To avoid cumbersome expressions, let us adopt the notation

$$\omega \oplus \theta := \pi_1^*\omega \oplus \pi_2^*\theta. \tag{16.81}$$

If $\mathbb{R}^{n*}$ denotes the dual space of $\mathbb{R}^n$, then we define the space of tensors contravariant of degree r and covariant of degree s by

$$T^{r,s} := \mathbb{R}^n \otimes \overset{r}{\cdots} \otimes \mathbb{R}^n \otimes \mathbb{R}^{n*} \otimes \overset{s}{\cdots} \otimes \mathbb{R}^{n*}.$$

and the tensor representation $t^{r,s} : \mathrm{SO}(n) \to \mathrm{GL}(T^{r,s})$ is given by

$$\begin{aligned} &t^{r,s}(A)(v_1 \otimes \cdots \otimes v_r \otimes \eta_1 \otimes \cdots \otimes \eta_s) \\ &= Av_1 \otimes \cdots \otimes Av_r \otimes \left(\eta_1 \circ A^{-1}\right) \otimes \cdots \otimes \left(\eta_s \circ A^{-1}\right). \end{aligned} \tag{16.82}$$

Then we have a representation $\rho \otimes t^{r,s} : G \times \mathrm{SO}(n) \to \mathrm{GL}(W \otimes T^{r,s})$, and we may consider the spaces $\overline{\Omega}^k(P \times_f FM, W \otimes T^{r,s})$ of horizontal, equivariant $W \otimes T^{r,s}$-valued k-forms on $P \times_f FM$. We have the usual isomorphism (see (16.11), p. 373)

$$\overline{\Omega}^k(P \times_f FM, W \otimes T^{r,s}) \cong \Omega^k\left(M, (P \times_f FM) \times_{G \times \mathrm{SO}(n)} (W \otimes T^{r,s})\right) \tag{16.83}$$

between horizontal, equivariant k-forms and k-forms with values in the associated vector bundle. More significantly for this section is the fact that either of the spaces in (16.83) can be identified with the subspace of elements of

$$\overline{\Omega}^0\left(P\times_f FM, W\otimes T^{r,s+k}\right)\cong C^\infty\left(M,(P\times_G W)\otimes T(M)^{r,s+k}\right)$$

which are antisymmetric in the last k slots. This is accomplished via standard horizontal vector fields (see Definition 16.44, p. 397). Recall that for $w\in\mathbb{R}^n$, the standard horizontal vector field $\overline{w}$ is defined on FM. However, we can take the horizontal lift of $\overline{w}$ to $P\times_f FM$, relative to the connection $\pi_1^*\omega$ on $\pi_2 : P\times_f FM\to FM$, in order to obtain a vector field $\widetilde{\overline{w}}$ on $P\times_f FM$. Let us just call $\widetilde{\overline{w}}$ the standard horizontal vector field on $P\times_f FM$ associated to $w\in\mathbb{R}^n$, and to avoid notational complication use the same notation $\overline{w}$ as we do for $\overline{w}$ on FM. Note that for $(g,A)\in G\times\mathrm{O}(n)$, we have

$$(R)_{(g,A)*}\left(\overline{w}_{(p,u)}\right)=\overline{\left(A^{-1}w\right)}_{(pg,u\circ A)}. \tag{16.84a}$$

(cf. (16.55), p. 397). Thus, regarding G and $\mathrm{O}(n)$ as subgroups of $G\times\mathrm{O}(n)$, note that $\overline{w}$ on $P\times_f FM$ is G-invariant, and although not $\mathrm{SO}(n)$-invariant, it transforms "nicely." For $\alpha\in\overline{\Omega}^k\left(P\times_f FM,W\otimes T^{r,s}\right)$, we define $\alpha'\in\overline{\Omega}^0\left(P\times_f FM,W\otimes T^{r,s+k}\right)$ by

$$\begin{aligned}&\alpha'(p,u)(\eta_1,\ldots,\eta_r,v_1,\ldots,v_{s+k})\\&:=\alpha_{(p,u)}(\eta_1,\ldots,\eta_r,v_1,\ldots,v_s)\left(\overline{v_{s+1}},\ldots,\overline{v_{s+k}}\right),\end{aligned}$$

where $\eta_1,\ldots,\eta_r\in\mathbb{R}^{n*}$, $v_1,\ldots,v_{s+k}\in\mathbb{R}^n$, and $\overline{v_{s+1}},\ldots,\overline{v_{s+k}}$ are the standard horizontal vector fields on $P\times_f FM$ associated with $v_{s+1},\ldots,v_{s+k}$. Note that $\alpha'\in\overline{\Omega}^0\left(P\times_f FM,W\otimes T^{r,s+k}\right)$, since (using (16.84a))

$$\begin{aligned}&\alpha'((p,u)(g,A))(\eta_1,\ldots,\eta_r,v_1,\ldots,v_{s+k})\\&=\alpha_{(pg,u\circ A)}(\eta_1,\ldots,\eta_r,v_1,\ldots,v_s)\left(\overline{v_{s+1}},\ldots,\overline{v_{s+k}}\right)\\&=\alpha_{(p,u)}(\eta_1,\ldots,\eta_r,v_1,\ldots,v_s)\left(R_{(g,A)*}\left(\overline{Av}_{s+1}\right),\ldots,R_{(g,A)*}\left(\overline{Av}_{s+k}\right)\right)\\&=\left(R^*_{(g,A)}\alpha\right)_{(p,u)}(\eta_1,\ldots,\eta_r,v_1,\ldots,v_s)\left(\overline{Av}_{s+1},\ldots,\overline{Av}_{s+k}\right)\\&=\left(\rho\otimes t^{r,s}\right)\left((g,A)^{-1}\right)\left(\alpha_{(p,u)}(\eta_1,\ldots,\eta_r,v_1,\ldots,v_s)\left(\overline{Av}_{s+1},\ldots,\overline{Av}_{s+k}\right)\right)\\&=\rho\left(g^{-1}\right)\left(\alpha_{(p,u)}(\eta_1\circ A,\ldots,\eta_r\circ A,Av_1,\ldots,Av_s)\right)\left(\overline{Av}_{s+1},\ldots,\overline{Av}_{s+k}\right)\\&=\rho\left(g^{-1}\right)\left(\alpha'(p,u)(\eta_1\circ A,\ldots,\eta_r\circ A,Av_1,\ldots,Av_{s+k})\right)\\&=\left(\rho\otimes t^{r,s+k}\right)\left((g,A)^{-1}\right)\left(\alpha'(p,u)\right)(\eta_1,\ldots,\eta_r,v_1,\ldots,v_{s+k}).\end{aligned}$$

In addition to the covariant *exterior* differentiation operator

$$D^{\omega\oplus\theta}:\overline{\Omega}^k\left(P\times_f FM,W\otimes T^{r,s}\right)\to\overline{\Omega}^{k+1}\left(P\times_f FM,W\otimes T^{r,s}\right), \tag{16.85}$$

we also have

$$\begin{aligned}\nabla^{\omega\oplus\theta}:=D^{\omega\oplus\theta}&:\overline{\Omega}^0\left(P\times_f FM,W\otimes T^{r,s+k}\right)\\&\to\overline{\Omega}^1\left(P\times_f FM,W\otimes T^{r,s+k}\right).\end{aligned} \tag{16.86}$$

Even though one can regard

$$\overline{\Omega}^k\left(P\times_f FM,W\otimes T^{r,s}\right)\subseteq\overline{\Omega}^0\left(P\times_f FM,W\otimes T^{r,s+k}\right)$$

and

$$\overline{\Omega}^{k+1}\left(P\times_f FM, W\otimes T^{r,s}\right)\subseteq\overline{\Omega}^{0}\left(P\times_f FM, W\otimes T^{r,s+k+1}\right)$$
$$\cong\overline{\Omega}^{1}\left(P\times_f FM, W\otimes T^{r,s+k}\right),$$

the operator in (16.85) is *not* a restriction of the operator in (16.86) if $k>0$. Hence, we have introduced a different notation, namely $\nabla^{\omega\oplus\theta}$ in (16.86). We also add that it is customary to denote the evaluation of the 1-form $\nabla^{\omega\oplus\theta}\alpha$ on a vector X by $\nabla^{\omega\oplus\theta}_X\alpha$, namely for $\alpha\in\overline{\Omega}^{0}\left(P\times_f FM, W\otimes T^{r,s+k}\right)$ and $X\in T\left(P\times_f FM\right)$,

$$\nabla^{\omega\oplus\theta}_X\alpha:=\left(\nabla^{\omega\oplus\theta}\alpha\right)(X)$$

Although not equal, $D^{\omega\oplus\theta}\alpha$ and $\nabla^{\omega\oplus\theta}\alpha$ are nevertheless related by

PROPOSITION 16.53. *For*

$$\alpha\in\overline{\Omega}^{k}\left(P\times_f FM, W\otimes T^{r,s}\right)\subseteq\overline{\Omega}^{0}\left(P\times_f FM, W\otimes T^{r,s+k}\right)$$

and $\overline{x_1},\ldots,\overline{x_{k+1}}$ *standard horizontal vector fields on* $P\times_f FM$ *associated with* $x_1,\ldots,x_{k+1}\in\mathbb{R}^n$, *we have*

$$\left(D^{\omega\oplus\theta}\alpha\right)\left(\overline{x_1},\ldots,\overline{x_{k+1}}\right)=\sum_{i=1}^{k+1}(-1)^{i+1}\left(\nabla^{\omega\oplus\theta}_{\overline{x_i}}\alpha\right)\left(x_1,\ldots,\widehat{x_i},\ldots,x_{k+1}\right).\tag{16.87}$$

PROOF. Using the fact that $[\overline{x_i},\overline{x_j}]$ is vertical (see (16.56)), we compute

$$\begin{aligned}
&\left(D^{\omega\oplus\theta}\alpha\right)\left(\overline{x_1},\ldots,\overline{x_{k+1}}\right)=(d\alpha)\left(\overline{x_1},\ldots,\overline{x_{k+1}}\right)\\
&=\sum_{i=1}^{k+1}(-1)^{i+1}\,\overline{x_i}\left[\alpha\left(\overline{x_1},\ldots,\widehat{\overline{x_i}},\ldots,\overline{x_{k+1}}\right)\right]\\
&+\sum_{1\le i<j\le k+1}^{k+1}(-1)^{i+j}\,\alpha\left(\left[\overline{x_i},\overline{x_j}\right],\overline{x_1},\ldots,\widehat{\overline{x_i}},\ldots,\widehat{\overline{x_j}},\ldots,\overline{x_{k+1}}\right)\\
&=\sum_{i=1}^{k+1}(-1)^{i+1}\,\overline{x_i}\left[\alpha\left(\overline{x_1},\ldots,\widehat{\overline{x_i}},\ldots,\overline{x_{k+1}}\right)\right]\\
&=\sum_{i=1}^{k+1}(-1)^{i+1}\,d\left(\alpha\left(\overline{x_1},\ldots,\widehat{\overline{x_i}},\ldots,\overline{x_{k+1}}\right)\right)\left(\overline{x_i}\right)\\
&=\sum_{i=1}^{k+1}(-1)^{i+1}\,d\left(\alpha\left(x_1,\ldots,\widehat{x_i},\ldots,x_{k+1}\right)\right)\left(\overline{x_i}\right)\\
&=\sum_{i=1}^{k+1}(-1)^{i+1}\left(\nabla^{\omega\oplus\theta}_{\overline{x_i}}\alpha\right)\left(x_1,\ldots,\widehat{x_i},\ldots,x_{k+1}\right).
\end{aligned}$$

□

Contractions of tensor fields are most efficiently displayed in terms of components. While one usually thinks that this involves the choice of some coordinate system or ad hoc choice of basis, one of the advantages of working on the frame bundle (or more generally $P\times_f FM$) is that one can use the standard horizontal vector fields $\overline{e_1},\ldots,\overline{e_n}$ corresponding to the *standard basis* $e_1,\ldots,e_n$ of $\mathbb{R}^n$ and *standard* *dual basis* $e^1, \cdots, e^n \in \mathbb{R}^{n*}$. Thus, there are standard W-valued components of $\alpha \in \overline{\Omega}^k (P \times_f FM, W \otimes T^{r,s})$ defined at each $(p,u) \in P \times_f FM$ by

$$\alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s; q_1 \cdots q_k} (p,u) := \alpha_{(p,u)} \left(\overline{e_{q_1}}, \cdots, \overline{e_{q_k}}\right) \left(e^{i_1}, \cdots, e^{i_r}, e_{j_1}, \cdots, e_{j_s}\right) \in W.$$

(If $r = s = 0$, we drop the semicolon.) Let $\varphi \in \overline{\Omega}^1 (FM, \mathbb{R}^n)$ be the canonical 1-form. Note that the components $\varphi^1, \ldots, \varphi^n$ of φ are $\mathbb{R}$-valued forms vanishing on vertical vectors. Using $\pi_2 : P \times_f FM \to FM$, we can pull back the form φ to a form $\pi_2^* (\varphi) \in \overline{\Omega}^m (P \times_f FM, \mathbb{R}^n)$, which we continue to denote by φ and whose components are still denoted by $\varphi^1, \ldots, \varphi^n$. For $\alpha \in \overline{\Omega}^k (P \times_f FM, W \otimes T^{r,s})$, we can write

$$\alpha \left(e^{i_1}, \cdots, e^{i_r}, e_{j_1}, \cdots, e_{j_s}\right) = \frac{1}{k!} \sum_{q_1, \ldots, q_k} \alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s; q_1 \cdots q_k} \varphi^{q_1} \wedge \cdots \wedge \varphi^{q_k}.$$

Also, for $\alpha \in \overline{\Omega}^0 (P \times_f FM, W \otimes T^{r,s})$, we use the notation

$$\alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s | q} := \left(\nabla^{\omega \oplus \theta}_{\overline{e_q}} \alpha\right)^{i_1 \cdots i_r}_{j_1 \cdots j_s} = \left(\nabla^{\omega \oplus \theta} \alpha\right)^{i_1 \cdots i_r}_{j_1 \cdots j_s; q}.$$

Note for any $f \in C^\infty (P \times_f FM, W)$, $(df)^H (\overline{e_i}) = df (\overline{e_i}) = \overline{e_i} [f]$, from which it follows that

$$\alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s | q} = d \left(\alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s}\right) (\overline{e_i}) = \overline{e_q} \left[\alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s}\right]. \tag{16.88}$$

In terms of this notation, (16.87) can be rewritten as

$$\left(D^{\omega \oplus \theta} \alpha\right)^{i_1 \cdots i_r}_{j_1 \cdots j_s; q_0 \cdots q_k} = \sum_{l=0}^{k} (-1)^l \alpha^{i_1 \cdots i_r}_{j_1 \cdots j_s; q_0 \cdots \widehat{q_l} \cdots q_k | q_l}. \tag{16.89}$$

Moreover, if h denotes the Riemannian metric on M, and $H \in \overline{\Omega}^0 \left(P \times_f FM, T^{0,2}\right)$ is defined by

$$H (p,u) (v_1, v_2) := h (u (v_1), u (v_2)),$$

then as $u : \mathbb{R}^n \to T_{\pi(u)} M$ is an isometry,

$$H (p,u) (e_i, e_j) = h (u (e_i), u (e_j)) = \delta_{ij}.$$

Thus, the components of H are constant functions. Thus, in view of (16.88), indices can be raised (or lowered) before or after applying $\nabla^{\omega \oplus \theta}$ or $D^{\omega \oplus \theta}$ producing the same result. We have already found (see Proposition 16.18, p. 377) the formal adjoint of $D^\omega : \Omega^m (P \times_G W) \to \Omega^{m+1} (P \times_G W)$ to be $\delta^\omega := -(-1)^{nm} * D^\omega *$, but for the purpose of stating and proving the B-W formulas, it is convenient to have a lifted version of δ^ω defined directly on $\overline{\Omega}^m (P \times_f FM, W)$ instead of $\Omega^m (P \times_G W)$. Since there is a composition of isomorphisms

$$\Psi : \Omega^m (M, P \times_G W) \cong \overline{\Omega}^m (P, W) \cong \overline{\Omega}^m (P \times_f FM, W),$$

all we really need is the lifted version $*$, say

$$\overline{*} := \Psi^{-1} \circ * \circ \Psi, \tag{16.90}$$

in which case the lifted version of δ^ω is given by

$$\delta^\omega = -(-1)^{nm} \overline{*} D^\omega \overline{*} : \overline{\Omega}^{m+1} (P \times_f FM, W) \to \overline{\Omega}^m (P \times_f FM, W). \tag{16.91}$$

For a $(\omega \oplus \theta)$-horizontal subspace $H \subseteq T_{(p,u)} (P \times_f FM)$, it is not hard to see that $\overline{*}$ is simply the usual star operator acting on the restrictions of horizontal forms to H, where H is given the metric and orientation which make $(\pi \times_f \pi_F)_* : H \to T_{\pi(p)}M$ an orientation-preserving isometry. In terms of the components $\alpha_{q_1 \cdots q_m}$ of $\alpha \in \overline{\Omega}^m (P, W)$, one can verify that $\overline{*}\alpha \in \overline{\Omega}^{n-m} (P, W)$ is given by

$$(\overline{*}\alpha)_{q_{m+1} \cdots q_n} = \frac{1}{m!} \alpha^{q_1 \cdots q_m} \varepsilon_{q_1 \cdots q_m q_{m+1} \cdots q_n}.$$

where $\varepsilon_{q_1 \cdots q_n}$ is antisymmetric in its indices with $\varepsilon_{1 \cdots n} = 1$.

Proposition 16.54. *For*

$$D^\omega : \overline{\Omega}^m (P \times_f FM, W) \to \overline{\Omega}^{m+1} (P \times_f FM, W)$$

and $\alpha \in \overline{\Omega}^m (P \times_f FM, W)$,

$$(D^\omega \alpha)_{q_0 \cdots q_m} = \left(D^{\omega \oplus \theta} \alpha\right)_{q_0 \cdots q_m} = \sum_{l=0}^m (-1)^l \alpha_{q_0 \cdots \widehat{q_l} \cdots q_m | q_l}. \tag{16.92}$$

For $\delta^\omega : \overline{\Omega}^{m+1} (P \times_f FM, W) \to \overline{\Omega}^m (P \times_f FM, W)$ *given by (16.91) and* $\beta \in \overline{\Omega}^{m+1} (P \times_f FM, W)$, *we have*

$$(\delta^\omega \beta)_{q_1 \cdots q_m} = \left(- (-1)^{nm} \overline{*} D^\omega (\overline{*}\beta)\right)_{q_1 \cdots q_m} = - \sum_{i=1}^m \beta^i{}_{q_1 \cdots q_m | i}. \tag{16.93}$$

Proof. The equation (16.92) is just a restatement of (16.87) in the special case $T^{r,s} = T^{0,0} = \mathbb{R}$. For (16.93) we compute as follows. From

$$(\overline{*}\beta)_{q_{m+2} \cdots q_n} = \frac{1}{(m+1)!} \beta_{q_1 \cdots q_{m+1}} \varepsilon^{q_1 \cdots q_{m+1}}{}_{q_{m+2} \cdots q_n}$$

we obtain

$$\begin{aligned} D^\omega (\overline{*}\beta)_{r_{m+1} \cdots r_n} &= \sum_{l=m+1}^n (-1)^{l-m-1} (\overline{*}\beta)_{r_{m+1} \cdots \widehat{r_l} \cdots r_n | r_l} \\ &= \frac{1}{(m+1)!} \sum_{l=m+1}^n (-1)^{l-m-1} \left(\beta_{q_1 \cdots q_{m+1} | r_l} \, \varepsilon^{q_1 \cdots q_{m+1}}{}_{r_{m+1} \cdots \widehat{r_l} \cdots r_n} \right). \end{aligned}$$

In the following, we use the identity

$$\varepsilon^{i_1 \cdots i_p k_{p+1} \cdots k_n} \varepsilon_{j_1 \cdots j_p k_{p+1} \cdots k_n} = (n-p)! \delta^{i_1 \cdots i_p}_{j_1 \cdots j_p}$$

where the generalized Kronecker delta $\delta^{i_1 \cdots i_p}_{j_1 \cdots j_p} := +1$ (or -1), depending on whether $j_1 \cdots j_p$ is an even (or odd) permutation of $i_1 \cdots i_p$, and 0 otherwise.

$$
\begin{aligned}
&(n-m-1)!\,(m+1)!\overline{*}D^{\omega}\left(\overline{*}\beta\right)_{r_1\cdots r_m}\\
&= D^{\omega}\left(\overline{*}\beta\right)^{r_{m+1}\cdots r_n}\varepsilon_{r_{m+1}\cdots r_n r_1\cdots r_m}\\
&= \left(\sum_{l=m+1}^{n}(-1)^{l-m-1}\beta_{q_1\cdots q_{m+1}|r_l}\varepsilon^{q_1\cdots q_{m+1}r_{m+1}\cdots\widehat{r_l}\cdots r_n}\right)\varepsilon_{r_{m+1}\cdots r_n r_1\cdots r_m}\\
&= \left(\sum_{l=m+1}^{n}(-1)^{m(n-m)}\beta_{q_1\cdots q_{m+1}|r_l}\varepsilon^{q_1\cdots q_{m+1}r_{m+1}\cdots\widehat{r_l}\cdots r_n}\right)\varepsilon_{r_1\cdots r_m r_l r_{m+1}\cdots\widehat{r_l}\cdots r_n}\\
&= (n-m-1)!\,(-1)^{mn-m}\sum_{l=m+1}^{n}\beta_{q_1\cdots q_{m+1}}{}^{|r_l}\delta^{q_1\cdots q_{m+1}}_{r_1\cdots r_m r_l}\\
&= (n-m-1)!\,(m+1)!\,(-1)^{mn-m}\sum_{l=m+1}^{n}\beta_{r_1\cdots r_m r_l}{}^{|r_l}\\
&= (n-m-1)!\,(m+1)!\,(-1)^{mn}\sum_{l=m+1}^{n}\beta^{r_l}{}_{r_1\cdots r_m|r_l},
\end{aligned}
$$

and multiplication by $-\frac{(-1)^{nm}}{(n-m-1)!(m+1)!}$ yields (16.93). □

If we regard $\overline{\Omega}^{k}(P,W)\cong\overline{\Omega}^{k}(P\times_f FM,W)\subseteq\overline{\Omega}^{0}\left(P\times_f FM,W\otimes T^{0,k}\right)$, then we have

$$\nabla^{\omega\oplus\theta}:\overline{\Omega}^{k}(P,W)\to\overline{\Omega}^{1}\left(P\times_f FM,W\otimes T^{0,k}\right).$$

We denote the formal adjoint of this map by $\left(\nabla^{\omega\oplus\theta}\right)^{*}$. We then have the so-called *connection Laplacian*

$$\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}:\overline{\Omega}^{k}(P,W)\to\overline{\Omega}^{k}(P,W).$$

Since $\nabla^{\omega\oplus\theta}$ is the same as $D^{\omega\oplus\theta}$ on $\overline{\Omega}^{0}\left(P\times_f FM,W\otimes T^{0,k}\right)$, we can use Proposition 16.54 (with $m=0$) to obtain

COROLLARY 16.55. *For* $\alpha\in\overline{\Omega}^{k}(P,W)\cong\overline{\Omega}^{k}(P\times_f FM,W)\cong\Omega^{k}(M,P\times_G W)$, *we have*

$$\left(\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}=-\alpha_{i_1\cdots i_k|}{}^{j}{}_{j}$$

We also have the *Hodge Laplacian*

$$\Delta^{\omega}:=\delta^{\omega}D^{\omega}+D^{\omega}\delta^{\omega}:\overline{\Omega}^{k}(P\times_f FM,W)\to\overline{\Omega}^{k}(P\times_f FM,W)$$

which does not depend on the Levi-Civita connection and only depends on the metric h on M via the Hodge star operator. The next result Bochner-Weitzenböck formula relates the two Laplacians $\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}$ and Δ^{ω}. First, there is some notation. Let

$$\Omega^{\omega}_{ij}=\left(\pi_1^{*}\Omega^{\omega}\right)\left(\overline{e_i},\overline{e_j}\right),$$

where $\pi_1:P\times_f FM\to P$ and Ω^{ω} is the curvature form of ω. Let

$$R_{ijkl}:=e_i\cdot\left(\Omega^{\theta}\left(\overline{e_k},\overline{e_l}\right)(e_j)\right).$$

These are the components of the Riemann curvature tensor R regarded as in $\overline{\Omega}^0\left(FM, T^{0,4}\left(\mathbb{R}^n\right)\right)$ or $\overline{\Omega}^0\left(P\times_f FM, T^{0,4}\left(\mathbb{R}^n\right)\right)$. The components of the Ricci curvature are then

$$R_{jl} := \delta^{ik} R_{ijkl} = R^i{}_{jil}.$$

THEOREM 16.56 (Bochner-Weitzenböck Formula). *For $\alpha\in\overline{\Omega}^k(P,W)$, we have*

$$\begin{aligned}(\Delta^\omega\alpha)_{i_1\cdots i_k} &= \left(\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}\\ &+\sum_{l=1}^k(-1)^l\rho'\left(\Omega^\omega_{i_l j}\right)\left(\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k}\right)-\sum_{l=1}^k(-1)^l R_{i_l j}\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k}\\ &-\sum_{l\neq m=1}^k(-1)^l R^{j_m}{}_{i_m i_l j}\alpha^j{}_{i_1\cdots i_{m-1}j_m i_{m+1}\cdots\widehat{i_l}\cdots i_k}.\end{aligned}\tag{16.94}$$

PROOF. Using Proposition 16.54, we have

$$\begin{aligned}(\delta^\omega D^\omega\alpha)_{i_1\cdots i_k} &= -\left(D^\omega\alpha\right)^j{}_{i_1\cdots i_k|j}\\ &= -\alpha_{i_1\cdots i_k|}{}^j{}_j-\sum_{l=1}^k(-1)^l\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|i_l j}\ \text{and}\\ (D^\omega\delta^\omega\alpha)_{i_1\cdots i_k} &= \sum_{l=1}^k(-1)^{l+1}\left(\delta^\omega\alpha\right)_{i_1\cdots\widehat{i_l}\cdots i_k|i_l}\\ &= -\sum_{l=1}^k(-1)^{l+1}\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|j i_l}.\end{aligned}$$

Thus,

$$(\Delta^\omega\alpha)_{i_1\cdots i_k} = -\alpha_{i_1\cdots i_k|}{}^j{}_j+\sum_{l=1}^k(-1)^l\left(\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|j i_l}-\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|i_l j}\right)$$

The first term $-\alpha_{i_1\cdots i_k|}{}^j{}_j=\left(\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}$. Regarding $\alpha\in\overline{\Omega}^0\left(P\times_f FM, W\otimes T^{0,k}\right)$ and using (16.92) and (16.13), we have

$$\begin{aligned}\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|j i_l}-\alpha^j{}_{i_1\cdots\widehat{i_l}\cdots i_k|i_l j} &= \left(D^{\omega\oplus\theta}\left(D^{\omega\oplus\theta}\alpha\right)\right)^j{}_{i_1\cdots\widehat{i_l}\cdots i_k;i_l j}\\ &= \left(\left(\rho\otimes t^{0,k}\right)'\left(\Omega^{\omega\oplus\theta}\right)(\alpha)\right)^j_{i_1\cdots\widehat{i_l}\cdots i_k;i_l j}.\end{aligned}$$

Recall that for $w\in W$, $\beta\in\Lambda^k\left(\mathbb{R}^{n*}\right)$, $g\in G$ and $B\in\mathrm{SO}(n)$, we have

$$\begin{aligned}&\left(\left(\left(\rho\otimes t^{0,k}\right)(g,A)\right)(w\otimes\beta)\right)_{i_1\cdots i_k}\\ &=\left(B^{-1}\right)^{j_1}{}_{i_1}\ldots\left(B^{-1}\right)^{j_k}{}_{i_k}\beta_{j_1\cdots j_k}\rho(g)(w).\end{aligned}$$

Thus, for $A\in\mathfrak{g}$ and $C\in\mathfrak{so}(n)$, we have

$$\begin{aligned}&\left(\left(\rho\otimes t^{0,k}\right)'(A,C)(w\otimes\beta)\right)_{i_1\cdots i_k}\\ &=\left(\left(\rho'\otimes I+I\otimes\left(t^{0,k}\right)'\right)(A,C)(w\otimes\beta)\right)_{i_1\cdots i_k}\\ &=\rho'(A)(w)\beta_{j_1\cdots j_k}-\left(C^{j_1}{}_{i_1}\beta_{j_1\cdots j_k}+\ldots+C^{j_k}{}_{i_k}\beta_{j_1\cdots j_k}\right)w.\end{aligned}$$

Hence,

$$\begin{aligned}&\left(\left(\rho\otimes t^{0,k}\right)'\left(\Omega^{\omega\oplus\theta}\right)(\alpha)\right)_{i_0 i_1\cdots\widehat{i_l}\cdots i_k;i_l j}\\&=\rho'\left(\Omega^{\omega}_{i_l j}\right)\left(\alpha_{i_0 i_1\cdots\widehat{i_l}\cdots i_k}\right)-R^{j_0}_{\ \ i_0 i_l j}\alpha_{j_0 i_1\cdots\widehat{i_l}\cdots i_k}\\&-\sum_{m\in\left\{1,\ldots,\widehat{l},\ldots,k\right\}}R^{j_m}_{\ \ i_m i_l j}\alpha_{i_0 i_1\cdots i_{m-1}j_m i_{m+1}\cdots\widehat{i_l}\cdots i_k},\end{aligned}$$

where in the sum, we do not mean to imply that $m<l$, only that $m\neq l$. Thus, raising i_0 and contracting i_0 with j, we obtain

$$\begin{aligned}&\left(\left(\rho\otimes t^{0,k}\right)'\left(\Omega^{\omega\oplus\theta}\right)(\alpha)\right)^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k;i_l j}\\&=\rho'\left(\Omega^{\omega}_{i_l j}\right)\left(\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k}\right)-R^{j_0}_{\ \ i_l}\alpha_{j_0 i_1\cdots\widehat{i_l}\cdots i_k}\\&-\sum_{m\in\left\{1,\ldots,\widehat{l},\ldots,k\right\}}R^{j_m}_{\ \ i_m i_l j}\alpha^{j}_{\ \ i_1\cdots i_{m-1}j_m i_{m+1}\cdots\widehat{i_l}\cdots i_k}\end{aligned}$$

Finally we obtain

$$\begin{aligned}&\left(\Delta^{\omega}\alpha\right)_{i_1\cdots i_k}=\left(\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}\\&+\sum_{l=1}^{k}(-1)^{l}\left(\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k|j i_l}-\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k|i_l j}\right)\\&=\left(\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}\\&+\sum_{l=1}^{k}(-1)^{l}\left(\left(\rho\otimes t^{0,k}\right)'\left(\Omega^{\omega\oplus\theta}\right)(\alpha)\right)^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k;i_l j}\\&=\left(\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}\\&+\sum_{l=1}^{k}(-1)^{l}\left(\begin{array}{c}\rho'\left(\Omega^{\omega}_{i_l j}\right)\left(\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k}\right)-R^{j_0}_{\ \ i_l}\alpha_{j_0 i_1\cdots\widehat{i_l}\cdots i_k}\\-\sum_{m\in\left\{1,\ldots,\widehat{l},\ldots,k\right\}}R^{j_m}_{\ \ i_m i_l j}\alpha^{j}_{\ \ i_1\cdots i_{m-1}j_m i_{m+1}\cdots\widehat{i_l}\cdots i_k}\end{array}\right)\\&=\left(\left(\nabla^{\omega\oplus\theta}\right)^{*}\nabla^{\omega\oplus\theta}\alpha\right)_{i_1\cdots i_k}+\sum_{l=1}^{k}(-1)^{l}\rho'\left(\Omega^{\omega}_{i_l j}\right)\left(\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k}\right)\\&-\sum_{l=1}^{k}(-1)^{l}R_{i_l j}\alpha^{j}_{\ \ i_1\cdots\widehat{i_l}\cdots i_k}-\sum_{l\neq m=1}^{k}(-1)^{l}R^{j_m}_{\ \ i_m i_l j}\alpha^{j}_{\ \ i_1\cdots i_{m-1}j_m i_{m+1}\cdots\widehat{i_l}\cdots i_k},\end{aligned}$$

as required. □

We consider some special cases.

1. ($k=1$) For 1-forms $\alpha\in\overline{\Omega}^{1}(P,W)\cong\Omega^{1}(M,P\times_{G}W)$, the last term in (16.94) is absent. Taking the L^2 inner product of $\Delta^{\omega}\alpha$ with α, we then have

$$\left(\Delta^{\omega}\alpha,\alpha\right)=\left\|\nabla^{\omega\oplus\theta}\alpha\right\|^{2}+\left(Ric\left(\alpha\right),\alpha\right)-\left(\rho'\left(\Omega^{\omega}\right)\alpha,\alpha\right).$$

It follows that if $Ric-\rho'\left(\Omega^{\omega}\right)\in\mathrm{End}\left(\Omega^{1}\left(M,P\times_{G}W\right)\right)$ is pointwise nonnegative, but not zero everywhere, then $\Delta^{\omega}\alpha=0\Rightarrow\alpha=0$. In particular, if G is trivial, we obtain S. Bochner's result (see [**Boc**]) that a compact, Riemannian manifold with positive Ricci curvature admits no nonzero harmonic 1-form.

2. $(k=2)$ For $\alpha \in \overline{\Omega}^2(P,W) \cong \overline{\Omega}^2(P,W) \cong \Omega^2(M, P\times_G W)$, we get

$$\begin{aligned}(\Delta^\omega\alpha)_{i_1i_2} &= \left(\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha\right)_{i_1i_2}\\ &-\left(\rho'\left(\Omega^\omega_{i_1j}\right)\left(\alpha^j{}_{i_2}\right)-\rho'\left(\Omega^\omega_{i_2j}\right)\left(\alpha^j{}_{i_1}\right)\right)\\ &+\left(R_{i_1j}\alpha^j{}_{i_2}-R_{i_2j}\alpha^j{}_{i_1}\right)+\left(R^m{}_{i_2i_1j}\alpha^j{}_m-R^m{}_{i_1i_2j}\alpha^j{}_m\right).\end{aligned}$$

In terms of the product $\vee$ (see (16.65), p. 399) and using the hat "$\wedge$" for converting curvature type tensors to operators on 2-forms (see (16.75), p. 402), we have

$$R_{i_1j}\alpha^j{}_{i_2}-R_{i_2j}\alpha^j{}_{i_1}=2\left((h\vee \mathrm{Ric})^\wedge(\alpha)\right)_{i_1i_2}.$$

Using the first Bianchi identity (16.53) and known symmetries of R,

$$\begin{aligned}&R^m{}_{i_2i_1j}\alpha^j{}_m-R^m{}_{i_1i_2j}\alpha^j{}_m=R^m{}_{i_2i_1j}\alpha^j{}_m+R^m{}_{i_1ji_2}\alpha^j{}_m\\ &=-R^m{}_{ji_2i_1}\alpha^j{}_m=R^m{}_{ji_1i_2}\alpha^j{}_m=-R_{i_1i_2jm}\alpha^{jm}=-2\widehat{R}(\alpha)_{i_1i_2},\end{aligned}$$

where $R(\alpha)$ is the image of α under the curvature operator R (see (16.75), p. 402). Moreover, it is convenient to define $[\rho'(\Omega^\omega),\alpha]$ via

$$[\rho'(\Omega^\omega),\alpha]_{i_1i_2}:=\rho'\left(\Omega^\omega_{i_1j}\right)\left(\alpha^j{}_{i_2}\right)-\rho'\left(\Omega^\omega_{i_2j}\right)\left(\alpha^j{}_{i_1}\right). \tag{16.95}$$

Then (for $\alpha\in\overline{\Omega}^2(P,W)$), we can write

$$\Delta^\omega\alpha=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+2(h\vee Ric)^\wedge(\alpha)-2\widehat{R}(\alpha).$$

When $n=\dim M=3$, note that by Proposition 16.50 (p. 16.50) and the fact that the Weyl tensor is zero for $n=3$,

$$R=\tfrac{1}{6}S(h\vee h)+2h\vee\left(Ric-\tfrac{1}{3}Sh\right)=h\vee\left(2Ric-\tfrac{1}{2}Sh\right)$$

Thus for $n=3$,

$$\begin{aligned}\Delta^\omega\alpha&=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+2(h\vee Ric)^\wedge(\alpha)-2\widehat{R}(\alpha)\\ &=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+2(h\vee Ric)^\wedge(\alpha)\\ &-2\left(h\vee\left(2\mathrm{Ric}-\tfrac{1}{2}Sh\right)\right)^\wedge(\alpha)\\ &=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]-2(h\vee Ric)^\wedge(\alpha)+S\alpha.\end{aligned}$$

If $n=4$, we have

$$R=\tfrac{1}{12}S(h\vee h)+h\vee\left(Ric-\tfrac{1}{4}Sh\right)+W,$$

and so

$$\begin{aligned}\Delta^\omega\alpha&=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+2(h\vee Ric)^\wedge(\alpha)-2\widehat{R}(\alpha)\\ &=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+2(h\vee Ric)^\wedge(\alpha)\\ &-2\left(\tfrac{1}{12}S(h\vee h)+h\vee\left(Ric-\tfrac{1}{4}Sh\right)+W\right)^\wedge(\alpha)\\ &=\left(\nabla^{\omega\oplus\theta}\right)^*\nabla^{\omega\oplus\theta}\alpha-[\rho'(\Omega^\omega),\alpha]+\tfrac{1}{3}S\alpha+W(\alpha).\end{aligned}$$

We have decompositions

$$\alpha=\alpha^++\alpha^-,\ \Omega^\omega=\Omega^{\omega+}+\Omega^{\omega-}\text{ and }W=W^++W^-$$

into self-dual and anti-self-dual parts. Under the $\mathrm{SO}(4)$-equivariant isomorphism $\mathfrak{so}(4) \cong \Lambda^2(\mathbb{R}^{4*})$ given by lowering an index, the irreducible, $\mathrm{SO}(4)$-invariant subspaces Λ^+ and Λ^-, correspond to irreducible, subspaces, say $\mathfrak{so}^+$ and $\mathfrak{so}^-$ which are $\mathrm{SO}(4)$-invariant with respect to the adjoint action. Thus,

$$\mathfrak{so}(4) \cong \mathfrak{so}^+ \oplus \mathfrak{so}^- \tag{16.96}$$

with $[\mathfrak{so}(4), \mathfrak{so}^\pm] \subseteq \mathfrak{so}^\pm$, and so $[\mathfrak{so}^-, \mathfrak{so}^+] \subseteq \mathfrak{so}^- \cap \mathfrak{so}^+ = \{0\}$ and $[\mathfrak{so}^\pm, \mathfrak{so}^\pm] \subseteq \mathfrak{so}^\pm$. It then follows from (16.95) that $[\rho'(\Omega^{\omega-}), \alpha^+] = [\rho'(\Omega^{\omega+}), \alpha^-] = 0$, where the fact that α is W-valued is irrelevant to the argument. Thus,

$$\begin{aligned}\Delta^\omega \alpha = {} & \left(\nabla^{\omega\oplus\theta}\right)^* \nabla^{\omega\oplus\theta}\alpha - \left[\rho'\left(\Omega^{\omega+}\right), \alpha^+\right] - \left[\rho'\left(\Omega^{\omega-}\right), \alpha^-\right] \\ & + \tfrac{1}{3} S\alpha + W^+\left(\alpha^+\right) + W^-\left(\alpha^-\right).\end{aligned}$$

The following direct consequence will be useful to us later.

Proposition 16.57. *When $\alpha \in \overline{\Omega}^2(P, W)$ is anti-self-dual ($\alpha^+ = 0$), Ω^ω is self-dual ($\Omega^{\omega-} = 0$) and h is a self-dual metric (i.e., $Weyl^- = 0$), we have*

$$\Delta^\omega \alpha = \left(\nabla^{\omega\oplus\theta}\right)^* \nabla^{\omega\oplus\theta}\alpha + \tfrac{1}{3} S\alpha.$$

Thus,

$$\left(\Delta^\omega \alpha, \alpha\right) = \left\|\nabla^{\omega\oplus\theta}\alpha\right\|^2 + \tfrac{1}{3}\left(S\alpha, \alpha\right).$$

If $S \geq 0$ and $\Delta^\omega \alpha = 0$, then either $S = 0$ and $\nabla^{\omega\oplus\theta}\alpha = 0$, or $S \neq 0$ and $\alpha = 0$.

7. Characteristic Classes and Curvature Forms

Let $E \to M$ be a C^∞ complex, Hermitian vector bundle of complex dimension m, with Hermitian inner product $\langle \cdot, \cdot \rangle$. For $x \in M$, a (unitary) *frame* of E_x is a linear, isometry $u : \mathbb{C}^m \to E_x$ (i.e., $\langle u(z), u(w)\rangle = z_1\overline{w_1} + \cdots + z_m\overline{w_m}$). If u is a frame and $A \in U(m)$, then $uA := u \circ A$ is a frame. The set $U(E)$ consisting of all frames at all points of M can be made into a C^∞ manifold, such that $\pi : U(E) \to M$ is a principal $U(m)$-bundle, namely the bundle of unitary frames of E. Without difficulty, one can prove that $U(E) \times_{U(m)} \mathbb{C}^m \cong E$, via $[u, w] \mapsto u(w)$, where the representation is inclusion $U(m) \to \mathrm{GL}(m, \mathbb{C})$.

Let ω be a connection 1-form on $U(E)$, which is a 1-form with values in the Lie algebra $\mathfrak{u}(m) = \{A \in \mathfrak{gl}(m, \mathbb{C}) : A^* = -A\}$ of $U(m)$, and with the required properties of Definition 16.4, p. 368. We will express the Chern classes of $E \to M$ in terms of the curvature Ω^ω of ω.

We begin by defining functions $s_k : \mathfrak{gl}(m, \mathbb{C}) \to \mathbb{C}$ by means of

$$\det(A + tI) = \sum_{k=0}^{m} s_k(A)\, t^{m-k}.$$

Note that $s_k(A)$ is a homogeneous polynomial of degree k in the entries of A, namely

$$s_k(A) = \frac{1}{k!} \sum_{(i),(j)} \delta^{j_1\cdots j_k}_{i_1\cdots i_k} a^{i_1}{}_{j_1} \cdots a^{i_k}{}_{j_k},$$

where $A = \left(a^i{}_j\right)$, $(i) := (i_1, \ldots, i_k)$ ranges over all sequences of k distinct elements of $\{1, \ldots, m\}$, and $\delta^{j_1\cdots j_k}_{i_1\cdots i_k} = +1$ (resp. -1), depending on whether (i) is an even (resp. odd) permutation of (j), and $\delta^{j_1\cdots j_k}_{i_1\cdots i_k} = 0$ if $\{i_1, \ldots, i_k\} \neq \{j_1, \ldots, j_k\}$. The s_k are invariant under the adjoint action of $\mathrm{GL}(m, \mathbb{C})$ on $\mathfrak{gl}(m, \mathbb{C})$, in the sense that $s_k\left(BAB^{-1}\right) = s_k(A)$ for all $A \in \mathfrak{gl}(m, \mathbb{C})$ and $B \in \mathrm{GL}(m, \mathbb{C})$, since

$$\det\left(BAB^{-1} + tI\right) = \det\left(B(A + tI)B^{-1}\right) = \det(A + tI).$$

The curvature $\Omega^\omega \in \overline{\Omega}^2(U(E), \mathfrak{u}(m))$ can be regarded as a matrix of $\mathbb{C}$-valued 2-forms, say $\Omega^\omega = \left(\Omega^i{}_j\right)$, such that $\Omega^i{}_j = -\overline{\Omega}^j{}_i$, and we define

$$s_k(\Omega^\omega) := \frac{1}{k!} \sum_{(i),(j)} \delta^{j_1\cdots j_k}_{i_1\cdots i_k} \Omega^{i_1}{}_{j_1} \wedge \cdots \wedge \Omega^{i_k}{}_{j_k} \in \Omega^{2k}(U(E), \mathbb{C}).$$

From the fact $R_g^*\Omega^\omega = ad_{g^{-1}}\Omega^\omega$ for all $g \in U(m)$ and the invariance of s_k under the adjoint action, it follows that $s_k(\Omega^\omega)$ is invariant under R_g^*. Since Ω^ω also vanishes on vertical vectors, we know that there is a $2k$-form, say $\sigma_k(\Omega^\omega) \in \Omega^{2k}(M, \mathbb{C})$, on M such that $s_k(\Omega^\omega) = \pi^*\sigma_k(\Omega^\omega)$. Note that

$$\pi^*(d\sigma_k(\Omega^\omega)) = d(\pi^*\sigma_k(\Omega^\omega)) = d(\pi^*\sigma_k(\Omega^\omega))^H = d(s_k(\Omega^\omega))^H$$

$$= \frac{1}{k!} \sum_{(i),(j)} \sum_{p=1}^{k} \delta^{j_1\cdots j_k}_{i_1\cdots i_k} \Omega^{i_1}{}_{j_1} \wedge \cdots \wedge (d\Omega^{i_p}{}_{j_p})^H \wedge \cdots \wedge \Omega^{i_k}{}_{j_k} = 0,$$

since $(\Omega^i{}_j)^H = \Omega^i{}_j$ and $(d\Omega^{i_p}{}_{j_p})^H = (D^\omega\Omega^\omega)^{i_p}{}_{j_p} = 0$ by the Bianchi identity. Thus, $\sigma_k(\Omega^\omega)$ is closed and determines a de Rham cohomology class $[\sigma_k(\Omega^\omega)] \in H^{2k}(M, \mathbb{C})$.

Definition 16.58. The **k-th Chern class** of the complex, Hermitian vector bundle E is defined as

$$c_k(E) := \left(\tfrac{i}{2\pi}\right)^k [\sigma_k(\Omega^\omega)] = \left[\sigma_k\left(\tfrac{i}{2\pi}\Omega^\omega\right)\right]$$

The form

$$c_k(E, \omega) := \sigma_k\left(\frac{i}{2\pi}\Omega^\omega\right) \in \Omega^{2k}(M, \mathbb{C})$$

is the **k-th Chern form** of the complex Hermitian vector bundle $E \to M$ for the connection 1-form ω on $U(E)$.

The factor i in $\frac{i}{2\pi}\Omega^\omega$ ensures that $\sigma_k\left(\frac{i}{2\pi}\Omega^\omega\right) \in \Omega^{2k}(M, \mathbb{R})$, since

$$\overline{i\Omega}^{i_q}{}_{j_q} = -i\overline{\Omega}^{i_q}{}_{j_q} = i\Omega^{j_q}{}_{i_q} \Rightarrow \overline{\sigma_k(i\Omega^\omega)} = \sigma_k\left(\overline{i\Omega^\omega}\right) = \sigma_k(i\Omega^\omega),$$

and so $c_k(E) \in H^{2k}(M, \mathbb{R})$. The factor of 2π in $\frac{i}{2\pi}\Omega^\omega$ is a normalization implying $\left[\sigma_k\left(\frac{i}{2\pi}\Omega^\omega\right)\right] \in H^{2k}(M, \mathbb{Z})$. A full proof of this would carry us too far afield. However, we will show that $[\sigma_k(\Omega^\omega)]$, and hence $c_k(E)$, is independent of the choice of ω. Indeed, let

$$\alpha := \omega_1 - \omega_0 \in \overline{\Omega}^1(U(E), \mathfrak{u}(m)) \quad \text{and} \quad \omega_t := \omega_0 + t\alpha,\ 0 \le t \le 1$$

Then ω_t is a connection with curvature $\Omega^{\omega_t} = d\omega_t + \frac{1}{2}[\omega_t, \omega_t]$. Let $\widetilde{s}_k$ be the k-linear symmetric form such that $s_k(A) = \widetilde{s}_k\left(A, \overset{k}{\ldots}, A\right)$. Note that the invariance of s_k and hence $\widetilde{s}_k$ under the adjoint action yields (at $u = 0$)

$$\begin{aligned} 0 &= \tfrac{d}{du}\left(\widetilde{s}_k\left(\exp(uB)A_1\exp(-uB), \overset{k}{\ldots}, \exp(uB)A_k\exp(-uB)\right)\right) \\ &= \widetilde{s}_k([B, A_1], A_2, \ldots, A_k) + \widetilde{s}_k(A_1, [B, A_2], A_3, \ldots, A_k) \\ &\quad + \ldots + \widetilde{s}_k(A_1, A_2, A_3, \ldots, [B, A_k]) \end{aligned}$$

Using this and the Bianchi identity $d\Omega^{\omega_t} + [\omega_t, \Omega^{\omega_t}] = 0$, we have

$$\begin{aligned}
&\frac{1}{k}\frac{d}{dt}s_k\left(\Omega^{\omega_t}\right) = \widetilde{s}_k\left(\tfrac{d}{dt}\left(\Omega^{\omega_t}\right), \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right)\\
&= \widetilde{s}_k\left(d\alpha + [\omega_t, \alpha], \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right)\\
&= \widetilde{s}_k\left(d\alpha, \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right) + \widetilde{s}_k\left([\omega_t, \alpha], \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right)\\
&= \widetilde{s}_k\left(d\alpha, \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right) - (k-1)\,\widetilde{s}_k\left(\alpha, [\omega_t, \Omega^{\omega_t}], \Omega^{\omega_t}, \overset{k-2}{\dots}, \Omega^{\omega_t}\right)\\
&= \widetilde{s}_k\left(d\alpha, \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right) + (k-1)\,\widetilde{s}_k\left(\alpha, d\Omega^{\omega_t}, \Omega^{\omega_t}, \overset{k-2}{\dots}, \Omega^{\omega_t}\right)\\
&= d\left(\widetilde{s}_k\left(\alpha, \Omega^{\omega_t}, \overset{k-1}{\dots}, \Omega^{\omega_t}\right)\right).
\end{aligned}$$

Thus,

$$s_k\left(\Omega^{\omega_1}\right) - s_k\left(\Omega^{\omega_0}\right) = kd\left(\int_0^1 \widetilde{s}_k\left(\alpha, \Omega^{\omega_t}, \dots, \Omega^{\omega_t}\right)\,dt\right).$$

Since, $\widetilde{s}_k\left(\alpha, \Omega^{\omega_t}, \dots, \Omega^{\omega_t}\right) \in \overline{\Omega}^{2k}\left(U(E), \mathbb{C}\right)$ is invariant under R_{g*},

$$\int_0^1 \widetilde{s}_k\left(\alpha, \Omega^{\omega_t}, \dots, \Omega^{\omega_t}\right)\,dt = \pi^*\beta,$$

for some unique form $\beta \in \Omega^{2k}(M, \mathbb{C})$, and $\sigma_k\left(\Omega^{\omega_1}\right) - \sigma_k\left(\Omega^{\omega_0}\right) = d\beta$. Hence, $[\sigma_k(\Omega^{\omega_0})] = [\sigma_k(\Omega^{\omega_1})]$, as required.

Remark 16.59. Suppose that $\mathrm{U}(E)$ is reducible to an $\mathrm{SU}(m)$-bundle, say we have a subprincipal $\mathrm{SU}(m)$-bundle $\mathrm{U}(E)_0 \to M$. Then we show that $c_1(E) = 0$. Let ω_0 be an arbitrary connection 1-form on $\mathrm{U}(E)_0$. We can extend the distribution of horizontal subspaces for ω_0 on $\mathrm{U}(E)_0$ to all of $\mathrm{U}(E)$ by the requiring that the distribution be R_{g*}-invariant for all $g \in \mathrm{U}(m)$. Let ω be the resulting connection on $\mathrm{U}(E)$. We know that ω_0 and Ω^{ω_0} are $\mathfrak{su}(m)$-valued. While ω is $\mathfrak{u}(m)$-valued and has values outside $\mathfrak{su}(m)$, we can show that Ω^ω is $\mathfrak{su}(m)$-valued. Indeed,

$$\Omega^\omega|_{\mathrm{U}(E)_0} = (d\omega)^H|_{\mathrm{U}(E)_0} = (d\omega_0)^H = \Omega^{\omega_0}.$$

Thus, Ω^ω has values in $\mathfrak{su}(m)$ on ω-horizontal subspaces at points of $\mathrm{U}(E)_0$. Since $\mathfrak{su}(m)$ is invariant under the adjoint action of $\mathrm{U}(m)$ and $R_g^*\Omega^\omega = ad_{g^{-1}}\Omega^\omega$, we know that Ω^ω is $\mathfrak{su}(m)$-valued throughout $\mathrm{U}(E)$. Since $s_1(A) = \mathrm{Tr}(A) = 0$ for $A \in \mathfrak{su}(m)$, we have $c_1(E) = 0$, when $\mathrm{U}(E)$ is reducible to an $\mathrm{SU}(m)$-bundle.

Once the Chern classes $c_k(E)$ are determined, the Chern character $\mathbf{ch}(E) \in H^*(M, \mathbb{Q})$ may be defined in terms of the $c_k(E)$. Alternatively, we can get $\mathbf{ch}(E)$ directly as follows. For $A \in \mathfrak{u}(m)$,

$$\mathrm{Tr}\left(\exp\left(t\frac{i}{2\pi}A\right)\right) = \sum_{k=0}^{\infty} r_k(A)\,t^k, \text{ for } r_k(A) := \frac{1}{k!}\mathrm{Tr}\left(\left(\frac{i}{2\pi}A\right)^k\right).$$

As with the s_k, the r_k are invariant under the adjoint action of $\mathrm{U}(m)$ on $\mathfrak{u}(m)$. Hence the horizontal form $r_k(\Omega^\omega)$ is R_g^*-invariant and $r_k(\Omega^\omega) = \pi^*(\rho_k(\Omega^\omega))$ for some closed form $\rho_k(\Omega^\omega) \in H^{2k}(M, \mathbb{R})$ (actually $H^{2k}(M, \mathbb{Q})$, where $\mathbb{Q}$ is the field of rational numbers) whose class $[\rho_k(\Omega^\omega)]$ is independent of ω. Then

$$\mathbf{ch}(E) := \sum_{k=0}^{\infty}[\rho_k(\Omega^\omega)] \in \bigoplus_{k=0}^{\left[\frac{1}{2}\dim M\right]} H^{2k}(M, \mathbb{Q}). \tag{16.97}$$

Note that $8\pi^2 r_2(A) = -Tr\left(A^2\right)$ for $A \in \mathfrak{u}(m)$, and so

$$ch(E)_2 = [\rho_2(\Omega^\omega)] = \frac{-1}{8\pi^2}[\mathrm{Tr}(\Omega^\omega \wedge \Omega^\omega)] \in H^4(M, \mathbb{Q}), \tag{16.98}$$

where we have regarded $\Omega^\omega \in \Omega^2(M, \mathrm{End}(E)) \subseteq \overline{\Omega}^2(U(E), \mathfrak{u}(m))$ and composition of endomorphisms is implicit in the wedge $\Omega^\omega \wedge \Omega^\omega$.

Other characteristic classes can be represented by forms. The Euler class of an oriented Riemannian $2m$-manifold is represented by the *Gauss-Bonnet form*

$$\mathrm{GB}\left(\Omega^\theta\right) := \frac{1}{2^{2m}\pi^m m!}\sum_{(i)} \varepsilon_{i_1\cdots i_{2m}}\Omega^\theta_{i_1 i_2}\wedge\cdots\wedge\Omega^\theta_{i_{2m-1}i_{2m}} \tag{16.99}$$

where Ω^θ is the curvature form of any connection θ (not necessarily Levi-Civita) on the principal $\mathrm{SO}(2m)$-bundle $\pi_F : FM \to M$ of oriented orthonormal frames. In (16.99), we may regard $\Omega^\theta \in \Omega^2(M, \mathrm{End}(TM))$ and the components Ω^θ_{ij} are relative to a locally defined orthonormal frame field. Alternatively, (16.99) may be regarded as the unique form on M which when pulled back to FM via π_F^* is the form given by the same formula, but where $\Omega^\theta_{ij} = \Omega^\theta(\overline{e_i}, \overline{e_j})$, the components of $\Omega^\theta \in \overline{\Omega}^2(FM, \mathfrak{so}(2m))$ relative to the standard, horizontal fields $\overline{e_1}, \ldots, \overline{e_{2m}}$. The form (16.99) arises from the homogenous polynomial of degree m on $\mathfrak{so}(2m)$ known as the Pfaffian, defined for $A \in \mathfrak{so}(2m)$, by

$$\mathrm{Pf}(A) := \frac{(-1)^m}{2^m m!}\sum_{(i)} \varepsilon_{i_1\cdots i_{2m}} A_{i_1 i_2}\cdots A_{i_{2m-1}i_{2m}}.$$

Thus, $\mathrm{GB}\left(\Omega^\theta\right) = \mathrm{Pf}\left(\frac{1}{2\pi}\Omega^\theta\right)$. The Pfaffian is invariant under the adjoint action of $\mathrm{SO}(2m)$ (but not $\mathrm{O}(2m)$). Indeed, for $B \in \mathrm{O}(2m)$, we have

$$\begin{aligned}
\mathrm{Pf}\left(BAB^{-1}\right) &= \mathrm{Pf}\left(BAB^T\right) \\
&= \frac{(-1)^m}{2^m m!}\sum_{(i)} \varepsilon_{i_1\cdots i_{2m}}\left(B_{i_1 j_1}A_{j_1 j_2}B_{i_2 j_2}\right)\cdots\left(B_{i_{2m-1}j_{2m-1}}A_{j_{2m-1}j_{2m}}B_{i_{2m}j_{2m}}\right) \\
&= \frac{(-1)^m}{2^m m!}\sum_{(i)} \varepsilon_{i_1\cdots i_{2m}} B_{i_1 j_1}B_{i_2 j_2}\cdots B_{i_{2m-1}j_{2m-1}}B_{i_{2m}j_{2m}}A_{j_1 j_2}\cdots A_{j_{2m-1}j_{2m}} \\
&= \det(B)\,\mathrm{Pf}(A).
\end{aligned}$$

The Gauss Bonnet Theorem, which is a special case of the Index Theorem, asserts that the integral of the Gauss-Bonnet form over (compact) M is $\chi(M)$.

We will also encounter the *Pontryagin classes* $p_k(M) \in H^{4k}(M, \mathbb{Z})$ of M. These may be defined in terms of the Chern classes of the complexified tangent bundle $T_{\mathbb{C}}M := \mathbb{C}\otimes TM$, which can be regarded as the associated bundle $FM \times_{\mathrm{SO}(n)} \mathbb{C}^n$ where the representation $\mathrm{SO}(n) \to \mathrm{U}(n)$ is just inclusion. Note that FM is a principal subbundle of the unitary frame bundle $U(T_{\mathbb{C}}M)$ of $T_{\mathbb{C}}M$, where the Hermitian metric H on $T_{\mathbb{C}}M$ is given in terms of the complex bilinear extension $h_{\mathbb{C}}$ of the Riemannain metric h via

$$H(X, Y) = h_{\mathbb{C}}\left(X, \bar{Y}\right) \text{ for } X, Y \in T_{\mathbb{C}}M.$$

A connection θ on FM determines a unique connection 1-form $\theta_c \in \Omega^1(U(T_{\mathbb{C}}M), \mathfrak{u}(n))$, such $\theta = \theta_c|_{FM}$. By definition, the Pontryagin class $p_k(M) \in H^{4k}(M, \mathbb{Z})$ is represented by the unique $4k$-form $\sigma_{2k}\left(\frac{1}{2\pi}\Omega^{\theta_c}\right) \in \Omega^{4k}(M)$, such that, for $\pi_c : U(T_{\mathbb{C}}M) \to M$,

$$\begin{aligned} p_k\left(\Omega^{\theta}\right) &:= {\pi_c}^* \sigma_{2k}\left(\tfrac{1}{2\pi}\Omega^{\theta_c}\right) = s_{2k}\left(\tfrac{1}{2\pi}\Omega^{\theta_c}\right) = (-i)^{2k} s_{2k}\left(\tfrac{i}{2\pi}\Omega^{\theta_c}\right) \\ &= (-1)^k s_{2k}\left(\tfrac{i}{2\pi}\Omega^{\theta_c}\right) = \frac{1}{(2\pi)^{2k}(2k)!}\sum_{(i),(j)} \delta^{j_1\cdots j_{2k}}_{i_1\cdots i_{2k}} \Omega^{\theta_c}_{i_1 j_1}\wedge\cdots\wedge\Omega^{\theta_c}_{i_{2k}j_{2k}}. \end{aligned} \tag{16.100}$$

In other words, since $c_{2k}(T_{\mathbb{C}}M) = \left[\sigma_{2k}\left(\frac{i}{2\pi}\Omega^{\theta_c}\right)\right]$, we have

$$p_k(M) = (-1)^k c_{2k}(T_{\mathbb{C}}M) \in H^{4k}(M,\mathbb{Z}). \tag{16.101}$$

Since $\Omega^{\theta} = \Omega^{\theta_c}|_{FM}$, it follows that $\sigma_{2k}\left(\Omega^{\theta_c}\right) = \sigma_{2k}\left(\Omega^{\theta}\right)$, where $\sigma_{2k}\left(\Omega^{\theta}\right)$ is the unique $4k$-form, such that for $\pi : FM \to M$,

$$\pi^*\sigma_{2k}\left(\Omega^{\theta}\right) = s_{2k}\left(\tfrac{1}{2\pi}\Omega^{\theta}\right) = \frac{1}{(2\pi)^{2k}(2k)!}\sum_{(i),(j)} \delta^{j_1\cdots j_{2k}}_{i_1\cdots i_{2k}} \Omega^{\theta}_{i_1 j_1}\wedge\cdots\wedge\Omega^{\theta}_{i_{2k}j_{2k}}$$

Note that Ω^{θ} has values in $\mathfrak{so}(n)$. For $A \in \mathfrak{so}(n)$ (so that $A^T = -A$), we have $s_k\left(\frac{1}{2\pi}\Omega^{\theta}\right) = 0$ for k odd, since

$$\begin{aligned} \sum_{k=0}^{n} s_k(A)\, t^{n-k} &= \det(A + tI) = \det\left(A^T + tI\right) = \det(-A + tI) \\ &= (-1)^n \det(A - tI) = (-1)^n \sum_{k=0}^{n} s_k(A)(-t)^{n-k} = \sum_{k=0}^{n} s_k(A)(-1)^k t^{n-k}. \end{aligned}$$

By the same argument, when E is a real Riemannian vector bundle we have $c_k(\mathbb{C}\otimes E) = 0$ for k odd. For more information on this approach to characteristic classes, see [**KN69**] and [**MilS**].

There are other characteristic classes that will arise in index theorems. Some of these are defined and manipulated more efficiently through the use of power series as follows. First note that for $A \in \mathfrak{gl}(\nu,\mathbb{C})$ with eigenvalues $\{\lambda_1,\ldots,\lambda_\nu\}$, there is $B \in \mathrm{GL}(m,\mathbb{C})$ such that BAB^{-1} is upper triangular with diagonal entries $\lambda_1,\ldots,\lambda_\nu$. Then

$$\sum_{k=0}^{\nu} s_k(A)\, t^{\nu-k} = \det(A + tI) = \prod_{j=1}^{\nu}(\lambda_j + t) = \sum_{k=0}^{\nu}\sigma_k(\lambda_1,\ldots,\lambda_\nu)\, t^{\nu-k},$$

where $\sigma_0 := 1$, $\sigma_1 := \sum_{i=1}^{\nu}\lambda_i$, and generally

$$\sigma_k(\lambda_1,\ldots,\lambda_\nu) := \sum\nolimits_{1\le i_1<\cdots<i_k\le\nu} \lambda_{i_1}\cdots\lambda_{i_k}$$

is the elementary symmetric polynomial of degree k in $x_1,\ldots,x_\nu$. We have seen that each $s_k(A)$ (and hence each σ_k) together with a Hermitian vector bundle $E \to M$ and connection gives rise to a characteristic form or class, namely the k-th Chern form $\sigma_k\left(\frac{i}{2\pi}\Omega^{\omega}\right)$ or class $c_k(E)$. Now any polynomial in the σ_k gives rise to a corresponding polynomial in the $\sigma_k\left(\frac{i}{2\pi}\Omega^{\omega}\right)$ where the multiplication is wedge product, or a polynomial in the $c_k(E)$ where the multiplication is cup product. By the Fundamental Theorem of Symmetric Polynomials, any symmetric polynomial in $(\lambda_1,\ldots,\lambda_\nu)$ can be expressed uniquely as a polynomial in $\sigma_1,\ldots,\sigma_\nu$. Hence, from a symmetric polynomial in $(\lambda_1,\ldots,\lambda_\nu)$, we can obtain "new" characteristic forms or classes, which are (however) ultimately polynomials in Chern forms or classes. One way to manufacture symmetric polynomials forming a symmetric function (e.g., the product) of a given power series in each λ_k, and considering the Taylor polynomials, as follows. Let $b(x) = \sum_{n=0}^{\infty} b_n x^n$ be a formal power series in the single variable x with $b_n \in \mathbb{C}$. We may form the product

(16.102)
$$b(x_1)\dots b(x_\nu) = \sum_{k=0}^{\infty} \sum_{k_1+\cdots+k_\nu=k}^{\infty} \beta_{k_1,\dots,k_\nu} x_1^{k_1}\cdots x_\nu^{k_\nu} = \sum_{k=0}^{\infty} B_k(\sigma_1,\dots,\sigma_k),$$

where $\sigma_j := \sigma_j(x_1,\dots,x_\nu)$ for $j = 1,\dots,\nu$. By the Fundamental Theorem of Elementary Polynomials, the $B_k(\sigma_1,\dots,\sigma_k)$ are uniquely determined by b. The sequence $\{B_k(\sigma_1,\dots,\sigma_k)\}$ is known as the *multiplicative sequence* determined by $b(x)$; the ideas here are due to F. Hirzebruch (see [**Hi66a**]). Given $A \in \mathfrak{gl}(\nu,\mathbb{C})$, and taking $\{x_1,\dots,x_\nu\} = \{i\lambda_1,\dots,i\lambda_\nu\}$ where $\{\lambda_1,\dots,\lambda_\nu\}$ is the set of eigenvalues of A, we obtain a function

$$A \mapsto B_k(A) := B_k(\sigma_1(i\lambda_1,\dots,i\lambda_\nu),\dots,\sigma_k(i\lambda_1,\dots,i\lambda_\nu))$$

which (like the set of eigenvalues of A) is invariant under the adjoint action. Note that if $A \in \mathfrak{u}(\nu)$, then the $i\lambda_j$ are real. If each σ_j in $B_k(\sigma_1,\dots,\sigma_k)$ is replaced by the Chern form $\sigma_k\left(\frac{i}{2\pi}\Omega^\omega\right)$ where Ω^ω is the curvature form of a connection 1-form ω on the unitary frame bundle $U(E)$ for a Hermitian vector bundle $E \to M$, then we obtain a closed $2k$-form

$$B_k(E,\omega) := B_k\left(\sigma_1\left(\tfrac{i}{2\pi}\Omega^\omega\right),\dots,\sigma_k\left(\tfrac{i}{2\pi}\Omega^\omega\right)\right) \in \Omega^{2k}(M,\mathbb{C})$$

which determines a cohomology class, say $B_k(E) \in H^{2k}(M,\mathbb{C})$. Hence, for each formal power series $b(x) = \sum_{n=0} b_n x^n$, there is an associated **total form** and **total class**, namely

$$\mathbf{B}(E,\omega) := B_0(E,\omega) + B_1(E,\omega) + \cdots \in \Omega^*(M,\mathbb{C}), \text{ and}$$
$$\mathbf{B}(E) := B_0(E) + B_1(E) + \cdots \in H^*(M,\mathbb{C}).$$

Since there seems to be no official notation to denote the assignments

$$(b(x),E,\omega) \mapsto \mathbf{B}(E,\omega) \in \Omega^*(M,\mathbb{C}) \text{ and } (b(x),E) \mapsto \mathbf{B}(E) \in H^*(M,\mathbb{C})$$

of a characteristic form (or class) to a formal power series and bundle with connection, let

$$\mathbf{MF}(b(x),E,\omega) := \mathbf{B}(E,\omega) \text{ and } \mathbf{MC}(b(x),E) := \mathbf{B}(E),$$

where **MF** stands for "muliplicative form" and **MC** stands for "muliplicative class". We use $\mathrm{MF}_k(b(x),E,\omega)$ and $\mathrm{MC}_k(b(x),E)$ for the homogeneous parts:

$$\mathbf{MF}(b(x),E,\omega) = \sum\nolimits_k \mathrm{MF}_k(b(x),E,\omega) \text{ and}$$
$$\mathbf{MC}(b(x),E) = \sum\nolimits_k \mathrm{MC}_k(b(x),E).$$

For two formal power series $b_1(x)$ and $b_2(x)$, we have

(16.103)
$$\mathbf{MC}(b_1(x)\,b_2(x),E) = \mathbf{MC}(b_1(x),E)\,\mathbf{MC}(b_2(x),E),$$

using $b_1(x_1)\,b_2(x_1)\cdots b_1(x_\nu)\,b_2(x_\nu) = b_1(x_1)\dots b_1(x_\nu)\,b_2(x_1)\dots b_2(x_\nu)$.

Given another Hermitian vector bundle $E' \to M$ of dimension ν', we may form the direct sum $E \oplus E' \to M$. Connections ω and ω' on $U(E)$ and $U(E')$ yield a connection $\omega \oplus \omega'$ on $U(E \oplus E')$, with curvature form $\Omega^{\omega \oplus \omega'} = \Omega^{\omega} \oplus \Omega^{\omega'}$. Note that for $\left(x''_1 \ldots, x''_\nu, x''_{\nu+1}, \ldots, x''_{\nu+\nu'}\right) = (x_1 \ldots, x_\nu, x'_1, \ldots, x'_{\nu'})$, we have

$$\begin{aligned}\sum_{k''=0}^{\infty} B_{k''}\left(\sigma''_1, \ldots, \sigma''_{k''}\right) &= b\left(x''_1\right) \ldots b\left(x''_\nu\right) b\left(x''_{\nu+1}\right) \ldots b\left(x''_{\nu+\nu'}\right) \\ &= b\left(x_1\right) \ldots b\left(x_\nu\right) b\left(x'_1\right) \ldots b\left(x'_{\nu'}\right) \\ &= \sum_{k=0}^{\infty} B_k\left(\sigma_1, \ldots, \sigma_k\right) \sum_{k'=0}^{\infty} B_{k'}\left(\sigma'_1, \ldots, \sigma'_k\right).\end{aligned}$$

Consequently, in the ring $H^*(M, \mathbb{C})$ we have

$$\mathbf{MC}\left(b(x), E \oplus E'\right) = \mathbf{MC}\left(b(x), E\right) \mathbf{MC}\left(b(x), E'\right). \tag{16.104}$$

Since the conjugate bundle $\overline{E}$ is the associated bundle $U(E) \times_{U(\nu)} \mathbb{C}^\nu$ relative to the conjugate representation, the curvature form for the conjugate bundle $\overline{E}$ is $-\frac{i}{2\pi}\overline{\Omega}^{\omega}$. If $A \in \mathfrak{u}(\nu)$, and $\{x_1, \ldots, x_\nu\} = \{i\lambda_1, \ldots, i\lambda_\nu\}$ where $\{\lambda_1, \ldots, \lambda_\nu\}$ is the set of eigenvalues of A, then the eigenvalues of $\overline{A}$ are $\{-\lambda_1, \ldots, -\lambda_\nu\}$ and $\{-i\lambda_1, \ldots, -i\lambda_\nu\} = \{-x_1, \ldots, -x_\nu\}$. Thus,

$$\mathbf{MC}\left(b(x), \overline{E}\right) = \mathbf{MC}\left(b(-x), E\right). \tag{16.105}$$

Note that

$$\text{the total Chern class of } E = \mathbf{c}(E) = \mathbf{MC}(1 + x, E).$$

A class which arises in the Hirzebruch-Riemann-Roch Theorem is

$$\text{the total Todd class of } E = \mathbf{Td}(E) := \mathbf{MC}\left(\frac{x}{1 - e^{-x}}, E\right) \in H^*(M, \mathbb{Q}),$$

where $\mathbb{Q}$ denotes the rationals. For a compact, complex manifold M, $\mathbf{Td}(TM)[M]$ is the *Todd genus* of M, which is a kind of holomorphic Euler characteristic of M. One can compute $\mathbf{Td}(E)$ in terms of Chern classes. Indeed, using <<Algebra 'SymmetricPolynomials' in *Mathematica*, we get (where $c_k = c_k(E)$)

$$\begin{aligned}&Td_0(E) = 1,\ Td_1(E) = \tfrac{1}{2}c_1,\ Td_2(E) = \tfrac{1}{12}\left(c_2 + c_1^2\right), \\ &Td_3(E) = \tfrac{1}{24}c_1c_2,\ Td_4(E) = \tfrac{1}{720}\left(-c_4 + c_3c_1 + 3c_2^2 + 4c_2c_1^2 - c_1^4\right), \ldots.\end{aligned}$$

In treating the case where E is the complexification of a real, even-dimensional Riemannian bundle $F \to M$ (i.e., $E = \mathbb{C} \otimes F$), we proceed as follows. If $A \in \mathfrak{so}(\nu, \mathbb{R})$, where $\nu = 2\mu$ is even, then the λ_j are not only pure imaginary, but they come in conjugate pairs, so that

$$(x_1, \ldots, x_\nu) = (i\lambda_1, \ldots, i\lambda_\nu) = (y_1, -y_1, \ldots, y_\mu, -y_\mu) \text{ for } y_l \in \mathbb{R}. \tag{16.106}$$

Hence it seems more appropriate to express $\sigma_j(x_1, \ldots, x_\nu)$ in terms of $\sigma_l\left(y_1^2, \ldots, y_\mu^2\right)$. To this end, note that

$$\begin{aligned}\sum\nolimits_{j=1}^{\nu} \sigma_j(x_1, \ldots, x_\nu) &= \prod\nolimits_{k=1}^{\nu}(1 + x_k) = \prod\nolimits_{l=1}^{\mu}(1 + y_l)(1 - y_l) \\ &= \prod\nolimits_{l=1}^{\mu}\left(1 - y_l^2\right) = \sum\nolimits_{l=1}^{\mu}(-1)^l \sigma_l\left(y_1^2, \ldots, y_\mu^2\right).\end{aligned}$$

Hence if (16.106) holds, then

$$\sigma_j(x_1, \ldots, x_\nu) = \begin{cases} 0 & \text{for } j \text{ odd} \\ (-1)^l \sigma_l\left(y_1^2, \ldots, y_\mu^2\right) & \text{for } j = 2l \text{ even}\end{cases}$$

Since $p_k(M) = (-1)^k c_{2k}(T_{\mathbb{C}}M)$ (16.101), it makes sense to define the Pontryagin classes of F by

$$p_k(F) := (-1)^k c_{2k}(\mathbb{C}\otimes F) \in H^{4k}(M,\mathbb{Z})$$

This implies that for $b(x)$ even, $\mathbf{MC}(b(x), \mathbb{C}\otimes F)$ can be obtained by writing

$$\begin{aligned} b(x_1)\dots b(x_\nu) &= b(y_1)\, b(-y_1)\dots b(y_\mu)\, b(-y_\mu) = b(y_1)^2\cdots b(y_\mu)^2 \\ &= \sum_{k=0}^{\infty} \widetilde{B}_k\left(\sigma_1\left(y_1^2,\dots,y_\mu^2\right),\dots,\sigma_k\left(y_1^2,\dots,y_\mu^2\right)\right), \end{aligned}$$

and then replacing $\sigma_j\left(y_1^2,\dots,y_\mu^2\right)$ by the j-th Pontryagin class $p_j(F) \in H^{4j}(M,\mathbb{Z})$. Thus, for real, even-dimensional, Riemannian bundles F and $b(x)$ even, we use the more direct notation

$$\mathbf{MC}\left(b(y)^2, F\right) := \mathbf{MC}(b(x), \mathbb{C}\otimes F).$$

There are various special cases arising in index theorems, which we now consider. For a real, 2μ-dimensional Riemannian bundle $F \to M$, we have

$$\text{the total } \widehat{A} \text{ class of } F = \widehat{\mathbf{A}}(F) := \mathbf{MC}\left(\frac{y/2}{\sinh(y/2)}, F\right), \tag{16.107}$$

which occurs in the index formula for the Dirac operator and its twists. We have

$$\prod_{j=1}^{\nu/2} \frac{y_j/2}{\sinh(y_j/2)} = \sum_{k=0}^{\infty} \widehat{A}_k\left(\sigma_1\left(y_1^2,\dots,y_\mu^2\right),\dots,\sigma_k\left(y_1^2,\dots,y_\mu^2\right)\right),$$

where (with the aid of *Mathematica* if desired)

$$\begin{aligned} &\widehat{A}_0 = 1,\ \widehat{A}_1(\sigma_1) = \frac{-1}{24}\sigma_1,\ \widehat{A}_2(\sigma_1,\sigma_2) = \frac{-4\sigma_2 + 7\sigma_1^2}{5760} \\ &\widehat{A}_3(\sigma_1,\sigma_2,\sigma_3) = \frac{-16\sigma_3 + 44\sigma_2\sigma_1 - 31\sigma_1^3}{967680} \\ &\widehat{A}_4(\sigma_1,\dots,\sigma_3) = \frac{-192\sigma_4 + 512\sigma_3\sigma_1 + 208\sigma_2^2 - 904\sigma_2\sigma_1^2 + 381\sigma_1^4}{464486400},\dots. \end{aligned} \tag{16.108}$$

To obtain $\widehat{A}_k(F)$, replace each σ_j by $p_j(F)$. In connection with the Hirzebruch-Signature Theorem, we have

$$\text{the total Hirzebruch } L \text{ class of } F = \mathbf{L}(F) := \mathbf{MC}\left(\frac{y}{\tanh y}, F\right), \text{ and}$$

$$\begin{aligned} &L_0 = 1,\ L_1(\sigma_1) = \frac{1}{3}\sigma_1,\ L_2(\sigma_1,\sigma_2) = \frac{7\sigma_2 - \sigma_1^2}{45} \\ &L_3(\sigma_1,\sigma_2,\sigma_3) = \frac{62\sigma_3 - 13\sigma_2\sigma_1 - 2\sigma_1^3}{945} \\ &L_4(\sigma_1,\dots,\sigma_3) = \frac{381\sigma_4 - 71\sigma_3\sigma_1 - 19\sigma_2^2 + 22\sigma_2\sigma_1^2 - 3\sigma_1^4}{14175},\dots. \end{aligned} \tag{16.109}$$

Suppose that $F \to M$ is the realification of complex bundle $F_{\mathbb{C}} \to M$ (i.e., we just restrict scalar multiplication for $F_{\mathbb{C}}$ to real scalars). If $J : F \to F$ is the map given by scalar multiplication by $\sqrt{-1}$, then there is $\mathbb{C}$-linear extension of J, say $J_{\mathbb{C}} : \mathbb{C}\otimes F \to \mathbb{C}\otimes F$. Since $J_{\mathbb{C}}^2 = -\operatorname{Id}$, we have $\mathbb{C}\otimes F = F^{1,0}\oplus F^{0,1}$, where $F^{1,0} := \{V - iJV : V \in F\}$ is the $+i$ eigenbundle of the $J_{\mathbb{C}}$ and $F^{0,1} := \{V + iJV : V \in F\}$

is the $-i$ eigenbundle of $J_{\mathbb{C}}$. Note that $F_{\mathbb{C}} \cong F^{1,0}$ via $V \mapsto V - iJV$, and $\overline{F}_{\mathbb{C}} \cong F^{0,1}$ via $V \mapsto V + iJV$. Since

$$\left(\frac{x/2}{\sinh(x/2)}\right)^2 = \frac{x^2}{\left(e^{x/2} - e^{-x/2}\right)^2} = \frac{x^2}{e^{x/2}\left(e^{x/2} - e^{-x/2}\right) e^{-x/2}\left(e^{x/2} - e^{-x/2}\right)}$$
$$= \frac{x^2}{\left(e^x - 1\right)\left(1 - e^{-x}\right)} = \frac{x}{1 - e^{-x}} \frac{-x}{1 - e^x},$$

it follows from (16.103), (16.105) and (16.104) that

$$\widehat{\mathbf{A}}(F)^2 = \mathbf{Td}(F_{\mathbb{C}})\,\mathbf{Td}\left(\overline{F_{\mathbb{C}}}\right) = \mathbf{Td}\left(F_{\mathbb{C}} \oplus \overline{F_{\mathbb{C}}}\right) = \mathbf{Td}(\mathbb{C} \otimes F).$$

Some characteristic classes are not expressible in terms of $\mathbf{MC}(b(x), E)$ for some formal power series $b(x)$, but they can be described in a similar way. The Chern character $\mathbf{ch}(E)$ is not $\mathbf{MC}(e^x, E)$. However, *adding* (instead of multiplying as in (16.102)) one obtains

$$e^{x_1} + \cdots + e^{x_\nu} = \sum_{k=0}^{\infty} ch_k(\sigma_1, \ldots, \sigma_k)$$

If each σ_j in $ch_k(\sigma_1, \ldots, \sigma_k)$ is replaced by the Chern class $c_j(E)$, then we obtain

$$\mathbf{ch}(E) = ch_0(E) + ch_1(E) + \cdots \in H^*(M, \mathbb{R}). \tag{16.110}$$

We compute

$$\begin{aligned}
&ch_0(E) = \dim E,\ ch_1(E) = c_1,\ ch_2(E) = \tfrac{1}{2}c_1^2 - c_2,\\
&ch_3(E) = \tfrac{1}{6}\left(3c_3 - 3c_2c_1 + c_1^3\right),\\
&ch_4(E) = \tfrac{1}{24}\left(-4c_4 + 4c_3c_1 + 2c_2^2 - 4c_2c_1^2 + c_1^4\right), \ldots .
\end{aligned}$$

While we do *not* have $\mathbf{ch}(E \oplus E') = \mathbf{ch}(E)\,\mathbf{ch}(E')$ as in (16.104),

$$\mathbf{ch}(E \oplus E') = \mathbf{ch}(E) + \mathbf{ch}(E') \text{ and } \mathbf{ch}(E \otimes E') = \mathbf{ch}(E)\,\mathbf{ch}(E').$$

In terms of curvature forms, the first relation is clear from $\Omega^{\omega \oplus \omega'} = \Omega^{\omega} \oplus \Omega^{\omega'}$, while the second follows from the fact that the curvature form for $E \otimes E'$ is $(\Omega^{\omega} \otimes \mathrm{Id}) \oplus \left(\mathrm{Id} \otimes \Omega^{\omega'}\right)$ together with

$$\sum\nolimits_{k=1}^{\nu} \sum\nolimits_{k'=1}^{\nu'} e^{x_k + x'_{k'}} = \sum\nolimits_{k=1}^{\nu} e^{x_k} \sum\nolimits_{k'=1}^{\nu'} e^{x'_{k'}}.$$

More generally, one could consider elementary polynomals $\sigma_k(b(x_1), \ldots, b(x_\nu))$ for k other than 1 or ν, although finding uses for such might be a challenge.

The Euler class of a real, oriented Riemannian bundle F of dimension 2μ is not generally expressible in terms of Pontriagin classes of $\mathbb{C} \otimes F$. We proceed as follows. If $A \in \mathfrak{so}(2\mu, \mathbb{R})$, then A is $\mathrm{SO}(2\mu)$-similar to a matrix of the form

$$\bigoplus_{k=1}^{\mu} \begin{bmatrix} 0 & -y_k \\ y_k & 0 \end{bmatrix}.$$

The eigenvalues λ_j of A come in pure imaginary conjugate pairs $\pm iy_k$, so that

$$\begin{aligned}
&(x_1, \ldots, x_\nu) := (i\lambda_1, \ldots, i\lambda_\nu) = (y_1, -y_1, \ldots, y_\mu, -y_\mu) \text{ for } y_l \in \mathbb{R}, \text{ and} \\
&\mathrm{Pf}\,(A) := \frac{(-1)^\mu}{2^\mu \mu!} \sum_{(i)} \varepsilon_{i_1 \cdots i_{2\mu}} A_{i_1 i_2} \cdots A_{i_{2\mu-1} i_{2\mu}} \\
&= (-1)^\mu A_{12} \cdots A_{2\mu-1,2\mu} = (-1)^\mu (-y_1) \cdots (-y_\mu) = y_1 \cdots y_\mu.
\end{aligned}$$

Note that $y_1 \cdots y_\mu$ is not a symmetric polynomial in $y_1^2, \cdots, y_\mu^2$. However, if ω is a connection on the bundle $\pi : \mathrm{SO}\,(F) \to M$ of *oriented* frames of F and Ω^ω is the curvature, then

$$\pi^* \chi\,(F, \omega) = (-1)^\mu\, \mathrm{Pf}\left(\tfrac{1}{2\pi} \Omega^\omega\right) \tag{16.111}$$

for a unique closed, 2μ-form $\chi\,(F, \omega) \in \Omega^{2\mu}(M, \mathbb{R})$, the *Euler form* of E *relative to* ω, which by definition represents the *Euler class*

$$\chi\,(F) := [\chi\,(F, \omega)] \in H^{2\mu}(M, \mathbb{R})\,. \tag{16.112}$$

In the case $F = TM$ and $\omega = \theta$, note that

$$\begin{aligned}
&\pi^*\, \mathrm{GB}\left(\Omega^\theta\right) = (-1)^m\, \mathrm{Pf}\left(\tfrac{1}{2\pi}\Omega^\theta\right) = \pi^* \chi\,(TM, \theta) \\
&\Rightarrow \mathrm{GB}\left(\Omega^\theta\right) = \chi\,(TM, \theta)\,.
\end{aligned} \tag{16.113}$$

Remark 16.60. We prefer to write $\chi\,(TM)$ as $\mathrm{GB}\,(TM)$. There is a good reason for this. The Gauss-Bonnet-Chern Theorem states that the Euler characteristic $\chi\,(M)$ of M, defined as the alternating sum of numbers of faces (or Betti numbers) of M, is given by

$$\chi\,(M) = \int_M \mathrm{GB}\left(\Omega^\theta\right) = \mathrm{GB}\,(TM)\,[M]\,.$$

On the other hand, for a generic characteristic class, say $C\,(TM) \in H^*(M)$, frequently one *defines* $C\,(M)$ to be $C\,(TM)\,[M]$. The Gauss-Bonnet-Chern Theorem is not simply a definition, and yet that is exactly what it looks like if one writes it as $\chi\,(M) = \chi\,(TM)\,[M]$. Thus, we prefer to use $\mathrm{GB}\,(TM)$ in place of $\chi\,(TM)$, although admittedly changing established notation is a losing battle, no matter how noble the cause.

Observe that $A \in \mathfrak{so}\,(2\mu)$ is the realification of $B = diag\,(iy_1, \ldots, iy_\mu) \in \mathfrak{su}\,(\mu)$, and

$$\det\,(iB) = (-1)^\mu\, y_1 \cdots y_\mu = (-1)^\mu\, \mathrm{Pf}\,(A)\,. \tag{16.114}$$

For a manifold M, $J \in \mathrm{End}\,(TM)$ is an **almost-complex structure** if $J^2 = -I$. If J exists, then TM becomes a complex vector bundle by defining $(a + ib)\,X = aX + bJX$ and M is an **almost-complex manifold**. The formula (16.114) will now be used to show that for a compact, almost-complex, manifold M with $\dim_{\mathbb{R}} M = 2m$, we have $c_m\,(TM) = \mathrm{GB}\,(TM) = \chi\,(TM)$. If h_0 is any Riemannian metric on M, then $h\,(X, Y) := h_0\,(X, Y) + h_0\,(JX, JY)$ is compatible with the complex structure, say J, on TM (i.e., $h\,(JX, JY) = h\,(X, Y)$). We then have a Hermitian metric $\langle X, Y\rangle := h\,(X, Y) + ih\,(X, JY)$ on TM regarded as a complex vector space. Note that

$$\begin{aligned}\langle JX,Y\rangle &= h(JX,Y)+ih(JX,JY) = h(JX,Y)+ih(X,Y)\\&= -h(X,JY)+ih(X,Y) = i(h(X,Y)+ih(X,JY)) = i\langle X,Y\rangle\\\langle X,JY\rangle &= h(X,JY)+ih\left(X,J^2Y\right) = h(X,JY)-ih(X,Y)\\&= -i(h(X,Y)+ih(X,JY)) = -i\langle X,Y\rangle.\end{aligned}$$

The unitary frame bundle $UM := U(TM)$ (relative to $\langle\cdot,\cdot\rangle$) is then a subbundle of FM. We may regard FM as the principal bundle $UM\times_{\mathrm{U}(m)}\mathrm{O}(2m)$ associated to UM via the inclusion $\mathrm{U}(m)\to\mathrm{O}(2m)$; see 16.19, p.378. If the Levi-Civita connection for h on FM restricts to a connection, say ω, on UM (i.e., the horizontal subspaces at points in UM are contained in $T(UM)$), then M is Kähler (by one definition). However, if this is not the case, then we can still uniquely extend *any* connection ω on UM to a connection, say θ, on FM (see 16.21, 379). If we regard ω as $\mathfrak{u}(m)$-valued, then $\Omega^\theta|_{UM}$ is just the realification of the curvature Ω^ω. Now,

$$\begin{aligned}\mathrm{GB}\left(\Omega^\theta\right) &= \frac{1}{2^{2m}\pi^m m!}\sum\nolimits_{(i)}\varepsilon_{i_1\cdots i_{2m}}\Omega^\theta_{i_1i_2}\wedge\cdots\wedge\Omega^\theta_{i_{2m-1}i_{2m}}\\&= \frac{1}{2^m m!}\left(\frac{1}{2\pi}\right)^m\sum\nolimits_{(i)}\varepsilon_{i_1\cdots i_{2m}}\Omega^\theta_{i_1i_2}\wedge\cdots\wedge\Omega^\theta_{i_{2m-1}i_{2m}}\\&= (-1)^m\,\mathrm{Pf}\left(\tfrac{1}{2\pi}\Omega^\theta\right),\ \text{and}\end{aligned}$$

$$\begin{aligned}\sigma_m\left(\frac{i}{2\pi}\Omega^\omega\right) &= \frac{1}{m!}\left(\frac{i}{2\pi}\right)^m\sum_{(i),(j)}\delta^{j_1\cdots j_m}_{i_1\cdots i_m}\left(\Omega^\omega\right)^{i_1}{}_{j_1}\wedge\cdots\wedge\left(\Omega^\omega\right)^{i_m}{}_{j_m}\\&= \det\left(\tfrac{i}{2\pi}\Omega^\omega\right).\end{aligned}$$

Hence, in view of (16.114) and restricting to UM, we have

$$\mathrm{GB}\left(\Omega^\theta\right) = (-1)^M\,\mathrm{Pf}\left(\tfrac{1}{2\pi}\Omega^\theta\right) = \det\left(\tfrac{i}{2\pi}\Omega^\omega\right) = \sigma_m\left(\tfrac{i}{2\pi}\Omega^\omega\right),\ \text{and}$$

$$c_m(TM) = \mathrm{GB}(TM) = \chi(TM). \tag{16.115}$$

It often happens that a Hermitian bundle E arises an associated bundle $E = P\times_G W$, relative to some unitary representation $r: G\to U(W)$, for some principal G-bundle $P\to M$ which is not necessarily the unitary frame bundle $U(E)$. Since $r: G\to U(W)$ is a homomorphism of the Lie group for P to $U(W)$, there is an associated principal $U(W)$-bundle (see Proposition 16.19, p. 378)

$$P' = P\times_G U(W) := \frac{P\times U(W)}{G} = \left\{[p,g'] : p\in P, g'\in U(W)\right\},$$

and an r-equivariant map $\Gamma: P\to P'$. There is also a map $F: P'\to U(E)$ given by

$$F([p,g'])(w) = [p,g'(w)],\ \text{for } p\in P,\ g'\in U(W),\ w\in W.$$

Note that F is well-defined and equivariant, since

$$\begin{aligned}F\left(\left[pg, r\left(g^{-1}\right)g'\right]\right)(w) &= \left[pg, r\left(g^{-1}\right)g'(w)\right] = [p,g'(w)],\ \text{and for } h'\in U(W),\\F([p,g']h')(w) &= F([p,g'\circ h'])(w) = [p,g'(h'(w))]\\&= F([p,g'])(h'(w)) = (F([p,g'])\circ h')(w).\end{aligned}$$

As F is also bijective, it is an isomorphism of principal bundles. Thus we have a morphism

$$\mathcal{R} := F \circ \Gamma : P \to U(E)$$

and $\mathcal{R}(P)$ may be regarded as a principal $r(G)$-subbundle of the $U(W)$-bundle $U(E)$. Using Proposition 16.21 (p.16.21), if ω is a connection on P, then there is a unique connection ω' on P' such that $\omega = \Gamma^*\omega'$. For the connection $\omega_E := F^{-1*}\omega'$ on $U(E)$, we then have

$$r' \circ \omega = \mathcal{R}^{-1*}\omega_E \text{ and } r' \circ \Omega^{\omega} = \mathcal{R}^{-1*}\Omega^{\omega_E}. \tag{16.116}$$

In the notation of (16.102), let

$$\begin{aligned} B_k(P,\omega,r') &:= B_k\left(\sigma_1\left(\tfrac{i}{2\pi}r'\circ\Omega^{\omega}\right),\dots,\sigma_k\left(\tfrac{i}{2\pi}r'\circ\Omega^{\omega}\right)\right), \text{ and} \\ \mathbf{MF}(b(x),P,\omega,r') &:= B_0(\omega,r') + B_1(\omega_0,r') + \cdots \in \Omega^*(M,\mathbb{C}). \end{aligned}$$

From $r' \circ \Omega^{\omega} = \mathcal{R}^{-1*}\Omega^{\omega_E}$, we get $B_k(P,\omega,r') = B_k(E,\omega_E)$. Thus,

$$\begin{aligned} &\mathbf{MF}(b(x),E,\omega_E) = \mathbf{MF}(b(x),P,\omega,r') \text{ and} \\ &\mathbf{MC}(b(x),E) = \mathbf{MC}(b(x),P,r') := \text{class of } \mathbf{MF}(b(x),P,\omega,r'). \end{aligned}$$

With similar notation, one also has

$$\begin{aligned} \mathbf{c}(E,\omega_E) = \mathbf{c}(P,\omega,r') \text{ and } \mathbf{c}(E) = \mathbf{c}(P,r'), \text{ and similarly} \\ \mathbf{ch}(E,\omega_E) = \mathbf{ch}(P,\omega,r') \text{ and } \mathbf{ch}(E) = \mathbf{ch}(P,r'). \end{aligned} \tag{16.117}$$

Consider the special case $G = \mathrm{U}(\nu) = \mathrm{U}(\mathbb{C}^{\nu})$ and $r : \mathrm{U}(\nu) \to \mathrm{U}(W)$. If $E_0 := P \times_{\mathrm{U}(\nu)} \mathbb{C}^{\nu}$, then we can often find $\mathbf{MC}(b(x),E) = \mathbf{MC}(b(x),\omega,r')$ or other characteristic of E classes (such as $\mathbf{ch}(E)$) in terms of the Chern classes $c_k(E_0)$. If possible, one just expresses the elementary symmetric polynomials in the eigenvalues of $r'(A)$ in terms of those for $A \in \mathfrak{u}(\nu)$. For example, consider the exterior product bundle $E = \Lambda^p(E_0)$, where $r : \mathrm{U}(\nu) \to \mathrm{U}(\Lambda^p(\mathbb{C}^{\nu}))$. If $A \in \mathfrak{u}(\nu)$ has eigenvectors $e_1,\dots,e_{\nu}$ with eigenvalues $\lambda_1,\dots,\lambda_{\nu}$, then an eigenbasis of $r'(A)$ is

$$\left\{e_{i_1}\wedge\cdots\wedge e_{i_p} : 1 \le i_1 < \cdots < i_p \le n\right\},$$

and the associated eigenvalues of $r'(A)$ are $\lambda_{i_1}+\cdots+\lambda_{i_p}$. For $\{x_1,\dots,x_{\nu}\} = \{i\lambda_1,\dots,i\lambda_{\nu}\}$, the j-th elementary symmetric polynomial in the $x_{i_1}+\cdots+x_{i_p}$ $(1 \le i_1 < \cdots < i_p \le n)$ is the coefficient of t^j in the expansion of

$$\prod_{(i)^p}\left(1+\left(x_{i_1}+\cdots+x_{i_p}\right)t\right),$$

where the multi-index $(i)^p$ ranges over $\{1 \le i_1 < \cdots < i_p \le n\}$. These coefficients can in turn be expressed as polynomials in the $\sigma_k(x_1,\dots,x_{\nu})$. The Chern classes of $\Lambda^p(E_0)$ are then the same polynomials in the $c_k(E_0)$. Since

$$\prod\nolimits_{k=1}^{\nu}\left(1+e^{x_k}\right) = \sum\nolimits_{p=1}^{\nu}\sum\nolimits_{(i)^p}\exp\left(x_{i_1}+\cdots+x_{i_p}\right),$$

the Chern character $ch_j(\Lambda^*(E_0)) \in H^{2j}(M,\mathbb{Q})$ can be found by expanding the product on the left and writing the symmetric, homogeneous j-th degree part of the power series as a polynomial in the $\sigma_k(x_1,\dots,x_{\nu})$, regarded as $c_k(E_0)$.

When $E_0 = \mathbb{C}\otimes F_0$ for some real, Riemannian bundle F_0 of dimension 2μ, then $\mathbf{ch}\left(\Lambda^*\left(\mathbb{C}\otimes F_0\right)\right)$ may be obtained by expanding

$$\begin{aligned}\prod_{k=1}^{\mu}\left(1+e^{y_k}\right)\left(1+e^{-y_k}\right) &= \prod_{k=1}^{\mu} e^{y_k/2}\left(e^{-y_k/2}+e^{y_k/2}\right)e^{-y_k/2}\left(e^{y_k/2}+e^{-y_k/2}\right)\\ &= \prod_{k=1}^{\mu} 4\cosh^2\left(y_k/2\right) \qquad (16.118)\end{aligned}$$

in terms of $\sigma_k\left(y_1^2,\ldots,y_\mu^2\right)$ which are then replaced by the Pontryagin classes $p_k\left(F_0\right)$. In other words,

$$\mathbf{ch}\left(\Lambda^*\left(\mathbb{C}\otimes F_0\right)\right) = \mathbf{MC}\left(4\cosh^2\left(y/2\right), F_0\right). \qquad (16.119)$$

8. Holonomy

In this section, we assume that M is connected. Let ω be a connection 1-form on the principal G-bundle $\pi : P \to M$. Fix a point $p_0 \in P$, and let P_0 be the set of all points $p \in P$ which can be joined to p_0 by a smooth horizontal curve $\gamma : [a,b] \to P$, say $\gamma(a) = p_0$, $\gamma(b) = p$ and $\omega(\gamma'(t)) = 0$ for $t \in (a,b)$. The *holonomy group* of ω with reference point p_0 is $G_0 := \{g \in G : p_0 g \in P_0\}$. It can be proven that (see [**KN63**, 83-85]) that P_0 is an immersed submanifold of P, and $\pi|_{P_0} : P_0 \to M$ is a principal G_0-bundle, which is known as the *holonomy bundle* of ω through p_0. If G_0 is a proper subgroup of G, then ω is said to be *reducible* to G_0. If $G_0 = G$, then ω is *irreducible.*

The *isotropy subgroup* at ω for the action of the group $\mathrm{GA}(P)$ of gauge transformations on the space $\mathcal{C}(P)$ of connection 1-forms on P will be denoted by $I_\omega := \{F \in \mathrm{GA}(P) : F\cdot\omega = \omega\}$. Under the isomorphism $\Phi : C(P,G) \to \mathrm{GA}(P)$ of Proposition 16.27 (p. 382), we can identify I_ω with the subgroup $\Phi^{-1}(I_\omega)$ of $C(P,G)$.

Proposition 16.61. *The homomorphism $I_\omega \to G$ given by $\Phi(f) \mapsto f(p_0)$ maps I_ω isomorphically onto the centralizer of G_0 in G, namely*

$$Z(G_0) := \{g \in G : gg_0 = g_0 g \text{ for all } g_0 \in G_0\}.$$

Proof. Let $\Phi(f) \in I_\omega$ and $g_0 \in G_0$. To prove that $f(p_0) \in Z(G_0)$, we need to show that $f(p_0)\, g_0 = g_0 f(p_0)$. Let γ be a horizontal curve joining p_0 to $p_0 g_0$. Then $\gamma\cdot f(p_0)$ is a horizontal curve joining $p_0 f(p_0)$ to $p_0 g_0 f(p_0)$. Since $\Phi(f)^*\omega = \omega$, $\Phi(f)\circ\gamma$ is a horizontal curve joining $p_0 f(p_0)$ to $\Phi(f)(p_0 g_0) = \Phi(f)(p_0)\, g_0 = p_0 f(p_0)\, g_0$. Since

$$\pi\left(\gamma(t)\cdot f(p_0)\right) = \pi\left(\Phi(f)(\gamma(t))\right),$$

the curves $\gamma\cdot f(p_0)$ and $\Phi(f)\circ\gamma$ are horizontal lifts of the same curve in M, and they have the same initial point these curves have $p_0 f(p_0)$. By the uniqueness of horizontal lifts with the same starting point (see [**KN63**, 69]), the endpoints $p_0 g_0 f(p_0)$ and $p_0 f(p_0)\, g_0$ must agree, whence $g_0 f(p_0) = f(p_0)\, g_0$ (i.e., $f(p_0) \in Z(G_0)$). To see that $I_\omega \to G$ is injective, we use (16.27) on p. 382, namely

$$\left(\Phi(f)\cdot\omega\right)(X) = f(p)\left(f^{-1}\right)_{*p}(X) + f(p)\,\omega(X)\, f(p)^{-1}, \qquad (16.120)$$

for $X \in T_pP$. If $\Phi(f) \in I_\omega$, then $\Phi(f)\cdot\omega = \omega$ and (16.120) implies that if X is horizontal, then $\left(f^{-1}\right)_{*p}(X) = 0$. Then f^{-1} (and hence f) is constant on all horizontal curves and in particular f is the constant $f(p_0)$ on P_0. Since P_0 meets each fiber of P and f is equivariant, f is uniquely determined by $f(p_0)$. Thus, $I_\omega \to G$ is injective. To prove that $I_\omega \to Z(G_0)$ is onto, let $g' \in Z(G_0)$, and let $f(q_0) = g'$ for all $q_0 \in P_0$. For arbitrary $p \in P$, there is some $g \in G$ such that $pg \in P_0$, and we define

$$f(p) := gf(pg)\,g^{-1} = g^{-1}g'g.$$

To show that f is well-defined, suppose that $ph \in P_0$. Then $gh^{-1} \in G_0$ and

$$g' \in Z(G_0) \Rightarrow g'\left(gh^{-1}\right) = \left(gh^{-1}\right)g' \Rightarrow g^{-1}g'g = h^{-1}g'h.$$

Note that by the definition of P_0, the horizontal subspace H_{q_0} of ω at any $q_0 \in P_0$ is contained in $T_{q_0}P_0$. Since $f|_{P_0}$ is constant, for each $X \in H_{q_0}$, we have

$$(\Phi(f)\cdot\omega)(X) = f(q_0)\,\omega(X)\,f(q_0)^{-1} = 0.$$

Thus, the horizontal subspace of $\Phi(f)\cdot\omega$ at any $q_0 \in P_0$ coincides with H_{q_0}. Since P_0 meets each fiber of P and horizontal subspaces are R_{g*}-invariant, the horizontal subspaces of $\Phi(f)\cdot\omega$ and ω coincide at all points of P. Thus, $\Phi(f) \in I_\omega$. □

PROPOSITION 16.62. *The curvature $\Omega^\omega \in \overline{\Omega}^2(P,\mathfrak{g})$ at any point of the holonomy bundle P_0 of ω (with reference point p_0) has values in the Lie algebra $\mathfrak{g}_0$ of the holonomy group G_0.*

PROOF. Note that $\omega|_{P_0}$ clearly has values in $\mathfrak{g}_0$. Thus, $(d\omega)|_{P_0} = d(\omega|_{P_0})$ has values in $\mathfrak{g}_0$. Since the horizontal subspace of ω at any $q_0 \in P_0$ is contained in $T_{q_0}P_0$, we have $\Omega^\omega(X,Y) = d\omega\left(X^H, Y^H\right) \in \mathfrak{g}_0$ for all $X, Y \in T_{q_0}P$. □

PROPOSITION 16.63. *Let $\pi : P \to M$ be a principal $\mathrm{U}(1)$-bundle with a connection 1-form ω, and let $\rho : V \to P$ be a local trivialization of P. Let $D = \left\{re^{it} : r \in [0,1],\, t \in \mathbb{R}\right\}$ be the closed unit disk in $\mathbb{C}$, and let $f : D \to V$ be the restriction of a smooth immersion of a larger open disk. Then there is a unique function $h : [0,2\pi] \to \mathbb{R}$, such that $\widetilde{\gamma}(t) := \rho(\gamma(t))\,e^{ih(t)}$ defines an ω-horizontal lift of $\gamma(t)$ and $h(0) = 0$. If $\Omega^\omega = d\omega \in \overline{\Omega}^2(P, i\mathbb{R})$ is the curvature 2-form of ω, then the element $e^{ih(2\pi)}$ of the holonomy group of ω at $p_0 = \rho(f(1))$ determined by $\widetilde{\gamma}$ is given by*

$$e^{ih(2\pi)} = \exp\left(-\int_D (\rho\circ f)^*\,\Omega^\omega\right). \tag{16.121}$$

PROOF. For $\gamma : [0,2\pi] \to V$ given by $\gamma(t) = f\left(e^{it}\right)$, a curve $\widetilde{\gamma} : [0,2\pi] \to P$ with $\pi\circ\widetilde{\gamma} = \gamma$ and $\widetilde{\gamma}(0) = \rho(\gamma(0))$ has the the form $\widetilde{\gamma}(t) = \rho(\gamma(t))\,e^{ih(t)}$ for some $h : [0,2\pi] \to \mathbb{R}$ with $h(0) = 0$. The curve $\widetilde{\gamma}$ is a horizontal lift of γ if and only if $\omega(\widetilde{\gamma}'(t)) = 0$ for all $t \in [0,2\pi]$. Note that

$$\begin{aligned}\widetilde{\gamma}'(t_0) &= \tfrac{d}{dt}\left(\rho(\gamma(t))\,e^{ih(t)}\right) = \tfrac{d}{dt}\left(\rho(\gamma(t))\,e^{ih(t_0)}\right) + \tfrac{d}{dt}\left(\rho(\gamma(t_0))\,e^{ih(t)}\right)\\ &= R_{e^{ih(t_0)}*}\left(\rho_*\left(\gamma'(t_0)\right)\right) + \left(ih'(t_0)\right)^*_{\rho(\gamma(t_0))}, \text{ and so}\end{aligned}$$

$$\begin{aligned}\omega\left(\widetilde{\gamma}'(t_0)\right) &= \omega\left(R_{e^{ih(t_0)}*}\left(\rho_*\left(\gamma'(t_0)\right)\right)\right) + \omega\left(ih'(t_0)^*_{\rho(\gamma(t_0))}\right)\\ &= ad_{e^{ih(t_0)}}\omega\left(\rho_*\left(\gamma'(t_0)\right)\right) + ih'(t_0)\\ &= \omega\left(\rho_*\left(\gamma'(t_0)\right)\right) + ih'(t_0).\end{aligned}$$

Thus, $\widetilde{\gamma}$ is a horizontal lift of γ (i.e., $\omega\left(\widetilde{\gamma}'(t)\right) = 0$) if and only if

$$\begin{aligned}&h'(t) = i\omega\left(\rho_*\left(\gamma'(t)\right)\right) = i\left(\rho^*\omega\right)\left(\gamma'(t)\right);\text{ i.e.,}\\ &h(t) = h(t) - h(0) = i\int_0^t \left(\rho^*\omega\right)\left(\gamma'(\tau)\right)\,d\tau = i\int_0^t \left(\rho^*\omega\right)\left(f_*\left(\tfrac{d}{d\tau}e^{i\tau}\right)\right)\,d\tau\\ &\qquad = i\int_0^t \left(\rho^*\omega\right)\left(f_*\left(ie^{i\tau}\right)\right)\,d\tau = i\int_0^t \left(\left(\rho\circ f\right)^*\omega\right)\left(ie^{i\tau}\right)\,d\tau.\end{aligned}$$

We then have (16.121), since using Stokes' Theorem,

$$\begin{aligned}\int_D \left(\rho\circ f\right)^*\Omega^\omega &= \int_D d\left(\left(\rho\circ f\right)^*\omega\right) = \int_{\partial D}\left(\rho\circ f\right)^*\omega\\ &= \int_0^{2\pi}\left(\left(\rho\circ f\right)^*\omega\right)\left(ie^{i\tau}\right)\,d\tau = -ih(2\pi).\end{aligned}$$

□

REMARK 16.64. Note that since $U(1)$ is abelian, we have that $\Omega^\omega \in \overline{\Omega}^2(P, i\mathbb{R})$ is right-invariant as well as horizontal, and so there is a unique $F^\omega \in \Omega^2(M,\mathbb{R})$ such that $\pi^*(F^\omega) = i\Omega^\omega$. Thus,

$$h(2\pi) = i\int_D \left(\rho\circ f\right)^*\Omega^\omega = \int_D f^*F^\omega,$$

note that $\int_D f^*F^\omega$ is defined even if $f(D) \subsetneq V$. Moreover, the element say g_γ of the holonomy group of ω at $p_0 = \rho(f(1))$ determined by a horizontal lift $\widetilde{\gamma}$ of γ is also defined even if $f(D) \subsetneq V$. Thus, it makes sense to ask whether

$$g_\gamma = \exp\left(i\int_D f^*F^\omega\right).$$

This is the case, and it can be proven by first establishing a version of Proposition 16.63 for domains with corners and then partitioning D into such domains each of which is mapped by F into an open set over which P is trivial. Incidentally, if we define $A_V \in \Omega^1(V,\mathbb{R})$ by $\rho^*(\omega) = iA_V$, then A_V *does* depend on ρ, and it has the interpretation of being an electromagnetic gauge potential. Then $F^\omega = \rho^*(i\Omega^\omega) = \rho^*(id\omega) = d(i\rho^*(\omega)) = -dA_V$, is the electromagnetic field, regarded as a 2-form.

Let $\pi_E : E \to M$ be a Hermitian line bundle over a compact surface M ($\dim_{\mathbb{R}} M = 2$, without boundary) and let $\pi : P \to M$ be the principal $\mathrm{U}(1)$-bundle of unitary frames with a connection 1-form ω. Suppose that $\psi \in C^\infty(E)$ with finite zero set $Z := \{z \in M : \psi(z) = 0\} = \{z_1, z_2, \ldots, z_n\}$. Let $D = \left\{re^{it} : r \in [0,1],\, t \in \mathbb{R}\right\}$ be the closed unit disk in $\mathbb{C}$ and for $k \in \{1,\ldots,m\}$, let $f_k : D \to M$ be the restriction of a smooth embedding of a larger open disk, such that P is trivial over $f_j(D)$. Assume that $f_j(D) \cap f_k(D) = \phi$ for $j \neq k$. Then for $M_1 := M - \cup_{j=1}^n f_j(D)$, we have a section $\psi_0 := \psi/|\psi|$ and a section $\rho : M_1 \to P$, where $\rho(x) : \mathbb{C} \to E_x$ is the frame given by $\rho(x)(z) := z\psi_0(x) \in E_x$. Since P is trivial over $f_j(D)$, we have a local section $\rho_j : f_j(D) \to P$. Let $g_j : \mathrm{U}(1) \to \mathrm{U}(1)$ be defined by

$$(\rho \circ f_j)\left(e^{it}\right) = \rho\left(f_j\left(e^{it}\right)\right) = (\rho_j \circ f_j)\left(e^{it}\right) g_j\left(e^{it}\right) =: ((\rho_j \circ f_j) \cdot g_j)\left(e^{it}\right).$$

The **degree** of the zero z_j of ψ, denoted $\deg(\psi; z_j)$, is defined to be the degree (or winding number) of g_j. There is a function $\widetilde{g}_j : \mathbb{R} \to \mathbb{R}$ (unique up to an additive multiple of 2π), such that $g_j\left(e^{it}\right) = e^{i\widetilde{g}_j(t)}$ and $2\pi \deg(\psi; z_j) = \widetilde{g}_j(2\pi) - \widetilde{g}_j(0)$. Since $\frac{d}{dt}\left(e^{i\widetilde{g}_j(t)}\right) = e^{i\widetilde{g}_j(t)} i\widetilde{g}_j'(t)$, we then have

$$2\pi \deg(\psi; z_j) = \int_0^{2\pi} \widetilde{g}_j'(t)\, dt = \int_0^{2\pi} -ig_j\left(e^{it}\right)^{-1} \tfrac{d}{dt} g_j\left(e^{it}\right) dt. \tag{16.122}$$

THEOREM 16.65. *As above, let $\pi_E : E \to M$ be a Hermitian line bundle over a compact, orientable 2-manifold M, let ω be a connection on the $\mathrm{U}(1)$-bundle of P of unitary frames of E, and let $\psi \in C^\infty(E)$ have finitely many zeros. For $F^\omega \in \Omega^2(M, \mathbb{R})$ determined by $\pi^*(F^\omega) = i\Omega^\omega$, we have*

$$c_1(E)[M] = \frac{1}{2\pi}\int_M F^\omega = \sum\nolimits_{j=1}^n \deg(\psi; z_j). \tag{16.123}$$

PROOF. Recall from Definition 16.58 that $c_1(E)[M] := \frac{i}{2\pi}\int_M \sigma_1(\Omega^\omega)$ which is $\frac{1}{2\pi}\int_M F^\omega$. Using the notation above and $A := \rho^*(-i\omega)$, we compute

$$\begin{aligned}
\int_{M_1} F^\omega &= \int_{M_1} -dA = -\int_{\partial M_1} A = \int_{\partial M_1} \rho^*(i\omega) \\
&= -\sum\nolimits_{j=1}^n \int_{\partial f_j(D)} \rho^*(i\omega) = -\sum\nolimits_{j=1}^n \int_{\partial D} (\rho \circ f_j)^*(i\omega) \\
&= -\sum\nolimits_{j=1}^n \int_{\partial D} ((\rho_j \circ f_j) \cdot g_j)^*(i\omega) \\
&= -\sum\nolimits_{j=1}^n \int_0^{2\pi} i\omega\left(((\rho_j \circ f_j) \cdot g_j)_*\left(ie^{it}\right)\right) dt. \qquad (16.124)
\end{aligned}$$

We have

$$\begin{aligned}
&((\rho_j \circ f_j) \cdot g_j)_*\left(ie^{it_0}\right) = \tfrac{d}{dt}\left((\rho_j \circ f_j)\left(e^{it}\right) g_j\left(e^{it}\right)\right)\Big|_{t=t_0} \\
&= \tfrac{d}{dt}\left((\rho_j \circ f_j)\left(e^{it}\right)\right)\Big|_{t=t_0} g_j\left(e^{it_0}\right) \\
&+ \tfrac{d}{dt}\left((\rho_j \circ f_j)\left(e^{it_0}\right) g_j\left(e^{it_0}\right) g_j\left(e^{i(t-t_0)}\right)\right)\Big|_{t=t_0} \\
&= R_{g_j(e^{it_0})*}\left((\rho_j \circ f_j)_*\left(ie^{it_0}\right)\right) + \left(g_j\left(ie^{it_0}\right)^{-1} g_{j*}\left(ie^{it_0}\right)\right)^*_{((\rho_j \circ f_j) \cdot g_j)(e^{it_0})},
\end{aligned}$$

Since ω is R_g-invariant and $\omega(B^*) = B$ for $B \in \mathfrak{u}(1) = i\mathbb{R}$, we then have

$$i\omega\left(((\rho_j \circ f_j) \cdot g_j)_*\left(ie^{it}\right)\right) = i\omega\left((\rho_j \circ f_j)_*\left(ie^{it_0}\right)\right) + ig_j\left(ie^{it_0}\right)^{-1} g_{j*}\left(ie^{it_0}\right).$$

Thus, using (16.122) and (16.124),

$$\int_{M_1} F^\omega = -\sum\nolimits_{j=1}^n \int_{\partial D} (\rho_j \circ f_j)^* i\omega + 2\pi \sum\nolimits_{j=1}^n \operatorname{ord}(\psi; z_j). \tag{16.125}$$

For $r \in (0,1]$, let $D_r = rD$ and let $f_{j,r} : D \to M$ be given by $f_{j,r}(z) := f_j(rz)$. Since $\omega\left((\rho_j \circ f_j)_*\left(ie^{it}\right)\right)$ is bounded on $D \setminus \{0\}$ and $\left(f_{j,r}^*\right)_{e^{it}}\left(ie^{it}\right) =$

$r\left(f_j^*\right)_{re^{it}}\left(ie^{it}\right)$, we have

$$\int_{\partial D}\left(\rho_j\circ f_{j,r}\right)^*\omega=\int_{\partial D}f_{j,r}^*\left(\rho_j^*\omega\right)=\int_{\partial D_r}rf_j^*\left(\rho_j^*\omega\right)=\mathrm{O}\left(r\right).$$

Thus, using $f_{j,r}$ in place of f_j in (16.125) and letting $r\to 0^+$, we obtain

$$\int_{M_1}F^\omega=2\pi\sum\nolimits_{j=1}^n\deg\left(\psi;z_j\right).$$

□

REMARK 16.66. In Theorem 16.65 if the image $\psi\left(M\right)\subset E$ intersects the image $0\left(M\right)\cong M$ of the zero section $0\in C^\infty\left(E\right)$ transversally, then $\sum_{j=1}^n\deg\left(\psi;z_j\right)$ is the **intersection number** of the surfaces $0\left(M\right)$ and $\psi\left(M\right)$ in E (i.e., the algebraic number of signed intersections, where the sign is ± 1, depending on whether the combined orientation of surfaces at an intersection point agrees with that of E). In particular, if M is a compact, orientable, embedded surface M in an orientable 4-manifold X and NM is the normal bundle with the orientation induced by those X and M. Then the intersection number of M and the exponential of a small section of $C^\infty\left(NM\right)$ transverse to the zero section is the **self-intersection number** of M in X, and by Theorem 16.65 this is $c_1\left(NM\right)\left[M\right]$.

CHAPTER 17

Gauge Theoretic Instantons

1. The Yang-Mills Functional

In order to define the Yang-Mills functional on the space $\mathcal{C}(P)$ of connections on a principal G-bundle $\pi : P \to M$, we need an inner product K on the Lie algebra $\mathfrak{g}$ of G. This inner product needs to be invariant under the adjoint action of G on $\mathfrak{g}$, namely

$$K(A,B) = K(ad_g(A), ad_g(B)) = K\left(gAg^{-1}, gBg^{-1}\right)$$

for all $g \in G$ and $A, B \in \mathfrak{g}$, where we continue to assume that G is a matrix group, so that $ad_g(A) = gAg^{-1}$. When G is $\mathrm{O}(n)$ or $\mathrm{SO}(n)$, $\mathfrak{g} = \mathfrak{so}(n) = \left\{A \in \mathrm{GL}(n,\mathbb{R}) : A^T = -A\right\}$, we can use the clearly ad-invariant

$$K(A,B) = -Tr(AB) = Tr\left(AB^T\right).$$

Note that $K(A,B) = Tr\left(AA^T\right) = \sum_{i,j=1}^n A_{ij}^2 > 0$ for $A \neq 0$. When G is $\mathrm{U}(n)$, $\mathfrak{g} = \mathfrak{u}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^* = -A\}$, one can take

$$K(A,B) = -\operatorname{Re}(Tr(AB)) = \operatorname{Re}(Tr(AB^*)).$$

Note that $K(A,A) = \operatorname{Re}(Tr(AA^*)) = \sum_{i,j=1}^n A_{ij}\overline{A_{ij}} = \sum_{i,j=1}^n |A_{ij}|^2 > 0$ for $A \neq 0$. If these are the only cases which are of interest, the reader may skip the next two paragraphs.

For a compact Lie group G, such an ad-invariant K can be produced via integration as follows. Select an arbitrary nonzero $v_0 \in \Lambda^m(\mathfrak{g},\mathbb{R})$. A volume form $v \in \Omega^m(G,\mathbb{R})$ is defined at any $g' \in G$ by

$$v_{g'}(A_1,\ldots,A_m) := v_0\left(R_{g'*}^{-1}(A_1),\ldots,R_{g'*}^{-1}(A_m)\right).$$

Note that v is right-invariant in the sense that $R_g^* v = v$. Indeed, noting that

$$R_{g'g} = R_g \circ R_{g'} \Rightarrow R_{g'g*} = R_{g*} \circ R_{g'*} \Rightarrow R_{g'g*}^{-1} = R_{g'*}^{-1} \circ R_{g*}^{-1},$$

we have

$$\begin{aligned}
\left(R_g^* v\right)_{g'}(A_1,\ldots,A_m) &= v_{g'g}(R_{g*}A_1,\ldots,R_{g*}A_m) \\
&= v_0\left(R_{g'g*}^{-1}R_{g*}A_1,\ldots,R_{g'g*}^{-1}R_{g*}A_m\right) \\
&= v_0\left(R_{g'*}^{-1}R_{g*}^{-1}R_{g*}(A_1),\ldots,R_{g'*}^{-1}R_{g*}^{-1}R_{g*}(A_m)\right) \\
&= v_0\left(R_{g'*}^{-1}(A_1),\ldots,R_{g'*}^{-1}(A_m)\right) \\
&= v_{g'}(A_1,\ldots,A_m).
\end{aligned}$$

For $A, B \in \mathfrak{g}$ and an arbitrary inner product K_0 on $\mathfrak{g}$, consider the form $\alpha_{(A,B)} \in \Omega^m(G, \mathbb{R})$, defined by

$$\alpha_{(A,B)}(g') := K_0\left(ad_{g'}(A), ad_{g'}(B)\right) v_{g'}.$$

Then let

$$K(A,B) := \int_G \alpha_{(A,B)}.$$

We have $\alpha_{(ad_g A, ad_g B)} = R_g^*\left(\alpha_{(A,B)}\right)$, since

$$\begin{aligned}
\alpha_{(ad_g A, ad_g B)}(g_1) &= K_0\left(ad_{g_1}(ad_g A), ad_{g_1}(ad_g B)\right) v_{g_1} \\
&= K_0\left(ad_{g_1 g} A, ad_{g_1 g}(B)\right) R_g^*(v)(g_1) \\
&= R_g^*\left(\alpha_{(A,B)}\right)(g_1).
\end{aligned}$$

The ad-invariance of K then follows, since

$$\begin{aligned}
K(ad_g A, ad_g B) &= \int_G \alpha_{(ad_g A, ad_g B)} = \int_G R_g^*\left(\alpha_{(A,B)}\right) \\
&= \int_G \alpha_{(A,B)} = K(A,B).
\end{aligned}$$

There are other ways of producing ad-invariant inner products K on $\mathfrak{g}$. Recall that $\mathfrak{ad} : \mathfrak{g} \to End(\mathfrak{g})$ is the derivative of ad at the identity, and is given by $\mathfrak{ad}(A)(B) = [A, B]$. The *Killing form* β is a symmetric bilinear form on $\mathfrak{g}$ defined for $A, B \in \mathfrak{g}$ by

$$\beta(A,B) := Tr\left(\mathfrak{ad}(A) \circ \mathfrak{ad}(B)\right),$$

Note that $\mathfrak{ad}\left(gAg^{-1}\right) = ad_g \circ \mathfrak{ad}(A) \circ ad_{g^{-1}}$, since

$$\begin{aligned}
\mathfrak{ad}\left(gAg^{-1}\right)(B) &= \left[gAg^{-1}, B\right] = gAg^{-1}B - BgAg^{-1} \\
&= gA\left(g^{-1}Bg\right)g^{-1} - g\left(g^{-1}Bg\right)Ag^{-1} \\
&= ad_g\left(\mathfrak{ad}(A)\left(ad_{g^{-1}}(B)\right)\right).
\end{aligned}$$

Then β is invariant under the adjoint action, since

$$\begin{aligned}
&\beta\left(gAg^{-1}, gBg^{-1}\right) = Tr\left(\mathfrak{ad}\left(gAg^{-1}\right) \circ \mathfrak{ad}\left(gBg^{-1}\right)\right) \\
&= Tr\left(ad_g \circ \mathfrak{ad}(A) \circ ad_{g^{-1}} \circ ad_g \circ \mathfrak{ad}(B) \circ ad_{g^{-1}}\right) \\
&= Tr\left(ad_g \circ \mathfrak{ad}(A) \circ \mathfrak{ad}(B) \circ ad_{g^{-1}}\right) \\
&= Tr\left(\mathfrak{ad}(A) \circ \mathfrak{ad}(B)\right).
\end{aligned}$$

Of course, β need not be definite; e.g., $\beta = 0$ for abelian groups such as $U(1)$. However, for compact, semi-simple G, Cartan's criterion guarantees that $-\beta$ is positive definite, and hence would serve as an ad-invariant inner product K. If $ad : G \to \mathrm{GL}(\mathfrak{g})$ is irreducible, then any two ad-invariant inner products, say K_1 and K_2, on $\mathfrak{g}$ must agree up to a multiplicative constant, since any eigenspace of K_1 relative to K_2 would be invariant. In what follows, we will assume that G is compact, connected and semi-simple, in which case we can and do take K on $\mathfrak{g}$ to be $-\beta$.

Definition 17.1. Let $\pi : P \to M$ be principal G-bundle, where M has a Riemannian metric h, $\mathfrak{g}$ has an ad-invariant inner product K, and M and G are compact. Let $\mathcal{C}(P)$ denote the set of connection 1-forms on P. The *Yang-Mills functional*

$$YM : \mathcal{C}(P) \to \mathbb{R}^+ := [0, \infty)$$

is defined by

$$YM(\omega) := \tfrac{1}{2}\|\Omega^\omega\|^2 = \tfrac{1}{2}\int_M |\Omega^\omega|^2 \; v_h,$$

where

$$\Omega^\omega = d\omega + \omega\wedge\omega = d\omega + \tfrac{1}{2}[\omega,\omega] \in \overline{\Omega}^2(P,\mathfrak{g}) \cong \Omega^2(M, P\times_G \mathfrak{g}).$$

is the curvature of the connection $\omega \in \mathcal{C}(P)$, and $|\Omega^\omega|^2 = \langle \Omega^\omega, \Omega^\omega\rangle$ in terms of the pairing (16.15) on p. 375.

Recall from Remark 16.11, p. 371, that $\mathcal{C}(P)$ has the structure of an affine space based on the vector space $\overline{\Omega}^1(P,\mathfrak{g})$ which can then be regarded as a formal tangent space of $\mathcal{C}(P)$ at any $\omega\in\mathcal{C}(P)$. For $\tau \in \overline{\Omega}^1(P,\mathfrak{g})$ and $t\in\mathbb{R}$, note that $\omega_t := \omega + t\tau \in \mathcal{C}(P)$. At $t=0$,

$$\begin{aligned}\tfrac{d}{dt}\Omega^{\omega_t} &= \tfrac{d}{dt}\left(d\omega_t + \tfrac{1}{2}[\omega_t,\omega_t]\right) = d\tau + \tfrac{1}{2}([\tau,\omega]+[\omega,\tau])\\ &= d\tau + [\omega,\tau] = d\tau + \mathfrak{ad}(\omega)\wedge\tau = D^\omega\tau.\end{aligned}$$

Thus at $t=0$, we have

$$\tfrac{d}{dt}YM(\omega_t) = \tfrac{d}{dt}\tfrac{1}{2}(\Omega^\omega,\Omega^\omega) = (D^\omega\tau, \Omega^\omega) = (\tau, \delta^\omega\Omega^\omega),$$

where

$$\delta^\omega = (-1)^{n+1} * D^\omega * : \Omega^2(M, P\times_G\mathfrak{g}) \to \Omega^1(M, P\times_G \mathfrak{g})$$

is the formal adjoint of

$$D^\omega : \Omega^1(M, P\times_G\mathfrak{g}) \to \Omega^2(M, P\times_G\mathfrak{g}),$$

namely δ^ω is the covariant codifferential (see Proposition 16.18, p. 377). Since all inner products involved are positive definite, all the "directional derivatives" $\frac{d}{dt}YM(\omega+t\tau)\big|_{t=0}$ will be zero precisely when

$$\delta^\omega\Omega^\omega = 0.$$

This is called the (source-free) *Yang-Mills equation*. Formally, it characterizes the critical points (connections) for the Yang-Mills functional.

We mention that $YM : \mathcal{C}(P) \to \mathbb{R}^+$ is invariant under the action on $\mathcal{C}(P)$ of the group $\mathrm{GA}(P)$ of gauge transformations, since for $F\in \mathrm{GA}(P)$, $F\cdot\omega = \left(F^{-1}\right)^*\omega$ yields

$$\begin{aligned}\Omega^{F\cdot\omega} &= d(F\cdot\omega) + (F\cdot\omega)\wedge(F\cdot\omega) = \left(F^{-1}\right)^*(d\omega + \omega\wedge\omega)\\ &= \left(F^{-1}\right)^*\Omega^\omega = F\cdot\Omega^\omega,\end{aligned}$$

and then $|F\cdot\Omega^\omega|^2 = |\Omega^\omega|^2$ by Corollary 16.30.

If $\dim M = 4$ and M is oriented, then we have the Hodge $*$ operator on $\Omega^2(M, P\times_G\mathfrak{g})$, and it makes sense to speak of Ω^ω as being self-dual ($*\Omega^\omega = \Omega^\omega$) or anti-self-dual ($*\Omega^\omega = -\Omega^\omega$). For these, the Yang-Mills equation always holds, since

$$\delta^\omega\Omega^\omega = - * D^\omega * \Omega^\omega = - * D^\omega * \pm\Omega^\omega = \mp * (D^\omega\Omega^\omega) = 0,$$

by the Bianchi identity $D^\omega\Omega^\omega = 0$ (see Proposition 16.16, p. 373).

Proposition 17.2. *In the notation of Definition 17.1, if Ω^ω is self-dual or anti-self-dual, then ω is an absolute minimum of the functional YM.*

PROOF. Let $\mathfrak{g}_{\mathbb{C}} = \mathbb{C} \otimes \mathfrak{g}$ be the complexification of $\mathfrak{g}$. The inner product K on $\mathfrak{g}$ extends to a Hermitian inner product $K_{\mathbb{C}}$ on $\mathfrak{g}_{\mathbb{C}}$ via

$$K_{\mathbb{C}}(z \otimes A, w \otimes B) = z\overline{w}K_{\mathbb{C}}(A, B) \text{ for } A, B \in \mathfrak{g} \text{ and } z, w \in \mathbb{C}.$$

Then the orthogonal (relative to K) representation $ad : G \to \mathrm{SO}(\mathfrak{g})$ extends to $ad_{\mathbb{C}} : G \to \mathrm{SU}(\mathfrak{g}_{\mathbb{C}})$, and we may form the associated complex vector bundle $E := P \times_G \mathfrak{g}_{\mathbb{C}}$ which has a Hermitian structure inherited from $K_{\mathbb{C}}$. Let $U(E) \to M$ be the bundle of unitary frames. Each $p \in P$ gives rise to a unitary mapping $\widehat{p} : \mathfrak{g} \to E_{\pi(p)}$, simply via

$$\widehat{p}(A) = [p, A].$$

In this way we have an embedding $P \to U(E)$ given by $p \mapsto \widehat{p}$, which makes P a principal subbundle of $U(E)$. A connection ω on P uniquely extends to a connection ω_E on $U(E)$ in such a way that the horizontal subspaces of ω_E at points of P are those of ω. The second Chern character class of E is (see 16.98, p. 417)

$$ch(E)_2 = \frac{-1}{8\pi^2}\left[Tr\left(\Omega^{\omega_E} \wedge \Omega^{\omega_E}\right)\right] \in H^4(M, \mathbb{Q})$$

Note that $\Omega^{\omega_E}|_P \in \overline{\Omega}^k(P, \mathfrak{u}(\mathfrak{g}_{\mathbb{C}})) \cong \Omega^2(M, End(E))$ is related to Ω^{ω} via

$$\Omega^{\omega_E}|_P = \mathfrak{ad}(\Omega^{\omega}).$$

Hence, for $X_i \in T_pP$,

$$\begin{aligned} &Tr\left(\Omega^{\omega_E}(X_1, X_2) \circ \Omega^{\omega_E}(X_3, X_4)\right) \\ &= Tr\left(\mathfrak{ad}\left(\Omega^{\omega}(X_1, X_2)\right) \circ \mathfrak{ad}\left(\Omega^{\omega}(X_1, X_2)\right)\right) \\ &= -K\left(\Omega^{\omega}(X_1, X_2), \Omega^{\omega}(X_1, X_2)\right). \end{aligned}$$

Thus, we have

$$\begin{aligned} &Tr\left(\Omega^{\omega_E} \wedge \Omega^{\omega_E}\right)(X_1, X_2, X_3, X_4) \\ &= \frac{1}{2!2!}\sum_{\sigma}(-1)^{\sigma}\, Tr\left(\Omega^{\omega_E}(X_{\sigma_1}, X_{\sigma_2}) \circ \Omega^{\omega_E}(X_{\sigma_3}, X_{\sigma_4})\right) \\ &= -\tfrac{1}{4}\sum_{\sigma}(-1)^{\sigma} K\left(\Omega^{\omega}(X_{\sigma_1}, X_{\sigma_2}), \Omega^{\omega}(X_{\sigma_3}, X_{\sigma_4})\right) \\ &= -K\left(\Omega^{\omega} \wedge \Omega^{\omega}\right)(X_1, X_2, X_3, X_4) \\ &= -K\left(\Omega^{\omega} \wedge * * \Omega^{\omega}\right)(X_1, X_2, X_3, X_4) \\ &= -K\left(\Omega^{\omega}, *\Omega^{\omega}\right)\, \pi^*(\nu_h)(X_1, X_2, X_3, X_4). \end{aligned}$$

Hence,

$$ch(E)_2[M] = \frac{-1}{8\pi^2}\int_M Tr\left(\Omega^{\omega_E} \wedge \Omega^{\omega_E}\right) = \frac{1}{8\pi^2}\int_M \left\langle \Omega^{\omega}, *\Omega^{\omega}\right\rangle \nu_h.$$

Incidentally, since $ad_{\mathbb{C}} : G \to \mathrm{SU}(\mathfrak{g}_{\mathbb{C}})$ we have $c_1(E) = 0$, in which case

$$ch(E)_2 = \tfrac{1}{2}\left(c_1(E)^2 - 2c_2(E)\right) = -c_2(E).$$

Writing $\Omega^{\omega} = \Omega^{\omega+} + \Omega^{\omega-}$, we then have

$$8\pi^2 ch(E)_2[M] = \int_M \left\langle \Omega^{\omega}, *\Omega^{\omega}\right\rangle \nu_h = \int_M \left|\Omega^{\omega+}\right|^2 \nu_h - \int_M \left|\Omega^{\omega-}\right|^2 \nu_h$$

Thus,

$$
\begin{aligned}
YM(\omega) &= \tfrac{1}{2}\int_M |\Omega^\omega|^2\ v_h = \tfrac{1}{2}\int_M \left|\Omega^{\omega+}\right|^2\ \nu_h + \tfrac{1}{2}\int_M \left|\Omega^{\omega-}\right|^2\ \nu_h \\
&= \tfrac{1}{2}\int_M \left|\Omega^{\omega+}\right|^2\ \nu_h + \tfrac{1}{2}\left(\int_M \left|\Omega^{\omega+}\right|^2\ \nu_h - 8\pi^2 ch\,(E)_2\,[M]\right) \\
&= \int_M \left|\Omega^{\omega+}\right|^2\ \nu_h - 4\pi^2 ch\,(E)_2\,[M]\,, \qquad (17.1)
\end{aligned}
$$

and similarly

$$
YM(\omega) = \tfrac{1}{2}\int_M |\Omega^\omega|^2\ v_h = \int_M \left|\Omega^{\omega-}\right|^2\ \nu_h + 4\pi^2 ch\,(E)_2\,[M]\,. \qquad (17.2)
$$

Thus, if Ω^ω self-dual or anti-self-dual, then ω will furnish an absolute minimum for YM. □

Taking the difference of (17.1) and (17.2) yields

COROLLARY 17.3. *In the notation of the proof of Proposition 17.2,*

$$
8\pi^2 ch\,(E)_2\,[M] = \int_M \left|\Omega^{\omega+}\right|^2\ \nu_h - \int_M \left|\Omega^{\omega-}\right|^2\ \nu_h. \qquad (17.3)
$$

Thus if $ch\,(E)_2\,[M] > 0$, then there is no connection $\omega \in \mathcal{C}(P)$ with anti-self-dual curvature Ω^ω (i.e., $\Omega^{\omega+} = 0$). If $ch\,(E)_2\,[M] < 0$,then there is no $\omega \in \mathcal{C}(P)$ with self-dual curvature Ω^ω (i.e., $\Omega^{\omega-} = 0$). If $ch\,(E)_2\,[M] = 0$, then $\Omega^{\omega+} = 0 \Rightarrow \Omega^\omega = 0$, and $\Omega^{\omega-} = 0 \Rightarrow \Omega^\omega = 0$ (i.e., all $\omega \in \mathcal{C}(P)$ with self-dual or anti-self-dual curvature are flat).

By a convenient abuse of terminology, connections ω for which Ω^ω is self-dual (resp. anti-self-dual) are known as *self-dual* (resp. *anti-self-dual*) *connections*. Such connections are also known as *instantons*, particularly when the base M is S^4. This deserves some explanation. We saw in Section 15.15.2 (p. 361-363) that, generally speaking, an instanton is a solution of a Euclidean action principle, which minimizes the Euclidean action among all suitable paths joining two minima for a potential function. When the general concept is applied to the configuration space of connections (gauge potentials) modulo gauge transformations, an instanton is a minimum for the Yang-Mills functional $\frac{1}{2}\int_{\mathbb{R}^4} |\Omega^\omega|^2\ v_h$ for Euclidean $\mathbb{R}^4$ with the standard metric h, where certain asymptotic conditions at ∞ are imposed on ω which make the integral converge. Moreover, the functional is defined on the quotient space of connections modulo the group of gauge transformations defined on $\mathbb{R}^4$ which tend to the identity at infinity. Recall that the star operator on 2-forms in dimension 4 is conformally invariant, so that self-duality and anti-self-duality are preserved under conformal changes of metric. Moreover, the invariance of $*$ implies that the functional YM is invariant under conformal changes, since

$$
\tfrac{1}{2}\int_{\mathbb{R}^4} |\Omega^\omega|^2\ v_h = \tfrac{1}{2}\int_{\mathbb{R}^4} K\left(\Omega^\omega \wedge *\Omega^\omega\right).
$$

Alternatively, note that replacing h by λh, $|\Omega^\omega|^2$ and v_h acquire factors of λ^{-2} and λ^2 respectively. Since $\mathbb{R}^4$ is conformally equivalent to $S^4 - \{\infty\}$ via stereographic projection, we can work over $S^4 - \{\infty\}$. It turns out that via a crucial theorem in [?], the physically reasonable asymptotic conditions imposed on ω and the gauge transformations at ∞ are precisely those that enable one to extend (over the point

$\infty \in S^4$) a self-dual connection ω for the bundle $\mathbb{R}^4 \times G \to \mathbb{R}^4$ to a self-dual connection ω for a principal G-bundle $\pi : P \to S^4$. The extension of the bundle $\mathbb{R}^4 \times G \to \mathbb{R}^4$ is generally nontrivial, depending on the homotopy class of gauge transformations over the slice $S^3 = \left(\mathbb{R}^3 \cup \{\infty\}\right) \times \{0\}$. For $G = \mathrm{SU}(2)$ this homotopy class is determined by the degree of $\phi' : S^3 \to \mathrm{SU}(2) \cong S^3$, discussed at the end of Section 15.2. Of course, one can work over compact, orientable Riemannian 4-manifolds M other than S^4.

2. Instantons on Euclidean 4-Space

In the sections that follow, we will construct manifolds of gauge-equivalence classes of self-dual connections (instantons) and compute their dimensions. However, these results require the existence of such connections, or else they might just be sophisticated statements about the empty set. Here our primary goal is to produce self-dual connections for principal $\mathrm{SU}(2)$-bundles $P \to S^4$ with arbitrary nonpositive Chern number $-k := c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)[S^4]$. Our construction will be motivated from the standpoint of the Riemannian geometry of conformally related metrics on S^4 or $\mathbb{R}^4$, but first we give a brief history.

The first such instanton that came from the physics community was the BPST instanton, named after A. A. Belavin, A. M. Polykov, A. S. Schwarz and Y. S. Tyupkin (see [**BPST**]). The standard BPST instanton was given on $\mathbb{R}^4$, in the sense that it was presented as an $\mathfrak{su}(2)$-valued 1-form on $\mathbb{R}^4$, which can be regarded as the pullback of a connection 1-form, say ω_0, on the trivial principal $\mathrm{SU}(2)$-bundle $\mathbb{R}^4 \times \mathrm{SU}(2)$ by the global section $x \mapsto (x, I)$. It is not difficult to extend ω_0 to a connection 1-form ω on nontrivial principal $\mathrm{SU}(2)$-bundle $P \to S^4$, namely the quaternionic Hopf bundle $S^7 \to S^4$. Then ω turns out to be the well-known universal connection on this bundle. Conformal transformations of S^4 act on the space of (anti-)self-dual connections on S^4. While the conformal transformations which are isometries of S^4 preserve ω, the 5-parameter family of "boosts" acts effectively on ω to produce a 5-parameter family of instantons. Here, a "boost" is a conformal transformation of a sphere which contracts toward one point of the sphere and dilates about the antipode. The standard example is the function $z \mapsto \alpha z$ on the extended complex plane (Riemann sphere). For S^4, four parameters suffice to locate the pair of antipodal points and the dilation factor is the fifth parameter. The Chern number of the bundle $S^7 \to S^4$ is -1, and so the 5-parameter BPST family of instantons pertains to this case. Instantons for arbitrary negative Chern number $-k$ were exhibited in [**Wit**], and $5k$-parameter families of such instantons were produced by t'Hooft (unpublished), in [**Wil**], and in [**CF**]. In [**JNR**] R. Jackiw, C. Nohl and C. Rebbi used the fact that these families could be augmented to conformally invariant families, to enlarge the number of parameters to $5k+4$. In [**AHS1**], M. F. Atiyah, N. J. Hitchen and I. M. Singer applied the index theorem to find that the maximal number of effective parameters is $8k - 3$ ($k \geq 1$), a result which was also derived in [**Schw79**]. Not long after, the analysis was applied to arbitrary principal G-bundles over S^4 (G simple and compact); see [**AHS2**] and [**BCGW**]. The problem of actually constructing the most general $8k - 3$ family of solutions was solved in [**AW**] using the techniques originating with the twistor theory of Roger Penrose and some algebraic geometry. A construction involving "only" linear algebra was finally developed by M. F. Atiyah, V. G. Drinfield, N. J.

Hitchen and Y. I. Manin in [**ADHM**]. We will describe this construction later in this section. The interested reader may augment this brief history by consulting the excellent surveys [**EGH**] and [**Mad**].

In what follows, we will derive the Jackiw-Nohl-Rebbi $(5k+4)$-parameter family of instantons in a very natural way from the viewpoint of Riemannian geometry. Recall from (16.96) that corresponding to the decomposition $\Lambda^2\left(\mathbb{R}^4\right) \cong \Lambda^+ \oplus \Lambda^-$ into self-dual and anti-self-dual 2-forms, there is a Lie algebra decomposition $\mathfrak{so}(4) \cong \mathfrak{so}^+ \oplus \mathfrak{so}^-$ which we can make explicit as follows. Under the index-lowering isomorphism $\Lambda^2\left(\mathbb{R}^4\right) \cong \mathfrak{so}(4)$, we have $e^2\wedge e^3 \pm e^1\wedge e^4$, $e^3\wedge e^1 \pm e^2 \wedge e^4$, and $e^1 \wedge e^2 \pm e^3 \wedge e^4$ corresponding to

$$\begin{bmatrix} 0 & 0 & 0 & \pm 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ \mp 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & \pm 1 \\ 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pm 1 \\ 0 & 0 & \mp 1 & 0 \end{bmatrix} \tag{17.4a}$$

respectively. Choosing $+$ in $\pm$ (and $-$ in $\mp$), these are the *self-dual 't Hooft matrices* η_1, η_2 and η_3, which form a basis of $\mathfrak{so}^+$. Choosing $-$ in $\pm$ (and $+$ in $\mp$), we have the *anti-self-dual 't Hooft matrices* $\overline{\eta}_1$, $\overline{\eta}_2$ and $\overline{\eta}_3$ which form a basis of $\mathfrak{so}^-$. One readily verifies that $[\eta_a, \eta_b] = -2\varepsilon_{abc}\eta_c$ and $[\overline{\eta}_a, \overline{\eta}_b] = -2\varepsilon_{abc}\overline{\eta}_c$. For the Hermitian Pauli matrices

$$\sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

we have $i\sigma_a \in \mathfrak{su}(2)$ with $[i\sigma_a, i\sigma_b] = -2\varepsilon_{abc} i\sigma_c$. Thus, $\eta_a \mapsto i\sigma_a$ defines an isomorphism $\mathfrak{so}^+ \cong \mathfrak{su}(2)$, while $\overline{\eta}_a \mapsto i\sigma_a$ yields $\mathfrak{so}^- \cong \mathfrak{su}(2)$.

Recall from Definition 16.39 that the Levi-Civita connection θ for a Riemannian 4-manifold M is an $\mathfrak{so}(4)$-valued connection 1-form on the bundle of FM of orthonormal frames. Relative to the isomorphisms $\mathfrak{so}(4) \cong \mathfrak{so}^+ \oplus \mathfrak{so}^- \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, the connection θ splits into two $\mathfrak{su}(2)$-valued forms θ^+ and θ^-. Suppose that M is simply $\mathbb{R}^4$ with some metric tensor h. Then there is a global section $\sigma : \mathbb{R}^4 \to F\mathbb{R}^4$ (e.g., apply the Gram-Schmidt procedure to the standard coordinate fields $\partial_i := \frac{\partial}{\partial x^i}$, $i = 1, \cdots, n$). We may pull back θ^+ and θ^- to obtain the $\mathfrak{su}(2)$-valued forms $A^\pm := \sigma^*\theta^\pm$ on $\mathbb{R}^4$. Let

$$F^\pm := \sigma^*\left(\Omega^{\theta^\pm}\right) = dA^\pm + \tfrac{1}{2}\left[A^\pm, A^\pm\right]$$

be the field strengths. It is not difficult to find conditions on the metric h such that $*F^+ = F^+$ or $*F^- = F^-$. Indeed, $F^+ \oplus F^-$ is the decomposition of the curvature operator $\widehat{R} \in End\left(\Omega^2(M)\right)$ according to the decomposition of its values in $\Omega^2(M) = \Omega^+(M) \oplus \Omega^-(M)$. Writing

$$\widehat{R} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},\ * = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix},$$

as in (16.76), we have $F^+ = \begin{bmatrix} A & B \end{bmatrix}$ while $F^- = \begin{bmatrix} B^T & C \end{bmatrix}$. Then

$$\begin{aligned} *F^+ &= \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = \begin{bmatrix} A & -B \end{bmatrix} \\ *F^- &= \begin{bmatrix} B^T & C \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = \begin{bmatrix} B^T & -C \end{bmatrix}. \end{aligned} \tag{17.5}$$

Hence,

$$\begin{aligned} F^- \text{ is self-dual} &\Leftrightarrow C = 0 \\ F^+ \text{ is anti-self-dual} &\Leftrightarrow A = 0 \\ F^+ \text{ is self-dual} &\Leftrightarrow B = 0 \Leftrightarrow F^- \text{is anti-self-dual.} \end{aligned}$$

From (16.79) we have

$$A = \frac{s(R)}{12} + \widetilde{A} \ \text{ and } \ C = \frac{s(R)}{12} + \widetilde{C}$$

and the Weyl tensor (as a curvature type operator) is

$$W = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{C} \end{bmatrix}.$$

Thus, if $W = 0$ (i.e., h is conformally flat) and $s(R) = 0$, then $A = 0$ and $C = 0$. Hence, in this case F^- is self-dual and F^+ is anti-self-dual. Recall that $B = 0$ if the traceless Ricci tensor $R_{ij} - \frac{1}{4}s(R)\,h = 0$, in which case h is an Einstein metric. It is not easy to produce Einstein metrics, but it is trivial to produce conformally flat metrics, namely take

$$h = f^2 ds^2 := f^2\left(\left(dx^1\right)^2 + \cdots (dx^n)^2\right) \ \text{ for } \ 0 < f \in C^\infty(\mathbb{R}^n). \tag{17.6}$$

The next proposition provides the Levi-Civita connection and the curvature tensor of such h for arbitrary n. We also need $s(R) = 0$ to conclude that F^- is self-dual and F^+ is anti-self-dual. In the case $n = 4$, this proposition says that $s(R) = 0 \Leftrightarrow \Delta f = 0$ (i.e., f is harmonic).

Proposition 17.4. *Let $F\mathbb{R}^n \to \mathbb{R}^n$ be the frame bundle of $\mathbb{R}^n$ with the metric $h = f^2 ds^2$, and let $\tau : \mathbb{R}^n \to F\mathbb{R}^n$ be the global section $\tau(x) = \frac{1}{f(x)}(\partial_1, \cdots, \partial_n)$, where $\partial_i = \partial/\partial x^i$ are the standard coordinate vector fields. If $\theta \in \mathcal{C}(F\mathbb{R}^n) \subseteq \Omega^1(F\mathbb{R}^n, \mathfrak{so}(4))$ is the Levi-Civita connection of h, then $\tau^*\theta \in \Omega^1(\mathbb{R}^n, \mathfrak{so}(4))$ is given by*

$$(\tau^*\theta)^i{}_j = \frac{1}{f}\left(\partial_j f \ dx^i - \partial_i f \ dx^j \ \right).$$

Let $\Omega^\theta \in \overline{\Omega}^2(\mathbb{R}^n, \mathfrak{so}(4))$ be the curvature of θ. With

$$\begin{aligned} \left(\tau^*\Omega^\theta\right)^h{}_i &= \tfrac{1}{2}\sum \left(\tau^*\Omega^\theta\right)^h{}_{ijk}\, dx^j \wedge dx^k \\ &= \tfrac{1}{2}\sum R^h{}_{ijk} dx^j \wedge dx^k, \end{aligned}$$

the curvature tensor $R \in C^\infty\left(\mathbb{R}^n, T^{0,4}\mathbb{R}^n\right)$ with $R_{hijk} := f^2 R^h{}_{ijk}$ is given in terms of the K-N product $\vee$ (see (16.65)) by

$$\begin{aligned} R &= -2f\,\nabla^2 f \vee I + 4\,(df \otimes df) \vee I - |df|^2\, I \vee I \\ &= \left(-2f\,\nabla^2 f + 4\,(df \otimes df) - |df|^2\, I\right) \vee I, \end{aligned} \tag{17.7}$$

where $\left(\nabla^2 f\right)_{ij} = \partial_i \partial_j f$, $(df \otimes df)_{ij} = \partial_i f \partial_j f$, and $|df|^2 = \sum_{i=1}^n (\partial_i f)^2$. The scalar curvature $s(R)$ of h is

$$s(R) = f^{-4}(n-1)\left(-2f\,\Delta f + (4-n)\,|df|^2\right),$$

which is $-6f^{-3}\Delta f$ when $n = 4$.

PROOF. Regarding $\tau(x)$ as an isometry $\mathbb{R}^n \to T_x\mathbb{R}^n$ we have $\tau(x)(\mathbf{e}_i) = f(x)^{-1}(\partial_i)_x$ where $\mathbf{e}_i$ is the i-th standard unit vector in $\mathbb{R}^n$. Recall that the canonical 1-form $\varphi \in \overline{\Omega}^1(F\mathbb{R}^n, \mathbb{R}^n)$ is given by $\varphi_u(X) = u^{-1}\pi_*(X)$. Thus,

$$\begin{aligned}&(\tau^*\varphi)_x\left(f(x)^{-1}(\partial_i)_x\right) = \varphi\left(\tau_*\left(f(x)^{-1}(\partial_i)_x\right)\right)\\&= \tau(x)^{-1}\left(\pi_*\tau_*\left(f(x)^{-1}(\partial_i)_x\right)\right) = \tau(x)^{-1}\left(f(x)^{-1}(\partial_i)_x\right) = \mathbf{e}_i.\end{aligned}$$

Hence,

$$\left(\phi^1, \ldots, \phi^n\right) := \phi := \tau^*\varphi \in \Omega^1(\mathbb{R}^n, \mathbb{R}^n)$$

is the dual coframe field for $\left(f^{-1}\partial_1, \ldots, f^{-1}\partial_n\right)$ in the sense that

$$\phi^j\left(f^{-1}\partial_i\right) = \left((\tau^*\varphi)\left(f(x)^{-1}(\partial_i)_x\right)\right)^j = (\mathbf{e}_i)^j = \delta_i^j.$$

In other words, $\phi^i = f dx^i$. The Levi-Civita connection 1-form $\theta \in \mathcal{C}(F\mathbb{R}^n) \subseteq \Omega^1(F\mathbb{R}^n, \mathfrak{so}(n))$ on $F\mathbb{R}^n$ is uniquely determined by the condition that the torsion vanishes; i.e., $0 = D^\theta\varphi = d\varphi + \theta \wedge \varphi$. We have

$$\begin{aligned}0 &= \tau^*\left(D^\theta\varphi\right) = \tau^*(d\varphi + \theta \wedge \varphi)\\&= d(\tau^*\varphi) + \tau^*(\theta) \wedge (\tau^*\varphi) = d\phi + \tau^*(\theta) \wedge \phi.\end{aligned}$$

Writing $\tau^*(\theta)$ as a skew-symmetric matrix $(\theta^i{}_j)$ of 1-forms $\theta^i{}_j$, this becomes $d\phi^i = -\sum_{j=1}^n \theta^i{}_j \wedge \phi^j$. On the other hand,

$$\begin{aligned}d\phi^i &= d\left(f dx^i\right) = df \wedge dx^i = \sum_{j=1}^n \partial_j f \; dx^j \wedge dx^i\\&= -\sum_{j=1}^n \partial_j f \; dx^i \wedge dx^j = -\sum_{j=1}^n f^{-1}\partial_j f \; dx^i \wedge \phi^j\\&= -\sum_{j=1}^n f^{-1}\left(\partial_j f \; dx^i - \partial_i f \; dx^j\right) \wedge \phi^j.\end{aligned}$$

Thus, by the uniqueness of θ, we have

$$\theta^i{}_j = f^{-1}\left(\partial_j f \; dx^i - \partial_i f \; dx^j\right).$$

Since the curvature $\Omega^\theta \in \overline{\Omega}^2 \left(F\mathbb{R}^n, \mathfrak{so}\left(n\right)\right)$ is given by $\Omega^\theta = d\theta + \theta \wedge \theta$,

$$
\begin{aligned}
&\tau^* \left(\Omega^\theta\right)^h_{\ i} = d\theta^h_{\ i} + \sum_{p=1}^{n} \theta^h_{\ p} \wedge \theta^p_{\ i} \\
&= d\left(f^{-1}\left(\partial_i f \ dx^h - \partial_h f \ dx^i\right)\right) \\
&+ f^{-2} \sum_{p=1}^{n} \left(\partial_p f \ dx^h - \partial_h f \ dx^p\right) \wedge \left(\partial_i f \ dx^p - \partial_p f \ dx^i\right) \\
&= d\left(f^{-1}\right) \wedge \left(\partial_i f \ dx^h - \partial_h f \ dx^i\right) + f^{-1} d\left(\partial_i f \ dx^h - \partial_h f \ dx^i\right) \\
&+ f^{-2} \sum_{p=1}^{n} \left(\partial_p f \ dx^h - \partial_h f \ dx^p\right) \wedge \left(\partial_i f \ dx^p - \partial_p f \ dx^i\right) \\
&= -f^{-2} \sum_{q=1}^{n} \left(\partial_q f \ \partial_i f \ dx^q \wedge \ dx^h - \partial_q f \ \partial_h f \ dx^q \wedge dx^i\right) \\
&+ f^{-1} \sum_{q=1}^{n} \left(\partial_q \partial_i f \ dx^q \wedge dx^h - \partial_q \partial_h f \ dx^q \wedge dx^i\right) \\
&+ f^{-2} \sum_{p=1}^{n} \Big(\partial_p f \ \partial_i f \ dx^h \wedge dx^p + \partial_h f \ \partial_p f \ dx^p \wedge dx^i - \left(\partial_p f\right)^2 \ dx^h \wedge dx^i\Big) \\
&= f^{-1} \sum_{q=1}^{n} \left(\partial_q \partial_i f \ dx^q \wedge dx^h - \partial_q \partial_h f \ dx^q \wedge dx^i\right) \\
&+ f^{-2} \sum_{p=1}^{n} \Big(2\partial_p f \ \partial_i f \ dx^h \wedge dx^p + 2\partial_h f \ \partial_p f \ dx^p \wedge dx^i - \left(\partial_p f\right)^2 \ dx^h \wedge dx^i\Big).
\end{aligned}
$$

Thus, using the fact $\left(dx^q \wedge dx^h\right)\left(\partial_j, \partial_k\right) = \delta^q_j \delta^h_k - \delta^q_k \delta^h_j$,

$$
\begin{aligned}
R_{hijk} &= f^2 R^h_{\ ijk} = \tau^* \left(\Omega^\theta\right)^h_{\ i} \left(\partial_j, \partial_k\right) \\
&= f \sum_{q=1}^{n} \left(\partial_q \partial_i f \ \left(\delta^q_j \delta^h_k - \delta^q_k \delta^h_j\right) - \partial_q \partial_h f \ \left(\delta^q_j \delta^i_k - \delta^q_k \delta^i_j\right)\right) \\
&+ \sum_{p=1}^{n} \left(2\partial_p f \ \partial_i f \ \left(\delta^h_j \delta^p_k - \delta^h_k \delta^p_j\right) + 2\partial_h f \ \partial_p f \ \left(\delta^p_j \delta^i_k - \delta^p_k \delta^i_j\right)\right) \\
&- \sum_{p=1}^{n} \left(\partial_p f\right)^2 \ \left(\delta^h_j \delta^i_k - \delta^h_k \delta^i_j\right)
\end{aligned}
$$

$$
\begin{aligned}
&= f\left(\partial_j \partial_i f \ \delta^h_k - \partial_k \partial_i f \ \delta^h_j - \left(\partial_j \partial_h f \ \delta^i_k - \partial_k \partial_h f \ \delta^i_j\right)\right) \\
&+ 2 \sum_{p=1}^{n} \left(\left(\partial_k f \ \partial_i f \ \delta^h_j - \partial_j f \ \partial_i f \ \delta^h_k\right) + \left(\partial_h f \ \partial_j f \ \delta^i_k - \partial_h f \ \partial_k f \ \delta^i_j\right)\right) \\
&+ (\sum_{p=1}^{n} - \left(\partial_p f\right)^2) \left(\delta^h_j \delta^i_k - \delta^h_k \delta^i_j\right).
\end{aligned}
$$

In terms of the K-N product (see (16.65))

$$(Q \vee I)_{hijk} = \tfrac{1}{2}\left(Q_{ik}\delta_{hj} - Q_{hk}\delta_{ij} + Q_{hj}\delta_{ik} - Q_{ij}\delta_{hk}\right),$$

we can write

$$\begin{aligned} R &= -2f\,\nabla^2 f \vee I + 4\,(df \otimes df) \vee I - |df|^2\, I \vee I \\ &= \left(-2f\,\nabla^2 f + 4\,(df \otimes df) - |df|^2\, I\right) \vee I, \end{aligned} \tag{17.8}$$

where $\left(\nabla^2 f\right)_{ij} = \partial_i \partial_j f$, $(df \otimes df)_{ij} = \partial_i f \partial_j f$, and $|df|^2 = \sum_{i=1}^n (\partial_i f)^2$. The scalar curvature of h is then

$$\begin{aligned} s(R) &= f^{-4}\,(n-1)\,Tr\left(-2f\,\nabla^2 f + 4\,(df \otimes df) - |df|^2\, I\right) \\ &= f^{-4}\,(n-1)\left(-2f\,\Delta f + (4-n)\,|df|^2\right), \end{aligned}$$

where the factor of f^{-4} comes from raising two indices using h (i.e., $s(R) = R^{hi}{}_{hi}$). □

We can express the connection $A^- := \tau^*\theta^- \in \Omega^1\left(\mathbb{R}^4, \mathfrak{su}(2)\right)$ and field strength $F^- := \tau^*\left(\Omega^{\theta-}\right) \in \Omega^-\left(\mathbb{R}^4, \mathfrak{su}(2)\right)$ using the 't Hooft and Pauli matrices as follows. The projection of $\tau^*\theta \in \Omega^1\left(\mathbb{R}^4, \mathfrak{so}(4)\right)$ onto $\Omega^1\left(\mathbb{R}^4, \mathfrak{so}^-\right)$ is

$$\tfrac{1}{4}\sum_{a=1}^{3}\left(\sum_{i,j=1}^{4} (\tau^*\theta)^{ij}\,(\overline{\eta}_a)_{ij}\right)\overline{\eta}_a,$$

where the $\frac{1}{4}$ comes from the fact that $\sum_{i,j=1}^4 (\overline{\eta}_a)_{ij}\,(\overline{\eta}_a)_{ij} = 4$ for each a. Under the isomorphism $\mathfrak{so}^- \cong \mathfrak{su}(2)$ given by $\overline{\eta}_a \mapsto i\sigma_a$,

$$\begin{aligned} A^- &= \frac{1}{4}\sum_{a=1}^{3}\left(\sum_{k,j=1}^{4} (\tau^*\theta)^{kj}\,(\overline{\eta}_a)_{kj}\right) i\sigma_a \\ &= \frac{1}{4}\sum_{a=1}^{3}\left(\sum_{k,j=1}^{4} f^{-1}\left(\partial_j f\ dx^k - \partial_k f\ dx^j\right)(\overline{\eta}_a)_{kj}\right) i\sigma_a \\ &= \frac{1}{2}\sum_{a=1}^{3}\sum_{k,j=1}^{4} (\overline{\eta}_a)_{kj}\,\partial_j\,(\log f)\,dx^k\,(i\sigma_a). \end{aligned} \tag{17.9}$$

There is an alternate expression in terms of vector notation. First, for $\left(\mathbf{u}, u^4\right) := \left(u^1, u^2, u^3, u^4\right)$, we have (using the Einstein summation convention)

$$\begin{aligned} u^k v^j \overline{\eta}^a{}_{kj}\sigma_a &= u^k v^j\left(\varepsilon_{abc}\delta^b_k\delta^c_j - \left(\delta^a_k\delta^4_j - \delta^a_j\delta^4_k\right)\right)\sigma_a \\ &= \left(\varepsilon_{abc}\delta^b_k\delta^c_j u^k v^j - \left(\delta^a_k\delta^4_j u^k v^j - \delta^a_j\delta^4_k u^k v^j\right)\right)\sigma_a \\ &= \left(\varepsilon_{abc}u^b v^c - \left(v^4 u^a - v^a u^4\right)\right)\sigma_a \\ &= \left(\mathbf{u} \times \mathbf{v} + u^4\mathbf{v} - v^4\mathbf{u}\right)\cdot\sigma, \end{aligned} \tag{17.10}$$

where

$$\mathbf{w}\cdot\sigma := w^1\sigma_1 + w^2\sigma_2 + w^3\sigma_3.$$

Readers who have worked with the algebra $\mathbb{H}$ of quaternions

$$u = u^4 + u^1\mathbf{i} + u^2\mathbf{j} + u^3\mathbf{k},$$

where $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$, $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$, etc., will sense their involvement in (17.10). As ensuing developments are more easily expressed with quaternions, we recall some basic facts. By definition

$$\begin{aligned}&\operatorname{Im}(u) := u^1\mathbf{i} + u^2\mathbf{j} + u^3\mathbf{k} = \mathbf{u}\cdot(\mathbf{i},\mathbf{j},\mathbf{k}) \quad \text{and}\\ &\overline{u} := u^4 - u^1\mathbf{i} - u^2\mathbf{j} - u^3\mathbf{k} = u^4 - \operatorname{Im}(u) = u^4 - \mathbf{u}\cdot(\mathbf{i},\mathbf{j},\mathbf{k}),\end{aligned}$$

One computes that the quaternion product $u\overline{v}$ of $u \in \mathbb{H}$ with $\overline{v} \in \mathbb{H}$ is given by

$$u\overline{v} = u \cdot v - \left(\mathbf{u} \times \mathbf{v} + u^4\mathbf{v} - v^4\mathbf{u}\right) \cdot (\mathbf{i},\mathbf{j},\mathbf{k}),$$

where $u \cdot v$ is the ordinary dot product of u and v as vectors in $\mathbb{R}^4$. Note that $u\overline{u} = u \cdot u$ and

$$\begin{aligned}v\overline{u} &= v \cdot u - \left(\mathbf{v} \times \mathbf{u} + v^4\mathbf{u} - u^4\mathbf{v}\right) \cdot (\mathbf{i},\mathbf{j},\mathbf{k})\\ &= u \cdot v + \left(\mathbf{u} \times \mathbf{v} + u^4\mathbf{v} - v^4\mathbf{u}\right) \cdot (\mathbf{i},\mathbf{j},\mathbf{k})\\ &= \overline{u\overline{v}}\end{aligned}$$

There is an algebra isomorphism $\mathbb{H} \cong \mathbb{R}I + i\mathfrak{su}(2)$ given by

$$u^4 + u^1\mathbf{i} + u^2\mathbf{j} + u^3\mathbf{k} \longleftrightarrow u^4 I - u^1 i\sigma_1 - u^2 i\sigma_2 - u^2 i\sigma_3 = u^4 I - \mathbf{u}\cdot i\sigma.$$

Under this isomorphism, we have

$$\begin{aligned}u^k v^j \overline{\eta}^a{}_{kj}\,(i\sigma)_a &= \left(\mathbf{u} \times \mathbf{v} + u^4\mathbf{v} - v^4\mathbf{u}\right) \cdot i\sigma\\ &\longleftrightarrow -\left(\mathbf{u} \times \mathbf{v} + u^4\mathbf{v} - v^4\mathbf{u}\right) \cdot (\mathbf{i},\mathbf{j},\mathbf{k}) = \operatorname{Im}(u\overline{v}).\end{aligned}$$

Hence, we obtain

$$\begin{aligned}A^- &= \frac{1}{2}\sum_{a=1}^{3}\sum_{k,j=1}^{4} (\overline{\eta}_a)_{kj}\,\partial_j(\log f)\,dx^k\,(i\sigma_a)\\ &\longleftrightarrow \tfrac{1}{2}\operatorname{Im}\left(dx\overline{\partial(\log f)}\right) = -\tfrac{1}{2}\operatorname{Im}\left(\partial(\log f)\,d\overline{x}\right),\end{aligned} \tag{17.11}$$

where the quaternion differential dx is

$$dx = dx^4 + dx^1\mathbf{i} + dx^2\mathbf{j} + dx^3\mathbf{k},$$

and the quaternion-valued function $\partial(\log f)$ is given by

$$\begin{aligned}\partial(\log f) &:= \partial_4(\log f) + \nabla(\log f)\cdot(\mathbf{i},\mathbf{j},\mathbf{k})\\ &:= \partial_4(\log f) + \partial_1(\log f)\,\mathbf{i} + \partial_2(\log f)\,\mathbf{j} + \partial_3(\log f)\,\mathbf{k}.\end{aligned}$$

Alternatively, in non-quaternionic notation with $u = dx = \left(d\mathbf{x}, dx^4\right)$ and $v = \partial \log f = (\nabla(\log f), \partial_4(\log f))$, we have

$$A^- = \frac{i}{2}\left(d\mathbf{x} \times \nabla(\log f) + dx^4\nabla(\log f) - \partial_4(\log f)\,d\mathbf{x}\right)\cdot\sigma, \tag{17.12}$$

which is considerably less tidy than (17.11). Before considering the field strength F^-, we mention that a direct computation yields

$$
\begin{aligned}
-\tfrac{1}{2}d\overline{x}\wedge dx = &\ \left(dx^2\wedge dx^3 + dx^1\wedge dx^4\right)\mathbf{i}\\
&+\left(dx^3\wedge dx^1 + dx^2\wedge dx^4\right)\mathbf{j}\\
&+\left(dx^1\wedge dx^2 + dx^3\wedge dx^4\right)\mathbf{k}\ \text{ and}\\
-\tfrac{1}{2}dx\wedge d\overline{x} = &\ \left(dx^2\wedge dx^3 - dx^1\wedge dx^4\right)\mathbf{i}\\
&+\left(dx^3\wedge dx^1 - dx^2\wedge dx^4\right)\mathbf{j}\\
&+\left(dx^1\wedge dx^2 - dx^3\wedge dx^4\right)\mathbf{k}.
\end{aligned}
\tag{17.13}
$$

Thus, $d\overline{x}\wedge dx$ is self-dual and $dx\wedge d\overline{x}$ is anti-self-dual.

According to (17.8),

$$F^-_{jk} = \frac{i}{4}\sum_{a=1}^{3}\left(\sum_{h,i=1}^{4}(\overline{\eta}_a)_{hi}\,(P_f\vee I)_{hijk}\right)\sigma_a,$$

where

$$P_f := -2f\,\nabla^2 f + 4\,(df\otimes df) - |df|^2\,I. \tag{17.14}$$

We now consider some obvious choices for the harmonic dilation factor f. Of course for $f=1$, we obtain the standard flat metric ds^2 with $R=0$, $A^\pm=0$ and $F^\pm=0$. For $f(x)=\lambda^2 r^{-2}$ (where $0<\lambda\in\mathbb{R}$ and $r:=|x|>0$) one also finds that $R=0$, either by direct computation of P_f, or by verifying that f^2ds^2 is the pull-back of ds^2 under the inversion $x\mapsto\lambda^2r^{-2}x$ of $\mathbb{R}^4$ in the sphere $r=\lambda$. If we add, taking

$$f = 1+\lambda^2r^{-2} = 1+f_0(x)\,, \tag{17.15}$$

then

$$
\begin{aligned}
P_f &= -2\,(1+f_0)\,\nabla^2 f_0 + 4\,(df_0\otimes df_0) - |df_0|^2\,I\\
&= -2\nabla^2 f_0 = -4\lambda^2 r^{-6}\left(4x\otimes x - r^2 I\right),
\end{aligned}
$$

since

$$
\begin{aligned}
\left(\nabla^2 f_0\right)_{ij} &= \partial_j\partial_i f_0 = -2\lambda^2\partial_j\left(r^{-4}x^i\right)\\
&= -2\lambda^2\partial_j\left(r^{-4}\right)x^i - 2\lambda^2 r^{-4}\delta_{ij}\\
&= 4\lambda^2\left(r^2\right)^{-3}\partial_j\left(r^2\right)x^i - 2\lambda^2 r^{-4}\delta_{ij}\\
&= 2\lambda^2 r^{-6}\left(4x^jx^i - r^2\delta_{ij}\right).
\end{aligned}
\tag{17.16}
$$

Thus, for $f=1+\lambda^2r^{-2}$, we have a nonzero field strength. Moreover, in this case

$$
\begin{aligned}
\partial_i(\log f) &= \partial_i\left(\log\left(1+\lambda^2r^{-2}\right)\right) = \frac{\lambda^2\partial_i\left(r^{-2}\right)}{1+\lambda^2r^{-2}}\\
&= \frac{-2\lambda^2r^{-4}x^i}{1+\lambda^2r^{-2}} = -2\frac{\lambda^2}{r^2}\frac{x^i}{r^2+\lambda^2},
\end{aligned}
$$

or in quaternion notation,

$$\partial(\log f) = -2\frac{\lambda^2}{r^2}\frac{x}{r^2+\lambda^2}\quad\text{and}\quad\overline{\partial(\log f)} = -2\frac{\lambda^2}{r^2}\frac{\overline{x}}{r^2+\lambda^2}.$$

Hence,

$$\begin{aligned} A^-(x) &\longleftrightarrow \tfrac{1}{2}\mathrm{Im}\left(dx\overline{\partial(\log f)}\right) = \tfrac{1}{2}\mathrm{Im}\left(dx\left(-2\frac{\lambda^2}{r^2}\frac{\overline{x}}{r^2+\lambda^2}\right)\right) \\ &= -\mathrm{Im}\left(dx\left(\frac{\lambda^2}{r^2}\frac{\overline{x}}{r^2+\lambda^2}\right)\right) = \mathrm{Im}\left(\frac{\lambda^2}{r^2}\frac{xd\overline{x}}{r^2+\lambda^2}\right) \\ &= \frac{\lambda^2}{r^2}\frac{1}{r^2+\lambda^2}\mathrm{Im}(xd\overline{x}) \in \mathbb{H}, \end{aligned}$$

or alternatively,

$$\begin{aligned} A^-(x) &= -\frac{\lambda^2}{r^2}\frac{1}{r^2+\lambda^2}\left((d\mathbf{x})\times\mathbf{x} + dx^4\mathbf{x} - x^4 d\mathbf{x}\right)\cdot i\sigma \\ &= \frac{\lambda^2}{r^2}\frac{1}{r^2+\lambda^2}\left(\mathbf{x}\times d\mathbf{x} + x^4 d\mathbf{x} - dx^4\mathbf{x}\right)\cdot i\sigma. \end{aligned}$$

Note that $f = 1+\lambda^2 r^{-2}$ and A^- are singular at $x=0$. However, we will find that a gauge transformation can be applied to A^- to yield a local connection form which extends smoothly over $x=0$. For now, we will compute the value $YM(A^-)$ of the Yang-Mills functional at A^-.

PROPOSITION 17.5. *We have*

$$YM\left(A^-\right) = \tfrac{1}{2}\int_{\mathbb{R}^4}\left|F^-\right|_h^2\ \nu_h = 16\pi^2.$$

PROOF. To avoid confusion in what follows, we denote norms computed with $h = \left(1+\lambda^2 r^{-2}\right)^2 ds^2$ by $|\cdot|_h$ and norms computed with ds^2 by $|\cdot|_e$. Note that due to conformal invariance of the YM functional,

$$\int_{\mathbb{R}^4}\left|F^-\right|_h^2\ \nu_h = \int_{\mathbb{R}^4}\left|F^-\right|_e^2\ d^4x,$$

As stated earlier, we are also using the metric $K=-\beta$ (minus the Killing form on $\mathfrak{su}(2)$). To compute this, we use

$$\mathfrak{ad}(i\sigma_b)(i\sigma_c) = [i\sigma_b, i\sigma_c] = -2\sum_{d=1}^{3}\varepsilon_{bcd}i\sigma_d \quad \text{and}$$

$$\mathfrak{ad}(i\sigma_a)\,\mathfrak{ad}(i\sigma_b)(i\sigma_c) = -2\sum_{d=1}^{3}\varepsilon_{dbc}\mathfrak{ad}(i\sigma_a)(i\sigma_d) = -4\sum_{d,e=1}^{3}\varepsilon_{dbc}\varepsilon_{dae}(i\sigma_e)$$

to deduce that

$$K(i\sigma_a, i\sigma_b) = -Tr\left(\mathfrak{ad}(i\sigma_a)\,\mathfrak{ad}(i\sigma_b)\right) = 4\sum_{c,d=1}^{3}\varepsilon_{dbc}\varepsilon_{dac} = 8\delta_{ab}.$$

The inner product on $\mathfrak{so}^-$ under the isomorphism $\mathfrak{so}^- \cong \mathfrak{su}(2)$, given by $\overline{\eta}_a \mapsto i\sigma_a$, is then twice the contraction inner product $(\overline{\eta}_a)_{ij}(\overline{\eta}_b)^{ij} = 4\delta_{ab}$. Since $A=C=0$

in (17.5), F^+ and F^- have equal magnitude. Thus, we find

$$\begin{aligned}
\left|F^-\right|_e^2 &= \tfrac{1}{2}\sum_{i,j=1}^{4} K\left(F_{ij}^-, F_{ij}^-\right) = \tfrac{1}{4}\sum_{i,j=1}^{4}\left(K\left(F_{ij}^-, F_{ij}^-\right) + K\left(F_{ij}^+, F_{ij}^+\right)\right) \\
&= \tfrac{1}{4}2\sum_{h,k,i,j=1}^{4}\left({R^h}_{kij}\right)^2 = \tfrac{1}{2}f^{-4}\sum_{h,k,i,j=1}^{4}\left(R_{hkij}\right)^2 = \tfrac{1}{2}f^{-4}\left|R\right|_e^2 \\
&= \tfrac{1}{2}f^{-4}\left|P_f \vee I\right|_e^2. \qquad (17.17)
\end{aligned}$$

Using (16.72) with $n = 4$ and (17.16), we have

$$\begin{aligned}
\left|P_f \vee I\right|_e^2 &= 2\left|P_f\right|_e^2 + Tr\left(P_f\right)^2 = 2\left|P_f\right|_e^2 \\
&= 2\left|-4\lambda^2 r^{-6}\left(4x \otimes x - r^2 I\right)\right|_e^2 \\
&= 32\lambda^4 r^{-12}\left|4x \otimes x - r^2 I\right|_e^2 \\
&= 32\lambda^4 r^{-12}\left(4x_i x_j - r^2\delta_{ij}\right)\left(4x^i x^j - r^2\delta^{ij}\right) \\
&= 32\lambda^4 r^{-12}\left(16x_i x_j x^i x^j - 8r^2 x_i x_j \delta^{ij} + r^4\delta_{ij}\delta^{ij}\right) \\
&= 32\lambda^4 r^{-12}\left(16r^4 - 8r^4 + 4r^4\right) \\
&= 384\lambda^4 r^{-8}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left|F^-\right|_e^2 &= \tfrac{1}{2}f^{-4}\left|P_f \vee I\right|_e^2 = \tfrac{1}{2}f^{-4}\left(384\lambda^4 r^{-8}\right) \\
&= \left(1 + \lambda^2 r^{-2}\right)^{-4}\left(192\lambda^4 r^{-8}\right) = \frac{192\lambda^4}{\left(r^2 + \lambda^2\right)^4}, \qquad (17.18)
\end{aligned}$$

and

$$\begin{aligned}
\int_{\mathbb{R}^4}\left|F^-\right|_e^2\, d^4x &= \mathrm{Vol}\left(S^3\right)\int_0^\infty \frac{192\lambda^4}{\left(r^2+\lambda^2\right)^4} r^3 dr \\
&= 2\pi^2 \cdot 192\lambda^4 \int_0^\infty \frac{r^3}{\left(r^2+\lambda^2\right)^4} dr.
\end{aligned}$$

Using the substitution $y = r^2 + \lambda^2$ (with $dy = 2rdr$), we compute

$$\begin{aligned}
\int_0^\infty \frac{r^3}{\left(r^2+\lambda^2\right)^4} dr &= \tfrac{1}{2}\int_{\lambda^2}^\infty \frac{r^2}{\left(r^2+\lambda^2\right)^4} dy \\
&= \tfrac{1}{2}\int_{\lambda^2}^\infty \frac{y - \lambda^2}{y^4} dy \\
&= \tfrac{1}{2}\int_{\lambda^2}^\infty y^{-3} - \lambda^2 y^{-4} dy \\
&= -\tfrac{1}{4}y^{-2} + \tfrac{1}{6}\lambda^2 y^{-3}\Big|_{\lambda^2}^\infty \\
&= \frac{1}{12\lambda^4}.
\end{aligned}$$

Hence,

$$\int_{\mathbb{R}^4}\left|F^-\right|_h^2\,\nu_h = \int_{\mathbb{R}^4}\left|F^-\right|_e^2\, d^4x = \frac{2\pi^2 \cdot 192\lambda^4}{12\lambda^4} = 32\pi^2.$$

$\square$

REMARK 17.6. If we had only integrated over the ball $r \leq \lambda$, we would have obtained half this value, since

$$-\tfrac{1}{4}y^{-2} + \tfrac{1}{6}\lambda^2 y^{-3}\Big|_{2\lambda^2}^{\infty} = \frac{1}{24\lambda^4} = \frac{1}{2}\frac{1}{12\lambda^4}.$$

Thus, the total field strength $\int_{\mathbb{R}^4} |F^-|_h^2 \, \nu_h$ of the instanton is $32\pi^2$, and as half of this is in the ball $r \leq \lambda$, we say that the **size** of the instanton is λ.

Let $\pi_0 : P_0 = \left(\mathbb{R}^4 \setminus \{0\}\right) \times \mathrm{SU}(2) \to \mathbb{R}^4 \setminus \{0\}$ be the trivial principal SU (2)-bundle, and let $\rho : \mathbb{R}^4 \setminus \{0\}$ be the section $\rho(x) = (x, I)$. There is a unique connection 1-form $\omega_0 \in \mathcal{C}(P)$, such that $\rho^*\omega_0 = A^-$. Since we have found that $YM(A^-) < \infty$, the results in [**?**] imply that the P_0 can be extended over 0 and ∞ to a bundle $P \to S^4$ in such a way that ω_0 extends to a smooth connection 1-form ω on P. However, it is instructive to do this explicitly.

We can always extend $\pi_0 : P_0 \to \mathbb{R}^4 \setminus \{0\}$ trivially to $\pi : \mathbb{R}^4 \times \mathrm{SU}(2) \to \mathbb{R}^4$ with $\pi(x, B) = x$, but ω_0 will not extend smoothly. Another possibility is to take a trivial bundle $\pi_1 : P_1 \to \mathbb{R}^4$ but identify it nontrivially with $\pi_0 : P_0 \to \mathbb{R}^4 \setminus \{0\}$ over $\mathbb{R}^4 \setminus \{0\}$. In other words, identify $(x, B) \in P_1$ with $(x, g(x) B) \in P_0$ for some function $g : \mathbb{R}^4 \setminus \{0\} \to \mathrm{SU}(2)$. This identification mapping $Q : P_1 \to P_0$ given by

$$Q(x, B) := (x, g(x) B) = (x, B)\left(B^{-1} g(x) B\right),$$

defines a gauge transformation of P_0, since for $C \in \mathrm{SU}(2)$, we have

$$Q((x, B) C) = Q(x, BC) = (x, g(x) BC) = Q(x, B) C.$$

The associated function $q \in C(P_0, \mathrm{SU}(2))$ of Proposition 16.27 is given by

$$q(x, B) = B^{-1} g(x) B.$$

Under the identification Q the connection form ω_0 on P_0 when viewed on $P_1|_{(\mathbb{R}^4 \setminus \{0\})}$ is just $Q^*\omega_0$. Our task is then to find $g : \mathbb{R}^4 \setminus \{0\} \to \mathrm{SU}(2)$ such that $Q^*\omega_0$ extends smoothly to all of P_1. According to (16.27) in Proposition 16.28,

$$Q^*\omega_0 = q^{-1}dq + q^{-1}\omega_0 q.$$

Let $\rho_1 : \mathbb{R}^4 \setminus \{0\} \to P_1$ denote the section $\rho_1(x) = (x, I)$. If $\widetilde{\rho} := Q \circ \rho_1$, then using the fact $q(\rho_1(x)) = q(x, I) = I^{-1} g(x) I = g(x)$, we have

$$\begin{aligned}
\widetilde{\rho}^*\omega_0 &= \rho_1^*(Q^*\omega_0) = \rho_1^*\left(q^{-1}dq + q^{-1}\omega_0 q\right) \\
&= (q \circ \rho_1)^{-1} d(q \circ \rho_1) + (q \circ \rho_1)^{-1} \rho_1^*\omega_0 (q \circ \rho) \\
&= g^{-1}dg + g^{-1}A^- g. \qquad (17.19)
\end{aligned}$$

We claim that there is $g : \mathbb{R}^4 \setminus \{0\} \to \mathrm{SU}(2)$, such that $\widetilde{\rho}^*\omega_0 = \rho_1^*(Q^*\omega_0)$ extends smoothly at $0 \in \mathbb{R}^4$. Indeed, take

$$g(x) := \frac{1}{r}(x_4 I - i\mathbf{x} \cdot \sigma) \longleftrightarrow \frac{x}{r}, \qquad (17.20)$$

We compute

$$\begin{aligned}
g^{-1}dg &= \frac{\overline{x}}{r} d\left(\frac{x}{r}\right) = \frac{\overline{x}}{r}\left(\frac{dx}{r} - \frac{x r dr}{r^3}\right) = \frac{\overline{x}}{r}\left(\frac{dx}{r} - \frac{x(x \cdot dx)}{r^3}\right) \\
&= \frac{1}{r^2}(\overline{x}dx - (x \cdot dx)) = \frac{1}{r^2}\mathrm{Im}(\overline{x}dx),
\end{aligned}$$

and similarly

$$gd\left(g^{-1}\right)=\frac{1}{r^2}\operatorname{Im}\left(xd\overline{x}\right).$$

Thus,

$$A^-=\frac{\lambda^2}{r^2}\frac{1}{r^2+\lambda^2}\operatorname{Im}\left(xd\overline{x}\right)=\frac{\lambda^2}{r^2+\lambda^2}gd\left(g^{-1}\right).$$

Hence, using the fact $0=dI=d\left(g^{-1}g\right)=d\left(g^{-1}\right)g+g^{-1}dg$, we have

$$\begin{aligned}\widetilde{\rho}^*\omega_0&=g^{-1}dg+g^{-1}A^-g=g^{-1}dg+g^{-1}\left(\frac{\lambda^2}{r^2+\lambda^2}gd\left(g^{-1}\right)\right)g\\&=g^{-1}dg+\frac{\lambda^2}{r^2+\lambda^2}d\left(g^{-1}\right)g=g^{-1}dg-\frac{\lambda^2}{r^2+\lambda^2}g^{-1}dg\\&=\frac{r^2}{r^2+\lambda^2}g^{-1}dg=\frac{r^2}{r^2+\lambda^2}\frac{1}{r^2}\operatorname{Im}\left(\overline{x}dx\right)\\&=\frac{1}{r^2+\lambda^2}\operatorname{Im}\left(\overline{x}dx\right),\end{aligned}$$

which is smooth at $0\in\mathbb{R}^4$. We can show A^- itself (without gauge transformation) extends smoothly at ∞ as follows. Let $J:\mathbb{R}^4\setminus\{0\}\cong\mathbb{R}^4\setminus\{0\}$ be the inversion map $J(x)=\frac{1}{x}=\frac{\overline{x}}{r^2}$. For A^- to extend smoothly at ∞, we need to check that J^*A^- extends smoothly across 0. We have $J^*g=g^{-1}$ (and $J^*\left(g^{-1}\right)=g$), since

$$g\left(J\left(x\right)\right)=g\left(\frac{\overline{x}}{r^2}\right)=\frac{\frac{\overline{x}}{r^2}}{\left|\frac{\overline{x}}{r^2}\right|}=\frac{\overline{x}}{r}=g\left(x\right)^{-1}.$$

Then J^*A^- extends smoothly over $x=0$ by the computation

$$\begin{aligned}J^*A^-&=J^*\left(\frac{\lambda^2}{r^2+\lambda^2}gd\left(g^{-1}\right)\right)=\frac{\lambda^2}{J^*\left(r^2\right)+\lambda^2}J^*\left(g\right)dJ^*\left(g^{-1}\right)\\&=\frac{1}{r^{-2}+\lambda^2}g^{-1}dg=\frac{1}{r^{-2}+\lambda^2}\frac{1}{r^2}\operatorname{Im}\left(\overline{x}dx\right)=\frac{1}{1+r^2\lambda^2}\operatorname{Im}\left(\overline{x}dx\right).\end{aligned}$$

The function $g:\mathbb{R}^4\setminus\{0\}\to\mathrm{SU}(2)\cong S^3$ is used to identify (clutch) P_0 and P_1 over $\mathbb{R}^4\setminus\{0\}$ to obtain $\pi:P\to S^4$. Since g restricts to the (degree 1) identity map $S^3\to S^3$ on $S^3\subseteq\mathbb{R}^4\setminus\{0\}$, we suspect that the Chern number $c_2\left(E'\right)\left[S^4\right]$ of the associated bundle $E':=P\times_{\mathrm{SU}(2)}\mathbb{C}^2$ is ±1, where the representation $r:\mathrm{SU}(2)\to\mathrm{GL}\left(\mathbb{C}^2\right)$ is just inclusion. We now verify that $c_2\left(E'\right)\left[S^4\right]=-1$, using the result $\int_{\mathbb{R}^4}\left|F^-\right|_h^2\nu_h=32\pi^2$ of Proposition 17.5.

Proposition 17.7. *Let $\pi:P\to M$ be an arbitrary principal* $\mathrm{SU}(2)$*-bundle over a compact 4-manifold M. If $E':=P\times_{\mathrm{SU}(2)}\mathbb{C}^2$ and $E:=P\times_{\mathrm{SU}(2)}\mathfrak{su}(2)_{\mathbb{C}}$, then*

$$4c_2\left(E'\right)=c_2\left(E\right)=-ch_2\left(E\right). \tag{17.21}$$

For the bundle $\pi:P\to S^4$ defined above by clutching P_0 and P_1 over $\mathbb{R}^4\setminus\{0\}$, we have $c_2\left(E'\right)\left[S^4\right]=-1$.

Proof. Note that any $C\in\mathrm{End}\left(\mathbb{C}^2\right)\cong\mathbb{C}^2\otimes\mathbb{C}^{2*}$ can be written uniquely in the form $C=zB+wI$ where $B\in\mathfrak{su}(2)$ and $z,w\in\mathbb{C}$, so that $\mathrm{End}\left(\mathbb{C}^2\right)=\mathfrak{su}(2)_{\mathbb{C}}\oplus\mathbb{C}I$ and the summands are $\mathrm{SU}(2)$-invariant subspaces of the representation $\rho:\mathrm{SU}(2)\to\mathrm{End}\left(\mathbb{C}^2\right)$ given by

$$\rho\left(B\right)\left(\xi\right)=B\circ\xi\circ B^{-1}.$$

Hence,

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^2 \otimes \mathbb{C}^{2*} = \operatorname{End}\left(\mathbb{C}^2\right) = \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathbb{C}I$$

Here we have used the fact that $r : \mathrm{SU}(2) \to \mathrm{GL}\left(\mathbb{C}^2\right)$ and $r^* : \mathrm{SU}(2) \to \mathrm{GL}\left(\mathbb{C}^{2*}\right)$ are equivalent. This follows from the fact that the irreducible representations of $\mathrm{SU}(2)$ are determined by dimension $2s+1$ ($s = \text{spin} = 0, \frac{1}{2}, 1, \cdots$), but a direct equivalence can be exhibited as follows. There is a standard skew-symmetric bilinear form ε on $\mathbb{C}^2$ given by

$$\varepsilon\left(\left(v_1, v_2\right), \left(w_1, w_2\right)\right) := v_1 w_2 - v_2 w_1 = \det \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix}.$$

Define $\mathbb{C}^2 \to \mathbb{C}^{2*}$ by $v \mapsto \varepsilon(v, \cdot)$. Then for any $B \in \mathrm{SU}(2)$,

$$\begin{aligned} B(v) &\mapsto \varepsilon(B(v), \cdot) = \varepsilon\left(B^{-1}B(v), B^{-1}(\cdot)\right) \\ &= \det\left(B^{-1}\right)\varepsilon\left(v, B^{-1}(\cdot)\right) = \varepsilon\left(v, B^{-1}(\cdot)\right), \end{aligned}$$

since $\det\left(B^{-1}\right) = \det(B)^{-1} = 1$, and so $v \mapsto \varepsilon(v, \cdot)$ is an equivalence. In terms of the bundles

$$E' := P \times_{\mathrm{SU}(2)} \mathbb{C}^2 \text{ and } E = P \times_{\mathrm{SU}(2)} \mathfrak{su}(2)_{\mathbb{C}}$$

we have (where $1_{\mathbb{C}}$ is the trivial line bundle over S^4)

$$E \oplus 1_{\mathbb{C}} \cong E' \otimes E'$$

We have observed (see Remark 16.59) that for associated $\mathrm{SU}(2)$-bundles V, $ch_1(V) = c_1(V) = 0$ and $ch_2(V) = \frac{1}{2}c_1(V)^2 - c_2(V) = -c_2(V)$. Hence,

$$\begin{aligned} &ch(E) \oplus ch\left(1_{\mathbb{C}}\right) = 4 - 2c_2(E) \text{ and} \\ &\quad ch\left(E' \otimes E'\right) = ch\left(E'\right)^2 = \left(2 - c_2\left(E'\right)\right)^2 = 4 - 4c_2\left(E'\right) \end{aligned}$$

imply that $4c_2(E') = c_2(E) = -ch_2(E)$. In the case of the clutched bundle $\pi : P \to S^4$, we have found that

$$\begin{aligned} 8\pi^2 ch(E)_2\left[S^4\right] &= \int_{S^4} \left|\Omega^{\omega+}\right|^2 \nu_h - \int_{S^4} \left|\Omega^{\omega-}\right|^2 \nu_h \\ &= \int_{S^4} \left|\Omega^{\omega+}\right|^2 \nu_{S^4} = \int_{\mathbb{R}^4} \left|F^-\right|_h^2 \nu_h \\ &= \int_{\mathbb{R}^4} \left|F^-\right|_e^2 \nu_e = 32\pi^2. \end{aligned}$$

Hence $ch(E)_2\left[S^4\right] = 4$ and $4c_2(E')\left[S^4\right] = -ch_2(E)\left[S^4\right] = -4$, and so $c_2(E')\left[S^4\right] = -1$. □

Before going on to find instantons for bundles $\pi : P \to S^4$ with arbitrary negative $c_2(E')\left[S^4\right]$, we will compute the field strength $F^- = \tau^*\Omega^{\theta^-}$ rather easily by using the gauge equivalent local connection form $\widetilde{A} := \widetilde{\rho}^*\omega_0 = \frac{1}{r^2+\lambda^2}\operatorname{Im}(\overline{x}dx)$ instead of A^-.

Proposition 17.8. *The local field strength* $\widetilde{F} := \widetilde{\rho}^*\Omega^{\omega_0} = d\widetilde{A} + \widetilde{A} \wedge \widetilde{A}$ *is given by*

$$\widetilde{F}(x) = \frac{\lambda^2}{\left(r^2+\lambda^2\right)^2}\, d\overline{x} \wedge dx, \tag{17.22}$$

where $x \in \mathbb{H} \cong \mathbb{R}^4$. *Moreover,* $F^- = \tau^* \Omega^{\theta^-}$ *is given by*

$$F^-(x) = g(x)\,\widetilde{F}(x)\,g(x)^{-1} = \frac{\lambda^2}{\left(r^2+\lambda^2\right)^2}\;\frac{x}{r}\,(d\overline{x}\wedge dx)\,\frac{\overline{x}}{r}. \tag{17.23}$$

Proof. We have

$$\begin{aligned}
F' &= dA' + A'\wedge A' \\
&= d\left(\frac{1}{r^2+\lambda^2}\operatorname{Im}(\overline{x}dx)\right) + \frac{1}{r^2+\lambda^2}\operatorname{Im}(\overline{x}dx)\wedge\frac{1}{r^2+\lambda^2}\operatorname{Im}(\overline{x}dx) \\
&= \frac{1}{\left(r^2+\lambda^2\right)^2}\begin{pmatrix} -\operatorname{Im}(d(\overline{x}x)\wedge\overline{x}dx) \\ +\left(r^2+\lambda^2\right)\operatorname{Im}(d\overline{x}\wedge dx) + \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)\end{pmatrix} \\
&= \frac{1}{\left(r^2+\lambda^2\right)^2}\begin{pmatrix} -\operatorname{Im}((d(\overline{x})x+\overline{x}dx)\wedge\overline{x}dx) \\ +\left(r^2+\lambda^2\right)\operatorname{Im}(d\overline{x}\wedge dx) + \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)\end{pmatrix} \\
&= \frac{1}{\left(r^2+\lambda^2\right)^2}\begin{pmatrix} -r^2\operatorname{Im}(d\overline{x}\wedge dx) - \operatorname{Im}(\overline{x}dx\wedge\overline{x}dx) \\ +\left(r^2+\lambda^2\right)\operatorname{Im}(d\overline{x}\wedge dx) + \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)\end{pmatrix} \\
&= \frac{1}{\left(r^2+\lambda^2\right)^2}\left(\lambda^2\operatorname{Im}(d\overline{x}\wedge dx) + \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx) - \operatorname{Im}(\overline{x}dx\wedge\overline{x}dx)\right)
\end{aligned}$$

To obtain (17.22), it remains to show that

$$\operatorname{Im}(\overline{x}dx\wedge\overline{x}dx) = \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx). \tag{17.24}$$

For this, we have

$$\begin{aligned}
&\overline{x}dx\wedge\overline{x}dx \\
&= (\operatorname{Re}(\overline{x}dx)+\operatorname{Im}(\overline{x}dx))\wedge(\operatorname{Re}(\overline{x}dx)+\operatorname{Im}(\overline{x}dx)) \\
&= \operatorname{Re}(\overline{x}dx)\wedge\operatorname{Re}(\overline{x}dx)+\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx) \\
&+ \operatorname{Im}(\overline{x}dx)\wedge\operatorname{Re}(\overline{x}dx)+\operatorname{Re}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx) \\
&= \operatorname{Re}(\overline{x}dx)\wedge\operatorname{Re}(\overline{x}dx)+\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx),
\end{aligned} \tag{17.25}$$

where $\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Re}(\overline{x}dx)+\operatorname{Re}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)=0$ since real and pure imaginary quaternions commute, while 1-forms anticommute. Note that

$$\begin{aligned}
\overline{\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)} &= -\overline{\operatorname{Im}(\overline{x}dx)}\wedge\overline{\operatorname{Im}(\overline{x}dx)} \\
&= -(-\operatorname{Im}(\overline{x}dx))\wedge(-\operatorname{Im}(\overline{x}dx)) \\
&= -\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx).
\end{aligned}$$

So that $\operatorname{Im}(\overline{x}dx)\wedge\operatorname{Im}(\overline{x}dx)$ is pure imaginary, and taking the imaginary part of (17.25) yields (17.24). We have (17.23), either by direct computation using (17.19) or by using the general result (16.28). □

Remark 17.9. Recall (see (17.13)) that

$$-\tfrac{1}{2}d\overline{x}\wedge dx = \left(dx^2\wedge dx^3 + dx^1\wedge dx^4\right)\mathbf{i}+\cdots.$$

Since the 2-forms $dx^2\wedge dx^3 + dx^1\wedge dx^4,\ldots$ are each of norm-square 2 and **i, j,** and **k** have norm-square 8 with the Killing metric on $S^3\cong \mathrm{SU}(2)$, we have that $|d\overline{x}\wedge dx|^2 = 4(3\cdot(2\cdot 8)) = 192$. Thus,

$$\left|F^-\right|_e^2 = |F'|_e^2 = \frac{192\lambda^4}{\left(r^2+\lambda^2\right)^4},$$

in agreement with (17.18).

Suppose that we have found a principal $\mathrm{SU}(2)$-bundle $\pi : P \to S^4$ with $c_2(E')[S^4] = -k$ (recall $E' := P \times_{\mathrm{SU}(2)} \mathbb{C}^2$). Then (see (17.3)) for any connection 1-form ω on P,
(17.26)
$$-4k = 4c_2(E')[S^4] = -ch_2(E)[S^4] = \frac{1}{8\pi^2}\left(\int_{S^4} \left|\Omega^{\omega-}\right|^2 \nu_h - \int_{S^4} \left|\Omega^{\omega+}\right|^2 \nu_h\right)$$
Thus, we can only have self-dual connections ($\Omega^{\omega-} = 0$) if $k \geq 0$ and only anti-self-dual connections ($\Omega^{\omega+} = 0$) if $k \leq 0$. If $k = 0$ (e.g., $P = S^4 \times \mathrm{SU}(2)$), then all such (anti-)self-dual connections are flat, namely $\Omega^\omega = 0$. We produced self-dual connections in the case $k = 1$, by considering the $\mathfrak{so}^-(\cong \mathfrak{su}(2))$-valued component θ^- of the Levi-Civita connection θ for the metric $h = f^2 ds^2$ on $\mathbb{R}^4$ by taking $f = 1 + \lambda^2 r^{-2}$. If we had considered θ^+ for such h, then we would have found anti-self-dual connections for the case $k = -1$. It is reasonable (and correct) that for arbitrary $k \geq 1$, we should consider the harmonic function
$$f(x) := 1 + \sum_{i=1}^{k} \frac{\lambda_i^2}{|x - x_i|^2}$$
for nonzero $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ and distinct points $x_1, \ldots, x_k \in \mathbb{R}^4$. In order to show that the self-dual connection θ^- with
$$A^- = \tau^*\theta^- = -\tfrac{1}{2}\mathrm{Im}\left(\partial(\log f)\, d\overline{x}\right) \tag{17.27}$$
yields a self-dual connection on some principal $\mathrm{SU}(2)$-bundle $\pi : P \to S^4$ with $c_2(E')[S^4] = -k$, we proceed as follows. From (17.17),
$$\left|F^-\right|_e^2 = \tfrac{1}{2} f^{-4} \left\|P_f \vee I\right\|_e^2 = f^{-4} \left\|P_f\right\|_e^2. \tag{17.28}$$
Letting $f_i(x) = \lambda_i^2 |x - x_i|^{-2}$, we have $f = 1 + \sum_i f_i$ $(1 \leq i \leq k)$ and
$$\begin{aligned}
P_f &= -2f\,\nabla^2 f + 4(df \otimes df) - |df|^2 I \\
&= -2\left(1 + \sum\nolimits_i f_i\right)\nabla^2\left(1 + \sum\nolimits_j f_i\right) + 4\left(\sum\nolimits_i df_i \otimes \sum\nolimits_j df_j\right) - \left|\sum\nolimits_i df_i\right|^2 I \\
&= -2\sum\nolimits_i \nabla^2 f_i + \sum\nolimits_{i,j}\left(-2f_j\nabla^2 f_i + 4df_i \otimes df_j - \langle df_i, df_j\rangle I\right) \\
&= -2\sum\nolimits_i \nabla^2 f_i + \sum\nolimits_{i\neq j}\left(-2f_j\nabla^2 f_i + 4df_i \otimes df_j - \langle df_i, df_j\rangle I\right),
\end{aligned}$$
where in the last equality we have used the fact that $f_i^2 ds^2$ is flat (so that $P_{f_i} = 0$) to deduce that the terms with $i = j$ vanish. Setting $r_i(x) := |x - x_i|$, we have
$$\begin{aligned}
f_i &= \lambda_i^2 r_i^{-2}, \\
df_i &= \lambda_i^2 d\left(r_i^{-2}\right) = \lambda_i^2 d\left(\left(r_i^2\right)^{-1}\right) = -\lambda_i^2\left(r_i^2\right)^{-2} d\left(r_i^2\right) \\
&= -\lambda_i^2\left(r_i^2\right)^{-2} 2(x - x_i)\cdot dx, \text{ and} \\
\left(\nabla^2 f_i\right)_{pq} &= \partial_p\partial_q f_i = \partial_p\partial_q\left(\lambda_i^2 r_i^{-2}\right) = \lambda_i^2 \partial_p\partial_q\left(r_i^{-2}\right) \\
&= 2\lambda_i^2 r_i^{-6}\left(4\left(x^p - x_i^p\right)\left(x^q - x_i^q\right) - r_i^2\delta_{pq}\right), \text{ and so}
\end{aligned}$$
$$\begin{aligned}
|df_i|^2 &= \left|-\lambda_i^2\left(r_i^2\right)^{-2} 2(x - x_i)\cdot dx\right|^2 = 4\lambda_i^4 r_i^{-6} \\
\left|\nabla^2 f_i\right|^2 &= 48\lambda^4 r_i^{-8}.
\end{aligned}$$

For $j \neq i$, the quantities $f_j, |df_j|$ and $\left|\nabla^2 f_j\right|$ are bounded about x_i. Thus as $r_i \to 0$,

$$\|P_f\|_e^2 = 4\left|\nabla^2 f_j\right|^2 + \mathrm{O}\left(r_i^{-6}\right) = 192\lambda_i^4 r_i^{-8} + \mathrm{O}\left(r_i^{-6}\right).$$

Then

$$\begin{aligned}\left|F^-\right|_e^2 &= f^{-4}\left|P_f\right|_e^2 = \left(1+\sum\nolimits_i \lambda_i^2 r_i^{-2}\right)^{-4}\left|P_f\right|_e^2 \\ &= \left(1+\lambda_i^2 r_i^{-2} + \sum\nolimits_{i\neq j}\lambda_j^2 r_j^{-2}\right)^{-4}\left(192\lambda_i^4 r_i^{-8} + \mathrm{O}\left(r_i^{-6}\right)\right) \\ &= 192\lambda_i^4\left(r_i^2+\lambda_i^2+\sum\nolimits_{i\neq j}\lambda_j^2\left(r_i/r_j\right)^2\right)^{-4} + \mathrm{O}\left(r_i^2\right),\end{aligned}$$

which is bounded near x_i. Also, as $r := |x| \to \infty$, we have

$$\left|F^-\right|_e^2 = f^{-4}\left|P_f\right|_e^2 = 4\sum\nolimits_i\left|\nabla^2 f_i\right|^2 + \mathrm{O}\left(r^{-10}\right) \leq Cr^{-8}.$$

for some constant C which tends to 0 as $(\lambda_1,\ldots,\lambda_k) \to \mathbf{0}$. Thus, $|F^-|_e^2$ is integrable. By K. Uhlenbeck's theorem (or by directly using clutching functions g_i as in (17.20) about the x_i), one can deduce that A^-, defined by (17.27) on $\mathbb{R}^4 - \{x_1,\ldots,x_k\}$, lifts to a smooth connection 1-form, say ω, for some principal $\mathrm{SU}(2)$-bundle $\pi : P \to S^4$. By (17.26), we can determine the Chern number $c_2(E')\left[S^4\right]$, where $E' := P \times_{\mathrm{SU}(2)} \mathbb{C}^2$. Indeed,

$$c_2(E')\left[S^4\right] = \frac{-1}{32\pi^2}\int_{S^4}\left|\Omega^{\omega+}\right|^2 \nu_h = \frac{-1}{32\pi^2}\int_{\mathbb{R}^4}\left|F^-\right|_e^2 \nu_e.$$

This integral is independent of the nonzero $\lambda_1,\ldots,\lambda_k \in \mathbb{R}$ and the positions distinct points $x_1,\ldots,x_k \in \mathbb{R}^4$. We can evaluate the integral by first choosing $x_1,\ldots,x_k$ far enough apart so that if B_i is the Euclidean ball of radius 1 about x_i, then $r_i/r_j \leq \varepsilon$ in B_i for all $j \neq i$. Then

$$\int_{B_i}\left|F^-\right|_e^2 \nu_e = 192\lambda_i^4\int_{B_i}\left(r_i^2+\lambda_i^2\right)^{-4}\nu_e + \mathrm{O}(\varepsilon).$$

Since $\int_{\mathbb{R}^4 - \cup_i B_i}|F^-|_e^2\,\nu_e \to 0$ as $(\lambda_1,\ldots,\lambda_k) \to \mathbf{0}$, and

$$\begin{aligned}&\lim_{\lambda_i\to 0} 192\lambda_i^4\int_{B_i}\left(r_i^2+\lambda_i^2\right)^{-4}\nu_e = 2\pi^2\cdot 192\lim_{\lambda\to 0}\lambda^4\int_0^1\frac{r^3}{\left(r^2+\lambda^2\right)^4}dr \\ &= 2\pi^2\cdot 192\lim_{\lambda\to 0}\lambda^4\left(-\tfrac{1}{4}y^{-2}+\tfrac{1}{6}\lambda^2 y^{-3}\right)\Big|_{\lambda^2}^1 \\ &= 2\pi^2\cdot 192\cdot\frac{1}{12} = 32\pi^2,\end{aligned}$$

we obtain, as $\varepsilon \to 0$ and $(\lambda_1,\ldots,\lambda_k) \to \mathbf{0}$,

$$c_2(E')\left[S^4\right] = \frac{-1}{32\pi^2}\int_{\mathbb{R}^4}\left|F^-\right|_e^2 \nu_e \to -k$$

But, $c_2(E')\left[S^4\right]$ is independent of ε and $\lambda_1,\ldots,\lambda_k$. Thus, $c_2(E')\left[S^4\right] = -k$.

We can extend the $k+4k = 5k$-parameter family of harmonic functions

$$1+\sum_{i=1}^k\frac{\lambda_i^2}{\left|x-x_i\right|^2}$$

to the $(5k+5)$-parameter family

$$\sum_{i=0}^{k} \frac{\lambda_i^2}{|x-x_i|^2}.$$

The first family is a limiting case of the second if we take $\lambda_0 = |x_0|$ and let $x_0 \to \infty$. Moreover, the analysis of A^- and F^- about the points x_i is the same as before so that the singularities of A^- are removable and A^- lifts and extends to a smooth connection on a principal SU(2)-bundle $\pi : P \to S^4$. The first family is a limiting case of the second if we take $\lambda_0 = |x_0|$ and let $x_0 \to \infty$. Thus, the Chern number $c_2(E')[S^4]$ for the associated bundle $E' = P \times_{\mathrm{SU}(2)} \mathbb{C}^2$ for functions of the second family will also be $-k$. Not all of the parameters of the $5k+5$ are effective, since multiplying f by a constant does not change A^- by (17.9) or (17.12). Thus, there are at most $5k+4$ effective parameters. Actually, there are $5k+4$ effective parameters only for $k \geq 3$. For $k=1$, there are only 5 effective parameters corresponding to the position and size of the BPST instanton, while for $k=2$ there are 13 effective parameters (see [**JNR**]). In the next section, we prove that the correct number of parameters is $8k-3$ for all k. There was really no guarantee that the procedure that we used (i.e., generating connections from conformally-flat metrics $f^2 ds^2$ with vanishing scalar curvature) would give us all of the instantons. There is a procedure (the ADHM procedure) developed by M. F. Atiyah, V. G. Drinfield, N. J. Hitchen, and Y. I. Manin in [**ADHM**] for capturing all instantons, which we describe below. However, we mention that the singular metrics

$$f^2 ds^2 = \left(\sum_{i=0}^{k} \frac{\lambda_i^2}{|x-x_i|^2}\right)^2 ds^2 \tag{17.29}$$

are interesting in their own right. We have noticed that for $f(x) = \lambda^2 r^{-2}$ (i.e., the case $k=1$) one finds that $f^2 ds^2$ is the pull-back of ds^2 under the inversion $x \mapsto \lambda^2 r^{-2} x$ of $\mathbb{R}^4$ in the sphere $r = \lambda$. In other words, when a ball about the origin is given the metric $f^2 ds^2$ it becomes isometric to the *exterior* of a ball of $\mathbb{R}^4$ with the flat metric. While $f^2 ds^2$ in (17.29) is not perfectly flat in a ball about an x_i, the influence of the other terms is comparatively slight, so that the metric is asymptotically flat as one approaches x_i. Thus, the manifold $\mathbb{R}^4 - \{x_1, \ldots, x_k\}$ with metric $f^2 ds^2$ joins together k asymptotically flat regions, and serves to motivate the study of gravitational instantons.

It is not difficult to *describe* the instantons constructed by the ADHM procedure. The difficulty lies in proving that all instantons are captured by the procedure. We will content ourselves with a suitably motivated description. Recall that for a harmonic function f on $\mathbb{R}^4$, a $\mathfrak{su}(2)$-valued connection 1-form with self-dual curvature is defined by

$$A_f = -\tfrac{1}{2}\mathrm{Im}\left(\partial(\log f)\, d\overline{x}\right),$$

where we have identified $\mathfrak{su}(2)$ with the pure imaginary quaternions. When

$$f(x) = 1 + \sum_{j=1}^{k} \frac{\lambda_j^2}{|x-x_j|^2},$$

$$
\begin{aligned}
A_f(x) &= -\tfrac{1}{2}\mathrm{Im}\left(\partial\left(\log f\right)d\overline{x}\right) = -\tfrac{1}{2}\mathrm{Im}\left(\frac{\partial f}{f}d\overline{x}\right) \\
&= \mathrm{Im}\left(\frac{1}{f(x)}\sum_{j=1}^{k}\lambda_j^2\frac{x-x_j}{|x-x_j|^4}d\left(\overline{x}-\overline{x}_j\right)\right)
\end{aligned}
\tag{17.30}
$$

For a quaternion variable x, we apply the convenient identity

$$
\mathrm{Im}\left(\frac{x}{|x|^4}d\overline{x}\right) = \mathrm{Im}\left(\overline{\left(\frac{1}{x}\right)}d\left(\frac{1}{x}\right)\right),
$$

which is derived from the computation

$$
\begin{aligned}
\overline{\left(\frac{1}{x}\right)}d\left(\frac{1}{x}\right) &= \overline{\left(\frac{\overline{x}}{|x|^2}\right)}d\left(\frac{\overline{x}}{|x|^2}\right) \\
&= \frac{x}{|x|^2}d\left(\frac{\overline{x}}{|x|^2}\right) = \frac{x}{|x|^2}\left(\overline{x}d\left(\frac{1}{|x|^2}\right)+\frac{1}{|x|^2}d\overline{x}\right) \\
&= \frac{x}{|x|^2}\overline{x}d\left(\frac{1}{|x|^2}\right)+\frac{x}{|x|^2}\frac{1}{|x|^2}d\overline{x} \\
&= d\left(|x|^{-2}\right)+\frac{x}{|x|^4}d\overline{x},
\end{aligned}
$$

using the fact that $d\left(|x|^{-2}\right)$ is real. Thus, we may rewrite (17.30) as

$$
A_f(x) = \mathrm{Im}\left(\frac{\sum_{j=1}^{k}\overline{\left(\frac{\lambda_j}{x-x_j}\right)}d\left(\frac{\lambda_j}{x-x_j}\right)}{1+\sum_{j=1}^{k}\left(\frac{\lambda_j}{x-x_j}\right)\overline{\left(\frac{\lambda_j}{x-x_j}\right)}}\right).
$$

We can generalize this (admittedly with some hindsight) as follows. Although the λ_j were taken to be real, we now let them be quaternions. For a column matrix $\lambda = [\lambda_1,\ldots,\lambda_k]^T \in \mathbb{H}^k$ and B a $k\times k$ matrix of quaternions, define $u:\mathbb{H}\to\mathbb{H}^k$ by

$$
u(x) := (B - xI_k)^{-1}\lambda.
$$

For

$$
\begin{aligned}
&u^* := \overline{u(x)}^T = \left[\overline{u_1(x)},\ldots,\overline{u_k(x)}\right], \\
&u^*du := \sum_{j=1}^{k}\overline{u_j(x)}du_j, \text{ and} \\
&|u(x)|^2 := u^*u = \sum_{j=1}^{k}|u_j(x)|^2, \text{ let}
\end{aligned}
$$

$$
A_{\lambda,B}(x) := \mathrm{Im}\left(\frac{u^*du}{1+|u(x)|^2}\right).
$$

Note that $A_f(x)$ in (17.30) is the special case where the B is the diagonal $k\times k$ matrix with diagonal entries $x_1,\ldots,x_k\in\mathbb{R}^4\cong\mathbb{H}$. In order that $A_{\lambda,B}$ define a connection with self-dual curvature and removable singularities, we need to assume that the following two conditions hold (where $B^* = \overline{B}^T$):

$$\text{(17.31)}\qquad \begin{array}{l} \text{(I) } B \text{ is symmetric and } \lambda\lambda^* + BB^* \text{ is a real } k \times k \text{ matrix.} \\ \text{(II) For any } x \in \mathbb{H}, \text{ and } \xi \in \mathbb{H}^k, \text{ we have} \\ \quad ((B - xI_k)\,\xi = 0 \text{ and } \lambda^*\xi = 0) \Rightarrow \xi = 0. \end{array}$$

In the special case that B is diagonal then (I) is obviously met. If, in addition, we assume that all λ_i are nonzero and B has *distinct* diagonal entries, then (II) is met. Indeed, with these assumptions,

$$\begin{aligned} &(B - x)\,\xi = 0 \text{ and } \lambda^T\xi = 0 \\ &\Rightarrow \text{for each } i,\ (b_{ii} - x)\,\xi_i = 0 \text{ and } \lambda^T\xi = 0 \\ &\Rightarrow \text{for each } i, \xi_i = 0 \text{ or } x = b_{ii} \text{ and } \lambda^T\xi = 0 \\ &\Rightarrow \xi = 0, \text{ or there is at most one } i, \text{ with } \xi_i \neq 0 \text{ and } \lambda_i\xi_i = \lambda^T\xi = 0 \\ &\Rightarrow \xi = 0. \end{aligned}$$

THEOREM 17.10. *If (λ, B) satisfies Conditions (I) and (II), then $A_{\lambda,B}$ is an instanton (i.e., the curvature $F_{\lambda,B}$ of $A_{\lambda,B}$ is self-dual). Also, $A_{\lambda,B}$ lifts and extends to a connection on a principal* SU(2)*-bundle $\pi : P \to S^4$. Moreover, if $\lambda' = T\lambda q$ and $B' = TBT^{-1}$ for some unit $q \in \mathbb{H}$ (i.e., $q \in S^3$) and some $T \in$ SO(k), then (λ', B') also satisfies Conditions (I) and (II), and*

$$A_{\lambda',B'} = q^{-1}A_{\lambda,B}q,$$

so that $A_{\lambda',B'}$ is equivalent (via a gauge transformation) to $A_{\lambda,B}$.

PROOF. First we introduce the following notation

$$\begin{aligned} \sigma(x) &:= \sqrt{1 + |u(x)|^2} \\ U(x) &:= \frac{1}{\sigma(x)}\begin{bmatrix} 1 \\ u(x) \end{bmatrix} = \frac{1}{\sigma(x)}\begin{bmatrix} 1 \\ (B - xI_k)^{-1}\lambda \end{bmatrix}. \end{aligned}$$

Note that

$$\begin{aligned} U^*dU &= \sigma^{-1}\begin{bmatrix} 1 & u^* \end{bmatrix} d\left(\sigma^{-1}\begin{bmatrix} 1 \\ u \end{bmatrix}\right) \\ &= \sigma^{-1}\begin{bmatrix} 1 & u^* \end{bmatrix}\left(\sigma^{-1}\begin{bmatrix} 0 \\ du \end{bmatrix} + d\left(\sigma^{-1}\right)\begin{bmatrix} 1 \\ u \end{bmatrix}\right) \\ &= \sigma^{-2}u^*du + \sigma^{-1}d\left(\sigma^{-1}\right)\begin{bmatrix} 1 & u^* \end{bmatrix}\begin{bmatrix} 1 \\ u \end{bmatrix} \\ &= \sigma^{-2}u^*du + \sigma d\left(\sigma^{-1}\right). \end{aligned}$$

Since $\sigma d\left(\sigma^{-1}\right)$ is real, we then have

$$A_{\lambda,B}(x) := \operatorname{Im}\left(\frac{u^*du}{1 + |u(x)|^2}\right) = \operatorname{Im}\left(U^*dU\right).$$

Moreover,

$$\begin{aligned}F_{\lambda,B} &= dA_{\lambda,B} + A_{\lambda,B}\wedge A_{\lambda,B}\\ &= d\,\mathrm{Im}\,(U^*dU) + \mathrm{Im}\,(U^*dU)\wedge \mathrm{Im}\,(U^*dU)\\ &= \mathrm{Im}\,(d\,(U^*dU)) + \mathrm{Im}\,((U^*dU)\wedge (U^*dU))\,,\end{aligned}$$

since d commutes with the algebraic projection Im, and $\mathrm{Im}\,(U^*dU)\wedge \mathrm{Im}\,(U^*dU) = \mathrm{Im}(U^*dU)\wedge \mathrm{Im}(U^*dU)$ by the computation

$$\begin{aligned}&U^*dU\wedge U^*dU\\ &= (\mathrm{Re}\,(U^*dU) + \mathrm{Im}\,(U^*dU))\wedge (\mathrm{Re}\,(U^*dU) + \mathrm{Im}\,(U^*dU))\\ &= \mathrm{Im}\,(U^*dU)\wedge \mathrm{Im}\,(U^*dU) + \mathrm{Re}\,(U^*dU)\wedge \mathrm{Re}\,(U^*dU)\\ &+ \mathrm{Im}\,(U^*dU)\wedge \mathrm{Re}\,(U^*dU) + \mathrm{Re}\,(U^*dU)\wedge \mathrm{Im}U^*dU\\ &= \mathrm{Re}\,(U^*dU)\wedge \mathrm{Re}\,(U^*dU) + \mathrm{Im}\,(U^*dU)\wedge \mathrm{Im}\,(U^*dU)\,,\end{aligned}$$

where the cross terms cancel, since 1-forms anti-commute, while real and pure imaginary quaternions commute. Thus,

$$\begin{aligned}F_{\lambda,B} &= \mathrm{Im}\,(d\,(U^*dU) + (U^*dU)\wedge (U^*dU))\\ &= \mathrm{Im}\,(dU^*\wedge dU + (U^*dU)\wedge (U^*dU))\\ &= \mathrm{Im}\,(F_U)\,,\end{aligned}$$

where

$$F_U := dU^*\wedge dU + (U^*dU)\wedge (U^*dU)\,.$$

Let $M_{m,n}\,(\mathbb{H})$ denote the set of $m\times n$ matrices with quaternion entries. While $U^*U = 1$, we have

$$P := UU^* \in M_{k+1,k+1}\,(\mathbb{H})\,.$$

Note that $P^2 = UU^*UU^* = P$ and $P : \mathbb{H}^{k+1}\to \mathbb{H}^{k+1}$ is a projection onto

$$\mathit{span}\,(U) := \{Uq : q\in \mathbb{H}\}\,.$$

We have

$$F_U = U^*\,(dP\wedge dP)\,U$$

by the computation (where we use $d\,(U^*U) = d\,(1) = 0$)

$$\begin{aligned}U^*\,(dP\wedge dP)\,U &= U^*\,(d\,(UU^*)\wedge d\,(UU^*))\,U\\ &= U^*\,(((dU)\,U^* + U dU^*)\wedge ((dU)\,U^* + U dU^*))\,U\\ &= (U^*\,(dU)\,U^* + dU^*)\wedge (dU + U\,(dU^*)\,U)\\ &= dU^*\wedge dU + U^*dU\wedge U^*dU\\ &+ ((dU^*)\,U + U^*\,(dU))\wedge (dU^*)\,U\\ &= dU^*\wedge dU + U^*dU\wedge U^*dU + d\,(U^*U)\wedge (dU^*)\,U\\ &= F_U.\end{aligned}$$

Hence, the self-duality of F_U will follow, once the self-duality of $dP\wedge dP$ is shown. To demonstrate that $dP\wedge dP$ is self-dual, it turns out to be easier to work with the complementary projection $Q := I - P$. Note that

$$Q := I - P \Rightarrow dP\wedge dP = d\,(I-P)\wedge d\,(I-P) = dQ\wedge dQ.$$

In order to find an appropriate formula for Q, we first seek $v \in M_{k+1,k}(\mathbb{H})$, whose columns span a subspace, say $U^{\perp}$, of $M_{k+1}(\mathbb{H})$ which is orthogonal to U; i.e., $U^*v = 0$. Writing

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in M_{k+1,k}(\mathbb{H}),$$

where $v_1 \in M_{1,k}(\mathbb{H})$ and $v_2 \in M_{k,k}(\mathbb{H})$, we want

$$\begin{aligned} 0 = \sigma U^* v &= \begin{bmatrix} 1 & u^* \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 + u^* v_2 \\ &= v_1 + \left((B - xI_k)^{-1} \lambda \right)^* v_2 \\ &= v_1 + \lambda^* (B - xI_k)^{*-1} v_2. \end{aligned}$$

This is achieved by taking $v_2 = (B - xI_k)^*$ and $v_1 = -\lambda^*$. Thus, we take

$$v = \begin{bmatrix} -\lambda^* \\ (B - xI_k)^* \end{bmatrix} \in M_{k+1,k}(\mathbb{H}).$$

Note that Condition II says precisely that v has rank k. Then $v^*v = M_{k,k}(\mathbb{H})$ is invertible, and the orthogonal projection Q of $\mathbb{H}^{k+1}$ onto $U^{\perp}$ is given by the quaternionic version of the usual formula, namely

$$Q = v (v^*v)^{-1} v^*.$$

Indeed,

$$\begin{aligned} QU &= v (v^*v)^{-1} (v^*U) = 0 \ \text{ and} \\ Qv &= v (v^*v)^{-1} v^*v = v. \end{aligned}$$

Now,

$$\begin{aligned} U^* (dP \wedge dP) U &= U^* d (1 - Q) \wedge d (1 - Q) U \\ &= U^* dQ \wedge (dQ) U \end{aligned}$$

Since $U^*v = 0$, we have

$$\begin{aligned} U^* dQ &= U^* d \left(v (v^*v)^{-1} v^* \right) = U^* \left((dv) (v^*v)^{-1} v^* + v d \left((v^*v)^{-1} v^* \right) \right) \\ &= U^* (dv) (v^*v)^{-1} v^* + U^* v d \left((v^*v)^{-1} v^* \right) \\ &= U^* (dv) (v^*v)^{-1} v^*. \end{aligned}$$

Similarly, since $v^*U = (U^*v)^* = 0$, we have

$$(dQ) U = d \left(v (v^*v)^{-1} v^* \right) U = v (v^*v)^{-1} (dv^*) U.$$

Thus,

$$\begin{aligned} F_U &= U^* dQ \wedge (dQ) U = U^* (dv) (v^*v)^{-1} v^* \wedge v (v^*v)^{-1} (dv^*) U \\ &= U^* (dv) (v^*v)^{-1} \wedge (dv^*) U. \end{aligned}$$

Note that

$$\begin{aligned} dv &= d \begin{bmatrix} -\lambda^* \\ (B - xI_k)^* \end{bmatrix} = \begin{bmatrix} 0 \\ -(d\overline{x}) I_k \end{bmatrix} = \begin{bmatrix} 0 \\ -I_k \end{bmatrix} d\overline{x} \\ dv^* &= d [-\lambda, B - xI_k] = dx [0, -I_k]. \end{aligned}$$

Hence,

$$\begin{aligned}F_U &= U^* dQ \wedge (dQ)\, U = U^*\,(dv)\,(v^*v)^{-1} \wedge (dv^*)\, U\\&= U^* \begin{bmatrix} 0 \\ -I_k \end{bmatrix} d\overline{x}\,(v^*v)^{-1} \wedge dx\,[0, -I_k]\, U\\&= \sigma^{-1} \begin{bmatrix} 1 & u^* \end{bmatrix} \begin{bmatrix} 0 \\ -I_k \end{bmatrix} d\overline{x}\,(v^*v)^{-1} \wedge dx\,[0, -I_k]\,\sigma^{-1} \begin{bmatrix} 1 \\ u \end{bmatrix}\\&= \sigma^{-2} u^*\,(d\overline{x})\,(v^*v)^{-1} \wedge (dx)\, u\end{aligned}$$

By (17.13), $d\overline{x} \wedge dx$ is a self-dual, quaternion-valued form. The intervening factor $(v^*v)^{-1}$ in $(d\overline{x})\,(v^*v)^{-1} \wedge (dx)$ prevents us from asserting that F_U is self-dual. On the other hand, if $(v^*v)^{-1} \in M_{k,k}(\mathbb{H})$ has *real* entries, then $(v^*v)^{-1}$ commutes with dx, and

$$F_U = \sigma^{-2} u^* \left((v^*v)^{-1}\, d\overline{x} \wedge dx\right) u$$

is indeed self-dual. It is precisely Condition (I) which guarantees that v^*v (and hence $(v^*v)^{-1}$) is real. To show this, first note that

$$\begin{aligned}v^*v &= \begin{bmatrix} -\lambda & B - xI_k \end{bmatrix} \begin{bmatrix} -\lambda^* \\ (B - xI_k)^* \end{bmatrix}\\&= \lambda\lambda^* + (B - xI_k)\,(B - xI_k)^*\\&= \lambda\lambda^* + BB^* - (B\overline{x} + xB^*) + |x|^2\, I_k\end{aligned}$$

This is real for all $x \in \mathbb{H} \Leftrightarrow \lambda\lambda^* + BB^*$ is real and $B\overline{x} + xB^*$ is real for all $x \in \mathbb{H}$. We show that $B\overline{x}+xB^*$ is real for all $x \in \mathbb{H} \Leftrightarrow B$ symmetric. We have B symmetric $\Leftrightarrow B^T = B \Leftrightarrow B^* = \overline{B}$. Assume that $B\overline{x} + xB^*$ is real for all $x \in \mathbb{H}$. Then

$$xB^* = \overline{(B\overline{x})} = x\overline{B},$$

and in particular for $x = 1$, $B^* = \overline{B}$ (i.e., $B^T = B$). Conversely, for B symmetric, we have $B^* = \overline{B}$ and so for all $x \in \mathbb{H}$,

$$B\overline{x} + xB^* = B\overline{x} + x\overline{B} = \overline{\left(x\overline{B} + B\overline{x}\right)} = \overline{(B\overline{x} + xB^*)},$$

namely $B\overline{x} + xB^*$ is real.

It is easy to check that F_U is pure imaginary. Thus,

$$F_{\lambda,B} = \operatorname{Im}(F_U) = F_U = \sigma^{-2} u^* \left((v^*v)^{-1}\, d\overline{x} \wedge dx\right) u.$$

It is also easy to verify that as $|x| \to \infty$

$$|F_{\lambda,B}|^2 = \mathrm{O}\left(|x|^{-8}\right),$$

so that $|F_{\lambda,B}|^2 \in L^2\left(\mathbb{R}^4\right)$, and Uhlenbeck's Theorem then implies that $A_{\lambda,B}$ lifts and extends to a connection on a principal SU(2)-bundle $\pi : P \to S^4$.

The verification of the fact that (λ', B') also satisfies Conditions (I) and (II) is routine. Note that B' is symmetric, since

$$\begin{aligned}B'^T &= \left(RBR^{-1}\right)^T = R^{-1T} B^T R^T\\&= RB^T R^{-1} = B'.\end{aligned}$$

Moreover $B'B'^{*}+\lambda'\lambda'^{*}$ is real, since

$$\begin{aligned} B'B'^{*}+\lambda'\lambda'^{*} &= RBR^{-1}\left(RBR^{-1}\right)^{*}+(R\lambda q)(R\lambda q)^{*} \\ &= RBR^{-1}\left(R^{-1*}B^{*}R^{*}\right)+(R\lambda q)(q^{*}\lambda^{*}R^{*}) \\ &= RBR^{-1}\left(RB^{*}R^{T}\right)+R\lambda\lambda^{*}R^{T} \\ &= RBB^{*}R^{T}+R\lambda\lambda^{*}R^{T} \\ &= R\left(BB^{*}+\lambda\lambda^{*}\right)R^{T}. \end{aligned}$$

Finally,

$$\begin{aligned} &(B'-xI_k)\,\xi=0 \text{ and } \lambda'^{*}\xi=0 \\ &\qquad\Rightarrow \left(RBR^{-1}-xI_k\right)\xi=0 \text{ and } (R\lambda q)^{*}\,\xi=0 \\ &\qquad\Rightarrow R\left(B-xI_k\right)R^{-1}\xi=0 \text{ and } q^{*}\lambda^{*}R^{T}\xi=0 \\ &\qquad\Rightarrow (B-xI_k)\left(R^{-1}\xi\right)=0 \text{ and } \lambda^{*}R^{-1}\xi=0 \\ &\qquad\Rightarrow R^{-1}\xi=0\Rightarrow \xi=0. \end{aligned}$$

Finally note that for $u_{\lambda,B}(x):=(B-xI_k)^{-1}\lambda$, we have

$$\begin{aligned} u_{\lambda',B'}(x) &= (B'-xI_k)^{-1}\lambda'=\left(TBT^{-1}-xI_k\right)^{-1}T\lambda q \\ &= \left(T\left(B-xI_k\right)T^{-1}\right)^{-1}T\lambda q \\ &= (B-xI_k)^{-1}\lambda q \\ &= u_{\lambda,B}(x)\,q, \end{aligned}$$

and so

$$\begin{aligned} A_{\lambda',B'}(x) &= \operatorname{Im}\left(\frac{(uq)^{*}\,d\,(uq)}{1+|uq(x)|^{2}}\right)=\operatorname{Im}\left(\frac{\overline{q}u^{*}\,(du)\,q}{1+|u(x)|^{2}}\right) \\ &= \overline{q}\operatorname{Im}\left(\frac{u^{*}\,(du)}{1+|u(x)|^{2}}\right)q=\overline{q}A_{\lambda,B}(x)\,q. \end{aligned}$$

□

REMARK 17.11. The principal $\mathrm{SU}(2)$-bundle $\pi : P \to S^4$ in Theorem 17.10 has $c_2(E')[S^4]=-k$, for k in (17.31) and $E'=P\times_{\mathrm{SU}(2)}\mathbb{C}^2$. We have shown this for the special case where the B is the diagonal $k\times k$ matrix with distinct diagonal entries. If the set of (λ,B) satisfying Conditions (I) and (II) is connected, then $c_2(E')[S^4]=-k$ for all such (λ,B) by continuity. A much more difficult task would be to show that every instanton arises in this fashion, as stated in Theorem 17.12 below. Using methods from algebraic geometry and twistor theory, this was accomplished in [**ADHM**]. An excellent, expanded presentation is found in [**?**], in which Theorem 17.12 appears on p. 26. The reader will notice some differences between our presentation and Atiyah's, since Atiyah writes his quaternions as $x=x_1+x_2\mathbf{i}+x_3\mathbf{j}+x_4\mathbf{k}$, whereas we have written $x=x_4+x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k}$. This causes a reversal of orientation. It also (for our own good) made us feel obligated to go through the entire construction in the uncommon detail found in the proof of Theorem 17.10.

THEOREM 17.12. *Every instanton on* $\mathbb{R}^4$*, which extends to an instanton on a principal* SU (2)*-bundle* $\pi : P \to S^4$ *with* $c_2\left(P \times \mathbb{C}^2\right)[S^4] = -k$*, is of the form* $A_{\lambda,B}$ *for* (λ, B) *satisfying Conditions (I) and (II). Moreover,* $A_{\lambda,B}$ *is equivalent (via a gauge transformation) to* $A_{\lambda',B'} \Leftrightarrow \lambda' = T\lambda q$ *and* $B' = TBT^{-1}$ *for some unit* $q \in \mathbb{H}$ *(i.e.,* $q \in S^3$*) and some* $T \in \mathrm{SO}(k)$*.*

Given Theorem 17.12, we can count the number of independent parameters in the space of instantons modulo gauge transformations. For a fixed k, the real dimension of the space of (λ, B) for which B is symmetric, is

$$4k + 4 \cdot \tfrac{1}{2}k\,(k+1) = 2k^2 + 6k.$$

Since $(B^*B + \lambda^*\lambda)^* = B^*B + \lambda^*\lambda$, we have $(B^*B + \lambda^*\lambda)^T = \overline{B^*B + \lambda^*\lambda}$, and so

$$\left(\mathrm{Im}\,(B^*B + \lambda^*\lambda)\right)^T = \mathrm{Im}\left(\overline{B^*B + \lambda^*\lambda}\right) = -\,\mathrm{Im}\,(B^*B + \lambda^*\lambda)\,.$$

Thus, $\mathrm{Im}\,(B^*B + \lambda^*\lambda)$ is skew symmetric and setting it equal to 0 eliminates at most $3\cdot\frac{1}{2}k\,(k-1)$ dimensions. Quotienting by the group $S^3 \times \mathrm{SO}(k)$ of gauge equivalences among the $A_{\lambda,B}$ further reduces the dimension by at most $3 + \frac{1}{2}k\,(k-1)$. Thus, the dimension of the space of $A_{\lambda,B}$ satisfying conditions (I) and (II) is at least

$$2k^2 + 6k - 3\cdot\tfrac{1}{2}k\,(k-1) - \left(3 + \tfrac{1}{2}k\,(k-1)\right) = 8k - 3.$$

In the next section, we apply the index theorem to prove that the space of instantons modulo gauge transformations is a manifold of dimension $8k - 3$. Thus, the naive dimension count is actually correct. In other words, the conditions imposed are actually all independent, in case there was any doubt.

3. Linearization of the Manifold of Moduli of Self-dual Connections

In Section 17.1 we observed that $YM : \mathcal{C}(P) \to \mathbb{R}^+$ is invariant under the action on $\mathcal{C}(P)$ of the group $\mathrm{GA}(P)$ of gauge transformations, since for $F \in \mathrm{GA}(P)$, $F \cdot \omega = \left(F^{-1}\right)^* \omega$ yields

$$\begin{aligned}\Omega^{F\cdot\omega} &= d\,(F\cdot\omega) + (F\cdot\omega)\wedge(F\cdot\omega) = \left(F^{-1}\right)^*(d\omega + \omega\wedge\omega)\\ &= \left(F^{-1}\right)^*\Omega^\omega = F\cdot\Omega^\omega,\end{aligned}$$

and then $|F\cdot\Omega^\omega|^2 = |\Omega^\omega|^2$ by Corollary 16.30. In particular, the set of critical points of YM is preserved by this action, as well as the (possibly empty) set of (anti-) self-dual connections (absolute minima of YM, if they exist). Thus, the quotient space $\mathcal{M} := \mathcal{C}(P)\,/\,\mathrm{GA}(P)$ is a natural object for study. In particular, one would like to know the extent to which may regard $\mathcal{M}$ as a manifold, say modeled on some infinite dimensional Fréchet space. Also, if $\mathcal{C}(P)^+$ is the space of self-dual connections (i.e., with self-dual curvature), it would be of interest to compute the dimension of "the space of moduli of self-dual connections" $\mathcal{M}^+ := \mathcal{C}(P)^+\,/\,\mathrm{GA}(P)$, at least where it is a submanifold of $\mathcal{M}$. In the heuristic discussion which follows, we will talk rather loosely about infinite dimensional manifolds, and their tangent spaces and normal spaces. However, this discussion will motivate a precise theorem with a precise proof. In Section 17.4, we will provide the framework within which it makes sense to speak of $\mathcal{M}^+$ as being a submanifold of $\mathcal{M}$.

Heuristic discussion: For $\omega \in \mathcal{C}(P)$, the orbit of the action of $\mathrm{GA}(P)$ on $\mathcal{C}(P)$ is

$$\mathrm{GA}(P) \cdot \omega := \{F \cdot \omega : F \in \mathrm{GA}(P)\} = \{\Phi(f) \cdot \omega : f \in C(P,G)\}$$

in the notation of Proposition 16.27, p. 382. Formally, the Lie algebra of $C(P,G)$ is $C(P,\mathfrak{g})$ and the infinitesimal action of $C(P,\mathfrak{g})$ on $\mathcal{C}(P)$ is given (see Proposition 16.32, p. 384), for $s \in C(P,\mathfrak{g})$ and $\omega \in \mathcal{C}(P)$, by

$$s \cdot \omega = \frac{d}{dt}\left(\Phi(\mathrm{Exp}(ts)) \cdot \omega\right)\Big|_{t=0} = -\left(ds + [\omega, s]\right) = -D^{\omega} s \in \overline{\Omega}^{1}(P,\mathfrak{g}).$$

Thus the "formal tangent space" at $\omega \in \mathcal{C}(P)$ of the orbit $\mathrm{GA}(P) \cdot \omega$ is

$$T_{\omega}(\mathrm{GA}(P) \cdot \omega) := \left\{D^{\omega} s \in \overline{\Omega}^{1}(P,\mathfrak{g}) : s \in C(P,\mathfrak{g})\right\}.$$

Note that $T_{\omega}(\mathrm{GA}(P) \cdot \omega) \subseteq \overline{\Omega}^{1}(P,\mathfrak{g})$ which is the vector space on which the affine space $\mathcal{C}(P)$ is modeled (see Remark 16.11, p. 371). For $\tau \in \overline{\Omega}^{1}(P,\mathfrak{g})$ and $s \in C(P,\mathfrak{g})$, we have

$$(\tau, D^{\omega} s) = (\delta^{\omega} \tau, s),$$

where δ^{ω} is covariant codifferential, the formal adjoint of D^{ω} (see Proposition 16.18, p. 377). Thus, the "formal normal space" to $\mathrm{GA}(P) \cdot \omega$ at ω is

$$N_{\omega}(\mathrm{GA}(P) \cdot \omega) := \left\{\tau \in \overline{\Omega}^{1}(P,\mathfrak{g}) : \delta^{\omega} \tau = 0\right\}. \tag{17.32}$$

The "formal slice of the action" is

$$S_{\omega} := \{\omega + \tau : \tau \in N_{\omega}(\mathrm{GA}(P) \cdot \omega)\} = \{\omega + \tau : \delta^{\omega} \tau = 0\}. \tag{17.33}$$

Intuitively, we expect that every connection ω' in a "suitably small neighborhood" of ω will be gauge-equivalent to a connection in S_{ω}; i.e., there is some $F \in \mathrm{GA}(P)$, such that $F \cdot \omega' \in S_{\omega}$. ■

Recall from Section 16.8, that the isotropy subgroup of ω is

$$I_{\omega} := \{F \in \mathrm{GA}(P) : F \cdot \omega = \omega\}.$$

Note that I_{ω} leaves S setwise fixed, since, for $\tau \in N_{\omega}(\mathrm{GA}(P) \cdot \omega)$,

$$F \in I_{\omega} \Rightarrow F \cdot (\omega + \tau) = \omega + F \cdot \tau \text{ and}$$
$$\delta^{\omega}(F \cdot \tau) = F \cdot \delta^{F \cdot \omega} \tau = F \cdot \delta^{\omega} \tau = 0,$$

where we have used Corollary 16.31 on p. 384 Unless I_{ω} acts trivially on S_{ω}, there will be gauge-equivalent connections in S_{ω}, and so S_{ω} will not parametrize $\mathcal{M} := \mathcal{C}(P) / \mathrm{GA}(P)$ in a 1-1 fashion even locally about ω. Note that if $g \in Z(G) :=$ the center of G, then $R_g : P \to P$ is in $\mathrm{GA}(P)$ and for any $\omega' \in \mathcal{C}(P)$, we have $R_g^{*} \omega' = \omega'$ by (16.27), p. 382. A condition which implies that I_{ω} consists of only these "central" gauge transformations is that ω be irreducible; this is immediate from Proposition 16.61, p. 426, which asserts that I_{ω} is isomorphic to the centralizer $Z(G_0)$ (in G) of the holonomy group G_0. The center of a semi-simple, compact G must be finite. In this case, for irreducible ω, there can be no nontrivial one-parameter subgroup in I_{ω}.

PROPOSITION 17.13. *Let $\alpha \in C(P,\mathfrak{g})$. Then $\Phi(\operatorname{Exp}(t\alpha)) \in I_\omega$ for all real t, if and only if $D^\omega \alpha = 0$. Thus, the following are equivalent (where G_0 is the holonomy group of ω with arbitrary reference point $p \in P$) :*

(i) $\operatorname{Ker}\left(D^\omega : C(P,\mathfrak{g}) \to \overline{\Omega}^1(P,\mathfrak{g})\right) = \{0\}$
(ii) $Z(G_0) = \{g \in G : gg_0 = g_0 g \text{ for all } g_0 \in G_0\}$ *is discrete.*

PROOF. Using (16.32) on p. 384, for any real t_0, we have

$$\begin{aligned}
&\tfrac{d}{dt}\left(\Phi(\operatorname{Exp}(t\alpha))\cdot\omega\right)\Big|_{t=t_0} = \tfrac{d}{dt}\left(\Phi(\operatorname{Exp}((t_0+t)\alpha))\cdot\omega\right)\Big|_{t=0}\\
&= \tfrac{d}{dt}\left(\Phi(\operatorname{Exp}(t_0\alpha)\operatorname{Exp}(t\alpha))\cdot\omega\right)\Big|_{t=0}\\
&= \tfrac{d}{dt}\left(\left(\Phi(\operatorname{Exp}(t_0\alpha))\cdot\Phi(\operatorname{Exp}(t\alpha))\right)\cdot\omega\right)\Big|_{t=0}\\
&= \tfrac{d}{dt}\left(\Phi(\operatorname{Exp}(t_0\alpha))\cdot\left(\Phi(\operatorname{Exp}(t\alpha))\cdot\omega\right)\right)\Big|_{t=0}\\
&= \tfrac{d}{dt}\left(\Phi\left(\operatorname{Exp}(t_0\alpha)^{-1}\right)^*\left(\Phi(\operatorname{Exp}(t\alpha))\cdot\omega\right)\right)\Big|_{t=0}\\
&= \Phi(\operatorname{Exp}(-t_0\alpha))^*\left(\tfrac{d}{dt}\Phi(\operatorname{Exp}(t\alpha))\cdot\omega\Big|_{t=0}\right)\\
&= \Phi(\operatorname{Exp}(-t_0\alpha))^*\left(-D^\omega\alpha\right).
\end{aligned}$$

Since $\Phi(\operatorname{Exp}(0\cdot\alpha))\cdot\omega = \omega$, $\Phi(\operatorname{Exp}(t\alpha))\cdot\omega = \omega$ for all $t \Leftrightarrow D^\omega\alpha = 0$. □

DEFINITION 17.14. Let $\pi : P \to M$ principal G-bundle where G is a compact, semi-simple Lie group. A connection 1-form ω on P is called **weakly irreducible** if (i), or equivalently (ii), holds in Proposition 17.13.

If $\omega \in \mathcal{C}(P)^+$ is a weakly-irreducible, self-dual connection, then

$$\omega' \in \mathcal{C}(P)^+ \cap S_\omega \Leftrightarrow *\Omega^{\omega'} = \Omega^{\omega'} \text{ and } \delta^\omega(\omega'-\omega) = 0.$$

Writing $\tau = \omega' - \omega$ (or $\omega' = \omega + \tau$), we have

$$\begin{aligned}
\Omega^{\omega'} &= d\omega' + \tfrac{1}{2}[\omega',\omega'] = d\omega + \tfrac{1}{2}[\omega,\omega] + d\tau + \mathfrak{ad}(\omega)\wedge\tau + \tfrac{1}{2}[\tau,\tau]\\
&= \Omega^\omega + D^\omega\tau + \tfrac{1}{2}[\tau,\tau].
\end{aligned}$$

Thus, $*\Omega^{\omega'} = \Omega^\omega \Leftrightarrow (1-*)\left(D^\omega\tau + \frac{1}{2}[\tau,\tau]\right) = 0$. Hence,

$$\omega' \in \mathcal{C}(P)^+ \cap S_\omega \Leftrightarrow \begin{cases} \text{(A) } \delta^\omega\tau = 0 \text{ and}\\ \text{(B) } (1-*)\left(D^\omega\tau + \frac{1}{2}[\tau,\tau]\right) = 0.\end{cases} \tag{17.34}$$

Condition (A) is linear and the linearization of the quadratic Condition (B) is $(1-*)D^\omega\tau = 0$. Hence, the "formal tangent space" of $\mathcal{C}(P)^+ \cap S_\omega$ at ω is

$$T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right) := \left\{\tau \in \overline{\Omega}^1(P,\mathfrak{g}) : \delta^\omega\tau = 0 \text{ and } (1-*)D^\omega\tau = 0\right\} \tag{17.35}$$

The preceding motivates the following key result which specifies the dimension of the hypothetical manifold $\mathcal{C}(P)^+\cap S_\omega$ near a weakly-irreducible self-dual connection ω for a suitable principal G-bundle $\pi : P \to M$. The original proof in [**AHS2**] makes use of the Dirac operator on spinor fields and the $\widehat{A}$ genus, along with the Index Theorem. The proof here is spinor-free, and is essentially Hodge-theoretic. The Index Theorem is used mainly to handle the "twist" in the bundle $P \times_G \mathfrak{g}_{\mathbb{C}}$.

THEOREM 17.15. *Let G be a compact, semi-simple Lie group and let $\omega \in \mathcal{C}(P)^+$ be a weakly-irreducible, self-dual connection for a principal G-bundle $\pi : P \to M$ over a self-dual, compact, connected, oriented Riemannian 4-manifold M with non-negative scalar curvature $S \neq 0$. For $T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right)$ defined by (17.35), we have*
(17.36)

$$\dim T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right) = 2ch\left(P \times_G \mathfrak{g}_{\mathbb{C}}\right)[M] - \tfrac{1}{2}\dim(G)\left(\chi(M) - \operatorname{sig}(M)\right),$$

where $P \times_G \mathfrak{g}_{\mathbb{C}} \to M$ is the vector bundle associated to $\pi : P \to M$ via the complex adjoint representation $ad_{\mathbb{C}} : G \to \mathrm{GL}(\mathfrak{g}_{\mathbb{C}})$, $\mathfrak{g}_{\mathbb{C}} := \mathbb{C} \otimes \mathfrak{g}$, $\chi(M)$ is the Euler characteristic of M, and $\operatorname{sig}(M)$ is the signature of M.

PROOF. Let $E := P \times_G \mathfrak{g}_{\mathbb{C}}$, and let $\Omega^k(E)$ denote the space of E-valued k-forms on M. By (16.11), p. 373, we may (and often do) make the identification

$$\Omega^k(E) \cong \overline{\Omega}^k(P, \mathfrak{g}_{\mathbb{C}}).$$

Let $\Omega^2_-(E)$ denote the space of anti-self-dual E-valued 2-forms on M. Define $\mathcal{T} : \Omega^1(E) \to \Omega^0(E) \oplus \Omega^2_-(E)$, for $\tau \in \Omega^k(E)$, by

$$\mathcal{T}(\tau) := \left(\delta^\omega \tau, \tfrac{1}{2}(1 - *) D^\omega \tau\right). \tag{17.37}$$

Note that $\operatorname{Ker}\mathcal{T} = T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right)$. Of course, $\Omega^0(E) \oplus \Omega^2_-(E)$ can be regarded as the space of sections of the single bundle $E \otimes \left(\Lambda^0(T^*M) \oplus \Lambda^2_-(T^*M)\right)$, and so $\mathcal{T}$ is a differential operator (of order 1) mapping $C^\infty\left(E \otimes \Lambda^1(T^*M)\right)$ to $C^\infty\left(E \otimes \left(\Lambda^0(T^*M) \oplus \Lambda^2_-(T^*M)\right)\right)$. Since

$$\dim \Lambda^0(T^*M) = 4 = 1 + 3 = \dim\left(\Lambda^0(T^*M) \oplus \Lambda^2_-(T^*M)\right),$$

we know that $\mathcal{T}$ will be elliptic if its symbol $\sigma(\mathcal{T})$ is injective. From the local formulas for D^ω and δ^ω (see (16.22) and (16.24), p. 378), it follows that $\sigma(\mathcal{T}) = \mathrm{Id}_E \otimes \sigma(\mathcal{T}_0)$, where $\sigma(\mathcal{T}_0)$ is the symbol of the "untwisted" (i.e., coefficients no longer in E) operator $\mathcal{T}_0 : \Omega^1(M) \to \Omega^0(M) \oplus \Omega^2_-(M)$ given by

$$\mathcal{T}_0(\gamma) := \left(\delta\gamma, \tfrac{1}{2}(1 - *) d\gamma\right).$$

Let $S(M) \subseteq T^*M$ be the unit cosphere bundle. Then

$$\sigma(\mathcal{T}_0) : S(M) \to \operatorname{Hom}\left(T^*M, \Lambda^0(T^*M) \oplus \Lambda^2_-(T^*M)\right)$$

is given, for $\xi \in S(M)$ and $\eta \in T^*_xM$, by

$$\sigma_\xi(\eta) = \left(-\langle \xi, \eta\rangle, \tfrac{1}{2}(1 - *)\,\xi \wedge \eta\right).$$

Now, $\sigma(\mathcal{T}_0)_\xi(\eta) = 0 \Rightarrow \langle \xi, \eta\rangle = 0$ and $*(\xi \wedge \eta) = \xi \wedge \eta$. If $*(\xi \wedge \eta) = \xi \wedge \eta$ and ν_x is the volume element at x, then

$$|\xi \wedge \eta|^2 \nu_x = \xi \wedge \eta \wedge *(\xi \wedge \eta) = \xi \wedge \eta \wedge \xi \wedge \eta = 0,$$

and so $\eta = c\xi$ for some constant c. However, then $0 = \langle \xi, \eta\rangle = c\langle \xi, \xi\rangle = c$, in which case $\eta = 0$. Thus, $\sigma(\mathcal{T}_0)_\xi$ is injective for all $\xi \in S(M)$, and $\mathcal{T}$ is elliptic.

Since $\operatorname{Ker}\mathcal{T} = T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right)$, our goal is to compute $\dim\operatorname{Ker}\mathcal{T}$. We show that $\operatorname{Ker}\mathcal{T}^* = 0$, and then go on to compute

$$\dim\operatorname{Ker}\mathcal{T} = \dim\operatorname{Ker}\mathcal{T} - \dim\operatorname{Ker}\mathcal{T}^* = \operatorname{index}\mathcal{T}$$

via the index formula. To obtain a formula for $\mathcal{T}^*$, we compute (where $\alpha \in \Omega^0(E)$, $\beta \in \Omega^2_-(E)$, and $\tau \in \Omega^1(E)$)

$$\begin{aligned}(\mathcal{T}^*(\alpha,\beta),\tau) &= ((\alpha,\beta),\mathcal{T}(\tau)) = \left((\alpha,\beta),\left(\delta^\omega\tau,\tfrac{1}{2}(1-*)D^\omega\tau\right)\right)\\ &= (\alpha,\delta^\omega\tau) + \left(\beta,\tfrac{1}{2}(1-*)D^\omega\tau\right)\\ &= (D^\omega\alpha,\tau) + (\beta,D^\omega\tau) = (D^\omega\alpha+\delta^\omega\beta,\tau),\end{aligned} \tag{17.38}$$

where we note that $\frac{1}{2}(1-*)$ is projection onto $\Omega^2_-(E)$ in which β resides, so that $\left(\beta,\frac{1}{2}(1-*)D^\omega\tau\right) = (\beta,D^\omega\tau)$. Thus,

$$\mathcal{T}^*(\alpha,\beta) = D^\omega\alpha + \delta^\omega\beta.$$

We now prove that $\operatorname{Ker}\mathcal{T}^* = 0$. Suppose that $\mathcal{T}^*(\alpha,\beta)=0$, so that $D^\omega\alpha = -\delta^\omega\beta$. Since $\omega \in \mathcal{C}(P)^+$ (i.e., $\Omega^\omega \in \Omega^2_+(E)$) and $\beta \in \Omega^2_-(E) \subseteq \Omega^2_+(E)^\perp$, we have

$$\|D^\omega\alpha\|^2 = (-\delta^\omega\beta, D^\omega\alpha) = -(\beta, D^\omega D^\omega\alpha) = -(\beta,[\Omega^\omega,\alpha]) = 0.$$

Thus, $\mathcal{T}^*(\alpha,\beta) = 0 \Rightarrow -\delta^\omega\beta = D^\omega\alpha = 0$. Since we have assumed that ω is weakly-irreducible, $D^\omega\alpha = 0 \Rightarrow \alpha = 0$. As $\beta \in \Omega^2_-(E)$,

$$\begin{aligned}&\delta^\omega\beta = 0 \Rightarrow D^\omega\beta = -D^\omega*\beta = *(*D^\omega*\beta) = -*\delta^\omega\beta = 0\\ &\Rightarrow \Delta^\omega\beta = \delta^\omega D^\omega\beta + D^\omega\delta^\omega\beta = 0 \Rightarrow \beta = 0,\end{aligned}$$

by Proposition 16.57, p. 414 which applies under our assumptions on ω and M. Hence, $\operatorname{Ker}\mathcal{T}^* = 0$, and $\dim\operatorname{Ker}\mathcal{T} = \operatorname{index}\mathcal{T}$.

To compute $\operatorname{index}\mathcal{T}$, we use the formula of (III.4.A)

$$\begin{aligned}\operatorname{index}\mathcal{T} &= \left(\phi^{-1}ch(\sigma(\mathcal{T})) \smile \tau(TM\otimes\mathbb{C})\right)[M]\\ &= \left(\phi^{-1}ch(\operatorname{Id}_E\otimes\sigma(\mathcal{T}_0)) \smile \tau(TM\otimes\mathbb{C})\right)[M]\\ &= \left(ch(E)\smile\phi^{-1}(ch(\sigma(\mathcal{T}_0)))\smile\tau(TM\otimes\mathbb{C})\right)[M],\end{aligned}$$

where we recall that $\sigma(\mathcal{T}) = \operatorname{Id}_E\otimes\sigma(\mathcal{T}_0)$, ϕ is the Thom isomorphism, and $\tau(TM\otimes\mathbb{C})$ is the Todd class. The unitary frame bundle $U(E)$ of $E = P\times_G\mathfrak{g}_\mathbb{C}$ is reducible to an $\mathrm{SU}(N)$ bundle ($N = \dim\mathfrak{g}$), since the orthogonal (relative to K) representation $ad: G\to\mathrm{SO}(\mathfrak{g})$ extends to $ad_\mathbb{C}: G\to\mathrm{SU}(\mathfrak{g}_\mathbb{C})$ which serves to define $E = P\times_G\mathfrak{g}_\mathbb{C}$. By the Remark 16.59, p. 416, we then have $ch_1(E) = c_1(E) = 0$, and then $ch(E) = \dim\mathfrak{g} + ch(E)_2$. Note that

$$\left(\phi^{-1}(ch(\sigma(T_0)))\smile\tau(TM\otimes\mathbb{C})\right)_2[M] = \operatorname{index}\mathcal{T}_0.$$

We will save the verification that

$$\left(\phi^{-1}(ch(\sigma(\mathcal{T}_0)))\smile\tau(TM\otimes\mathbb{C})\right)_0 = 2$$

for last, but assuming that this is correct, we have

$$\operatorname{index}\mathcal{T} = 2ch_2(E)[M] + (\dim\mathfrak{g})\operatorname{index}\mathcal{T}_0. \tag{17.39}$$

We now show that

$$\operatorname{index}\mathcal{T}_0 = -\tfrac{1}{2}(\chi(M) - \operatorname{sig}(M)). \tag{17.40}$$

By Hodge theory (see III.4.D) the cohomology space $H^k(M,\mathbb{R})$ can be identified with the space of harmonic k-forms:

$$\beta\in H^k(M,\mathbb{R}) \Leftrightarrow \Delta\beta = 0 \Leftrightarrow d\beta = 0 \text{ and } \delta\beta = 0.$$

Let $b_k := \dim H^k(M,\mathbb{R})$ denote the k-th Betti number. It is easy to check that $*$ commutes with Δ. Thus $H^2(M,\mathbb{R})$ is preserved by $*$ and splits into the ±1

eigenspaces of $*$, say $H^2_\pm(M,\mathbb{R})$ of dimensions $b_2^\pm$, with $b_2 = b_2^+ + b_2^-$. For $\alpha \in H^2_+(M,\mathbb{R})$ and $\beta \in H^2_-(M,\mathbb{R})$, we have $\alpha\wedge\alpha = \alpha\wedge *\alpha = |\alpha|^2\,\nu$, $\beta\wedge\beta = -\beta\wedge *\beta = -\,|\beta|^2\,\nu$, and

$$\begin{aligned}\alpha\wedge\beta &= -\alpha\wedge *\beta = -\left\langle \alpha,\beta\right\rangle \nu = -\left\langle \beta,\alpha\right\rangle \nu\\ &= -\beta\wedge *\alpha = -\beta\wedge\alpha \Rightarrow \alpha\wedge\beta = 0,\end{aligned}$$

since 2-forms commute. Thus, $\operatorname{sig}(M) = b_2^+ - b_2^-$.Since $* : H^1(M,\mathbb{R}) \cong H^3(M,\mathbb{R})$, we have $b_3 = b_1$, and

$$\begin{aligned}&-\tfrac{1}{2}\left(\chi(M) - \operatorname{sig}(M)\right)\\ &= -\tfrac{1}{2}\left(b_0 - b_1 + b_2 - b_3 + b_4 - \left(b_2^+ - b_2^-\right)\right)\\ &= -\tfrac{1}{2}\left(2 - 2b_1 + 2b_2^-\right) = b_1 - \left(1 + b_2^-\right).\end{aligned}$$

Thus, it suffices to prove that $\dim\operatorname{Ker}\mathcal{T}_0 = b_1$and $\dim\operatorname{Ker}\mathcal{T}_0^* = 1 + b_2^-$. We claim $\operatorname{Ker}\mathcal{T}_0 = H^1(M,\mathbb{R})$. Since $\mathcal{T}_0(\gamma) = \left(\delta\gamma, \tfrac{1}{2}(1-*)\,d\gamma\right)$, it is clear that $H^1(M,\mathbb{R}) \subseteq \operatorname{Ker}\mathcal{T}_0$. If $\gamma \in \operatorname{Ker}\mathcal{T}_0$, then $\delta\gamma = 0$ and $(1-*)\,d\gamma = 0$. Hence,

$$\begin{aligned}0 = \delta\left((1-*)\,d\gamma\right) &= \delta d\gamma - \delta * d\gamma = \delta d\gamma + *d**d\gamma\\ &= \delta d\gamma + *d^2\gamma = \delta d\gamma, \text{ and}\\ \delta d\gamma = 0 \Rightarrow 0 &= (\delta d\gamma,\gamma) = (d\gamma, d\gamma) \Rightarrow d\gamma = 0.\end{aligned}$$

Thus, $\operatorname{Ker}\mathcal{T}_0 = H^1(M,\mathbb{R})$, and $\dim\operatorname{Ker}\mathcal{T}_0 = b_1$. We now prove that $\operatorname{Ker}\mathcal{T}_0^* = H^0(M,\mathbb{R}) \oplus H^2_-(M,\mathbb{R})$. By a computation strictly analogous to (17.38), $\mathcal{T}_0^*(\alpha,\beta) = d\alpha + \delta\beta$. Thus, clearly $H^0(M,\mathbb{R}) \oplus H^2_-(M,\mathbb{R}) \subseteq \operatorname{Ker}\mathcal{T}_0^*$. If $(\alpha,\beta) \in \operatorname{Ker}\mathcal{T}_0^*$, then $d\alpha + \delta\beta = 0$ and

$$0 = \delta\left(d\alpha + \delta\beta\right) = \delta d\alpha + \delta^2\beta = \delta d\alpha.$$

Hence, $0 = (\delta d\alpha, \alpha) = (d\alpha, d\alpha)$ and $d\alpha = 0$, so that $\alpha \in H^0(M,\mathbb{R})$. Then $d\alpha + \delta\beta = 0 \Rightarrow \delta\beta = 0$, and since $\beta \in \Omega^2_-(M,\mathbb{R})$,

$$0 = *\delta\beta = -**d*\beta = d*\beta = -d\beta.$$

Thus, $\beta \in H^2_-(M,\mathbb{R})$, and $\operatorname{Ker}\mathcal{T}_0^* = H^0(M,\mathbb{R}) \oplus H^2_-(M,\mathbb{R})$.

We now prove that the degree 0 part of $\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right) \smile \tau\left(TM\otimes\mathbb{C}\right)$ is 2. The degree 0 part of $\tau\left(TM\otimes\mathbb{C}\right)$ is 1 (directly from the definition in Section 14.14.1), and so we need to show that the degree 0 part of $\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)$ is 2. We use the fact that the Thom class $U \in H^4(BM,SM)$ and Euler class $\chi(TM) \in H^4(M)$ are related by $\pi^*\chi(TM) = i^*U$, where $\pi : BM \to M$ and $i : (BM,\phi) \to (BM,SM)$. For $V := \Lambda^1(T^*M)_{\mathbb{C}}$ and $F := \Lambda^0(T^*M)_{\mathbb{C}} \oplus \Lambda^2_-(T^*M)_{\mathbb{C}}$, we prove

$$\chi(TM) \smile \phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)_0 = ch(V)_2 - ch(F)_2\,. \tag{17.41}$$

We have the commutative diagram (where the coefficients are rational)

$$\begin{array}{ccccccc} & & H^0(M) & \overset{\smile\chi(TM)}{\to} & H^4(M) & & \\ \pi_B^{*,0}\swarrow & & \pi_{B,S}^{*,0}\downarrow & \overset{\phi}{\searrow} & \downarrow\pi_{B,S}^{*,4} & & \searrow\pi_B^{*,4} \\ H^0(BM) & \overset{i^{*,0}}{\leftarrow} & H^0(BM,SM) & \overset{\smile U}{\to} & H^4(BM,SM) & \overset{i^{*,4}}{\to} & H^4(BM)\,. \end{array}$$

Since ϕ^{-1} lowers degree by 4, $\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)_0 = \phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)_2\right)$. Thus,

$$\begin{aligned}
&\pi_B^{*,4}\left(\chi\left(TM\right) \smile \phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)_0\right)\\
&= \pi_B^{*,4}\left(\chi\left(TM\right)\right) \smile \pi_B^{*,0}\left(\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)_0\right)\\
&= \pi_B^{*,4}\left(\chi\left(TM\right)\right) \smile \pi_B^{*,0}\left(\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)_2\right)\right)\\
&= i^{*,4}U \smile i^{*,0}\pi_{B,S}^{*,0}\left(\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)_2\right)\right)\\
&= i^{*,4}\left(U \smile \pi_{B,S}^{*,0}\left(\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)_2\right)\right)\right) = i^{*,4}\left(\phi\phi^{-1}ch\left(\sigma(\mathcal{T}_0)\right)_2\right)\\
&= i^{*,4}\left(ch\left(\sigma(\mathcal{T}_0)\right)_2\right) = ch\left(\pi_B^*V\right)_2 - ch\left(\pi_B^*F\right)_2\\
&= \pi_B^{*,4}\left(ch\left(V\right)_2 - ch\left(F\right)_2\right).
\end{aligned}$$

Since M is homotopic to BM, we know that $\pi_B^{*,4}$ is an isomorphism, and the identity (17.41) is proven. To show $\phi^{-1}\left(ch\left(\sigma(\mathcal{T}_0)\right)\right)_0 = 2$, it suffices to prove that $ch\left(V\right)_2 - ch\left(F\right)_2 = 2\chi\left(TM\right)$ when $\chi\left(TM\right) \neq 0$. This identity is conveniently proved by representing the characteristic classes in terms of forms, as in Section 16.7. The Hermitian vector bundles, $V = \Lambda^1\left(T^*M\right)_{\mathbb{C}}$ and $F = \Lambda^0\left(T^*M\right)_{\mathbb{C}} \oplus \Lambda^2_-\left(T^*M\right)_{\mathbb{C}}$, are complexifications of Riemannian bundles. Hence the principal unitary frame bundles of V and F reduce to principal orthogonal frame bundles, say $\mathrm{O}(V)$ and $\mathrm{O}(F)$. When the curvature forms of the unitary frame bundles are restricted to $\mathrm{O}(V)$ and $\mathrm{O}(F)$, they have values in spaces of skew-symmetric real matrices (i.e., Lie algebras of orthogonal groups). Hence, in verifying $ch\left(V\right)_2 - ch\left(F\right)_2 = 2\chi\left(TM\right)$, it suffices to work with real skew-symmetric matrices when checking the corresponding identity on the level of invariant polynomials. The representation of $\mathrm{SO}(4)$ associated with TM is just the identity $\mathrm{Id} : \mathrm{SO}(4) \to \mathrm{SO}(4)$. The dual representation is $B \mapsto \left(B^{-1}\right)^T$, which is also Id for $B \in \mathrm{SO}(4)$, hence the representation of $\mathrm{SO}(4)$ associated with $\Lambda^1\left(T^*M\right)$ is also Id. The representation for $\Lambda^0\left(T^*M\right) \oplus \Lambda^2_-\left(T^*M\right)$ is the direct sum of the trivial representation and the representation $\mathrm{SO}(4) \to \mathrm{GL}(\Lambda^2_-\left(\mathbb{R}^{n*}\right)) \cong \mathrm{GL}(\mathfrak{so}^-)$. On the Lie algebra level, the second representation can be expressed in terms of the 't Hooft matrices $\overline{\eta}_a \in \mathfrak{so}^-$ $(a = 1, 2, 3)$ given in (17.4a), p. 437. For $A \in \mathrm{SO}(4)$, we compute

$$\begin{aligned}
\left[A, \overline{\eta}_b\right] &= \tfrac{1}{4}\sum_{c=1}^{3}\left(A \cdot \overline{\eta}_c\right)\left[\overline{\eta}_c, \overline{\eta}_b\right] = \tfrac{1}{4}\sum_{a,c=1}^{3}\left(A \cdot \overline{\eta}_c\right)\left(-2\varepsilon_{cba}\overline{\eta}_a\right)\\
&= \tfrac{1}{2}\sum_{a,c=1}^{3}\left(A \cdot \overline{\eta}_c\right)\varepsilon_{abc}\overline{\eta}_a.
\end{aligned}$$

Defining the matrix $S\left(A\right) \in \mathfrak{so}\left(3\right)$ by $\left[A, \overline{\eta}_b\right] = \sum_{a=1}^{3} S\left(A\right)_{ab}\overline{\eta}_a$, we have

$$S\left(A\right)_{ab} = \tfrac{1}{2}\sum_{c=1}^{3}\left(A \cdot \overline{\eta}_c\right)\varepsilon_{abc}.$$

Thus,

$$\begin{aligned}
S\left(A\right)_{12} &= \tfrac{1}{2}A \cdot \overline{\eta}_3 = A_{12} - A_{34}\\
S\left(A\right)_{23} &= \tfrac{1}{2}A \cdot \overline{\eta}_1 = A_{23} - A_{14}, \text{ and}\\
S\left(A\right)_{31} &= \tfrac{1}{2}A \cdot \overline{\eta}_2 = A_{31} - A_{24}.
\end{aligned}$$

Hence for $A \in \mathfrak{so}(4)$,

$$\begin{aligned}
&\tfrac{1}{2}tr\left(-A^2\right) - \tfrac{1}{2}tr\left(-S(A)^2\right) \\
&= \sum_{i<j}(A_{ij})^2 - (A_{12}-A_{34})^2 - (A_{23}-A_{14})^2 - (A_{31}-A_{24})^2 \\
&= 2\left(A_{12}A_{34} + A_{23}A_{14} + A_{31}A_{24}\right) \\
&= \tfrac{1}{4}\sum_{(i)} \varepsilon_{i_1 i_2 i_3 i_4} A_{i_1 i_2} A_{i_3 i_4}.
\end{aligned}$$

Thus,

$$\tfrac{1}{8\pi^2}tr\left(-A^2\right) - \tfrac{1}{8\pi^2}tr\left(-S(A)^2\right) = 2\tfrac{1}{32\pi^2}\sum_{(i)} \varepsilon_{i_1 i_2 i_3 i_4} A_{i_1 i_2} A_{i_3 i_4},$$

which yields the result $ch(V)_2 - ch(F)_2 = 2\chi(TM)$, upon replacing A by the curvature form Ω^θ of the Levi-Civita connection θ (or actually any connection) on FM. □

If $G = \mathrm{SU}(2)$, by (17.26), p. 450, we have

(17.42) $$c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)\left[S^4\right] = -k \Leftrightarrow ch_2\left(P \times_{\mathrm{SU}(2)} \mathfrak{su}(2)\right)\left[S^4\right] = 4k,$$

Also, for $M = S^4$, we have $\chi(M) = 2$ and $\mathrm{sig}(M) = 0$. Thus, (17.36) becomes (17.43) below. Moreover for $k \geq 1$, in the proof of the following, we show that any $\omega \in \mathcal{C}(P)^+$ is irreducible.

Corollary 17.16. *Let $P \to S^4$ be a principal* SU(2)*-bundle with*

$$c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)\left[S^4\right] = -k, \text{ for } k \geq 1.$$

If ω is a self-dual connection on P (i.e., $\omega \in \mathcal{C}(P)^+$), then ω is irreducible and

(17.43) $$\dim T_\omega\left(\mathcal{C}(P)^+ \cap S_\omega\right) = 8k - 3,$$

in the notation of (17.35).

Proof. Once ω is shown to be irreducible, (17.43) follows from (17.36) and (17.42). Suppose that the holonomy group of such ω is a subgroup $G_0 \subsetneqq \mathrm{SU}(2)$. The Lie algebra $\mathfrak{g}_0$ of G_0 has dimension at most 1, since two independent elements of $\mathfrak{su}(2)$ have a bracket outside of their span, and $\mathfrak{g}_0$ cannot be $\mathfrak{su}(2)$ since then $\mathrm{SU}(2) = \exp(\mathfrak{g}_0) \subseteq G_0$. By Proposition 16.62, p. 427, the restriction $\Omega^\omega|_{P_0}$ has values in $\mathfrak{g}_0$. Thus, $\dim \mathfrak{g}_0 \neq 0$, since

$$\Omega^\omega|_{P_0} = 0 \Rightarrow \Omega^\omega = 0 \Rightarrow -k = c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right) = 0.$$

Hence, $\dim \mathfrak{g}_0 = 1$, and the connected component of G_0 containing Id is a circle. Thus, $ad : G_0 \to \mathrm{GL}(\mathfrak{g}_0)$ is trivial. Then for $g \in G_0$, we have $R_g^*(\Omega^\omega|_{P_0}) = ad_{g^{-1}}\Omega^\omega|_{P_0} = \Omega^\omega|_{P_0}$ (i.e., $\Omega^\omega|_{P_0}$ is invariant under G_0). Hence, $\Omega^\omega|_{P_0} = \pi^*\alpha$ for a unique $\mathfrak{g}_0$-valued 2-form α on M. For a local section $\sigma : S^4 \to P_0$, we have (by the Bianchi identity $d\Omega^\omega + [\omega, \Omega^\omega] = 0$)

$$\begin{aligned}
d\alpha &= d\left(\sigma^*\pi^*\alpha\right) = \sigma^* d\left(\Omega^\omega|_{P_0}\right) \\
&= \sigma^*\left(d\Omega^\omega|_{P_0}\right) = \sigma^*\left(-\left[\omega, \Omega^\omega\right]|_{P_0}\right) = 0,
\end{aligned}$$

since $\omega|_{P_0}$ and $\Omega^\omega|_{P_0}$ have values in $\mathfrak{g}_0$ which has a trivial bracket. Since $*\Omega^\omega = \Omega^\omega$, it follows that $*\alpha = \alpha$ and $\delta\alpha = -*d*\alpha = -*d\alpha = 0$, so that α is a $\mathfrak{g}_0$-valued, harmonic 2-form on S^4. However, any such form on S^4 is zero, since $H^2(S^4) = 0$. Finally, by the transitivity of the SU(2) action on P and the equivariance of Ω^ω,

$$\alpha = 0 \Rightarrow \Omega^\omega|_{P_0} = \pi^* \alpha = 0 \Rightarrow \Omega^\omega = 0 \Rightarrow k = 0.$$

□

4. Manifold Structure for Moduli of Self-dual connections

Let $\mathcal{C}(P)^+$ be the set of weakly-irreducible, self-dual connections on a principal G-bundle P over a compact, oriented, Riemannian 4-manifold M. We assume that G is semi-simple, $\mathcal{C}(P)^+ \neq \phi$, and that there are no Δ^ω-harmonic forms in $\Omega^2_-(P \times_G \mathfrak{g})$ (e.g., M is self-dual with nonnegative scalar curvature $S \neq 0$; see Proposition 16.57, p. 414). Let $\omega \in \mathcal{C}(P)^+$ and recall from (17.33)

$$S_\omega := \left\{ \omega + \tau : \tau \in \overline{\Omega}^1 (P, P \times_G \mathfrak{g}) \text{ and } \delta^\omega \tau = 0 \right\}. \tag{17.44}$$

In this section we wil1 prove that (in a suitable sense) $\mathcal{C}(P)^+ \cap S_\omega$ is a submanifold of the affine space $\mathcal{C}(P)$ in a neighborhood of ω, and this submanifold has dimension

$$\dim T_\omega \left(\mathcal{C}(P)^+ \cap S_\omega \right) := 2\mathrm{ch}\left(P \times_G \mathfrak{g}_{\mathbb{C}}\right)[M] - \tfrac{1}{2} \dim(G) \left(\chi(M) - \mathrm{sig}(M)\right),$$

in the notation of Theorem 17.15. We introduce the space $\mathcal{C}(P)^+_m \subseteq \mathcal{C}(P)^+$ of "mildly-irreducible" self-dual connections, and show that $\mathcal{C}(P)^+_m / \mathrm{GA}(P)$ can be made into a Hausdorff topological space, such that each class $[\omega] \in \mathcal{C}(P)^+_m / \mathrm{GA}(P)$ possesses a neighborhood U which is homeomorphic to a neighborhood of ω in S_ω. These homeomorphisms are shown to constitute an atlas which makes $\mathcal{C}(P)^+_m / \mathrm{GA}(P)$ a C^∞ manifold. In doing all of this, we need to state some basic results about L^p-Sobolev Banach spaces of sections of vector bundles. Proofs or references to proofs can be found in the excellent survey article [**Can**, 1981] and [**Pal68**]. In the following review of these results, M is a compact Riemannian n-manifold with metric h, where n is not necessarily 4 until further notice.

Let $E \to M$ be a C^∞ Hermitian vector bundle, say $E = P \times_G W$, for a principal G-bundle $\pi : P \to M$, and unitary representation $\rho : G \to \mathrm{U}(W)$. Let θ be the Levi-Civita connection on FM and let ω_0 be a connection 1-form on P. In the notation of Section 16.6, there is a connection $\omega_0 \oplus \theta$ on $P \times_f FM$ (see (16.81), p. 405), and a covariant derivative operator

$$\nabla^{\omega_0 \oplus \theta} : C^\infty \left(E \otimes T^{r,s}(M)\right) \to C^\infty \left(E \otimes T^{r,s+1}(M)\right)$$

For this, see (16.86), p. 406, with $k = 1$, and note that

$$C^\infty \left(E \otimes T^{r,s}(M)\right) \cong \overline{\Omega}^0 \left(P \times_f FM, W \otimes T^{r,s}\right), \text{ while}$$
$$C^\infty \left(E \otimes T^{r,s+1}(M)\right) \cong \overline{\Omega}^1 \left(P \times_f FM, W \otimes T^{r,s}\right).$$

To shorten notation, we write $\nabla^{\omega_0 \oplus \theta}$ simply as ∇. For a nonnegative integer k and $p \in [1, \infty)$, and $u \in C^\infty(E) = C^\infty\left(E \otimes T^{0,0}(M)\right)$, we set

$$\|u\|_{p,k} := \left(\sum_{j=0}^{k} \int_M \left|\nabla^j u\right|^p \nu_h \right)^{\frac{1}{p}}, \tag{17.45}$$

where ν_h is the volume element of the Riemannian metric h, $\nabla^0 u = u$ and ∇^k is the k-fold composition

$$\nabla^k u := \left(\nabla \circ \overset{k \text{ times}}{\cdots} \circ \nabla\right) u \in C^\infty\left(E \otimes T^{0,k}(M)\right).$$

In (17.45), $|\nabla^k u|$ is computed using the Hermitian structure on E and the metric h on M.

DEFINITION 17.17. The **Sobolev space** $W^{p,k}(E)$ is the completion of $C^\infty(E)$ with the norm $\|\cdot\|_{p,k}$.

REMARK 17.18. The Banach space $W^{p,k}(E)$ coincides with the subspace of $L^p(E) := W^{p,0}(E)$ consisting of sections that have "weak derivatives" of orders $\leq k$ in L^p. More precisely, $u \in W^{p,k}(E)$, if for each $j \leq k$, there is $v_j \in L^p(E \otimes T^{0,j}(M))$, such that for all $w \in C^\infty(E \otimes T^{0,j}(M))$, we have

$$\int_M \langle v_j, w\rangle \,\nu_h = \int_M \left\langle u, (\nabla^*)^j\, w\right\rangle \nu_h,$$

where ∇^* is the formal adjoint of ∇. We say that "$\nabla^j u = v_j$ in the weak (or distributional sense)".

We have the following standard results (see [**Can**] and [**Pal68**]).

PROPOSITION 17.19. *Let $D : C^\infty(E) \to C^\infty(F)$ be a linear differential operator of order m (E, F Hermitian vector bundles over M). For each $k \geq 0$, D has a unique continuous extension*

$$D_{p,k+m} : W^{p,k+m}(E) \to W^{p,k}(E) \text{ with } \|D_{p,k+m}(\alpha)\|_{p,k} \leq K\,\|\alpha\|_{p,k+m},$$

for some $K > 0$, independent of $\alpha \in W^{p,k+m}(E)$.

PROPOSITION 17.20. *Let $C^m(E)$ be the Banach space of m-times (strongly) differentiable sections of E with norm*

$$\|u\|_{C^m} := \sum_{j=0}^{m} \sup_{x\in M} \left|\left(\nabla^j u\right)_x\right|.$$

Let $n = \dim M$, $k \in \{0, 1, 2, \ldots\}$ and $p \in [1, \infty)$. For $0 \leq m < k - \frac{n}{p}$, we have a compact (i.e., completely continuous) inclusion

$$W^{p,k}(E) \subseteq C^m(E). \tag{17.46}$$

For $k > m$ and $k - \frac{n}{p} > m - \frac{n}{q}$, we also have a compact inclusion

$$W^{p,k}(E) \subseteq W^{q,m}(E). \tag{17.47}$$

PROPOSITION 17.21 (Fundamental Elliptic Estimate). *Assume that $D : C^\infty(E) \to C^\infty(F)$ is an elliptic of order m, with formal adjoint $D^* : C^\infty(F) \to C^\infty(E)$. Suppose that for some $u \in L^p(E)$, we have $v \in W^{p,k}(F)$ such that*

$$\int_M \langle v, w\rangle\ \nu_h = \int_M \langle u, D^* w\rangle\,\nu_h,$$

for all $w \in C^\infty(F)$ (i.e., $Du := v$ exists weakly in $W^{p,k}(F)$). Then $u \in W^{p,k+m}(E)$. Moreover, for each $k \geq 0$, there is a constant $C_k > 0$ independent of u, such that

$$\|u\|_{p,k+m} \leq C_k \left(\|Du\|_{p,k} + \|u\|_p\right). \tag{17.48}$$

If $Du \in C^\infty(F)$, then for all $k \geq 0$, $\|Du\|_{p,k} < \infty$, and we have $\|u\|_{p,k+m} < \infty$, in which case $u \in C^\infty(E)$ by (17.46). In particular, weak solutions of $Du = 0$ are C^∞.

PROPOSITION 17.22 (Sobolev Multiplication). *Let $Q : E_1 \times E_2 \to E_3$ be a smooth, bilinear map of Riemannian (or Hermitian) vector bundles over M (compact with $\dim(M) = n$). Then subject to the conditions below, the induced map $\mathcal{Q} : C^\infty(E_1) \times C^\infty(E_2) \to C^\infty(E_3)$ extends uniquely to a bounded bilinear map $\overline{\mathcal{Q}} : L^{p_1,k_1}(E_1) \times L^{p_2,k_2}(E_2) \to L^{p_3,k_3}(E_3)$, i.e.,*

$$\left\|\overline{\mathcal{Q}}(x_1, x_2)\right\|_{p_3,k_3} \leq C \|x_1\|_{p_1,k_1} \|x_2\|_{p_2,k_2}$$

for C depending only on Q, the connections ∇^{E_1}, ∇^{E_2}, ∇^{E_3}, and the Riemannian metric on M. The conditions are that $k_3 \leq \min\{k_1, k_2\}$ and

$$k_3 - \tfrac{n}{p_3} \leq \begin{cases} \left(k_1 - \tfrac{n}{p_1}\right) + \left(k_2 - \tfrac{n}{p_2}\right) & \text{if } \max\left\{k_1 - \tfrac{n}{p_1},\ k_2 - \tfrac{n}{p_2}\right\} < 0 \\ \min\left\{k_1 - \tfrac{n}{p_1},\ k_2 - \tfrac{n}{p_2}\right\} & \text{if } \max\left\{k_1 - \tfrac{n}{p_1},\ k_2 - \tfrac{n}{p_2}\right\} > 0. \end{cases}$$

If $\max\left\{k_1 - \tfrac{n}{p_1},\ k_2 - \tfrac{n}{p_2}\right\} = 0$, then it suffices that

$$k_3 - \tfrac{n}{p_3} < \min\left\{k_1 - \tfrac{n}{p_1},\ k_2 - \tfrac{n}{p_2}\right\} = \left(k_1 - \tfrac{n}{p_1}\right) + \left(k_2 - \tfrac{n}{p_2}\right),$$

where the inequality is strict. When $p_1 = p_2 = p_3 = p$ and $k_3 = k_2 = j \leq k = k_1 > n/p$, we obtain the boundedness of

$$\overline{\mathcal{Q}} : W^{p,k}(E_1) \times W^{p,j}(E_2) \to W^{p,j}(E_3).$$

(Note that by (17.46) we have $W^{p,k}(E_1) \subseteq C^0(E_1)$ for $k > n/p$, whence the induced map $\overline{\mathcal{Q}}$ may be defined pointwise.)

PROPOSITION 17.23 (Elliptic Decomposition). *Let $D : C^\infty(E) \to C^\infty(F)$ be a differential operator of order m with a symbol which is injective or surjective, and let D^* be the formal adjoint of D. If $D^{*p,k+m} : L^{p,k+m}(F) \to L^{p,k}(E)$ is the Sobolev extension of D^*, then we have the following direct sum decompositions into closed subspaces for $k \geq 0$ and $p \geq 2n/(2m+n)$ (e.g., $p \geq 2$),*

$$W^{p,k+m}(E) = \operatorname{Ker}(D^{p,k+m}) \oplus \operatorname{Im}(D^{*p,k+2m}) \quad \text{and}$$
$$W^{p,k+m}(F) = \operatorname{Ker}(D^{*p,k+m}) \oplus \operatorname{Im}(D^{p,k+2m}).$$

If the symbol of D is injective, then $D^* \circ D$ is elliptic and

$$\operatorname{Ker}(D^{p,k+m}) = \operatorname{Ker}(D) = \operatorname{Ker}(D^* \circ D) \subseteq C^\infty(E)$$

is finite-dimensional. If the symbol of D is surjective, then $D \circ D^*$ is elliptic and

$$\operatorname{Ker}(D^{*p,k+m}) = \operatorname{Ker}(D^*) = \operatorname{Ker}(D \circ D^*) \subseteq C^\infty(F)$$

is finite-dimensional. In particular, if D is elliptic (i.e., with injective and surjective symbol), then both $\operatorname{Ker}(D)$ and $\operatorname{Ker}(D^*)$ are finite-dimensional.

Given the estimate (17.48), this last result is not difficult to prove (e.g., see [**Can**]). Proposition 17.23 is indispensable in verifying the hypotheses of the following implicit function theorems for Banach spaces in applications where the differential of the map is an elliptic operator.

Theorem 17.24 (Implicit Function Theorem I). *Let B_1 and B_2 be Banach spaces and let $F : B_1 \to B_2$ be C^k $(1 \leq k \leq \infty)$. Assume that $F_{*x} : T_xB_1 \to T_{F(x)}B_2$ is a surjection and $T_xB_1 = \operatorname{Ker}(F_{*x}) \oplus H$, where H is closed. Then $F^{-1}(F(x))$ is a C^k submanifold of B_1 in a neighborhood of x, and its tangent space at x is $\operatorname{Ker}(F_{*x})$.*

We will also need the following version, whose proof is also in [**Lan**].

Theorem 17.25 (Implicit Function Theorem II). *Let B_1, B_2 and B_3 be Banach spaces and let $F : B_1 \times B_2 \to B_3$ be a C^k map $(1 \leq k \leq \infty)$ with $F(x_1, x_2) = x_3$. Suppose that the partial derivative of F in the B_2-direction at (x_1, x_2) (i.e., $(D_2F)_{(x_1,x_2)} : B_2 \to B_3$) . Then there are neighborhoods U_1 of x_1 and U_2 of x_2, such that there is a unique C^k map $G : U_1 \to U_2$ whose graph $\{(x, G(x)) : x \in U_1\}$ is $F^{-1}(x_3) \cap U_1 \times U_2$. Indeed, there is a neighborhood U_3 of x_3 and a unique C^k function $H : U_1 \times U_3 \to U_2$ such that for all $z \in U_3$ we have $F(x, H(x, z)) = z$.*

When $B_1 = \{0\}$, we obtain a special case of Theorem 17.25 which is usually established beforehand, namely

Theorem 17.26 (Inverse Function Theorem). *$F : B_2 \to B_3$ be a C^k map $(1 \leq k \leq \infty)$ with $F(x_2) = x_3$. If $(DF)_{x_2} : B_2 \to B_3$ is a bicontinuous isomorphism, then there are neighborhoods U_1 of x_1 and U_3 of x_3 and a unique C^k function $H : U_3 \to U_2$, such that for all $z \in U_3$ we have $F(H(z)) = z$.*

In the rest of this section, the assumptions of the opening paragraph are in force (e.g., $\dim M = 4$). We set $E := P \times_G \mathfrak{g}$ and make the identification of the space $\overline{\Omega}^k(P, \mathfrak{g})$ (of equivariant forms on P) with $\Omega^k(E)$. For any (smooth) $\omega \in \mathcal{C}(P)$, recall that $\mathcal{C}(P) = \omega + \overline{\Omega}^1(P, \mathfrak{g})$. Thus, we can identify $\mathcal{C}(P)$ with the set of equivalence classes of pairs in $\mathcal{C}(P) \times \overline{\Omega}^1(P, \mathfrak{g})$, where

$$(\omega_1, \alpha_1) \equiv (\omega_2, \alpha_2) \Leftrightarrow \omega_1 + \alpha_1 = \omega_2 + \alpha_2 \Leftrightarrow \omega_1 - \omega_2 = \alpha_2 - \alpha_1 \tag{17.49}$$

Since $\overline{\Omega}^1(P, \mathfrak{g}) \cong C^\infty(E \otimes \Lambda^1(M))$, say $\alpha \longleftrightarrow \widetilde{\alpha}$, we can also regard $\mathcal{C}(P)$ as the set of equivalence classes of pairs in $\mathcal{C}(P) \times C^\infty(E \otimes \Lambda^1(M))$, where

$$(\omega_1, \widetilde{\alpha}_1) \equiv (\omega_2, \widetilde{\alpha}_2) \Leftrightarrow \widetilde{(\omega_1 - \omega_2)} = \widetilde{\alpha}_2 - \widetilde{\alpha}_1.$$

We define $\mathcal{C}(P)^{p,k}$ as the set of equivalence classes of $\mathcal{C}(P) \times W^{p,k}\left(E \otimes \Lambda^1(M)\right)$, where the equivalence relation is still defined by (17.49), and ω_1 and ω_2 are still smooth, but $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ are not necessarily smooth elements of $W^{p,k}\left(E \otimes \Lambda^1(M)\right)$ (although $\widetilde{\alpha}_2 - \widetilde{\alpha}_1$ is smooth). One may wish to think of $\mathcal{C}(P)^{p,k}$ as the set of all "generalized" connection 1-forms on P which differ from a smooth connection by an element of $W^{p,k}\left(E \otimes \Lambda^1(M)\right)$. We have only given a more precise meaning to this thought. In what follows, for simplicity we will not adhere to the notation $\widetilde{\alpha}$, but simply use α. Note that each $\omega \in \mathcal{C}(P)$ defines a bijection

$$\phi_\omega : \mathcal{C}(P)^{p,k} \to W^{p,k}\left(E \otimes \Lambda^1(M)\right), \text{ where } \phi_\omega([(\omega, \alpha)]) = \alpha.$$

For $\omega_1, \omega_2 \in \mathcal{C}(P)$, $\phi_{\omega_1} \circ \phi_{\omega_2}^{-1}$ is a translation of $W^{p,k}\left(E \otimes \Lambda^1(M)\right)$, since

$$\begin{aligned}\left(\phi_{\omega_1} \circ \phi_{\omega_2}^{-1}\right)(\alpha) &= \phi_{\omega_1}([(\omega_2, \alpha)]) \\ &= \phi_{\omega_1}([(\omega_1, \alpha + (\omega_2 - \omega_1))]) = \alpha + (\omega_2 - \omega_1)\end{aligned}$$

Thus, the ϕ_ω serve as a collection of charts making $\mathcal{C}(P)^{p,k}$ a C^∞ Banach manifold.

THEOREM 17.27. *Let $\omega \in \mathcal{C}(P)^+$ and assume that there are no nonzero Δ^ω-harmonic forms in $\Omega^2_-(E)$ $(E = P \times_G \mathfrak{g})$. Let S_ω be defined by (17.44). For $2 \le k \in \mathbb{Z}$ and $p \in [1,\infty)$ with $pk > 4$, we have that $\mathcal{C}(P)^+ \cap S_\omega$ is a C^∞ submanifold of $\mathcal{C}(P)^{p,k+1}$.*

PROOF. Recall from (17.34) that

$$\omega + \tau \in \mathcal{C}(P)^+ \cap S_\omega \Leftrightarrow 0 = Q(\tau) := \left(\delta^\omega \tau, (1-*)\left(D^\omega \tau + \tfrac{1}{2}[\tau,\tau]\right)\right).$$

Thus, $\mathcal{C}(P)^+ \cap S_\omega = Q^{-1}(0)$ and we wish to apply Implicit Function Theorem I (say, IFT I) to a "Sobolev extension" of Q. To this end, we first show that $Q : C^\infty\left(E \otimes \Lambda^1(M)\right) \to C^\infty\left(E \otimes \left(\Lambda^0(M) \oplus \Lambda^2_-(M)\right)\right)$ has a C^∞ extension to

$$Q_{p,k+1} : W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \to W^{p,k}\left(E \otimes \left(\Lambda^0(M) \oplus \Lambda^2_-(M)\right)\right), \tag{17.50}$$

provided that $p(k+1) > 4$; the stronger inequality $pk > 4$ will be used later. Note that Q is a sum of the first-order elliptic differential operator $\tau \mapsto (\delta^\omega \tau, (1-*)D^\omega \tau)$ and the quadratic map $\tau \mapsto (1-*)\frac{1}{2}[\tau,\tau]$. The differential operator has an extension to $W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$ by Proposition 17.19. To show that the quadratic map also extends, first note since $p(k+1) > 4$, Proposition 17.22 yields a bounded bilinear extension

$$W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \times W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \to W^{p,k+1}\left(E \otimes \Lambda^2(M)\right)$$

of the bilinear function $(\tau_1,\tau_2) \mapsto [\tau_1,\tau_2]$. Since $\tau \mapsto (\tau,\tau)$ is clearly bounded and linear and $(1-*)$ is a differential operator of order 0 so that Proposition 17.19 applies, the composition

$$\tau \mapsto (\tau,\tau) \mapsto (1-*)\tfrac{1}{2}[\tau,\tau]$$

defines a bounded, quadratic map

$$W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \to W^{p,k+1}\left(E \otimes \Lambda^2_-(M)\right) \subseteq W^{p,k}\left(E \otimes \Lambda^2_-(M)\right).$$

Thus, Q has the extension $Q_{p,k+1}$ in (17.50). The differential $(Q_{p,k+1})_{*0}$ of $Q_{p,k+1}$ at $\tau = 0$ is the linear part given for $\tau' \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$ by

$$(Q_{p,k+1})_{*0}(\tau') := \left((\delta^\omega)_{p,k+1}\tau', (1-*)(D^\omega)_{p,k+1}\tau'\right).$$

This is just the Sobolev extension $T_{p,k+1}$ of the elliptic operator

$$T(\tau') := \left(\delta^\omega \tau', \tfrac{1}{2}(1-*)D^\omega \tau'\right)$$

used in the proof of Theorem 17.15, p. 462. There we showed that $\operatorname{Ker} T^* = \{0\}$, under the assumptions that ω is weakly-irreducible and the space of Δ^ω-harmonic forms in $\Omega^2_-(E)$ is trivial. With the goal of applying IFT I to $Q_{p,k+1}$, we wish to use Proposition 17.23 with $D = T^*$ to deduce that $(Q_{p,k+1})_{*0}$ is onto, but we need to show that T^* is elliptic first. Since $T^*(\alpha,\beta) = D^\omega\alpha + \delta^\omega\beta$ (see (17.38), p. 463), the symbol of T^* is given by

$$\sigma(T^*)_\xi(\alpha,\beta) = \alpha\xi - \beta(\xi^\#,\cdot),$$

where $\xi^\#$ is defined by $\xi = h(\xi^\#,\cdot)$. If $\xi \ne 0$ and $\sigma(T^*)_\xi(\alpha,\beta) = 0$, then

$$0 = \alpha\xi(\xi^\#) - \beta(\xi^\#,\xi^\#) = \alpha|\xi|^2 \Rightarrow \alpha = 0,$$

and so $\beta(\xi^{\#}, \cdot) = 0$. This means that β is a sum of bicovectors in the 3-dimensional orthogonal complement of ξ, in which case

$$0 = \beta \wedge \beta = -\beta \wedge *\beta = -|\beta|^2 \nu_h,$$

since $\beta \in \Lambda^2_- (TM^*)$. Thus, $\sigma(T^*)_\xi$ is injective, and an isomorphism for dimensional reasons. Hence T^* is elliptic, and $(Q_{p,k+1})_*$ is onto by Proposition 17.23 with $D = T^*$. Using the ellipticity of T, we have the spliting

$$W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) = \operatorname{Ker} T \oplus \operatorname{Im}\left(T^*_{p,k+2}\right)$$

of the domain of $(Q_{p,k+1})_{*0}$. Then we may then finally apply IFT I to deduce that in a neighborhood U of ω, $(Q_{p,k+1})^{-1}(0)$ is a submanifold of $W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$ of dim $\operatorname{Ker} T$. Since $\mathcal{C}(P)^+ \cap U \subseteq (Q_{p,k+1})^{-1}(0) \cap U$ is clear, it remains to show that $(Q_{p,k+1})^{-1}(0) \cap U \subseteq \mathcal{C}(P)^+ \cap U$, where U is possibly replaced by a smaller neighborhood of ω. For $\tau \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$ with $Q_{p,k+1}(\tau) = 0$, we have

$$T_{p,k+1}(\tau) = -\left(0, (1-*)\tfrac{1}{2}[\tau,\tau]\right) \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right),$$

and then using (17.48),

$$\begin{aligned} \|\tau\|_{p,k+2} &\leq C\left(\|T_{p,k+1}(\tau)\|_{p,k+1} + \|\tau\|_p\right) \\ &= C\left(\left\|\tfrac{1}{2}[\tau,\tau]\right\|_{p,k+1} + \|\tau\|_p\right) \\ &\leq C\left(C' \|\tau\|^2_{p,k+1} + \|\tau\|_p\right). \end{aligned}$$

Replacing k by $k+1$, $k+2, \ldots$, we have $\|\tau\|_{p,j} < \infty$ for all j. Thus, $\tau \in C^\infty\left(E \otimes \Lambda^1(M)\right)$ by (17.46) and $\omega + \tau$ is a C^∞ self-dual connection. It remains to show that U can be chosen so that the C^∞ elements of U are weakly-irreducible. Let $\mathcal{L}$ be the Banach space of bounded (continuous) linear transformations from $W^{p,k+1}(E)$ to $W^{p,k}\left(E \otimes \Lambda^1(M)\right)$. For $\tau \in C^\infty\left(E \otimes \Lambda^1(M)\right)$, by Proposition 17.19, $(D^{\omega+\tau})_{p,k+1} \in \mathcal{L}$. Indeed, for $\alpha \in W^{p,k+1}(E)$, we have

$$\begin{aligned} \left\|(D^{\omega+\tau})_{p,k+1}(\alpha)\right\|_{p,k} &= \left\|(D^{\omega})_{p,k+1}(\alpha) + [\tau,\alpha]\right\|_{p,k} \\ &\leq \left\|(D^{\omega})_{p,k+1}(\alpha)\right\|_{p,k} + \|[\tau,\alpha]\|_{p,k+1} \\ &\leq C_1 \|\alpha\|_{p,k+1} + C_2 \|\tau\|_{p,k+1} \|\alpha\|_{p,k+1} \\ &\leq \left(C_1 + C_2 \|\tau\|_{p,k+1}\right) \|\alpha\|_{p,k+1}, \end{aligned}$$

by Propositions 17.19 and 17.22. Indeed, this shows that the function

$$\tau \mapsto \left(D^{\omega+\tau}\right)_{p,k+1} \in \mathcal{L}$$

extends to a continuous affine map

$$\Phi : W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \to \mathcal{L}.$$

We know that $\operatorname{Ker}\left(D^\omega : C^\infty(E) \to C^\infty\left(E \otimes \Lambda^1(M)\right)\right) = \{0\}$, since we have assumed that ω is weakly-irreducible. Moreover, $\operatorname{Ker}(D^\omega) = \operatorname{Ker}(\delta^\omega D^\omega)$ and $\delta^\omega D^\omega$ is elliptic. Suppose that $(D^\omega)_{p,k+1}(\alpha) = 0$, for $\alpha \in W^{p,k+1}(E)$. By Proposition

17.20, $\alpha \in C^1(E)$ for $1 < k + 1 - \frac{4}{p}$ or $kp > 4$. Then $(\delta^\omega D^\omega)(\alpha) = 0$ weakly, since for $\alpha \in C^1(E)$ and $\beta \in C^\infty(E)$, we can perform integration by parts:

$$\int_M \left\langle \alpha, (\delta^\omega D^\omega)^* \beta \right\rangle \nu_h = \int_M \langle \alpha, \delta^\omega D^\omega \beta \rangle \nu_h = \int_M \langle D^\omega \alpha, D^\omega \beta \rangle \nu_h = 0,$$

where the last equality holds since $0 = (D^\omega)_{p,k+1}(\alpha) = D^\omega \alpha$ for $\alpha \in W^{p,k+1}(E) \subseteq C^1(E)$. Thus, for $\alpha \in W^{p,k+1}(E)$ with $pk > 4$, we may apply Proposition 17.21 to obtain

$$\begin{aligned}(D^\omega)_{p,k+1}(\alpha) = 0 &\Rightarrow (\delta^\omega D^\omega)(\alpha) = 0 \text{ weakly}\\ &\Rightarrow \alpha \in C^\infty(E) \text{ and } (\delta^\omega D^\omega)(\alpha) = 0\\ &\Rightarrow \alpha \in C^\infty(E) \text{ and } \|D^\omega(\alpha)\|^2 = (\delta^\omega D^\omega \alpha, \alpha) = 0\\ &\Rightarrow \alpha \in \operatorname{Ker}(D^\omega) \Rightarrow \alpha = 0,\end{aligned}$$

by the assumption that ω is weakly-irreducible. Thus, $\operatorname{Ker}(D^\omega)_{p,k+1} = \{0\}$, and so $\Phi(0)$ is in the open set, say $\mathcal{L}_0$, of injective elements of $\mathcal{L}$. Since Φ is continuous, there is some neighborhood U_0 of 0 in $W^{p,k+1}(E)$ such that $\Phi(\tau) \in \mathcal{L}_0$ for $\tau \in U_0$. Thus replacing U with $U \cap U_0$, we are assured that the C^∞ elements of U are weakly-irreducible. □

THEOREM 17.28. *Let $\omega \in \mathcal{C}(P)$ be weakly-irreducible. For $1 \le p < \infty$ and $k \in \mathbb{Z}$ with $pk > 4$, there are positive constants C_1 and C_2, such that for every $\tau \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$ with $\|\tau\|_{p,k+1} < C_1$, there is a unique $\sigma(\tau) \in W^{p,k+2}(E)$ with $\|\sigma(\tau)\|_{p,k+2} < C_2$ and*

$$(\delta^\omega)_{p,k+1}\left(\Phi(\operatorname{Exp}(\sigma(\tau))) \cdot (\omega + \tau) - \omega\right) = 0. \tag{17.51}$$

Here, the function $(\tau, s) \mapsto \Phi(\operatorname{Exp}(s)) \cdot (\omega + \tau) - \omega$, which has meaning for smooth pairs (τ, s), is proved to analytically extend to a function on $W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \times W^{p,k+2}(E)$ with values in $W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$, so that (17.51) makes sense. In other words, $\Phi(\operatorname{Exp}(\sigma(\tau))) \cdot (\omega + \tau) \in S_\omega^{p,k+1}$, where

$$S_\omega^{p,k+1} := \left\{\omega + \tau' \in \mathcal{C}(P)^{p,k+1} : (\delta^\omega)_{p,k+1}(\tau') = 0\right\}. \tag{17.52}$$

Moreover, $\tau \mapsto \sigma(\tau)$ is a C^∞ function.

PROOF. Let

$$\mathcal{Q} : C^\infty\left(E \otimes \Lambda^1(M)\right) \times C^\infty(E) \to C^\infty(E) \tag{17.53}$$

be defined by

$$\mathcal{Q}(\tau, s) := \delta^\omega\left(\Phi(\operatorname{Exp}(s)) \cdot (\omega + \tau) - \omega\right).$$

As in (16.30), p. 384, let $f = \operatorname{Exp}(s) \in C(P, G) \subseteq C\left(P, \operatorname{GL}\left(\mathbb{C}^N\right)\right)$, where we continue to assume that G is a matrix group. According to (16.27), p. 382, we have

$$\Phi(f) \cdot \omega = f d\left(f^{-1}\right) + f \omega f^{-1}, \text{ and so}$$

$$\Phi(\operatorname{Exp}(s)) \cdot (\omega + \tau) = \operatorname{Exp}(s)\, d(\operatorname{Exp}(-s)) + \operatorname{Exp}(s)(\omega + \tau)\operatorname{Exp}(-s).$$

For $A, B \in \mathfrak{g}$, we will use the following identity:

$$\operatorname{Exp}(A)\, B \operatorname{Exp}(-A) = \sum_{i=0}^\infty \frac{1}{i!}(\mathfrak{ad}_A)^i(B),$$

where $\mathfrak{ad}_A(B) = [A, B]$. This follows from

$$\frac{1}{i!}(\mathfrak{ad}_A)^i B = \sum_{m+n=i} \frac{1}{n!m!} A^n B A^m,\ i \geq 0,$$

which can be proved by straightforward induction. Similarly, we have

$$\frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (ds) = -\sum_{m+n=i+1} \frac{1}{n!m!} s^n d(s^m),$$

for $s \in C(P, \mathfrak{g}) \cong C^\infty(E)$, and this yields

$$\operatorname{Exp}(s)\, d(\operatorname{Exp}(-s)) = -\sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (ds),\ \text{and}$$

$$\begin{aligned}
&\Phi(\operatorname{Exp}(s)) \cdot (\omega + \tau) = \operatorname{Exp}(s)\, d(\operatorname{Exp}(-s)) + \operatorname{Exp}(s)(\omega+\tau)\operatorname{Exp}(-s) \\
&= -\sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (ds) + \sum_{i=0}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^i (\omega+\tau) \\
&= -\sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (ds) + \omega - \sum_{i=1}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^{i-1}([\omega, s]) + \sum_{i=0}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^i (\tau) \\
&= \omega - \sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (ds + [\omega, s]) + \sum_{i=0}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^i (\tau) \\
&= \omega - \sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (D^\omega s) + \sum_{i=0}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^i (\tau).
\end{aligned}$$

Hence for $p(k+1) > 4$, by repeated use of Proposition 17.22,

$$\begin{aligned}
(\tau, s) &\mapsto \Phi(\operatorname{Exp}(s)) \cdot (\omega+\tau) - \omega \\
&= -\sum_{i=0}^{\infty} \frac{1}{(i+1)!}(\mathfrak{ad}_s)^i (D^\omega s) + \sum_{i=0}^{\infty} \frac{1}{i!}(\mathfrak{ad}_s)^i (\tau),
\end{aligned}$$

extends to a C^∞ (indeed, analytic) map

$$W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \times W^{p,k+2}(E) \to W^{p,k+1}\left(E \otimes \Lambda^1(M)\right).$$

Composing this map with $(\delta^\omega)_{p,k+1}$ gives us an extension of $\mathcal{Q}$ in (17.53), say

$$\mathcal{Q}_{p,k+1} : W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \times W^{p,k+2}(E) \to W^{p,k}\left(E \otimes \Lambda^1(M)\right).$$

The derivative of $\mathcal{Q}_{p,k+1}$ at $(0,0)$ is

$$(\mathcal{Q}_{p,k+1})_{*(0,0)}(\tau', s') = (\delta^\omega)_{p,k+1}\left(-(D^\omega)_{p,k+2}(s') + \tau'\right).$$

As we wish to apply Implicit Function Theorem II (IFT II) to $\mathcal{Q}_{p,k+1}$, we note that the partial derivative $D_2(\mathcal{Q}_{p,k+1})_{(0,0)}$ of $\mathcal{Q}_{p,k+1}$ in the $W^{p,k+2}(E)$ direction is the Sobolev extension $(-\Delta^\omega)_{p,k+2}$ of the formally self-adjoint elliptic operator $-\Delta^\omega = -\delta^\omega D^\omega$. By Proposition 17.23,

$$\begin{aligned}
W^{p,k}(E) &= \operatorname{Ker}(-\Delta^\omega) \oplus (-\Delta^\omega)_{p,k+2}\left(W^{p,k+2}(E)\right) \\
&= (-\Delta^\omega)_{p,k+2}\left(W^{p,k+2}(E)\right),
\end{aligned}$$

since ω weakly-irreducible implies that $\operatorname{Ker}(-\Delta^\omega) = \operatorname{Ker} D^\omega = 0$. Thus, our $D_2\left(\mathcal{Q}_{p,k+1}\right)_{(0,0)} = (-\Delta^\omega)_{p,k+2}$ is onto with trivial kernel, and hence is a bicontinuous isomorphism by the Open Mapping Theorem. Thus, the IFT II applies to give us the constants C_1 and C_2, and C^∞ function σ. □

Theorem 17.28 provides us with a "local slice" for the action of a "Sobolev extension" of $\mathrm{GA}(P)$ on the space $\mathcal{C}(P)^{p,k+1}$. In other words, every connection $\omega' \in \mathcal{C}(P)^{p,k+1}$ in a suitably small neighborhood of $\omega \in \mathcal{C}(P)$ is gauge-equivalent to a connection in $S^{p,k+1}_\omega$ via some "generalized" gauge transformation $\operatorname{Exp} s$ where $\|s\|_{p,k+2}$ small. Of course, one would like to precisely define a Sobolev extension $\mathrm{GA}(P)^{p,k+2}$ of $\mathrm{GA}(P)$. Since $\mathrm{GA}(P)$ (or equivalently $C(P,G)$) is not the space of sections of a *vector* bundle, some elaboration is needed. We have assumed that G is a matrix group, say a Lie subgroup of $\mathrm{GL}\left(\mathbb{C}^N\right)$. Now $\mathrm{GL}\left(\mathbb{C}^N\right)$ (and hence G) is contained in the vector space $\mathfrak{gl}\left(\mathbb{C}^N\right)$ of all linear endomorphisms of $\mathbb{C}^N$. Thus, $C(P,G)$ (and hence $\mathrm{GA}(P)$) can identified with a subset of $\overline{\Omega}^0\left(P, \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$, where the representation $G \to \mathrm{GL}\left(\mathfrak{gl}\left(\mathbb{C}^N\right)\right)$ is the adjoint representation (i.e., $g \cdot A = gAg^{-1}$, $g \in G$ and $A \in \mathfrak{gl}\left(\mathbb{C}^N\right)$). Now $\overline{\Omega}^0\left(P, \mathfrak{gl}\left(\mathbb{C}^N\right)\right) \cong C^\infty\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$ which has a Sobolev extension $W^{p,k}\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$. Thus, it makes sense to define $C(P,G)^{p,k}$ to be the closure in $W^{p,k}\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$ of the subset corresponding to $C(P,G) \subseteq \overline{\Omega}^0\left(P, \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$. Since (by Proposition 17.20) we have a continuous inclusion $W^{p,k}\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right) \subseteq C^0\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$ for $pk > 4$, it follows that $C(P,G)^{p,k}$ consists of continuous, *ad*-equivariant, G-valued functions on P. Note that $C(P,G)^{p,k}$ may then be identified with a certain set, say $\mathrm{GA}(P)^{p,k}$, of "continuous gauge transformations" (i.e., continuous (as opposed to C^∞), equivariant, fiber-preserving homeomorphisms of P). A proof that (for $pk > \dim M$) $\mathrm{GA}(P)^{p,k}$ is actually a Lie group, modelled on the Banach space $W^{p,k}(E)$ in such a way that $\operatorname{Exp} : W^{p,k}(E) \to \mathrm{GA}(P)^{p,k}$ is a local diffeomorphism, can be found in [**MV**] for the case $p = 2$, but their proof works for $pk > \dim M$ as well. Although as a map between Banach spaces, σ in Theorem 17.28 is a C^∞ function of τ, we still need to prove that $\sigma(\tau)$ is C^∞ if τ is C^∞ as the next result states.

Theorem 17.29. *In Theorem 17.28, $\sigma(\tau)$ is C^∞ if τ is C^∞, provided C_1 is chosen small enough.*

Proof. From the proof of Theorem 17.28, we know that $s := \sigma(\tau) \in W^{p,k+2}(E)$ obeys the equation

$$\begin{aligned} 0 &= (\delta^\omega)_{p,k+1}\left(\Phi\left(\operatorname{Exp}(s)\right) \cdot (\omega + \tau) - \omega\right) \\ &= (\delta^\omega)_{p,k+1}\left(\operatorname{Exp}(s)\, d\left(\operatorname{Exp}(-s)\right) + \operatorname{Exp}(s)\,(\omega+\tau)\operatorname{Exp}(-s) - \omega\right) \end{aligned}$$

$$= (\delta^\omega)_{p,k+1}\left(-\sum_{i=0}^{\infty} \frac{1}{(i+1)!}\left(\mathfrak{ad}_s\right)^i\left(\left(D^\omega\right)_{p,k+2}(s)\right) + \sum_{i=0}^{\infty}\frac{1}{i!}\left(\mathfrak{ad}_s\right)^i(\tau)\right)$$

$$= (\delta^\omega)_{p,k+1}\left(\begin{array}{l} -\left(D^\omega\right)_{p,k+2}(s) - \sum_{i=1}^{\infty}\frac{1}{(i+1)!}\left(\mathfrak{ad}_s\right)^i\left(\left(D^\omega\right)_{p,k+2}(s)\right) \\ +\sum_{i=0}^{\infty}\frac{1}{i!}\left(\mathfrak{ad}_s\right)^i(\tau) \end{array}\right)$$

Thus,

$$\begin{aligned}(\Delta^{\omega})_{p,k+2}(s) &= (\delta^{\omega})_{p,k+1}(D^{\omega})_{p,k+2}(s)\\ &= (\delta^{\omega})_{p,k+1}\left(R\left(s,(D^{\omega})_{p,k+2}(s),\tau\right)\right), \text{ where}\end{aligned}$$

$$\begin{aligned}R\left(s,(D^{\omega})_{p,k+2}(s),\tau\right) = &-\sum_{i=1}^{\infty}\frac{1}{(i+1)!}(\mathfrak{ad}_s)^i\left((D^{\omega})_{p,k+2}(s)\right)\\ &+\sum_{i=0}^{\infty}\frac{1}{i!}(\mathfrak{ad}_s)^i(\tau).\end{aligned}$$

Note that

$$\begin{aligned}(\delta^{\omega})_{p,k+1}\left(R\left(s,(D^{\omega})_{p,k+2}(s),\tau\right)\right) &= -\sum_{i=1}^{\infty}\frac{1}{(i+1)!}(\mathfrak{ad}_s)^i\left((\Delta^{\omega})_{p,k+2}(s)\right)\\ +\mathcal{P}\left(s,(\delta^{\omega})_{p,k+1}(s),(D^{\omega})_{p,k+2}(s),\tau,(\delta^{\omega})_{p,k+1}(\tau)\right),\end{aligned}$$

where $\mathcal{P}$ is a power series in its arguments and contains terms which result when $(\delta^{\omega})_{p,k+1}$ is passed through the "factor" $(\mathfrak{ad}_s)^i$ in $(\mathfrak{ad}_s)^i\left((D^{\omega})_{p,k+2}(s)\right)$. Thus,

$$\begin{aligned}&\left(1+\sum_{i=1}^{\infty}\frac{1}{(i+1)!}(\mathfrak{ad}_s)^i\right)(\Delta^{\omega})_{p,k+2}(s)\\ &=\mathcal{P}\left(s,(\delta^{\omega})_{p,k+1}(s),(D^{\omega})_{p,k+2}(s),\tau,(\delta^{\omega})_{p,k+1}(\tau)\right)\end{aligned}$$

For arbitrarily small C_2 in Theorem 17.28, the constant C_1 exists. By (17.46), we may then assume that C_2 is small enough so that $\|s\|_{p,k+2} < C_2 \Rightarrow \|s\|_{C^0}$ is small enough so that the endomorphism

$$\Psi(s) := \left(1+\sum_{i=1}^{\infty}\frac{1}{(i+1)!}(\mathfrak{ad}_s)^i\right) : W^{p,k}(E)\to W^{p,k}(E)$$

can be inverted. Thus,

$$(\Delta^{\omega})_{p,k+2}(s) = \Psi(s)^{-1}\mathcal{P}\left(s,(\delta^{\omega})_{p,k+1}(s),(D^{\omega})_{p,k+2}(s),\tau,(\delta^{\omega})_{p,k+1}(\tau)\right).$$

Since the $W_{p,k}$ norm of $\Psi(s)^{-1}$ is finite, as well as the $W_{p,k}$ norms of all of the arguments of Q (recall that $\tau \in C^{\infty}\left(E\otimes\Lambda^1(M)\right)$), it follows that $\left\|(\Delta^{\omega})_{p,k+2}(s)\right\|_{p,k} < \infty$. Hence by (17.48), $s \in W^{p,k+4}(E)$, and repeating the argument with k replaced by k by $k+2$, etc., yields $s \in C^{\infty}(E)$. □

We have proved that every C^{∞} connection in a sufficiently small neighborhood (in $\mathcal{C}(P)^{p,k+1}$, $pk > 4$) of a weakly-irreducible $\omega \in \mathcal{C}(P)$ is gauge equivalent, via a unique small (i.e., close to Id in $\mathrm{GA}(P)^{p,k+2}$) gauge transformation which is necessarily smooth, to a connection in S_ω. We have yet to introduce hypotheses which will enable us to prove that no two smooth connections near ω in S_ω are gauge equivalent by a possibly large gauge transformation. For this we need a stronger condition on ω than weak-irreducibility. If we insist that ω be irreducible, then at some later point, we would have to face the problem of proving that every connection near an irreducible connection is irreducible. While this is quite believable, there seems to be no simple proof. To get around the difficulty, we introduce a milder irreducibility condition which is still stronger than weak-irreducibility. To this end, note that the adjoint representation $ad : G \to \mathrm{GL}(\mathfrak{g})$ induces a representation $r = ad \otimes ad^* : G \to \mathrm{GL}(\mathrm{End}\,\mathfrak{g})$ given by $r(g)(h) = ad_g \circ h \circ ad_{g^{-1}}$. There is an inner product $k \otimes k^*$ on $\mathrm{End}\,\mathfrak{g}$ induced from k (minus the Killing form) on $\mathfrak{g}$, and r is orthogonal with respect to $k \otimes k^*$. Now $\mathrm{End}\,\mathfrak{g} = H_1 \oplus H_1^\perp$, where H_1 is the invariant subspace

$$H_1 := \{h \in \mathrm{End}\,(\mathfrak{g}) : r\,(g)\,(h) = h, \text{ for all } g \in G\}\,. \tag{17.54}$$

PROPOSITION 17.30. *Suppose that G is connected and semi-simple. If $ad_{g'} \in H_1$ for some $g' \in G$, then $g' \in Z\,(G) :=$ the center of G.*

PROOF. For all $g \in G$, we have $ad_{g'} = r\,(g)\,(ad_{g'}) = ad_g \circ ad_{g'} \circ ad_{g^{-1}}$ or $ad_{gg'g^{-1}g'^{-1}} = \mathrm{Id}$. We claim that $gg'g^{-1}g'^{-1} \in Z\,(G)$. Indeed, since G is connected, any $g'' \in G$ can be written as $\exp A$. Using the fact that $\exp \circ ad_g = Ad_g \circ \exp$ for all $g \in G$, we then have

$$\begin{aligned} Ad_{gg'g^{-1}g'^{-1}}\,(g'') &= Ad_{gg'g^{-1}g'^{-1}}\,(\exp A) \\ &= \exp\left(ad_{gg'g^{-1}g'^{-1}}A\right) = \exp A = g''. \end{aligned}$$

Thus, $gg'g^{-1}g'^{-1} \in Z\,(G)$. Since G is semi-simple, $Z\,(G)$ is discrete. It follows that $gg'g^{-1}g'^{-1}$ is the identity of G, because g (and hence $gg'g^{-1}g'^{-1}$) can be connected to the identity by a path in G. Thus, $gg' = g'g$ and $g' \in Z\,(G)$. □

DEFINITION 17.31. We call $\omega \in \mathcal{C}(P)$ **mildly-irreducible** if there is no nonzero element of $H_1^\perp$, which is fixed by all transformations $r(g_0)$ as g_0 ranges over the holonomy group G_0 of ω.

Clearly, ω irreducible (i.e., $G_0 = G$) $\Rightarrow$ ω mildly-irreducible (i.e., $Z(G_0)$ is discrete). Moreover, we have

PROPOSITION 17.32. *Let $\pi : P \to M$ be a principal G-bundle, where G is connected and semi-simple, let $\omega \in \mathcal{C}(P)$. For $E := P \times_G \mathfrak{g}$, let $\mathrm{End}\,(E) = \mathrm{End}_1\,(E) \oplus \mathrm{End}_1\,(E)^\perp$ be the orthogonal decomposition arising from $\mathrm{End}\,\mathfrak{g} = H_1 \oplus H_1^\perp$ in (17.54). We have ω mildly-irreducible, if and only if*

$$\mathrm{Ker}\left(D^\omega : C^\infty(\mathrm{End}_1\,(E)^\perp) \to \Omega^1\left(\mathrm{End}_1\,(E)^\perp\right)\right) = 0.$$

Moreover, if ω is mildly-irreducible, then

1. *$Z(G_0) = Z\,(G)$*
2. *The isotropy group I_ω of ω is $\{R_g : g \in Z\,(G)\}$*
3. *ω is weakly-irreducible (see Definition 17.14, p. 461).*

PROOF. Suppose that $D^\omega \alpha = 0$ for $\alpha \in C^\infty(\mathrm{End}_1\,(E)^\perp)$ where ω is mildly-irreducible. Then α (regarded as in $C(P, H_1^\perp)$) is constant on the holonomy bundle P_0 of ω through some $p_0 \in P$ (see Section 16.16.8). Since α is equivariant as well, we have (for $g_0 \in G_0 =$ holonomy group of ω) $r(g_0^{-1})\alpha(p_0) = \alpha(p_0 g_0) = \alpha(p_0)$, whence $\alpha(p_0) \in H_1^\perp$ is fixed by all members of $r(G_0)$. Thus, $\alpha(p_0) = 0$ since ω is mildly-irreducible, and so $\alpha = 0$. Conversely, if ω is not mildly-irreducible, then there is $0 \neq h \in H_1^\perp$ invariant under $r(G_0)$. Let α be the constant function h on

P_0, and note that α extends by equivariance to a nonzero element of $C(P, H_1^\perp)$ in $\operatorname{Ker} D^\omega$.

For (1), suppose that $g_c \in Z(G_0)$. Then $ad_{g_c} \in \operatorname{End}(\mathfrak{g})$ is invariant under all elements of $r(G_0)$, since for all $A \in \mathfrak{g}$ and $g_0 \in G_0$,

$$\begin{aligned}(r(g_0)(ad_{g_c}))(A) &= ad_{g_0} \circ ad_{g_c} \circ ad_{g_0^{-1}}(A) \\ &= g_0 g_c \left(g_0^{-1} A g_0\right) g_c{}^{-1} g_0^{-1} = g_c A g_c{}^{-1} = ad_{g_c}(A).\end{aligned}$$

The same is true of the projection of ad_{g_c} onto $H_1^\perp$. Hence, ω mildly-irreducible implies that $ad_{g_c} \in H_1$. Then $g_c \in Z(G)$ by Proposition 17.30. For (2) note that any $z \in Z(G)$, $R_z : P \to P$ is a gauge transformation which acts trivially on $\mathcal{C}(P)$, since $R_z \cdot \omega = R^*_{z^{-1}}\omega = ad_z\omega = \omega$ for all $\omega \in \mathcal{C}(P)$. By Proposition 16.61, p. 426, the homomorphism $I_\omega \to G$ given by $\Phi(f) \mapsto f(p_0)$ maps the isotropy group I_ω isomorphically onto the centralizer $Z(G_0)$. Thus, for ω mildly-irreducible, it follows from $Z(G_0) = Z(G)$ that $I_\omega = \{R_z : z \in Z(G)\}$. For (3), note that since $Z(G_0) = Z(G)$ is discrete by (1), ω is weakly-irreducible by Proposition 17.13, p. 461. □

In distinction to Proposition 17.28, the following result essentially states that near a mildly-irreducible connection ω, the set

$$S_\omega = \left\{\omega + \tau : \tau \in \overline{\Omega}^1(P, P \times_G \mathfrak{g}) \text{ and } \delta^\omega \tau = 0\right\}$$

serves as a *global* slice. In other words, no two distinct connections in S_ω near a mildly-irreducible ω are gauge equivalent by a possibly *large* gauge transformation. Some key ideas we use in the proof were inspired by [**AHS2**].

THEOREM 17.33. *Let $\omega \in \mathcal{C}(P)$ be mildly-irreducible. Then there is a constant $C > 0$, such that if $\|\omega - \omega'\|_{p,k+1} \le C$ for $\omega' \in \mathcal{C}(P)$, then ω' is mildly-irreducible. Also, if $pk > 4$, ω_1, $\omega_2 \in S_\omega$ with $\|\omega - \omega_i\|_{p,k+1} \le C$ $(i = 1, 2)$, and there is $F \in \mathrm{GA}(P)^{p,k+2} \subseteq W^{p,k+2}\left(P \times_G \operatorname{End}\left(\mathbb{C}^N\right)\right)$ with $\omega_2 = F^{-1*}\omega_1$, then $F \in I_\omega = \{R_z : z \in Z(G)\}$, and hence $\omega_2 = \omega_1$.*

PROOF. By Proposition 17.32, we know that $\omega' \in \mathcal{C}(P)$ is mildly-irreducible if and only if

$$\operatorname{Ker}\left(D^{\omega'} : C^\infty(\operatorname{End}_1(E)^\perp) \to \Omega^1\left(\operatorname{End}_1(E)^\perp\right)\right) = 0.$$

Using Proposition 17.22, note that $\omega' \mapsto D^{\omega'}$ extends to a C^∞ map

$$\mathcal{C}(P)^{p,k+1} \to L(W^{p,k+1}(\operatorname{End}_1(E)^\perp), W^{p,k}(\operatorname{End}_1(E)^\perp \otimes \Lambda^1(M)),$$

where $L(V_1, V_2) :=$ bounded linear maps from V_1 to V_2. Since the subset of injective bounded linear maps is open, we see that the mildly-irreducible connections form an open subset of $\mathcal{C}(P)^{p,k+1}$. This proves the first assertion of Theorem 17.33.

Let $\tau = \omega_2 - \omega_1 \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$, and let

$$f \in C(P, G)^{p,k+2} \subseteq W^{p,k+2}\left(P \times_G \mathfrak{gl}\left(\mathbb{C}^N\right)\right)$$

correspond to $F^{-1} \in \mathrm{GA}(P)^{p,k+2}$ (i.e., $f = \Phi\left(F^{-1}\right)$ or $F(p) = pf(p)$ for all $p \in P$). Note that $2 < (k+2) - 4/p$ for $pk > 4$, whence f is C^2 by (17.46), p. 468.

By (16.27), p. 382, we have

$$\begin{aligned} &\omega_2 = F^{-1*}\omega_1 \\ &\quad\Leftrightarrow \omega_2 = f^{-1}df + f^{-1}\omega_1 f \\ &\quad\quad = f^{-1}(df + \omega_1 f) = f^{-1}(df + \omega_1 f - f\omega_1) + \omega_1 \\ &\quad\Leftrightarrow \tau = \omega_2 - \omega_1 = f^{-1}(df + [\omega_1, f]) = f^{-1}(D^{\omega_1} f). \end{aligned} \tag{17.55}$$

The idea is to prove that if C is sufficiently small (so that $\|\tau\|_{p,k+1}$ will also be small), then F will be close to some R_z ($z \in Z(G)$) or equivalently $z^{-1}f$ will be close enough to the constant map (to the identity Id of G) so that we can apply Theorem 17.28 to conclude (since $\omega_1, \omega_2 \in S_\omega$) that $R_{z^{-1}} \circ F = \mathrm{Id}$ (or $F = R_z$).

Since $\mathrm{Ker}\left(D^{\omega_1}|_{\mathrm{End}_1(E)^\perp}\right) = 0$, we know that the positive self-adjoint elliptic operator

$$\Delta^{\omega_1} = \delta^{\omega_1} D^{\omega_1} : C^\infty\left(\mathrm{End}_1(E)^\perp\right) \to C^\infty\left(\mathrm{End}_1(E)^\perp\right)$$

has a positive smallest eigenvalue λ. Thus, for $\eta \in C^\infty\left(\mathrm{End}_1(E)^\perp\right)$,

$$\|D^{\omega_1}\eta\|^2 = (\Delta^{\omega_1}\eta, \eta) \geq \lambda \|\eta\|^2. \tag{17.56}$$

Recall from (17.55) that $\tau = f^{-1}(D^{\omega_1} f) \in W^{p,k+1}\left(E \otimes \Lambda^1(M)\right)$, where $f \in C(P,G)^{p,k+2}$. Via the representation $ad : G \to \mathrm{O}(\mathfrak{g}) \subset \mathrm{End}(\mathfrak{g})$, we may associate to each $f \in C(P,G)^{p,k+2}$, some $f_{\mathrm{End}} \in W^{p,k+2}(\mathrm{End}(E))$. Note also that $\mathfrak{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ induces a map

$$W^{p,k+1}\left(E \otimes \Lambda^1(M)\right) \to W^{p,k+1}\left(\mathrm{End}(E) \otimes \Lambda^1(M)\right)$$

Applying this to both sides of the equation $\tau = f^{-1}(D^{\omega_1} f)$, we obtain an equation, say

$$\tau_{\mathrm{End}} = \left(f^{-1}(D^{\omega_1} f)\right)_{\mathrm{End}} = f_{\mathrm{End}}^{-1}(D^{\omega_1} f_{\mathrm{End}}),$$

where both sides are in $W^{p,k+1}\left(\mathrm{End}(E) \otimes \Lambda^1(M)\right)$. According to $\mathrm{End}(E) = \mathrm{End}_1(E) \oplus \mathrm{End}_1(E)^\perp$, we can orthogonally decompose

$$\left(f^{-1}(D^{\omega_1} f)\right)_{\mathrm{End}} = \left(f^{-1}(D^{\omega_1} f)\right)_{\mathrm{End}_1} + \left(f^{-1}(D^{\omega_1} f)\right)_{\mathrm{End}_1^\perp}.$$

Then, pointwise,

$$|\tau_{\mathrm{End}}|^2 = \left|\left(f^{-1}(D^{\omega_1} f)\right)_{\mathrm{End}}\right|^2 = |(D^{\omega_1} f)_{\mathrm{End}}|^2 \geq \left|(D^{\omega_1} f)_{\mathrm{End}_1^\perp}\right|^2.$$

Upon integrating this, (17.56) yields

$$\|\tau_{\mathrm{End}}\|^2 \geq \left\|(D^{\omega_1} f)_{\mathrm{End}_1^\perp}\right\|^2 \geq \lambda \left\|f_{\mathrm{End}_1^\perp}\right\|^2. \tag{17.57a}$$

From Proposition 17.20, we know that there is a constant K so that

$$\max(|\tau_{\mathrm{End}}|) \leq K \|\tau_{\mathrm{End}}\|_{p,k+1}. \tag{17.58a}$$

Since $d\left(\left|f_{\mathrm{End}_1^\perp}\right|^2\right) = 2\left\langle f_{\mathrm{End}_1^\perp}, D^{\omega_1} f_{\mathrm{End}_1^\perp}\right\rangle$,

$$\begin{aligned} \left|d\left(\left|f_{\mathrm{End}_1^\perp}\right|^2\right)\right| &\leq 2\left|D^{\omega_1}\left(f_{\mathrm{End}_1^\perp}\right)\right| \left|f_{\mathrm{End}_1^\perp}\right| \\ &= 2|\tau_{\mathrm{End}}| \left|f_{\mathrm{End}_1^\perp}\right| \leq K \|\tau_{\mathrm{End}}\|_{p,k+1} \left|f_{\mathrm{End}_1^\perp}\right|. \end{aligned}$$

Hence, for $\|\tau\|_{p,k+1}$ sufficiently small, can assure that $\left|d\left(\left|f_{\mathrm{End}_1^\perp}\right|^2\right)\right| < \epsilon$ on M. This implies that $\left|f_{\mathrm{End}_1^\perp}\right|$ can be made arbitrarily close to a constant function on M as $\|\tau\|_{p,k+1} \to 0$. In fact, by (17.57a) and (17.58a), we know this constant function must be zero. Consequently, $f_{\mathrm{End}} = f_{\mathrm{End}_1} + f_{\mathrm{End}_1^\perp}$ can be made arbitrarily C^0 close to the subbundle $\mathrm{End}\,(E)_1$. By Proposition 17.30 and the definition of f_{End}, it follows that $f \in C\,(P,G)^{p,k+2}$ can be made arbitrarily C^0 close to the constant map $P \to \{z\}$ for some $z \in Z\,(G)$ for $\|\tau\|_{p,k+1}$ sufficiently small. Then $zf^{-1} \in C\,(P,G)^{p,k+2}$ is C^0 close to the constant map with value $I \in G$. Thus, $zf^{-1} = \mathrm{Exp}\, s$ for some $s \in C\,(P,G)^{p,k+2}$, and

$$\begin{aligned}(\mathrm{Exp}\, s)\cdot\omega_1 &= \left(zf^{-1}\right)\cdot\omega_1 = \left(zf^{-1}\right) d\left(\left(zf^{-1}\right)^{-1}\right) + \left(zf^{-1}\right)\omega_1\left(zf^{-1}\right)^{-1} \\ &= f^{-1}df + f^{-1}\omega_1 f = F\cdot\omega_1 = \omega_2.\end{aligned}$$

By the same computation as in the proof of Theorem 17.28, the equation

$$(\mathrm{Exp}\, s)\cdot\omega_1 - \omega_1 = \omega_2 - \omega_1 = \tau$$

can be expanded to yield

$$\left(1 + \sum_{n=1}^{\infty}\frac{1}{(n+1)!}\mathfrak{ad}_s^n\right) D^{\omega_1} s = \tau.$$

Since $\max|s|$ can be made arbitrarily small for $\|\tau\|_{p,k+1}$ sufficiently small. Thus, for sufficiently small $\|\tau\|_{p,k+1}$, $|\mathfrak{ad}_s|$ will be small enough so that the series can be inverted, say

$$\left(1 + \sum_{n=1}^{\infty}\frac{1}{(n+1)!}\mathfrak{ad}_s^n\right)^{-1} = 1 + \sum_{n=1}^{\infty} c_n \mathfrak{ad}_s^n$$

for constants c_n. We then have $D^{\omega_1}s = (1 + \sum_{n=1}^{\infty} c_n\mathfrak{ad}_s^n)\tau$, and by Proposition 17.22

$$\|D^{\omega_1}s\|_{p,j} \le (1 + \sum_{n=1}^{\infty} c_n K^n \|s\|_{p,j}^n)\,\|\tau\|_{p,k+1}\,, \tag{17.59}$$

for each $j \in \{0,\ldots,k+1\}$ and some constant $K > 0$, independent of s and τ. As a consequence of the definition of Sobolev spaces, we have

$$\|s\|_{p,j+1} \le C\left(\|D^{\omega_1}s\|_{p,j} + \|s\|_{p,0}\right). \tag{17.60}$$

As $\|\tau\|_{p,k+1} \to 0$, we have $\|s\|_{p,0} \to 0$ (since then $\max|s| \to 0$ and M has finite volume). By (17.59), as $\|\tau\|_{p,k+1} \to 0$, we then have $\|D^{\omega_1}s\|_{p,0} \to 0$, in which case $\|s\|_{p,1} \to 0$ by (17.60). We then see inductively that if $\|s\|_{p,j} \to 0$ as $\|\tau\|_{p,k+1} \to 0$, then $\|s\|_{p,j+1} \to 0$ as $\|\tau\|_{p,k+1} \to 0$ for $j \in \{0,\ldots,k+1\}$. Thus, $\|s\|_{p,k+2} \to 0$ as $\|\tau\|_{p,k+1} \to 0$. Hence for $\|\tau\|_{p,k+1}$ sufficiently small, we will have $\|s\|_{p,k+2}$ small enough so that if $\mathrm{Exp}\, s\cdot\omega_1 = \omega_2$, then $\mathrm{Exp}\, s = \mathrm{Id}$ by Theorem 17.28 since ω_1 and ω_2 are in S_ω. □

Let $\mathcal{C}\,(P)_m^+$ denote the set (assumed nonvoid) of mildly irreducible self-dual C^∞ connections on P. We now show how to use the local slices S_ω, $\omega \in \mathcal{C}\,(P)_m^+$

to introduce a C^∞ atlas on $\mathcal{M}^+ := \mathcal{C}(P)^+_m / \mathrm{GA}(P)$ in such a way that the topology induced on $\mathcal{M}^+$ by the atlas is Hausdorff and $\dim(\mathcal{M}^+) = 2ch(E)[M] - \dim G(\chi(M) - \mathrm{sig}(M))$ found in Theorem 17.15, p. 462.

For each $\omega \in \mathcal{C}(P)^+_m$, we choose a neighborhood of ω, say $U_\omega \subseteq S_\omega \cap \mathcal{C}(P)^+_m$, such that the function $U_\omega \to \mathcal{M}^+$ given by

$$\omega' \mapsto [\omega'] := \{F \cdot \omega' : F \in \mathrm{GA}(P)\}$$

is injective; U_ω exists by Theorem 17.33. Let $[U_\omega]$ denote the image of U_ω in $\mathcal{M}^+$, and let $\phi_\omega : [U_\omega] \to U_\omega$ be the inverse. We know from Theorem 17.27, that we can (and do) choose U_ω small enough so that U_ω is a C^∞ submanifold of $\mathcal{C}(P)^{p,k+1}$. To show that the ϕ_ω constitute an atlas, we need to prove that

$$\phi_\omega \circ \phi_{\omega'}^{-1} : \phi_{\omega'}([U_\omega] \cap [U_{\omega'}]) \to \phi_\omega([U_\omega] \cap [U_{\omega'}])$$

is C^∞ if $[U_\omega] \cap [U_{\omega'}]$ is nonempty. Select $\omega'' \in U_{\omega'}$, such that $[\omega''] \in [U_\omega]$. Then, there is $F \in \mathrm{GA}(P)$, such that $F \cdot \omega'' \in U_\omega$. The map $\tilde{\omega} \mapsto F \cdot \tilde{\omega}$ is a C^∞ diffeomorphism of $\mathcal{C}(P)^{p,k+1}$, as one can verify using Proposition 16.28, p. 382, and Proposition 17.22. Hence $F \cdot U_{\omega'}$ is a C^∞ submanifold of $\mathcal{C}(P)^{p,k+1}$ containing $F \cdot \omega'' \in U_\omega$. A neighborhood of $F \cdot \omega''$ in $F \cdot U_{\omega'}$ will be contained in the ball of radius C_1 about ω in Theorem 17.28, provided we choose the U_ω small enough. Then Theorem 17.28 provides us with a C^∞ map $\omega + \tau \mapsto \mathrm{Exp}(\sigma(\tau))(\omega + \tau)$ which will carry this neighborhood of $F \cdot \omega''$ smoothly into U_ω, proving that $\phi_\omega \circ \phi_{\omega'}^{-1}$ is C^∞ at the arbitrary $\omega'' \in \phi_{\omega'}([U_\omega] \cap [U_{\omega'}])$. Note that Theorem 17.29 is needed to ensure that if $\omega + \tau$ is C^∞, then $\mathrm{Exp}(\sigma(\tau))$ given by Theorem 17.28 is in $\mathrm{GA}(P)$ (i.e., $\mathrm{Exp}(\sigma(\tau))$ is C^∞). The topology on $\mathcal{M}^+$ is the smallest topology that makes the maps ϕ_ω continuous. To show that $\mathcal{M}^+$ is Hausdorff, select ω and ω' with $[\omega] \neq [\omega']$, and set $\tau = \omega' - \omega \in \Omega^1(E)$. We argue as in [**AHS2**] by first noting that $[\omega] \neq [\omega']$ means that for all $f \in C(P, G) \cong \mathrm{GA}(P)$

$$\begin{aligned} 0 &\neq f^{-1} \cdot \omega - \omega' = f^{-1} df + f^{-1} \omega f - \omega + \omega - \omega' \\ &= f^{-1}(df + \omega f - f\omega) + \omega - \omega' = f^{-1} D^\omega f - \tau. \end{aligned} \tag{17.61}$$

As in the proof of Theorem 17.33, we can consider the $f_{\mathrm{End}} \in C^\infty(\mathrm{End}\, E)$ ($E := P \times_G \mathfrak{g}$) associated with $f \in C(P, G)$; and $\tau_{\mathrm{End}} \in \Omega^1(\mathrm{End}\, E)$ associated with $\tau \in \Omega^1(E)$. If we can show that

$$\left\| f_{\mathrm{End}}^{-1} D^\omega(f_{\mathrm{End}}) - \tau_{\mathrm{End}} \right\|_{2,0} > \epsilon > 0. \tag{17.62}$$

as f varies, over $C(P, G)$, then $\left\| f^{-1} \cdot \omega - \omega' \right\|_{2,0}$ will be bounded away from 0. For $pk > 4$ (indeed $p(k+1) > 2$), by (17.47) we would then have $\left\| f^{-1} \cdot \omega - \omega' \right\|_{p,k+1}$ bounded away from 0. Thus, the orbits of ω and ω' will be bounded away from each other in $\mathcal{C}(P)^{p,k+1}$, and we will have that $\mathcal{M}^+$ is Hausdorff. In order to establish (17.62), we introduce

$$\begin{aligned} &D^\omega - \tau_{\mathrm{End}} : C^\infty(\mathrm{End}\, E) \to \Omega^1(\mathrm{End}\, E) \text{ given by} \\ &(D^\omega - \tau_{\mathrm{End}})(f_{\mathrm{End}}) := D^\omega(f_{\mathrm{End}}) - f_{\mathrm{End}} \tau_{\mathrm{End}}, \text{ and} \\ &\Delta^{\omega,\tau} := (D^\omega - \tau_{\mathrm{End}})^* (D^\omega - \tau_{\mathrm{End}}) : C^\infty(\mathrm{End}\, E) \to C^\infty(\mathrm{End}\, E). \end{aligned}$$

Now, since f_{End} is an isometry of $\mathrm{End}\, E$, we have

$$\begin{aligned}\left\| f_{\mathrm{End}}^{-1} D^{\omega}\left(f_{\mathrm{End}}\right) - \tau_{\mathrm{End}}\right\|_{2,0} &= \left\| D^{\omega}\left(f_{\mathrm{End}}\right) - f_{\mathrm{End}}\tau_{\mathrm{End}}\right\|_{2,0} \\ &= \left\| \left(D^{\omega} - \tau_{\mathrm{End}}\right)\left(f_{\mathrm{End}}\right)\right\|_{2,0} \geq \lambda \left\| f_{\mathrm{End}}^{\perp}\right\|_{2,0},\end{aligned}$$

where $f_{\mathrm{End}}^{\perp}$ is the projection of f_{End} onto the orthogonal complement of $K := \mathrm{Ker}\left(\Delta^{\omega,\tau}\right)$ and λ is the smallest positive eigenvalue of $\Delta^{\omega,\tau}$. We need to show that $\left\| f_{\mathrm{End}}^{\perp}\right\|_{2,0}$ is bounded away from 0. Suppose, on the contrary, that there is a sequence $f_n \in \mathrm{GA}(P)$, such that $\left\| \left(f_n\right)_{\mathrm{End}\, 2,0}^{\perp} \right\| \to 0$. Since $\left(f_n\right)_{\mathrm{End}}$ is an isometry, $\left\|\left(f_n\right)_{\mathrm{End}}\right\|_{2,0}$ is constant, and so $\left(f_n\right)_{\mathrm{End}} - \left(f_n\right)_{\mathrm{End}}^{\perp}$ is a bounded sequence in K, with $\dim K < \infty$ since $\Delta^{\omega,\tau}$ is elliptic. Thus, by passing to a subsequence, we may assume that $\left(f_n\right)_{\mathrm{End}} - \left(f_n\right)_{\mathrm{End}}^{\perp}$ converges. Calling the limit $h_{\infty} \in K$, we have

$$\left\| h_{\infty} - \left(f_n\right)_{\mathrm{End}}\right\|_{2,0}^{2} = \left\| h_{\infty} - \left(\left(f_n\right)_{\mathrm{End}} - \left(f_n\right)_{\mathrm{End}}^{\perp}\right)\right\|_{2,0}^{2} + \left\| \left(f_n\right)_{\mathrm{End}}^{\perp}\right\|_{2,0}^{2} \to 0,$$

whence $\left(f_n\right)_{\mathrm{End}} \to h_{\infty}$ in $L^2\left(\mathrm{End}\, E\right)$. We remark that $h_{\infty} = \left(f_{\infty}\right)_{\mathrm{End}}$ for some $f_{\infty} \in \mathrm{GA}\left(P\right)$. To see this, note that $ad : G \to \mathrm{O}\left(\mathfrak{g}\right)$ has discrete kernel $Z\left(G\right)$. Thus, if $\left(f\right)_{\mathrm{End}} = \left(f'\right)_{\mathrm{End}}$, then $f' = zf$ for some $z \in Z\left(G\right)$. Hence, f_n may be replaced by $z_n f_n$ which converges in L^2 to some $f_{\infty} \in \mathrm{GA}\left(P\right)$ for which $\left(f_{\infty}\right)_{\mathrm{End}} = \lim_{n\to\infty}\left(z_n f_n\right)_{\mathrm{End}} = \lim_{n\to\infty}\left(f_n\right)_{\mathrm{End}} = h_{\infty}$. Thus, $D^{\omega} f_{\infty} - f_{\infty}\tau = 0$ or $f_{\infty}^{-1} D^{\omega} f_{\infty} - \tau = 0$ in violation of (17.61). In summary, we have proved the following.

THEOREM 17.34. *Let $\pi : P \to M$ be a principal G-bundle with G compact and semi-simple over a compact, oriented, self-dual, Riemannian 4-manifold with scalar curvature $S \geq 0$ and $S \neq 0$. Then the space $C(P)_m^{+}/\,\mathrm{GA}(P)$ of moduli of mildly-irreducible self-dual connections has (if non-empty) the structure of a Hausdorff C^{∞} manifold of dimension*

$$2ch\left(P \times_G \mathfrak{g}_{\mathbb{C}}\right)[M] - \tfrac{1}{2}\dim\left(G\right)\left(\chi\left(M\right) - \mathrm{sig}\left(M\right)\right).$$

REMARK 17.35. If

$$\mathcal{C}_0\left(P\right)_m^{+} := \left\{\omega \in \mathcal{C}\left(P\right)_m^{+} : \mathrm{Ker}\left(\Delta^{\omega} : \Omega_{-}^{2}\left(P \times_G \mathfrak{g}_{\mathbb{C}}\right) \hookleftarrow\right) = \{0\}\right\}$$

is nonempty, then the same conclusion holds for $\mathcal{C}_0\left(P\right)_m^{+}/\,\mathrm{GA}(P)$, and we may drop the self-duality and positive scalar curvature assumptions on M. It is likely a proof that $\mathcal{C}_0\left(P\right)_m^{+} = \mathcal{C}\left(P\right)_m^{+}$ for a generic class of metrics on M can be constructed along the lines found [**FU**].

CHAPTER 18

The Local Index Theorem for Twisted Dirac Operators

One of our main goals in this chapter, will be to show that the classical geometric operators such as the signature operator, the de Rham operator, the Dolbeaut operator and even the Yang-Mills operator can all be locally expressed in terms of twisted Dirac operators. The index of any of these operators (and *their* twists) can then be obtained from the Local Index Theorem for twisted Dirac operators which is proved in unusual detail. This theorem supplies a globally-defined n-form on M, whose integral is the index of an operator which is perhaps only locally of the form of a twisted Dirac operator, as with the classical geometric operators. This n-form (or "index density") is expressed in terms of curvature forms of characteristic classes. The Index Theorem thus obtained then becomes a formula that relates a global invariant quantity, namely the index of an operator, to the integral of a local quantity involving curvature. This is in the spirit of the Gauss-Bonnet Theorem which is a special case.

1. Clifford Algebras and Spinors

Let V be a real vector space with positive-definite inner product $\langle\cdot,\cdot\rangle$.

Definition 18.1. The **Clifford algebra** $Cl\,(V)$ is the real algebra generated by V and $\mathbb{R}$ with the relation

$$vw + wv = -2\,\langle v, w\rangle\,, \text{ for all } v, w \in V.$$

Note that the product of v and w in $Cl\,(V)$ is denoted by the plain juxtaposition vw. Also, $v^2 := vv = -\langle v, v\rangle = -\|v\|^2$, and $vw = -wv$ if $\langle v, w\rangle = 0$. In the following, we take $\{e_1, \ldots, e_n\}$ to be an orthonormal basis of V.

Example 18.2. If $\dim V = 1$, then $e_1^2 := e_1 e_1 = -1$, and $Cl\,(V)$ is isomorphic to the algebra $\mathbb{C}$ of complex numbers, via $\alpha_0 + \alpha_1 e_1 \mapsto \alpha_0 + i\alpha_1$, for $\alpha_0, \alpha_1 \in \mathbb{R}$.

Example 18.3. If $\dim V = 2$, then it is easy to check that

$$\alpha_0 + \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_1 e_2 \mapsto \alpha_0 + \alpha_1 \mathbf{i} + \alpha_2 \mathbf{j} + \alpha_3 \mathbf{k}$$

defines an isomorphism $Cl\,(V)$ with the algebra $\mathbb{H}$ of quaternions. Note that $(e_1 e_2)^2 = e_1 e_2 e_1 e_2 = -e_1 e_1 e_2 e_2 = -1$, and $(e_1 e_2)\, e_1 = -e_1^2 e_2 = e_2$, etc..

Example 18.4. For $\dim V = 3$, one can check that there is an isomorphism $Cl\,(V) \cong \mathbb{H} \oplus \mathbb{H}$ determined by

$$e_1 \mapsto (-\mathbf{i}, \mathbf{i})\,,\ e_2 \mapsto (-\mathbf{j}, \mathbf{j})\,,\ e_3 \mapsto (-\mathbf{k}, \mathbf{k})\,.$$

EXAMPLE 18.5. If $\dim V = 4$, we have an isomorphism, $Cl(V) \cong \mathbb{H}(2) :=$ the algebra of 2×2 quaternionic matrices, determined by

$$e_1 \mapsto \begin{bmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{bmatrix}, e_2 \mapsto \begin{bmatrix} 0 & \mathbf{j} \\ \mathbf{j} & 0 \end{bmatrix}, e_3 \mapsto \begin{bmatrix} 0 & \mathbf{k} \\ \mathbf{k} & 0 \end{bmatrix}, e_4 \mapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{18.1}$$

For $\dim V = n$, we write $Cl(V) = Cl_n$. In [**LaMi**], it is shown that there is the following table of algebra isomorphisms
(18.2)

n	0	1	2	3	4	5	6	7	8
$Cl_n \cong$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H} \oplus \mathbb{H}$	$\mathbb{H}(2)$	$\mathbb{C}(4)$	$\mathbb{R}(8)$	$\mathbb{R}(8) \oplus \mathbb{R}(8)$	$\mathbb{R}(16)$

Here $\mathbb{R}(k)$, $\mathbb{C}(k)$ and $\mathbb{H}(k)$ denote the algebras of $k \times k$ matrices with entries in $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ respectively. Moreover, it is also proven that there is periodicity relation $Cl_{n+8} \cong Cl_8 \otimes Cl_n = \mathbb{R}(16) \otimes Cl_n$, so that this table can be extended indefinitely. The case of non-degenerate indefinite inner products with signature (r, s) is also handled in [**LaMi**]; we have only considered $(n, 0)$.

Let $\Lambda^*(V) = \oplus_{k=1}^n \Lambda^k(V)$ be the exterior algebra of V. While $\Lambda^*(V)$ is not isomorphic to $Cl(V)$ as an algebra, there is a linear isomorphism of vector spaces

$$\begin{aligned} &\mathcal{L} : \Lambda^*(V) \cong Cl(V) \text{ determined by} \\ &\mathcal{L}(e_{i_1} \wedge \cdots \wedge e_{i_k}) := e_{i_1} \cdots e_{i_k}, \quad (i_1 < \cdots < i_k), \end{aligned}$$

where we continue to let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of V. It can be shown that $\mathcal{L}$ is $\mathrm{O}(n)$-equivariant and independent of the choice of orthonormal basis. Moreover via $\mathcal{L}$, the natural inner product on $\Lambda^*(V)$ gives us an inner product and norm on $Cl(V)$. There is an exponential map $\exp : Cl(V) \to Cl(V)$ given by

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \quad x \in Cl(V),$$

which converges, since $\left\|x^k\right\| \le c^k \|x\|^k$ for some constant c depending on n but not on x. Indeed, for $x, y \in Cl(V)$, each of the 2^n components of xy (relative to the basis $\{e_1 \cdots e_k : i_1 < \cdots < i_k\}$) can be no larger than $2^n \|x\| \|y\|$, and so $\|xy\| \le \sqrt{2^n (2^n \|x\| \|y\|)^2} = 2^{3n/2} \|x\| \|y\|$.

We define the *bracket* (or *commutator*) of any $x, y \in Cl(V)$, by $[x, y] = xy - yx$. The linear subspace $\mathcal{L}\left(\Lambda^2(V)\right)$ is closed under bracket, since

$$\begin{aligned} [e_i e_j, e_h e_k] &= e_i e_j e_h e_k - e_h e_k e_i e_j = e_i e_j e_h e_k + e_h (e_i e_k + 2\delta_{ik}) e_j \\ &= e_i e_j e_h e_k + e_h e_i e_k e_j + 2\delta_{ik} e_h e_j \\ &= e_i e_j e_h e_k - (e_i e_h + 2\delta_{ih}) e_k e_j + 2\delta_{ik} e_h e_j \\ &= e_i e_j e_h e_k - e_i e_h e_k e_j - 2\delta_{ih} e_k e_j + 2\delta_{ik} e_h e_j \\ &= e_i e_j e_h e_k + e_i e_h (e_j e_k + 2\delta_{jk}) - 2\delta_{ih} e_k e_j + 2\delta_{ik} e_h e_j \\ &= e_i e_j e_h e_k + e_i e_h e_j e_k + 2\delta_{jk} e_i e_h - 2\delta_{ih} e_k e_j + 2\delta_{ik} e_h e_j \\ &= e_i e_j e_h e_k - e_i (e_j e_h + 2\delta_{jh}) e_k + 2\delta_{jk} e_i e_h - 2\delta_{ih} e_k e_j + 2\delta_{ik} e_h e_j \\ &= -2\delta_{jh} e_i e_k + 2\delta_{jk} e_i e_h - 2\delta_{ih} e_k e_j + 2\delta_{ik} e_h e_j. \end{aligned}$$

For $A = [a_{ij}]$, $B = [b_{ij}] \in \mathfrak{so}(n)$ (i.e., A, B anti-symmetric), we have (where we implicitly sum over all indices)

$$\begin{aligned}
&[a_{ij}e_ie_j, b_{hk}e_he_k] = a_{ij}b_{hk}[e_ie_j, e_he_k] \\
&= -2a_{ij}b_{hk}\delta_{jh}e_ie_k + 2\delta_{jk}a_{ij}b_{hk}e_ie_h - 2\delta_{ih}a_{ij}b_{hk}e_ke_j + 2\delta_{ik}a_{ij}b_{hk}e_he_j \\
&= -2(AB)_{ik}e_ie_k - 2(AB)_{ih}e_ie_h + 2(BA)_{kj}e_ke_j + 2(BA)_{hj}e_he_j \\
&= -4[A,B]_{ij}e_ie_j.
\end{aligned}$$

Thus,

$$[-\tfrac{1}{4}\sum_{i,j} a_{ij}e_ie_j, -\tfrac{1}{4}\sum_{h,k} b_{hk}e_he_k] = -\tfrac{1}{4}\sum_{i,j}[A,B]_{ij}e_ie_j,$$

which implies that

$$c' : \mathcal{L}\left(\Lambda^2(V)\right) \cong \mathfrak{so}(n), \text{ given by } c'(-\tfrac{1}{4}\sum_{i,j} a_{ij}e_ie_j) := A \tag{18.3}$$

is an isomorphism $\mathcal{L}\left(\Lambda^2(V)\right) \cong \mathfrak{so}(n)$ of Lie algebras. We define

$$\mathrm{Spin}(n) := \exp\left(\mathcal{L}\left(\Lambda^2(V)\right)\right).$$

Before showing that

$$c : \mathrm{Spin}(n) \to \mathrm{SO}(n), \text{ given by } c(\exp x) := \exp(c'(x))$$

is a well defined double cover, we consider some examples.

EXAMPLE 18.6. For Cl_2, $\mathcal{L}\left(\Lambda^2(\mathbb{R}^2)\right) = \{te_1e_2 : t \in \mathbb{R}\}$, and

$$\begin{aligned}
\exp(te_1e_2) &= \sum_{k=1}^{\infty}\frac{1}{k!}(te_1e_2)^k \\
&= 1 + te_1e_2 + \tfrac{1}{2}t^2(e_1e_2)^2 + \tfrac{1}{6}t^2(e_1e_2)^3 + \cdots \\
&= 1 - \tfrac{1}{2}t^2 + \cdots + \left(t - \tfrac{1}{6}t^2 + \cdots\right)e_1e_2 = \cos(t) + \sin(t)e_1e_2.
\end{aligned}$$

Thus, $\mathrm{Spin}(2) = \{\cos(t) + \sin(t)e_1e_2 : t \in \mathbb{R}\}$.

EXAMPLE 18.7. For Cl_3, $\mathcal{L}\left(\Lambda^2(\mathbb{R}^3)\right) = \{a_1e_2e_3 + a_2e_3e_1 + a_3e_1e_2 : a_1, a_2, a_2 \in \mathbb{R}\}$. If $\mathbf{a} := a_1e_2e_3 + a_2e_3e_1 + a_3e_1e_2$, then

$$\begin{aligned}
\mathbf{a}^2 &= (a_1e_2e_3 + a_2e_3e_1 + a_3e_1e_2)^2 \\
&= a_1^2(e_2e_3)^2 + a_2^2(e_3e_1)^2 + a_3^2(e_1e_2)^2 \\
&+ a_1a_2(e_2e_3e_3e_1 + e_3e_1e_2e_3) + \cdots \\
&= -\left(a_1^2 + a_2^2 + a_3^2\right) = -\|\mathbf{a}\|^2.
\end{aligned}$$

Thus, with $\frac{\sin t}{t} := 1 - \frac{1}{6}t^2 + \cdots$ (analytic), it follows that

$$\begin{aligned}
\exp(\mathbf{a}) &= \cos(\|\mathbf{a}\|) + \left(\mathbf{a} - \tfrac{1}{6}\mathbf{a}^3 + \cdots\right) \\
&= \cos(\|\mathbf{a}\|) + \frac{\sin(\|\mathbf{a}\|)}{\|\mathbf{a}\|}\mathbf{a}, \text{ and}
\end{aligned}$$

$$\mathrm{Spin}(3) = \left\{\alpha_0 + \alpha_1e_2e_3 + \alpha_2e_3e_1 + \alpha_3e_1e_2 : \sum\nolimits_{k=0}^{3}\alpha_k^2 = 1\right\},$$

which may be regarded as the 3-sphere of unit quaternions.

EXAMPLE 18.8. To exhibit Spin (4), it is convenient to utilize the duality decomposition $\Lambda^2(\mathbb{R}^4) = \Lambda^+ \oplus \Lambda^-$. Then $\mathcal{L}(\Lambda^2(\mathbb{R}^4)) = \mathcal{L}(\Lambda^+) \oplus \mathcal{L}(\Lambda^-)$ and we set

$$\begin{aligned}\mathbf{a} &= a_1\tfrac{1}{2}(e_2e_3 + e_1e_4) + a_2\tfrac{1}{2}(e_3e_1 + e_2e_4) + a_3\tfrac{1}{2}(e_1e_2 + e_3e_4) \in \mathcal{L}(\Lambda^+)\\ \mathbf{b} &= b_1\tfrac{1}{2}(e_2e_3 - e_1e_4) + b_2\tfrac{1}{2}(e_3e_1 - e_2e_4) + b_3\tfrac{1}{2}(e_1e_2 - e_3e_4) \in \mathcal{L}(\Lambda^-).\end{aligned}$$

Under the isomorphism $F : Cl_4 \cong \mathbb{H}(2)$ determined by (18.1), it is easy to check that

$$\begin{aligned}F(\mathbf{a}+\mathbf{b}) &= \begin{bmatrix} a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k} & 0 \\ 0 & b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k} \end{bmatrix}, \text{ and so}\\ F(\exp(\mathbf{a}+\mathbf{b})) &= \begin{bmatrix} \exp(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}) & 0 \\ 0 & \exp(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}) \end{bmatrix}\\ &= \begin{bmatrix} \cos(\|\mathbf{a}\|) + \frac{\sin(\|\mathbf{a}\|)}{\|\mathbf{a}\|}\mathbf{a} & 0 \\ 0 & \cos(\|\mathbf{b}\|) + \frac{\sin(\|\mathbf{b}\|)}{\|\mathbf{b}\|}\mathbf{b} \end{bmatrix}.\end{aligned}$$

Thus, we have $F : \mathrm{Spin}(4) \cong S^3 \times S^3$. To delineate Spin (4) itself, first note that for $v_4 = e_1e_2e_3e_4$,

$$F(v_4) = F(e_1e_2e_3e_4) = \begin{bmatrix} -1 & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}.$$

Hence, Spin (4) consists of all elements of Cl_4 of the form

$$\begin{aligned}&\exp(\mathbf{a}+\mathbf{b}) = F^{-1}(F(\exp(\mathbf{a}+\mathbf{b})))\\ &= \tfrac{1}{2}(1-\nu_4)\cos\|\mathbf{a}\| + \frac{\sin(\|\mathbf{a}\|)}{\|\mathbf{a}\|}\mathbf{a} + \tfrac{1}{2}(1+\nu_4)\cos\|\mathbf{b}\| + \frac{\sin(\|\mathbf{b}\|)}{\|\mathbf{b}\|}\mathbf{b}\\ &= \tfrac{1}{2}(\cos\|\mathbf{a}\| + \cos\|\mathbf{b}\|) + \frac{\sin(\|\mathbf{a}\|)}{\|\mathbf{a}\|}\mathbf{a}\\ &+ \frac{\sin(\|\mathbf{b}\|)}{\|\mathbf{b}\|}\mathbf{b} + \tfrac{1}{2}(\cos\|\mathbf{b}\| - \cos\|\mathbf{a}\|)\nu_4.\end{aligned}$$

Note that here $\exp(\mathbf{a})\exp(\mathbf{b}) = \exp(\mathbf{a}+\mathbf{b}) = \exp(\mathbf{a}) + \exp(\mathbf{b})$!

The "vector representation" or double cover $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ is defined by means of the next result. Here $\dim(V) = n$, and $\mathcal{L}(\Lambda^1(V)) \subset Cl(V)$ are identified.

PROPOSITION 18.9. *For $g \in \mathrm{Spin}(n)$ and $v \in V = \mathcal{L}(\Lambda^1(V))$, let*

$$c(g)(v) := gvg^{-1} \in Cl(V). \tag{18.4}$$

Then $c(g)(v) \in \mathcal{L}(\Lambda^1(V)) = V$. Also, $c(g) \in \mathrm{SO}(n) := \mathrm{SO}(V)$ and

$$c : \mathrm{Spin}(n) \to \mathrm{SO}(n). \tag{18.5}$$

is a double covering homomorphism (universal for $n \geq 3$). Moreover, for c' defined as in (18.3), we have

$$c(\exp a) = \exp(c'(a)), \tag{18.6}$$

and so c' is the Lie algebra homomorphism for c.

PROOF. We write $g = \exp(a) = \exp\left(-\frac{1}{4}\sum_{i,j} a_{ij}e_ie_j\right)$ for $a \in \mathcal{L}(\Lambda^2(V))$. To show that $gvg^{-1} \in V$, it suffices to verify that

$$\exp(ta)\,v\exp(-ta) = \exp(tA)\,v, \tag{18.7}$$

where $A = c'(a) \in \mathfrak{so}(n)$. Since each side of (18.7) is a $Cl(V)$-valued power series in t with infinite radius of convergence, we need only check that all derivatives of both sides agree at $t = 0$; i.e.,

$$\frac{d^k}{dt^k}\left(\exp(ta)\, v \exp(-ta)\right)\Big|_{t=0} = A^k(v)\,,\ \ k = 0, 1, 2, \ldots, \tag{18.8}$$

In verifying this, we will use the identity

$$\begin{aligned} e_i e_j e_k - e_k e_i e_j &= e_i\left(-e_k e_j - 2\delta_{kj}\right) - e_k e_i e_j = -e_i e_k e_j - 2\delta_{kj} e_i - e_k e_i e_j \\ &= -\left(-e_k e_i - 2\delta_{ki}\right) e_j - 2\delta_{kj} e_i - e_k e_i e_j \\ &= 2\left(\delta_{ki} e_j - \delta_{kj} e_i\right). \end{aligned}$$

At $t = 0$,

$$\begin{aligned} &\tfrac{d}{dt}\left(\exp(ta)\, v \exp(-ta)\right) = [a, v] = -\tfrac{1}{4}\sum_{i,j} a_{ij}\left[e_i e_j, v\right] \\ &= -\tfrac{1}{4}\sum_{i,j,k} a_{ij}\left[e_i e_j, v_k e_k\right] = -\tfrac{1}{4}\sum_{i,j,k} a_{ij} v_k \left(e_i e_j e_k - e_k e_i e_j\right) \\ &= -\tfrac{1}{4}\sum_{i,j,k} a_{ij} v_k 2\left(\delta_{ki} e_j - \delta_{kj} e_i\right) = -\tfrac{1}{2}\sum_{i,j,k}\left(a_{ij} v_k \delta_{ki} e_j - a_{ij} v_k \delta_{kj} e_i\right) \\ &= -\tfrac{1}{2}\sum_{i,j}\left(a_{ij} v_i e_j - a_{ij} v_j e_i\right) = \sum_{i,j} a_{ij} v_j e_i = A(v)\,. \end{aligned}$$

At arbitary t,

$$\begin{aligned} &\tfrac{d}{dt}\left(\exp(ta)\, v \exp(-ta)\right) = \tfrac{d}{du}\left(\exp\left((t+u)\, a\right) v \exp\left(-\left(t+u\right) a\right)\right)\big|_{u=0} \\ &= \exp(ta)\, \tfrac{d}{du}\left(\exp(ua)\, v \exp(-ua)\right)\big|_{u=0} \exp(-ta) \\ &= \exp(ta)\, A(v) \exp(-ta)\,. \end{aligned}$$

Hence,

$$\tfrac{d^k}{dt^k}\left(\exp(ta)\, v \exp(-ta)\right) = \exp(ta)\, A^k(v) \exp(-ta)\,,$$

and evaluating both sides at $t = 0$ yields (18.8).

Now, $c(g) \in \mathrm{SO}(n)$, since

$$\begin{aligned} -\left\|c(g)\, v\right\|^2 = -\left\|g v g^{-1}\right\|^2 &= \left(g v g^{-1}\right)\left(g v g^{-1}\right) \\ &= g v v g^{-1} = -\left\|v\right\|^2 g^{-1} g = -\left\|v\right\|^2. \end{aligned}$$

Thus, $c(\mathrm{Spin}(n)) \subseteq \mathrm{SO}(n)$. Note that (18.6) does in fact hold by (18.7) with $t = 1$. Since $\mathrm{Spin}(n)$ and $\mathrm{SO}(n)$ are connected and $c' : \mathfrak{spin}(n) \to \mathfrak{so}(n)$ is an isomorphism, it follows that $c(\mathrm{Spin}(n))$ is the connected component of $I \in \mathrm{SO}(n)$, namely $\mathrm{SO}(n)$ itself, and $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ is a covering homomorphism. Since $\pi_1(\mathrm{SO}(n)) \cong \mathbb{Z}_2$ for $n > 2$, it follows from covering space theory that $\pi_1(\mathrm{Spin}(n)) \cong \pi_1(\mathrm{SO}(n)) / \operatorname{Ker} c$. Then for $n > 2$, $\mathrm{Spin}(n)$ is the universal, (simply-connected) covering space of $\mathrm{SO}(n)$ if $\pm 1 \in \operatorname{Ker} c$. Certainly, $1 \in \operatorname{Ker} c$, and if $-1 \in \mathrm{Spin}(n)$, then $-1 \in \operatorname{Ker} c$, since $c(-1)(v) = -1 v (-1)^{-1} = v$. Thus, it remains to check that $-1 \in \mathrm{Spin}(n)$, but this is immediate from $\exp(t e_1 e_2) = \cos(t) + \sin(t)\, e_1 e_2$ with $t = \pi$. For $n = 2$, $c : \mathrm{Spin}(2) \to \mathrm{SO}(2)$ is still a double cover, since

$$\begin{aligned}
&c(\exp(te_1e_2))(e_1) = \exp(te_1e_2)\, e_1 \exp(-te_1e_2) \\
&= (\cos(t) + \sin(t)\, e_1e_2)\, e_1 (\cos(t) - \sin(t)\, e_1e_2) \\
&= \cos^2(t)\, e_1 - \cos(t)\sin(t)\, e_1e_1e_2 \\
&+ \sin(t)\cos(t)\, e_1e_2e_1 - \sin^2(t)\, e_1e_2e_1e_1e_2 \\
&= \left(\cos^2(t) - \sin^2(t)\right) e_1 + 2\cos(t)\sin(t)\, e_2 \\
&= \cos(2t)\, e_1 + \sin(2t)\, e_2.
\end{aligned}$$

However, $\pi_1(\mathrm{SO}(2)) \cong \mathbb{Z}$ so that the covering is not universal for $n = 2$. □

Besides the vector representation $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$, there are fundamental spinor representations, which we will describe. Since the index of an elliptic operator on a compact, odd-dimensional manifold is always 0, for simplicity we assume that n is even, say $n = 2m$. Then there is a unique (up to equivalence) irreducible representation (homomorphism of algebras over $\mathbb{R}$)

$$\rho : Cl_{2m} \to \mathrm{End}(\Sigma_{2m}),$$

where $\mathrm{End}(\Sigma_{2m})$ is the algebra of $\mathbb{C}$-linear endomorphisms of some complex vector space Σ_{2m}, the elements of which are called *spinors*. Here "irreducible" means that Σ_{2m} has no proper subspace which is invariant under all operators in $\rho(Cl_{2m})$. In the following we give an explicit construction of Σ_{2m} and ρ.

Let $\langle \cdot, \cdot \rangle$ be the standard Hermitian inner product on $\mathbb{C}^m$ given by

$$\langle z, w \rangle = \sum\nolimits_{k=1}^{m} z_k \overline{w_k}.$$

We identify $\mathbb{C}^m$ with $\mathbb{R}^n = \mathbb{R}^{2m}$, and for $w \in \mathbb{C}^m$, we have $\mathbb{C}$-linear function

$$w\wedge : \Lambda^k(\mathbb{C}^m) \to \Lambda^{k+1}(\mathbb{C}^m)$$

given by $\alpha \mapsto w \wedge \alpha$ for $\alpha \in \Lambda^k(\mathbb{C}^m)$. Moreover, there is a $\mathbb{C}$-linear function

$$w\llcorner : \Lambda^k(\mathbb{C}^m) \to \Lambda^{k-1}(\mathbb{C}^m), \text{ defined via}$$

$$w\llcorner(v_1 \wedge \cdots \wedge v_k) := \sum_{j=1}^{k} (-1)^{j+1} \langle v_j, w \rangle\, v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_k,$$

where $\widehat{v_j}$ means that the factor v_j is omitted. While $w\llcorner$ is $\mathbb{C}$-linear, the function $\mathbb{C}^m \to \mathrm{End}(\Lambda^*(\mathbb{C}^m))$ given by $w \mapsto w\llcorner$ is $\mathbb{R}$-linear (but $\mathbb{C}$-conjugate linear). There is a Hermitian inner product $\langle \cdot, \cdot \rangle$ on $\Lambda^k(\mathbb{C}^m)$ induced by that on $\mathbb{C}^m$, such that $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 < \cdots < i_k\}$ is an orthonormal basis for $\Lambda^k(\mathbb{C}^m)$ if $e_1, \ldots, e_m$ is an orthonormal basis for $\mathbb{C}^m$. Relative to this inner product, $w\llcorner$ and $w\wedge$ are adjoints, since for all $v_1, \cdots, v_k, u_1, \cdots, u_{k-1} \in \mathbb{C}^m$ we have

$$\begin{aligned}
&\langle v_1 \wedge \cdots \wedge v_k, w \wedge u_1 \wedge \cdots \wedge u_{k-1} \rangle \\
&= \sum_{j=1}^{k} (-1)^{j+1} \langle v_j, w \rangle \langle v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_k, u_1 \wedge \cdots \wedge u_{k-1} \rangle \\
&= \left\langle \sum\nolimits_{j=1}^{k} (-1)^{j+1} \langle v_j, w \rangle\, v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_k, u_1 \wedge \cdots \wedge u_{k-1} \right\rangle \\
&= \langle w\llcorner(v_1 \wedge \cdots \wedge v_k), u_1 \wedge \cdots \wedge u_{k-1} \rangle .
\end{aligned}$$

Proposition 18.10. *Let $\rho_1 : \mathbb{C}^m \to \operatorname{End}(\Lambda^*(\mathbb{C}^m))$ be given by*

$$\rho_1(w)(\alpha) := (w \wedge - w\llcorner)(\alpha) = w \wedge \alpha - w\llcorner\alpha.$$

Then ρ_1 uniquely extends to an $\mathbb{R}$-linear homomorphism

$$\rho : Cl_{2m} \to \operatorname{End}(\Lambda^*(\mathbb{C}^m)),$$

of algebras over $\mathbb{R}$.

Proof. Using $w \wedge w \wedge \alpha = 0$ and $w\llcorner(w\llcorner\alpha) = 0$, we obtain

$$\begin{aligned}(\rho_1(w) \circ \rho_1(w))(\alpha) &= w \wedge (w \wedge \alpha - w\llcorner\alpha) - w\llcorner(w \wedge \alpha - w\llcorner\alpha)\\ &= w \wedge w \wedge \alpha - w \wedge (w\llcorner\alpha) - w\llcorner(w \wedge \alpha) + w\llcorner(w\llcorner\alpha)\\ &= -w \wedge (w\llcorner\alpha) - w\llcorner(w \wedge \alpha)\\ &= -\langle w, w\rangle \alpha, \end{aligned} \tag{18.9}$$

where in the last equality we used

$$\begin{aligned}&w\llcorner(w \wedge (v_1 \wedge \cdots \wedge v_k))\\ &= \langle w, w\rangle v_1 \wedge \cdots \wedge v_k - \sum\nolimits_{j=1}^{k} (-1)^{j+1} \langle v_j, w\rangle w \wedge v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_k\\ &= \langle w, w\rangle v_1 \wedge \cdots \wedge v_k - w \wedge (w\llcorner(v_1 \wedge \cdots \wedge v_k)).\end{aligned}$$

It follows from (18.9) that

$$\begin{aligned}&\rho_1(w_1) \circ \rho_1(w_2) + \rho_1(w_1) \circ \rho_1(w_2)\\ &= -(\langle w_1, w_2\rangle + \langle w_2, w_1\rangle) \operatorname{Id} = -2 \operatorname{Re}\langle w_1, w_2\rangle \operatorname{Id}.\end{aligned}$$

Since $\operatorname{Re}\langle w_1, w_2\rangle$ is the standard inner product on $\mathbb{R}^{2m} \cong \mathbb{C}^m$, ρ_1 extends uniquely to an $\mathbb{R}$-linear homomorphism

$$\rho : Cl_{2m} \to \operatorname{End}(\Lambda^*(\mathbb{C}^m)),$$

of algebras over $\mathbb{R}$. □

Let $\mathbb{C}l_{2m} := \mathbb{C} \otimes_{\mathbb{R}} Cl_{2m}$ be the complex Clifford algebra. One of our goals is to prove that the complex linear extension of ρ, say

$$\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}(\Lambda^*(\mathbb{C}^m)),$$

is an isomorphism of algebras over $\mathbb{C}$. For this and other reasons, it is convenient to introduce more notation. Let $(f_1, \cdots, f_m)$ be an orthonormal basis of $\mathbb{C}^m$, then

$$(e_1, \cdots, e_{2m}) := (f_1, if_1, \cdots, f_m, if_m) \tag{18.10}$$

is an oriented, orthonormal basis of $\mathbb{R}^{2m}$, and the *complex volume element* is

$$\omega_{\mathbb{C}} := i^m e_1 \cdots e_{2m} \in \mathbb{C}l_{2m}; \tag{18.11}$$

this is independent of the choice of oriented orthonormal basis of $\mathbb{R}^{2m}$. We have

$$\begin{aligned}\omega_{\mathbb{C}}^2 &= (-1)^m e_1 \cdots e_{2m} e_1 \cdots e_{2m}\\ &= (-1)^m (-1)^{2m} (-1)^{(2m-1)+(2m-2)+\cdots+1}\\ &= (-1)^{m+2m(2m-1)/2} = (-1)^{m+m(2m-1)} = (-1)^{2m^2} = 1, \text{ and so}\end{aligned}$$

$$\rho_{\mathbb{C}}(\omega_{\mathbb{C}})^2 = \rho_{\mathbb{C}}(\omega_{\mathbb{C}}^2) = \rho_{\mathbb{C}}(1) = \operatorname{Id}. \tag{18.12}$$

Since $\rho_1(w) = w\wedge - w\llcorner$ is the difference between an operator and its adjoint, $\rho_1(w)$ is skew-adjoint. We define skew-adjoint operators

$$\gamma_j := \rho_{\mathbb{C}}(e_j) = \rho(e_j)\,, \text{ for } j \in \{1,\dots,2m\}\,.$$

In harmony with the physics literature, we set $\gamma_{n+1} = \gamma_{2m+1} := \gamma_1 \cdots \gamma_n$, so that $\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) = i^m \gamma_{2m+1}$. The case $n = 4$ abounds in physics, as does $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4$, but due to the indefiniteness of the metric in relativity, $\gamma_4^2 = -\gamma_i^2$ for $i = 1,2,3$. We will remain in the Euclidean category. Using the fact that the γ_j are skew-adjoint, it is straightforward (but tedious) to show directly that $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is self-adjoint. But, we can easily get this from the fact that $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is clearly either self-adjoint or skew-adjoint, and, *unlike* $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$, the square of a skew-adjoint transformation has nonpositive eigenvalues (squares of pure imaginaries). Since $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})^2 = \mathrm{Id}$, the eigenvalues of $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ are ± 1. As $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is self-adjoint, the eigenspaces of $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$, say

$$\Sigma_{2m}^{+} := (\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) + \mathrm{Id})\,\Lambda^*(\mathbb{C}^m) \text{ and } \Sigma_{2m}^{-} := (\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) - \mathrm{Id})\,\Lambda^*(\mathbb{C}^m)\,, \tag{18.13}$$

are orthogonal. Using the notation (18.10),

$$\begin{aligned}
\rho(e_1e_2) &= (e_1\wedge - e_1\llcorner)\circ(e_2\wedge - e_2\llcorner)\\
&= (f_1\wedge - f_1\llcorner)\circ(if_1\wedge - (if_1)\llcorner)\\
&= (f_1\wedge - f_1\llcorner)\circ(if_1\wedge + i(f_1)\llcorner)\\
&= (i(f_1\wedge)\circ(f_1\llcorner) - i(f_1\llcorner)\circ(f_1\wedge))\\
&= i((f_1\wedge)\circ(f_1\llcorner) - (f_1\llcorner)\circ(f_1\wedge))\,, \text{ and in general}\\
\rho(e_{2j-1}e_{2j}) &= i((f_j\wedge)\circ(f_j\llcorner) - (f_j\llcorner)\circ(f_j\wedge))\,.
\end{aligned}$$

Thus, the action of $\rho(e_{2j-1}e_{2j})$ on $f_{j_1}\wedge f_{j_2}\wedge\cdots\wedge f_{j_l}$ is given by

$$\begin{aligned}
&\rho(e_{2j-1}e_{2j})(f_{j_1}\wedge f_{j_2}\wedge\cdots\wedge f_{j_l})\\
&= \begin{cases} i(f_{j_1}\wedge f_{j_2}\wedge\cdots\wedge f_{j_l}) & \text{if } j_k = j \text{ for some } k\\ -i(f_{j_1}\wedge f_{j_2}\wedge\cdots\wedge f_{j_l}) & \text{if } j_k \neq j \text{ for all } k.\end{cases}
\end{aligned} \tag{18.14}$$

Hence, when restricted to $\Lambda^l(\mathbb{C}^m)$, $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is

$$i^m\rho(e_1\cdots e_{2m})\,|_{\Lambda^l(\mathbb{C}^m)} = i^m i^l(-i)^{m-l}\,\mathrm{Id} = (-1)^l\,\mathrm{Id}\,, \text{ and so}$$

$$\begin{aligned}
\Sigma_{2m}^{+} &= \Lambda^{\text{even}}(\mathbb{C}^m) := \bigoplus\nolimits_{l \text{ even}} \Lambda^l(\mathbb{C}^m)\,, \text{ while}\\
\Sigma_{2m}^{-} &= \Lambda^{\text{odd}}(\mathbb{C}^m) := \bigoplus\nolimits_{l \text{ odd}} \Lambda^l(\mathbb{C}^m)\,.
\end{aligned}$$

Using this (or the easy fact that for $j \in \{1,\dots,2m\}$, $\gamma_{2m+1}\circ\gamma_j = -\gamma_{2m+1}\circ\gamma_j$), we have $\gamma_j\left(\Sigma_{2m}^{\pm}\right) = \Sigma_{2m}^{\mp}$, and indeed $\gamma_j : \Sigma_{2m}^{\pm} \cong \Sigma_{2m}^{\mp}$ with inverse $-\gamma_j$. Moreover, for $j,k \in \{1,\dots,2m\}$, the spaces $\Sigma_{2m}^{\pm}$ are each invariant under the compositions $\gamma_j\circ\gamma_k$, and so $\rho(\mathfrak{spin}(2m))\left(\Sigma_{2m}^{\pm}\right) \subset \Sigma_{2m}^{\pm}$. For $j \neq k$,

$$\begin{aligned}
(\gamma_j\circ\gamma_k)^* &= \gamma_k^*\circ\gamma_j^* = -\gamma_k\circ-\gamma_j = -\gamma_j\circ\gamma_k \text{ and}\\
\mathrm{Tr}(\gamma_j\circ\gamma_k) &= -\mathrm{Tr}(\gamma_k\circ\gamma_j) = -\mathrm{Tr}(\gamma_j\circ\gamma_k)\,.
\end{aligned}$$

Thus, the elements of $\rho(\mathfrak{spin}(2m))$ are skew-adjoint and traceless, and so $\rho(\mathrm{Spin}(2m)) \subset \mathrm{SU}(\Lambda^*(\mathbb{C}^m))$. In summary, $\rho : \mathrm{Spin}(2m) \to \mathrm{SU}(\Lambda^*(\mathbb{C}^m))$ is the orthogonal direct sum of two special unitary "half-spinor" or "chiral" representations

$$\rho^{\pm} : \mathrm{Spin}(2m) \to \mathrm{SU}\left(\Sigma_{2m}^{\pm}\right)\,. \tag{18.15}$$

DEFINITION 18.11. Let

$$\Sigma_{2m} := \Lambda^* \left(\mathbb{C}^m\right) \text{ and } \pi^{\pm} := \tfrac{1}{2} \left(\rho_{\mathbb{C}} \left(\omega_{\mathbb{C}}\right) \pm \mathrm{Id}\right) : \Sigma_{2m} \to \Sigma_{2m}^{\pm}.$$

The **supertrace** of an endomorphism $A \in \mathrm{End}\left(\Sigma_{2m}\right)$ is

$$\begin{aligned} \mathrm{Str}\,(A) &:= Tr\left(\pi^{+} \circ \left(A|_{\Sigma_{2m}^{+}}\right)\right) - Tr\left(\pi^{-} \circ \left(A|_{\Sigma_{2m}^{-}}\right)\right) \\ &= Tr\left(A \circ \rho_{\mathbb{C}}\left(\omega_{\mathbb{C}}\right)\right) = i^m Tr\left(A \circ \gamma_{2m+1}\right). \end{aligned}$$

The following result will be crucial in evaluating the local index density of the twisted Dirac operator.

PROPOSITION 18.12. *For $k \in \{1, \ldots, 2m\}$ with $j_1, j_2, \ldots, j_k$ distinct, we have*

$$\begin{aligned} &\mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = 0, \text{ and} \\ &\mathrm{Str}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = i^m\, \mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\gamma_{2m+1}\right) \\ &\qquad = \begin{cases} 0 & \text{if } k < 2m \\ (-2i)^m \varepsilon_{j_1\cdots j_{2m}} & \text{if } k = 2m. \end{cases} \end{aligned}$$

The 2^n endomorphisms consisting of Id *and those $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}$ with $j_1 < j_2 < \cdots < j_k$, $k \in \{1, \ldots, 2m\}$, form a basis of* $\mathrm{End}\left(\Sigma_{2m}\right)$.

PROOF. Since $\gamma_j\left(\Sigma_{2m}^{\pm}\right) \subseteq \Sigma_{2m}^{\mp}$, we have $\mathrm{Tr}\left(\gamma_j\right) = 0$. More generally, for k odd and $j_1, j_2, \ldots, j_k$ distinct, we have $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\left(\Sigma_{2m}^{\pm}\right) \subseteq \Sigma_{2m}^{\mp}$ and $\mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = 0$. For k even and $j_1, j_2, \ldots, j_k$ distinct, we have

$$\begin{aligned} &\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k} = (-1)^{k-1}\gamma_{j_2}\cdots\gamma_{j_k}\gamma_{j_1} = -\left(\gamma_{j_2}\cdots\gamma_{j_k}\right)\gamma_{j_1}, \text{ and so} \\ &\mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = -\mathrm{Tr}\left(\left(\gamma_{j_2}\cdots\gamma_{j_k}\right)\gamma_{j_1}\right) = -\mathrm{Tr}\left(\gamma_{j_1}\left(\gamma_{j_2}\cdots\gamma_{j_k}\right)\right) = 0. \end{aligned}$$

If $k < 2m$, $j_1, j_2, \ldots, j_k$ are distinct, and the complementary set of indices is $\{h_1, \ldots, h_{2m-k}\} := \{1, \ldots, 2m\} - \{j_1, j_2, \ldots, j_k\}$, then

$$\begin{aligned} \mathrm{Str}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) &= i^m\, \mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\gamma_{2m+1}\right) \\ &= \pm i^m\, \mathrm{Tr}\left(\gamma_{h_1}\gamma_{h_2}\cdots\gamma_{h_{2m-k}}\right) = 0. \end{aligned}$$

If $k = 2m$ and $j_1, j_2, \ldots, j_k$ are distinct, then

$$\begin{aligned} &\mathrm{Str}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_{2m}}\right) = \varepsilon_{j_1\cdots j_{2m}}\, \mathrm{Str}\left(\gamma_1\gamma_2\cdots\gamma_{2m}\right) \\ &= \varepsilon_{j_1\cdots j_{2m}} i^m\, \mathrm{Tr}\left(\left(\gamma_{2m+1}\right)^2\right) = \varepsilon_{j_1\cdots j_{2m}} i^m\, \mathrm{Tr}\left(\left(i^{-m}\rho_{\mathbb{C}}\left(\omega_{\mathbb{C}}\right)\right)^2\right) \\ &= \varepsilon_{j_1\cdots j_{2m}} i^m (-1)^m\, \mathrm{Tr}\,(\mathrm{Id}) = \varepsilon_{j_1\cdots j_{2m}} (-i)^m\, 2^m = \varepsilon_{j_1\cdots j_{2m}} (-2i)^m. \end{aligned}$$

Since $\dim \mathrm{End}\left(\Sigma_{2m}\right) = \left(\dim \Sigma_{2m}\right)^2 = 2^{2m}$, the 2^n endomorphisms consisting of Id and the $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}$ with $j_1 < j_2 < \cdots < j_k$ will form a basis of $\mathrm{End}\left(\Sigma_{2m}\right)$, if they are shown to be independent. For this, let

$$c_0 I + \sum\nolimits_{j_1 < j_2 < \cdots < j_k} c_{j_1} \ldots_{j_k}\, \gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k} = 0,$$

for $c_0, c_{j_1}\dots{}_{j_k} \in \mathbb{C}$ and let $\{h_1, \dots, h_{2m-k}\} := \{1, \dots, 2m\} - \{l_1, l_2, \dots, l_k\}$ for some $l_1 < l_2 < \dots < l_k$. Then

$$\begin{aligned} 0 &= \operatorname{Str}\left(\gamma_{h_1}\gamma_{h_2}\cdots\gamma_{h_{2m-k}}\left(c_0 I + \sum\nolimits_{j_1<j_2<\cdots<j_k} c_{j_1}\dots{}_{j_k}\,\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right)\right) \\ &= c_{l_1}\dots{}_{l_k}\operatorname{Str}\left(\gamma_{h_1}\gamma_{h_2}\cdots\gamma_{h_{2m-k}}\gamma_{l_1}\gamma_{l_2}\cdots\gamma_{l_k}\right) = \pm\left(-2i\right)^m c_{l_1}\dots{}_{l_k}\,, \end{aligned}$$

with the convention that $c_{l_1}\dots{}_{l_k} = c_0$ if $k = 0$. □

COROLLARY 18.13. *The representation $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}(\Sigma_{2m})$ is irreducible and an isomorphism of complex algebras. Moreover, the representations $\rho^{\pm} : \operatorname{Spin}(2m) \to \operatorname{End}\left(\Sigma_{2m}^{\pm}\right)$ are irreducible and inequivalent.*

PROOF. The last statement of Proposition 18.12 implies that $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}(\Sigma_{2m})$ is an isomorphism, so that in particular $\rho_{\mathbb{C}}(\mathbb{C}l_{2m}) = \operatorname{End}(\Sigma_{2m})$. As $\operatorname{End}(\Sigma_{2m})$ acts transitively on the set of subspaces of Σ_{2m} of a given dimension, $\operatorname{End}(\Sigma_{2m})$ leaves no proper subspace of Σ_{2m} invariant, and hence $\rho_{\mathbb{C}}$ is irreducible. Let

$$\begin{aligned} \operatorname{End}^0(\Sigma_{2m}) &:= \left\{A \in \operatorname{End}(\Sigma_{2m}) : A\left(\Sigma_{2m}^{\pm}\right) \subseteq \Sigma_{2m}^{\pm}\right\} \text{ and} \\ \operatorname{End}^1(\Sigma_{2m}) &:= \left\{A \in \operatorname{End}(\Sigma_{2m}) : A\left(\Sigma_{2m}^{\pm}\right) \subseteq \Sigma_{2m}^{\mp}\right\}. \end{aligned}$$

The linear isomorphism $\mathcal{L} : \Lambda^*\left(\mathbb{R}^{2m}\right) \cong Cl_{2m}$ extends to $\mathcal{L}_{\mathbb{C}} : \Lambda^*\left(\mathbb{C}^{2m}\right) \cong \mathbb{C}l_{2m}$. We have

$$\begin{aligned} &\mathbb{C}l_{2m} = \mathcal{L}_{\mathbb{C}}\left(\Lambda^{\text{even}}\left(\mathbb{C}^{2m}\right)\right) \oplus \mathcal{L}_{\mathbb{C}}\left(\Lambda^{\text{odd}}\left(\mathbb{C}^{2m}\right)\right), \\ &\rho_{\mathbb{C}}\left(\mathcal{L}\left(\Lambda^{\text{even}}\left(\mathbb{C}^{2m}\right)\right)\right) \subseteq \operatorname{End}^0(\Sigma_{2m}), \text{ and} \\ &\rho_{\mathbb{C}}\left(\mathcal{L}\left(\Lambda^{\text{odd}}\left(\mathbb{C}^{2m}\right)\right)\right) \subseteq \operatorname{End}^1(\Sigma_{2m}). \end{aligned}$$

As $\rho_{\mathbb{C}}$ is an isomorphism, these last two inclusions are equalities. In particular, the restrictions

$$\pi^{\pm} \circ \rho_{\mathbb{C}} : \mathcal{L}\left(\Lambda^{\text{even}}\left(\mathbb{C}^{2m}\right)\right) \to \operatorname{End}\left(\Sigma_{2m}^{\pm}\right)$$

are irreducible representations of the subalgebra $\mathcal{L}\left(\Lambda^{\text{even}}\left(\mathbb{C}^{2m}\right)\right) \subset \mathbb{C}l_{2m}$. Since $\mathcal{L}\left(\Lambda^{\text{even}}\left(\mathbb{C}^{2m}\right)\right)$ is generated by $\mathcal{L}\left(\Lambda^2\left(\mathbb{R}^{2m}\right)\right) = \mathfrak{spin}(2m)$, there is also no proper subspace of Σ_{2m}^+ or Σ_{2m}^- which is invariant under $\mathfrak{spin}(2m)$ or under $\operatorname{Spin}(2m) = \exp(\mathfrak{spin}(2m))$. Thus, $\rho^{\pm} : \operatorname{Spin}(2m) \to \operatorname{SU}\left(\Sigma_{2m}^{\pm}\right)$ are irreducible representations of $\operatorname{Spin}(2m)$. Using the computation in Example 18.6, each of $e_1e_2, \dots, e_{2m-1}e_{2m}$ are in $\operatorname{Spin}(2m)$. Hence $e_1e_2\cdots e_{2m-1}e_{2m} \in \operatorname{Spin}(2m)$. Since $\rho\left(e_1e_2\cdots e_{2m-1}e_{2m}\right) = \gamma_{2m+1} = \pm i^{-m}$ on $\Sigma_{2m}^{\pm}$, the representations $\pi^+ \circ \rho$ and $\pi^- \circ \rho$ are inequivalent. □

PROPOSITION 18.14. *Let $R : \mathbb{C}l_{2m} \to \operatorname{End}(V)$ be a finite-dimensional representation. Then $V = \bigoplus_{k=1}^{N} W_k$ where $W_1, \dots, W_N$ of V are invariant subspaces, such that $R_k : \mathbb{C}l_{2m} \to \operatorname{End}(W_k)$ defined by $R_k(\alpha) = R(\alpha)|_{W_k}$ is equivalent to $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}(\Lambda^*(\mathbb{C}^m))$ for all $k = \{1, \dots, N\}$. In particular, all irreducible representations of $\mathbb{C}l_{2m}$ are equivalent to $\rho_{\mathbb{C}}$. Moreover, let*

$$\operatorname{Hom}_0(\Sigma_{2m}, V) := \left\{F \in \operatorname{Hom}(\Sigma_{2m}, V) : F\left(\rho_{\mathbb{C}}(\alpha)(w)\right) = R(\alpha)\left(F(w)\right)\right\}$$

be the subspace of $\operatorname{Hom}(\Sigma_{2m}, V)$ of $\mathbb{C}l_{2m}$-equivariant linear maps; note that $\mathbb{C}l_{2m}$ acts trivially on $\operatorname{Hom}_0(\Sigma_{2m}, V)$. There is then an isomorphism of $\mathbb{C}l_{2m}$-modules

$$\Phi : \operatorname{Hom}_0(\Sigma_{2m}, V) \otimes \Sigma_{2m} \cong V \text{ given by } \Phi(\phi \otimes \psi) := \phi(\psi);$$

Proof. Let $e_1, \ldots, e_{2m}$ be an oriented, orthonormal basis for $\mathbb{R}^{2m}$. Let $\sigma_j = ie_{2j-1}e_{2j} \in \mathbb{C}l_{2m}$ for $j \in \{1, \ldots, m\}$. Note that $[e_{2j-1}e_{2j}, e_{2k-1}e_{2k}] = 0$ for all $j, k \in \{1, \ldots, m\}$, and so $[\sigma_j, \sigma_k] = 0$ and $[R(\sigma_j), R(\sigma_k)] = 0$. Since $\sigma_j^2 = ie_{2j-1}e_{2j}ie_{2j-1}e_{2j} = e_{2j-1}e_{2j-1}e_{2j}e_{2j} = 1$, the eigenvalues of $R(\sigma_j)$ are ± 1. Thus, there are simultaneous eigenspaces of the $R(\sigma_j)$, $j \in \{1, \ldots, m\}$, indexed by

$$\begin{aligned} &\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m) \in \mathbb{Z}_2^m := \{1, -1\} \times \overset{m}{\cdots} \times \{1, -1\}, \text{ namely} \\ &V(\varepsilon) := \{v \in V : R(\sigma_j)(v) = \varepsilon_j v\}. \end{aligned}$$

We have $V = \bigoplus_{\varepsilon \in \mathbb{Z}_2^m} V(\varepsilon)$. Let

$$\alpha(\varepsilon) = \tfrac{1}{2}(1 - \varepsilon_1) e_2 \cdots \tfrac{1}{2}(1 - \varepsilon_m) e_{2m} = \prod_{j=1}^{m} \tfrac{1}{2}(1 - \varepsilon_j) e_{2j} = \prod\nolimits_{\{j : \varepsilon_j = -1\}} e_{2j}.$$

Then

$$\sigma_k \alpha(\varepsilon) = ie_{2k-1}e_{2k}\alpha(\varepsilon) = \begin{cases} \alpha(\varepsilon)\sigma_k & \text{if } \varepsilon_k = 1 \\ -\alpha(\varepsilon)\sigma_k & \text{if } \varepsilon_k = -1 \end{cases} = \varepsilon_k \alpha(\varepsilon)\sigma_k$$

For $1_m := (1, \ldots, 1) \in \mathbb{Z}_2^m$, we claim

$$R(\alpha(\varepsilon)) : V(1_m) \to V(\varepsilon)$$

is a well-defined isomorphism. Indeed, for $v \in V(1_m)$, we have $R(\alpha(\varepsilon))v \in V(\varepsilon)$, since

$$\begin{aligned} R(\sigma_k)(R(\alpha(\varepsilon))v) &= R(\sigma_k\alpha(\varepsilon))v = R(\varepsilon_k\alpha(\varepsilon)\sigma_k)v \\ &= \varepsilon_k R(\alpha(\varepsilon))R(\sigma_k)v = \varepsilon_k(R(\alpha(\varepsilon))v), \end{aligned}$$

and $\alpha(\varepsilon)^2 = \pm 1 \Rightarrow R(\alpha(\varepsilon))^2 = \pm\,\mathrm{Id}$. Thus,

$$\dim V = \dim(\bigoplus\nolimits_{\varepsilon \in \mathbb{Z}_2^m} V(\varepsilon)) = 2^m \dim V(1_m).$$

Let $\{v_1, \ldots, v_N\}$ be a basis for $V(1_m)$ and for $k \in \{1, \ldots, N\}$, let

$$W_k = \operatorname{span}\{R(\alpha(\varepsilon))v_k : \varepsilon \in \mathbb{Z}_2^m\}.$$

We claim that W_k is a $\mathbb{C}l_{2m}$-module. For this first note that $\mathbb{C}l_{2m}$ is generated by

$$\{\alpha(\varepsilon) : \varepsilon \in \mathbb{Z}_2^m\} \cup \{\sigma_j : j \in \{1, \ldots, m\}\},$$

since for any $j \in \{1, \ldots, m\}$, $e_{2j} \in \{\alpha(\varepsilon) : \varepsilon \in \mathbb{Z}_2^m\}$, $\sigma_j e_{2j} = -e_{2j-1}$, and $\{e_1, \ldots, e_{2m}\}$ generate $\mathbb{C}l_{2m}$. Thus, it suffices to show that $R(\alpha(\varepsilon))(W_k) \subset W_k$ for all $\varepsilon \in \mathbb{Z}_2^m$, and $R(\sigma_j)(W_k) \subset W_k$. To this end,

$$\begin{aligned} &\varepsilon, \varepsilon' \in \mathbb{Z}_2^m \Rightarrow \alpha(\varepsilon')\alpha(\varepsilon) = \pm\alpha(\varepsilon'') \text{ for some } \varepsilon'' \in \mathbb{Z}_2^m \\ &\quad \Rightarrow R(\alpha(\varepsilon'))R(\alpha(\varepsilon))v_k = \pm R(\varepsilon'')v_k \\ &\quad \Rightarrow R(\alpha(\varepsilon))(W_k) \subset W_k \text{ for all } \varepsilon \in \mathbb{Z}_2^m\text{; moroever,} \\ &v_k \in V(1_m) \Rightarrow R(\sigma_j)R(\alpha(\varepsilon))v_k = \pm R(\alpha(\varepsilon))R(\sigma_j)v_k = \pm R(\alpha(\varepsilon))v_k \\ &\quad \Rightarrow R(\sigma_j)(W_k) \subset W_k. \end{aligned}$$

Thus, $V = \bigoplus_{k=1}^N W_k$. Note that $\dim W_k \le 2^m$, since $\{R(\alpha(\varepsilon))v_k : \varepsilon \in \mathbb{Z}_2^m\}$ spans W_k. Since W_k is a $\mathbb{C}l_{2m}$-module, the same proof as that of Proposition 18.12 yields that the 2^{2m} endomorphisms consisting of Id and those $R(e_{j_1} \cdots e_{j_k})$ with $1 \le j_1 < j_2 < \cdots < j_k \le 2m$, $k \in \{1, \ldots, 2m\}$, form a linearly independent subset of $\operatorname{End}(W_k)$. Since $\dim(\operatorname{End}(W_k)) = (\dim W_k)^2 = 2^{2m}$, we obtain $\operatorname{End}(W_k) = \{R(\alpha)|_{W_k} : \alpha \in \mathbb{C}l_{2m}\}$. Indeed

$$\mathbb{C}l_{2m} \cong \operatorname{End}(W_k) \text{ via } \alpha \mapsto R(\alpha)|_{W_k},$$

and hence each W_k is an irreducible $\mathbb{C}l_{2m}$-module and W_k is isomorphic to the specific module $\Lambda^*(\mathbb{C}^m)$. Note that $\Phi : \operatorname{Hom}_0(\Sigma_{2m}, V) \otimes \Sigma_{2m} \to V$ is indeed a morphism, since

$$\begin{aligned}\Phi(\alpha \cdot (\phi \otimes \psi)) &= \Phi\left(\left(R(\alpha) \circ \phi \circ \rho_{\mathbb{C}}(\alpha)^{-1} \otimes \rho_{\mathbb{C}}(\alpha)(\psi)\right)\right) \\ &= \left(R(\alpha) \circ \phi \circ \rho_{\mathbb{C}}(\alpha)^{-1}\right)(\rho_{\mathbb{C}}(\alpha)(\psi)) = R(\alpha)(\phi(\psi)).\end{aligned}$$

Since Σ_{2m} is irreducible, $\operatorname{Hom}_0(\Sigma_{2m}, \Sigma_{2m}) = \mathbb{C}\operatorname{Id}$. Since $V \cong \bigoplus_{k=1}^N \Sigma_{2m}$, we then have (where $\pi_k : \bigoplus_{k=1}^N \Sigma_{2m} \to \Sigma_{2m}$ is the projection)

$$\operatorname{Hom}_0(\Sigma_{2m}, V) \cong \operatorname{Hom}_0\left(\Sigma_{2m}, \bigoplus_{k=1}^N \Sigma_{2m}\right) = \bigoplus_{k=1}^N \mathbb{C}\pi_k.$$

and hence $\dim(\operatorname{Hom}_0(\Sigma_{2m}, V) \otimes \Sigma_{2m}) = N \cdot 2^m = \dim(V)$. For $z_k \in \mathbb{C}$,

$$\Phi\left(\bigoplus_{k=1}^N z_k \pi_k \otimes \psi_k\right) = (z_1\psi_1, \ldots, z_N\psi_N),$$

and Φ is then onto, and an isomorphism for dimensional reasons. □

REMARK 18.15. Alternatively, it is known (see [**We46**]) that, up to equivalence, the only irreducible representation of the algebra $\operatorname{End}(W)$ for any complex or real vector space W is the defining representation, namely $\operatorname{Id} : \operatorname{End}(W) \to \operatorname{End}(W)$. Since $\mathbb{C}l_{2m} \cong \operatorname{End}(\Sigma_{2m})$, it follows that, up to equivalence, $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}(\Sigma_{2m})$ is the only irreducible representation of $\mathbb{C}l_{2m}$. Note that $\rho_{\mathbb{C}}$ restricts to $\rho : Cl_{2m} \to \operatorname{End}(\Sigma_{2m})$. For dimensional reasons, ρ is not an isomorphism of real algebras. However, it is clearly an irreducible complex representation of Cl_{2m}, since any invariant subspace for ρ would also be invariant for $\rho_{\mathbb{C}}$. Additional considerations found in [**LaMi**] imply that $\rho : Cl_{2m} \to \operatorname{End}(\Sigma_{2m})$ is the unique (up to isomorphism) irreducible real representation of Cl_{2m}, but we will not be using this fact.

2. Spin Structures and Twisted Dirac Operators

Let M be a compact, oriented Riemannian n-manifold. Until further notice, we do not assume that n is even. Let FM be the principal $\operatorname{SO}(n)$-bundle of oriented, orthonormal frames. A *spin structure* for M consists of a principal $\operatorname{Spin}(n)$-bundle $P \to M$ and a map $C : P \to FM$ which is equivariant in the sense that $C(pg) = C(p)c(g)$, where $c : \operatorname{Spin}(n) \to \operatorname{SO}(n)$ is the double cover of (18.4), namely $c(g)(v) = gvg^{-1}$. Spin structures do not always exist, but given a coordinate ball $U \subseteq M$ and a local trivialization $T : FM|_U \cong U \times \operatorname{SO}(n)$, there is the obvious local spin structure

$$T \circ (\operatorname{Id} \times c) : U \times \operatorname{Spin}(n) \to U \times \operatorname{SO}(n) \cong FM|_U.$$

Our immediate goal is to establish the meaning and sketch the proof Proposition 18.18 below which exhibits the obstruction to finding a (global) spin structure for

M. Suppose that $\mathcal{U} = \{U_\alpha : \alpha \in J\}$ is an open covering of M, such that each intersection $U_{\alpha_1} \cap \cdots \cap U_{\alpha_k}$ of finitely many of the U_α is contractible (e.g., take each U_α to be a convex, normal coordinate ball). Such a covering is known as a *Leray covering*. There is then a local trivialization $T_\alpha : FM|_{U_\alpha} \cong U_\alpha \times \mathrm{SO}(n)$, say

$$T_\alpha(u) = (\pi(u), s_\alpha(u)),$$

where $s_\alpha(ug) = s_\alpha(u)g$ for all $g \in \mathrm{SO}(n)$. Since

$$s_\alpha(ug)s_\beta(ug)^{-1} = s_\alpha(u)g(s_\beta(u)g)^{-1} = s_\alpha(u)gg^{-1}s_\beta(u)^{-1} = s_\alpha(u)s_\beta(u)^{-1},$$

there is a well defined *transition function*

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{SO}(n) \text{ given by } g_{\alpha\beta}(\pi(u)) := s_\alpha(u)s_\beta(u)^{-1}.$$

Note that $g_{\beta\alpha} = g_{\alpha\beta}^{-1}$ and, more generally, we have the so-called cocycle condition

$$g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = I.$$

Conversely, if we are given functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{SO}(n)$, satisfying the cocycle condition (together with $g_{\alpha\alpha} \equiv I$), then a principal $\mathrm{SO}(n)$-bundle over M can be constructed as the set of equivalence classes for the relation on the disjoint union of the $U_\alpha \times \mathrm{SO}(n)$, where we declare that $(x, A_\alpha) \in U_\alpha \times \mathrm{SO}(n)$ is equivalent to $(x, A_\beta) \in U_\beta \times \mathrm{SO}(n)$ if $A_\alpha = g_{\alpha\beta}(x)A_\beta$ for $x \in U_\alpha \cap U_\beta$. Since, $g_{\alpha\alpha} \equiv I$, $g_{\beta\alpha} = g_{\alpha\beta}^{-1}$, and $g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = I$, this is an equivalence relation. As each of the $U_\alpha \cap U_\beta$ is contractible, we can find a lift $g'_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Spin}(n)$ of $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{SO}(n)$ with $c \circ g'_{\alpha\beta} = g_{\alpha\beta}$, where $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ is the double covering (vector representation). We can (and do) choose the collection of $g'_{\alpha\beta}$ so that $g'_{\beta\alpha} = (g'_{\alpha\beta})^{-1}$. Now, on $U_\alpha \cap U_\beta \cap U_\gamma$, we have

$$c\left(g'_{\alpha\beta}g'_{\beta\gamma}g'_{\gamma\alpha}\right) = g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = I \Rightarrow g'_{\alpha\beta}g'_{\beta\gamma}g'_{\gamma\alpha} = \pm 1 \in \mathbb{Z}_2 \subseteq \mathrm{Spin}(n).$$

Let

$$E'_{\alpha\beta\gamma} := g'_{\alpha\beta}g'_{\beta\gamma}g'_{\gamma\alpha} : U_\alpha \cap U_\beta \cap U_\gamma \to \mathbb{Z}_2.$$

Note that $E'_{\alpha\beta\gamma}$ is symmetric (as well as antisymmetric, as $E'_{\alpha\beta\gamma} \in \mathbb{Z}_2$) in α, β, γ, since

$$\begin{aligned} g'_{\alpha\beta}g'_{\beta\gamma}g'_{\gamma\alpha} = \pm 1 &\Rightarrow g'_{\alpha\beta}g'_{\beta\gamma} = \pm g'_{\alpha\gamma} \\ &\Rightarrow E'_{\alpha\beta\gamma} = \left(g'_{\alpha\beta}g'_{\beta\gamma}\right)g'_{\gamma\alpha} = g'_{\gamma\alpha}\left(g'_{\alpha\beta}g'_{\beta\gamma}\right) = E'_{\gamma\alpha\beta}, \text{ and} \\ E'_{\alpha\beta\gamma} = \left(E'_{\alpha\beta\gamma}\right)^{-1} &= \left(g'_{\gamma\alpha}\right)^{-1}\left(g'_{\beta\gamma}\right)^{-1}\left(g'_{\alpha\beta}\right)^{-1} = g'_{\alpha\gamma}g'_{\gamma\beta}g'_{\beta\alpha} = E'_{\alpha\gamma\beta}. \end{aligned}$$

The collection $E' = \left\{E'_{\alpha\beta\gamma}\right\}$ is an example of a Čech 2-cochain (with values in $\mathbb{Z}_2$) relative to the cover $\mathcal{U}$, the group of which we denote by $C^2(\mathcal{U}, \mathbb{Z}_2)$. The coboundary of E' is the Čech 3-cochain in $C^3(\mathcal{U}, \mathbb{Z}_2)$ defined by

$$(\delta E')_{\alpha\beta\gamma\delta} := E'_{\beta\gamma\delta} \cdot \left(E'_{\alpha\gamma\delta}\right)^{-1} \cdot E'_{\alpha\beta\delta} \cdot \left(E'_{\alpha\beta\gamma}\right)^{-1}$$

The following shows that $\delta E' = 1$:

$$\begin{aligned} (\delta E')_{\alpha\beta\gamma\delta} &= E'_{\beta\gamma\delta}E'_{\alpha\gamma\delta}E'_{\alpha\beta\delta}E'_{\alpha\beta\gamma} = \left(E'_{\delta\beta\gamma}E'_{\delta\gamma\alpha}\right)\left(E'_{\beta\delta\alpha}E'_{\beta\alpha\gamma}\right) \\ &= \left(g'_{\delta\beta}g'_{\beta\gamma}g'_{\gamma\alpha}g'_{\alpha\delta}\right)\left(g'_{\beta\delta}g'_{\delta\alpha}g'_{\alpha\gamma}g'_{\gamma\beta}\right) \\ &= \left(g'_{\delta\beta}g'_{\beta\gamma}g'_{\gamma\alpha}g'_{\alpha\delta}\right)\left(g'_{\delta\alpha}g'_{\alpha\gamma}g'_{\gamma\beta}g'_{\beta\delta}\right) = 1, \end{aligned}$$

meaning that E' is a Čech 2-cocycle, the group of which is denoted by $Z^2(\mathcal{U}, \mathbb{Z}_2)$. Suppose that we choose a different set of lifts of the $g_{\alpha\beta}$, say $g''_{\alpha\beta}$, say $g''_{\alpha\beta} = h_{\alpha\beta} g'_{\alpha\beta}$ for $h = \{h_{\alpha\beta}\} \in C^1(\mathcal{U}, \mathbb{Z}_2)$. Then for $E''_{\alpha\beta\gamma} := g''_{\alpha\beta} g''_{\beta\gamma} g''_{\gamma\alpha}$,

$$\begin{aligned} E''_{\alpha\beta\gamma} \left(E'_{\alpha\beta\gamma}\right)^{-1} &= \left(g''_{\alpha\beta} g''_{\beta\gamma} g''_{\gamma\alpha}\right) \left(g'_{\alpha\beta} g'_{\beta\gamma} g'_{\gamma\alpha}\right)^{-1} \\ &= \left(g''_{\alpha\beta} g''_{\beta\gamma} g''_{\gamma\alpha}\right) \left(g'_{\gamma\alpha}\right)^{-1} \left(g'_{\beta\gamma}\right)^{-1} \left(g'_{\alpha\beta}\right)^{-1} \\ &= h_{\alpha\beta} h_{\beta\gamma} h_{\gamma\alpha} = h_{\beta\gamma} h_{\gamma\alpha} h_{\alpha\beta} =: (\delta h)_{\alpha\beta\gamma}, \end{aligned}$$

so that E'' and E' differ by δh. The group of 2-coboundaries is

$$B^2(\mathcal{U}, \mathbb{Z}_2) := \left\{\delta h : h \in C^1(\mathcal{U}, \mathbb{Z}_2)\right\},$$

and it is easy to check that for any $h \in C^1(\mathcal{U}, \mathbb{Z}_2)$, $\delta\delta h = 1$, so that $B^2(\mathcal{U}, \mathbb{Z}_2) \subset Z^2(\mathcal{U}, \mathbb{Z}_2)$. Hence, we have shown that E' and E'' determine the same Čech cohomology class

$$[E'] = [E''] \in H^2(\mathcal{U}, \mathbb{Z}_2) := \frac{Z^2(\mathcal{U}, \mathbb{Z}_2)}{B^2(\mathcal{U}, \mathbb{Z}_2)}.$$

Of course $H^k(\mathcal{U}, \mathbb{Z}_2)$ can be defined for $k = 0, 1, 2, \ldots$. It can be shown that for Leray coverings $\mathcal{U}$, $H^k(\mathcal{U}, \mathbb{Z}_2)$ is naturally isomorphic to the usual (say, singular) cohomology group $H^k(M, \mathbb{Z}_2)$, with $\mathbb{Z}_2$-coefficients.

DEFINITION 18.16. Let $E' \in Z^2(\mathcal{U}, \mathbb{Z}_2)$ be the Čech 2-cocycle given by

$$E'_{\alpha\beta\gamma} := g'_{\alpha\beta} g'_{\beta\gamma} g'_{\gamma\alpha} : U_\alpha \cap U_\beta \cap U_\gamma \to \mathbb{Z}_2,$$

where the $g'_{\alpha\beta} : U_\alpha \cap U_\beta \to \operatorname{Spin}(n)$ are lifts of the transition functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \operatorname{SO}(n)$ for the oriented frame bundle FM of a compact, oriented Riemannian n-manifold relative to a *Leray covering* $\mathcal{U} = \{U_\alpha : \alpha \in J\}$. The class $w_2(M) := [E'] \in H^2(M, \mathbb{Z}_2)$ is known as the **second Stiefel-Whitney class of** M. More generally, by using transition functions, any equivalence class of a principal $\operatorname{SO}(k)$-bundle $P \to M$ (where k is not necessarily the dimension of M) can be identified with some $[P] \in H^1(M, \operatorname{SO}(k))$ and a function

$$w_2 : H^1(M, \operatorname{SO}(k)) \to H^2(M, \mathbb{Z}_2)$$

may be defined in the same way as $w_2(M)$ was defined in the case of FM. Thus, $w_2(M)$ is $w_2([P])$ in the special case $P = FM$, but for convenience we write $w_2(M)$ instead of $w_2([FM])$.

REMARK 18.17. Given a principal $\operatorname{SO}(k_1)$-bundle $P_1 \to M$ and a principal $\operatorname{SO}(k_2)$-bundle $P_2 \to M$, one can define a principal $\operatorname{SO}(k_1) \times \operatorname{SO}(k_2)$ bundle $P_1 \times P_2 \to M$ (fibered product) which determines an $\operatorname{SO}(k_1 + k_2)$-bundle $P \to M$ by means of the injection $\operatorname{SO}(k_1) \times \operatorname{SO}(k_2) \to \operatorname{SO}(k_1 + k_2)$ as in Proposition 16.19, p.378. Since transition functions for $P \to M$ can be taken to be products of transition functions for $P_1 \to M$ (with values in $\operatorname{SO}(k_1) \times \mathrm{Id}$) and transition functions for $P_2 \to M$ (with values in $\mathrm{Id} \times \operatorname{SO}(k_2)$). Since such products commute, it is clear from our construction that $w_2([P]) = w_2([P_1])\, w_2([P_2])$, or regarding $H^2(M, \mathbb{Z}_2)$ as an additive group (as in usually the case) we have

$$w_2([P]) = w_2([P_1]) + w_2([P_2]). \tag{18.16}$$

This formula may seem wrong to those already familiar with Stiefel-Whitney classes, since generally there is also a cup product term $w_1([P_1]) \smile w_1([P_2])$ on the right. However, $w_1([P])$ is trivial for $\mathrm{SO}(k)$-bundles, which sufficient for our purposes.

PROPOSITION 18.18. *Let M be a compact, oriented Riemannian n-manifold. Then M admits a spin structure if and only if $w_2(M) = 0$ (i.e., $w_2(M)$ is the identity of $H^2(M, \mathbb{Z}_2)$). In this case $H^1(M, \mathbb{Z}_2)$ acts freely on the set of inequivalent spin structures.*

PROOF. If $w_2(M) = 0$, then $E' \in B^2(\mathcal{U}, \mathbb{Z}_2)$ and so E' is the Čech coboundary of a Čech 1-cochain, say $F = \{F_{\alpha\beta}\}$; i.e.,

$$E'_{\alpha\beta\gamma} = (\delta F)_{\alpha\beta\gamma} = F_{\beta\gamma}(F_{\alpha\gamma})^{-1} F_{\alpha\beta} = F_{\beta\gamma}F_{\gamma\alpha}F_{\alpha\beta} = F_{\alpha\beta}F_{\beta\gamma}F_{\gamma\alpha}.$$

We construct a spin structure $C : P \to FM$ as follows. Let $\widetilde{g}'_{\alpha\beta} := F_{\alpha\beta}g'_{\alpha\beta} \in \mathrm{Spin}(n)$, and note that while the $g'_{\alpha\beta}$ did not necessarily satisfy the cocyle condition, the $\widetilde{g}'_{\alpha\beta}$ do:

$$\begin{aligned}\widetilde{g}'_{\alpha\beta}\widetilde{g}'_{\beta\gamma}\widetilde{g}'_{\gamma\alpha} &= F_{\alpha\beta}g'_{\alpha\beta}F_{\beta\gamma}g'_{\beta\gamma}F_{\gamma\alpha}g'_{\gamma\alpha} \\ &= F_{\alpha\beta}F_{\beta\gamma}F_{\gamma\alpha}g'_{\alpha\beta}g'_{\beta\gamma}g'_{\gamma\alpha} = \left(E'_{\alpha\beta\gamma}\right)^2 = 1.\end{aligned}$$

Thus, a principal $\mathrm{Spin}(n)$-bundle $P \to M$ can be constructed from the transition functions $\widetilde{g}'_{\alpha\beta}$. Since $c(\widetilde{g}'_{\alpha\beta}) = c(F_{\alpha\beta}g'_{\alpha\beta}) = c(\pm g'_{\alpha\beta}) = g_{\alpha\beta}$, where $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ is the double cover, we obtain a spin structure $C : P' \to FM$. Conversely, given a spin structure $C : P' \to FM$, with transition functions $\widetilde{g}'_{\alpha\beta}$, there is $F \in C^1(\mathcal{U}, \mathbb{Z}_2)$ with $\widetilde{g}'_{\alpha\beta} = F_{\alpha\beta}g'_{\alpha\beta}$. However, a choice F', differing from F, with $\delta F' = \delta F = E'$ might lead to different spin stucture $P' \to M$, defined via transition functions $\widetilde{g}''_{\alpha\beta} := F'_{\alpha\beta}g'_{\alpha\beta}$. While F and F' are not cocycles, their "difference" $F'F^{-1}$ is a cocycle, since $\delta\left(F'F^{-1}\right) = E'E'^{-1} = 1$. If $F'F^{-1}$ is a coboundary, say $F'F^{-1} = \delta k$ for some $k \in C^0(\mathcal{U}, \mathbb{Z}_2)$, then

$$F'_{\beta\alpha}F_{\alpha\beta} = k_\beta k_\alpha^{-1} \quad \text{or} \quad F_{\alpha\beta} = k_\alpha^{-1}F'_{\alpha\beta}k_\beta$$

This condition yields a well defined principal bundle isomorphism $\phi : P \to P'$ determined locally via the maps $\phi_\alpha : U_\alpha \times \mathrm{Spin}(n) \to U_\alpha \times \mathrm{Spin}(n)$ defined by

$$\phi_\alpha(x, a_\alpha) = (x, k_\alpha a_\alpha).$$

To see that the ϕ_α yeild a well defined $\phi : P \to P'$, note that (where $\equiv'$denotes the equivalence relation used in defining P' from the $\widetilde{g}''_{\alpha\beta}$)

$$\begin{aligned}&(x, k_\alpha a_\alpha) \equiv' (x, k_\beta a_\beta) \Leftrightarrow k_\alpha a_\alpha = \widetilde{g}''_{\alpha\beta}k_\beta a_\beta \\ &\Leftrightarrow k_\alpha a_\alpha = F'_{\alpha\beta}g'_{\alpha\beta}k_\beta a_\beta \Leftrightarrow a_\alpha = k_\alpha^{-1}F'_{\alpha\beta}k_\beta g'_{\alpha\beta}a_\beta \\ &\Leftrightarrow a_\alpha = F_{\alpha\beta}g'_{\alpha\beta}a_\beta \Leftrightarrow a_\alpha = \widetilde{g}'_{\alpha\beta}a_\beta \Leftrightarrow (x, a_\alpha) \equiv (x, a_\beta).\end{aligned}$$

That $\phi : P \to P'$ is $\mathrm{Spin}(n)$-equivariant and $\phi \circ C = C' \circ \phi$ for the coverings $C : P \to FM$ and $C' : P' \to FM$, follows from these obvious properties for the ϕ_α (recall $k_\alpha \in \mathbb{Z}_2 = \{1, -1\}$, so that $c(a_\alpha) = c(k_\alpha a_\alpha)$). Thus, if $F'F^{-1} = \delta k$ (i.e., F' and F differ by a coboundary), then F and F' define equivalent spin structures and the converse also holds. Note that

$$\widetilde{g}''_{\alpha\beta} = F'_{\alpha\beta}g'_{\alpha\beta} = F'_{\alpha\beta}F^{-1}_{\alpha\beta}F_{\alpha\beta}g'_{\alpha\beta} = F'_{\alpha\beta}F^{-1}_{\alpha\beta}\widetilde{g}'_{\alpha\beta} = \left(F'F^{-1}\right)_{\alpha\beta}\widetilde{g}'_{\alpha\beta}.$$

In other words, for a given F, there is a one-to-one correspondence between $H^1(M,\mathbb{Z}_2)$ and the set $\mathcal{S}(M)$ of inequivalent spin structures, induced by $F'F^{-1} \mapsto \left(F'F^{-1}\right)\widetilde{g}'$. Alternatively, there is a free action of $H^1(M,\mathbb{Z}_2)$ on $\mathcal{S}(M)$, given simply by multiplication, namely $(z \cdot \widetilde{g}')_{\alpha\beta} := z_{\alpha\beta}\widetilde{g}'_{\alpha\beta}$ for $z \in H^1(M,\mathbb{Z}_2)$. □

PROPOSITION 18.19. *Let P be a principal* $\mathrm{U}(1)$*-bundle over a connected, oriented 2-manifold M. Since* $\mathrm{U}(1) \cong \mathrm{SO}(2)$, *we may regard P as a principal* $\mathrm{SO}(2)$*-bundle. Thus, P possesses a Stiefel-Whitney class $w_2(P) \in H^2(M;\mathbb{Z}_2)$ as well as a Chern class $c_1(P) \in H^2(M;\mathbb{Z})$. We have*

$$w_2(P)[M] = c_1(P)[M] \bmod 2.$$

In the special case $P = FM$, we have (see (16.115), p.424) $c_1(FM)[M] = \chi[M] = 2 - 2\,\mathrm{genus}(M)$ which is even and so $w_2(M) := w_2(FM) = 0$.

PROOF. Let $\omega \in \Lambda^1(P,\mathfrak{u}(1))$ be a connection 1-form for P as a $\mathrm{U}(1)$-bundle, and let $\mathcal{U} = \{U_\alpha : \alpha \in J\}$ be a Leray covering by smoothly embedded disks so that we have trivalizing local sections $\sigma_\alpha : U_\alpha \to P|_{U_\alpha}$. For a $\mathrm{U}(1)$-bundle, the formula (see Exercise 16.33, p. 385) relating $\sigma_\alpha^*\omega$ to $\sigma_\beta^*\omega$ simplifies:

$$\sigma_\beta{}^*\omega = g_{\alpha\beta}^{-1}\left(\sigma_\alpha^*\omega\right)g_{\alpha\beta} + g_{\alpha\beta}^{-1}dg_{\alpha\beta} = \sigma_\alpha^*\omega + g_{\alpha\beta}^{-1}dg_{\alpha\beta} \text{ or}$$
$$\sigma_\beta{}^*\omega - \sigma_\alpha^*\omega = g_{\alpha\beta}^{-1}dg_{\alpha\beta} = d\left(\log g_{\alpha\beta}\right)$$

Let $iA_\alpha := \sigma_\alpha^*\omega$. The local curvature forms $\Omega_\alpha = d\sigma_\alpha^*\omega = idA_\alpha$ and $\Omega_\beta = d\sigma_\beta^*\omega = idA_\beta$ agree on the overlaps $U_\alpha \cap U_\beta$ to yield a well-defined 2-form $F \in \Omega^2(M,\mathbb{R})$ given locally by $F := -dA_\alpha$, and $\frac{1}{2\pi}F$ represents $c_1(P) \in H^2(M;\mathbb{Z})$ in de Rham cohomology. The isomorphism

$$H^2(M,\mathbb{R}) \to H^2(\mathcal{U},\mathbb{R}) := \frac{Z^2(\mathcal{U},\mathbb{R})}{B^2(\mathcal{U},\mathbb{R})}$$

from de Rham cohomology to Čech cohomology is obtained via the intermediate isomorphisms

$$H^2(M,\mathbb{R}) = \frac{H^0\left(\mathcal{U},\mathcal{Z}^2\right)}{dH^0\left(\mathcal{U},\mathcal{A}^1\right)} \cong H^1\left(\mathcal{U},\mathcal{Z}^1\right) \cong H^2(\mathcal{U},\mathbb{R}),$$

which we describe as follows. For a Leray cover $\mathcal{U}$ of M and for $0 \le j,k \le 2$, let $C^j\left(\mathcal{U},\mathcal{A}^k\right)$ be the group of Čech j-cochains c which assign to an ordered $(j+1)$-tuple $\left(U_{\alpha_0},\ldots,U_{\alpha_j}\right)$, with $U_{\alpha_i} \in \mathcal{U}$, a k-form $c_{\alpha_0\ldots\alpha_j} \in \Omega^k\left(U_{\alpha_0} \cap \ldots \cap U_{\alpha_j},\mathbb{R}\right)$. Similarly, let

$$C^j\left(\mathcal{U},\mathcal{Z}^k\right) := \left\{c \in C^j\left(\mathcal{U},\mathcal{A}^k\right) : dc_{\alpha_0\ldots\alpha_j} = 0 \text{ for all } (\alpha_0,\ldots,\alpha_j)\right\} \text{ and}$$
$$C^j\left(\mathcal{U},\mathcal{B}^k\right) := \left\{\begin{array}{l} c \in C^j\left(\mathcal{U},\mathcal{A}^k\right) : c_{\alpha_0\ldots\alpha_j} = db_{\alpha_0\ldots\alpha_j} \text{ for some} \\ b_{\alpha_0\ldots\alpha_j} \in \Omega^{k-1}\left(U_{\alpha_0} \cap \ldots \cap U_{\alpha_j},\mathbb{R}\right)\end{array}\right\}$$

The short exact sequences of cochain complexes

$$0 \to C^*\left(\mathcal{U},\mathcal{Z}^0 \cong \mathbb{R}\right) \xrightarrow{i} C^*\left(\mathcal{U},\mathcal{A}^0\right) \xrightarrow{d} C^*\left(\mathcal{U},\mathcal{Z}^1\right) \to 0 \text{ and}$$
$$0 \to C^*\left(\mathcal{U},\mathcal{Z}^1\right) \xrightarrow{i} C^*\left(\mathcal{U},\mathcal{A}^1\right) \xrightarrow{d} C^*\left(\mathcal{U},\mathcal{Z}^2\right) \to 0$$

give rise (in the standard way) to long exact sequences of cohomology groups

$$\begin{aligned}
&0 \to H^0\left(\mathcal{U},\mathcal{Z}^0\right) \to H^0\left(\mathcal{U},\mathcal{A}^0\right) \to H^0\left(\mathcal{U},\mathcal{Z}^1\right) \to H^1\left(\mathcal{U},\mathcal{Z}^0\right) \to H^1\left(\mathcal{U},\mathcal{A}^0\right)\\
&\quad \to H^1\left(\mathcal{U},\mathcal{Z}^1\right) \overset{\delta_1}{\to} H^2\left(\mathcal{U},\mathcal{Z}^0\right) \to H^2\left(\mathcal{U},\mathcal{A}^0\right)\cdots \text{ and}\\
&0 \to H^0\left(\mathcal{U},\mathcal{Z}^1\right) \to H^0\left(\mathcal{U},\mathcal{A}^1\right) \to H^0\left(\mathcal{U},\mathcal{Z}^2\right) \overset{\delta_0}{\to} H^1\left(\mathcal{U},\mathcal{Z}^1\right) \to H^1\left(\mathcal{U},\mathcal{A}^1\right)\cdots
\end{aligned} \tag{18.17}$$

Since the $\mathcal{A}^j$ are fine sheaves (admitting partitions of unity), we have (see [**Wells**, Theorem 3.11, p.56]) $H^i\left(\mathcal{U},\mathcal{A}^j\right) = 0$ for $i > 0$ and $j \geq 0$. Thus,

$$H^2_{\text{de Rham}}(M,\mathbb{R}) \cong \frac{H^0\left(\mathcal{U},\mathcal{Z}^2\right)}{dH^0\left(\mathcal{U},\mathcal{A}^1\right)} \overset{\delta_0}{\cong} H^1\left(\mathcal{U},\mathcal{Z}^1\right) \overset{\delta_1}{\cong} H^2\left(\mathcal{U},\mathcal{Z}^0\right) = H^2_{\check{\text{C}}\text{ech}}(\mathcal{U},\mathbb{R}).$$

The first isomorphism is obtained by noting that

$$H^0\left(\mathcal{U},\mathcal{Z}^2\right) = Z^0\left(\mathcal{U},\mathcal{Z}^2\right) = \left\{ \begin{array}{l} c \in C^0\left(\mathcal{U},\mathcal{A}^2\right) : dc_{\alpha_0} = 0 \text{ and} \\ 0 = (\delta c)_{\alpha_0\alpha_1} = c_{\alpha_1} - c_{\alpha_0} \text{ for all } \alpha_0,\alpha_1 \end{array} \right\},$$

and so any $c \in H^0\left(\mathcal{U},\mathcal{Z}^2\right)$ gives rise to a globally defined, closed 2-form on M. Similarly $H^0\left(\mathcal{U},\mathcal{A}^1\right)$ can be identified with the space of globally defined 1-forms on M, and $dH^0\left(\mathcal{U},\mathcal{A}^1\right) \cong$ the space of exact 2-forms on M. The isomorphism δ_0 in (18.17) is given as follows. Let $[F] \in H^2_{\text{de Rham}}(M,\mathbb{R})$ for a closed 2-form $F \in \Omega^2(M,\mathbb{R})$. We have that $F|_{U_\alpha}$ is exact, say $F|_{U_\alpha} = -dA_\alpha$ for some $A_\alpha \in \Omega^1(U_\alpha)$ or $A \in C^0\left(\mathcal{U},\mathcal{A}^1\right)$. We have $(\delta A)_{\alpha\beta} = A_\beta - A_\alpha$. Now

$$d\left((\delta A)_{\alpha\beta}\right) = d\omega_\beta - d\omega_\alpha = F|_{(U_\alpha\cap U_\beta)} - F|_{(U_\alpha\cap U_\beta)} = 0,$$

so that $\delta A \in C^1\left(\mathcal{U},\mathcal{Z}^1\right)$ and $\delta(\delta A) = 0$ so that $\delta A \in Z^1\left(\mathcal{U},\mathcal{Z}^1\right)$ and $[\delta A] \in H^1\left(\mathcal{U},\mathcal{Z}^1\right)$. Note that $[\delta A]$ is not necessarily 0, since $A \in C^0\left(\mathcal{U},\mathcal{A}^1\right)$ is not necessarily in $C^0\left(\mathcal{U},\mathcal{Z}^1\right)$. At any rate, δ_0 in (18.17) is given by $\delta_0[F] := [-\delta A]$. We now define δ_1 in (18.17). For $[B] \in H^1\left(\mathcal{U},\mathcal{Z}^1\right)$ we have $dB_{\alpha_0\alpha_1} = 0$ and $0 = \delta B \in C^2\left(\mathcal{U},\mathcal{Z}^1\right)$, and so there is $C_{\alpha_0\alpha_1} \in C^1\left(\mathcal{U},\mathcal{A}^0\right)$, such that $dC_{\alpha_0\alpha_1} = B_{\alpha_0\alpha_1}$. Moreover,

$$\begin{aligned}
\delta(C)_{\alpha_0\alpha_1\alpha_2} &= C_{\alpha_1\alpha_2} - C_{\alpha_0\alpha_2} + C_{\alpha_0\alpha_1}\\
d\left(\delta(C)_{\alpha_0\alpha_1\alpha_2}\right) &= dC_{\alpha_1\alpha_2} - dC_{\alpha_0\alpha_2} + dC_{\alpha_0\alpha_1} = B_{\alpha_1\alpha_2} - B_{\alpha_0\alpha_2} + B_{\alpha_0\alpha_1}\\
&= (\delta B)_{\alpha_0\alpha_1\alpha_2} = 0.
\end{aligned}$$

Thus, $\delta C \in C^2\left(\mathcal{U},\mathcal{Z}^0\right)$ and $\delta(\delta C) = 0$ so that $\delta C \in Z^2\left(\mathcal{U},\mathcal{Z}^0\right)$ and $\delta_1([B]) := [\delta C] \in H^2\left(\mathcal{U},\mathcal{Z}^0\right)$. We wish compute

$$\delta_1\delta_0 c_1(P) = \delta_1\delta_0\left(\left[\tfrac{1}{2\pi}F\right]\right)$$

Recall that $F|_{U_\alpha} = -dA_\alpha$, and

$$iA_\beta - iA_\alpha = {\sigma_\beta}^*\omega - \sigma_\alpha^*\omega = g_{\alpha\beta}^{-1}dg_{\alpha\beta} = d\log g_{\alpha\beta},$$

where the simple-connectedness of $U_\alpha \cap U_\beta$ yields a well-defined function $\log g_{\alpha\beta} \in \Omega^0(U_\alpha \cap U_\beta, i\mathbb{R})$ (unique up to additive multiples of $2\pi i$), such that $\exp(\log g_{\alpha\beta}) = g_{\alpha\beta}$.

$$\delta_1\delta_0 c_1(P) = \delta_1\delta_0\left(\left[\tfrac{1}{2\pi}F\right]\right) = \delta_1\left(\left[-\tfrac{1}{2\pi}\delta A\right]\right) = \left[\delta\left(\tfrac{1}{2\pi i}\log g\right)\right] \in H^2(\mathcal{U},\mathbb{R})$$

In fact $\delta_1\delta_0 c_1(P) \in H^2(\mathcal{U},\mathbb{Z})$, since $g_{\alpha_1\alpha_2} g_{\alpha_0\alpha_2}^{-1} g_{\alpha_0\alpha_1} = 1$ implies

$$\begin{aligned}\delta\left(\tfrac{1}{2\pi i}\log g\right)_{\alpha_0\alpha_1\alpha_2} &= \tfrac{1}{2\pi i}\left(\log g_{\alpha_1\alpha_2} - \log g_{\alpha_0\alpha_2} + \log g_{\alpha_0\alpha_1}\right)\\ &= \tfrac{1}{2\pi i}\log\left(g_{\alpha_1\alpha_2} g_{\alpha_0\alpha_2}^{-1} g_{\alpha_0\alpha_1}\right) \in \mathbb{Z}.\end{aligned}$$

Thus, $\left[\delta\left(\frac{1}{2\pi i}\log g\right)\right] \in H^2(\mathcal{U},\mathbb{Z})$ is the Čech cohomological version of $c_1(P) \in H^2(M;\mathbb{Z})$. The Stiefel-Whitney class $w_2(P) \in H^2(\mathcal{U};\mathbb{Z}_2)$ is $[E'] = [\delta g']$, where we may take $g'_{\alpha\beta} = \sqrt{g_{\alpha\beta}} := \exp\left(\frac{1}{4\pi i}\log g_{\alpha\beta}\right)$ (see ()). Then $\delta\left(\frac{1}{2\pi i}\log g\right)_{\alpha_0\alpha_1\alpha_2}$ is an odd integer on $U_{\alpha_0}\cap U_{\alpha_1}\cap U_{\alpha_2} \Leftrightarrow (\delta g')_{\alpha_0\alpha_1\alpha_2}$ is -1, and it follows that the mod 2 reduction of $\left[\delta\left(\frac{1}{2\pi i}\log g\right)\right]$ is $[\delta g']$. Thus, $w_2(P) \in H^2(\mathcal{U};\mathbb{Z}_2)$ is the mod 2 reduction of $c_1(P) \in H^2(M;\mathbb{Z})$, and $w_2(P)[M] = c_1(P)[M] \bmod 2$. □

Until further notice, we again assume that $2m = n = \dim M$ is even, and moreover that there is a spin structure $P \xrightarrow{C} FM \to M$. We may then form the Hermitian positive and negative spinor bundles $\Sigma^{\pm}(M) := P \times_{\mathrm{Spin}(n)} \Sigma_{2m}^{\pm}$ associated to the half-spinor representations $\rho^{\pm} : \mathrm{Spin}(n) \to \mathrm{SU}\left(\Sigma_{2m}^{\pm}\right)$ of (18.15), p. 490. Recall that $\rho^{+}\oplus\rho^{-} : \mathrm{Spin}(n) \to \mathrm{SU}(\Sigma_{2m})$ is the restriction of $\rho : Cl_{2m} \to \mathrm{End}(\Sigma_{2m})$. For $v \in \mathbb{R}^{2m} \subset Cl_{2m}$, we have $\rho(v) : \Sigma_{2m}^{\pm} \to \Sigma_{2m}^{\mp}$, with

$$\rho(c(a)v) = \rho\left(ava^{-1}\right) = \rho(a)\rho(v)\rho(a)^{-1} \quad \text{for } a \in \mathrm{Spin}(n).$$

Thus, $v \mapsto \rho(v)$ induces a well defined vector bundle morphism $TM \to \mathrm{End}(\Sigma(M))$ where $\Sigma(M) := \Sigma^{+}(M) \oplus \Sigma^{-}(M)$, or equivalently a so-called *Clifford multiplication*

$$cl : TM \otimes \Sigma(M) \to \Sigma(M).$$

with $cl\left(TM \otimes \Sigma^{\pm}(M)\right) \subseteq \Sigma^{\mp}(M)$. The Riemannian metric on M gives an identification of TM^* with TM and hence we may regard $cl : TM^* \otimes \Sigma(M) \to \Sigma(M)$, which induces a map (still denoted by cl) on the level of sections

$$cl : \Omega^1(M) \otimes C^\infty(\Sigma(M)) \to C^\infty(\Sigma(M)).$$

The Levi-Civita conection $\theta \in \Omega^1(FM,\mathfrak{so}(n))$ on FM pulls back via $C : P \to FM$ to a form $C^*(\theta) \in \Omega^1(P,\mathfrak{so}(n))$, which when composed with the isomorphism $c'^{-1} : \mathfrak{so}(n) \cong \mathfrak{spin}(n)$, gives us a form $\omega = c'^{-1}(C^*(\theta)) \in \Omega^1(P,\mathfrak{spin}(n))$. It follows from the equivariance of $C : P \to FM$ with respect to $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ (i.e., $C(pg) = C(p)c(g)$), that ω is a connection 1-form for $P \to M$. Thus, we have a covariant differentiation operator associated with ω, say

$$\nabla^{\Sigma} : C^\infty(\Sigma(M)) \to \Omega^1(M) \otimes C^\infty(\Sigma(M)).$$

Definition 18.20. For an oriented Riemannian manifold M with spin structure, the (standard) **Dirac operator** (also known as the **Atiyah-Singer operator** in [**LaMi**]), is the composition

$$\mathcal{D}^{\Sigma} := cl \circ \nabla^{\Sigma} : C^\infty(\Sigma(M)) \to C^\infty(\Sigma(M)).$$

For many applications, it will be necessary to use twisted Dirac operators which are defined as follows. Let $E \to M$ be a Hermitian vector bundle, and let $U(E) \to M$ be the principal bundle of unitary frames of E. We equip $U(E)$ with any connection 1-form, say $\varepsilon \in \mathcal{C}(U(E))$, and then we have an associated covariant differentiation operator $\nabla^E : C^\infty(E) \to \Omega^1(M) \otimes C^\infty(E)$. For the bundle $E \otimes$

$\Sigma(M)$ (which is an associated bundle of the fibered product $U(E) \times_f P$), we have a covariant differentiation operator

$$\nabla := \nabla^E \otimes 1 + 1 \otimes \nabla^\Sigma : C^\infty(E \otimes \Sigma(M)) \to \Omega^1(M) \otimes C^\infty(E \otimes \Sigma(M)).$$

This corresponds to the connection 1-form on $U(E) \times_f P$ which is the direct sum of the pull-backs of the connection 1-forms on $U(E)$ and P to a $U(E) \times_f P$.

Definition 18.21. For an oriented Riemannian manifold M with spin structure, and a Hermitian vector bundle $E \to M$ with unitary connection $\varepsilon \in \mathcal{C}(U(E))$, the **twisted Dirac operator $\mathcal{D}$ associated with** (E, ε) is

$$\mathcal{D} := (1 \otimes cl) \circ \nabla : C^\infty(E \otimes \Sigma(M)) \to C^\infty(E \otimes \Sigma(M)). \tag{18.18}$$

Note that $\nabla(C^\infty(E \otimes \Sigma^\pm(M))) \subseteq \Omega^1(M) \otimes C^\infty(E \otimes \Sigma^\pm(M))$, while

$$(1 \otimes cl)\left(\Omega^1(M) \otimes C^\infty\left(E \otimes \Sigma^\pm(M)\right)\right) \subseteq C^\infty\left(E \otimes \Sigma^\mp(M)\right).$$

Thus,

$$\begin{aligned} &\mathcal{D} = \mathcal{D}^+ \oplus \mathcal{D}^-, \text{ where} \\ &\mathcal{D}^\pm : C^\infty\left(E \otimes \Sigma^\pm(M)\right) \to C^\infty\left(E \otimes \Sigma^\mp(M)\right). \end{aligned}$$

The symbol of the first-order differential operator $\mathcal{D}$ is computed as follows. For $\phi \in C^\infty(M)$ with $\phi(x) = 0$ and $\psi \in C^\infty(E \otimes \Sigma(M))$, we have at x

$$\begin{aligned} (1 \otimes cl) \circ \nabla(\phi\psi) &= (1 \otimes cl) \circ ((d\phi)\psi + \phi\nabla\psi) = (1 \otimes cl) \circ (d\phi)\psi \\ &= (1 \otimes cl(d\phi))\psi. \end{aligned}$$

Thus, the symbol $\sigma(\mathcal{D}) : T_xM^* \to End(\Sigma(M))$ at the covector $\xi_x \in T_xM^*$ is given by

$$\sigma(\mathcal{D})(\xi_x) = 1 \otimes cl(\xi_x) : \Sigma_x \to \Sigma_x.$$

For $\xi_x \neq 0$, $\sigma(\mathcal{D})(\xi_x)$ is an isomorphism, since

$$\sigma(\mathcal{D})(\xi_x) \circ \sigma(\mathcal{D})(\xi_x) = 1 \otimes cl(\xi_x)^2 = -|\xi_x|^2 \operatorname{Id}.$$

Thus, $\mathcal{D}$ is an elliptic operator. Moreover, since $\sigma(\mathcal{D}^+)$ and $\sigma(\mathcal{D}^-)$ are restrictions of $\sigma(\mathcal{D})$, it follows that $\mathcal{D}^+$ and $\mathcal{D}^-$ are elliptic. In what follows, we show that $\mathcal{D}$ is formally self-adjoint, and $\mathcal{D}^+$ and $\mathcal{D}^-$ are formal adjoints of each other.

At times it best to express $\mathcal{D}\psi$ in terms of a local orthonormal frame field $E_1, \ldots, E_n$ on M. If $\varphi_1, \ldots, \varphi_n$ is the dual coframe, then, for any vector field X, we have

$$\begin{aligned} &(\nabla\psi)(X) = (\nabla\psi)(\sum_j \varphi_j(X) E_j)\psi \\ &= \sum_j \varphi_j(X)(\nabla\psi)(E_j) = \sum_j \varphi_j(X) \nabla_{E_j}\psi; \text{ i.e.,} \\ &\nabla\psi = \sum_j \varphi_j \otimes \nabla_{E_j}\psi. \end{aligned}$$

Thus, with "$\cdot$" denoting Clifford multipication $(1 \otimes cl)$ and using $T^*M \cong TM$, we have

$$\mathcal{D}\psi = \sum_j \varphi_j \cdot \nabla_{E_j}\psi = \sum_j E_j \cdot \nabla_{E_j}\psi. \tag{18.19}$$

Note that $\mathcal{D}\psi$ is independent of the choice of local orthonormal frame field. Moreover, in terms of local coordinates, say $y_1, \ldots, y_n$ with associated coordinate fields $\partial_i := \partial/\partial y^i$, we have

$$\mathcal{D}\psi = \sum_{i,j} h^{ij} \partial_i \cdot \nabla_{\partial_j} \psi, \tag{18.20}$$

where h^{ij} are the entries of the inverse of the matrix $[h_{ij}] = [h(\partial_i, \partial_j)]$. In the proof of Proposition 18.23 below, we take advantage of the fact that we can always choose $E_1, \ldots, E_n$ in a neighborhood U of a point $x \in M$ so that $\nabla_{E_j} E_k = 0$ at x, where ∇ is the covariant derivative for the Levi-Civita connection θ. More generally we have

PROPOSITION 18.22. *Let $\sigma : U \to FM$ be a local section of the frame bundle of a Riemannian manifold M, where $U \subseteq M$ is open. Let $E_1, \ldots, E_n$ be the orthonormal frame field on U given, at $y \in U$, by $(E_j)_y := \sigma(y)(e_j)$, where e_j is the j-th standard unit vector in $\mathbb{R}^n$. If $\theta \in \Omega^1(FM, \mathfrak{so}(n))$ is a connection 1-form on FM, with associated covariant differentiation operator ∇, then we have*

$$\nabla_{E_j} E_k = \sum_{i=1}^{n} (\sigma^*\theta)_{ik}(E_j) E_i = \sum_{i=1}^{n} \theta_{ik}(\sigma_{*y}(E_j)) E_i. \tag{18.21}$$

For $x \in M$ and $u \in \pi_F^{-1}(x)$, let N be an embedded submanifold of FM with $T_u N = H_u := \operatorname{Ker}\theta_u$. There is a neighborhood V of x, such that there is a unique section $\sigma : V \to FM$ with $\sigma(x) = u$, $\sigma(V) \subseteq N$ and $\sigma_{*x}(T_x M) = H_u$. In this case, $\nabla_{E_j} E_k = 0$ at x by (18.21).

PROOF. For a vector field Y on M, the corresponding equivariant function $\overline{\Omega}^0(FM, \mathbb{R}^n)$ is $\varphi(\widetilde{Y})$, where $\widetilde{Y}$ is the horizontal lift of Y relative to the connection θ and $\varphi \in \overline{\Omega}^1(FM, \mathbb{R}^n)$ is the canonical 1-form $(\varphi_u(Y) := u^{-1}(\pi_{F*}Y))$. For horizontal lifts $\widetilde{Y}, \widetilde{W}$ of vector fields Y and W on M, we have

$$\varphi(\widetilde{\nabla_Y W}) = D^\theta(\varphi(\widetilde{W}))\left(\widetilde{Y}\right) = \widetilde{Y}[\varphi(\widetilde{W})].$$

Let $\widetilde{E}_1, \ldots, \widetilde{E}_n$ be the horizontal lifts of $E_1, \ldots, E_n$. Note that $\varphi(\widetilde{E}_k)$ is constant on $\sigma(U)$, since

$$\varphi_{\sigma(y)}(\widetilde{E}_k) = \sigma(y)^{-1}(\pi_{F*}(\widetilde{E}_k)) = \sigma(y)^{-1}((E_k)_y) = e_k,$$

and so $d(\varphi(\widetilde{E}_k))(\sigma_{*y}(E_j)) = 0$. Then, as $\widetilde{E}_j = \sigma_{*x}(E_j)_H$ and $D^\theta = d + \theta$, we have

$$\begin{aligned} D^\theta(\varphi(\widetilde{E}_k))(\widetilde{E}_j) &= D^\theta(\varphi(\widetilde{E}_k))(\sigma_{*y}(E_j)) \\ &= d(\varphi(\widetilde{E}_k))(\sigma_{*y}(E_j)) + \theta(\sigma_{*y}(E_j))(\varphi(\widetilde{E}_k)) \\ &= \theta(\sigma_{*y}(E_j))(\varphi(\widetilde{E}_k)). \end{aligned}$$

Hence, it follows that

$$\begin{aligned} \widetilde{\nabla_{E_j} E_k} &= (\varphi|_H)^{-1}\left(D^\theta(\varphi(\widetilde{E}_k))(\widetilde{E}_j)\right) \\ &= \sum_{i=1}^{n} \theta_{ik}(\sigma_{*x}(E_j))(\varphi|_H)^{-1}(e_i) = \sum_{i=1}^{n} \theta_{ik}(\sigma_{*y}(E_j))\widetilde{E}_i \end{aligned}$$

and (18.21) follows by applying $(\pi_F)_*$ to both sides. The remaining assertions follow directly from the implicit function theorem. □

For $\psi, \psi' \in C^\infty(E \otimes \Sigma(M))$ and ν the volume form for (M, h), let

$$(\psi, \psi') := \int_M \langle \psi, \psi' \rangle \, \nu$$

PROPOSITION 18.23. *The twisted Dirac operator $\mathcal{D}$ of (18.18) is formally self-adjoint, namely*

$$(\mathcal{D}\psi_1, \psi_2) = (\psi_1, \mathcal{D}\psi_2). \tag{18.22}$$

Moreover, $\mathcal{D}^+$ is the formal adjoint of $\mathcal{D}^-$; i.e., for $\psi_1^+ \in C^\infty(E \otimes \Sigma^+(M))$ and $\psi_2^- \in C^\infty(E \otimes \Sigma^-(M))$, we have

$$\left(\mathcal{D}^+\psi_1^+, \psi_2^-\right) = \left(\psi_1^+, \mathcal{D}^-\psi_2^-\right). \tag{18.23}$$

PROOF. Let $\psi_1, \psi_2 \in C^\infty(E \otimes \Sigma(M))$. Since ∇ is the covariant differentiation for a connection on $U(E) \times_f P$ with group $\mathrm{U}(N) \times \mathrm{Spin}(n)$ which preseves the Hermitian structure on $\mathbb{C}^N \otimes \Sigma_n$, it follows that

$$d\langle \psi_1, \psi_2 \rangle = \langle \nabla\psi_1, \psi_2 \rangle + \langle \psi_1, \nabla\psi_2 \rangle. \tag{18.24}$$

Moreover, for $\rho : Cl_n \to \mathrm{End}(\Sigma_n)$, $a \in \mathrm{Spin}(n)$, $v \in \mathbb{R}^n$, and $\sigma \in \Sigma_n$, we have

$$\begin{aligned} \rho(a)(\rho(v)\sigma) &= \left(\rho(a) \circ \rho(v) \circ \rho(a)^{-1} \circ \rho(a)\right)(\sigma) \\ &= \rho\left(ava^{-1}\right)(\rho(a)\sigma) = \rho(c(a)v)(\rho(a)\sigma). \end{aligned} \tag{18.25}$$

Replacing a with $\exp(ta')$ for $a' \in \mathfrak{spin}(n)$ and differentiating (18.25) with respect to t at $t = 0$, we get

$$\rho'(a')(\rho(v)\sigma) = \rho(c'(a')v)(\sigma) + \rho(v)(\rho'(a')\sigma) \tag{18.26}$$

From this it follows that, for $\varphi \in \Omega^1(M)$ and $\psi \in C^\infty(E \otimes \Sigma(M))$,

$$\nabla(\varphi \cdot \psi) = (\nabla^\theta \varphi) \cdot \psi + \varphi \cdot \nabla\psi.$$

For a local orthonormal frame field $E_1, \ldots, E_n$, chosen so that $\nabla_{E_j} E_j = 0$ at some fixed $x \in M$, using (18.24) and (18.26), we have (at x)

$$\begin{aligned} \langle \mathcal{D}\psi_1, \psi_2 \rangle &= \sum\nolimits_j \left\langle E_j \cdot \nabla_{E_j}\psi_1, \psi_2 \right\rangle = \sum\nolimits_j -\left\langle \nabla_{E_j}\psi_1, E_j \cdot \psi_2 \right\rangle \\ &= \sum\nolimits_j -\left(E_j\left[\langle \psi_1, E_j \cdot \psi_2 \rangle\right] - \left\langle \psi_1, \nabla_{E_j}(E_j \cdot \psi_2) \right\rangle\right) \\ &= \sum\nolimits_j -E_j\left[\langle \psi_1, E_j \cdot \psi_2 \rangle\right] + \left\langle \psi_1, \left(\left(\nabla_{E_j} E_j\right) \cdot \psi_2 + E_j \cdot \nabla_{E_j}\psi_2\right) \right\rangle \\ &= \sum\nolimits_j -E_j\left[\langle \psi_1, E_j \cdot \psi_2 \rangle\right] + \left\langle \psi_1, E_j \cdot \nabla_{E_j}\psi_2 \right\rangle \\ &= \sum\nolimits_j -E_j\left[\langle \psi_1, E_j \cdot \psi_2 \rangle\right] + \langle \psi_1, \mathcal{D}\psi_2 \rangle. \end{aligned}$$

Consider the 1-form α given by

$$\alpha(Y) = \langle \psi_1, Y \cdot \psi_2 \rangle$$

We claim that at x,

$$\delta\alpha := *d(*\alpha) = -\sum\nolimits_j E_j\left[\alpha(E_j)\right] = -\sum\nolimits_j E_j\left[\langle \psi_1, E_j \cdot \psi_2 \rangle\right].$$

The j-th component of the related $\widetilde{\alpha} \in \Omega^0\left(FM, T^{0,1}\right)$ on FM is given at $u \in FM$ by $\widetilde{\alpha}_j = \widetilde{\alpha}(u)(e_j) = \alpha(\pi_{F*}(\overline{e}_j))$, where $\overline{e}_j$ is the j-th standard horizontal vector field at u. According to (16.93) on p. 409,

$$\delta\alpha = -\sum\nolimits_j \overline{e}_j\left[\widetilde{\alpha}_j\right],$$

which is constant on each fiber of FM. For any $y \in U$, we have $\overline{e}_j(\sigma(y)) = \widetilde{E}_j(\sigma(y))$, since both sides are horizontal and

$$\begin{aligned}\varphi\left(\widetilde{E}_j(\sigma(y))\right) &= \sigma(y)^{-1}\left(\pi_{F*}\widetilde{E}_j(\sigma(y))\right)\\ &= \sigma(y)^{-1}(E_j(y)) = e_j = \varphi(\overline{e}_j(\sigma(y))),\end{aligned}$$

where $\varphi \in \overline{\Omega}^1(FM, \mathbb{R}^n)$ is the canonical 1-form. For any $y \in U$,

$$\begin{aligned}(\sigma^*\widetilde{\alpha}_j)(y) &= \widetilde{\alpha}_j(\sigma(y)) = \alpha(\pi_{F*}\overline{e}_j(\sigma(y)))\\ &= \alpha\left(\pi_{F*}\widetilde{E}_j(\sigma(y))\right) = \alpha(E_j(y)), \text{ and so}\\ E_j[\alpha(E_j)] &= d(\alpha(E_j))(E_j) = d(\sigma^*\widetilde{\alpha}_j)(E_j) = \sigma^*(d\widetilde{\alpha}_j)(E_j)\\ &= d\widetilde{\alpha}_j(\sigma_*(E_j)) = \sigma_*(E_j)[\widetilde{\alpha}_j].\end{aligned}$$

However at $u = \sigma(x)$, we have $\sigma_*(E_j)[\widetilde{\alpha}_j] = \overline{e}_j[\widetilde{\alpha}_j]$. Thus,

$$\delta\alpha = -\sum\nolimits_j \overline{e}_j[\widetilde{\alpha}_j] = -\sum\nolimits_j E_j[\alpha(E_j)] \text{ at } x.$$

Then (18.22) follows, since

$$\begin{aligned}&\langle\mathcal{D}\psi_1, \psi_2\rangle\nu - \langle\psi_1, \mathcal{D}\psi_2\rangle\nu = -\sum\nolimits_j E_j\langle\psi_1, E_j\cdot\psi_2\rangle\nu\\ &= (\delta\alpha)\nu = *(\delta\alpha) = -**d(*\alpha) = -d(*\alpha)\\ &\Rightarrow (\mathcal{D}\psi_1, \psi_2) - (\mathcal{D}\psi_2, \psi_1) = \int_M(\langle\mathcal{D}\psi_1, \psi_2\rangle - \langle\psi_1, \mathcal{D}\psi_2\rangle)\nu\\ &= -\int_M d(*\alpha) = 0.\end{aligned}$$

Applying (18.22) when ψ_1 has values in $E \otimes \Sigma^+(M)$ and ψ_2 has values in the orthogonal bundle $E \otimes \Sigma^-(M)$, we obtain (18.23). □

The index of the self-adjoint operator $\mathcal{D}$ is 0, but the index of $\mathcal{D}^+$ (or $\mathcal{D}^-$) is not necessarily 0. Using the fact that $\mathcal{D}^+$ and $\mathcal{D}^-$ are adjoints, we have

$$\begin{aligned}\operatorname{index}\left(\mathcal{D}^+\right) &= \dim\operatorname{Ker}\left(\mathcal{D}^+\right) - \dim\operatorname{Coker}\left(\mathcal{D}^+\right)\\ &= \dim\operatorname{Ker}\left(\mathcal{D}^+\right) - \dim\operatorname{Ker}\left(\mathcal{D}^-\right)\\ &= \dim\operatorname{Ker}\left(\mathcal{D}^-\circ\mathcal{D}^+\right) - \dim\operatorname{Ker}\left(\mathcal{D}^+\circ\mathcal{D}^-\right).\end{aligned}$$

Moreover, since $\mathcal{D}^2 = (\mathcal{D}^-\circ\mathcal{D}^+) \oplus (\mathcal{D}^+\circ\mathcal{D}^-)$, it is convenient to define

$$\begin{aligned}\mathcal{D}_+^2 &:= \mathcal{D}^2|_{C^\infty(E\otimes\Sigma^+(M))} = \mathcal{D}^-\circ\mathcal{D}^+\\ \mathcal{D}_-^2 &:= \mathcal{D}^2|_{C^\infty(E\otimes\Sigma^-(M))} = \mathcal{D}^+\circ\mathcal{D}^-. \qquad (18.27)\end{aligned}$$

A detailed study of $\mathcal{D}^2$ is desirable, since the above yields

$$\operatorname{index}\left(\mathcal{D}^+\right) = \dim\operatorname{Ker}\left(\mathcal{D}_+^2\right) - \dim\operatorname{Ker}\left(\mathcal{D}_-^2\right).$$

To develop a suitable formula for $\mathcal{D}^2$, we compute (at $x \in M$ where $\nabla_{E_j} E_k = 0$)

$$\begin{aligned}
\mathcal{D}^2\psi &= \mathcal{D}\left(\mathcal{D}\psi\right) = \mathcal{D}(\sum_j E_j \cdot \nabla_{E_j}\psi) = \sum_j E_j \cdot \nabla_{E_j}(\sum_k E_k \cdot \nabla_{E_k}\psi) \\
&= \sum_j E_j \cdot \sum_k \nabla_{E_j}\left(E_k \cdot \nabla_{E_k}\psi\right) \\
&= \sum_{j,k} E_j \cdot \left(\nabla_{E_j} E_k \cdot \nabla_{E_k}\psi + E_k \cdot \nabla_{E_j}\left(\nabla_{E_k}\psi\right)\right) \\
&= \sum_{j,k} E_j \cdot E_k \cdot \nabla_{E_j}\left(\nabla_{E_k}\psi\right) \\
&= \sum_j E_j \cdot E_j \cdot \nabla_{E_j}\left(\nabla_{E_k}\psi\right) + \sum_{j\neq k} E_j \cdot E_k \cdot \nabla_{E_j}\left(\nabla_{E_k}\psi\right) \\
&= -\sum_j \nabla_{E_j}\left(\nabla_{E_j}\psi\right) + \tfrac{1}{2}\sum_{j\neq k} E_j \cdot E_k \cdot \left(\nabla_{E_j}\left(\nabla_{E_k}\psi\right) - \nabla_{E_k}\left(\nabla_{E_j}\psi\right)\right).
\end{aligned}$$

There is an *invariant second-derivative* $\nabla^2\psi \in C^\infty\left(M, E \otimes \Sigma(M) \otimes T^{0,2}(M)\right)$ defined by

$$\nabla^2_{X,Y}\psi = \nabla_X\left(\nabla_Y\psi\right) - \nabla_{\nabla_X Y}\psi.$$

The value of $\nabla^2_{(X,Y)}\psi$ at $x \in M$ depends on X_x but is independent of how X_x is extended, this is also true for Y_x, but less obviously so. Indeed,

$$\begin{aligned}
\nabla^2_{X,Y}\psi - \nabla^2_{Y,X}\psi &= \nabla_X\left(\nabla_Y\psi\right) - \nabla_Y\left(\nabla_X\psi\right) - \nabla_{\nabla_X Y - \nabla_Y X} \\
&= \nabla_X\left(\nabla_Y\psi\right) - \nabla_Y\left(\nabla_X\psi\right) - \nabla_{[X,Y]}\psi \\
&= \left(\left(D^{\varepsilon\oplus\omega}\right)^2\psi\right)(X,Y) = \Omega^{\varepsilon\oplus\omega}(X,Y)\,\psi,
\end{aligned}$$

and we know that $\Omega^{\varepsilon\oplus\omega}(X,Y)$ is independent of extensions. The trace of $\nabla^2\psi$ is the *connection Laplacian,* whichs since $\nabla_{E_j} E_k = 0$ at x, is given at x by

$$\Delta\psi := \sum_j \nabla^2_{E_j,E_j}\psi = \nabla_{E_j}\left(\nabla_{E_j}\psi\right) - \nabla_{\nabla_{E_j} E_j}\psi = \nabla_{E_j}\left(\nabla_{E_j}\psi\right). \tag{18.28}$$

REMARK 18.24. The *connection Laplacian* Δ *coincides with*

$$-\delta^{\varepsilon\oplus\omega} \circ D^{\varepsilon\oplus\omega} : \Omega^0\left(M, E \otimes \Sigma(M)\right) \to \Omega^0\left(M, E \otimes \Sigma(M)\right)$$

where $D^{\varepsilon\oplus\omega} : \Omega^0\left(M, E \otimes \Sigma(M)\right) \to \Omega^1\left(M, E \otimes \Sigma(M)\right)$ is the covariant derivative and $\delta^{\varepsilon\oplus\omega}$ is its formal adjoint, namely the covariant codifferential given in Proposition 16.18.

We have $\nabla_{E_j}\left(\nabla_{E_k}\psi\right) - \nabla_{E_k}\left(\nabla_{E_j}\psi\right) = \Omega^{\varepsilon\oplus\omega}\left(E_j, E_k\right)\psi$, and so

$$\begin{aligned}
\mathcal{D}^2\psi &= -\Delta\psi + \tfrac{1}{2}\sum_{j\neq k} E_j \cdot E_k \cdot \Omega^{\varepsilon\oplus\omega}\left(E_j, E_k\right)\psi \\
&= -\Delta\psi + \tfrac{1}{2}\sum_{j\neq k} E_j \cdot E_k \cdot \left(\left(\Omega^{\varepsilon}\left(E_j, E_k\right) \otimes \mathrm{Id}\right)(\psi) - \tfrac{1}{4}\textstyle\sum_{h,i} R_{hijk} E_h \cdot E_i \cdot \psi\right) \\
&= -\Delta\psi + \tfrac{1}{2}\sum_{j,k} \Omega^{\varepsilon}_{jk} E_j \cdot E_k\psi - \tfrac{1}{8}\sum_{h,i,j,k} R_{hijk} E_h \cdot E_i \cdot E_j \cdot E_k \cdot \psi,
\end{aligned}$$

where $\Omega^{\varepsilon}_{jk} = \Omega^{\varepsilon}(E_j, E_k) \otimes \mathrm{Id} \in \Omega^2(M, \mathrm{End}(E \otimes \Sigma(M)))$. We show that

$$-\tfrac{1}{8} \sum_{h,i,j,k} R_{hijk} E_j \cdot E_k \cdot E_h \cdot E_i \psi = \tfrac{1}{4} S\psi, \tag{18.29}$$

where $S = \sum_{h,i} R_{hihi}$ is the scalar curvature of M. For i, j, k distinct $E_i \cdot E_j \cdot E_k$ is invariant under cyclic permutation. Hence, by the first Bianchi identity,

$$\sum_{i,j,k \text{ distinct}} R_{hijk} E_i E_j E_k = \tfrac{1}{3} \sum_{i,j,k \text{ distinct}} (R_{hijk} + R_{hkij} + R_{hjki}) E_i E_j E_k = 0.$$

Thus, in view of the symmetries of R_{hijk}, we may assume that in the sum on the left of (18.29), no three indices are distinct. Then (18.29) follows from

$$\begin{aligned} \sum_{h,i,j,k} R_{hijk} E_j E_k E_h E_i &= 2 \sum_{h,i} R_{hihi} E_h E_i E_h E_i \\ &= -2 \sum_{h,i} R_{hihi} E_h E_h E_i E_i = -2 \sum_{h,i} R_{hihi} = -2S. \end{aligned}$$

In summary, we have

PROPOSITION 18.25. *Let $\Omega^{\varepsilon} \in \Omega^2(M, \mathrm{End}(E))$ denote the curvature of the connection ε on $U(E)$ and let S be the scalar curvature of M. For $\psi \in C^{\infty}(E \otimes \Sigma(M))$, and an orthonormal frame $E_1, \ldots, E_n$ at $x \in M$, let*

$$(\mathfrak{R}^{\varepsilon}\psi)(x) := \tfrac{1}{2} \sum_{j,k} \Omega^{\varepsilon}_{jk} E_j \cdot E_k \cdot \psi(x),$$

where $\Omega^{\varepsilon}_{jk} = \Omega^{\varepsilon}(E_j, E_k) \otimes \mathrm{Id} \in \Omega^2(M, \mathrm{End}(E \otimes \Sigma(M)))$. We have

$$\mathcal{D}^2\psi = -\Delta\psi + \mathfrak{R}^{\varepsilon}\psi + \tfrac{1}{4} S\psi. \tag{18.30}$$

COROLLARY 18.26. *Let M be a compact Riemannian manifold with spin structure and let E be a Hermitian vector bundle over M. If the symmetric transformation*

$$\left(\mathfrak{R}^{\varepsilon} + \tfrac{1}{4} S\right)(x) \in \mathrm{End}(E_x \otimes \Sigma_x(M)) \tag{18.31}$$

is nonnegative semi-definite at each $x \in M$, then $\mathcal{D}\psi = 0 \Rightarrow \nabla\psi = 0$;i.e., all harmonic twisted spinors on M are parallel. If M is connected and (18.31) is nonnegative semi-definite at each $x \in M$ and positive definite at some $x_0 \in M$, then $\mathrm{Ker}\,\mathcal{D} = 0$ (i.e., there are no nonzero twisted harmonic spinors on M). In particular (taking $E = 0$), if $S \geq 0$ and $S \neq 0$, then there are no nonzero harmonic spinors on M.

PROOF. By Proposition 18.23 and Remark 18.24,

$$\begin{aligned} \|\mathcal{D}\psi\|^2 &= \left(\mathcal{D}^2\psi, \psi\right) = \left(-\Delta\psi + \mathfrak{R}^{\varepsilon}\psi + \tfrac{1}{4} S\psi, \psi\right) \\ &= \left(\delta^{\varepsilon\oplus\omega} \circ D^{\varepsilon\oplus\omega}\psi, \psi\right) + \left(\left(\mathfrak{R}^{\varepsilon} + \tfrac{1}{4} S\right)\psi, \psi\right) \\ &= \|\nabla\psi\|^2 + \left(\left(\mathfrak{R}^{\varepsilon} + \tfrac{1}{4} S\right)\psi, \psi\right) \geq 0, \end{aligned} \tag{18.32}$$

where the inequality is strict if $\nabla\psi \neq 0$. Thus, $\mathcal{D}\psi = 0 \Rightarrow \nabla\psi = 0$. Now

$$\begin{aligned} \nabla\psi = 0 &\Rightarrow d\left(|\psi|^2\right) = \langle \nabla\psi, \psi \rangle + \langle \psi, \nabla\psi \rangle = 0 \\ &\Rightarrow |\psi|^2 \text{ constant} \Rightarrow \psi = 0 \text{ or } \psi(x) \neq 0 \text{ for all } x \in M. \end{aligned}$$

Hence, if $\mathcal{D}\psi = 0$ for $\psi \neq 0$, then $\psi(x_0) \neq 0$ and assuming that $\left(\mathfrak{R}^\varepsilon + \frac{1}{4}S\right)(x_0)$ is positive definite, we have $\left\langle \left(\mathfrak{R}^\varepsilon + \frac{1}{4}S\right)\psi, \psi\right\rangle > 0$ at x_0 and so $\left(\left(\mathfrak{R}^\varepsilon + \frac{1}{4}S\right)\psi, \psi\right) > 0$, which contradicts (18.32). Thus, $\mathcal{D}\psi = 0 \Rightarrow \psi = 0$, if $\left(\mathfrak{R}^\varepsilon + \frac{1}{4}S\right)(x_0)$ is positive definite for some x_0. □

3. The Spinorial Heat Kernel

Recall from (18.27) that we have a pair of self-adjoint elliptic operators

$$\mathcal{D}_+^2 := \mathcal{D}^2|_{C^\infty(E\otimes\Sigma^+(M))} = \mathcal{D}^- \circ \mathcal{D}^+$$
$$\mathcal{D}_-^2 := \mathcal{D}^2|_{C^\infty(E\otimes\Sigma^-(M))} = \mathcal{D}^+ \circ \mathcal{D}^-.$$

For $\lambda \in \mathbb{C}$, let

$$V_\lambda\left(\mathcal{D}_\pm^2\right) := \left\{\psi \in C^\infty\left(E \otimes \Sigma^\pm(M)\right) : \mathcal{D}_\pm^2\psi = \lambda\psi\right\}.$$

From the general theory of formally self-adjoint, elliptic operators on compact manifolds, we know that

$$\operatorname{Spec}\left(\mathcal{D}_\pm^2\right) = \left\{\lambda \in \mathbb{C} : V_\lambda\left(\mathcal{D}_\pm^2\right) \neq \{0\}\right\}.$$

consists of the eigenvalues of $\mathcal{D}_\pm^2$ and is a discrete subset of $[0,\infty)$, the eigenspaces $V_\lambda\left(\mathcal{D}_\pm^2\right)$ are finite-dimensional, and an $L^2\left(E\otimes\Sigma^\pm(M)\right)$-complete orthonormal set of vectors can be selected from the $V_\lambda\left(\mathcal{D}_\pm^2\right)$. Note that $\mathcal{D}^+\left(V_\lambda\left(\mathcal{D}_+^2\right)\right) \subseteq V_\lambda\left(\mathcal{D}_-^2\right)$, since for $\psi \in V_\lambda\left(\mathcal{D}_+^2\right)$

$$\begin{aligned}\mathcal{D}_-^2\left(\mathcal{D}^+\psi\right) &= \left(\mathcal{D}^+ \circ \mathcal{D}^-\right)\left(\mathcal{D}^+\psi\right) = \mathcal{D}^+\left(\left(\mathcal{D}^- \circ \mathcal{D}^+\right)(\psi)\right)\\ &= \mathcal{D}^+\left(\mathcal{D}_+^2(\psi)\right) = \mathcal{D}^+(\lambda\psi) = \lambda\mathcal{D}^+(\psi),\end{aligned}$$

and similarly $\mathcal{D}^-\left(V_\lambda\left(\mathcal{D}_-^2\right)\right) \subseteq V_\lambda\left(\mathcal{D}_+^2\right)$. For $\lambda \neq 0$,

$$\mathcal{D}^\pm|_{V_\lambda\left(\mathcal{D}_\pm^2\right)} : V_\lambda\left(\mathcal{D}_\pm^2\right) \to V_\lambda\left(\mathcal{D}_\mp^2\right)$$

is an isomorphism, since it has inverse $\frac{1}{\lambda}\mathcal{D}^\mp$. Thus the set of nonzero eigenvalues (and their multiplicities) of $\mathcal{D}_+^2$ coincides with that of $\mathcal{D}_-^2$. However, in general

$$\dim V_0\left(\mathcal{D}_+^2\right) - \dim V_0\left(\mathcal{D}_-^2\right) = \dim\operatorname{Ker}\left(\mathcal{D}_+^2\right) - \dim\operatorname{Ker}\left(\mathcal{D}_-^2\right) = \operatorname{index}\left(\mathcal{D}^+\right) \neq 0.$$

Since $\dim V_\lambda\left(\mathcal{D}_+^2\right) - \dim V_\lambda\left(\mathcal{D}_-^2\right) = 0$ for $\lambda \neq 0$, obviously

$$\begin{aligned}\operatorname{index}\left(\mathcal{D}^+\right) &= \dim V_0\left(\mathcal{D}_+^2\right) - \dim V_0\left(\mathcal{D}_-^2\right)\\ &= \sum\nolimits_{\lambda\in\operatorname{Spec}\left(\mathcal{D}_+^2\right)} e^{-t\lambda}\left(\dim V_\lambda\left(\mathcal{D}_+^2\right) - \dim V_\lambda\left(\mathcal{D}_-^2\right)\right).\end{aligned}$$

The point is that the sum can be expressed as the integral of the supertrace of the heat kernel for the spinorial heat equation $\frac{\partial\psi}{\partial t} = -\mathcal{D}^2\psi$, from which the local index theorem for $\mathcal{D}^+$ will eventually follow. However, first we need to establish the existence of the heat kernel.

Let the *positive* eigenvalues of $\mathcal{D}_\pm^2$ be placed in a sequence $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$ where each eigenvalue is repeated according to its multiplicity. Let $u_1^\pm, u_2^\pm, \ldots$ be an L^2-orthonormal sequence in $C^\infty\left(E\otimes\Sigma^+(M)\right)$ with $\mathcal{D}_\pm^2\left(u_j^\pm\right) = \lambda_j u_j^\pm$ (i.e., $u_j^\pm \in V_{\lambda_j}\left(\mathcal{D}_\pm^2\right)$). We let $u_{0_1}^+, \ldots, u_{0_{n^+}}^+$ be an L^2-orthonormal basis of $\operatorname{Ker}\mathcal{D}_+^2 = \operatorname{Ker}\mathcal{D}_+$, and $u_{0_1}^-, \ldots, u_{0_{n^-}}^-$ be an L^2-orthonormal basis of $\operatorname{Ker}\mathcal{D}_-^2 = \operatorname{Ker}\mathcal{D}_-$. We can pull back the bundle $E \otimes \Sigma^{\pm}(M)$ via either of the projections $M \times M \times (0,\infty) \to M$ given by $\pi_1(x,y,t) := x$ and $\pi_2(x,y,t) := y$ and take the tensor product of the results to form a bundle

$$\mathcal{K}^{\pm} := \pi_1^*\left(E \otimes \Sigma^{\pm}(M)\right) \otimes \pi_2^*\left(E \otimes \Sigma^{\pm}(M)\right) \to M \times M \times (0,\infty).$$

PROPOSITION 18.27. *For $t > t_0 > 0$, the series $k'^{\pm}$, defined by*

$$k'^{\pm}(x,y,t) := \sum_{j=1}^{\infty} e^{-\lambda_j t} u_j^{\pm}(x) \otimes u_j^{\pm}(y),$$

converges uniformly in $C^q(\mathcal{K}^{\pm}|_{M\times M\times(t_0,\infty)})$ for all $q \geq 0$. Hence $k'^{\pm} \in C^{\infty}(\mathcal{K}^{\pm})$, and (for $t > 0$)

$$\frac{\partial}{\partial t} k^{\pm}(x,y,t) = -\sum_{j=1}^{\infty} \lambda_j e^{-\lambda_j t} u_j^{\pm}(x) \otimes u_j^{\pm}(y) = -\mathcal{D}_{\pm}^2 k^{\pm}(x,y,t). \tag{18.33}$$

PROOF. Recall from (17.46), p. 468, that there are constants $c_k > 0$ such that

$$\left\|u_j^{\pm}\right\|_{C^q} \leq c_k \left\|u_j^{\pm}\right\|_{2,k} \text{ for } 0 \leq q < k - \frac{n}{2}$$

Thus for

$$k(q) := q + \frac{n}{2} + \tau,$$

where $\tau > 0$ so that $q < k - \frac{n}{2}$, we have

$$\left\|e^{-\lambda_j t} u_j^{\pm} \otimes u_j^{\pm}\right\|_{C^q} \leq e^{-\lambda_j t} \left\|u_j^{\pm}\right\|_{C^q} \left\|u_j^{\pm}\right\|_{C^q} \leq e^{-\lambda_j t} c_{k(q)}^2 \left\|u_j^{\pm}\right\|_{2,k(q)}^2.$$

Moreover, by Proposition 17.21, p. 468, there are constants C_k independent of j, such that

$$\begin{aligned}\left\|u_j^{\pm}\right\|_{2,k+2} &\leq C_k \left(\left\|\mathcal{D}_{\pm}^2 u_j^{\pm}\right\|_{2,k} + \left\|u_j^{\pm}\right\|_2\right) \\ &= C_k \left(\left\|\lambda_j u_j^{\pm}\right\|_{2,k} + \left\|u_j^{\pm}\right\|_2\right) \leq C_k (\lambda_j + 1) \left\|u_j^{\pm}\right\|_{2,k}.\end{aligned}$$

As $\left\|u_j^{\pm}\right\|_{2,0} = 1$, iteration yields contants $C'_{k(q)}$, such that

$$\left\|u_j^{\pm}\right\|_{2,k(q)} \leq C'_{k(q)} (\lambda_j + 1)^{[k(q)/2]} \left\|u_j^{\pm}\right\|_{2,0} = C'_{k(q)} (\lambda_j + 1)^{[k(q)/2]} \text{ and}$$

$$\left\|u_j^{\pm}\right\|_{C^q} \leq c_{k(q)} \left\|u_j^{\pm}\right\|_{2,k(q)} \leq c_{k(q)} C'_{k(q)} (\lambda_j + 1)^{[k(q)/2]}. \tag{18.34}$$

Combining the above estimates, we then have (for j sufficiently large)

$$\begin{aligned}&\left\|e^{-\lambda_j t} u_j^{\pm} \otimes u_j^{\pm}\right\|_{C^q} \leq e^{-\lambda_j t} c_{k(q)}^2 \left\|u_j^{\pm}\right\|_{2,k(q)}^2 \\ &\leq e^{-\lambda_j t} c_{k(q)}^2 C'^2_{k(q)} (\lambda_j + 1)^{k(q)} \leq e^{-\lambda_j t} c_{k(q)}^2 C'^2_{k(q)} (2\lambda_j)^{k(q)} \\ &\leq C''_{k(q)} e^{-\lambda_j t} \lambda_j^{k(q)}, \text{ where } C''_{k(q)} := 2^{k(q)} c_{k(q)}^2 C'^2_{k(q)}.\end{aligned}$$

In order to apply the Weierstrass M-test to deduce the (uniform) convergence of $\sum_{j=1}^{\infty} e^{-\lambda_j t} u_j^{\pm} \otimes u_j^{\pm}$ in $C^q(\mathcal{K}|_{M\times M\times(t,\infty)})$, we need to show that $\sum_{j=1}^{\infty} e^{-\lambda_j t} \lambda_j^{k(q)} < \infty$; note that the inclusion of the arbitrary positive parameter τ in the definition of $k(q)$ is designed to handle the uniform C^q convergence in t as well. It is easy to check that

$$e^{-x/2} x^k \leq D_k := e^{-k} (2k)^k,$$

by noting that

$$\tfrac{d}{dx}\left(e^{-x/2}x^k\right) = x^{k-1}\left(-\tfrac{1}{2}x + k\right)e^{-\frac{1}{2}x} = 0 \text{ for } x = 2k.$$

Thus,

$$\begin{aligned}
&e^{-\lambda_j t/2}\left(\lambda_j t\right)^{k(q)} \leq D_{k(q)},\\
&e^{-\lambda_j t}\lambda_j^{k(q)} \leq D_{k(q)} t^{-k(q)} e^{-\lambda_j t/2}, \text{ and}\\
&\sum_{j=1}^{\infty} e^{-\lambda_j t}\lambda_j^{k(q)} \leq D_{k(q)} t^{-k(q)} \sum_{j=1}^{\infty} e^{-\lambda_j t/2}.
\end{aligned}$$

If we can show that $\lambda_j \geq C j^{\alpha}$ for some positive constants C and α, then

$$\sum_{j=1}^{\infty} e^{-\lambda_j t/2} \leq \sum_{j=1}^{\infty} e^{-C j^{\alpha} t/2} < \infty,$$

by comparison with an integral of the form (for α,$\beta \in (0,\infty)$)

$$\int_0^{\infty} e^{-\beta x^{\alpha}}\,dx = \frac{1}{\alpha}\beta^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right) < \infty.$$

Thus, it remains to show that $\lambda_j \geq C j^{\alpha}$ for some positive constants C and α. Let $V_j^{\pm} := \left\{\sum_{i=1}^{j} a_i u_i^{\pm} : a_i \in \mathbb{C}\right\}$. For $u^{\pm} = \sum_{i=1}^{j} a_i u_i^{\pm} \in V_n^{\pm}$ and $k > \frac{n}{2}$, we have (using Proposition 17.21, p. 468)

$$\begin{aligned}
\left\|u^{\pm}\right\|_{C^0} &\leq c_k \left\|u^{\pm}\right\|_{2,k} \leq c_k C_k \left\|\left(\mathcal{D}_{\pm}^2 + 1\right)^{[k/2]}\left(u^{\pm}\right)\right\|_{2,0}\\
&\leq c_k C_k \left(\lambda_j + 1\right)^{[k/2]} \left\|u^{\pm}\right\|_{2,0}\\
&= C_k'' \left(\lambda_j + 1\right)^{[k/2]} \left(\sum_{i=1}^{j} \left|a_i\right|^2\right)^{\frac{1}{2}},
\end{aligned}$$

where we have used the characterization

$$\left(\lambda_j + 1\right)^{[k/2]} = \sup_{u^{\pm} \in V_j^{\pm}} \frac{\left\|\left(\mathcal{D}_{\pm}^2 + 1\right)^{[k/2]}\left(u^{\pm}\right)\right\|_{2,0}}{\left\|u^{\pm}\right\|_{2,0}}.$$

Hence, for any $x \in M$ and $a_i \in \mathbb{C}$,

$$\left|\sum_{i=1}^{j} a_i u_i^{\pm}(x)\right| \leq C_k'' \left(\lambda_j + 1\right)^{[k/2]} \left(\sum_{i=1}^{j} \left|a_i\right|^2\right)^{\frac{1}{2}}.$$

Let $F_1(x), \ldots, F_N(x)$ ($N = \dim E \cdot \dim \Sigma_n^{\pm}$) be an orthonormal basis for $E_x \otimes \Sigma^{\pm}(M)_x$, and for *fixed* $h \in \{1, \ldots, N\}$, we make the specific choice $a_j = \left\langle u_j^{\pm}(x), F_h(x)\right\rangle^*$

for the a_j. Then, *not* summing over h,

$$\begin{aligned}
&\sum\nolimits_{i=1}^{j}\left|\left\langle u_i^{\pm}(x),F_h(x)\right\rangle\right|^2=\sum\nolimits_{i=1}^{j}a_i\left\langle u_i^{\pm}(x),F_h(x)\right\rangle\left|F_h(x)\right|\\
&=\left|\sum\nolimits_{i=1}^{j}a_i\left\langle u_i^{\pm}(x),F_h(x)\right\rangle F_h(x)\right|\le\left|\sum\nolimits_{i=1}^{j}a_iu_i^{\pm}(x)\right|\\
&\le C_k''(\lambda_j+1)^{[k/2]}\left(\sum\nolimits_{i=1}^{j}|a_i|^2\right)^{\frac{1}{2}}\\
&=C_k''(\lambda_j+1)^{[k/2]}\left(\sum\nolimits_{i=1}^{j}\left|\left\langle u_i^{\pm}(x),F_h(x)\right\rangle\right|^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Dividing this by $\left(\sum_{i=1}^{j}\left|\left\langle u_i^{\pm}(x),F_h(x)\right\rangle\right|^2\right)^{\frac{1}{2}}$ and squaring, we have

$$\sum\nolimits_{i=1}^{j}\left|\left\langle u_j^{\pm}(x),F_h(x)\right\rangle\right|^2\le C_k''^2(\lambda_j+1)^k.$$

Now summing over h, we get

$$\sum_{i=1}^{j}\left|u_i^{\pm}(x)\right|^2=\sum\nolimits_{h=1}^{N}\sum\nolimits_{i=1}^{j}\left|\left\langle u_i^{\pm}(x),F_h(x)\right\rangle\right|^2\le NC_k''^2(\lambda_j+1)^k.$$

Integrating over M, we obtain

$$j=\sum_{i=1}^{j}\left\|u_i^{\pm}\right\|_{2,0}^2\le NC_k''^2(\lambda_j+1)^k\operatorname{Vol}(M).$$

Thus, for $k>n/2$ and j sufficiently large, we have the desired result

$$\lambda_j\ge\left(\frac{j}{C_k''^2\operatorname{Vol}(M)}\right)^{\frac{1}{k}}-1\ge2\left(\frac{1}{C_k''^2\operatorname{Vol}(M)}\right)^{\frac{1}{k}}j^{\frac{1}{k}}. \tag{18.35}$$

□

DEFINITION 18.28. The **positive and negative twisted spinorial heat kernels** (or the **heat kernels for $\mathcal{D}_{\pm}^2$**) $k^{\pm}\in C^{\infty}(\mathcal{K}^{\pm})$ are given by

$$k^{\pm}(x,y,t):=\sum_{i=1}^{n^{\pm}}u_{0_i}^{\pm}(x)\otimes u_{0_i}^{\pm}(y)+\sum_{j=1}^{\infty}e^{-\lambda_jt}u_j^{\pm}(x)\otimes u_j^{\pm}(y).$$

The **total twisted spinorial heat kernel** (or the **heat kernel for $\mathcal{D}^2$**) is

$$\begin{aligned}
&k=\left(k^+,k^-\right)\in C^{\infty}\left(\mathcal{K}^+\right)\oplus C^{\infty}\left(\mathcal{K}^-\right)\cong C^{\infty}\left(\mathcal{K}^+\oplus\mathcal{K}^-\right)\subseteq C^{\infty}(\mathcal{K}),\\
&\text{where }\mathcal{K}:=\pi_1^*(E\otimes\Sigma(M))\otimes\pi_2^*(E\otimes\Sigma(M)).
\end{aligned} \tag{18.36}$$

The terminology is justified in view of the following

PROPOSITION 18.29. *Let* $\psi_0^{\pm}\in C^{\infty}\left(E\otimes\Sigma^{\pm}(M)\right)$ *and let*

$$\psi^{\pm}(x,t)=\int_M\left\langle k^{\pm}(x,y,t),\psi_0^{\pm}(y)\right\rangle_y\ \nu(y).$$

Then for $t > 0$, $\psi^{\pm}$ solves the heat equation with initial spinor $\psi_0^{\pm}$:

$$\frac{\partial \psi^{\pm}}{\partial t} = -\mathcal{D}_{\pm}^2 \psi \text{ and}$$
$$\lim_{t \to 0^+} \psi^{\pm}(\cdot, t) = \psi_0^{\pm} \text{ in } C^q \text{ for all } q \geq 0.$$

Moreover, for $\psi_0 \in C^{\infty}(E \otimes \Sigma(M))$ and

$$\psi(x,t) := \int_M \langle k(x,y,t), \psi_0 \rangle \ \nu,$$

we have $\dfrac{\partial \psi}{\partial t} = -\mathcal{D}^2 \psi$ and $\lim_{t \to 0^+} \psi(\cdot, t) = \psi_0(\cdot)$ in C^q for all $q \geq 0$.

Proof. Since $k^{\pm}$ is C^{∞}, we may differentiate under the integral and use (18.33) to deduce that $\dfrac{\partial \psi^{\pm}}{\partial t} = -\mathcal{D}_{\pm}^2 \psi$. We now show that $\lim_{t \to 0^+} \psi^{\pm}(\cdot, t) = \psi_0^{\pm}$ in C^q for all $q \geq 0$. First note that

$$\begin{aligned}
\psi^{\pm}(x,t) &= \int_M \left\langle k^{\pm}(x,y,t), \psi_0^{\pm}(y) \right\rangle_y \ \nu(y) \\
&= \int_M \left\langle \sum_{i=1}^{n^{\pm}} u_{0_i}^{\pm}(x) \otimes u_{0_i}^{\pm}(y) + \sum_{j=1}^{\infty} e^{-\lambda_j t} u_j^{\pm}(x) \otimes u_j^{\pm}(y), \psi_0^{\pm}(y) \right\rangle \ \nu(y) \\
&= \sum_{i=1}^{n^{\pm}} \left(\int_M \left\langle u_{0_i}^{\pm}(y), \psi_0^{\pm}(y) \right\rangle \ \nu(y) \right) u_{0_i}^{\pm}(x) \\
&+ \sum_{j=1}^{\infty} e^{-\lambda_j t} \left(\int_M \left\langle u_j^{\pm}(y), \psi_0^{\pm}(y) \right\rangle \ \nu(y) \right) u_j^{\pm}(x) \\
&= \sum_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm} \right) u_{0_i}^{\pm}(x) + \sum_{j=1}^{\infty} e^{-\lambda_j t} \left(u_j^{\pm}, \psi_0^{\pm} \right) u_j^{\pm}(x),
\end{aligned}$$

where we have used the uniform convergence of the series to interchange sum and integral. It remains to prove that in the C^q norm, the limit as $t \to 0^+$ may be taken under the infinite sum. This is permitted if

$$\sum_{j=1}^{\infty} \left\| \left(u_j^{\pm}, \psi_0^{\pm} \right) u_j^{\pm}(x) \right\|_{C^q} < \infty.$$

Since $\psi_0^{\pm}$ is C^{∞}, for any $l = 0, 1, \ldots$, we have

$$\left(\mathcal{D}_{\pm}^2 \right)^l \psi_0^{\pm} \in C^{\infty}\left(E \otimes \Sigma^{\pm}(M) \right) \subseteq L^2\left(E \otimes \Sigma^{\pm}(M) \right), \text{ and}$$
$$\left(u_j^{\pm}, \left(\mathcal{D}_{\pm}^2 \right)^l \psi_0^{\pm} \right)_{2,0} = \left(\left(\mathcal{D}_{\pm}^2 \right)^l u_j^{\pm}, \psi_0^{\pm} \right)_{2,0} = \lambda_j^l \left(u_j^{\pm}, \psi_0^{\pm} \right)_{2,0}.$$

Thus, for any $l = 0, 1, \ldots$,

$$\sum_{j=1}^{\infty} \lambda_j^{2l} \left| \left(u_j^{\pm}, \psi_0^{\pm} \right)_{2,0} \right|^2 = \left\| \left(\mathcal{D}_{\pm}^2 \right)^l \psi_0^{\pm} \right\|^2 < \infty, \text{ and so}$$
$$\left| \left(u_j^{\pm}, \psi_0^{\pm} \right)_{2,0} \right| \leq K_l \lambda_j^{-l} \text{ for some } K_l > 0.$$

Using $\left\|u_j^{\pm}\right\|_{C^q} \le c_{k(q)} C'_{k(q)} \left(\lambda_j + 1\right)^{[k(q)/2]}$ (see 18.34), we have

$$\sum_{j=1}^{\infty} \left\|\left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}\right\|_{C^q} \le \sum_{j=1}^{\infty} K_l \lambda_j^{-l} \left\|u_j^{\pm}\right\|_{C^q}$$

$$\le \sum_{j=1}^{\infty} K_l c_{k(q)} C'_{k(q)} \lambda_j^{-l} \left(\lambda_j + 1\right)^{[k(q)/2]} < \infty,$$

since we can choose l arbitrarily large and we have shown (see (18.35)) that $\lambda_j \ge C j^{\alpha}$ for some positive constants C and α. Thus, the sum

$$\sum_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm}\right) u_{0_i}^{\pm} + \sum_{j=1}^{\infty} \left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm},$$

which is known to converge in $L^2\left(E \otimes \Sigma^{\pm}(M)\right)$ to $\psi_0^{\pm}$, in fact converges in $C^q\left(E \otimes \Sigma^{\pm}(M)\right)$ to $\psi_0^{\pm}$. By dominated convergence,

$$\sum_{j=1}^{\infty} \left\|e^{-\lambda_j t}\left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}\right\|_{C^q} \le \sum_{j=1}^{\infty} \left\|\left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}\right\|_{C^q} < \infty,$$

implies that we have have the following limit in C^q:

$$\lim_{t \to 0^+} \psi^{\pm}(\cdot, t) = \sum_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm}\right) u_{0_i}^{\pm} + \sum_{j=1}^{\infty} \lim_{t \to 0^+} e^{-\lambda_j t}\left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}$$

$$= \sum_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm}\right) u_{0_i}^{\pm} + \sum_{j=1}^{\infty} \left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm} = \psi_0^{\pm}.$$

The analogous assertions for ψ_0 and ψ are proved in the same way. □

Note that for $x \in M$, the Hermitian inner product $\langle\ ,\ \rangle_x$ on $\left(E \otimes \Sigma(M)\right)_x$ gives us a conjugate-linear map $\psi \mapsto \psi^*(\cdot) := \langle \cdot, \psi\rangle_x$ from $\left(E \otimes \Sigma(M)\right)_x$ to its dual $\left(E \otimes \Sigma(M)\right)_x^*$. Thus, for $t > 0$, we may regard

$$k(x, y, t) \in \operatorname{Hom}\left(\left(E \otimes \Sigma(M)\right)_x, \left(E \otimes \Sigma(M)\right)_y\right),$$

and simailarly for $k^{\pm}(x, y, t)$. For any finite dimensional Hermitian vector space $(V, \langle\cdot,\cdot\rangle)$ with orthonormal basis $e_1, \ldots, e_N$, we have (for $v \in V$)

$$\operatorname{Tr}\left(v^* \otimes v\right) = \sum\nolimits_{i=1}^{N} \left\langle \left(v^* \otimes v\right)\left(e_i\right), e_i\right\rangle = \sum\nolimits_{i=1}^{N} \left\langle v^*\left(e_i\right) v, e_i\right\rangle$$

$$= \sum\nolimits_{i=1}^{N} \left\langle \left\langle e_i, v\right\rangle v, e_i\right\rangle = \sum\nolimits_{i=1}^{N} \left\langle e_i, v\right\rangle \left\langle v, e_i\right\rangle = \sum\nolimits_{i=1}^{N} \left|\left\langle e_i, v\right\rangle\right|^2 = |v|^2.$$

In particular, $k^{\pm}(x, x, t) \in \operatorname{End}\left(\left(E \otimes \Sigma^{\pm}(M)\right)_x\right)$ and

$$\operatorname{Tr}\left(k^{\pm}(x, x, t)\right) = \sum_{i=1}^{n^{\pm}} \left|u_{0_i}^{\pm}(x)\right|^2 + \sum_{j=1}^{\infty} e^{-\lambda_j t}\left|u_j^{\pm}(x)\right|^2.$$

Since this series converges uniformly and $\left\|u_{0_i}^{\pm}\right\|_{2,0} = \left\|u_j^{\pm}\right\|_{2,0} = 1$, we have

$$\int_M Tr\left(k^{\pm}(x, x, t)\right)\ \nu_x = n^{\pm} + \sum_{j=1}^{\infty} e^{-\lambda_j t} < \infty.$$

For $t > 0$, we define the bounded linear operator $e^{-t\mathcal{D}_{\pm}^{2}} \in \mathrm{End}\left(L^{2}\left(E \otimes \Sigma^{\pm}(M)\right)\right)$ by

$$e^{-t\mathcal{D}_{\pm}^{2}}\left(\psi^{\pm}\right) = \sum_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm}\right) u_{0_i}^{\pm} + \sum_{j=1}^{\infty} e^{-\lambda_j t} \left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}.$$

Note that $e^{-t\mathcal{D}_{\pm}^{2}}$ is of trace class, since

$$\mathrm{Tr}\left(e^{-t\mathcal{D}_{\pm}^{2}}\right) = n^{\pm} + \sum_{j=1}^{\infty} e^{-\lambda_j t} = \int_M \mathrm{Tr}\left(k^{\pm}(x,x,t)\right)\ \nu_x < \infty.$$

Now, we have

$$\begin{aligned}
\mathrm{index}\left(\mathcal{D}^{+}\right) &= \dim V_0\left(\mathcal{D}_{+}^{2}\right) - \dim V_0\left(\mathcal{D}_{-}^{2}\right) \\
&= n^{+} - n^{-} + \sum_{j=1}^{\infty}\left(e^{-\lambda_j t} - e^{-\lambda_j t}\right) \\
&= n^{+} + \sum_{j=1}^{\infty} e^{-\lambda_j t} - \left(n^{-} + \sum_{j=1}^{\infty} e^{-\lambda_j t}\right) \\
&= \int_M \mathrm{Tr}\left(k^{+}(x,x,t)\right)\ \nu_x - \int_M \mathrm{Tr}\left(k^{-}(x,x,t)\right)\ \nu_x \\
&= \int_M \left(\mathrm{Tr}\left(k^{+}(x,x,t)\right) - \mathrm{Tr}\left(k^{-}(x,x,t)\right)\right)\ \nu_x. \qquad (18.37)
\end{aligned}$$

Since $\mathcal{D}^2 = \mathcal{D}_{+}^{2} \oplus \mathcal{D}_{-}^{2}$, we also have the trace-class operator $e^{-t\mathcal{D}^2} \in \mathrm{End}\left(L^2\left(E \otimes \Sigma(M)\right)\right)$, whose trace is given by

$$\mathrm{Tr}\left(e^{-t\mathcal{D}^2}\right) = \int_M \mathrm{Tr}\left(k(x,x,t)\right)\ \nu_x = \int_M \left(\mathrm{Tr}\left(k^{+}(x,x,t)\right) + \mathrm{Tr}\left(k^{-}(x,x,t)\right)\right)\ \nu_x.$$

The *supertrace* of $k(x,x,t)$ is defined by

$$\mathrm{Str}\left(k(x,x,t)\right) := \mathrm{Tr}\left(k^{+}(x,x,t)\right) - \mathrm{Tr}\left(k^{-}(x,x,t)\right),$$

and in view of (18.37), we have

$$\mathrm{index}\left(\mathcal{D}^{+}\right) = \int_M \mathrm{Str}\left(k(x,x,t)\right)\ \nu_x. \qquad (18.38)$$

The left side is independent of t and so the right side is also independent of t. The main task now is to determine the behavior of $\mathrm{Str}\left(k(x,x,t)\right)$ as $t \to 0^{+}$. We suspect that for each $x \in M$, as $t \to 0^{+}$, $k(x,x,t)$ and $\mathrm{Str}\left(k(x,x,t)\right)$ are influenced primarily by the geometry (e.g., curvature of M and E) near x, since the heat sources of points far from x are not felt very strongly at x for small t. In the sections to follow, we show that

$$\lim_{t\to 0^{+}} \mathrm{Str}\left(k(x,x,t)\right) = \left\langle \mathrm{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right) \wedge \det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh\left(i\Omega^{\theta}/4\pi\right)}\right)^{\frac{1}{2}}, \nu_x \right\rangle, \qquad (18.39)$$

where the meaning of the right side is explained in the following digression.

The curvature form of the connection ε on E is denoted by Ω^{ε}, while Ω^{θ} is the curvature form of the Levi-Civita connection θ for FM. We have (recall $2m =$

$\dim M$)

$$e^{i\Omega^{\varepsilon}/2\pi} := \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{i}{2\pi}\right)^{k} \Omega^{\varepsilon} \wedge \overset{k}{\cdots} \wedge \Omega^{\varepsilon} = \sum_{k=0}^{m} \frac{1}{k!} \left(\frac{i}{2\pi}\right)^{k} \Omega^{\varepsilon} \wedge \overset{k}{\cdots} \wedge \Omega^{\varepsilon}, \tag{18.40}$$

where $\Omega^{\varepsilon} \wedge \overset{k}{\cdots} \wedge \Omega^{\varepsilon} \in \Omega^{2k}(\mathrm{End}(E))$. Also $\mathrm{Tr}\left(\Omega^{\varepsilon} \wedge \overset{k}{\cdots} \wedge \Omega^{\varepsilon}\right) \in \Omega^{2k}(M)$ and

$$\mathrm{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right) \in \bigoplus_{k=1}^{m} \Omega^{2k}(M).$$

This (by one of many equivalent definitions) is a representative of the total Chern character $\mathbf{ch}(E) \in \bigoplus_{k=1}^{m} H^{2k}(M, \mathbb{Q})$. Now $\Omega^{\theta} \in \Omega^{2}(\mathrm{End}(TM))$ has values in the skew-symmetric endomorphisms of TM. A skew-symmetric endomorphism of $\mathbb{R}^{2m}$, say $B \in \mathfrak{so}(n)$, has pure imaginary eigenvalues $\pm i r_k$, where $r_k \in \mathbb{R}$ $(1 \le k \le m)$. Thus, iB has real eigenvalues $\pm r_k$. Now $\frac{z/2}{\sinh(z/2)}$ is a power series in z with radius of convergence 2π. Thus, $\frac{isB/2}{\sinh(isB/2)}$ is defined for s sufficiently small and has eigenvalues $\frac{r_k s/2}{\sinh(r_k s/2)}$ each repeated twice. Hence

$$\det\left(\frac{isB/2}{\sinh(isB/2)}\right) = \prod_{k=1}^{m} \left(\frac{r_k s/2}{\sinh(r_k s/2)}\right)^{2} \text{ and}$$

$$\det\left(\frac{isB/2}{\sinh(isB/2)}\right)^{\frac{1}{2}} = \prod_{k=1}^{m} \frac{r_k s/2}{\sinh(r_k s/2)}.$$

The last product is a power series in s of the form

$$\prod_{k=1}^{m} \frac{r_k s/2}{\sinh(r_k s/2)} = \sum_{j=0}^{\infty} a_k\left(r_1^2, \dots, r_m^2\right) s^{2k}, \tag{18.41}$$

where the coefficient $a_k\left(r_1^2, \dots, r_m^2\right)$ is a homogeneous, symmetric polynomial in $r_1^2, \dots, r_m^2$ of degree k. One can always express any such a symmetric polynomial as a polynomial in the elementary symmetric polynomials $\sigma_1, \dots, \sigma_m$ in $r_1^2, \dots, r_m^2$, where

$$\sigma_1 = \sum\nolimits_{i=1}^{m} r_i^2,\ \sigma_2 = \sum\nolimits_{i<j}^{m} r_i^2 r_j^2,\ \sigma_2 = \sum\nolimits_{i<j<k}^{m} r_i^2 r_j^2 r_k^2, \dots.$$

These in turn may be expressed in terms of $\mathrm{SO}(n)$-invariant polynomials in the entries of $B \in \mathfrak{so}(n)$ via

$$\det(\lambda I - B) = \prod_{j=1}^{m} (\lambda + i r_j)(\lambda - i r_j) = \prod_{j=1}^{m} \left(\lambda^2 + r_j^2\right)$$
$$= \sum_{k=1}^{m} \sigma_k\left(r_1^2, \dots, r_m^2\right) \lambda^{2(m-k)}.$$

On the other hand,

$$\det(\lambda I - B) = \sum_{k=1}^{m} \left(\frac{1}{(2k)!} \sum_{(i),(j)} \delta_{i_1 \cdots i_{2k}}^{j_1 \cdots j_{2k}} B_{j_1}^{i_1} \cdots B_{j_{2k}}^{i_{2k}}\right) \lambda^{2(m-k)}, \text{ and so}$$

$$\sigma_k\left(r_1^2, \dots, r_m^2\right) = \frac{1}{(2k)!} \sum_{(i),(j)} \delta_{i_1 \cdots i_{2k}}^{j_1 \cdots j_{2k}} B_{j_1}^{i_1} \cdots B_{j_{2k}}^{i_{2k}},$$

where $(i) = (i_1, \cdots, i_{2k})$ is an ordered $2k$-tuple of distinct elements of $\{1, \ldots, 2m\}$ and (j) is a permutation of (i) with sign $\delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}}$. If we replace $B^i{}_j$ with the 2-form $\frac{1}{2\pi}\left(\Omega^\theta\right)^i{}_j$ relative to an orthonormal frame field, we obtain the Pontryagin forms

$$p_k\left(\Omega^\theta\right) = \frac{1}{(2\pi)^{2k}(2k)!}\sum_{(i),(j)} \delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}} \Omega^\theta_{i_1 j_1} \wedge \cdots \wedge \Omega^\theta_{i_{2k} j_{2k}},$$

which represent the Pontryagin classes of the $\mathrm{SO}(n)$ bundle FM. Note that $p_k\left(\Omega^\theta\right)$ is independent of the choice of framing by the *ad*-invariance of the polynomials σ_k. Getting back to (18.39), if we express the $a_k\left(r_1^2, \ldots, r_m^2\right)$ as polynomials, say $\mathcal{A}_k(\sigma_1, \ldots, \sigma_k)$, in the σ_j $(j \le k)$, we can ultimately write

$$\det\left(\frac{isB/2}{\sinh(isB/2)}\right)^{\frac{1}{2}} = \sum_{k=0}^{\infty} \mathcal{A}_k(\sigma_1, \ldots, \sigma_k)\, s^{2k}.$$

Formally replacing B by $\frac{1}{2\pi}\Omega^\theta$, we finally have motivated the definition

$$\det\left(\frac{i\Omega^\theta/4\pi}{\sinh(i\Omega^\theta/4\pi)}\right)^{\frac{1}{2}} := \sum_{k=0}^{\infty} \mathcal{A}_k\left(p_1\left(\Omega^\theta\right), \ldots, p_k\left(\Omega^\theta\right)\right),$$

where the $p_j\left(\Omega^\theta\right)$ are multiplied via wedge product when evaluating the $\mathcal{A}_k\left(p_1\left(\Omega^\theta\right), \ldots, p_k\left(\Omega^\theta\right)\right)$; the order of multiplication does not matter since $p_j\left(\Omega^\theta\right)$ is of even degree $4j$. Also, since $\mathcal{A}_k\left(p_1\left(\Omega^\theta\right), \ldots, p_k\left(\Omega^\theta\right)\right)$ is a $4k$-form, there are only a finite number of nonzero terms in the infinite sum. Abbreviating $p_j\left(\Omega^\theta\right)$ simply by p_j, one finds

$$\begin{aligned}\det\left(\frac{i\Omega^\theta/4\pi}{\sinh(i\Omega^\theta/4\pi)}\right)^{\frac{1}{2}} &= 1 - \frac{1}{24}p_1 + \frac{1}{5760}\left(7p_1^2 - 4p_2\right) \\ &\quad - \frac{1}{967\,680}\left(31p_1^3 - 44p_1p_2 + 16p_3\right) + \cdots .\end{aligned} \tag{18.42}$$

This (by one definition) represents the total $\widehat{A}$-class of M, denoted by

$$\widehat{\mathbf{A}}(M) \in \bigoplus_{k=1}^{m} H^{2k}(M, \mathbb{Q}), \tag{18.43}$$

where actually $\widehat{\mathbf{A}}(M)$ has only nonzero components in $H^{2k}(M, \mathbb{Q})$ when k is even (or $2k \equiv 0 \bmod 4$). In (18.39) the multi-degree forms (18.40) and (18.42) have been wedged, and the top ($2m$-degree) component (relative to the volume form) has been harvested. In terms of characteristic classes, (18.38) and (18.39) yield

$$\operatorname{index}\left(\mathcal{D}^+\right) = \left(\mathbf{ch}(E) \smile \widehat{\mathbf{A}}(M)\right)[M], \tag{18.44}$$

which is the *index formula for the twisted Dirac operator* $\mathcal{D}^+$.

In the next section, we give a different construction of the total twisted spinorial heat kernel, which yields its asymptotic expansion as $t \to 0^+$. In order to deduce that this construction actually yields the same result as (18.36), we introduce the definition of a general heat kernel for $\mathcal{D}^2$. Propositions 18.27 and 18.29 imply that the heat kernel (18.36) satisfies the criteria in this definition. We prove below that any two general heat kernels are the same. Hence, if a differently constructed function also fits this definition, it must agree with the heat kernel (18.36). We set

$$\mathcal{H} := \pi_1^*(E \otimes \Sigma(M))^* \otimes \pi_2^*(E \otimes \Sigma(M)) \to M \times M \times (0, \infty),$$

where $\pi_1(x,y,t) := x$ and $\pi_2(x,y,t) := y$.

DEFINITION 18.30. A **general heat kernel** for $\mathcal{D}^2$ is a section $\kappa \in C^0(\mathcal{H})$, where $\kappa(x,y,t) \in \hom\left((E\otimes\Sigma^{\pm}(M))_x, (E\otimes\Sigma^{\pm}(M))_y\right)$ is C^2 in x, C^1 in t, and $\left(\partial_t + \mathcal{D}_x^2\right)\kappa(x,y,t) = 0$. Moreover, for all $\psi \in C^0(E\otimes\Sigma^{\pm}(M))$, we require

$$\lim_{t\to 0^+}\int_M \kappa(y,x,t)\,\psi(y)\; v_h(y) = \psi(x)\,. \tag{18.45}$$

REMARK 18.31. Here and elsewhere, the phrase "C^2 in x" means not only that $\kappa(x,y,t)$ is C^2 in x for fixed t, but that the second partials with respect to local coordinates for x are jointly continuous in (x,y,t). Also "C^1 in t" means that $\partial_t\kappa(x,y,t)$ is jointly continuous in (x,y,t).

LEMMA 18.32. *Let $\pi : M\times[0,\infty)\to M$ be given by $\pi(x,t) = x$, and let ψ_1, $\psi_2 \in C^0(\pi^*(E\otimes\Sigma(M)))$ be C^2 in x and C^1 in t. Suppose that $\left(\partial_t+\mathcal{D}^2\right)\psi_1 = 0$ and $\left(\partial_t+\mathcal{D}^2\right)\psi_2 = 0$. Then for any $t>0$, we have*

$$\frac{d}{ds}\int_M \langle\psi_1(x,t-s),\psi_2(x,s)\rangle_x\; v_h(x) = 0 \text{ for } s\in(0,t)\,.$$

In other words, $\int_M \langle\psi_1(x,t-s),\psi_2(x,s)\rangle_x\, v_h(x)$ is independent of s in $[0,t]$.

PROOF. For $L := \partial_t + \mathcal{D}^2$,

$$\begin{aligned}
0 &= \langle (L\psi_1)(x,t-s),\psi_2(x,s)\rangle - \langle\psi_1(x,t-s),(L\psi_2)(x,s)\rangle\\
&= \langle -\partial_s(\psi_1(x,t-s)),\psi_2(x,s)\rangle - \langle\psi_1(x,t-s),\partial_s(\psi_2(x,s))\rangle\\
&+ \left\langle \mathcal{D}_x^2\psi_1(x,t-s),\psi_2(x,s)\right\rangle - \left\langle\psi_1(x,t-s),\mathcal{D}_x^2\psi_2(x,s)\right\rangle\\
&= \tfrac{\partial}{\partial s}\left(\langle(\psi_1(x,t-s)),\psi_2(x,s)\rangle\right)\\
&+ \left\langle \mathcal{D}_x^2\psi_1(x,t-s),\psi_2(x,s)\right\rangle - \left\langle\psi_1(x,t-s),\mathcal{D}_x^2\psi_2(x,s)\right\rangle.
\end{aligned}$$

Since $\mathcal{D}^2$ is formally self-adjoint, integrating over M, we then have

$$\begin{aligned}
&\frac{d}{ds}\int_M \langle\psi_1(x,t-s),\psi_2(x,s)\rangle_x\; v_h(x)\\
&= \int_M \tfrac{\partial}{\partial s}\langle\psi_1(x,t-s),\psi_2(x,s)\rangle_x\; v_h(x) = 0,
\end{aligned}$$

since the integrand $\langle\psi_1(x,t-s),\psi_2(x,s)\rangle_x$ is C^1 in t. □

PROPOSITION 18.33. *If κ is a general heat kernel for $\mathcal{D}^2$, then $\kappa(y,x,t) = \kappa(x,y,t)^*$. Moreover, if κ_1 and κ_2 are general heat kernels for $\mathcal{D}^2$, then $\kappa_1 = \kappa_2$.*

PROOF. For any fixed $\alpha\in(E\otimes\Sigma^{\pm}(M))_x$ and $\beta\in(E\otimes\Sigma^{\pm}(M))_y$, let $\psi_1(z,t) := \kappa_1(z,x,t)^*(\alpha)$ and $\psi_2(z,t) := \kappa_2(z,y,t)^*(\beta)$ for $(z,t)\in M\times(0,\infty)$. Then

$$\begin{aligned}
\left(\partial_t+\mathcal{D}_z^2\right)\psi_1 &= \left(\partial_t+\mathcal{D}_z^2\right)\left(\kappa_1(z,x,t)^*(\alpha)\right)\\
&= \left(\left(\partial_t+\mathcal{D}_z^2\right)\kappa_1(z,x,t)\right)^*(\alpha) = 0,
\end{aligned}$$

and similarly $\left(\partial_t + \mathcal{D}_z^2\right)\psi_2 = 0$. According to (18.45),

$$\begin{aligned}
\langle \psi_1(y,t), \beta \rangle &= \left\langle \lim_{s\to 0^+} \int_M \kappa_2(z,y,s)\, \psi_1(z,t-s)\ v_h(z), \beta \right\rangle_y \\
&= \lim_{s\to 0^+} \int_M \langle \kappa_2(z,y,s)\, \psi_1(z,t-s), \beta \rangle_y\ v_h(z) \\
&= \lim_{s\to 0^+} \int_M \left\langle \psi_1(z,t-s), \kappa_2(z,y,s)^*(\beta) \right\rangle_z\ v_h(z) \\
&= \lim_{s\to 0^+} \int_M \langle \psi_1(z,t-s), \psi_2(z,s) \rangle_z\ v_h(z), \text{ and}
\end{aligned}$$

$$\begin{aligned}
\langle \alpha, \psi_2(x,t) \rangle &= \left\langle \alpha, \lim_{s\to t^-} \int_M \kappa_1(z,x,t-s)\, \psi_2(z,s)\ v_h(z) \right\rangle_x \\
&= \lim_{s\to t^-} \int_M \langle \alpha, \kappa_1(z,x,t-s)\, \psi_2(z,s) \rangle_x\ v_h(z) \\
&= \lim_{s\to t^-} \int_M \left\langle \kappa_1(z,x,t-s)^*(\alpha), \psi_2(z,s) \right\rangle_z\ v_h(z) \\
&= \lim_{s\to t^-} \int_M \langle \psi_1(z,t-s), \psi_2(z,s) \rangle_z\ v_h(z).
\end{aligned}$$

Hence, using Lemma 18.32,

$$\begin{aligned}
&\left\langle \kappa_1(y,x,t)^*(\alpha), \beta \right\rangle = \langle \psi_1(y,t), \beta \rangle \\
&= \lim_{s\to 0^+} \int_M \langle \psi_1(z,t-s), \psi_2(z,s) \rangle_z\ v_h(z) \\
&= \lim_{s\to t^-} \int_M \langle \psi_1(z,t-s), \psi_2(z,s) \rangle_z\ v_h(z) \\
&= \langle \alpha, \psi_2(x,t) \rangle = \left\langle \alpha, \kappa_2(x,y,t)^*(\beta) \right\rangle,
\end{aligned}$$

from which we have $\kappa_2(x,y,t) = \kappa_1(y,x,t)^*$. In the case $\kappa_1 = \kappa_2 = \kappa$, we have $\kappa(x,y,t) = \kappa(y,x,t)^*$, and so $\kappa_1(x,y,t) = \kappa_1(y,x,t)^* = \kappa_2(x,y,t)$. □

4. The Asymptotic Formula for the Heat Kernel

The well-known heat kernel (or fundamental solution) for the ordinary heat equation $u_t = \Delta u$ in Euclidean space $\mathbb{R}^n$, is given by

$$e(x,y,t) = (4\pi t)^{-n/2} \exp\left(-|x-y|^2/4t\right). \tag{18.46}$$

Since $H(x,y,t)$ only depends on $r = |x-y|$ and t, it is convenient to write

$$e(x,y,t) = \mathcal{E}(r,t) := (4\pi t)^{-n/2} \exp\left(-r^2/4t\right). \tag{18.47}$$

We do not expect such a simple expression for the heat kernel $k = (k^+, k^-)$ of Definition 18.28 on p. 510. However, we will show that for $x, y \in M$ (of even dimension $n = 2m$) with $r = d(x,y) :=$ Riemannian distance from x to y sufficiently small, we have an asymptotic expansion as $t \to 0^+$ for $k(x,y,t)$ of the form

$$k(x,y,t) \sim H_Q(x,y,t) := \mathcal{E}(r,t) \sum_{j=0}^{Q} h_j(x,y)\, t^j,$$

for any fixed integer $Q > m + 4$, where

$$h_j(x,y) \in \operatorname{Hom}\left((E \otimes \Sigma(M))_x, (E \otimes \Sigma(M))_y\right), \ j \in \{0, 1, \ldots, Q\}.$$

The meaning of $k(x,y,t) \sim H_Q(x,y,t)$ is that for $d(x,y)$ and t sufficiently small,

$$|k(x,y,t) - k_Q(x,y,t)| \le C_Q \mathcal{E}(r,t)\, t^{Q+1} \le C_Q t^{Q-m+1},$$

where C_Q is a constant, independent of (x,y,t). We then will have

$$k(x,x,t) \sim (4\pi t)^{-m} \sum_{j=0}^{Q} h_j(x,x)\, t^j = (4\pi)^{-m} \sum_{j=0}^{Q} h_j(x,x)\, t^{j-m}. \tag{18.48}$$

This may seem a bit odd, since we know from (18.38) that for any $t > 0$,

$$\int_M \operatorname{Str}(k(x,x,t))\ \nu_h(x) = \operatorname{index}\left(\mathcal{D}^+\right),$$

which is constant, independent of t. As $t \to 0^+$, we deduce from (18.48) that

$$\int_M \operatorname{Str}(h_j(x,x))\ \nu_h(x) = 0 \ \text{ for } \ j \in \{0,1,\ldots,m-1\}, \ \text{ while}$$

$$(4\pi)^{-m} \int_M \operatorname{Str}(h_m(x,x))\ \nu_h(x) = \operatorname{index}\left(\mathcal{D}^+\right).$$

Thus, to prove the Local Index Formula, it will suffice to show that

$$(4\pi)^{-m} \operatorname{Str}(h_m(x,x)) = \left\langle \operatorname{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right) \wedge \det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh(i\Omega^{\theta}/4\pi)}\right)^{\frac{1}{2}}, \nu_h(x) \right\rangle.$$

While this may not be the intellectual equivalent of climbing Mount Everest, it is not for the faint of heart. We will find some shortcuts, and we encourage the reader to find other approaches to the summit. In what follows, we develop machinery to a obtain local formula $\mathcal{D}^2$ about a point $x \in M$ using a local section (known as the radial gauge) of the frame bundle, which is as simple as possible. We will then use this to construct what ought to be the asymptotic expansion for the heat kernel of the spinorial heat equation $\psi_t = \mathcal{D}^2\psi$. The proof that this is indeed the asymptotic expansion for the heat kernel will then be carried out. This will be the foundation for the proof of the Local Index Formula in the next chapter.

In this paragraph, we construct the radial gauge. For any $v \in \mathbb{R}^n$, we denote the standard horizontal vector field on FM (with respect to the Levi-Civita connection θ) by $\overline{v}$. Recall that $\overline{v}$ is determined by the conditions

$$\theta(\overline{v}) = 0 \ \text{ and } \ v = \varphi(\overline{v}_u) = u^{-1}(\pi_{F*}(\overline{v}_u)), \text{ for all } u \in FM,$$

where $\varphi \in \overline{\Omega}^1(P, \mathbb{R}^n)$ is the canonical 1-form. Let $C : P \to FM$ be the spin structure for M and let $\pi_{U(E)} : U(E) \to M$ be the unitary frame bundle of E with connection ε. We then have the fibered-product bundle $\pi_{U(E)} \times_f (\pi_F \circ C) : U(E) \times_f P \to M$ with connection $\varepsilon \oplus \widetilde{\theta}$, where $\widetilde{\theta} := c'^{-1}(C^*\theta)$. We may pullback the canonical form φ to a form in $\overline{\Omega}^1(U(E) \times_f P, \mathbb{R}^n)$ via the map $U(E) \times_f P \to P \to FM$. For simplicity, we will denote this pullback of φ by the same symbol φ. Moreover, we can define the notion of the standard horizontal vector field $\overline{v}$ on $U(E) \times_f P$ for $v \in \mathbb{R}^n$ by means of the conditions

$$\left(\varepsilon \oplus \widetilde{\theta}\right)(\overline{v}) = 0 \text{ and } v = \varphi(\overline{v}).$$

For fixed $x \in M$, select a frame $u_x \in (FM)_x$ and some $u'_x \in P$ with $C(u'_x) = u_x$; u'_x is sometimes called a *spinor frame* at x. Also select a (unitary) frame $w'_x \in U(E)_x$. Then let $w_x := (w'_x, u'_x) \in (U(E) \times_f P)_x$. For $v \in \mathbb{R}^n$, let $\eta_v : \mathbb{R} \to U(E) \times_f P$ be the integral curve of $\overline{v}$ with initial point $\eta_v(0) = w_x$. Thus, for $\pi_f : U(E) \times_f P \to M$,

$$\eta'_v(t) = \overline{v}_{\eta_v(t)} \quad \text{and} \quad \varphi(\eta'_v(t)) = \varphi\left(\overline{v}_{\eta_v(t)}\right) = u^{-1}\left(\pi_{f*}\left(\overline{v}_{\eta_v(t)}\right)\right) = v.$$

Define $\zeta : \mathbb{R}^n \to U(E) \times_f P$ by $\zeta(v) = \eta_v(1)$. Note that $\zeta(tv) = \eta_{tv}(1) = \eta_v(t)$ and $\zeta(0) = \eta_v(0) = w_x$. By the implicit function theorem, there is a ball $B(0, r_0) \subseteq \mathbb{R}^n$ ($r_0 > 0$) about $0 \in \mathbb{R}^n$, such that $\zeta|_{B(0,r_0)}$ defines a smoothly embedded submanifold of $U(E) \times_f P$ through w_x and $\pi_f \circ \zeta|_{B(0,r_0)} : B(0, r_0) \to M$ is a diffeomorphism onto its image, say $\mathcal{B}$. There is a unique local section $\sigma : \mathcal{B} \to U(E) \times_f P$ whose image is $\zeta(B(0, r_0))$. This local section σ is known as the *radial gauge for the choice* $(w'_x, u'_x) \in (U(E) \times_f P)_x$; it depends on the connection $\varepsilon \oplus \widetilde{\theta}$. Since $U(E) \times_f P \subset U(E) \times P$, the map σ has two component local sections; i.e.,

$$\sigma = \sigma_1 \times \sigma_2, \text{ where } \sigma_1 : \mathcal{B} \to U(E) \text{ and } \sigma_2 : \mathcal{B} \to P.$$

Using the spin structure $C : P \to FM$, we also have a local section

$$u := C \circ \sigma_2 : \mathcal{B} \to FM \tag{18.49}$$

of the frame bundle (i.e., a frame field). For $y \in \mathcal{B}$ and the standard basis vectors $e_k \in \mathbb{R}^n$, we set

$$E_k(y) := ((C \circ \sigma_2)(y))(e_k). \tag{18.50}$$

PROPOSITION 18.34. *For $v \in \mathbb{R}^n$, the curve $\xi_v := \pi_f \circ \eta_v$ is a geodesic relative to the metric h on M with $\xi'_v(0) = u_x(v)$. If $\xi_v(1) \in \mathcal{B}$, the map*

$$\phi := \left(\pi_f \circ \zeta|_{B(0,r_0)}\right)^{-1} : \mathcal{B} \to B(0, r_0)$$

assigns to the point $\xi_v(1) \in \mathcal{B}$, the components relative to the frame u_x of $\xi'_v(0)$, namely

$$\phi(\xi_v(1)) = u_x^{-1}(\xi'_v(0)) = v. \tag{18.51}$$

In other words, $\phi : \mathcal{B} \to B(0, r_0)$ is a normal coordinate system.

PROOF. We claim that the E_k are parallel along the each of the curves $\xi_v := \pi_f \circ \eta_v$ in the sense that $\nabla^\theta_{\xi'_v} E_k = 0$. Let $u = C \circ \sigma_2$ as in (18.49) and let $\widetilde{E}_k(u(y)) \in T_{u(y)}FM$ be the θ-horizontal lift of $E_k(y)$. We have that $\varphi_{\eta_v(t)}(\widetilde{E}_k) = e_k$, since

$$\begin{aligned} \varphi_{u(\xi_v(t))}(\widetilde{E}_k) &= u(\xi_v(t))^{-1}(\pi_*(\widetilde{E}_k)) = u(\xi_v(t))^{-1}(E_k(\xi_v(t))) \\ &= u(\xi_v(t))^{-1}(u(\xi_v(t))(e_k)) = e_k. \end{aligned}$$

Then using the definition $\nabla^\theta_X Y := \pi_*(\varphi^{-1}(\widetilde{X}[\varphi(\widetilde{Y})]))$,

$$\begin{aligned} \nabla^\theta_{\xi'_v(t)} E_k &= \pi_*\left(\varphi^{-1}_{u(\xi_v(t))}\left(\widetilde{\xi}'_v(t)_{u(\xi_v(t))}\left[\varphi_{u(\xi_v(\cdot))}(\widetilde{E}_k(\eta_v(\cdot)))\right]\right)\right) \\ &= \pi_*\left(\varphi^{-1}_{u(\xi_v(t))}\left(\widetilde{\xi}'_v(t)_{u(\xi_v(t))}[e_k]\right)\right) = 0. \end{aligned}$$

In particular, the curve ξ_v is a geodesic (for the metric h on M), with $\xi_v'(0) = u_x(v)$, since its tangent vector field $\xi_v'(t)$ is a linear combination (with constant coefficients) of the parallel vector fields E_k; explicitly,

$$\begin{aligned}\xi_v'(t) &= (\pi_f \circ \eta_v)'(t) = \pi_{f*}(\eta_v'(t)) = \pi_{f*}\left(\overline{v}_{\eta_v(t)}\right) = \pi_{F*}\left(\overline{v}_{u(\xi_v(t))}\right)\\ &= u(\xi_v(t))(v) \ \ (\text{where we used } v = \varphi(\overline{v}_u) = u^{-1}(\pi_{F*}(\overline{v}_u))\,)\\ &= u(\xi_v(t))\left(\sum\nolimits_{k=1}^n v_k e_k\right) = \sum\nolimits_{k=1}^n v_k u(\xi_v(t))(e_k)\\ &= \sum\nolimits_{k=1}^n v_k (E_k)_{\xi_v(t)}\,. \qquad (18.52)\end{aligned}$$

We get (18.51) from

$$\begin{aligned}&\zeta(v) = \eta_v(1) \Rightarrow (\pi_f \circ \zeta)(v) = (\pi_f \circ \eta_v)(1) = \xi_v(1)\\ &\qquad \Rightarrow \phi(\xi_v(1)) = v = u_x^{-1}(\xi_v'(0))\,.\end{aligned}$$

In other words, ϕ is the restriction of the exponential map from T_xM to M. These are called the *normal coordinates about x relative to the frame u_x*. □

We denote the Cartesian coordinates of points in $B(0, r_0) \subseteq \mathbb{R}^n$ by $\left(y^1, \ldots, y^n\right)$, and we let $r^2 = \left(y^1\right)^2 + \cdots + (y^n)^2$. In what follows, we identify points $y \in \mathcal{B}$ with their coordinate points in $B(0, r_0)$. In particular the original point x, about which all of these constructions started, is $0 \in B(0, r_0)$, and the curve ξ_v is simply given by

$$\xi_v(t) = tv = t(v_1, \ldots, v_n)$$

We use the notation $\partial_j := \partial/\partial y^j$ for the standard coordinate vector fields. By (18.52) we have $u(\xi_v(t))(v) = \xi_v'(t)$, and so

$$E_k(0) = u_x(e_k) = u(\xi_{e_k}(0))(e_k) = \xi_{e_k}'(0) = \tfrac{d}{dt}(te_k)\big|_{t=0} = \partial_k|_{y=0}$$

Thus, the orthonormal vector fields $E_1(y), \ldots, E_n(y)$ coincide with $\partial_1, \ldots, \partial_n$ at $y = 0$, but not necessarily elsewhere in $B(0, r_0)$ unless the metric h is flat, as we shall see. In general, the functions

$$h_{ij}(y) := h_y(\partial_i, \partial_j)$$

are not the constant functions $\delta_{ij} = h_y(E_i, E_j)$, and we will eventually need to determine them to second order about $y = 0$.

The *radial vector field* ∂_r on $\mathcal{B} - \{x\}$ (or $B(0, r_0) \setminus \{0\}$) is defined by

$$(\partial_r)_y := \frac{1}{r}\sum\nolimits_{j=1}^n y^j (\partial_j)_y = \frac{1}{r}\ \tfrac{d}{dt}(y + ty)\big|_{t=0} = \frac{1}{r}\ \tfrac{d}{dt}(1+t)\,y\big|_{t=0} = \frac{1}{r}\xi_y'(1)\,.$$

More generally, for $t > 0$, we have

$$(\partial_r)_{\xi_y(t)} = \frac{1}{|y|}\xi_y'(t)\,,$$

since for $|y| > 0$,

$$\begin{aligned}(\partial_r)_{\xi_y(t)} &= \frac{1}{|\xi_y(t)|}\sum\nolimits_{j=1}^n \xi_y(t)^j (\partial_j)_{\xi_y(t)} = \sum\nolimits_{j=1}^n \frac{\xi_y(t)^j}{|\xi_y(t)|}(\partial_j)_{\xi_y(t)}\\ &= \sum\nolimits_{j=1}^n \frac{ty_j}{|ty|}(\partial_j)_{\xi_y(t)} = \frac{1}{|y|}\sum\nolimits_{j=1}^n y^j (\partial_j)_{\xi_y(t)} = \frac{1}{|y|}\tfrac{d}{dt}(ty) = \frac{1}{|y|}\xi_y'(t)\,.\end{aligned}$$

If $\sigma : U \to U(E) \times_f P$ is a radial gauge for the connection $\varepsilon \oplus \widetilde{\theta}$, then

$$\sigma^* \left(\varepsilon \oplus \widetilde{\theta}\right)(\partial_r) = 0, \text{ since for } y \neq 0,$$

$$\sigma^* \left(\varepsilon \oplus \widetilde{\theta}\right)\left((\partial_r)_y\right) = \left(\varepsilon \oplus \widetilde{\theta}\right)(\sigma_* \partial_r) = \left(\varepsilon \oplus \widetilde{\theta}\right)\left(\sigma_* \left(\frac{1}{r}\xi'_y(1)\right)\right)$$

$$= \frac{1}{r}\left(\varepsilon \oplus \widetilde{\theta}\right)\left(\sigma_*\left(\xi'_y(1)\right)\right) = \frac{1}{r}\left(\varepsilon \oplus \widetilde{\theta}\right)\left(\eta'_y(1)\right) = 0. \tag{18.53}$$

This fact simplifies covariant differentiation of twisted spinor fields in the radial direction. The next pair of basic results will also be very useful.

LEMMA 18.35 (Gauss's Lemma). *Let $t \in \mathbb{R}$ and let $u, v \in \mathbb{R}^n$ be nonzero orthogonal vectors relative to the standard inner product. Then at any point $tu \in B(0, r_0)$, the vectors $\sum_{i=1}^n u^i \partial_i$ and $\sum_{j=1}^n v^j \partial_j$ are orthogonal relative to h. In other words, for $tu \in B(0, r_0)$,*

$$\sum\nolimits_{i,j=1}^n \delta_{ij} u^i v^j = \sum\nolimits_{i,j=1}^n h_{ij}(0) u^i v^j = 0 \Rightarrow \sum\nolimits_{i,j=1}^n h_{ij}(tu) u^i v^j = 0.$$

PROOF. We may assume that $|u| = |v| = 1$. Define $V : \mathbb{R} \times [0, r_0) \to B(0, r_0)$ by

$$V(s,t) = t(u \cos s + v \sin s).$$

Then for any s, the curve $t \mapsto V(s,t)$ is a unit speed geodesic. Let $\partial_t V = V_*(\partial_t)$ and $\partial_s V = V_*(\partial_s)$. These are well-defined vector fields on *span* $(u,v) \cap B(0,r_0) \backslash \{0\}$ with $[\partial_t V, \partial_s V] = V_*([\partial_t, \partial_s]) = 0$. Since $t \mapsto V(s,t)$ is a geodesic for any fixed s, we have $\nabla_{\partial_t V} \partial_t V = 0$. Thus, as the Levi-Civita connection is a metric connection,

$$\partial_t V[h(\partial_t V, \partial_t V)] = 2h(\nabla_{\partial_t V}\partial_t V, \partial_t V) = 0, \text{ so that}$$
$$h_{V(s,t)}(\partial_t V, \partial_t V) = h_{V(s,0)}(\partial_t V, \partial_t V) = 1.$$

Since the Levi-Civita connection is also torsion-free, we then have

$$\partial_t V[h(\partial_t V, \partial_s V)] = h(\nabla_{\partial_t V}\partial_t V, \partial_s V) + h(\partial_t V, \nabla_{\partial_t V}\partial_s V)$$
$$= h(\partial_t V, \nabla_{\partial_t V}\partial_s V) = h(\partial_t V, \nabla_{\partial_s V}\partial_t V + [\partial_t V, \partial_s V])$$
$$= h(\partial_t V, \nabla_{\partial_s V}\partial_t V) = \tfrac{1}{2}\partial_s V[h(\partial_t V, \partial_t V)] = \tfrac{1}{2}\partial_s V[1] = 0.$$

Writing $\sum_{i=1}^n (\cos s) u^i \partial_i + (\sin s) v^i \partial_i \sin s$ simply as $u \cos s + v \sin s$, we have

$$0 = \partial_t V[h(\partial_t V, \partial_s V)] = \tfrac{d}{dt} h_{V(s,t)}((u\cos s + v \sin s), t(-u \sin s + v \cos s))$$
$$\Rightarrow h_{V(s,t)}((u\cos s + v\sin s), t(-u\sin s + v\cos s))$$
$$= h_{V(s,0)}(u\cos s + v\sin s, 0) = 0.$$

In particular, for all $t \in [0, r_0)$, we have $0 = h_{V(0,t)}(u, tv) = t h_{ut}(u,v)$, and so $h_{ut}(u,v) = 0$, as required. □

PROPOSITION 18.36. *If $y = (y^1, \ldots, y^n)$ is a normal coordinate system on $B(0, r_0)$ about $x \in M$ relative to the Riemannian metric h, $\partial_1, \ldots, \partial_n$ are the coordinate vector fields, and $h_{ij}(y) := h_y(\partial_i, \partial_j)$ for $y \in B(0, r_0)$, then*

$$\sum\nolimits_{j=1}^n h_{ij}(y) y^j = y^i. \tag{18.54}$$

If $R_{ikjl}(0)$ denotes the component $R(\partial_i, \partial_k, \partial_j, \partial_l)$ of the Riemann curvature tensor for the Levi-Civita connection at $y = 0$, then

$$h_{ij}(y) = \delta_{ij} - \tfrac{1}{3}\sum_{k,l=1}^{n} R_{ikjl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right),$$

$$h^{ij}(y) = \delta_{ij} + \tfrac{1}{3}\sum_{k,l=1}^{n} R_{ikjl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right),$$

$$h^{\alpha} := (\det \mathbf{h})^{\alpha} = 1 - \tfrac{1}{3}\alpha \sum_{k,l=1}^{n} R_{kl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right). \tag{18.55}$$

Proof. Since $t \mapsto ty$ is a geodesic with speed $|y|$, we have

$$|y|^2 = \sum_{i,j=1}^{n} h_{ty}\left(y^i \partial_i, y^j \partial_j\right) = \sum_{i,j=1}^{n} h_{ij}(ty)\, y^i y^j \text{ for all } t.$$

By Gauss's Lemma, we have

$$\sum_{i,j=1}^{n} h_{ij}(y)\, y^i v^j = 0 \text{ for any } v \in \mathbb{R}^n \text{ with } \langle y, v\rangle = 0.$$

We may write any $w \in \mathbb{R}^n$ as $w = \alpha y + v$ for some $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}\sum_{i,j=1}^{n} h_{ij}(y)\, y^i w^j &= \alpha \sum_{i,j=1}^{n} h_{ij}(y)\, y^i y^j + \sum_{i,j=1}^{n} h_{ij}(y)\, y^i v^j = \alpha |y|^2 \\ &= \langle y, \alpha y + v\rangle = \langle y, w\rangle = \sum_{j=1}^{n} y^j w^j,\end{aligned}$$

from which (18.54) is immediate.

To obtain (18.55), we first differentiate (18.54):

$$\begin{aligned}\delta_{ik} &= \partial_k\left(y^i\right) = \partial_k\left(\sum_{j=1}^{n} h_{ij}(y)\, y^j\right) \\ &= \sum_{j=1}^{n}\left(\partial_k\left(h_{ij}(y)\right) y^j + h_{ij}(y)\, \partial_k y^j\right),\end{aligned}$$

Using the convenient notation $h_{ij,k}(y) = \partial_k\left(h_{ij}(y)\right)$, we then have

$$\delta_{ik} = h_{ik}(y) + \sum_{j=1}^{n} h_{ij,k}(y)\, y^j.$$

Setting $y = 0$, we obtain $h_{ik}(0) = \delta_{ik}$. Differentiating again,

$$\begin{aligned}0 &= \partial_l\left(h_{ik}(y) + \sum_{j=1}^{n} h_{ij,k}(y)\, y^j\right) \\ &= h_{ik,l}(y) + \sum_{j=1}^{n}\left(h_{ij,kl}(y)\, y^j + h_{ij,k}(y)\, \delta_{lj}\right) \\ &= h_{ik,l}(y) + h_{il,k}(y) + \sum_{j=1}^{n} h_{ij,kl}(y)\, y^j,\end{aligned} \tag{18.56}$$

and setting $y = 0$, we obtain

$$h_{il,k}(0) + h_{ik,l}(0) = 0.$$

This implies $h_{il,k}(0) = 0$, since

$$\begin{aligned}h_{il,k}(0) &= h_{li,k}(0) = -h_{lk,i}(0) = -h_{kl,i}(0) \\ &= h_{ki,l}(0) = h_{ik,l}(0) = -h_{il,k}(0).\end{aligned}$$

Thus,

$$h_{ij}(y) = \delta_{ij} + \tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} h_{ij,kl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right).$$

Note that

$$\begin{aligned}
\delta_i^p &= \sum\nolimits_{j=1}^{n} h_{ij}(y)\, h^{jp}(y) \\
&= \sum\nolimits_{j=1}^{n} \left(\delta_{ij} + \tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} h_{ij,kl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right)\right) h^{jp}(y) \\
&= h^{ip}(y) + \tfrac{1}{2}\sum\nolimits_{j,k,l=1}^{n} h_{ij,kl}(0)\, y^k y^l h^{jp}(y) + \mathrm{O}\left(|y|^3\right) \\
&= h^{ip}(y) + \tfrac{1}{2}\sum\nolimits_{j,k,l=1}^{n} h_{ij,kl}(0)\, h^{jp}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right)
\end{aligned}$$

$$\Rightarrow h^{ip}(y) = \delta^{ip} - \tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} h_{ip,kl}(0)\, y^k y^l + \mathrm{O}\left(|y|^3\right). \tag{18.57}$$

Differentiation of (18.56) yields

$$\begin{aligned}
0 &= \partial_p\left(\sum\nolimits_{j=1}^{n} h_{ij,kl}(y)\, y^j + h_{il,k}(y) + h_{ik,l}(y)\right) \\
&= \sum\nolimits_{j=1}^{n} h_{ij,klp}(y)\, y^j + h_{ip,kl}(y) + h_{il,kp}(y) + h_{ik,lp}(y), \text{ and so}
\end{aligned}$$

$$h_{ip,kl} + h_{il,pk} + h_{ik,lp} = 0 \quad \text{at } y = 0. \tag{18.58}$$

This implies that $h_{ij,kl}(0) = h_{kl,ij}(0)$, since at $y = 0$,

$$\begin{aligned}
2h_{ij,kl} &= h_{ij,kl} + h_{ji,kl} = -\left(h_{ik,lj} + h_{il,jk}\right) - \left(h_{jk,li} + h_{jl,ik}\right) \\
&= -\left(h_{ki,lj} + h_{kj,il}\right) - \left(h_{li,jk} + h_{lj,ki}\right) \\
&= h_{kl,ji} + h_{lk,ij} = 2h_{kl,ij}.
\end{aligned} \tag{18.59}$$

We will use this to show that

$$\tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} h_{ij,kl}(0)\, y^k y^l = -\tfrac{1}{3}\sum\nolimits_{k,l=1}^{n} R_{ikjl}(0)\, y^k y^l.$$

We also need to write $R_{ikjl}(0)$ in terms of the derivatives of the h_{ij} at 0. If $\omega \in \Omega^1(B, GL(n))$ denotes the pull-back of the Levi-Civita connection 1-form on LM via the y-coordinate frame field, then

$$R_{ikjl} = (d\omega + \omega \wedge \omega)_{ik}(\partial_j, \partial_l).$$

By (16.47)

$$\omega(\partial_i)^l_j = \Gamma^l_{ij} := \tfrac{1}{2}h^{lk}\left(\partial_i[h_{jk}] + \partial_j[h_{ik}] - \partial_k[h_{ij}]\right), \tag{18.60}$$

which is 0 at $y = 0$, since we have shown $h_{il,k}(0) = 0$. Thus, $\omega \wedge \omega = 0$ at $y = 0$ and

$$R_{ikjl}(0) = R^i{}_{kjl}(0) = \left(d\omega^i_k\right)(\partial_j, \partial_l) = \partial_j\left[\omega(\partial_l)^i_k\right] - \partial_l\left[\omega(\partial_j)^i_k\right].$$

Since $h^{lk} = \delta^{lk} + \mathrm{O}\left(|y|^2\right)$ by (18.57), (18.60) yields

$$\omega(\partial_i)^l_j = \tfrac{1}{2}\left(h_{jl,i} + h_{il,j} - h_{ij,l}\right) + \mathrm{O}\left(|y|^2\right).$$

Thus (where henceforth all terms are evaluated at $y = 0$),

$$\begin{aligned}
R_{ikjl} &= \partial_j \left[\omega\left(\partial_l\right)^i_k\right] - \partial_l \left[\omega\left(\partial_j\right)^i_k\right] \\
&= \partial_j \left[\tfrac{1}{2}\left(h_{ki,l} + h_{li,k} - h_{lk,i}\right)\right] - \partial_l \left[\tfrac{1}{2}\left(h_{ki,j} + h_{ji,k} - h_{jk,i}\right)\right] \\
&= \tfrac{1}{2}\left(h_{ki,lj} + h_{li,kj} - h_{lk,ij}\right) - \tfrac{1}{2}\left(h_{ki,jl} + h_{ji,kl} - h_{jk,il}\right) \\
&= \tfrac{1}{2}\left(h_{li,kj} - h_{lk,ij}\right) - \tfrac{1}{2}\left(h_{ji,kl} - h_{jk,il}\right) \\
&= \tfrac{1}{2}\left(h_{li,kj} - h_{ji,kl} + h_{jk,il} - h_{lk,ij}\right) \\
&= h_{li,kj} - h_{ji,kl}.
\end{aligned}$$

using (18.59). We show

$$h_{ij,lk} = -\tfrac{1}{3}\left(R_{ikjl} + R_{iljk}\right)$$

Indeed,

$$\begin{aligned}
R_{ikjl} + R_{iljk} &= \left(h_{li,kj} - h_{ji,kl}\right) + \left(h_{ki,lj} - h_{ji,lk}\right) \\
&= -2h_{ji,kl} + h_{li,kj} + h_{ki,lj} = -2h_{ji,kl} + h_{il,jk} + h_{ik,lj} \\
&= -2h_{ji,kl} - h_{ij,kl} = -3h_{ji,kl}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} h_{ij,kl} y^k y^l &= -\tfrac{1}{2}\tfrac{1}{3}\sum\nolimits_{k,l=1}^{n}\left(R_{ikjl} + R_{iljk}\right) y^k y^l \\
&= -\tfrac{1}{3}\sum\nolimits_{k,l=1}^{n} R_{ikjl} y^k y^l.
\end{aligned}$$

Now

$$\begin{aligned}
\left(\det \mathbf{h}\right)^{\alpha} &= \det\left(\delta_{ij} - \tfrac{1}{3}R_{ikjl}\left(0\right) y^k y^l + \mathrm{O}\left(\left|y\right|^3\right)\right) \\
&= \det\left(\delta_{ij}\right) - \operatorname{Tr}\left(\tfrac{1}{3}R_{ikjl}\left(0\right) y^k y^l\right) + \mathrm{O}\left(\left|y\right|^3\right) \\
&= 1 - \tfrac{1}{3}R_{ikil}\left(0\right) y^k y^l + \mathrm{O}\left(\left|y\right|^3\right) \text{ and so} \\
\left(\det \mathbf{h}\right)^{\alpha} &= 1 - \tfrac{1}{3}\alpha R_{ikil}\left(0\right) y^k y^l + \mathrm{O}\left(\left|y\right|^3\right).
\end{aligned}$$

□

The Levi-Civita connection $\theta \in \Omega^1\left(FM, \mathfrak{so}\left(n\right)\right)$ is a form on FM and it is the restriction of the linear connection $\omega \in \Omega^1\left(LM, \mathfrak{gl}\left(n\right)\right)$ on LM. For

$$u := C \circ \sigma_2 : \mathcal{B} \to FM,$$

the framing $E_1 := u\left(e_1\right), \ldots, E_n := u\left(e_n\right)$, is a local section of FM, while the coordinate framing $\partial_1, \ldots, \partial_n$ is a local section, say $\mathcal{I} : \mathcal{B} \to LM$. We have the pull-backs $u^*\theta \in \Omega^1\left(\mathcal{B}, \mathfrak{so}\left(n\right)\right)$ and $\mathcal{I}^*\omega \in \Omega^1\left(\mathcal{B}, \mathfrak{gl}\left(n\right)\right)$. Moreover, if ∇^θ denotes the covariant derivative operator for the Levi-Civita connection, we have

$$\nabla^{\omega}_{\partial_i}\partial_j = \sum\nolimits_{k=1}^{n}\left(\left(\mathcal{I}^*\omega\right)\left(\partial_i\right)\right)^k_j \partial_k \text{ and}$$

$$\nabla^{\theta}_{E_i}E_j = \sum\nolimits_{k=1}^{n}\left(\left(u^*\theta\right)\left(E_i\right)\right)^k_j E_k. \tag{18.61}$$

We verify the second of the equations (18.61). The proof of the first equation is completely analogous, and we have shown that

$$\nabla^{\omega}_{\partial_i}\partial_j = \sum\nolimits_{k=1}^{n}\Gamma^k_{ij}\partial_j,$$

where the Γ_{ij}^k are the Christoffel symbols given by (16.47), namely

$$((\mathcal{I}^*\omega)(\partial_i))^k_{\ j} = \Gamma_{ij}^k = \tfrac{1}{2} h^{lk} \left(\partial_i [h_{jk}] + \partial_j [h_{ik}] - \partial_k [h_{ij}] \right).$$

By definition (see (16.34)), for vector fields X and Y,

$$\left(\nabla_X^\theta Y\right)_y := u(y) (D^\theta \left(\varphi(\widetilde{Y})\right)_{u(y)} (\widetilde{X})).$$

Using

$$\begin{aligned} u^* \left(\varphi(\widetilde{E}_j)\right)_y &= \varphi_{u(y)}(\widetilde{E}_j) = u(y)^{-1} \pi_*(\widetilde{E}_j) \\ &= u(y)^{-1} \pi_*(u_{*y}(E_j)) = u(y)^{-1} (E_j) = e_j, \end{aligned}$$

we then have

$$\begin{aligned} \left(\nabla_{E_i}^\theta E_j\right)_y &= u(y) (D^\theta \left(\varphi(\widetilde{E}_j)\right)_{u(y)} \left(\widetilde{E}_i\right)) \\ &= u(y) (D^\theta(\varphi(\widetilde{E}_j))_{u(y)} (u_{*y}(E_i))) \\ &= u(y) (d(\varphi(\widetilde{E}_j))_{u(y)} (u_{*y}(E_i)) + \theta_{u(y)} (u_{*y}(E_i)) \varphi_{u(y)}(\widetilde{E}_j)) \\ &= u(y) \left((u^* d(\varphi(\widetilde{E}_j))_y (E_i) \right) + u(y) \left((u^*\theta)_y (E_i) e_j \right) \\ &= u(y) \left((d(u^*(\varphi(\widetilde{E}_j)))_y (E_i) \right) + u(y) \left((u^*\theta)_y (E_i) e_j \right) \\ &= u(y) ((d(e_j)_y (E_i)) + u(y) \left((u^*\theta)_y (E_i) e_j \right) \\ &= u(y) \left((u^*\theta)_y (E_i) e_j \right) \\ &= u(y) \left(\sum_{k=1}^n \left((u^*\theta)_y (E_i) \right)^k_{\ j} e_k \right) \\ &= \sum\nolimits_{k=1}^n ((u^*\theta)(E_i))^k_{\ j} (E_k)_y . \end{aligned}$$

PROPOSITION 18.37. *We use the notation of Proposition 18.36. Let $\mathcal{I}^*\omega \in \Omega^1(B, GL(n))$ denote the pull-back of the Levi-Civita connection 1-form on LM via the y-coordinate frame field $\partial_1, \ldots, \partial_n$, and let ∇ denote the covariant derivative for the Levi-Civita connection so that*

$$\nabla_{\partial_i} \partial_j = \sum\nolimits_{k=1}^n ((\mathcal{I}^*\omega)(\partial_i))^k_{\ j} \partial_k = \sum_{k=1}^n \Gamma_{ij}^k(y) \partial_k.$$

Then the Christoffel symbols $\Gamma_{ij}^k(y) := \omega(\partial_i)^k_{\ j}$ obey

$$\Gamma_{ij}^k(y) = \tfrac{1}{3} \sum\nolimits_{p=1}^n \left(R_{ipjk}(0) + R_{ikjp}(0) \right) y^p + \mathrm{O}\left(|y|^2 \right). \tag{18.62}$$

Moreover, if $E_j(y) := u(e_j) := (C \circ \sigma_2)_y (e_j)$ $(j = 1, \ldots n)$ is the radial framing (which equals $\partial_1, \ldots, \partial_n$ at $y = 0$ and is parallel along the geodesics through $y = 0$), then

$$E_j(y) = \sum\nolimits_{k=1}^n \left(\delta_j^k + \tfrac{1}{6} \sum\nolimits_{i,l=1}^n R_{iklj}(0) y^i y^l + O\left(|y|^3 \right) \right) \partial_k, \tag{18.63}$$

and

$$\nabla_{E_i} E_j = \sum\nolimits_{k=1}^{n} \left((u^*\theta)_y (E_i) \right)_j^k E_k$$
$$= \tfrac{1}{2} \sum\nolimits_{k,l=1}^{n} R_{kjli}(0)\, y^l E_k + O\left(|y|^2\right). \tag{18.64}$$

PROOF. Using the summation convention and (18.58), we have

$$\begin{aligned}
\Gamma_{ij}^k &= \tfrac{1}{2} h^{lk} \left(\partial_i [h_{jl}] + \partial_j [h_{il}] - \partial_l [h_{ij}] \right)(y) \\
&= \tfrac{1}{2} h^{lk} \left(h_{jl,i} + h_{il,j} - h_{ij,l} \right)(y) \\
&= \tfrac{1}{2} \delta^{lk} \left(h_{jl,ip} + h_{il,jp} - h_{ij,lp} \right) y^p + \mathrm{O}\left(|y|^2\right) \\
&= \tfrac{1}{2} \left(h_{jk,ip} + h_{ik,jp} - h_{ij,kp} \right) y^p + \mathrm{O}\left(|y|^2\right) \\
&= \tfrac{1}{2} \left(-h_{kp,ij} - h_{ij,kp} \right) y^p + \mathrm{O}\left(|y|^2\right) \\
&= -h_{kp,ij} y^p + \mathrm{O}\left(|y|^2\right) = -h_{ij,kp} y^p + \mathrm{O}\left(|y|^2\right) \\
&= \tfrac{1}{3} \left(R_{ipjk}(0) + R_{ikjp}(0) \right) y^p + \mathrm{O}\left(|y|^2\right),
\end{aligned}$$

and so (18.62) holds.

Since $E_j = \partial_j$ at $y = 0$, there are functions b_{jl}^k, such that

$$E_j(y) = \left(\delta_j^k + b_{jl}^k(y)\, y^l \right) \partial_k.$$

As the E_j are parallel in the ∂_r direction, for all $y \in B(0, r_0)$ we have

$$\begin{aligned}
0 &= r\nabla_{\partial_r} E_j(y) = r\nabla_{\partial_r} \left(\partial_j + b_{jl}^k(y)\, y^l \partial_k \right) \\
&= \nabla_{y^i \partial_i} \left(\partial_j + b_{jl}^k(y)\, y^l \partial_k \right) \\
&= y^i \nabla_{\partial_i} \left(\partial_j + b_{jl}^k(y)\, y^l \partial_k \right) \\
&= y^i \left(\nabla_{\partial_i} \partial_j + \nabla_{\partial_i} \left(b_{jl}^k(y)\, y^l \partial_k \right) \right) \\
&= \Gamma_{ij}^k(y)\, y^i \partial_k + b_{jl}^k(y)\, y^i y^l \nabla_{\partial_i} \partial_k + \partial_i \left(b_{jl}^k(y)\, y^l \right) y^i \partial_k \\
&= \Gamma_{ij}^k(y)\, y^i \partial_k + b_{jl}^k(y)\, y^i y^l \Gamma_{ik}^p(y)\, \partial_p + \left(\partial_i \left(b_{jl}^k(y) \right) y^i y^l + b_{ji}^k(y)\, y^i \right) \partial_k \\
&= \left(b_{ji}^k(y)\, y^i + \Gamma_{ij}^k(y)\, y^i + \partial_i \left(b_{jl}^k(y) \right) y^i y^l + b_{jl}^q(y)\, \Gamma_{iq}^k(y)\, y^i y^l \right) \partial_k.
\end{aligned}$$

Since $\Gamma_{ij}^k(y) = \mathrm{O}(|y|)$, all terms except possibly $b_{ji}^k(y)\, y^i$ are $\mathrm{O}\left(|y|^2\right)$. As the entire expression vanishes to all orders, we must have $b_{ji}^k(0) = 0$, and so

$$b_{ji}^k(y) = b_{ji,l}^k(0)\, y^l + \mathrm{O}\left(|y|^2\right).$$

Considering the second-order part and using (18.62), we get

$$\begin{aligned}
0 &= \left(\tfrac{1}{3} \left(R_{iljk}(0) + R_{ikjl}(0) \right) + b_{ji,l}^k(0) + b_{jl,i}^k(0) \right) y^i y^l \\
&= \left(-\tfrac{1}{3} R_{iklj}(0) + b_{ji,l}^k(0) + b_{jl,i}^k(0) \right) y^i y^l, \text{ and so} \\
b_{ji,l}^k(0) &+ b_{jl,i}^k(0) = \tfrac{1}{6} \left(R_{iklj}(0) + R_{lkij}(0) \right).
\end{aligned}$$

Hence,

$$\begin{aligned}
E_j(y) &= \left(\delta_j^k + b_{jl}^k(y)\, y^l\right)\partial_k = \left(\delta_j^k + b_{jl,i}^k(0)\, y^i y^l + O\left(|y|^3\right)\right)\partial_k \\
&= \left(\delta_j^k + \tfrac{1}{2}\left(b_{jl,i}^k(0) + b_{ji,l}^k(0)\right) y^i y^l + O\left(|y|^3\right)\right)\partial_k \\
&= \left(\delta_j^k + \tfrac{1}{6} R_{iklj}\, y^i y^l + O\left(|y|^3\right)\right)\partial_k,
\end{aligned}$$

and we have (18.63). As for (18.64), we have, modulo O $\left(|y|^2\right)$,

$$\begin{aligned}
\nabla_{E_i} E_j &= \nabla_{\partial_i} E_j \\
&= \nabla_{\partial_i}\partial_j + \nabla_{\partial_i}\left(\left(\tfrac{1}{6}\sum\nolimits_{k,p,l=1}^{n} R_{pklj}(0)\, y^p y^l\right)\partial_k\right) \\
&= \nabla_{\partial_i}\partial_j + \partial_i\left(\tfrac{1}{6}\sum\nolimits_{k,p,l=1}^{n} R_{pklj}(0)\, y^p y^l\right)\partial_k \\
&= \nabla_{\partial_i}\partial_j + \left(\tfrac{1}{6}\sum\nolimits_{k,l=1}^{n} R_{iklj}(0)\, y^l + \tfrac{1}{6}\sum\nolimits_{k,p=1}^{n} R_{pkij}(0)\, y^p\right)\partial_k \\
&= \nabla_{\partial_i}\partial_j + \tfrac{1}{6}\sum\nolimits_{k,l=1}^{n}\left(R_{iklj}(0) + R_{lkij}(0)\right) y^l \partial_k \\
&= \sum\nolimits_{k=1}^{n} \Gamma_{ij}^k(y)\,\partial_k + \tfrac{1}{6}\sum\nolimits_{k,l=1}^{n}\left(R_{iklj}(0) + R_{lkij}(0)\right) y^l \partial_k \\
&= \sum\nolimits_{k=1}^{n} \tfrac{1}{3}\sum\nolimits_{l=1}^{n}\left(R_{iljk}(0) + R_{ikjl}(0)\right) y^l \partial_k \\
&+ \tfrac{1}{6}\sum\nolimits_{k,l=1}^{n}\left(R_{iklj}(0) + R_{lkij}(0)\right) y^l \partial_k \\
&= \sum\nolimits_{k,l=1}^{n}\left(\tfrac{1}{3}\left(R_{iljk}(0) + R_{ikjl}(0)\right) + \tfrac{1}{6}\left(R_{iklj}(0) + R_{lkij}(0)\right)\right) y^l \partial_k \\
&= \sum\nolimits_{k,l=1}^{n}\left(\tfrac{1}{3} R_{iljk}(0) + \tfrac{1}{6} R_{ikjl}(0) + \tfrac{1}{6} R_{lkij}(0)\right) y^l \partial_k \\
&= \sum\nolimits_{k,l=1}^{n}\left(\tfrac{1}{3} R_{kjli}(0) - \tfrac{1}{6} R_{kijl}(0) - \tfrac{1}{6} R_{klij}(0)\right) y^l \partial_k \\
&= \sum\nolimits_{k,l=1}^{n}\left(\tfrac{1}{2} R_{kjli}(0) - \tfrac{1}{6}\left(R_{kjli}(0) + R_{kijl}(0) + R_{klij}(0)\right)\right) y^l \partial_k \\
&= \tfrac{1}{2}\sum\nolimits_{k,l=1}^{n} R_{kjli}(0)\, y^l E_k
\end{aligned}$$

□

A section $\psi \in C^\infty(E \otimes \Sigma(M))$ can be identified with an equivariant function $\widetilde{\psi} \in \overline{\Omega}^0\left(U(E) \times_f P, \mathbb{C}^N \otimes \Sigma_{2m}\right)$. The local section $\sigma : \mathcal{B} \to U(E) \times_f P$ allows us form the pullback $\sigma^* \widetilde{\psi} = \widetilde{\psi} \circ \sigma \in C^\infty\left(\mathcal{B}, \mathbb{C}^N \otimes \Sigma_{2m}\right)$ and pulling back once more via $\phi^{-1} : B(0, r_0) \to \mathcal{B}$, we have $\widetilde{\psi} \circ \sigma \circ \phi^{-1} \in C^\infty\left(B(0, r_0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$. This allows us to locally treat the twisted spinor field $\psi \in C^\infty(E \otimes \Sigma(M))$ as a function on $B(0, r_0)$ with values in the single vector space $\mathbb{C}^N \otimes \Sigma_{2m}$. In particular, expressions such as $\partial_i \psi$, which are encountered in the following results then make sense. For $y \in B(0, r_0)$, let $\mathbf{h}(y)$ denote the matrix whose entries are $h_{ij}(y) = h_y(\partial_i, \partial_j)$. It is customary and convenient to use the abusive notation

$$\sqrt{h(y)} := \sqrt{\det \mathbf{h}(y)} = \sqrt{\det[h_y(\partial_i, \partial_j)]},$$

so that $\sqrt{h}dy^1 \wedge \cdots \wedge dy^n$ is the volume element for h on $B(0, r_0)$. To save vertical space, we set

$$h^{-1/2} := \frac{1}{\sqrt{h}}$$

Moreover, we adhere to the summation convention whereby sums are automatically taken over repeated indices if they are not summed explicitly.

Recall that in terms of a local orthonormal framing $E_1, \ldots, E_n$ we have

$$\mathcal{D}\psi = \sum_j \varphi_j \cdot \nabla_{E_j} \psi = \sum_j E_j \cdot \nabla_{E_j} \psi.$$

This is independent of the local orthonormal framing. However, if take $E_1, \ldots, E_n$ to be the radial framing and identify ψ and $\mathcal{D}\psi$ with $\widetilde{\psi} \circ \sigma \circ \phi^{-1}$ and $\widetilde{\mathcal{D}\psi} \circ \sigma \circ \phi^{-1} \in C^\infty\left(B(0, r_0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$, then

$$\widetilde{\mathcal{D}\psi} \circ \sigma \circ \phi^{-1} = \sum_j \gamma^j \left(d\left(\widetilde{\psi} \circ \sigma \circ \phi^{-1}\right)(E_j) + \sigma^*\left(\varepsilon \oplus \widetilde{\theta}\right)(E_j)\left(\widetilde{\psi} \circ \sigma \circ \phi^{-1}\right)\right).$$

Of course it would be cumbersome to maintain the notation $\widetilde{\psi} \circ \sigma \circ \phi^{-1}$, and so we will simply denote $\widetilde{\psi} \circ \sigma \circ \phi^{-1} \in C^\infty\left(B(0, r_0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$ by ψ, in which case

$$\begin{aligned}\mathcal{D}\psi &= \sum_j \gamma^j E_j[\psi] + \gamma^j \sigma^*\left(\varepsilon \oplus \widetilde{\theta}\right)(E_j)\psi \\ &= \sum_j \gamma^j E_j[\psi] + \gamma^j (\sigma^*\varepsilon)(E_j)\psi + \gamma^j \left(\sigma^*\widetilde{\theta}\right)(E_j)\psi.\end{aligned}$$

Note that ψ can be written locally a sum $\psi(y) = \sum_{p=1}^N \sum_{q=1}^{2^m} a_{p,q}(y)\, v_p \otimes \phi_q$, where $a_{p,q} \in C^\infty(B(0, r_0), \mathbb{C})$, $v_p \in \mathbb{C}^N$ and $\phi_q \in \Sigma_{2m}$. Also, $(\sigma^*\varepsilon)_y(\partial_j) \in \operatorname{End}\left(\mathbb{C}^N\right)$ acts only on v_p in $v_p \otimes \phi_q$, while $\left(\sigma^*\widetilde{\theta}\right)_y(\partial_j) \in \operatorname{End}(\not\prec_{2m})$, as well as γ^j, act only on ϕ_q in $v_p \otimes \phi_q$.

At $y = 0$, $E_j = \partial_j$ and $\sigma^*\left(\varepsilon(E_j) \oplus \widetilde{\theta}(E_j)\right) = 0$. Hence we simply have

$$(\mathcal{D}\psi)(0) = \sum_j \gamma^j (\partial_j \psi)(0)$$

We will need the lead-order terms of $(\sigma^*\varepsilon)_y(E_j)$ and $\left(\sigma^*\widetilde{\theta}\right)_y(E_j)$ as functions of y. Since $(\sigma^*\varepsilon)_0(\partial_j) = 0$ in the radial gauge, we have

$$(\sigma^*\varepsilon)_y(\partial_j) = \sum\nolimits_{i=1}^n \varepsilon_{ij} y^i + O\left(|y|^2\right) \quad \text{or} \quad (\sigma^*\varepsilon)_y = \sum\nolimits_{i=1}^n \left(\varepsilon_{ij} y^i + O\left(|y|^2\right)\right) dy^j$$

for some $\varepsilon_{ij} \in \operatorname{End}\left(\mathbb{C}^N\right)$. At $y = 0$, the pull-back of the curvature is

$$\begin{aligned}(\sigma^*\Omega^\varepsilon)_0 &= d(\sigma^*\varepsilon)_0 + (\sigma^*\varepsilon)_0 \wedge (\sigma^*\varepsilon)_0 = d(\sigma^*\varepsilon)_0 \\ &= \sum\nolimits_{i=1}^n d\left(\varepsilon_{ij} y^i dy^j\right) = \sum\nolimits_{i,j=1}^n \varepsilon_{ij} dy^i \wedge dy^j.\end{aligned}$$

Denoting the pull-back $\sigma^*\Omega^\varepsilon$ by F, we then have

$$\varepsilon_{ij} = F_{ij} := (\sigma^*\Omega^\varepsilon)_0(\partial_i, \partial_j).$$

Since $E_j = \partial_j + O\left(|y|^2\right)$ by (18.63),

$$(\sigma^*\varepsilon)_y(E_j) = (\sigma^*\varepsilon)_y(\partial_j) = \sum\nolimits_{i=1}^n F_{ij} y^i + O\left(|y|^2\right). \tag{18.65}$$

Using (18.3) and (18.64), we have

$$
\begin{aligned}
\left(\sigma^*\widetilde{\theta}\right)_y(E_j) &= \left(\sigma_2^*\left(c^{-1}C^*\theta\right)\right)_y(E_j) = c^{-1}\left(\left(\sigma_2^*\circ C^*\right)\theta\right)_y(E_j)\\
&= c^{-1}\left(\left(C\circ\sigma_2\right)^*\theta\right)_y(E_j) = c^{-1}\left(u^*\theta\right)_y(E_j)\\
&= -\sum\nolimits_{k,l=1}^{n}\tfrac{1}{4}\gamma^k\gamma^l\left(\left(u^*\theta\right)_y(E_j)\right)_{kl}\\
&= -\sum\nolimits_{k,l=1}^{n}\tfrac{1}{4}\gamma^k\gamma^l\tfrac{1}{2}\sum\nolimits_{i=1}^{n}R_{klij}(0)\,y^i\\
&= -\tfrac{1}{8}\sum\nolimits_{i,k,l=1}^{n}\gamma^k\gamma^l R_{klij}(0)\,y^i. \qquad (18.66)
\end{aligned}
$$

PROPOSITION 18.38. *For* $\psi\in C^\infty\left(B(0,r_0),\mathbb{C}^N\otimes\Sigma_{2m}\right)$, *we have*

$$\Delta\psi = h^{-1/2}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\nabla_{\partial_i}\psi\right).$$

PROOF. Using (16.49), p. 393, we have

$$
\begin{aligned}
\Delta\psi &= h^{ij}\nabla^2_{\partial_i,\partial_j}\psi = h^{ij}\nabla_{\partial_j}\left(\nabla_{\partial_i}\psi\right) - h^{ij}\nabla\psi\left(\nabla_{\partial_j}\partial_i\right)\\
&= h^{ij}\nabla_{\partial_j}\left(\nabla_{\partial_i}\psi\right) - \nabla\psi\left(h^{ij}\nabla_{\partial_j}\partial_i\right)\\
&= h^{ij}\partial_j\left(\nabla_{\partial_i}\psi\right) + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\nabla_{\partial_i}\psi\right) + \nabla\psi\left(h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}\right)\partial_i\right)\\
&= h^{ij}\partial_j\left(\nabla_{\partial_i}\psi\right) + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\nabla_{\partial_i}\psi\right) + h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}\right)\nabla_{\partial_i}\psi\\
&= h^{-1/2}\left(h^{ij}\sqrt{h}\partial_j\left(\nabla_{\partial_i}\psi\right) + \partial_j\left(h^{ij}\sqrt{h}\right)\nabla_{\partial_i}\psi\right)\\
&+ h^{-1/2}h^{ij}\sqrt{h}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\nabla_{\partial_i}\psi\right)\\
&= h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}\nabla_{\partial_i}\psi\right) + h^{-1/2}h^{ij}\sqrt{h}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\nabla_{\partial_i}\psi\right)\\
&= h^{-1/2}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\nabla_{\partial_i}\psi\right).
\end{aligned}
$$

□

We will generalize this a bit (see Proposition 18.40 below), but first we need

PROPOSITION 18.39. *Let* $F: B(0,r_0)\to\mathbb{R}$ *be of the form* $F(y) = f(\|y\|) = f(r)$. *Then*

$$\Delta F := h^{ij}\nabla^2_{\partial_i,\partial_j}(F) = f''(r) + \left(h^{-1/2}\partial_r[\sqrt{h}] + \frac{n-1}{r}\right)f'(r). \qquad (18.67)$$

Proof. By the same computation as in the proof of Proposition 18.38, we have $\Delta F = h^{-1/2}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\nabla_{\partial_i}F\right) = h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}\partial_i F\right)$. Then

$$
\begin{aligned}
&h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}\partial_i F\right) = h^{-1/2}\partial_j\left(h^{ij}\sqrt{h}f'(r)\frac{y^i}{r}\right)\\
&= h^{-1/2}\partial_j\left(y^j\sqrt{h}\frac{f'(r)}{r}\right) \text{ (using (18.36))}\\
&= h^{-1/2}y^j\partial_j\left(\sqrt{h}\frac{f'(r)}{r}\right) + h^{-1/2}\partial_j\left(y^j\right)\left(\sqrt{h}\frac{f'(r)}{r}\right)\\
&= y^j\partial_j\left(\frac{f'(r)}{r}\right) + h^{-1/2}y^j\partial_j\left(\sqrt{h}\right)\frac{f'(r)}{r} + \frac{nf'(r)}{r}\\
&= y^j\left(\frac{rf''(r)-f'(r)}{r^2}\frac{y_j}{r}\right) + h^{-1/2}y^j\partial_j\left(\sqrt{h}\right)\frac{f'(r)}{r} + \frac{nf'(r)}{r}\\
&= f''(r) - \frac{f'(r)}{r} + h^{-1/2}\partial_r\left(\sqrt{h}\right)f'(r) + \frac{nf'(r)}{r}\\
&= f''(r) + \left(h^{-1/2}\partial_r[\sqrt{h}] + \frac{n-1}{r}\right)f'(r).
\end{aligned}
$$

□

Proposition 18.40. *Let $F : B(0,r_0) \to \mathbb{R}$ be of the form $F(y) = f(\|y\|) = f(r)$. Then for $\psi \in C^\infty\left(B(0,r_0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$ as above, we have (where $r = |y|, y \in B(0,r_0)$)*

$$
\begin{aligned}
\Delta(F\psi)(y) &= (\Delta F)\psi + 2h^{ij}\left(\nabla_{\partial_i}F\right)\nabla_{\partial_j}\psi + F\Delta\psi\\
&= f''(r)\psi(y) + \left(h^{-1/2}\partial_r[\sqrt{h}] + \frac{n-1}{r}\right)f'(r)\psi(y)\\
&\quad + 2f'(r)\partial_r\psi + f(r)\Delta\psi.
\end{aligned}
\tag{18.68}
$$

Proof. By repeated use of the product rule,

$$
\begin{aligned}
\Delta(F\psi) &= h^{-1/2}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\nabla_{\partial_i}(F\psi)\right)\\
&= h^{-1/2}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\left(\nabla_{\partial_i}F\right)\psi + h^{ij}\sqrt{h}F\nabla_{\partial_i}\psi\right)\\
&= (\Delta F)\psi + 2h^{ij}\left(\nabla_{\partial_i}F\right)\nabla_{\partial_j}\psi + F\Delta\psi.
\end{aligned}
$$

In view of Proposition 18.39, it remains to compute

$$
\begin{aligned}
&h^{ij}(y)\left(\nabla_{\partial_i}F\right)\nabla_{\partial_j}\psi = h^{ij}(y)\left(\partial_r(f)\partial_i r\right)\nabla_{\partial_j}\psi = h^{ij}(y)\left(\partial_r(f)y^i/r\right)\nabla_{\partial_j}\psi\\
&= h^{ij}(y)y^i\frac{\partial_r(f)}{r}\nabla_{\partial_j}\psi = \partial_r(f)\nabla_{\frac{y^j}{r}\partial_j}\psi = \partial_r(f)\nabla_{\partial_r}\psi,
\end{aligned}
$$

where we have used Proposition 18.36. Note that (18.54) says that y is an eigenvector of the matrix $[h_{ij}(y)]$, and hence y is also eigenvector of the inverse matrix $[h^{ij}(y)]$. Finally note that $\nabla_{\partial_r}\psi = \partial_r\psi + \sigma^*\left(\varepsilon \oplus \widetilde{\theta}\right)(\partial_r)\psi = \partial_r\psi$ in the radial gauge by (18.53). □

Let $\Delta_e = \partial_1^2 + \cdots + \partial_n^2$ be the usual Laplace operator in $\mathbb{R}^n$ with coordinates $\left(y^1,\ldots,y^n\right)$ and $\partial_i := \partial/\partial y^i$. Setting $x = 0$ in (18.47), p. 517), the fundamental solution of $u_t = \Delta_e u$ in Euclidean n-space is

$$\mathcal{E}(r,t) := (4\pi t)^{-n/2} \exp\left(-\tfrac{1}{4} r^2/t\right), \text{ for } t > 0. \tag{18.69}$$

where $r^2 = \left(y^1\right)^2 + \ldots + (y^n)^2$. Using Proposition 18.39,

$$\partial_t \mathcal{E} = \Delta_e \mathcal{E} = \partial_r^2 \mathcal{E} + \frac{n-1}{r} \partial_r \mathcal{E}. \tag{18.70}$$

If Δ denotes the Laplace operator on $\mathcal{B}$ for the metric h, then by (18.67),

$$\Delta \mathcal{E} = \partial_r^2 \mathcal{E} + \left(h^{-1/2} \partial_r[\sqrt{h}] + \frac{n-1}{r} \right) \partial_r \mathcal{E} = \Delta_e \mathcal{E} + h^{-1/2} \partial_r[\sqrt{h}] \partial_r \mathcal{E}. \tag{18.71}$$

For $0 \le Q \in \mathbb{Z}$, let $\Psi_Q \in C^\infty\left(\mathcal{B} \times (0,\infty), \mathbb{C}^N \otimes \Sigma_{2m}\right)$ be of the form

$$\Psi_Q(y,t) := \mathcal{E}(r,t) \sum\nolimits_{k=0}^{Q} U_k(y)\, t^k,$$

where $U_k \in C^\infty\left(\mathcal{B}, \mathbb{C}^N \otimes \Sigma_{2m}\right)$. If $U_0(0) \in \mathbb{C}^N \otimes \Sigma_{2m}$ is arbitrarily specified, we seek a formula for $U_k(y)$, $k = 0, \ldots, Q$, such that

$$\left(\mathcal{D}^2 + \partial_t\right) \Psi_Q(y,t) = \mathcal{E}(r,t)\, t^Q \mathcal{D}^2 (U_Q)(y), \tag{18.72}$$

where the square of the Dirac operator $\mathcal{D}$ is given by

$$\mathcal{D}^2 \psi = -\Delta \psi + \tfrac{1}{2} \sum\nolimits_{j,k} \Omega^\varepsilon_{jk} E_j \cdot E_k \cdot \psi + \tfrac{1}{4} S \psi,$$

as computed in Proposition 18.25, p. 506. It is convenient to define the 0-th order operator $\mathcal{F}$ on $C^\infty\left(\mathcal{B}, \mathbb{C}^N \otimes \Sigma_{2m}\right)$ via

$$\begin{aligned} &\mathcal{F}[\psi] := \tfrac{1}{2} \sum\nolimits_{j,k} \Omega^\varepsilon_{jk} E_j \cdot E_k \cdot \psi, \text{ so that} \\ &\mathcal{D}^2 = -\Delta \psi + \left(\mathcal{F} + \tfrac{1}{4} S\right)[\psi]. \end{aligned}$$

The desired formula for the $U_k(y)$ involves the operator A on $C^\infty\left(\mathcal{B}, \mathbb{C}^N \otimes \Sigma_{2m}\right)$ given (where $h^{1/4} := (\sqrt{h})^{1/2}$ and $h^{-1/4} := (h^{-1/2})^{1/2}$) by

$$A[\psi] := -h^{1/4} \mathcal{D}^2 \left[h^{-1/4} \psi\right] = h^{1/4} \Delta \left[h^{-1/4} \psi\right] - \left(\mathcal{F} + \tfrac{1}{4} S\right)[\psi].$$

For $s \in [0,1]$, let

$$A_s[\psi](y) := A[\psi](sy).$$

Proposition 18.41. *Let $U_0(0) \in \mathbb{C}^N \otimes \Sigma_{2m}$, and let $V_0 \in C^\infty\left(\mathcal{B}, \mathbb{C}^N \otimes \Sigma_{2m}\right)$ be the constant function $V_0(y) \equiv U_0(0)$. Then the $U_k(y)$ which satisfy (18.72) are given by*

$$\begin{aligned} &U_k(y) = h(y)^{-1/4} V_k(y), \text{ where} \\ &V_k(y) = \int_{I^k} \prod\nolimits_{i=0}^{k-1} (s_i)^i \left(A_{s_{k-1}} \circ \cdots \circ A_{s_0}[V_0]\right)(y)\, ds_0 \ldots ds_{k-1}, \end{aligned} \tag{18.73}$$

and where $I^k = \{(s_0, \ldots s_k) : s_i \in [0,1], i \in \{0, \ldots, k-1\}\ \}$.

PROOF. Let $\Sigma(y,t) := \sum_{k=0}^{Q} U_k(y)\, t^k$. Using (18.68) and (18.71), we compute

$$\begin{aligned}
&\left(\partial_t + \mathcal{D}^2\right)[\Psi_Q(y,t)] = \left(\partial_t - \Delta + \mathcal{F} + \tfrac{1}{4}S\right)[\mathcal{E}(r,t)\,\Sigma(y,t)] \\
&= \partial_t[\mathcal{E}(r,t)]\,\Sigma(y,t) + \mathcal{E}(r,t)\,\partial_t\Sigma(y,t) \\
&- \Delta[\mathcal{E}(r,t)\,\Sigma(y,t)] + \left(\mathcal{F} + \tfrac{1}{4}S\right)\mathcal{E}(r,t)\,\Sigma(y,t) \\
&= \partial_t[\mathcal{E}(r,t)]\,\Sigma(y,t) + \mathcal{E}(r,t)\,\partial_t\Sigma(y,t) \\
&- \Delta(\mathcal{E}(r,t))\,\Sigma(y,t) - 2\partial_r(\mathcal{E}(r,t))\,\partial_r\Sigma(y,t) - \mathcal{E}(r,t)\,\Delta(\Sigma(y,t)) \\
&+ \left(\mathcal{F} + \tfrac{1}{4}S\right)\mathcal{E}(r,t)\,\Sigma(y,t) \\
&= (\partial_t - \Delta_e)[\mathcal{E}(r,t)]\,\Sigma(y,t) - h^{-1/2}\partial_r\left[\sqrt{h}\right]\partial_r[\mathcal{E}(r,t)]\,\Sigma(y,t) \\
&- 2\partial_r(\mathcal{E}(r,t))\,\partial_r\Sigma(y,t) + \mathcal{E}(r,t)\,(\partial_t\Sigma(y,t) - \Delta\Sigma(y,t)) \\
&+ \left(\mathcal{F} + \tfrac{1}{4}S\right)\mathcal{E}(r,t)\,\Sigma(y,t) \\
&= -\partial_r[\mathcal{E}(r,t)]\left(h^{-1/2}\partial_r\left[\sqrt{h}\right]\Sigma(y,t) + 2\partial_r\Sigma(y,t)\right) \\
&+ \mathcal{E}(r,t)\left(\partial_t\Sigma(y,t) - \Delta\Sigma(y,t) + \left(\mathcal{F} + \tfrac{1}{4}S\right)\Sigma(y,t)\right).
\end{aligned}$$

Since

$$\begin{aligned}
\partial_r[\mathcal{E}(r,t)] &= (4\pi t)^{-n/2}\,\partial_r\left[\exp\left(-\tfrac{1}{4}r^2/t\right)\right] \\
&= (4\pi t)^{-n/2}\exp\left(-\tfrac{1}{4}r^2/t\right)\partial_r\left[-\tfrac{1}{4}r^2/t\right] = -\tfrac{r}{2t}\mathcal{E}(r,t),
\end{aligned}$$

we have

$$\begin{aligned}
&\mathcal{E}(r,t)^{-1}\left(\partial_t + \mathcal{D}^2\right)[\Psi_Q(y,t)] \\
&= \frac{r}{2t}h^{-1/2}\partial_r\left[\sqrt{h}\right]\Sigma(y,t) + 2\partial_r\Sigma(y,t) \\
&+\ \partial_t\Sigma(y,t) - \Delta\Sigma(y,t) + \left(\mathcal{F} + \tfrac{1}{4}S\right)\Sigma(y,t) \\
&= \sum_{k=0}^{Q}\begin{pmatrix} \frac{r}{2t}h^{-1/2}\partial_r\left[\sqrt{h}\right]U_k(y) + \frac{r}{t}\partial_r U_k(y) \\ +\frac{k}{t}U_k(y) - \Delta U_k(y) + \left(\mathcal{F} + \tfrac{1}{4}S\right)U_k(y) \end{pmatrix} t^k \\
&= \sum_{k=0}^{Q}\begin{pmatrix} \left(k + \tfrac{r}{2}h^{-1/2}\partial_r\left[\sqrt{h}\right]\right)U_k(y) + r\partial_r U_k(y) \\ +\left(-\Delta + \mathcal{F} + \tfrac{1}{4}S\right)[U_{k-1}(y)] \end{pmatrix} t^{k-1} \\
&+ \left(-\Delta + \mathcal{F} + \tfrac{1}{4}S\right)[U_Q(y)]\,t^Q,
\end{aligned}$$

where $U_{-1}(y) := 0$. Thus, to achieve (18.72), we need

$$r\partial_r U_k(y) + \left(k + \frac{r}{2}h^{-1/2}\partial_r\left[\sqrt{h}\right]\right)U_k(y) = \left(\Delta - \mathcal{F} - \tfrac{1}{4}S\right)[U_{k-1}(y)].$$

For fixed $0 \neq y_0 \in \mathcal{B}$, we define

$$u_k(s) := U_k(sy_0) \text{ for } s \in [0,1].$$

At sy_0, we have $r = s\,\|y_0\|$ and so $r\partial_r = s\,\|y_0\|\,(d/ds)\,\partial_r s = sd/ds$. Thus, we obtain the first-order, linear ODE

$$s\frac{du_k}{ds} + \left(k + \frac{s}{2}h^{-1/2}\frac{d}{ds}\left[\sqrt{h(sy_0)}\right]\right)u_k(s) = \left(\Delta - \mathcal{F} - \tfrac{1}{4}S\right)[u_{k-1}(sy_0)].$$

The integrating factor is

$$\exp \int \left(\frac{k}{s} + \tfrac{1}{2} h^{-1/2} \frac{d}{ds} \left[\sqrt{h(sy_0)} \right] \right) ds = s^k h(sy_0)^{1/4}.$$

Hence, for $k > 0$, we obtain

$$\begin{aligned} \frac{d}{ds} \left(s^k h(sy_0)^{1/4} u_k(s) \right) &= s^{k-1} h(sy_0)^{1/4} \left(\Delta - \mathcal{F} - \tfrac{1}{4} S \right) [u_{k-1}(sy_0)] \text{ and} \\ s^k h(sy_0)^{1/4} u_k(s) &= \int_0^s \widetilde{s}^{k-1} h(\widetilde{s} y_0)^{1/4} \left(\Delta - \mathcal{F} - \tfrac{1}{4} S \right) [u_{k-1}(\widetilde{s} y_0)] \, d\widetilde{s}, \end{aligned}$$

where we note that both sides are 0 for $s = 0$ when $k > 0$. Setting $s = 1$, replacing y_0 by y, we have $U_k(y) = u_k(1)$ and (for $k > 0$),

$$U_k(y) = h(y)^{-1/4} \int_0^1 s^{k-1} h(sy)^{1/4} \left(\Delta - \mathcal{F} - \tfrac{1}{4} S \right) [U_{k-1}](sy) \, ds.$$

When $k = 0$,

$$\begin{aligned} &\frac{d}{ds} \left(h(sy_0)^{1/4} u_0(s) \right) = 0 \Rightarrow u_0(s) = h(sy_0)^{-1/4} u_0(0) \\ &\Rightarrow U_0(y) = h(y)^{-1/4} U_0(0). \end{aligned}$$

Since $U_k(y) = h(y)^{-1/4} V_k(y)$, we have $V_0(y) = U_0(0)$ and

$$\begin{aligned} V_k(y) &= \int_0^1 s^{k-1} h(sy)^{1/4} \left(\Delta - \mathcal{F} - \tfrac{1}{4} S \right) \left[h^{-1/4} V_{k-1} \right](sy) \, ds \\ &= \int_0^1 s^{k-1} A_s [V_{k-1}](y) \, ds = \int_0^1 (s_{k-1})^{k-1} A_{s_{k-1}} [V_{k-1}](y) \, ds_{k-1} \\ &= \int_0^1 (s_{k-1})^{k-1} A_{s_{k-1}} \left[\int_0^1 (s_{k-2})^{k-2} A_{s_{k-2}} [V_{k-2}] \; ds_{k-2} \right](y) \, ds_{k-1} \\ &= \int_0^1 \int_0^1 (s_{k-1})^{k-1} (s_{k-2})^{k-2} \left(A_{s_{k-1}} \circ A_{s_{k-2}} \right) [V_{k-2}](y) \, ds_{k-2} ds_{k-1} \\ &= \int_{I^k} \left(\prod_{i=0}^{k-1} (s_i)^i \right) \left(A_{s_{k-1}} \circ A_{s_{k-2}} \circ \cdots \circ A_{s_0} [V_0] \right)(y) \, ds_0 \ldots ds_{k-1}, \end{aligned}$$

where we have used the fact that the A_s are linear differential operators and all functions are C^∞ in order to bring the A_{s_j} inside the the s_i-integrals for $i \neq j$. □

Note that $U_0(0) \in \mathbb{C}^N \otimes \Sigma_{2m}$ may be arbitrarily specified, and once $U_0(0)$ is chosen, the $U_m(y)$ are uniquely determined via (18.73). Let $h_k(y) \in \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)$ be given by

$$h_k(y)(U_0(0)) := U_k(y) \tag{18.74}$$

(in particular, $h_0(0) = \mathrm{Id} \in \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)$), and

$$H_Q(0, y, t) := \mathcal{E}(r, t) \sum_{k=0}^{Q} h_k(y) t^k \in C^\infty\left(\mathcal{B}, \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)\right). \tag{18.75}$$

We may regard $H_Q(0, y, t)$ as

$$H_Q(x, y, t) \in \mathrm{Hom}\left(E_x \otimes \Sigma(M)_x, E_y \otimes \Sigma(M)_y\right),$$

where $x \in M$ is the point about which we have chosen normal coordinates. For y sufficiently close to x, we set

$$H_Q(x,y,t) := \mathcal{E}(d(x,y),t) \sum_{k=0}^{Q} h_k(x,y)\, t^k. \tag{18.76}$$

To obtain a globally defined version of H_Q we proceed as follows. For $r > 0$, let

$$\delta_r(M \times M) := \{(x,y) \in M \times M : d(x,y) < r\}.$$

For sufficiently small r, say $0 < r < r_M$ = injectivity radius of M, the $H_Q(x,y,t)$ yield a section of the bundle $\mathcal{H}|_{\delta_r(M\times M)\times[0,\infty)}$, where (as in Section 18.3)

$$\begin{aligned} \mathcal{H} &:= \pi_1^*(E \otimes \Sigma(M))^* \otimes (\pi_2^*(E \otimes \Sigma(M))) \\ &\cong \mathcal{K} := \pi_1^*(E \otimes \Sigma(M)) \otimes \pi_2^*(E \otimes \Sigma(M)), \end{aligned} \tag{18.77}$$

and $\pi_i : M \times M \times [0,\infty) \to M$ $(i = 1,2)$ are the projections. For some constants r_1, r_2 with $0 < r_1 < r_2 < r_M$, there is a C^∞ function $\rho : [0,\infty) \to [0,1]$ with $\rho|_{[0,r_1]} = 1$ and $\rho|_{[r_2,\infty)} = 0$. Then we have $\varphi \in C^\infty(M \times M, \mathbb{R})$ defined by

$$\varphi(x,y) := \rho(d(x,y)).$$

with $\varphi|_{\delta_{r_1}(M\times M)} = 1$ and $\varphi = 0$ on the complement of $\delta_{r_2}(M \times M)$. Then φH_Q extends by zero values to a section $G_Q \in C^\infty(\mathcal{H})$, namely,

$$G_Q(x,y,t) := \begin{cases} \varphi(x,y)\, H_Q(x,y,t) & \text{for } (x,y) \in \delta_{r_2}(M \times M),\ t > 0 \\ 0 & \text{for } (x,y) \notin \delta_{r_2}(M \times M),\ t > 0. \end{cases} \tag{18.78}$$

From $G_Q(x,y,t)$ we will give a different construction of the heat kernel $k(x,y,t)$ and obtain its asymptotic expansion as $t \to 0^+$. However, before the proof of the validity of the construction, it is best to provide some motivation for it, as follows.

One expects that at least for small t, $G_Q(x,y,t)$ is a good approximation for $k(x,y,t)$, since the effect at small time t of a heat source at x should not be appreciably felt at a distant point y. Nevertheless, unlike the true heat kernel, we do not expect that $\left(\partial_t + \mathcal{D}_x^2\right) G_Q(x,y,t) = 0$ exactly. Let

$$K_{Q,0}(x,y,t) := \left(\partial_t + \mathcal{D}_x^2\right) G_Q(x,y,t).$$

In order to begin the correction process, we consider the problem

$$\left(\partial_t + \mathcal{D}_x^2\right) \psi(x,y,t) = -K_{Q,0}(x,y,t), \text{ with } \psi(x,y,0) = 0.$$

The solution of this problem should be given approximately by

$$\begin{aligned} \psi(x,y,t) &\approx -\int_0^t \int_M G_Q(x,z,s)\, K_{Q,0}(z,y,t-s)\ v(z)\ ds \\ &= -\int_0^t \int_M G_Q(x,z,t-s)\, K_{Q,0}(z,y,s)\ v(z)\ ds. \end{aligned} \tag{18.79}$$

To compactify the expressions to follow, it is conventient to introduce

DEFINITION 18.42. For $f_1,\ f_2 \in C^0(\mathcal{H})$, the **convolution** $f_1 * f_2 \in C^0(\mathcal{H})$ is defined by

$$(f_1 * f_2)(x,y,t) := \int_0^t \int_M f_1(x,z,s) \circ f_2(z,y,t-s)\ \nu_h(z)\ ds. \tag{18.80}$$

By a change of variable, we also have the alternate expression

$$
\begin{aligned}
(f_1 * f_2)(x,y,t) &= \int_0^t \int_M f_1(x,z,s) \circ f_2(z,y,t-s)\ \nu_h(z)\ ds \\
&= \int_t^0 \int_M f_1(x,z,t-\sigma) \circ f_2(z,y,\sigma)\ \nu_h(z)\ (-d\sigma) \\
&= \int_0^t \int_M f_1(x,z,t-s) \circ f_2(z,y,s)\ \nu_h(z)\ ds. \qquad (18.81)
\end{aligned}
$$

It is convenient to record here that when f_1 and f_2 are C^1,

$$
\begin{aligned}
&\partial_{x^i}((f_1 * f_2)(x,y,t)) = (\partial_{x^i} f_1 * f_2)(x,y,t) \\
&\partial_{y^i}((f_1 * f_2)(x,y,t)) = \left(f_1 * \partial_{y^i} f_2\right)(x,y,t), \text{ and} \qquad (18.82)
\end{aligned}
$$

Moreover using (18.80) and (18.81),

$$
\begin{aligned}
\partial_t((f_1 * f_2)(x,y,t)) &= \int_M f_1(x,z,t) \circ f_2(z,y,0)\ \nu_h(z) \\
&\quad + \int_0^t \int_M f_1(x,z,s) \circ \partial_t f_2(z,y,t-s)\ \nu_h(z)\ ds, \text{ and} \\
\partial_t((f_1 * f_2)(x,y,t)) &= \int_M f_1(x,z,0) \circ f_2(z,y,t)\ \nu_h(z) \\
&\quad + \int_0^t \int_M \partial_t f_1(x,z,t-s) \circ f_2(z,y,s)\ \nu_h(z)\ ds.
\end{aligned}
$$

Thus, when $f_1(\cdot,\cdot,0) \equiv 0$ and $f_2(\cdot,\cdot,0) \equiv 0$, we have

$$
\partial_t((f_1 * f_2)(x,y,t)) = (f_1 * \partial_t(f_2))(x,y,t) = (\partial_t(f_1) * f_2)(x,y,t). \qquad (18.83)
$$

In terms of convolution, (18.79) can be written as

$$
\psi(x,y,t) \approx -(G_Q * K_{Q,0})(x,y,t).
$$

At least formally, we have

$$
\begin{aligned}
&\left(\partial_t + \mathcal{D}_x^2\right)\psi(x,y,t) = -\left(\partial_t + \mathcal{D}_x^2\right)(G_Q * K_{Q,0})(x,y,t) \\
&= -\left(\partial_t + \mathcal{D}_x^2\right) \int_0^t \int_M G_Q(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&= -\partial_t \int_0^t \int_M G_Q(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&- \mathcal{D}_x^2 \int_0^t \int_M G_Q(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&= -\lim_{s \to t^-} \int_M G_Q(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&- \int_0^t \int_M \left(\partial_t + \mathcal{D}_x^2\right) G_Q(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&= -K_{Q,0}(x,y,t) - \int_0^t \int_M K_{Q,0}(x,z,t-s) \circ K_{Q,0}(z,y,s)\ v(z)\ ds \\
&= -K_{Q,0}(x,y,t) - (K_{Q,0} * K_{Q,0})(x,y,t).
\end{aligned}
$$

Then one expects that, to greater accuracy than $\left(\partial_t + \mathcal{D}_x^2\right) G_Q(x,y,t) \approx 0$, we have

$$\left(\partial_t + \mathcal{D}_x^2\right)\left(G_Q - G_Q * K_{Q,0}\right)(x,y,t) \approx 0.$$

We compute

$$\begin{aligned}
&\left(\partial_t + \mathcal{D}_x^2\right)\left(G_Q - G_Q * K_{Q,0}\right)(x,y,t)\\
&= K_{Q,0}(x,y,t) + \left(\partial_t + \mathcal{D}_x^2\right)(\psi(x,y,t))\\
&= K_{Q,0}(x,y,t) - K_{Q,0}(x,y,t) - \left(K_{Q,0} * K_{Q,0}\right)(x,y,t)\\
&= -\left(K_{Q,0} * K_{Q,0}\right)(x,y,t).
\end{aligned}$$

Then we consider the problem

$$\left(\partial_t + \mathcal{D}_x^2\right)\psi(x,y,t) = \left(K_{Q,0} * K_{Q,0}\right)(x,y,t),$$

whose solution should be given approximately by

$$\begin{aligned}
\psi(x,y,t) &\approx \left(G_Q - G_Q * K_{Q,0}\right) * \left(K_{Q,0} * K_{Q,0}\right)(x,y,t)\\
&= G_Q * \left(K_{Q,0} * K_{Q,0}\right)(x,y,t) - G_Q * K_{Q,0} * \left(K_{Q,0} * K_{Q,0}\right)(x,y,t).
\end{aligned}$$

Adding this correction to $G_Q - G_Q * K_{Q,0}$, presumably gives a more accurate approximation to the fundamental solution, namely

$$G_Q - G_Q * K_{Q,0} + G_Q * \left(K_{Q,0} * K_{Q,0}\right) - G_Q * K_{Q,0} * \left(K_{Q,0} * K_{Q,0}\right).$$

Repeating, we conjecture that the exact fundamental solution is

$$G_Q + G_Q * \sum_{k=1}^{\infty} (-1)^k K_{Q,0} * \overset{k \text{ factors}}{\cdots} * K_{Q,0}.$$

To prepare the rigorous demonstration of this, we first establish some boundedness properties of the operator whose kernel is G_Q.

PROPOSITION 18.43. *For $\psi \in C^l(E \otimes \Sigma(M))$ $(l \geq 0)$ and*

$$\left(\mathcal{G}_Q \psi\right)(y,t) := \int_M G_Q(x,y,t)\,\psi(x)\ \nu_h(x),$$

there is a constant c_l independent of t, such that

$$\left\|\left(\mathcal{G}_Q \psi\right)(\cdot,t)\right\|_{C^l} \leq c_l \left\|\psi\right\|_{C^l}. \tag{18.84}$$

Also,

$$\lim_{t \to 0^+} \left(\mathcal{G}_Q \psi\right)(y,t) = \psi(y). \tag{18.85}$$

Moreover for $K \in C^l(\mathcal{H})$ ($\mathcal{H}$ as in (18.77)), we have

$$\left\|G_Q * K\right\|_{C^l(T)} \leq c_{l,T} \left\|K\right\|_{C^l(T)}, \tag{18.86}$$

where $C^l(T)$ is the C^l norm for restrictions of sections in $C^l(\mathcal{H})$ to $M \times M \times [0,T]$, where T may be chosen arbitrarily large, and $c_{l,T}$ is a constant depending on l and T.

PROOF. Making the change of variable $w = (x-y)/\sqrt{t},\ x = y + w\sqrt{t}$,

$$\begin{aligned}
(\mathcal{G}_Q\psi)(y,t) &= \int_M \varphi(x,y)\, H_Q(x,y,t)\, \psi(x)\ \nu_h(x) \\
&= \int_M \varphi(x,y)\, \mathcal{E}(r,t) \sum\nolimits_{k=0}^{Q} h_k(x,y)\, \psi(x)\, t^k\, \nu_h(x) \\
&= \int_M \varphi(x,y)\, (4\pi t)^{-n/2}\, e^{-\frac{1}{4}|y-x|^2/t} \sum\nolimits_{k=0}^{Q} h_k(x,y)\, \psi(x)\, t^k\, \nu_h(x) \\
&= (4\pi)^{-n/2} \sum\nolimits_{k=0}^{Q} t^k \int_{\mathbb{R}^n} \begin{pmatrix} \rho\left(|w|\sqrt{t}\right) e^{-\frac{1}{4}|w|^2}\cdot \\ \cdot h_k\left(y + w\sqrt{t}, y\right) \psi\left(y + w\sqrt{t}\right) \end{pmatrix} \nu_h(w).
\end{aligned}$$

Thus, as the h_k are C^∞, for any fixed $t > 0$,

$$\|(\mathcal{G}_Q\psi)(\cdot,t)\|_{C^l} \le c_l \|\psi\|_{C^l}.$$

for some constant c_l, and we have (18.84). For (18.85), note that each integrand in the above final expression for $(\mathcal{G}_Q\psi)(y,t)$ is bounded by an integrable function, namely a constant multiple of $e^{-\frac{1}{4}|w|^2}$. Thus, we may apply the Lebesgue Dominated Convergence Theorem to obtain (18.85):

$$\begin{aligned}
&\lim_{t\to 0^+} (\mathcal{G}_Q\psi)(y,t) \\
&= (4\pi)^{-n/2} \sum_{k=0}^{Q} \lim_{t\to 0^+} t^k \int_{\mathbb{R}^n} \begin{pmatrix} \rho\left(|w|\sqrt{t}\right) e^{-\frac{1}{4}|w|^2}\cdot \\ \cdot h_k\left(y + w\sqrt{t}, y\right) \psi\left(y + w\sqrt{t}\right) \end{pmatrix} \nu_h(w) \\
&= (4\pi)^{-n/2} \int_{\mathbb{R}^n} \lim_{t\to 0^+} \rho\left(|w|\sqrt{t}\right) e^{-\frac{1}{4}|w|^2} h_0\left(y + w\sqrt{t}, y\right) \psi\left(y + w\sqrt{t}\right) \nu_h(w) \\
&= (4\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-\frac{1}{4}|w|^2} h_0(y,y)\, \psi(y)\ \nu_e(w) = h_0(y,y)\, \psi(y) = \psi(y).
\end{aligned}$$

To see (18.86), we make the change of variables $w = (z-x)/\sqrt{s},\ z = x + w\sqrt{s}$ in computing

$$\begin{aligned}
(G_Q * K)(x,y,t) &= \int_0^t \int_{\mathbb{R}^n} \varphi(x,z)\, H_Q(x,z,s) \circ K(z,y,t-s)\ \nu_h(z)\ ds \\
&= \int_0^t \int_{\mathbb{R}^n} \varphi(x,z)\, (4\pi s)^{-m}\, e^{-\frac{1}{4}|z-x|^2/s} \sum\nolimits_{k=0}^{Q} h_k(x,z) \circ K(z,y,t-s)\, s^k\, \nu_h(z)\ ds
\end{aligned}$$

(18.87)
$$= \int_0^t \int_{\mathbb{R}^n} \begin{pmatrix} \varphi\left(x, x + w\sqrt{s}\right) (4\pi)^{-m}\, e^{-\frac{1}{4}w^2}\cdot \\ \cdot \sum_{k=0}^{Q} h_k\left(x, x + w\sqrt{s}\right) \circ K\left(x + w\sqrt{s}, y, t-s\right) s^k \end{pmatrix} \nu_h(w)\ ds.$$

Moreover, we have

$$\begin{aligned}
&\tfrac{\partial}{\partial t}(G_Q * K)(x,y,t) \\
&= \int_{\mathbb{R}^n} \begin{pmatrix} \varphi\left(x, x + w\sqrt{s}\right) (4\pi)^{-m} \exp\left(-\frac{1}{4}w^2\right)\cdot \\ \cdot \sum_{k=0}^{Q} h_k\left(x, x + w\sqrt{s}\right) \circ K\left(x + w\sqrt{s}, y, t-s\right) s^k \end{pmatrix} \nu_h(w) \\
&+ \int_0^t \int_{\mathbb{R}^n} \begin{pmatrix} \varphi\left(x, x + w\sqrt{s}\right) (4\pi)^{-m} \exp\left(-\frac{1}{4}w^2\right)\cdot \\ \cdot \sum_{k=0}^{Q} h_k\left(x, x + w\sqrt{s}\right) \circ \frac{\partial}{\partial t} K\left(x + w\sqrt{s}, y, t-s\right) s^k \end{pmatrix}\cdot
\end{aligned}$$

(18.88) $\quad \cdot\ \nu_h(w)\ ds.$

Using the fact that the h_k are fixed C^∞ sections, the forms of (18.87) and (18.88) imply that the derivatives of $G_Q * K$ of order $\leq l$ on $M \times M \times [0,T]$ have bounds (depending on T) in terms of those of K, so that $\|(G_Q * K)\|_{C^l(T)} \leq c_{l,T} \|K\|_{C^l(T)}$. □

We set

$$K_{Q,0}(x,y,t) := \left(\partial_t + \mathcal{D}_y^2\right) G_Q(x,y,t) \text{ and}$$
$$K_{Q,j}(x,y,t) := \int_0^t \int_M K_{Q,0}(x,z,s) \circ K_{Q,j-1}(z,y,t-s)\, \nu_h(z)\, ds$$
$$= \left(K_{Q,0} * K_{Q,j-1}\right)(x,y,t) \text{ for } j \geq 1.$$

The next result implies that for Q sufficiently large,

$$\kappa_Q := G_Q + \sum_{j=0}^{\infty} (-1)^{j+1} G_Q * K_{Q,j}; \text{ i.e.} \tag{18.89}$$

exists and is a general heat kernel for $\mathcal{D}^2$; in particular, κ_Q is independent of Q.

THEOREM 18.44. *For any integer $k \geq 2$, if we choose $Q > m + 2k$ and $T > t_0 > 0$, the series (18.89) defining $\kappa_Q(x,y,t)$ converges in $C^k\left(\mathcal{H}|_{M\times M\times[t_0,T]}\right)$, $\kappa_Q(x,y,t)$ satisfies $\left(\partial_t + \mathcal{D}_x^2\right)\kappa_Q = 0$, and*

$$\lim_{t\to 0^+} \int_M \kappa_Q(y,x,t)\, \psi(y)\, v_h(y) = \psi(x), \tag{18.90}$$

for all $\psi \in C^0\left(E \otimes \Sigma^{\pm}(M)\right)$ (i.e., κ_Q is a general heat kernel in the sense of Definition 18.30, p. 516). Moreover,

$$|\kappa_Q(x,y,t) - G_Q(x,y,t)| \leq C t^{Q-m+1}, \tag{18.91}$$

for some constant C independent of $(x,y,t) \in M \times M \times (0,T)$.

PROOF. Once we prove that for $k \geq 2$ and $Q > m + 2k + 1$ the series

$$K_Q(x,y,t) := \sum_{j=0}^{\infty} (-1)^{j+1} K_{Q,j}(x,y,t)$$

converges in $C^k\left(\mathcal{H}|_{M\times M\times[0,T]}\right)$, then according to Proposition 18.43,

$$G_Q + \sum_{j=0}^{\infty} (-1)^{j+1} G_Q * K_{Q,j} = G_Q + G_Q * \sum_{k=0}^{\infty} (-1)^{j+1} K_{Q,j}$$
$$= G_Q + G_Q * K_Q,$$

where the convergence of the first infinite sum is in $C^k\left(\mathcal{H}|_{M\times M\times[0,T]}\right)$. Thus, noting that although G_Q is not C^k at $t = 0$,

$$\kappa_Q := G_Q + \sum_{j=0}^{\infty} (-1)^{j+1} G_Q * K_{Q,j} \in C^k\left(\mathcal{H}|_{M\times M\times(0,\infty)}\right)$$

will exist and the convergence will be in $C^k\left(\mathcal{H}|_{M\times M\times[t_0,T]}\right)$ if $0<t_0<T<\infty$. Then for $k\geq 2$, $\left(\partial_t+\mathcal{D}_x^2\right)\kappa_Q$ can be computed via term-by-term differentiation:

$$\begin{aligned}
&\left(\partial_t+\mathcal{D}_x^2\right)\kappa_Q(x,y,t)\\
&=\left(\partial_t+\mathcal{D}_x^2\right)G_Q(x,y,t)\\
&+\sum_{j=0}^{\infty}(-1)^{j+1}\left(\partial_t+\mathcal{D}_x^2\right)\int_0^t\int_M G_Q(x,z,s)\circ K_{Q,j}(z,y,t-s)\,\nu_h(z)\,ds\\
&=K_{Q,0}(x,y,t)+\sum_{j=0}^{\infty}(-1)^{j+1}\lim_{s\to t^-}\int_M G_Q(x,z,t-s)\circ K_{Q,j}(z,y,s)\,\nu_h(z)\\
&+\sum_{j=1}^{\infty}(-1)^{j+1}\int_0^t\int_M K_{Q,0}(x,z,s)\circ K_{Q,j}(z,y,t-s)\,\nu_h(z)\,ds\\
&=K_{Q,0}(x,y,t)+\sum_{j=0}^{\infty}(-1)^{j+1}K_{Q,j}(x,y,t)+\sum_{k=0}^{\infty}(-1)^{j+1}K_{Q,j+1}(x,y,t)=0.
\end{aligned}$$

To prove that the series for $K_Q(x,y,t)$ converges in the $C^k\left(\mathcal{H}|_{M\times M\times[t_0,T]}\right)$, we will estimate the terms of this sum and their derivatives so that the Weierstrass M-test can be applied. Recall that

$$K_{Q,0}(x,y,t):=\left(\partial_t+\mathcal{D}_x^2\right)G_Q(x,y,t),\ \text{where}$$

$$G_Q(x,y,t):=\begin{cases}\varphi(x,y)\,H_Q(x,y,t) & \text{for } (x,y)\in\delta_{r_2}(M\times M),\ t>0\\ 0 & \text{for } (x,y)\notin\delta_{r_2}(M\times M),\ t>0,\end{cases}$$

and $\varphi(x,y):=\rho(d(x,y))$. Note that $G_Q(x,y,t)$ and all of its local derivatives vanish for $d(x,y)\geq r_2$, and so we assume that $d(x,y)<r_2$ in what follows. For $d(x,y)<r_1$, we have $\varphi(x,y)=\rho(d(x,y))=1$, and

$$\begin{aligned}
K_{Q,0}(x,y,t)&=\left(\partial_t+\mathcal{D}_x^2\right)G_Q(x,y,t)=\left(\partial_t+\mathcal{D}_x^2\right)(\varphi(x,y)\,H_Q(x,y,t))\\
&=\left(\partial_t+\mathcal{D}_x^2\right)H_Q(x,y,t)=\mathcal{E}(d(x,y),t)\,t^Q\mathcal{D}^2(U_Q)(x,y).
\end{aligned}$$

For $r:=d(x,y)\in[r_1,r_2)$,

$$\begin{aligned}
K_{Q,0}(x,y,t)&=\left(\partial_t+\mathcal{D}_x^2\right)G_Q(x,y,t)=\left(\partial_t+\mathcal{D}_x^2\right)(\varphi(x,y)\,H_Q(x,y,t))\\
&=\left(\partial_t+\mathcal{D}_x^2\right)H_Q(x,y,t)+\left(\partial_t+\mathcal{D}_x^2\right)((\varphi(x,y)-1)\,H_Q(x,y,t))\\
&=\mathcal{E}(r,t)\,t^Q\mathcal{D}^2(U_Q)(y)+\left(\partial_t+\mathcal{D}_x^2\right)((\varphi(x,y)-1)\,H_Q(x,y,t)),
\end{aligned}$$

and because of the factor $t^{-m}e^{-d(x,y)^2/4t}$ in $H_Q(x,y,t)$,

$$|K_{Q,0}(x,y,t)|\leq\mathcal{E}(r,t)\,t^Q\left|\mathcal{D}^2(U_Q)(y)\right|+C_0t^{-m-2}e^{-r_1^2/t}$$

for some (x,y,t)-independent constant C_0. Since $e^{-r_1^2/t}$ is $\mathrm{O}(t^k)$ for all $k>0$, putting this together with the result for $d(x,y)<r_1$, we have

$$|K_{Q,0}(x,y,t)|\leq C_Qe^{-r^2/4t}t^{Q-m}\leq C_Qt^{Q-m},$$

for all $(x,y,t)\in M\times M\times[0,\infty)$, for some (x,y,t)-independent constant C_Q. Using the same reasoning, and noting that differentiation of $e^{-d(x,y)^2/4t}$ with respect to

local coordinates x^i or y^i introduces factors of t^{-1}, while applying ∂_t introduces factors of t^{-2}, it is straightforward to see that
(18.92)

$$\left|(\partial_t)^s \partial_{x^{i_1}} \cdots \partial_{x^{i_p}} \partial_{y^{j_1}} \cdots \partial_{y^{j_q}} K_{Q,0}(x,y,t)\right| \leq C_Q(p+q+s)\, t^{Q-m-p-q-2s},$$

for some constant $C_Q(N)$ depending monotonically on $N = 0,1,2,\ldots$, say with $C_Q(0) = C_Q$. We will use this below. Now,

$$\begin{aligned}
|K_{Q,1}(x,y,t)| &= \left|\int_0^t \int_M K_{Q,0}(x,z,s) \circ K_{Q,0}(z,y,t-s)\ \nu_h(z)\ ds\right| \\
&\leq \int_0^t \int_M |K_{Q,0}(x,z,s) \circ K_{Q,0}(z,y,t-s)|\ \nu_h(z)\ ds \\
&\leq \int_0^t \int_M |K_{Q,0}(x,z,s)|\,|K_{Q,0}(z,y,t-s)|\ \nu_h(z)\ ds \\
&\leq \int_0^t \int_M {C_Q}^2 s^{Q-m}(t-s)^{Q-m}\ \nu_h(z)\ ds \\
&= V(M)\,{C_Q}^2 \int_0^t s^{Q-m}(t-s)^{Q-m}\ ds \\
&= V(M)\,{C_Q}^2 t^{2Q-2m+1} \frac{\Gamma(Q-m+1)^2}{\Gamma(2Q-2m+2)}.
\end{aligned}$$

For the last equality, recall that for $a, b \geq 0$,

$$\int_0^1 x^a (1-x)^b\, dx = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)}.$$

Thus,

$$\begin{aligned}
&\int_0^t s^a (t-s)^b\ ds = \int_0^t \left(\left(\frac{s}{t}\right)^a t^a\right) t^b \left(1-\frac{s}{t}\right)^b\ ds \\
&= t^{a+b} \int_0^t \left(\frac{s}{t}\right)^a \left(1-\frac{s}{t}\right)^b\ ds = \left(\text{with } x = \frac{s}{t},\ tdx = ds\right) \\
&= t^{a+b} \int_0^1 x^a (1-x)^b\ tdx = t^{a+b+1} \int_0^1 x^a (1-x)^b\ dx \\
&= t^{a+b+1} \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)}.
\end{aligned}$$

More generally, if $f_1,\ f_2 \in C^0(\mathcal{H})$ with

$$|f_1(x,y,t)| \leq C_1 t^{p_1} \text{ and } |f_2(x,y,t)| \leq C_2 t^{p_2} \text{ (for } p_1, p_2 \geq 0),$$

$$\begin{aligned}
|(f_1 * f_2)(x,y,t)| &= \left|\int_0^t \int_M f_1(x,z,s) \circ f_2(z,y,t-s)\ \nu_h(z)\ ds\right| \\
&\leq C_1 C_2 V(M) \int_0^t s^{p_1}(t-s)^{p_2}\ ds \\
&= C_1 C_2 \frac{\Gamma(p_1+1)\Gamma(p_2+1)}{\Gamma(p_1+p_2+2)} V(M)\, t^{p_1+p_2+1}.
\end{aligned}$$

Applying this, we get the following fact that will also be needed,

$$\begin{aligned}
|(f_1 * f_2 * f_3)(x,y,t)| &\le C_1\left(C_2C_3\frac{\Gamma(p_2+1)\Gamma(p_3+1)}{\Gamma(p_2+p_3+2)}V(M)\right)\cdot\\
&\cdot\frac{\Gamma(p_1+1)\Gamma(p_2+p_3+2)}{\Gamma(p_1+p_2+p_3+3)}V(M)\,t^{p_1+p_2+p_3+2}\\
&= C_1C_2C_3V(M)^2\frac{\Gamma(p_1+1)\Gamma(p_2+1)\Gamma(p_3+1)}{\Gamma(p_1+p_2+p_3+3)}t^{p_1+p_2+p_3+2}. \qquad (18.93)
\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|\left(*^3 f_1\right)(x,y,t)\right| &:= |(f_1 * f_1 * f_1)(x,y,t)|\\
&\le C_1^3 V(M)^2\frac{\Gamma(p_1+1)^3}{\Gamma(3(p_1+1))}t^{3p_1+2},
\end{aligned}$$

and by induction, we see that

$$\begin{aligned}
\left|\left(*^k f_1\right)(x,y,t)\right| &:= \left|\left(f_1 * \overset{k\text{ factors}}{\cdots} * f_1\right)(x,y,t)\right|\\
&\le C_1^k V(M)^{k-1}\frac{\Gamma(p_1+1)^k}{\Gamma((p_1+1)k)}t^{(p_1+1)k-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
|K_{Q,j}(x,y,t)| &= \left|*^{j+1}K_{Q,0}(x,y,t)\right|\\
&\le C_Q^{j+1}V(M)^j\frac{\Gamma(Q-m+1)^{j+1}}{\Gamma((Q-m+1)(j+1))}t^{(Q-m+1)(j+1)-1}.
\end{aligned}$$

Since we have assumed that $Q > m+2k$ and $k \ge 2$, all the powers $(Q-m+1)(j+1)-1$ of t are positive. Hence, the M-test will apply to give uniform convergence of the series $\sum_{j=0}^{\infty}(-1)^{j+1}K_{Q,j}(x,y,t)$ on $M \times M \times [0,T]$ if for every constant $R \ge 0$

$$\sum_{j=0}^{\infty}\frac{R^j}{\Gamma((Q-m+1)(j+1))} < \infty,$$

but as $Q > m$, this holds by comparison with $\sum_{j=0}^{\infty}\frac{R^j}{\Gamma(j+1)} = e^R$.

We now show the $C^k\left(\mathcal{H}|_{M\times M\times[0,T]}\right)$-convergence of

$$K_Q(x,y,t) = \sum_{j=0}^{\infty}(-1)^{j+1}K_{Q,j}(x,y,t),$$

for $Q > m+2k$. Through repeated use of (18.82) and (18.83), we have for $j \ge 3$

$$\begin{aligned}
&(\partial_t)^s\partial_{x^{i_1}}\cdots\partial_{x^{i_p}}\partial_{y^{j_1}}\cdots\partial_{y^{j_q}}K_{Q,j}(x,y,t)\\
&= \left(\left((\partial_t)^s\partial_{x^{i_1}}\cdots\partial_{x^{i_p}}K_{Q,0}\right)*\left(K_{Q,j-3}\right)*\left(\partial_{y^{j_1}}\cdots\partial_{y^{j_q}}K_{Q,0}\right)\right)(x,y,t).
\end{aligned}$$

For $N = p+q+s \le k$, we have

$$\begin{aligned}
&|(\partial_t)^s\partial_{x^{i_1}}\cdots\partial_{x^{i_p}}K_{Q,0}| \le C_Q(N)\,t^{Q-m-p-2s}\\
&|K_{Q,j-3}| \le C_Q^{j-2}(N)\,V(M)^{j-3}\cdot\\
&\cdot\frac{\Gamma(Q-m+1)^{j-2}}{\Gamma((Q-m+1)(j-2))}t^{(Q-m+1)(j-2)-1},\ \text{and}\\
&\left|\partial_{y^{j_1}}\cdots\partial_{y^{j_q}}K_{Q,0}\right| \le C_Q(N)\,t^{Q-m-q}, \qquad (18.94)
\end{aligned}$$

where all the powers of t are positive, since $Q > m + 2k \geq m + 2(p + q + s)$.

$$C_1 C_2 C_3 V(M)^2 \frac{\Gamma(p_1+1)\Gamma(p_2+1)\Gamma(p_3+1)}{\Gamma(p_1+p_2+p_3+3)} t^{p_1+p_2+p_3+2}$$

Using (18.82), (18.83), (18.93), and (18.94), we have

$$\begin{aligned}
&\left|(\partial_t)^s \partial_{x^{i_1}} \cdots \partial_{x^{i_p}} \partial_{y^{j_1}} \cdots \partial_{y^{j_q}} K_{Q,j}(x,y,t)\right| \\
&= \left|\left(\left((\partial_t)^s \partial_{x^{i_1}} \cdots \partial_{x^{i_p}} K_{Q,0}\right) * (K_{Q,j-3}) * \left(\partial_{y^{j_1}} \cdots \partial_{y^{j_q}} K_{Q,0}\right)\right)(x,y,t)\right| \\
&\leq C_Q (N)^j V(M)^{j-1} \frac{\Gamma(Q-m+1)^{j-2}}{\Gamma((Q-m+1)(j-2))} \cdot t^{(Q-m+1)j-(q+2s+p)-1} \cdot \\
&\frac{\Gamma(Q-m-p-2s+1)\Gamma((Q-m+1)(j-2))\Gamma(Q-m-q+1)}{\Gamma(Q-m-p-2s+(Q-m+1)(j-2)-1+Q-m-q+3)} \\
&\leq C_Q (N)^j V(M)^{j-1} \cdot t^{(Q-m+1)j-(q+2s+p)-1} \cdot \\
&\cdot \frac{\Gamma(Q-m-p-2s+1)\Gamma(Q-m+1)^{j-2}\Gamma(Q-m-q+1)}{\Gamma((Q-m+1)j-(p+2s+q))}
\end{aligned}$$

Since $\Gamma((Q-m+1)j-(p+2s+q)) \geq Cj!$ for some constant C, the series of these terms can be estimated from above, using $\sum R^j/j! < \infty$ as in the case $k = 0$. Thus, the M-test can be used to obtain the uniform convergence (for $Q > m + 2k$) on $M \times M \times [0,T]$ of

$$\sum_{j=0}^{\infty} (-1)^{j+1} (\partial_t)^s \partial_{x^{i_1}} \cdots \partial_{x^{i_p}} \partial_{y^{j_1}} \cdots \partial_{y^{j_q}} K_{Q,j}(x,y,t),$$

and hence the $C^k\left(\mathcal{H}|_{M\times M\times[0,T]}\right)$-convergence of $K_Q = \sum_{j=0}^{\infty} (-1)^{j+1} K_{Q,j}(x,y,t)$.

We now prove (18.91). From $\kappa_Q = G_Q + \sum_{j=0}^{\infty} (-1)^{j+1} G_Q * K_{Q,j}$, we have

$$|\kappa_Q(x,y,t) - G_Q(x,y,t)| = |(G_Q * K_Q)(x,y,t)|.$$

Moreover for some constant C,

$$\begin{aligned}
|K_Q(x,y,t)| &\leq \sum\nolimits_{j=0}^{\infty} |K_{Q,j}(x,y,t)| \\
&\leq \sum\nolimits_{j=0}^{\infty} C_Q^{j+1} V(M)^j \frac{\Gamma(Q-m+1)^{j+1}}{\Gamma((Q-m+1)(j+1))} t^{(Q-m+1)(j+1)-1} \\
&= t^{Q-m} \sum\nolimits_{j=0}^{\infty} C_Q^{j+1} V(M)^j \frac{\Gamma(Q-m+1)^{j+1}}{\Gamma((Q-m+1)(j+1))} t^{(Q-m+1)j} \\
&\leq Ct^{Q-m}. \qquad (18.95)
\end{aligned}$$

With $K = K_Q$ in (18.87), we obtain

$$\begin{aligned}
&(G_Q * K_Q)(x,y,t) = \\
&= \int_0^t \int_M \begin{pmatrix} \varphi(x, x+w\sqrt{s})(4\pi)^{-m} e^{-\frac{1}{4}w^2} \cdot \\ \cdot \sum_{k=0}^{Q} h_k(x, x+w\sqrt{s}) \circ K_Q(x+w\sqrt{s}, y, t-s) s^k \end{pmatrix} \nu_h(w)\, ds.
\end{aligned}$$

Then using (18.95), for some constants C' and C'', we have

$$\begin{aligned}&\left|\kappa_Q(x,y,t)-G_Q(x,y,t)\right| = \left|\left(G_Q * K_Q\right)(x,y,t)\right| \\ &\le C'\int_0^t \sum_{k=0}^{Q}(t-s)^{Q-m}s^k\,ds = \sum_{k=0}^{Q} C' t^{Q-m+k+1} \le C'' t^{Q-m+1} \text{ as } t\to 0^+,\end{aligned}$$

which yields (18.91). Finally, note that

$$\begin{aligned}&\left|\int_M \kappa_Q(y,x,t)\,\psi(y)\ v_y - \psi(x)\right| \\ &\le \left|\int_M \left(\kappa_Q(y,x,t)-G_Q(y,x,t)\right)\psi(y)\ v_y\right| + \left|\int_M G_Q(y,x,t)\,\psi(y)\ v_y - \psi(x)\right| \\ &\le Ct^{Q-m+1}\int_M |\psi(y)|\ v_y + \left|\int_M G_Q(y,x,t)\,\psi(y)\ v_y - \psi(x)\right|.\end{aligned}$$

Thus, (18.90) follows from (18.85) in Proposition 18.43, p. 536. □

5. The Local Index Formula

In the previous section, we showed that for $x, y \in M$ (of even dimension $n = 2m$) with $r = d(x,y) :=$ Riemannian distance from x to y sufficiently small, the heat kernel $k(x,y,t)$ of $\mathcal{D}^2$ has an asymptotic expansion as $t \to 0^+$ of the form

$$k(x,y,t) \sim H_Q(x,y,t) := \mathcal{E}(r,t)\sum\nolimits_{j=0}^{Q} h_j(x,y)\,t^j,$$

for any fixed integer $Q > m+4$, where

$$h_j(x,y) \in \operatorname{Hom}\left((E\otimes\Sigma(M))_x, (E\otimes\Sigma(M))_y\right),\ j\in\{0,1,\ldots,Q\}.$$

In particular, with $y = x$, we have

$$k(x,x,t) \sim (4\pi t)^{-m}\sum_{j=0}^{Q} h_j(x,x)\,t^j = (4\pi)^{-m}\sum_{j=0}^{Q} h_j(x,x)\,t^{j-m}. \tag{18.96}$$

From (18.38) we know that for any $t > 0$,

$$\int_M \operatorname{Str}(k(x,x,t))\ \nu_h(x) = \operatorname{index}\left(\mathcal{D}^+\right),$$

which is constant, independent of t. As $t \to 0^+$, we deduce from (18.96) that

$$\int_M \operatorname{Str}(h_j(x,x))\ \nu_h(x) = 0 \ \text{ for }\ j\in\{0,1,\ldots,m-1\},\ \text{while}$$

$$(4\pi)^{-m}\int_M \operatorname{Str}(h_m(x,x))\ \nu_h(x) = \operatorname{index}\left(\mathcal{D}^+\right). \tag{18.97}$$

We have a somewhat cumbersome formula for $h_m(x,x)$, and it is already clear from the formula that $h_m(x,x)$ is determined by the metric h and the connection ε on $U(E)$ in a neighborhood of the point x (indeed by finitely many derivatives of h and ε at x). One might regard the gist of the Index Formula for twisted Dirac operator as exhibiting the global quantity $\operatorname{index}(\mathcal{D}^+)$ as the integral of form which may be locally computed. From this perspective, (18.97) does the job. The direct computation of $h_m(x,x)$ is actually not very difficult for $m = 1$ and 2 (i.e., for $\dim(M) = 2$ or 4), and we will carry it out explicitly. However, for large values of m it is rather cumbersome and one would like a more tractable formula for $h_m(x,x)$. It is also desireable to express $h_m(x,x)$ in terms of curvature forms, thereby showing that $h_m(x,x)$ only depends on the 2-jet of the metric and the 1-jet of ε. Moreover, since $\operatorname{index}(\mathcal{D}^+)$ is insensitive to pertubations in h and ε, one expects that $h_m(x,x)\,\nu_h$ can be expressed in terms of curvature forms of characteristic classes for M and E. The Local Index Formula below does this. Moreover, since the Local Index Formula is a purely local result, it may be applied to obtain the (global) Index Formula for elliptic operators which are only locally of the form of twisted Dirac operator $\mathcal{D}^+$. Indeed, if $\mathcal{A}$ is such an operator (possibly on a nonspin manifold) and k is the heat kernel for $\mathcal{A}^*\mathcal{A}\oplus\mathcal{A}\mathcal{A}^*$, then from the spectral resolution of $\mathcal{A}$, we still have

$$\int_M \operatorname{Str}\left(h_m(x,x)\right)\,\nu_h(x) = \operatorname{index}(\mathcal{A}),$$

where the supertrace Str is defined in the natural way. Now, the crucial observation is that the Local Index Formula allows us to compute $\operatorname{Str}(h_m(x,x))$ once $\mathcal{A}$ is represented *locally* as a twisted Dirac operator. We will see a number of examples of such $\mathcal{A}$ in the next section. In fact, it is not easy to find any first-order elliptic operators of geometrical significance which are not locally twisted Dirac operators, or 0-th order perturbations thereof. Our goal in this section is to prove

THEOREM 18.45 (The Local Index Formula). *For* $\dim M = 2m$, *let* $h_m(x,x)$ *be as in the diagonal asymptotic expansion (18.96) of the heat kernel of an operator which in some neighborhood* B *about* $x\in M$ *is of the form* $\mathcal{D}^2$ *for a twisted Dirac operator* $\mathcal{D}: C^\infty(B, E(B)\otimes\Sigma(B)) \hookleftarrow$. *Here* $\mathcal{D}$ *is determined by a connection* ε *on* $E(B)$, *a metric* h *on* B *with Levi-Civita connection* θ, *and a spin structure over* FB. *If* Ω^ε *is the curvature of* ε *and* Ω^θ *is the curvature of* θ, *then*
(18.98)

$$(4\pi)^{-m}\operatorname{Str}\left(h_m(x,x)\right) = \left\langle \operatorname{Tr}\left(e^{i\Omega^\varepsilon/2\pi}\right)\wedge\det\left(\frac{i\Omega^\theta/4\pi}{\sinh\left(i\Omega^\theta/4\pi\right)}\right)^{\frac{1}{2}}, \nu_h(x)\right\rangle,$$

where the meaning of the right side was explained in the paragraph following (18.39), p. 513.

Using normal coordinates $\left(y^1,\ldots,y^{2m}\right)\in B(r_0,0)$ about $x\in M$ and the radial gauge, and selecting $V_0\in\mathbb{C}^N\otimes\Sigma_{2m}$, we have (see (18.76), (18.74) and Proposition 18.41, p. 531),
(18.99)

$$h_m(x,x)(V_0) = \int_{I^m}\prod\nolimits_{i=0}^{m-1}(s_i)^i\left(\left(A_{s_{m-1}}\circ\cdots\circ A_{s_0}\right)\left[\widetilde{V}_0\right]\right)(0)\,ds_0\ldots ds_{m-1},$$

where $\widetilde{V}_0\in C^\infty\left(B(r_0,0),\mathbb{C}^N\otimes\Sigma_{2m}\right)$ is the constant extension of V_0. Recall that for $\psi\in C^\infty\left(B(r_0,0),\mathbb{C}^N\otimes\Sigma_{2m}\right)$, we have

$$A_s[\psi](y) := A[\psi](sy),\ \text{where}$$
$$A[\psi] := h^{1/4}\Delta\left[h^{-1/4}\psi\right] - \left(\mathcal{F}+\tfrac{1}{4}S\right)[\psi].$$

While the right side of (18.99) may seem unwieldy, there is substantial simplification due to facts that $\left(A_{s_{m-1}}\circ\cdots\circ A_{s_0}\right)\left[\widetilde{V}_0\right](y)$ is evaluated at $y=0$ in (18.99) and that only those terms of $A_{s_{m-1}} \circ \cdots \circ A_{s_0}\left[\widetilde{V}_0\right](0)$ which involve $\gamma_{n+1} := \gamma_1 \cdots \gamma_{2m}$ will survive when the supertrace $\operatorname{Str}\left(h_m(x,x)\right)$ is taken (see Proposition 18.12, p. 491). To see how we may take advantage of these facts, we need to expand

$$
\begin{aligned}
A[\psi] &= -h^{1/4}\mathcal{D}^2\left[h^{-1/4}\psi\right] = h^{1/4}\Delta\left[h^{-1/4}\psi\right] - \left(\mathcal{F} + \tfrac{1}{4}S\right)[\psi] \\
&= h^{-1/4}\nabla_{\partial_j}\left(h^{ij}\sqrt{h}\nabla_{\partial_i}\left(h^{-1/4}\psi\right)\right) - \left(\mathcal{F} + \tfrac{1}{4}S\right)[\psi] \\
&= h^{-1/4}\begin{pmatrix} \partial_j\left(h^{ij}\sqrt{h}\nabla_{\partial_i}\left(h^{-1/4}\psi\right)\right) \\ +h^{ij}\sqrt{h}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\nabla_{\partial_i}\left(h^{-1/4}\psi\right)\end{pmatrix} - \left(\mathcal{F} + \tfrac{1}{4}S\right)[\psi] \\
&= h^{-1/4}\begin{pmatrix} \partial_j\left(h^{ij}\sqrt{h}\left(\partial_i\left(h^{-1/4}\psi\right) + \left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)h^{-1/4}\psi\right)\right) \\ +h^{ij}\sqrt{h}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\partial_i\left(h^{-1/4}\psi\right) + \left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)h^{-1/4}\psi\right)\end{pmatrix} \\
&\quad - \left(\mathcal{F} + \tfrac{1}{4}S\right)[\psi] \\
&= h^{-1/4}\partial_j\left(h^{ij}\sqrt{h}\partial_i\left(h^{-1/4}\psi\right)\right) + h^{-1/4}\partial_j\left(h^{1/4}h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\psi\right) \\
&\quad + h^{ij}h^{1/4}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\partial_i\left(h^{-1/4}\psi\right) + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\psi \\
&\quad - \tfrac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right)\psi - \tfrac{1}{4}S\psi.
\end{aligned}
$$

In order to exhibit the parts of A which are of pure order 0, 1 and 2, we expand further:

$$
\begin{aligned}
A[\psi] &= h^{ij}\partial_j\partial_i\psi + \left(h^{-1/4}\partial_j\left(h^{ij}h^{1/4}\right) + h^{1/4}h^{ji}\partial_j\left(h^{-1/4}\right)\right)\partial_i\psi \\
&\quad + h^{-1/4}\partial_j\left(h^{ij}\sqrt{h}\partial_i\left(h^{-1/4}\right)\right)\psi \\
&\quad + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\partial_j\psi + h^{-1/4}\partial_j\left(h^{1/4}h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\right)\psi \\
&\quad + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\partial_i\psi + h^{ij}h^{1/4}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\partial_i\left(h^{-1/4}\right)\psi \\
&\quad + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\psi - \tfrac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right)\psi - \tfrac{1}{4}S\psi \\
&= h^{ij}\partial_j\partial_i\psi \\
&\quad + \left(h^{-1/4}\partial_j\left(h^{ij}h^{1/4}\right) + h^{1/4}h^{ji}\partial_j\left(h^{-1/4}\right)\right)\partial_i\psi \\
&\quad + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\partial_j\psi + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\partial_i\psi \\
&\quad + h^{-1/4}\partial_j\left(h^{1/4}h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\right)\psi + h^{ij}h^{1/4}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\partial_i\left(h^{-1/4}\right)\psi \\
&\quad + h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\psi - \tfrac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right)\psi - \tfrac{1}{4}S\psi.
\end{aligned}
$$

Finally,

$$A[\psi] = h^{ij}\partial_j\partial_i\psi + \left(h^{-1/4}\partial_j\left(h^{ij}h^{1/4}\right) + h^{1/4}h^{ji}\partial_j\left(h^{-1/4}\right) + 2h^{ji}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\right)\partial_i\psi$$

(18.100)

$$+\begin{pmatrix} h^{-1/4}\partial_j\left(h^{1/4}h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)\right) + h^{ij}h^{1/4}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\,\partial_i\left(h^{-1/4}\right) \\ h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i) - \frac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right) - \frac{1}{4}S\end{pmatrix}\psi.$$

We wish to alter the operator A in such a way that the alteration does not affect $\operatorname{Str}(h_m(0,0))$, but the altered operator is much simpler. Note that in (18.100), differential operator A has three parts of orders 2, 1, and 0 which contain 1, 3, and 5 terms each. Correspondingly, there are $(1{+}3{+}5)^m = 9^m$ terms in the composition $A_{s_{m-1}}\circ\cdots\circ A_{s_0}$. However, any of these terms of $A_{s_{m-1}}\circ\cdots\circ A_{s_0}$ which involve fewer than $2m$ gamma matrices γ^i will not contribute to $\operatorname{Str}(h_m(0,0))$. The term of A which produces the most (four) γ^i is a subterm of $h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_i)$, namely

(18.101)

$$h^{ij}\widetilde{\theta}(\partial_j)\,\widetilde{\theta}(\partial_i)\,\psi = h^{ij}\left(R_{klpj}(0)\,\tfrac{1}{8}\gamma^k\gamma^l y^p\right)\left(R_{k'l'qi}(0)\,\tfrac{1}{8}\gamma^{k'}\gamma^{l'}\right)y^q\psi + O\left(|y|^3\right).$$

Although this subterm contributes four γ^i, it introduces two factors of y^i which must be differentiated by terms in subseqent factors of $A_{s_{m-1}}\circ\cdots\circ A_{s_0}$ in order to contribute to $\operatorname{Str}(h_m(0,0))$. We say that the *degree of contribution* ($\deg_c$) of the subterm (18.101) to $\operatorname{Str}(h_m(0,0))$ is $4-2$ or 2. Note that $h^{ij}\partial_j\partial_i\psi$ has no γ^i, but it contains two differentiations which can serve to eliminate two factors of y^i in previous terms. Hence, although $h^{ij}\partial_j\partial_i$ contains no γ^i, we take $\deg_c\left(h^{ij}\partial_j\partial_i\psi\right) = 2$. In general, we have

Definition 18.46. For $B(0,r_0) = \{y\in\mathbb{R}^n : |y|\le r_0\}$, let

$$T : C^\infty\left(B(0,r_0),\mathbb{C}^N\otimes\Sigma_{2m}\right)\to C^\infty\left(B(0,r_0),\mathbb{C}^N\otimes\Sigma_{2m}\right)$$

be an operator of the form

$$T(\psi)(y) = \sum_{(i),(j)} f^{i_1\cdots i_k}_{j_1\cdots j_p}(y)\,\gamma^{j_1}\cdots\gamma^{j_p}\left(\partial_{i_1}\cdots\partial_{i_k}\psi\right)(y),$$

where $f^{i_1\cdots i_k}_{j_1\cdots j_p}\in C^\infty(B(0,r_0),\mathbb{R})$ and (i) ranges over all multi-indices $(i_1,\ldots,i_k)\in\times^k\{1,\ldots,n\}$. Then the **degree of contribution** of T is

$$\deg_c(T) := p+q-r, \text{ where}$$

$$r := \min_{(i),(j)}\left\{d : f^{i_1\cdots i_k}_{j_1\cdots j_p}(y) = \mathrm{O}\left(|y|^d\right)\right\};$$

i.e., r is smallest of the degrees of the lead parts of the Taylor expansions of the coefficients $f^{i_1\cdots i_k}_{j_1\cdots j_p}$. For an integer $k\ge 0$, we use $\mathrm{O}_c(k)$ to denote an operator with $\deg_c\le k$.

Refering to (18.100), the only term in

$$\left(h^{-1/4}\partial_j\left(h^{ij}h^{1/4}\right) + h^{1/4}h^{ji}\partial_j\left(h^{-1/4}\right) + 2h^{ji}\left(\varepsilon\oplus\widetilde{\theta}\right)(\partial_j)\right)\partial_i$$

which has an $\deg_c$ of at least 2 is

$$2h^{ji}\widetilde{\theta}\left(\partial_j\right)\partial_i.$$

In (18.100), the sum of the terms of

$$\begin{pmatrix} h^{-1/4}\partial_j\left(h^{1/4}h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)\left(\partial_i\right)\right)+h^{ij}h^{1/4}\left(\varepsilon\oplus\widetilde{\theta}\right)\left(\partial_j\right)\partial_i\left(h^{-1/4}\right) \\ h^{ij}\left(\varepsilon\oplus\widetilde{\theta}\right)\left(\partial_j\right)\left(\varepsilon\oplus\widetilde{\theta}\right)\left(\partial_i\right)-\frac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right)-\frac{1}{4}S \end{pmatrix}$$

which have $\deg_c \geq 2$ is

$$h^{ij}\partial_j\left(\widetilde{\theta}\left(\partial_i\right)\right)+h^{ij}\widetilde{\theta}\left(\partial_j\right)\widetilde{\theta}\left(\partial_i\right)-\frac{1}{2}\left(\Omega^{\varepsilon}_{jk}\otimes\gamma^j\gamma^k\right).$$

Since there are only m factors in $A_{s_{m-1}}\circ\cdots\circ A_{s_0}$, the maximim number of γ^i in factors of the terms of $A_{s_{m-1}}\circ\cdots\circ A_{s_0}$ will be $2m$. To achieve this maximal number it is necessary (but not sufficient) that such a term be a composition of m factors each with the maximal $\deg_c$, namely 2. Thus, we can alter A without changing $\operatorname{Str}\left(h_m\left(0,0\right)\right)$ by retaining only those terms of A with $\deg_c = 2$. Moreover, only the lead part of the Taylor expansion in y of such terms need be retained. In the next lemma, we collect the relevant expansions that we have already found for the metric and the connections in terms of normal coordinates and the radial gauge, at least to the order we need.

Lemma 18.47. *We have*

$$\begin{aligned}
\varepsilon_y\left(\partial_j\right) &= \tfrac{1}{2}F_{ij}y^i+\mathrm{O}\left(|y|^2\right)=\tfrac{1}{2}\Omega^{\varepsilon}_{ij}\left(0\right)y^i+\mathrm{O}\left(|y|^2\right)\\
\widetilde{\theta}_y\left(\partial_j\right) &= \tfrac{1}{8}R_{klji}\left(0\right)\gamma^k\gamma^l y^i+\mathrm{O}\left(|y|^2\right)\\
h_{ij}\left(y\right) &= \delta_{ij}-\tfrac{1}{3}R_{ikjl}\left(0\right)y^k y^l+\mathrm{O}\left(|y|^3\right)\\
h^{ij}\left(y\right) &= \delta_{ij}+\tfrac{1}{3}R_{ikjl}\left(0\right)y^k y^l+\mathrm{O}\left(|y|^3\right)\\
h^{\alpha}:=\left(\det\mathbf{h}\right)^{\alpha} &= 1-\tfrac{1}{3}\alpha R_{kl}\left(0\right)y^k y^l+\mathrm{O}\left(|y|^3\right)\\
\tfrac{1}{2}\Omega^{\varepsilon}_{jk}\gamma^j\gamma^k\psi &= \tfrac{1}{2}\sum\nolimits_{i,j}\Omega^{\varepsilon}_{ij}\left(0\right)\otimes\gamma^i\gamma^j+\mathrm{O}\left(|y|\right)\\
\gamma^i &= h^{ij}\gamma_j=\gamma_i+\mathrm{O}\left(|y|^2\right).
\end{aligned}$$

We define

$$\widetilde{\theta}^1\left(\partial_j\right):=\tfrac{1}{8}R_{klji}\left(0\right)\gamma^k\gamma^l y^i \quad \text{and}$$

$$\mathcal{F}^0:=\tfrac{1}{2}\sum\nolimits_{i,j}F_{ij}\otimes\gamma^i\gamma^j=\tfrac{1}{2}\sum\nolimits_{i,j}\Omega^{\varepsilon}_{ij}\left(0\right)\otimes\gamma^i\gamma^j \tag{18.102}$$

Thus, a permissible alteration of A is

$$\begin{aligned}
A' &:= \sum\nolimits_i\left(\partial_i^2+2\widetilde{\theta}\left(\partial_i\right)\partial_i+\partial_i\left(\widetilde{\theta}\left(\partial_i\right)\right)+\widetilde{\theta}\left(\partial_i\right)^2\right)-\mathcal{F}^0\\
&= \sum\nolimits_i\left(\partial_i^2+2\widetilde{\theta}\left(\partial_i\right)\partial_i+\widetilde{\theta}\left(\partial_i\right)^2\right)-\mathcal{F}^0,
\end{aligned}$$

where we have used

$$\begin{aligned}
\partial_i\left(\widetilde{\theta}\left(\partial_i\right)\right) &= \partial_i\left(\tfrac{1}{8}R_{klij}\left(0\right)\gamma^k\gamma^l y^j\right)\\
&= \tfrac{1}{8}R_{klij}\left(0\right)\gamma^k\gamma^l\delta_i^j=\tfrac{1}{8}R_{klii}\left(0\right)\gamma^k\gamma^l=0.
\end{aligned}$$

We now show that even though

$$\deg_c\left(2\widetilde{\theta}\left(\partial_i\right)\partial_i\right) = \deg_c\left(\tfrac{1}{4}\sum\nolimits_j R_{klij}\left(0\right)\gamma^k\gamma^l y^j\partial_i\right) = 2,$$

the term $2\widetilde{\theta}\left(\partial_i\right)\partial_i$ may be dropped without changing $\operatorname{Str}\left(h_m\left(0,0\right)\right)$. Note that if this $2\widetilde{\theta}\left(\partial_i\right)\partial_i$ is on the far left in a term of $A'_{s_{m-1}} \circ \cdots \circ A'_{s_0}$, then the contribution of this term to $A'_{s_{m-1}} \circ \cdots \circ A'_{s_0}\left[V_0\right](0)$ will be zero, because of the factor of y^j in $2\widetilde{\theta}\left(\partial_i\right)\partial_i$ which is set to 0. Thus, it suffices to prove that $2\widetilde{\theta}\left(\partial_i\right)\partial_i$ commutes with $\partial_i^2, \widetilde{\theta}\left(\partial_i\right)^2$ and $\mathcal{F}^0$, modulo terms which do not affect $\operatorname{Str}\left(h_m\left(0,0\right)\right)$. In other words, we need to check that the commutators

$$\begin{aligned}
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \mathcal{F}^0\right],\\
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \sum\nolimits_j \partial_j^2\right], \text{ and}\\
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \sum\nolimits_j \widetilde{\theta}\left(\partial_j\right)^2\right].
\end{aligned} \tag{18.103}$$

each have $\deg_c < 4$. We have

$$\begin{aligned}
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \mathcal{F}^0\right]\\
&= \left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \tfrac{1}{2}\sum\nolimits_{jk}\left(F_{jk}\otimes\gamma^j\gamma^k\right)\right]\\
&= \tfrac{1}{2}\sum\nolimits_{i,j,k}\left[\widetilde{\theta}\left(\partial_i\right)\partial_i, \left(F_{jk}\otimes\gamma^j\gamma^k\right)\right]\\
&= \tfrac{1}{2}\sum\nolimits_{i,j,k,p}\left[\tfrac{1}{8}R_{pliq}\left(0\right)\gamma^p\gamma^l y^q\partial_i, \left(F_{jk}\otimes\gamma^j\gamma^k\right)\right]\\
&= \tfrac{1}{2}\sum\nolimits_{i,j,k}\tfrac{1}{8}y^q R_{pliq}\left(0\right)F_{jk}\otimes\left[\gamma^p\gamma^l, \gamma^j\gamma^k\right]\partial_i.
\end{aligned}$$

Since $\deg_c\left(\left[\gamma^p\gamma^l, \gamma^j\gamma^k\right]\right) \leq 2$, we have $\deg_c\left(\left[\sum_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \mathcal{F}^0\right]\right) \leq 2$. Moreover,

$$\begin{aligned}
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \sum\nolimits_j \partial_j^2\right]\\
&= \sum\nolimits_{i,j}\widetilde{\theta}\left(\partial_i\right)\partial_i\partial_j^2 - \partial_j^2\left(\widetilde{\theta}\left(\partial_i\right)\partial_i\right)\\
&= \sum\nolimits_{i,j}\widetilde{\theta}\left(\partial_i\right)\partial_i\partial_j^2 - \partial_j\left(\partial_j\left(\widetilde{\theta}\left(\partial_i\right)\right)\partial_i + \widetilde{\theta}\left(\partial_i\right)\partial_j\partial_i\right)\\
&= \sum\nolimits_{i,j}\widetilde{\theta}\left(\partial_i\right)\partial_i\partial_j^2 - 2\left(\partial_j^2\left(\widetilde{\theta}\left(\partial_i\right)\right)\partial_i + 2\partial_j\left(\widetilde{\theta}\left(\partial_i\right)\right)\partial_j\partial_i + \widetilde{\theta}\left(\partial_i\right)\partial_j^2\partial_i\right)\\
&= \sum\nolimits_{i,j} -4\partial_j\left(\widetilde{\theta}\left(\partial_i\right)\right)\partial_j\partial_i = \sum\nolimits_{i,j} -4\partial_j\left(-\tfrac{1}{8}R_{klji}\left(0\right)\gamma^k\gamma^l y^j\right)\partial_j\partial_i\\
&= \tfrac{1}{2}\sum\nolimits_{i,j} R_{klji}\left(0\right)\gamma^k\gamma^l\partial_j\partial_i = 0,
\end{aligned}$$

since R_{klji} is anti-symmetric in i and j, while $\partial_j\partial_i$ is symmetric. We show $\deg_c\left(\left[\sum_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \sum_j \widetilde{\theta}\left(\partial_j\right)^2\right]\right) \leq 2$ as follows:

$$
\begin{aligned}
&\left[\sum\nolimits_i \widetilde{\theta}\left(\partial_i\right)\partial_i, \sum\nolimits_j \widetilde{\theta}\left(\partial_j\right)^2\right] \\
&= \sum\nolimits_{i,j} \widetilde{\theta}\left(\partial_i\right)\partial_i \circ \left(\widetilde{\theta}\left(\partial_j\right)^2\right) - \widetilde{\theta}\left(\partial_j\right)^2 \widetilde{\theta}\left(\partial_i\right)\partial_i \\
&= \sum\nolimits_{i,j} \widetilde{\theta}\left(\partial_i\right)\partial_i \left(\widetilde{\theta}\left(\partial_j\right)^2\right) + \widetilde{\theta}\left(\partial_i\right)\widetilde{\theta}\left(\partial_j\right)^2 \partial_i - \widetilde{\theta}\left(\partial_j\right)^2 \widetilde{\theta}\left(\partial_i\right)\partial_i \\
&= \sum\nolimits_{i,j} \widetilde{\theta}\left(\partial_i\right)\partial_i \left(\widetilde{\theta}\left(\partial_j\right)^2\right) + \left[\widetilde{\theta}\left(\partial_i\right), \widetilde{\theta}\left(\partial_j\right)^2\right]\partial_i \\
&= \sum\nolimits_{i,j} \widetilde{\theta}\left(\partial_i\right)\partial_i \left(\widetilde{\theta}\left(\partial_j\right)^2\right) + \mathrm{O}_c\left(4-3+1\right) \\
&= \sum\nolimits_{i,j} \widetilde{\theta}\left(\partial_i\right)\partial_i \left(\widetilde{\theta}\left(\partial_j\right)^2\right) + \mathrm{O}_c\left(2\right).
\end{aligned}
$$

The first term appears to have $\deg_c$ 4 which is too high, but we now show that it actually has $\deg_c$ 2. Using

$$
\begin{aligned}
-8\widetilde{\theta}\left(\partial_j\right) &= R_{rsqj}\gamma^r\gamma^s y^q \text{ and} \\
8^2\widetilde{\theta}\left(\partial_j\right)^2 &= R_{rsqj}R_{r's'q'j}\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}y^q y^{q'},
\end{aligned}
$$

we have

$$
\begin{aligned}
&-8^3\widetilde{\theta}\left(\partial_i\right)\partial_i\left(\widetilde{\theta}\left(\partial_j\right)^2\right) \\
&= R_{klpi}\gamma^k\gamma^l y^p \partial_i\left(R_{rsqj}R_{r's'q'j}\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}y^q y^{q'}\right) \\
&= R_{klpi}\gamma^k\gamma^l y^p R_{rsqj}R_{r's'q'j}\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\left(\delta_{iq}y^{q'} + \delta_{iq'}y^q\right) \\
&= R_{klpi}R_{rsij}R_{r's'q'j}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^{q'} + R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q \\
&= R_{klpi}R_{rsij}R_{r's'qj}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q + R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q \\
&= R_{klpi}R_{r's'ij}R_{rsqj}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q \\
&+ R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q + \mathrm{O}_c\left(4-2\right) \\
&= 2R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^p y^q + \mathrm{O}_c\left(2\right)
\end{aligned}
$$

The first term which appears to be $\mathrm{O}_c(4)$ is actually $\mathrm{O}_c(2)$, since

$$\begin{aligned}
&R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= R_{klqi}R_{rspj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= -R_{klqi}R_{rspj}R_{r's'ji}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= -R_{klqj}R_{rspi}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= -R_{rspi}R_{klqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= -R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^r\gamma^s\gamma^k\gamma^l\gamma^{r'}\gamma^{s'}\right)y^py^q \\
&= -R_{klpi}R_{rsqj}R_{r's'ij}\left(\gamma^k\gamma^l\gamma^r\gamma^s\gamma^{r'}\gamma^{s'}\right)y^py^q + \mathrm{O}_c(4-2),
\end{aligned}$$

where we have used $\left[\gamma^r\gamma^s,\gamma^k\gamma^l\right] = \mathrm{O}_c(2)$ in the last equality.

Thus, we have shown that all the commutators in (18.103) have $\deg_c \le 2 < 4$. In summary, we have

Proposition 18.48. *Let*

$$\widetilde{\theta}^1(\partial_j) := \tfrac{1}{8}\sum\nolimits_{k,l,i} R_{klji}(0)\,\gamma^k\gamma^l y^i,$$

$$\mathcal{F}^0 := \tfrac{1}{2}\sum\nolimits_{i,j} F_{ij}\otimes\gamma^i\gamma^j = \tfrac{1}{2}\sum\nolimits_{i,j}\Omega^{\varepsilon}_{ij}(0)\otimes\gamma^i\gamma^j, \text{ and}$$

$$A^0 := \sum_i\left(\partial_i^2 + \widetilde{\theta}(\partial_i)^2\right) - \mathcal{F}_0. \tag{18.104}$$

For $V_0\in\mathbb{C}^N\otimes\Sigma_{2m}$, *define*

$$h_m^0(0,0)(V_0) := \int_{I^k}\prod\nolimits_{i=0}^{m-1}(s_i)^i\left(\left(A^0_{s_{m-1}}\circ\cdots\circ A^0_{s_0}\right)\left[\widetilde{V}_0\right]\right)(0)\,ds_0\ldots ds_{m-1}, \tag{18.105}$$

where $\widetilde{V}_0\in C^\infty\left(B(r_0,0),\mathbb{C}^N\otimes\Sigma_{2m}\right)$ *is the constant extension of* $V_0\in\mathbb{C}^N\otimes\Sigma_{2m}$. *Then*

$$\operatorname{Str}(h_m(0,0)) = \operatorname{Str}\left(h_m^0(0,0)\right). \tag{18.106}$$

In other words, in the computation of $\operatorname{Str}(h_m(0,0))$ *given by (18.99), we may replace* A *by* A^0.

At this point, we have that

$$\operatorname{index}\left(\mathcal{D}^+\right) = (4\pi)^{-m}\int_M \operatorname{Str}(h_m(x,x))\,\nu_h(x)$$

where the form $\operatorname{Str}(h_m(x,x))\,\nu_h(x)$ only depends on $\nu_h(x)$, the curvature of the Levi-Civita connection of h at x, and the curvature of the connection ε on $U(E)$ at x. Since the $\operatorname{index}(\mathcal{D}^+)$ is invariant under perturbations of the metric h and the connection ε, we suspect that in fact $\operatorname{Str}(h_m(x,x))\,\nu_h(x)$ can be expressed in terms of curvature forms of characteristic classes, whose integrals are also invariant. It remains to do this by establishing (18.98). Before doing this in general, we verify (18.98) in the cases $m=1$ and $m=2$ (i.e., for surfaces and 4-manifolds). For readers who have no use for the Local Index Theorem beyond dimension 4, this is sufficient. For $m = 1$ and $m = 2$, we do not need to use Mehler's Formula but rather we proceed directly using (18.105), (18.106) and

$$\begin{aligned} &A_s^0 [\psi] (y) := A^0 [\psi] (sy) \\ &= \sum_i \left(\partial_i^2 [\psi]\right) (sy) + s^2 \sum_i \widetilde{\theta} (\partial_i)^2 \psi (sy) - \tfrac{1}{2} \sum\nolimits_{j,k} \left(F_{jk} \otimes \gamma^j \gamma^k \psi (sy)\right). \end{aligned}$$

5.1. The case $m = 1$ (surfaces). We have

$$\begin{aligned} h_1^0 (0,0) (V_0) &= \int_0^1 (s_0)^0 A_{s_0}^0 [V_0] (0)\, ds_0 = \int_0^1 A_{s_0}^0 [V_0] (0)\, ds_0 \\ &= \int_0^1 A^0 [V_0] (s_0 0)\, ds_0 = \int_0^1 A^0 [V_0] (0)\, ds_0 = A^0 [V_0] (0) \\ &= \partial_i^2 [V_0] (0) + \widetilde{\theta} (\partial_i)^2 [V_0] (0) - \tfrac{1}{2} \left(\sum\nolimits_{j,k} F_{jk} \otimes \gamma^j \gamma^k\right) (V_0) \\ &= -\tfrac{1}{2} \sum\nolimits_{j,k} \left(F_{jk} \otimes \gamma^j \gamma^k\right) (V_0) \end{aligned}$$

For $m = 1$, $\operatorname{Str} (\gamma_{2m+1}) = (-2i)^m$ yields $\operatorname{Str} (\gamma_1 \gamma_2) = -2i$. Hence,

$$\begin{aligned} &(4\pi)^{-1} \operatorname{Str} (h_1 (0,0)) = (4\pi)^{-1} \operatorname{Str} \left(-\tfrac{1}{2} \sum\nolimits_{j,k} \left(F_{jk} \otimes \gamma^j \gamma^k\right)\right) \\ &= (4\pi)^{-1} \operatorname{Tr} (-F_{12}) \operatorname{Str} (\gamma_1 \gamma_2) = - (4\pi)^{-1} \operatorname{Tr} (\Omega_{12}^{\varepsilon} (0)) (-2i)^1 \\ &= \frac{i}{2\pi} \operatorname{Tr} \Omega_{12}^{\varepsilon} (0). \end{aligned}$$

Thus, we have (18.98) in the case $m = 1$, since

$$\begin{aligned} &\left\langle \operatorname{Tr} \left(e^{i\Omega^{\varepsilon}/2\pi}\right) \wedge \det \left(\frac{i\Omega^{\theta}/4\pi}{\sinh (i\Omega^{\theta}/4\pi)}\right)^{\frac{1}{2}}, \nu_h (0) \right\rangle \\ &= \langle \operatorname{Tr} (i\Omega^{\varepsilon} (0) /2\pi), \nu_h (0) \rangle = \frac{i}{2\pi} \operatorname{Tr} \Omega_{12}^{\varepsilon} (0). \quad \square \end{aligned}$$

5.2. The case $m=2$ (4-manifolds). For $m=2$, we have

$$h_2^0(0,0)(V_0)=\int_0^1\int_0^1 (s_1)^1(s_0)^0 A^0_{s_1}A^0_{s_0}[V_0](0)\,ds_0ds_1$$
$$=\int_0^1\int_0^1 (s_1)^1 A^0 A^0_{s_0}[V_0](s_10)\,ds_0ds_1$$
$$=\int_0^1\int_0^1 (s_1)^1 A^0\left[A^0_{s_0}[V_0]\right](0)\,ds_0ds_1$$
$$=\tfrac{1}{2}\int_0^1 A^0\left[A^0_{s_0}[V_0]\right](0)\,ds_0$$
$$=\tfrac{1}{2}\int_0^1 A^0\left[s_0^2\sum\nolimits_i\widetilde{\theta}(\partial_i)^2V_0-\tfrac{1}{2}\sum\nolimits_{j,k}\left(F_{jk}\otimes\gamma^j\gamma^kV_0\right)\right](0)\,ds_0$$
$$=\tfrac{1}{2}\int_0^1 s_0^2A^0\left[\sum\nolimits_i\widetilde{\theta}(\partial_i)^2V_0\right](0)\,ds_0-\tfrac{1}{2}A^0\left[\tfrac{1}{2}\sum\nolimits_{j,k}\left(F_{jk}\otimes\gamma^j\gamma^kV_0\right)\right](0)$$
$$=\tfrac{1}{6}A^0\left[\sum\nolimits_i\widetilde{\theta}(\partial_i)^2V_0\right](0)-\tfrac{1}{2}A^0\left[\tfrac{1}{2}\sum\nolimits_{j,k}\left(F_{jk}\otimes\gamma^j\gamma^kV_0\right)\right](0)$$
$$=\tfrac{1}{6}\sum\nolimits_{i,j}(\partial_j)^2\left[\widetilde{\theta}(\partial_i)^2V_0\right]+\tfrac{1}{2}\cdot\tfrac{1}{2}\sum\nolimits_{j,k,j',k'}\left(F_{j'k'}\otimes\gamma^{j'}\gamma^{k'}\right)\left[\tfrac{1}{2}\left(F_{jk}\otimes\gamma^j\gamma^kV_0\right)\right]$$
$$=\tfrac{1}{6}\sum\nolimits_{i,j}(\partial_j)^2\left[\widetilde{\theta}(\partial_i)^2V_0\right]+\tfrac{1}{8}\sum\nolimits_{j,k,j',k'}\left(F_{j'k'}\circ F_{jk}\otimes\gamma^{j'}\gamma^{k'}\gamma^j\gamma^k\right)[V_0]$$
$$=\tfrac{1}{6}\tfrac{1}{8^2}2\sum\nolimits_{i,p,k,l,k',l'}R_{klpi}(0)\,R_{k'l'pi}(0)\,\gamma^k\gamma^l\gamma^{k'}\gamma^{l'}[V_0]$$
$$+\tfrac{1}{8}\sum\nolimits_{j,k,j',k'}\left(\Omega^{\varepsilon}_{j'k'}(0)\circ\Omega^{\varepsilon}_{jk}(0)\otimes\gamma^{j'}\gamma^{k'}\gamma^j\gamma^k\right)[V_0],$$

where we have used

$$\tfrac{1}{6}\sum\nolimits_{i,j}(\partial_j)^2\left[\widetilde{\theta}(\partial_i)^2V_0\right]$$
$$=\tfrac{1}{6}\sum\nolimits_{i,j,p,k,l}(\partial_j)^2\left(R_{klip}(0)\left(\tfrac{1}{8}\gamma^k\gamma^l\right)y^p\right)^2[V_0]$$
$$=\tfrac{1}{6}\tfrac{1}{8^2}\sum\nolimits_{i,j,p,q,k,l,k',l'}(\partial_j)^2\left(R_{klip}(0)\,R_{k'l'iq}(0)\,\gamma^k\gamma^l\gamma^{k'}\gamma^{l'}y^py^q\right)[V_0]$$
$$=\tfrac{1}{6}\tfrac{1}{8^2}2\sum\nolimits_{i,p,k,l,k',l'}R_{klip}(0)\,R_{k'l'ip}(0)\,\gamma^k\gamma^l\gamma^{k'}\gamma^{l'}[V_0].$$

Since $\operatorname{Str}(\gamma_{2m+1}) = (-2i)^m \Rightarrow \operatorname{Str}\left(\gamma^k\gamma^l\gamma^{k'}\gamma^{l'}\right) = (-2i)^2\,\varepsilon_{klk'l'}$ when $m = 2$, we get

$$\begin{aligned}
&\operatorname{Str}(h_2(0,0))\\
&= \tfrac{1}{3}\tfrac{1}{8^2}\operatorname{Tr}(\mathrm{Id}_E)\sum\nolimits_{i,p,k,l,k',l'} R_{klip}(0)\,R_{k'l'ip}(0)\operatorname{Str}\left(\gamma^k\gamma^l\gamma^{k'}\gamma^{l'}\right)\\
&+ \tfrac{1}{8}\sum\nolimits_{j,k,j',k'} \operatorname{Tr}\left(\Omega^{\varepsilon}_{j'k'}(0)\circ\Omega^{\varepsilon}_{jk}(0)\right)\operatorname{Str}\left(\gamma^{j'}\gamma^{k'}\gamma^{j}\gamma^{k}\right)\\
&= \tfrac{1}{3}\tfrac{1}{8^2}\dim E\cdot\sum\nolimits_{i,p,k,l,k',l'} R_{klip}(0)\,R_{k'l'ip}(0)\,(-2i)^2\,\varepsilon_{klk'l'}\\
&+ \tfrac{1}{8}\sum\nolimits_{j,k,j',k'} \operatorname{Tr}\left(\Omega^{\varepsilon}_{j'k'}(0)\circ\Omega^{\varepsilon}_{jk}(0)\right)(-2i)^2\,\varepsilon_{j'k'jk}\\
&= \tfrac{1}{3}\tfrac{4}{8^2}\dim E\cdot\sum\nolimits_{i,p,k,l,k',l'} R_{ipkl}(0)\,R_{pik'l'}(0)\,\varepsilon_{klk'l'}\\
&- \tfrac{1}{2}\sum\nolimits_{j,k,j',k'} \operatorname{Tr}\left(\Omega^{\varepsilon}_{j'k'}(0)\circ\Omega^{\varepsilon}_{jk}(0)\right)\varepsilon_{j'k'jk}
\end{aligned}$$

We rewrite this as follows. In $\mathbb{R}^4$, for 2-forms $\alpha = \frac{1}{2}\sum_{k,l}\alpha_{kl}dy^k\wedge dy^l$ and $\beta = \frac{1}{2}\sum_{k,l}\beta_{k'l'}dy^{k'}\wedge dy^{l'}$,

$$\sum_{k,l,k',l'}\alpha_{kl}\beta_{k'l'}\varepsilon_{klk'l'} = 4\,\langle\alpha\wedge\beta,\nu_h\rangle\,.$$

Thus, we have

$$\sum\nolimits_{j,k,j',k'}\operatorname{Tr}\left(\Omega^{\varepsilon}_{j'k'}(0)\circ\Omega^{\varepsilon}_{jk}(0)\right)\varepsilon_{j'k'jk} = 4\,\langle\operatorname{Tr}(\Omega^{\varepsilon}\wedge\Omega^{\varepsilon}),\nu_h(0)\rangle\,.$$

and

$$\begin{aligned}
&\operatorname{Str}(h_2(0,0))\\
&= \tfrac{1}{3}\tfrac{4}{8^2}\dim E\cdot\sum\nolimits_{i,p,k,l,k',l'} R_{pikl}(0)\,R_{ipk'l'}(0)\,\varepsilon_{klk'l'} - 2\,\langle\operatorname{Tr}(\Omega^{\varepsilon}\wedge\Omega^{\varepsilon}),\nu_h(0)\rangle\,.
\end{aligned}$$

We can write the first sum as follows. Since

$$\Omega^{\theta}_{pi} = \tfrac{1}{2}\sum\nolimits_{k,l} R_{pikl}dy^k\wedge dy^l,\ \text{ we have}$$

$$\begin{aligned}
\left(\Omega^{\theta}\wedge\Omega^{\theta}\right)_{ji} &= \sum\nolimits_{p}\left(\Omega^{\theta}_{jp}\wedge\Omega^{\theta}_{pi}\right)\\
&= \tfrac{1}{4}\sum\nolimits_{p,k,l,k',l'} R_{jpkl}R_{pik'l'}dy^k\wedge dy^l\wedge dy^{k'}\wedge dy^{l'}\\
&= \tfrac{1}{4}\sum\nolimits_{p,k,l,k',l'}\left(R_{jpkl}R_{pik'l'}\varepsilon_{klk'l'}\right)\nu,\ \text{ and}\\
\operatorname{Tr}\left(\Omega^{\theta}\wedge\Omega^{\theta}\right) &= \sum\nolimits_{i}\left(\Omega^{\theta}\wedge\Omega^{\theta}\right)_{ii} = \tfrac{1}{4}\sum\nolimits_{i,p,k,l,k',l'}\left(R_{ipkl}R_{pik'l'}\varepsilon_{klk'l'}\right)\nu.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\operatorname{Str}(h_2(0,0))\\
&= \tfrac{1}{3}\tfrac{4}{8^2}\dim E\cdot 4\left\langle\operatorname{Tr}\left(\Omega^{\theta}\wedge\Omega^{\theta}\right),\nu_h(0)\right\rangle - 2\,\langle\operatorname{Tr}(\Omega^{\varepsilon}\wedge\Omega^{\varepsilon}),\nu_h(0)\rangle\\
&= \left\langle\tfrac{1}{12}\dim E\cdot\operatorname{Tr}\left(\Omega^{\theta}\wedge\Omega^{\theta}\right) - 2\operatorname{Tr}(\Omega^{\varepsilon}\wedge\Omega^{\varepsilon}),\nu_h(0)\right\rangle.
\end{aligned}$$

From the definition (16.100) of the Pontryagin forms,

$$\begin{aligned} p_1\left(\Omega^\theta\right) &= \frac{1}{(2\pi)^2\, 2!}\sum\nolimits_{i_1,i_2,j_1,j_2} \delta^{j_1 j_2}_{i_1 i_2}\Omega^\theta_{i_1 j_1}\wedge\Omega^\theta_{i_2 j_2} \\ &= \frac{-1}{8\pi^2}\sum\nolimits_{i_1,i_2}\Omega^\theta_{i_1 i_2}\wedge\Omega^\theta_{i_2 i_1} = \frac{-1}{8\pi^2}\operatorname{Tr}\left(\Omega^\theta\wedge\Omega^\theta\right). \end{aligned}$$

In dimension 4, we have

$$\begin{aligned} &\operatorname{Tr}\left(e^{i\Omega^\varepsilon/2\pi}\right) = \operatorname{Tr} I_E + \frac{i}{2\pi}\operatorname{Tr}\left(\Omega^\varepsilon\right) + \tfrac{1}{2}\left(\frac{i}{2\pi}\right)^2 \operatorname{Tr}\left(\Omega^\varepsilon\wedge\Omega^\varepsilon\right) \\ &= \dim E + \frac{i}{2\pi}\operatorname{Tr}\left(\Omega^\varepsilon\right) - \frac{1}{8\pi^2}\operatorname{Tr}\left(\Omega^\varepsilon\wedge\Omega^\varepsilon\right), \text{ and} \\ &\det\left(\frac{i\Omega^\theta/4\pi}{\sinh\left(i\Omega^\theta/4\pi\right)}\right)^{\frac{1}{2}} = 1 - \frac{1}{24}p_1\left(\Omega^\theta\right). \end{aligned}$$

Thus, as required,

$$\begin{aligned} &\left\langle \operatorname{Tr}\left(e^{i\Omega^\varepsilon/2\pi}\right)\wedge\det\left(\frac{i\Omega^\theta/4\pi}{\sinh\left(i\Omega^\theta/4\pi\right)}\right)^{\frac{1}{2}}, \nu_h(0)\right\rangle \\ &= \left\langle -\frac{1}{24}\left(\dim E\right)p_1\left(\Omega^\theta\right) - \frac{1}{8\pi^2}\operatorname{Tr}\left(\Omega^\varepsilon\wedge\Omega^\varepsilon\right), \nu_h(0)\right\rangle \\ &= \left\langle -\frac{1}{24}\left(\dim E\right)\left(\frac{-1}{8\pi^2}\operatorname{Tr}\left(\Omega^\theta\wedge\Omega^\theta\right)\right) - \frac{1}{8\pi^2}\operatorname{Tr}\left(\Omega^\varepsilon\wedge\Omega^\varepsilon\right), \nu_h(0)\right\rangle \\ &= \frac{1}{16\pi^2}\left\langle \tfrac{1}{12}\dim E\cdot\operatorname{Tr}\left(\Omega^\theta\wedge\Omega^\theta\right) - 2\operatorname{Tr}\left(\Omega^\varepsilon\wedge\Omega^\varepsilon\right), \nu_h(0)\right\rangle \\ &= \frac{1}{(4\pi)^2}\operatorname{Str}\left(h_2(0,0)\right). \quad \square \end{aligned}$$

Our starting point for the derivation of the Local Index Formula (18.98) for arbitrary even dimensions is the determination of the kernel, say $e_a(x,y,t)$, for the generalized 1-dimensional heat equation

$$u_t = u_{yy} - a^2y^2u, \quad u(y,t)\in\mathbb{R},\ (y,t)\in\mathbb{R}\times(0,\infty),$$

where $a\in\mathbb{R}$ is a given constant. If $a=0$, we have the familiar result

$$e_0(x,y,t) = \frac{1}{\sqrt{4\pi t}}e^{-\frac{1}{4t}(y-x)^2}.$$

For $a\neq 0$, we derive Mehler's formula
(18.107)

$$e_a(x,y,t) = \frac{1}{\sqrt{4\pi\frac{\sinh(2at)}{2a}}}\exp\left(-\frac{1}{4\frac{\sinh(2at)}{2a}}\left(\cosh(2at)\left(x^2+y^2\right)-2xy\right)\right).$$

Note that we recover $e_0(x,y,t)$ as $a\to 0$. We know that $e_a(x,y,t)$ should be symmetric in x and y, but not necessarily a function of $y-x$, since $u_t = u_{yy}-a^2y^2u$ is not translation invariant. The simplest suitable form is

$$e_a(x,y,t) = f(t)^{-\frac{1}{2}}\exp\left(\tfrac{1}{4}g(t)\left(x^2+y^2\right)+h(t)\,xy\right).$$

We have

$$\begin{aligned}(e_a)_y(x,y,t) &= e_a(x,y,t)\left(\tfrac{1}{2}g(t)\,y+h(t)\,x\right)\\ (e_a)_{yy}(x,y,t) &= (e_a)_y(x,y,t)\left(\tfrac{1}{2}g(t)\,y+h(t)\,x\right)+E_a\tfrac{1}{2}g(t)\\ &= e_a(x,y,t)\left(\left(\tfrac{1}{2}g(t)\,y+h(t)\,x\right)^2+\tfrac{1}{2}g(t)\right),\ \text{and}\end{aligned}$$

$$\begin{aligned}&e_a^{-1}\left((e_a)_t-(e_a)_{yy}+a^2y^2e_a\right)(x,y,t)\\ &=-\tfrac{1}{2}\frac{f'(t)}{f(t)}+\tfrac{1}{4}g'(t)\left(x^2+y^2\right)+h'(t)\,xy\\ &-\left(\tfrac{1}{4}g(t)^2y^2+g(t)\,h(t)\,xy+h(t)^2x^2+\tfrac{1}{2}g(t)\right)+a^2y^2\\ &=-\tfrac{1}{2}\left(\frac{f'(t)}{f(t)}+g(t)\right)+\tfrac{1}{4}\left(g'(t)-g(t)^2+4a^2\right)y^2\\ &+\left(\tfrac{1}{4}g'(t)-h(t)^2\right)x^2+\left(h'(t)-g(t)\,h(t)\right)xy.\end{aligned}$$

Equating to zero the coefficients of this quadratic polynomial in (x,y), we get

$$\begin{aligned}g'=g^2-4a^2 &\Rightarrow g(t)=-2a\coth(2at+C)\\ h(t)=\tfrac{1}{2}\sqrt{g'(t)} &= \tfrac{1}{2}\sqrt{(2a)^2\operatorname{csch}^2(2at+C)}=\frac{a}{\sinh(2at+C)}\\ \frac{f'}{f}=-g &\Rightarrow f(t)=C'\exp\left(2a\int\coth(2at+C)\,dt\right)\\ &= C'\exp(\ln\sinh(2at+C))=C'\sinh(2at+C)\,.\end{aligned}$$

Luckily (?), we also have $h'(t)=g(t)\,h(t)$. Hence,

$$\begin{aligned}&e_a(x,y,t)\\ &=\frac{1}{\sqrt{C'\sinh(2at+C)}}\exp\left(-\frac{1}{4\frac{\sinh(2at)}{2a}}\left(\cosh(2at+C)\left(x^2+y^2\right)-2xy\right)\right)\end{aligned}$$

The behavior is correct as a and t approach 0 if and only if $C=0$ and $C'=2\pi/a$, and we have (18.107).

Now suppose that B is a real, skew-symmetric $n\times n$ matrix where $n=2m$ is even. We will find the heat kernel for

(18.108)

$$u_t(y,t)=\Delta u-|By|^2u=\Delta u+\left\langle B^2y,y\right\rangle u,\quad u(y,t)\in\mathbb{R},\ (y,t)\in\mathbb{R}^n\times(0,\infty)\,,$$

where $\Delta u=\partial_1^2u+\cdots+\partial_n^2u$. The eigenvalues of B are of the form $\pm ir_1,\ldots,\pm ir_m$, and for some $R\in\mathrm{SO}(n)$,

$$R^{-1}B^2R=-diag\left(r_1^2,r_1^2,\ldots,r_m^2,r_m^2\right).$$

Let $y':=R^{-1}y$ and $u'(y',t):=u(y,t)=u(Ry',t)$. If $u_t=\Delta u+\left\langle B^2y,y\right\rangle u$, then

$$\begin{aligned}u_t'(y',t) &= u_t(Ry',t)=\Delta u(Ry',t)+\left\langle B^2Ry',Ry'\right\rangle u(Ry',t)\\ &=\Delta'u'(y',t)+\left\langle R^{-1}B^2Ry',y'\right\rangle u'(y',t)\end{aligned}$$

(18.109)
$$=\Delta'u'(y',t)-\sum\nolimits_{j=1}^{m}r_j^2\left(y_{2j-1}'^2+y_{2j}'^2\right)u'(y',t)\,.$$

Let $v_1 = v_1(y_1, t)$ with $v_{1t} = \partial_1^2 v_1 - a_1^2 v_1$, and $v_2 = v_2(y_2, t)$ with $v_{2t} = \partial_2^2 v_2 - a_2^2 v_2$. Then for $v(y_1, y_2, t) := v_1(y_1, t)\, v_2(y_2, t)$, we have

$$\begin{aligned} v_t(y_1, y_2, t) &= v_{1t}(y_1, t)\, v_2(y_2, t) + v_1(y_1, t)\, v_{2t}(y_2, t) \\ &= \left(\partial_1^2 v_1 - a_1^2 v_1\right) v_2(y_2, t) + v_1(y_1, t)\left(\partial_2^2 v_2 - a_2^2 v_2\right) \\ &= \left(\partial_1^2 v_1(y_1, t)\, v_2(y_2, t) - a_1^2 v_1(y_1, t)\, v_2(y_2, t)\right) \\ &+ \left(v_1(y_1, t)\, \partial_2^2 v_2 - v_1(y_1, t)\, a_2^2 v_2(y_2, t)\right) \\ &= \left(\partial_1^2 + \partial_2^2\right) v(y_1, y_2, t) - \left(a_1^2 + a_2^2\right) v(y_1, y_2, t). \end{aligned}$$

Thus, the heat kernel for (18.109) is

$$e'(x', y', t) := \prod_{j=1}^{m} e_{r_j}\left(x'_{2j-1}, y'_{2j-1}, t\right) e_{r_j}\left(x'_{2j}, y'_{2j}, t\right).$$

The heat kernel for (18.108) is then

$$e_B(x, y, t) = e'(x', y', t) = e'\left(R^{-1}x, R^{-1}y, t\right), \text{ and}$$

$$e_B(0, 0, t) = e'\left(R^{-1}0, R^{-1}0, t\right) = (4\pi)^{-m} \prod_{j=1}^{m} \frac{2r_j}{\sinh(2r_j t)}. \tag{18.110}$$

Let $A^B : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ be the operator defined by

$$\begin{aligned} &A^B[f](y) := \Delta u(y) - |By|^2 f(y), \text{ and let} \\ &A_s^B[f](y) := A^B[f](sy). \end{aligned}$$

With $\mathcal{E}(r, t) := (4\pi t)^{-n/2} \exp\left(-r^2/4t\right)$, just as before (only easier) we have an asymptotic expansion (as $t \to 0^+$)

$$\begin{aligned} &\mathcal{E}(|y|, t)^{-1} e_B(0, y, t) \sim \sum\nolimits_{j=0}^{Q} h_j^B(0, y)\, t^j, \text{ where} \\ &h_j^B(0, y) = \int_{I^m} \prod\nolimits_{i=0}^{j-1} (s_i)^i \left(\left(A_{s_{j-1}}^B \circ \cdots \circ A_{s_0}^B\right)[1]\right)(y)\, ds_0 \ldots ds_{j-1}. \end{aligned}$$

In particular,

$$\mathcal{E}(0, t)^{-1} e_B(0, 0, t) \sim \sum\nolimits_{j=0}^{Q} h_j^B(0, 0)\, t^j,$$

and in view of (18.110),

$$\mathcal{E}(0, t)^{-1} e_B(0, 0, t) = (4\pi t)^{n/2} (4\pi)^{-m} \prod_{j=1}^{m} \frac{2r_j}{\sinh(2r_j t)} = \prod_{j=1}^{m} \frac{2r_j t}{\sinh(2r_j t)},$$

which is analytic at $t = 0$. Thus for t sufficiently small,

$$\prod_{j=1}^{m} \frac{2r_j t}{\sinh(2r_j t)} = \sum\nolimits_{j=0}^{\infty} h_j^B(0, 0)\, t^j$$

Since the eigenvalues of iB are $\pm r_1, \ldots, \pm r_m$, we have

$$\prod_{j=1}^{m} \frac{2r_j t}{\sinh(2r_j t)} = \det\left(\frac{2tiB}{\sinh(2tiB)}\right)^{\frac{1}{2}}.$$

Recall (see (18.41)) that

$$\prod_{k=1}^{m} \frac{r_k s/2}{\sinh(r_k s/2)} = \sum_{j=0}^{\infty} a_k\left(r_1^2,\ldots,r_m^2\right) s^{2k} = \sum_{k=0}^{\infty} \mathcal{A}_k\left(\sigma_1,\ldots,\sigma_k\right) s^{2k},$$

where the coefficient $a_k\left(r_1^2,\ldots,r_m^2\right)$ is a homogeneous, symmetric polynomial in $r_1^2,\ldots,r_m^2$ of degree k. We have written $a_k\left(r_1^2,\ldots,r_m^2\right)$ as a polynomial $\mathcal{A}_k\left(\sigma_1,\ldots,\sigma_k\right)$ in the elementary symmetric polynomials $\sigma_k = \sigma_k\left(r_1^2,\ldots,r_m^2\right)$. With $s/2 = 2t$ so that $s = 4t$,

$$\begin{aligned} \det\left(\frac{2tiB}{\sinh(2tiB)}\right)^{\frac{1}{2}} &= \prod_{j=1}^{m} \frac{2r_j t}{\sinh(2r_j t)} \\ = \sum_{k=0}^{\infty} \mathcal{A}_k\left(\sigma_1,\ldots,\sigma_k\right)(4t)^{2k} &= \sum_{k=0}^{\infty} 4^{2k}\mathcal{A}_k\left(\sigma_1,\ldots,\sigma_k\right) t^{2k}. \end{aligned}$$

Thus, $h_j^B(0,0) = 0$ for j odd, and with

$$\begin{aligned} \mathcal{A}_k(B) &:= \mathcal{A}_k\left(\sigma_1,\ldots,\sigma_k\right), \text{ where} \\ \sigma_l &:= \frac{1}{(2l)!}\sum_{(i),(j)} \delta_{i_1\cdots i_{2l}}^{j_1\cdots j_{2l}} B_{j_1}^{i_1}\cdots B_{j_{2l}}^{i_{2l}}, \end{aligned}$$

we have

$$\begin{aligned} 4^{2k}\mathcal{A}_k(B) &= h_{2k}^B(0,0) \\ &= \int_{I^{2k}} \prod\nolimits_{i=0}^{2k-1} (s_i)^i \left(A_{s_{2k-1}}^B \circ\cdots\circ A_{s_0}^B\right)[1](0)\, ds_0\ldots ds_{2k-1}. \end{aligned}$$

Let $C \in \mathbb{R}$. If u satisfies (18.108), and $v(y,t) := e^{Ct}u(y,t)$, then v satisfies

$$v_t(y,t) = \Delta v - |By|^2 v + Cv \tag{18.111}$$

Thus, the kernel for the heat equation (18.111) is

$$e_{B,C}(x,y,t) := e^{Ct}e_B(x,y,t),$$

since this solution has the correct behavior as $t \to 0^+$. With

$$\begin{aligned} A^{B,C}[f](y) &:= \Delta f(y) - |By|^2 f(y) + Cf(y), \\ A_s^{B,C}[f](y) &:= A^{B,C}[f](sy), \text{ and} \end{aligned}$$

$$\begin{aligned} \mathcal{A}_j(B,C) &:= h_j^{B,C}(0,0) \\ &:= \int_{I^j} \prod\nolimits_{i=0}^{j-1} (s_i)^i \left(A_{s_{j-1}}^{B,C} \circ\cdots\circ A_{s_0}^{B,C}\right)[1](0)\, ds_0\ldots ds_{j-1}, \end{aligned}$$

we then have

$$\begin{aligned}
&\sum_{l=0}^{\infty} h_l^{B,C}(0,0)\, t^l = \mathcal{E}(|y|,t)^{-1} e_{B,C}(0,0,t) \\
&= e^{Ct} \prod_{j=1}^{m} \frac{2r_j t}{\sinh(2r_j t)} = e^{Ct} \sum_{j=0}^{\infty} h_j^B(0,0)\, t^j = e^{Ct} \sum_{k=0}^{\infty} 4^{2k} \mathcal{A}_k(B)\, t^{2k} \\
&= \sum_{j=0}^{\infty} \frac{C^j t^j}{j!} \sum_{k=0}^{\infty} \mathcal{A}_k(B)\, t^{2k} = \sum_{j,k=0}^{\infty} \frac{1}{j!} C^j 4^{2k} \mathcal{A}_k(B)\, t^{j+2k} \\
&= \sum_{l=0}^{\infty} \left(\sum_{j+2k=l} \frac{1}{j!} C^j 4^{2k} \mathcal{A}_k(B) \right) t^l, \text{ or}
\end{aligned}$$

(18.112)

$$\int_{I^l} \prod_{i=0}^{l-1} (s_i)^i \left(A_{s_{l-1}}^{B,C} \circ \cdots \circ A_{s_0}^{B,C} \right)[1](0)\ ds_0 \ldots ds_{l-1} = \sum_{j+2k=l} \frac{1}{j!} C^j 4^{2k} \mathcal{A}_k(B)\,.$$

Observe that both sides are ultimately homogeneous polynomials of degree l in the "variables" C and B_{ij}. Suppose that we replace these variables by operators $\widehat{C}$ and $\widehat{B}_{ij}$ in $\operatorname{End}(W)$ for a finite dimensional vector space W, specifically

$$\begin{aligned}
W &:= \mathbb{C}^N \otimes \Sigma_{2m} \\
\widehat{C} &:= -\tfrac{1}{2} \sum_{p,q} \Omega_{pq}^{\varepsilon}(0) \otimes \gamma^p \gamma^q, \text{ and} \\
\widehat{B}_{pj} &:= \tfrac{1}{8} \sum_{k,l} R_{klpj}(0) \left(\operatorname{Id}_{\mathbb{C}^N} \otimes i \gamma^k \gamma^l \right).
\end{aligned}$$

The replacement of B by $\widehat{B}$ and C by $\widehat{C}$ in $A^{B,C}$ gives us an operator $A^{\widehat{B},\widehat{C}}$: $C^\infty(\mathbb{R}^n, W) \hookleftarrow$, namely

$$\begin{aligned}
&A^{\widehat{B},\widehat{C}}(\psi) := \Delta\psi - \sum_{p,j,j'} \widehat{B}_{pj} y^j \widehat{B}_{pj'} y^{j'} \psi + \widehat{C}\psi \\
&= \sum_i \partial_i^2 + \sum_p \left(\tfrac{1}{8} \sum_{k,l,j} R_{klpj}(0) \gamma^k \gamma^l y^j \right) \left(\tfrac{1}{8} \sum_{k',l',j'} R_{k'l'pj'}(0) \gamma^{k'} \gamma^{l'} y^{j'} \right) \\
&- \tfrac{1}{2} \sum_{i,j} \Omega_{ij}^{\varepsilon}(0) \otimes \gamma^i \gamma^j,
\end{aligned}$$

which is the same as the operator A^0 in (18.104); note the presence of the factor of $i = \sqrt{-1}$ in the definition of $\widehat{B}_{pj}$ which accounts for a crucial sign change in order that $A^{\widehat{B},\widehat{C}} = A^0$. The right side of (18.112) still makes sense as an element of $\operatorname{End}(W)$ if multiplication is taken to be composition. But, since generally operators do not commute, one could get a different result unless some ordering in compositions of the factors $\widehat{C}$ and $\widehat{B}_{ij}$ in monomials is specified. With operator replacement, the left side may also be interpreted as in $\operatorname{End}(W)$, if for a given $w \in W$, we define

$$h_l^{\widehat{B},\widehat{C}}(0,0) := \left(\int_{I^l} \prod_{i=0}^{l-1} (s_i)^i \left(A_{s_{l-1}}^{\widehat{B},\widehat{C}} \circ \cdots \circ A_{s_0}^{\widehat{B},\widehat{C}} \right)[1](0)\, ds_0 \ldots ds_{l-1} \right)(w)$$

(18.113)

$$:= \left(\int_{I^l} \prod_{i=0}^{l-1} (s_i)^i \left(A_{s_{l-1}}^{\widehat{B},\widehat{C}} \circ \cdots \circ A_{s_0}^{\widehat{B},\widehat{C}} \right)[\widetilde{w}](0)\, ds_0 \ldots ds_{l-1} \right),$$

where $\widetilde{w}$ is the constant function in $C^\infty(\mathbb{R}^n, W)$ with value w; note that the 1 in [1] of (18.113) may be regarded as the map $w \mapsto \widetilde{w}$. There is already a definite ordering of the factors in the various terms of the expansion of

$$\int_{I^l} \prod_{i=0}^{l-1} (s_i)^i \left(A_{s_{l-1}}^{\widehat{B},\widehat{C}} \circ \cdots \circ A_{s_0}^{\widehat{B},\widehat{C}}\right)[1](0)\, ds_0 \ldots ds_{l-1}.$$

Since we know that the two sides of (18.112) agree as polynomials in commuting variables C and B_{ij}, if we replace these variables by operators, there is a corresponding reordering of operators within the terms of the right side so that both sides are the same element of $\mathrm{End}\,(W)$. In other words,

$$\begin{aligned} &\int_{I^l} \prod_{i=0}^{l-1} (s_i)^i \left(A_{s_{l-1}}^{\widehat{B},\widehat{C}} \circ \cdots \circ A_{s_0}^{\widehat{B},\widehat{C}}\right)[1](0)\ \ ds_0 \ldots ds_{l-1} \\ &= \sum_{j+2k=l} \frac{4^{2k}}{j!} \mathcal{R}(\widehat{C}^j \mathcal{A}_k(\widehat{B})), \end{aligned} \tag{18.114}$$

where $\mathcal{R}(\widehat{C}^j \mathcal{A}_k(\widehat{B}))$ is a reordering of the operators $\widehat{C}$ and $\widehat{B}_{ij}$ in the monomial terms of $\widehat{C}^j \mathcal{A}_k(\widehat{B})$ so that (18.114) holds. Now,

$$\begin{aligned} \widehat{C} \circ \widehat{B}_{pj} &= \left(-\tfrac{1}{2}\sum_{p,r} \Omega_{qr}^{\varepsilon}(0) \otimes \gamma^q \gamma^r\right) \circ \left(\tfrac{1}{8}\sum_{k,l} R_{klpj}(0) \left(\mathrm{Id}_{\mathbb{C}^N} \otimes i\gamma^k\gamma^l\right)\right) \\ &= -\tfrac{1}{16}\sum_{q,r,k,l} R_{klpj}(0)\ \ \Omega_{qr}^{\varepsilon}(0) \otimes i\gamma^q\gamma^r\gamma^k\gamma^l \\ &= -\tfrac{1}{16}\sum_{q,r,k,l} R_{klpj}(0)\ \ \Omega_{qr}^{\varepsilon}(0) \otimes i\gamma^k\gamma^l\gamma^q\gamma^r + \mathrm{O}_c(2) \\ &= \widehat{B}_{pj} \circ \widehat{C} + \mathrm{O}_c(2). \end{aligned}$$

Thus, for $j + 2k = m = n/2$,

$$Str\left(\mathcal{R}(\widehat{C}^j \mathcal{A}_k(\widehat{B}))\right) = Str\left(\widehat{C}^j \mathcal{R}(\mathcal{A}_k(\widehat{B}))\right),$$

where $\mathcal{R}(\mathcal{A}_k(\widehat{B}))$ is $\mathcal{A}_k(\widehat{B})$ with a possible reordering within the monimial terms of $\mathcal{A}_k(\widehat{B})$. However, these monomial terms are products of the operators

$$\widehat{B}_{pj} = \tfrac{1}{8}\sum_{k,l} R_{klpj}(0)\, i\gamma^k\gamma^l,$$

and since the $R_{klij}(0)$ are just scalars, such operators commute modulo terms of $\deg_c 2$. Hence, for $j + 2k = m = n/2$,

$$\mathrm{Str}\left(\mathcal{R}(\widehat{C}^j \mathcal{A}_k(\widehat{B}))\right) = \mathrm{Str}\left(\widehat{C}^j \mathcal{R}(\mathcal{A}_k(\widehat{B}))\right) = \mathrm{Str}\left(\widehat{C}^j \mathcal{A}_k(\widehat{B}))\right).$$

Thus, while a reordering $\mathcal{R}$ is needed to make (18.114) correct, when $l = m = n/2$, we have

$$\begin{aligned} \mathrm{Str}\left(h_m^{\widehat{B},\widehat{C}}(0,0)\right) &= \mathrm{Str}\left(\sum_{j+2k=m} \frac{4^{2k}}{j!} \mathcal{R}(\widehat{C}^j \mathcal{A}_k(\widehat{B}))\right) \\ &= \mathrm{Str}\left(\sum_{j+2k=m} \frac{4^{2k}}{j!} \widehat{C}^j \mathcal{A}_k(\widehat{B})\right). \end{aligned} \tag{18.115}$$

With $\nu := dy^1 \wedge \cdots \wedge dy^n$ and $n = 2m$,

$$\begin{aligned} \mathrm{Str}\left(\gamma^{j_1}\gamma^{j_2}\cdots\gamma^{j_{2k-1}}\gamma^{j_{2k}}\right) &= (-2i)^m \left\langle \left(dy^{j_1} \wedge dy^{j_2}\right) \wedge \cdots \wedge \left(dy^{j_{2k-1}} \wedge dy^{j_{2k}}\right), \nu \right\rangle \\ &= \left\langle \left(-2i dy^{j_1} \wedge dy^{j_2}\right) \wedge \cdots \wedge \left(-2i dy^{j_{2k-1}} \wedge dy^{j_{2k}}\right), \nu \right\rangle, \end{aligned}$$

where both sides are 0 if $k < m$. Thus, in (18.115) we may replace

$$\begin{aligned}
\widehat{C} &= -\tfrac{1}{2}\sum\nolimits_{p,q}\Omega^{\varepsilon}_{pq}\otimes\gamma^p\gamma^q \text{ by}\\
2i\Omega^{\varepsilon} &= -\tfrac{1}{2}\sum\nolimits_{p,q}\Omega^{\varepsilon}_{pq}\left(-2i dy^p\wedge dy^q\right),\ \text{and}\\
\widehat{B}_{pj} &= \tfrac{1}{8}\sum\nolimits_{k,l}R_{klpj}i\gamma^k\gamma^l \text{ by}\\
\tfrac{1}{2}\Omega^{\theta}_{pj} &= \tfrac{1}{8}\sum\nolimits_{k,l}R_{klpj}i\left(-2i dy^k\wedge dy^l\right)
\end{aligned}$$

provided we take the inner product, with ν, of the trace (in End $\left(\mathbb{C}^N\right)$) of the result. The following computation then completes the proof of the Local Index Theorem (i.e., Theorem 18.45, p. 544):

$$\begin{aligned}
&(4\pi)^{-m}\operatorname{Str}\left(h_m(0,0)\right) = (4\pi)^{-m}\operatorname{Str}(h_m^{\widehat{B},\widehat{C}}(0,0))\\
&= \operatorname{Str}\left((4\pi)^{-m}\sum\nolimits_{j+2k=m}\frac{1}{j!}(\widehat{C})^j 4^{2k}\mathcal{A}_k(\widehat{B})\right)\\
&= \left\langle\sum\nolimits_{j+2k=m}\frac{1}{j!}\wedge^j(\frac{2i}{4\pi}\Omega^{\varepsilon})\wedge\frac{4^{2k}}{(4\pi)^{2k}}\mathcal{A}_k(\tfrac{1}{2}\Omega^{\theta}),\nu\right\rangle\\
&= \left\langle\sum\nolimits_{j=0}^{\infty}\frac{1}{j!}\operatorname{Tr}\left(\wedge^j(\tfrac{i}{2\pi}\Omega^{\varepsilon}\right)\wedge\sum\nolimits_{k=0}^{\infty}\mathcal{A}_k(\tfrac{1}{2\pi}\Omega^{\theta}),\nu\right\rangle\\
&= \left\langle\operatorname{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right)\wedge\det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh\left(i\Omega^{\theta}/4\pi\right)}\right)^{\frac{1}{2}},\nu\right\rangle.
\end{aligned}$$

As a consequence of the *Local* Index Theorem for twisted Dirac Operator we have the "integrated version" itself, namely

THEOREM 18.49. *(Index Theorem for Twisted Dirac Operators) Let M be a compact, oriented Riemannian $2m$-manifold with a spin structure, and let E be a Hermitian vector bundle over M with unitary connection $\varepsilon\in\mathcal{C}(U(E))$. Then the index of the Dirac operator*

$$\mathcal{D}^+ : C^{\infty}\left(E\otimes\Sigma^+(M)\right)\to C^{\infty}\left(E\otimes\Sigma^-(M)\right)$$

is given by

$$\operatorname{index}\mathcal{D}^+ = \left(\mathbf{ch}(E)\smile\widehat{\mathbf{A}}(M)\right)[M],$$

where $\mathbf{ch}(E)\in H^*(M,\mathbb{Q})$ is the Chern character of E (see (16.97) or (16.110)) and $\widehat{\mathbf{A}}(M)\in H^*(M,\mathbb{Q})$ is the total $\widehat{\mathbf{A}}$ class of M (see (16.107) with $F = TM$ and $\widehat{\mathbf{A}}(M) := \widehat{\mathbf{A}}(TM)$ or the equivalent formulation (18.43)).

PROOF. We have implicitly proven this already in the discussion preceeding (18.44) on p. 515, but there we *assumed* the Local Index Theorem (p. 544). As this was done many pages ago, it is fitting to recall the computation, especially now that we have finally established the Local Index Theorem:

$$\begin{aligned}\text{index}\left(\mathcal{D}^{+}\right) &= \int_M \left(\operatorname{Tr}\left(k^{+}(x,x,t)\right) - \operatorname{Tr}\left(k^{-}(x,x,t)\right)\right)\ \nu_x \text{ by (18.37)}\\ &= \int_M \operatorname{Str}\left(k(x,x,t)\right)\ \nu_x = \int_M \lim_{t\to 0^+} \operatorname{Str}\left(k(x,x,t)\right)\ \nu_x\\ &= \int_M \langle \operatorname{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right) \wedge \det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh\left(i\Omega^{\theta}/4\pi\right)}\right)^{\frac{1}{2}}, \nu_x\rangle\, \nu_x \text{ by Theorem 18.45}\\ &= \left(\mathbf{ch}(E) \smile \widehat{\mathbf{A}}(M)\right)[M],\end{aligned}$$

where the final equality was explained on p. 515. □

Corollary 18.50. *Let M be a compact, oriented Riemannian $2m$-manifold with a spin structure and positive scalar curvature, then* $\widehat{A}(M) := \widehat{\mathbf{A}}(M)[M] = 0$.

Proof. Take $E = 0$. By Corollary 18.26, p. 506, $\operatorname{Ker}\mathcal{D}^{+} = \operatorname{Ker}\mathcal{D}^{-} = \{0\}$ so that $\widehat{A}(M) = \operatorname{index}\mathcal{D}^{+} = 0$. □

As an application, especially relevant to the study intersection forms of smooth 4-manifolds and Seiberg-Witten theory, we have

Theorem 18.51. *If M is a compact, orientable C^{∞} $4k$-manifold with a spin structure, then $\widehat{A}(M) := \widehat{\mathbf{A}}(M)[M]$ is an integer. Moreover, if k is odd, then $\widehat{A}(M)$ is even.*

Proof. By Theorem 18.49 with E trivial, $\widehat{A}(M) = \widehat{\mathbf{A}}(M)[M] = \operatorname{index}\mathcal{D}^{+}$ and so $\widehat{A}(M)$ is an integer. If $k = 2a+1$ is odd, then according to the table (18.2) and the periodicity $Cl_{n+8} \cong Cl_8 \otimes Cl_n = \mathbb{R}(16) \otimes Cl_n$, for the *real* Cl_{4k} we have

$$Cl_{4k} \cong Cl_{8a+4} \cong \left(\otimes^{a}\mathbb{R}(16)\right) \otimes Cl_4 \cong \mathbb{R}\left(2^{4a}\right) \otimes \mathbb{H}(2) \cong \mathbb{H}\left(2^{4a+1}\right).$$

Thus, Cl_{4k} can be identified with the algebra of $2^{4a+1} \times 2^{4a+1}$ matrices with entrees in $\mathbb{H}$ (the quaternions) which act via $\mathbb{R}$-linear transformations to the left on $\mathbb{H}^{2^{4a+1}}$. Of course, there is *right* action of $\mathbb{H}$ on $\mathbb{H}^{2^{4a+1}}$ via $\mathbb{H}$-scalar multiplication, and this right action commutes with the left action of $\mathbb{H}\left(2^{4a+1}\right)$ on $\mathbb{H}^{2^{4a+1}}$. If we define a complex on structure on $\mathbb{H}^{2^{4a+1}}$ via right multiplication by a pure imaginary unit quaternion (e.g., $-\mathbf{k}$), then $\mathbb{H}^{2^{4a+1}}$ becomes a complex vector space of dimension 2^{4a+2}. The left action of $\mathbb{H}\left(2^{4a+1}\right)$ on $\mathbb{H}^{2^{4a+1}}$ is $\mathbb{C}$-linear, and the complexified $\mathbb{C}l_{4k}$ is the full algebra ($\cong \mathbb{C}\left(2^{4a+2}\right)$) of $\mathbb{C}$-linear transformations of $\mathbb{H}^{2^{4a+1}}$. The complex spinor space Σ_{4k} is then $\mathbb{H}^{2^{4a+1}}$with its complex structure. We still have the $\mathbb{R}$-linear right action of $\mathbb{H}$ on $\mathbb{H}^{4a+1}$, and although this not $\mathbb{C}$-linear, this right action commutes with left multiplication by the volume element $\omega_{\mathbb{C}} = i^{2k} e_1 \cdots e_{4k} = (-1)^{k} e_1 \cdots e_{4k}$, since $\omega_{\mathbb{C}}$ is in the *real* $Cl_{4k} = \mathbb{H}\left(2^{4a+1}\right)$. Thus, the spinor spaces $\Sigma_{4k}^{\pm} := (1 \pm \omega_{\mathbb{C}})\,\Sigma_{4k}$ are invariant under the $\mathbb{R}$-linear right action of $\mathbb{H}$ on $\Sigma_{4k} \cong \mathbb{H}^{2^{4a+1}}$. In other words $\Sigma_{4k}^{\pm}$ are right $\mathbb{H}$-modules over $\mathbb{R}$. The spinor bundles $\Sigma_{4k}^{\pm}(M) := P \times_{\operatorname{Spin}(4k)} \Sigma_{4k}^{\pm}$ are also right $\mathbb{H}$-modules since the right $\mathbb{H}$-action on $\Sigma_{4k}^{\pm}$ commutes with the spinor representations $\rho^{\pm} : \operatorname{Spin}(4k) \to \operatorname{SU}\left(\Sigma_{4k}^{\pm}\right)$ which are just restrictions of Clifford multiplication to $\mathrm{Spin}(4k) \subset Cl_{4k}$. For the same reason, the $\mathbb{H}$-action on $\Sigma_{4k}^{\pm}(M)$ commutes with

$$\text{Clifford multiplication } cl : \Omega^1(M) \otimes C^\infty(\Sigma(M)) \to C^\infty(\Sigma(M)),$$

$$\text{covariant differentiation } \nabla^\Sigma : C^\infty(\Sigma(M)) \to \Omega^1(M) \otimes C^\infty(\Sigma(M)).$$

and hence with the Dirac operator $\mathcal{D} = cl \circ \nabla^\Sigma$. Thus, the spaces $\mathrm{Ker}\left(\mathcal{D}^\pm : C^\infty(\Sigma^\pm(M)) \to C^\infty(\Sigma^\mp(M))\right)$ are right $\mathbb{H}$-modules. The real dimension of any $\mathbb{H}$-module V is a multiple of 4 (see Exercise 18.53 below). Hence,

$$\widehat{A}(M) = \widehat{\mathbf{A}}(M)[M] = \mathrm{index}\,\mathcal{D}^+ = \dim_{\mathbb{C}} \mathrm{Ker}\,\mathcal{D}^+ - \dim_{\mathbb{C}} \mathrm{Ker}\,\mathcal{D}^- \text{ is even.}$$

□

Corollary 18.52. *(Rochlin's Theorem)The signature $\tau(M)$ of a compact, orientable C^∞ 4-manifold M with a spin structure is a multiple of 16.*

Proof. The Hirzebruch Signature Theorem (to be proven independently in Section 18.6; see Theorem 18.59, p. 573) states that $\tau(M) = \frac{1}{3}p_1(M)$. Since $\widehat{A}(M) = \frac{-1}{24}p_1(M)$ by 16.108, we obtain $\tau(M) = -8\widehat{A}(M)$, which is a multiple of 16 by Theorem 18.51. □

Exercise 18.53. Let $\phi : \mathbb{H} \to \mathrm{End}_{\mathbb{R}}(V)$ be a real right representation (i.e., $\phi(qq') = \phi(q')\phi(q)$) with $\dim V < \infty$. By completing parts (a),(b) and (c), show that $V = V_1 \oplus \cdots \oplus V_h$ where the V_i are invariant subspaces on which ϕ is equivalent to the right action of $\mathbb{H}$ on itself.
(a) Let $r : \mathbb{H} \to \mathrm{End}_{\mathbb{R}}(\mathbb{H})$ be given by $r(q)(q') = q'q$. Show that r is irreducible;i.e., $r(W) \subseteq W$ for some subspace $W \subseteq \mathbb{H}$ implies that W is $\{0\}$ or $\mathbb{H}$. [Hint. $\mathbb{H}$ is a division algebra.]
(b) For any nonzero $v \in V$, show that $\phi_v : \mathbb{H} \to \mathrm{End}(\phi(\mathbb{H})v)$ given by $\phi_v(q) = \phi(q)|_{\phi(\mathbb{H})v}$ is equivalent to r. [Hint. For $f_v : \mathbb{H} \to \phi(\mathbb{H})v$ given by $f_v(q) = \phi(q)v$, show first that $\mathrm{Ker}\, f_v$ is r-invariant.]
(c) Show that $V = \phi(\mathbb{H})v_1 \oplus \cdots \oplus \phi(\mathbb{H})v_h$ for some $v_1, \ldots, v_h \in V$.

6. The Index Theorem for Standard Geometric Operators

Our goal here is to obtain index formulas for the standard elliptic geometric operators and their twists. The standard elliptic geometric operators include

$$\text{the signature operator } d + \delta : (1 + *)\Omega^*(M) \to (1 - *)\Omega^*(M),$$

$$\text{the Euler-Dirac operator } d + \delta : \Omega^{even}(M) \to \Omega^{odd}(M), \text{ and}$$

$$\text{the Dolbeult-Dirac operator } \sqrt{2}\left(\bar{\partial} + \bar{\partial}^*\right) : \Omega^{0,even}(M) \to \Omega^{0,odd}(M).$$

Here, d denotes exterior derivative, δ denotes the exterior coderivative (i.e., the formal adjoint d^* of d), $*$ denotes the Hodge star operator, and $\bar{\partial}$ and its adjoint $\bar{\partial}^*$ denotes the complex analogs of d and d^* on complex manifolds, which, along with $\Omega^{0,*}(M)$, will be defined. The index formula obtained for the above operators yields the Hirzebruch Signature Theorem, the Chern-Gauss-Bonnet Theorem, and the Hirzebruch-Riemann-Roch Theorem, respectively. While these operators generally are not globally twisted Dirac operators, locally they expressible in these terms. Thus, even if the underlying Riemannian manifold M (still assumed to be oriented and and of even dimension $n = 2m$) does not admit a spin structure, we may still use the Local Index Theorem for twisted Dirac operators to compute the index density and hence the index of these operators. While it is possible to carry this out separately for each of the geometric operators, basically all of these theorems are consequences of an index theorem for generalized Dirac operators on Clifford module bundles (defined below). Using the Local Index Theorem for twisted Dirac operators, we prove this index theorem first, and then we apply it to obtain the geometric index theorems.

For a compact, oriented manifold M of even dimension $n = 2m$ with Riemannian metric h, let $\mathbb{C}l\,(T_xM)$ be the complexified Clifford algebra of T_xM with inner product h_x. Then $\mathbb{C}l\,(TM) := \cup_{x\in M}\mathbb{C}l\,(T_xM)$ is the total space of the so-called **complex Clifford bundle** $\mathbb{C}l\,(TM) \to M$. This bundle of algebras is defined whether M admits a spin structure or not. As a complex vector bundle, it is canonically isomorphic to $\Lambda^*(M,\mathbb{C})$, but the algebra structure is different. We can also describe $\mathbb{C}l\,(TM)$ as follows. Let $c : \mathrm{Spin}\,(n) \to \mathrm{SO}\,(n)$ be the covering given by $c\,(g)\,(v) = gvg^{-1}$, where $v \in \mathbb{R}^n$ and we identify $\mathbb{R}^n$ with $\mathcal{L}\left(\Lambda^1(\mathbb{R}^n)\right) \subset \mathbb{C}l_n$. Define the representation

$$\begin{aligned}&r : \mathrm{SO}\,(n) \to \mathrm{End}\,(\mathbb{C}l_n) \quad \text{by}\\ &r\,(c\,(g))\,(\alpha) = g\alpha g^{-1} \text{ for } g \in \mathrm{Spin}\,(n)\,.\end{aligned}$$

This is simply the extension to $\mathbb{C}l_n$ of the defining representation of $\mathrm{SO}\,(n)$ on $\mathbb{R}^n \cong \mathcal{L}\left(\Lambda^1(\mathbb{R}^n)\right) \subset \mathbb{C}l_n$. Relative to r, $\mathbb{C}l\,(TM)$ is then the associated bundle $FM \times_{\mathrm{SO}(n)} \mathbb{C}l_n$. Note that Clifford multiplication is $\mathrm{SO}\,(n)$-equivariant, in the sense that for $A \in \mathrm{SO}\,(n)$, $r\,(A) \in \mathrm{End}\,(\mathbb{C}l_n)$ is an algebra automorphism of $\mathbb{C}l_n$. Thus, regarding $\mathbb{C}l\,(TM)$ as $FM \times_{\mathrm{SO}(n)} \mathbb{C}l_n$, a well-defined algebra structure on each fiber is given by $[p,\alpha_1]\,[p,\alpha_2] := [p,\alpha_1\alpha_2]$, since

$$\left[pA, r\left(A^{-1}\right)(\alpha_1)\, r\left(A^{-1}\right)(\alpha_2)\right] = \left[pA, r\left(A^{-1}\right)(\alpha_1\alpha_2)\right] = [p,\alpha_1\alpha_2]$$

Let $Q : \mathbb{C}l_n \to \mathrm{End}\,(W)$ be an algebra representation where W is a complex vector space with $\dim(W) < \infty$. There is a Hermitian inner product $\langle\cdot,\cdot\rangle$ on W, such that for any unit vector $v \in S^n \subset \mathbb{R}^n$, we have $Q\,(v) \in U\,(W)$:=unitary group for W with $\langle\cdot,\cdot\rangle$. Indeed, for any Hermitian inner product $\langle\cdot,\cdot\rangle'$ on W, let

$$\langle w_1, w_2\rangle := \int_{S^n} \langle Q\,(v)\,w_1, Q\,(v)\,w_2\rangle'\; dv$$

Using $Q\,(v)^2 = -I$ for $v \in S^n$, we deduce that $Q\,(v)^* = -Q\,(v)$ from

$$\langle Q\,(v)\,w_1, w_2\rangle = \left\langle Q\,(v)^2\,w_1, Q\,(v)\,w_2\right\rangle = \langle -w_1, Q\,(v)\,w_2\rangle\,.$$

Definition 18.54. Let $q : \mathrm{Spin}\,(n) \to U\,(W)$ be a representation such that

$$q\,(g)\circ Q\,(\alpha)\circ q\left(g^{-1}\right) = Q\,(r\,(c\,(g))\,\alpha) \text{ for all } g \in \mathrm{Spin}\,(n) \text{ and } \alpha \in \mathbb{C}l_n. \tag{18.116}$$

If M admits a spin structure $C : P \to FM$, the representation q provides us with an associated bundle

$$W\,(M) := P \times_{\mathrm{Spin}(n)} W.$$

In the case $q : \mathrm{Spin}\,(n) \to U\,(W)$ is of the form $q_0 \circ c$, where $c : \mathrm{Spin}\,(n) \to \mathrm{SO}\,(n)$ is the double cover and $q_0 : \mathrm{SO}\,(n) \to U\,(W)$ is a representation, then we may form a bundle

$$W(M) := FM \times_{\mathrm{SO}(n)} W,$$

whether M admits a spin structure or not. In either case, we call a $W(M)$ **Clifford module bundle**.

We consider some examples. Note that $\mathbb{C}l(TM) \cong \Lambda^*(M, \mathbb{C})$ is a Clifford module bundle. Indeed, let $W = \mathbb{C}l_n$ and $Q : \mathbb{C}l_n \to \mathrm{End}(\mathbb{C}l_n)$ be left multiplication of $\mathbb{C}l_n$ on itself, and let $q : \mathrm{Spin}(n) \to \mathrm{U}(\mathbb{C}l_n)$ be given by $q(g)(\beta) = g\beta g^{-1}$. Then as required

$$\begin{aligned}&\left(q(g) \circ Q(\alpha) \circ q\left(g^{-1}\right)\right)(\beta) = g\left(\alpha\left(g^{-1}\beta g\right)\right)g^{-1}\\ &= g\alpha g^{-1}\beta = Q\left(g\alpha g^{-1}\right)\beta = Q(r(c(g))\alpha)(\beta).\end{aligned}$$

If M admits a spin structure, the primary example is the spin bundle $\Sigma(M)$. Here $W = \Sigma_n$, and $Q = \rho_{\mathbb{C}} : \mathbb{C}l_n \to \mathrm{End}(\Sigma_n)$ is the unique irreducible representation, which restricts to $q = \rho : \mathrm{Spin}(n) \to \mathrm{End}(\Sigma_n)$. Then

$$\begin{aligned}&\left(q(g) \circ Q(\alpha) \circ q\left(g^{-1}\right)\right)(\beta) = \rho(g) \circ \rho_{\mathbb{C}}(\alpha) \circ \rho\left(g^{-1}\right)(\beta)\\ &= \rho_{\mathbb{C}}\left(g\alpha g^{-1}\right)(\beta) = \rho_{\mathbb{C}}(r(c(g))\alpha)(\beta) = Q(r(c(g))\alpha)(\beta).\end{aligned}$$

Note that when M admits a spin structure, $P \to FM$, one can also consider the bundle $P \times_{\mathrm{Spin}(n)} W$ relative to the representation $Q|_{\mathrm{Spin}(n)} : \mathrm{Spin}(n) \to U(W)$, but since it may not be the case that $q = Q|_{\mathrm{Spin}(n)}$, this bundle is *not* necessarily isomorphic to $W(M)$ obtained from q. For example when $Q : \mathbb{C}l_n \to \mathrm{End}(\mathbb{C}l_n)$ is left muliplication, the bundle $P \times_{\mathrm{Spin}(n)} W$ relative to $Q|_{\mathrm{Spin}(n)}$ is not $\mathbb{C}l(TM)$, but rather it is designated by $\mathbb{C}l_{\mathrm{Spin}}(TM)$ which is actually isomorphic to a direct sum of 2^m copies of $\Sigma(M)$, $m = \frac{1}{2}\dim M$.

To show that a Clifford module bundle $W(M)$ is indeed a bundle of $\mathbb{C}l(TM)_x$-modules, we let

$$\begin{aligned}&cl : \mathbb{C}l(TM) \otimes W(M) \to W(M) \text{ be defined via}\\ &cl([p,\alpha] \otimes [p,w]) := [p, Q(\alpha)w].\end{aligned}$$

which is a well-defined "Clifford multiplication". Indeed, using (18.116),

$$\begin{aligned}&cl\left(\left[pA, r(A)^{-1}\alpha\right] \otimes \left[pA, q(A)^{-1}w\right]\right) = \left[pA, Q\left(r\left(A^{-1}\right)\alpha\right)\left(q(A)^{-1}w\right)\right]\\ &= \left[pA, q\left(A^{-1}\right)Q(\alpha)q(A)q(A)^{-1}(w)\right] = \left[pA, q\left(A^{-1}\right)Q(\alpha)(w)\right]\\ &= [p, Q(\alpha)(w)] = cl([p,\alpha] \otimes [p,w]).\end{aligned}$$

Moreover, for $\alpha \in \mathcal{L}\left(\Lambda^1(\mathbb{R}^n)\right) \subset \mathbb{C}l_n$, we have

$$cl([p,\alpha] \otimes cl([p,\alpha] \otimes [p,w])) = \left[p, Q(\alpha)^2 w\right] = -|[p,\alpha]|^2[p,w],$$

so that cl does in fact make $W(M)_x$ a $\mathbb{C}l(TM)_x$-module. As in the special case of $\Sigma(M)$, there is a decomposition $W(M) = W^+(M) \oplus W^-(M)$ defined as follows. If $E_1, \ldots, E_n$ is a positively oriented orthonormal frame at $x \in M$, let

$$\omega_{\mathbb{C}}(x) := i^m E_1 \cdots E_n \in \mathbb{C}l(TM)_x. \tag{18.117}$$

It is easy to check that this is independent of the choice of frame and $\omega_{\mathbb{C}}(x)^2 = 1$ (see (18.12), p. 489). Then

$$\mu_{\mathbb{C}}(x) := cl(\omega_{\mathbb{C}}(x)) \in \operatorname{End}(W(M)_x) \text{ and } \mu_{\mathbb{C}}(x)^2 = \mathrm{Id}, \text{ and}$$
$$W^{\pm}(M) := (I \pm \mu_{\mathbb{C}})W(M) \Rightarrow W(M) = W^{+}(M) \oplus W^{-}(M). \tag{18.118}$$

For any $0 \neq X \in T_xM$, we may take $E_1 = X/|X|$. Since $cl(E_1)$ is an isomorphism of $W_x(M)$ which anticommutes with $\mu_{\mathbb{C}}(x)$, we have

$$cl(X) = |X|\, cl(E_1) : W^{\pm}(M) \cong W^{\mp}(M). \tag{18.119}$$

Since $W(M)$ is an associated bundle of FM or the Spin (n)-bundle P of a spin structure $P \to FM$, the Levi-Civita connection θ on FM (or its lift to P) yields a covariant derivative operator

$$\nabla^W : C^\infty(W(M)) \to \Omega^1(W(M)) \cong C^\infty\left(\Omega^1(M) \otimes W(M)\right).$$

Moreover, $cl : \mathbb{C}l(TM) \otimes W(M) \to W(M)$ induces map on sections,

$$cl : C^\infty(\mathbb{C}l(TM)) \otimes C^\infty(W(M)) \to C^\infty(W(M)) \text{ (same notation).}$$

For $\alpha \in C^\infty(\mathbb{C}l(TM))$ and $\psi \in C^\infty(W(M))$, we have

$$\nabla^W(cl(\alpha \otimes \psi)) = cl(\nabla\alpha \otimes \psi) + cl\left(\alpha \otimes \nabla^W\psi\right), \tag{18.120}$$

where ∇ is the usual Levi-Civita covariant derivative on $C^\infty(\mathbb{C}l(TM)) \cong \Omega^*(M, \mathbb{C})$. Indeed, using the infinitesimal verison of (18.116), namely $q'(A)\,Q(\alpha) - Q(\alpha)\,q(A') = Q(r'(A)\,\alpha)$, it follows that on FM where $\alpha \otimes \psi \in \overline{\Omega}^0(\mathbb{C}l_{2m} \otimes W) :=$ equivariant $\mathbb{C}l_{2m} \otimes W$-valued functions on FM,

$$\begin{aligned}
&D^\theta(Q(\alpha \otimes \psi)) = d(Q(\alpha \otimes \psi)) + q'(\theta)\,Q(\alpha \otimes \psi)\\
&= Q(d\alpha \otimes \psi) + Q(\alpha \otimes d\psi) + Q((r'(\theta)\,\alpha) \otimes \psi + \alpha \otimes q'(\theta)\,\psi)\\
&= Q(d\alpha \otimes \psi) + Q(r'(\theta)(\alpha) \otimes \psi) + Q(\alpha \otimes d\psi) + Q(\alpha \otimes q'(\theta)\,\psi)\\
&= Q\left(D^\theta(\alpha) \otimes \psi\right) + Q\left(\alpha \otimes D^\theta\psi\right).
\end{aligned}$$

Via the Riemannian metric, $\Lambda^1(T_xM^*) \cong \Lambda^1(T_xM) \subset \mathbb{C}l(TM)$, and so

$$\Omega^1(M) \otimes C^\infty(W(M)) \subset C^\infty(\mathbb{C}l(TM)) \otimes C^\infty(W(M))$$

Thus, we have an operator $\mathcal{D}^W := cl \circ \nabla^W \in \operatorname{End}(C^\infty(W(M)))$,

$$\mathcal{D}^W : C^\infty(W(M)) \overset{\nabla^W}{\to} \Omega^1(M) \otimes C^\infty(W(M)) \overset{cl}{\to} C^\infty(W(M)),$$

which we call a **generalized Dirac operator**. In view of (18.119), we also have

$$\mathcal{D}^{W\pm} : C^\infty\left(W^{\pm}(M)\right) \to C^\infty\left(W^{\mp}(M)\right),$$

as with Dirac operators. We now show that $\mathcal{D}^W$ is locally a twisted Dirac operator. Indeed, we show that $\mathcal{D}^W$ is globally a twisted Dirac operator when M admits a spin structure $P \to FM$, from which the local statement follows since locally M admits spin structures. Recall (see Proposition 18.14, p. 492) that for any $\mathbb{C}l_{2m}$-module W, we have a $\mathbb{C}l_{2m}$-equivariant linear isomorphism

$$\Phi : \operatorname{Hom}_0(\Sigma_{2m}, W) \otimes \Sigma_{2m} \cong W \text{ via } \Phi(\phi \otimes \psi) := \phi(\psi),$$

where $\mathrm{Hom}_0(\Sigma_{2m}, W)$ consists of the $\mathbb{C}l_{2m}$-equivariant linear maps $\Sigma_{2m} \to W$. If M admits a spin structure $P \to FM$, then Φ yields an isomorphism

$$\Phi_M : W(M) \cong E \otimes \Sigma(M) \cong \mathrm{End}_0(\Sigma(M), W(M)) \otimes \Sigma(M), \text{ where}$$
$$E_0 := \mathrm{End}_0(\Sigma(M), W(M)) = P \times_{\mathrm{Spin}(n)} \mathrm{End}_0(\Sigma_{2m}, W).$$

Since $\phi \in \mathrm{Hom}_0(\Sigma_{2m}, W)$, we have

$$\Phi(\phi \otimes \rho(\omega_{\mathbb{C}})\psi) = \phi(\rho(\omega_{\mathbb{C}})\psi) = cl(\omega_{\mathbb{C}})\phi(\psi) = \mu_{\mathbb{C}}\phi(\psi), \text{ and so}$$

$$\Phi_M\left(W(M)^{\pm}\right) = E_0 \otimes \Sigma(M)^{\pm}. \tag{18.121}$$

As both $C^\infty(\Sigma(M))$ and $C^\infty(W(M))$ have covariant differentiation operators arising ultimately from the Levi-Civita connection, $C^\infty\left(W(M) \otimes \Sigma(M)^*\right)$ also has such an operator, say $\nabla^{W\otimes\Sigma^*}$. Now E_0 is a subbundle of $W(M) \otimes \Sigma(M)^*$, and $\nabla^{W\otimes\Sigma^*}(C^\infty(E_0)) \subset \Omega^1(E_0)$ as a consequence of (18.120). Indeed,

$$\begin{aligned}
&\phi \in C^\infty(E_0) \Rightarrow cl_W(v)\phi(\sigma) = \phi(cl_\Sigma(v)\sigma)\\
&\quad\Rightarrow \nabla_X(cl_W(v)\phi(\sigma)) = \nabla_X(\phi(cl_\Sigma(v)\sigma))\\
&\quad\Rightarrow cl_W(\nabla_X v)(\phi(\sigma)) + cl_W(v)\left(\nabla_X^{W\otimes\Sigma^*}\phi\right)(\sigma) + cl_W(v)(\phi(\nabla_X\sigma))\\
&\quad= \left(\nabla_X^{W\otimes\Sigma^*}\phi\right)(cl_\Sigma(v)\sigma) + \phi(cl_\Sigma(\nabla_X v)\sigma) + \phi(cl_\Sigma(v)\nabla_X\sigma)\\
&\quad\Rightarrow cl_W(v)\left(\nabla_X^{W\otimes\Sigma^*}\phi\right)(\sigma) = \left(\nabla_X^{W\otimes\Sigma^*}\phi\right)(cl_\Sigma(v)\sigma)\\
&\quad\Rightarrow \nabla_X\phi \in C^\infty(E_0).
\end{aligned}$$

For $\nabla^{E_0} := \nabla^{W\otimes\Sigma^*}|_{C^\infty(E_0)}$, we then have the twisted Dirac operator

$$\mathcal{D}^{E_0} := (1 \otimes cl_\Sigma) \circ \nabla : C^\infty(E_0 \otimes \Sigma(M)) \to C^\infty(E_0 \otimes \Sigma(M)), \text{ where}$$
$$\nabla := \nabla^{E_0} \otimes 1 + 1 \otimes \nabla^\Sigma : C^\infty(E_0 \otimes \Sigma(M)) \to \Omega^1(M) \otimes C^\infty(E_0 \otimes \Sigma(M)).$$

For $\Phi_M : C^\infty(W(M)) \cong C^\infty(E_0 \otimes \Sigma(M))$ corresponding to $W(M) \cong E_0 \otimes \Sigma(M)$, we then have

$$\mathcal{D}^W = \Phi_M^{-1} \circ \mathcal{D}^{E_0} \circ \Phi_M.$$

Since Φ_M is canonical, this shows that $\mathcal{D}^W$ is canonically equivalent to twisted Dirac operator in the case M has a spin structure, and in the general case $\mathcal{D}^W$ is locally such. If M admits a spin structure, then the Index Theorem for twisted Dirac operators yields

$$\mathrm{Index}\left(\mathcal{D}^W\right) = \left(\mathbf{ch}(E_0) \smile \widehat{\mathbf{A}}(M)\right)[M]$$

However, we will see that

$$\mathbf{ch}(E_0) \smile \widehat{\mathbf{A}}(M) = \mathbf{ch}(W(M)) \smile \widetilde{\mathbf{A}}(M), \tag{18.122}$$

where $\widetilde{\mathbf{A}}(M)$ is the total characteristic class defined by

$$\widetilde{\mathbf{A}}(M) := \mathbf{MC}\left(\tfrac{y/2}{\sinh(y)}, TM\right), \tag{18.123}$$

which is "almost" the same as $\widehat{\mathbf{A}}(M) = \mathbf{MC}\left(\frac{y/2}{\sinh(y/2)}, TM\right)$. A key observation is that the right side of (18.122) is defined even if M does not admit a spin structure. Shortly we will see that the standard total forms that represent the total classes $\mathbf{ch}(E_0) \smile \widehat{\mathbf{A}}(M)$ and $\mathbf{ch}(W(M)) \smile \widetilde{\mathbf{A}}(M)$ are identical. The local index density for $\mathcal{D}^W$ is computed using a local spin structure about a point. However, the result of that local computation is identical to the restriction of a globally-defined form which represents $\mathbf{ch}(W(M)) \smile \widetilde{\mathbf{A}}(M)$. In essense, the local spin structure is a computational aid, while the index density itself can be expressed without reference to spin structures. Thus, once (18.122) is shown on the level of forms, we have an index formula for the generalized Dirac operator $\mathcal{D}^{W+} : C^\infty(W^+(M)) \to C^\infty(W^-(M))$, namely

Theorem 18.55 (Index Theorem for Generalized Dirac Operators). *For a Clifford module bundle $W(M)$ over an oriented, compact Riemannian manifold M, we have*

$$\operatorname{Index}\left(\mathcal{D}^{W+}\right) = \left(\mathbf{ch}(W(M)) \smile \widetilde{\mathbf{A}}(M)\right)[M], \tag{18.124}$$

where $\widetilde{\mathbf{A}}(M)$ is defined in (18.123).

Proof. For a normal coordinate ball $V \subset M$, let $C : P_{\mathrm{Spin}} \to FM|_V$ be a spin structure. The Levi-Civita connection θ on FM lifts to a unique connection, say $\widetilde{\theta} := c'^{-1}(C^*(\theta)) \in \Omega^1(P, \mathfrak{spin}(n))$. Note that $\Sigma(V) = P \times_{\mathrm{Spin}(n)} \Sigma_n$. Let $\mathcal{R} : P_{\mathrm{Spin}} \to \mathrm{U}(\Sigma(V))$ be the morphism determined by the representation $\rho : \mathrm{Spin}(n) \to \mathrm{U}(\Sigma(V))$, as in (16.116), p. 425. To compute the Chern forms for $\Sigma(V)$ relative to the connection ω on $\mathrm{U}(\Sigma(V))$, such $\rho' \circ \widetilde{\theta} = \mathcal{R}^*\omega$, we make use of the result (see (16.117), p.425)

$$c_j(\Sigma(V), \omega) = c_j\left(\Sigma(V), \widetilde{\theta}, \rho'\right)$$

Hence we need to determine the eigenvalues of $\rho'(b)$ for $b \in \mathfrak{spin}(n)$. Any $b \in \mathfrak{spin}(n)$, can be written in the form

$$b = b_\lambda := \tfrac{1}{2}\sum\nolimits_{j=1}^{m} \lambda_j e_{2j-1}e_{2j}$$

for some oriented, orthonormal basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$. Recall that Proposition 18.10 (p.489) provides us with an explicit representation

$$\rho : Cl_{2m} \to \mathrm{End}(\Lambda^*(\mathbb{C}^m)).$$

If $(f_1, \cdots, f_m)$ is an orthonormal basis of $\mathbb{C}^m$, then

$$(e_1, \cdots, e_{2m}) := (f_1, if_1, \cdots, f_m, if_m)$$

is an oriented orthonormal basis of $\mathbb{R}^n$. By (18.14), p.490, we have

$$\begin{aligned} &\rho(e_{2j-1}e_{2j})\left(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_p}\right) \\ &= \begin{cases} i\left(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_p}\right) & \text{if } j_k = j \text{ for some } k \\ -i\left(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_p}\right) & \text{if } j_k \neq j \text{ for all } k \end{cases}. \end{aligned}$$

Hence, $f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_p}$ is an eigenvector of $\rho\left(\frac{1}{2}\sum_{j=1}^m \lambda_j e_{2j-1}e_{2j}\right)$ with eigenvalue

$$\tfrac{i}{2}\left(\sum\nolimits_{j \in \{j_1, \ldots, j_p\}} \lambda_j - \sum\nolimits_{j \notin \{j_1, \ldots, j_p\}} \lambda_j\right)$$

Thus, in order to find the total Chern character form $\mathbf{ch}(\Sigma(V), \omega)$, we compute

$$\begin{aligned} &\sum\nolimits_{p=1}^{m} \sum\nolimits_{(j)^p} e^{\frac{1}{2}\left(\sum_{j \in \{j_1, \ldots, j_p\}} \lambda_j - \sum_{j \notin \{j_1, \ldots, j_p\}} \lambda_j\right)} \\ &= \prod\nolimits_{k=1}^{m} \left(e^{\lambda_k/2} + e^{-i\lambda_k/2}\right) = \prod\nolimits_{k=1}^{m} 2\cosh\left(\tfrac{1}{2}\lambda_k\right). \end{aligned}$$

and replace each $\sigma_k\left(\lambda_1^2,\dots,\lambda_k^2\right)$ by $p_k\left(\Omega^\theta\right)$. Hence $ch_k\left(\Sigma\left(V\right),\omega\right)$, which is defined on V, coincides with polynomial in Pontryagin forms $p_j\left(\Omega^\theta\right)$ that are defined throughout M. Note also that since $W\left(M\right) := FM\times_{\mathrm{SO}(n)} W$ is associated to FM, $\mathbf{ch}\left(W\right)$ is expressible in terms of the gobally-defined curvature form Ω^θ. Thus, the same is true for $\mathbf{ch}\left(W\left(M\right)\otimes\Sigma\left(V\right)^*\right)$ and $\mathbf{ch}\left(E_0\right)$. Since

$$\begin{aligned}\prod_{k=1}^m \frac{\frac{1}{2}\lambda_i}{\sinh\left(\frac{1}{2}\lambda_i\right)} &= \prod_{k=1}^m \frac{\frac{1}{2}\lambda_i}{2\sinh\left(\frac{1}{2}\lambda_i\right)\cosh\left(\frac{1}{2}\lambda_i\right)}\prod_{k=1}^m 2\cosh\left(\tfrac{1}{2}\lambda_i\right)\\ &= \prod_{k=1}^m \frac{\frac{1}{2}\lambda_i}{\sinh\left(\lambda_i\right)}\prod_{k=1}^m 2\cosh\left(\tfrac{1}{2}\lambda_i\right),\end{aligned}$$

we have the following equality of forms

$$\widehat{\mathbf{A}}\left(V,\theta\right) = \widetilde{\mathbf{A}}\left(V,\theta\right)\wedge\mathbf{ch}\left(\Sigma\left(V\right),\theta\right).$$

The desired equality (18.122) on the level of *forms* then follows from

$$\begin{aligned}&\mathbf{ch}\left(E_0|_V,\theta\right)\wedge\widehat{\mathbf{A}}\left(V,\theta\right) = \mathbf{ch}\left(E_0|_V,\theta\right)\wedge\left(\widetilde{\mathbf{A}}\left(V,\theta\right)\wedge\mathbf{ch}\left(\Sigma\left(V\right),\theta\right)\right)\\ &= \left(\mathbf{ch}\left(E_0|_V,\theta\right)\wedge\mathbf{ch}\left(\Sigma\left(V\right),\theta\right)\right)\wedge\widetilde{\mathbf{A}}\left(V,\theta\right)\\ &= \mathbf{ch}\left(E_0|_V\otimes\Sigma\left(V\right),\theta\right)\wedge\widetilde{\mathbf{A}}\left(V,\theta\right) = \mathbf{ch}\left(W\left(M\right),\theta\right)\wedge\widetilde{\mathbf{A}}\left(V,\theta\right).\end{aligned}\tag{18.125}$$

In view of the preliminary discussion, we have shown (18.124). □

Recall that $W\left(M\right)$ is an associated bundle of FM or a covering P. Thus, Theorem 18.55 does not yet include the case of arbitrary twisting, say by a Hermitian vector bundle $E\to M$ with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}\left(E\right)$. To handle this, we proceed as follows. If we let $\mathbb{C}l\left(TM\right)$ act trivially on E, we have

$$\mathrm{Id}_E\otimes cl : C^\infty\left(\mathbb{C}l\left(TM\right)\right)\otimes C^\infty\left(E\otimes W\left(M\right)\right)\to C^\infty\left(E\otimes W\left(M\right)\right)$$

and can define a **twisted generalized Dirac operator**

$$\begin{aligned}\mathcal{D}^{E,W} : C^\infty\left(E\otimes W\left(M\right)\right) &\overset{\nabla^{E\otimes W}}{\to} \Omega^1\left(M\right)\otimes C^\infty\left(E\otimes W\left(M\right)\right)\\ &\overset{\mathrm{Id}_E\otimes cl}{\to} C^\infty\left(E\otimes W\left(M\right)\right),\end{aligned}$$

Moreover, we have a decomposition

$$\begin{aligned}&E\otimes W\left(M\right) = \left(E\otimes W\left(M\right)^+\right)\oplus\left(E\otimes W\left(M\right)^-\right)\text{ and operator}\\ &\mathcal{D}^{E,W+} : C^\infty\left(E\otimes W\left(M\right)^+\right)\to C^\infty\left(E\otimes W\left(M\right)^-\right).\end{aligned}$$

THEOREM 18.56 (Index Theorem for Twisted Generalized Dirac Operators). *For a Clifford module bundle $W\left(M\right)$ over an oriented, compact Riemannian manifold M and a Hermitian vector bundle $E\to M$ with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}\left(E\right)$, we have*

$$\begin{aligned}\mathrm{Index}\left(\mathcal{D}^{E,W+}\right) &= \left(\mathbf{ch}\left(E\right)\smile\left(\mathbf{ch}\left(W\left(M\right)\right)\smile\widetilde{\mathbf{A}}\left(M\right)\right)\right)\left[M\right]\\ &= \left(\mathbf{ch}\left(E\otimes W\left(M\right)\right)\smile\widetilde{\mathbf{A}}\left(M\right)\right)\left[M\right],\end{aligned}\tag{18.126}$$

where $\widetilde{\mathbf{A}}\left(M\right)$ is defined in (18.123).

PROOF. We proceed as in the proof of Theorem 18.55, noting that now $\mathcal{D}^{E,W}$ is locally a twisted Dirac operator just as $\mathcal{D}^{W}$ was, except that there is an additional twist by E. In other words, while the generalized Dirac operator $\mathcal{D}^{W}$ is locally the twisted Dirac operator $\mathcal{D}^{E_0}$, the twisted generalized Dirac operator $\mathcal{D}^{E,W}$ is locally the twisted Dirac operator $\mathcal{D}^{E\otimes E_0}$. The computation (18.125) generalizes to

$$\begin{aligned}
&\mathbf{ch}\left(E\otimes E_0|_V,\varepsilon\otimes\theta\right)\wedge\widehat{\mathbf{A}}\left(V,\theta\right)\\
&=\mathbf{ch}\left(E\otimes E_0|_V,\varepsilon\otimes\theta\right)\wedge\left(\widetilde{\mathbf{A}}\left(V,\theta\right)\wedge\mathbf{ch}\left(\Sigma\left(V\right),\theta\right)\right)\\
&=\mathbf{ch}\left(E|_V,\varepsilon\right)\wedge\left(\mathbf{ch}\left(E_0|_V,\theta\right)\wedge\mathbf{ch}\left(\Sigma\left(V\right),\theta\right)\right)\wedge\widetilde{\mathbf{A}}\left(V,\theta\right)\\
&=\mathbf{ch}\left(E|_V,\varepsilon\right)\wedge\mathbf{ch}\left(E_0|_V\otimes\Sigma\left(V\right),\theta\right)\smile\widetilde{\mathbf{A}}\left(V,\theta\right)\\
&=\mathbf{ch}\left(E|_V,\varepsilon\right)\wedge\mathbf{ch}\left(W\left(M\right),\theta\right)\wedge\widetilde{\mathbf{A}}\left(V,\theta\right),
\end{aligned}\tag{18.127}$$

which implies (18.126) as before. □

The Hirzebruch Signature Formula.

The standard irreducible representation of $\mathbb{C}l\left(n\right)$ is the $\mathbb{C}$-linear extension $\rho_{\mathbb{C}}:\mathbb{C}l_{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$ of $\rho:Cl_{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$ which is described as follows. Let $\rho_1:\mathbb{C}^m\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$ be given by

$$\rho_1\left(w\right)\left(\alpha\right):=\left(w\wedge-w\llcorner\right)\left(\alpha\right)=w\wedge\alpha-w\llcorner\alpha.$$

We think of $\mathbb{C}^m$ as $\mathbb{R}^{2m}$. By Proposition 18.10, p. 489, ρ_1 uniquely extends to an $\mathbb{R}$-linear homomorphism

$$\rho:Cl_{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right),$$

of algebras over $\mathbb{R}$. Then $\rho_{\mathbb{C}}:\mathbb{C}l_{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$ is the $\mathbb{C}$-linear extension of ρ. There is also a highly reducible representation $Q:\mathbb{C}l_{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^{2m}\right)\right)$ that is determined (where we regard $\mathbb{C}^{2m}\subset\mathbb{C}l_{2m}$) by

$$Q_1:\mathbb{C}^{2m}\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^{2m}\right)\right),$$

which is the same as $\rho_1:\mathbb{C}^m\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$ with m replaced by $2m$. However in the case of Q, we think of $\mathbb{C}^{2m}$ as being the complexification of $\mathbb{R}^{2m}$ instead of $\mathbb{R}^{2m}$ as the realification of $\mathbb{C}^m$. Let $q_0:\mathrm{SO}\left(2m\right)\to\operatorname{End}\left(\Lambda^*\left(\mathbb{C}^{2m}\right)\right)$ be the complex extension of the usual representation $\mathrm{SO}\left(2m\right)\to\operatorname{End}\left(\Lambda^*\left(\mathbb{R}^{2m}\right)\right)$. Since $\wedge$ and $\llcorner$ are $\mathrm{SO}\left(2m\right)$-invariant operations, for $A\in\mathrm{SO}\left(2m\right)$ and $v\in\mathbb{C}^{2m}$, we have

$$q_0\left(A\right)\circ Q_1\left(v\right)\circ q_0\left(A^{-1}\right)=Q_1\left(A\left(v\right)\right)=Q_1\left(r\left(A\right)v\right)$$

It follows that for $A\in\mathrm{SO}\left(2m\right)$, and $\alpha\in\mathbb{C}l_{2m}$, we have

$$\begin{aligned}
&q_0\left(A\right)\circ Q\left(\alpha\right)\circ q_0\left(A^{-1}\right)=Q\left(r\left(A\right)\alpha\right),\ \text{and}\\
&q\left(g\right)Q\left(\alpha\right)q\left(g^{-1}\right)=Q\left(r\left(c\left(g\right)\right)\alpha\right)\ \text{for all } g\in\operatorname{Spin}\left(n\right)\ \text{and } \alpha\in\mathbb{C}l_n,
\end{aligned}$$

where $q=q_0\circ c$. Hence, $\Lambda^*\left(T_{\mathbb{C}}M\right):=FM\times_{\mathrm{SO}(2m)}\Lambda^*\left(\mathbb{C}^{2m}\right)$ is a Clifford module bundle. We now determine $\Lambda^*\left(T_{\mathbb{C}}M\right)^+$ and $\Lambda^*\left(T_{\mathbb{C}}M\right)^-$. For $e_1,\ldots,e_n$ an oriented, orthonormal basis of $\mathbb{R}^n$ and $\omega_{\mathbb{C}}=i^m e_1\cdots e_n$, we show that

$$Q\left(\omega_{\mathbb{C}}\right)\left(e_{i_1}\wedge\cdots\wedge e_{i_k}\right)=i^{m+k(k-1)}*\left(e_{i_1}\wedge\cdots\wedge e_{i_k}\right),$$

where $*$ is the Hodge star operator on $\Lambda^*\left(\mathbb{C}^{2m}\right)$. Let

$$*\left(e_{i_1}\wedge\cdots\wedge e_{i_k}\right)=e_{j_1}\wedge\cdots\wedge e_{j_{n-k}}.$$

Since $e_{i_1}\cdots e_{i_k}e_{j_1}\cdots e_{j_{n-k}} = e_1\cdots e_n$ in $\mathbb{C}l_n$, we have

$$\begin{aligned}
&Q(\omega_{\mathbb{C}})(e_{i_1}\wedge\cdots\wedge e_{i_k}) = Q\left(i^m e_{i_1}\cdots e_{i_k}e_{j_1}\cdots e_{j_{n-k}}\right)(e_{i_1}\wedge\cdots\wedge e_{i_k})\\
&= i^m(-1)^{k(n-k)}Q\left(e_{j_1}\cdots e_{j_{n-k}}e_{i_1}\cdots e_{i_k}\right)(e_{i_1}\wedge\cdots\wedge e_{i_k})\\
&= i^m(-1)^{k(2m-k)}Q\left(e_{j_1}\cdots e_{j_{n-k}}\right)Q\left(e_{i_1}\cdots e_{i_k}\right)(e_{i_1}\wedge\cdots\wedge e_{i_k})\\
&= i^m(-1)^{-k^2}Q\left(e_{j_1}\cdots e_{j_{n-k}}\right)(-1)^{k(k-1)/2}Q\left(e_{i_k}\cdots e_{i_1}\right)(e_{i_1}\wedge\cdots\wedge e_{i_k})\\
&= i^m(-1)^{-k+k(k-1)/2}Q\left(e_{j_1}\cdots e_{j_{n-k}}\right)(-1)^k 1\\
&= i^m(-1)^{k(k-1)/2}e_{j_1}\wedge\cdots\wedge e_{j_{n-k}} = i^{m+k(k-1)}*(e_{i_1}\wedge\cdots\wedge e_{i_k}).
\end{aligned}$$

Hence if $*_k := *|_{\Lambda^k(\mathbb{C}^{2m})}$,

$$Q(\omega_{\mathbb{C}}) = \bigoplus_{k=0}^{n} i^{m+k(k-1)} *_k . \tag{18.128}$$

Usually one defines the star operator on forms, sections of exterior products of T^*M, instead of sections of exterior products of TM, but the Riemannian metric identifies T^*M with TM, and so we have a choice of leaving the above alone or dualizing it, replacing $\rho_1 : \mathbb{C}^m \to \mathrm{End}(\Lambda^*(\mathbb{C}^m))$ by

$$\begin{aligned}
&\rho_1^* : \mathbb{C}^m \to \mathrm{End}\left(\Lambda^*\left((\mathbb{C}^m)^*\right)\right) \text{ defined by}\\
&\rho_1^*(w)(\alpha) := (w_\flat\wedge - w\llcorner)(\alpha) = w\wedge\alpha - w\llcorner\alpha.
\end{aligned}$$

The advantage of dualizing is that we will find that then the Dirac operator $\mathcal{D}^W$ is the familiar sum $d+\delta$ of the exterior derivative d and its adjoint (the exterior codifferential δ), instead of the less familiar operators on $C^\infty(\Lambda^*(TM))$. Thus, we go ahead and dualize, in which case $W = \Lambda^*\left(\left(\mathbb{C}^{2m}\right)^*\right)$. Refering to (18.117), (18.118) and (18.128), we then have that $\mu_{\mathbb{C}}(x) := cl(\omega_{\mathbb{C}}(x)) \in \mathrm{End}(W(M)_x)$ is given by

$$\begin{aligned}
&\mu_{\mathbb{C}}(x) = \tau(x) := \bigoplus_{k=0}^{n}\tau_k := \bigoplus_{k=0}^{n} i^{m+k(k-1)} *_k \text{ at } x\in M, \text{ and}\\
&W^\pm(M) = \Lambda^\pm(M) := (1\pm\tau)\Lambda^*\left((T_{\mathbb{C}}M)^*\right).
\end{aligned}$$

We may directly check that $\tau(x)^2 = \mathrm{Id}$. Indeed,

$$\begin{aligned}
&\left(i^{m+(2m-k)(2m-k-1)}*_{2m-k}\right)\left(i^{m+k(k-1)}*_k\right)\\
&= i^{m+(2m-k)(2m-k-1)+m+k(k-1)}\left(*_{2m-k}*_k\right)\\
&= i^{2k^2+4m^2-4mk}(-1)^{k(2m-k)}\,\mathrm{Id} = (-1)^{k^2}(-1)^{-k^2}\,\mathrm{Id} = \mathrm{Id}.
\end{aligned}$$

Note that $\mathcal{D}^W = d+\delta$, since

$$\mathcal{D}^W\varphi = cl\left(E^k\right)(\nabla_{E_k}\alpha) = \varphi^k\wedge(\nabla_{E_k}\alpha) - E^k\llcorner(\nabla_{E_k}\alpha) = (d+\delta)\varphi.$$

In view of this and the theorem and corollary below, $d+\delta$ is sometimes called the **DeRham-Dirac operator**. A form $\varphi\in\Omega^k(M) := \Omega^k(M,\mathbb{C})$ is **harmonic** if $(d+\delta)\varphi = 0$. We let

$$\mathcal{H}^k(M) := \left\{\varphi\in\Omega^k(M,\mathbb{C}) : (d+\delta)\varphi = 0\right\}$$

denote the space of harmonic forms. Since $\mathcal{D}^W$ is elliptic, $\operatorname{Ker}\left(\mathcal{D}^W\right)$ is finite-dimensional. Moreover, we have

THEOREM 18.57 (Hodge-DeRham Decomposition). *For $0 \leq k \leq n$, there is an orthogonal decomposition*

$$\begin{aligned} \Omega^q(M) &= \mathcal{H}^q(M) \oplus d\left(\Omega^{q-1}(M)\right) \oplus \delta\left(\Omega^{q+1}(M)\right) \\ &= \mathcal{H}^q(M) \oplus d\delta\left(\Omega^q(M)\right) \oplus \delta d\left(\Omega^q(M)\right). \end{aligned} \tag{18.129}$$

PROOF. The operator

$$\Delta := d\delta + \delta d = (d+\delta)^2 : \Omega^{0,q}(M) \to \Omega^{0,q}(M)$$

is elliptic, and hence we have the orthogonal decomposition

$$\Omega^q(M) = \operatorname{Ker}\Delta \oplus \Delta\left(\Omega^q(M)\right).$$

Moreover, $\operatorname{Ker}\Delta = \mathcal{H}^q(M)$. Indeed,

$$\begin{aligned} \alpha \in \mathcal{H}^q(M) &\Rightarrow (d+\delta)\alpha = 0 \\ &\Rightarrow d\alpha = 0 \text{ and } \delta\alpha = 0 \Rightarrow (d\delta + \delta d)\alpha = 0, \end{aligned}$$

and conversely if $\alpha \in \operatorname{Ker}\Delta$, then

$$0 = \left((d\delta + \delta d)\alpha, \alpha\right) = \|\delta\alpha\|^2 + \|d\alpha\|^2 \Rightarrow d\alpha + \delta\alpha = 0.$$

It remains to prove that we have an orthogonal decompostion

$$\Delta\left(\Omega^q(M)\right) = d\left(\Omega^q(M)\right) \oplus \delta\left(\Omega^{q+1}(M)\right). \tag{18.130}$$

The summands are orthogonal since

$$\beta \in \Omega^{q-1}(M) \text{ and } \gamma \in \Omega^{q+1}(M) \Rightarrow (d\beta, \delta\gamma) = (d\beta, \gamma) = 0.$$

Moreover, for any $\alpha \in \Omega^q(M)$, we have

$$\Delta\alpha = d\delta\alpha + \delta d\alpha \in d\left(\Omega^{q-1}(M)\right) \oplus \delta\left(\Omega^{q+1}(M)\right), \tag{18.131}$$

$$\text{and so } \Delta\left(\Omega^q(M)\right) \subseteq d\left(\Omega^{q-1}(M)\right) \oplus \delta\left(\Omega^{q+1}(M)\right).$$

For the reverse inclusion, note that $d\left(\Omega^{0,q-1}(M)\right)$ and $\delta\left(\Omega^{0,q-1}(M)\right)$ are both in $\mathcal{H}^q(M)^{\perp}$, since

$$\alpha \in \mathcal{H}^q(M) \Rightarrow (d\beta, \alpha) = (\beta, \delta\alpha) = 0 = (\gamma, d\alpha) = (\delta\gamma, \alpha),$$

$$\text{and so } d\left(\Omega^{q-1}(M)\right) \oplus \delta\left(\Omega^{q+1}(M)\right) \subseteq \mathcal{H}^q(M)^{\perp} = \Delta\left(\Omega^q(M)\right).$$

Note that (18.131) and (18.130) then give the second equality of (18.129). □

COROLLARY 18.58. *Suppose that $d\gamma = 0$ for some $\gamma \in \Omega^q(M)$. There is a unique $\alpha \in \mathcal{H}^q(M)$, such that for some $\beta \in \Omega^{q-1}(M)$, $\gamma = \alpha + d\beta$. In other words, every cohomology class in the DeRham cohomology space*

$$H^q(M) := \frac{\operatorname{Ker}\left(d : \Omega^q(M) \to \Omega^{q+1}(M)\right)}{d\Omega^{q-1}(M)}$$

has a unique harmonic representative.

Proof. Theorem 18.57 yields a unique $\alpha \in \mathcal{H}^q(M)$ such that

$$\gamma = \alpha + d\beta + \delta\beta'$$

for some $\beta \in \Omega^{q-1}(M)$ and $\beta' \in \Omega^{q+1}(M)$. Now

$$\begin{aligned} 0 &= d\gamma = d\alpha + d^2\beta + d\delta\beta' = d\delta\beta' \\ &\Rightarrow (d\delta\beta', \beta') = 0 \Rightarrow \|\delta\beta'\|^2 = 0 \Rightarrow \delta\beta' = 0. \end{aligned}$$

□

If $\mathcal{H}^m(M,\mathbb{R})$ ($\cong H^m(M,\mathbb{R})$) denotes the space of $\mathbb{R}$-valued harmonic forms on M, we have a bilinear form

$$B : \mathcal{H}^m(M,\mathbb{R}) \times \mathcal{H}^m(M,\mathbb{R}) \to \mathbb{R} \text{ given by } B(\alpha,\beta) := \int_M \alpha \wedge \beta$$

If m is odd, B is anti-symmetric. If m is even (i.e. $\dim M = 2m \equiv 0 \bmod 4$), B is symmetric. Let $(\alpha,\beta) := \int_M \alpha \wedge *\beta = \int_M \langle \alpha,\beta\rangle\, \nu_M$. Since $(*_m)^2 = (-1)^{m(2m-m)}\,\mathrm{Id} = (-1)^m\,\mathrm{Id}$, we have

$$B(\alpha,\beta) = \int_M \alpha \wedge \beta = \int_M \alpha \wedge *_m(*_m\beta) = (\alpha, (-1)^m *_m \beta).$$

Note that $\tau_m = i^{m+m(m-1)} *_m = i^{m^2} *_m$. Thus,

$$m \text{ even} \Rightarrow \tau_m = *_m \Rightarrow B(\alpha,\beta) = (\alpha, *_m\beta) = (\alpha, \tau_m\beta).$$

Observe that $\alpha \in \mathcal{H}^k(M,\mathbb{R}) \Leftrightarrow *\alpha \in \mathcal{H}^{n-k}(M,\mathbb{R})$, since (using $\delta = \pm * d *$)

$$\begin{aligned} &d\alpha = 0 \text{ and } \delta\alpha = 0 \Leftrightarrow d\alpha = 0 \text{ and } d(*\alpha) = 0 \\ &\Leftrightarrow d(*(*\alpha)) = 0 \text{ and } d(*\alpha) = 0 \Leftrightarrow \delta\alpha = 0 \text{ and } d(*\alpha) = 0. \end{aligned}$$

Alternatively, we know on general grounds that $\tau = cl(\omega_{\mathbb{C}})$ and $d + \delta = \mathcal{D}^W$ anti-commute. At any rate,

$$\mathcal{H}^k(M,\mathbb{R}) \overset{*_k}{\cong} \mathcal{H}^{n-k}(M,\mathbb{R}) \text{ and } \mathcal{H}^k(M,\mathbb{C}) \overset{\tau_k}{\cong} \mathcal{H}^{n-k}(M,\mathbb{C})$$

Moreover, $B(\alpha,\beta) = (\alpha, *_m\beta) = (\alpha, \tau_m\beta)$ implies that

$$\begin{aligned} &\mathrm{sig}(M) := \text{signature of } B \\ &= \dim_{\mathbb{R}}((1+*)\mathcal{H}^m(M,\mathbb{R})) - \dim_{\mathbb{R}}((1-*)\mathcal{H}^m(M,\mathbb{R})) \\ &= \dim_{\mathbb{C}}((1+\tau)\mathcal{H}^m(M,\mathbb{C})) - \dim_{\mathbb{C}}((1-\tau)\mathcal{H}^m(M,\mathbb{C})); \end{aligned}$$

i.e., $\mathrm{sig}(M)$ is the difference between the dimensions of the spaces of self-dual and anti-self-dual harmonic forms. In the present case, the generalized Dirac operator $\mathcal{D}^{W+} = (d+\delta)^+$ is known as the **signature operator**, because

$$\begin{aligned} &\mathrm{Index}\left((d+\delta)^+ : C^\infty\left(\Lambda^+(M)\right) \to C^\infty\left(\Lambda^-(M)\right)\right) \\ &= \dim \mathrm{Ker}\,(d+\delta)^+ - \dim \mathrm{Ker}\,(d+\delta)^- \\ &= \dim((1+\tau)\mathcal{H}^*(M,\mathbb{C})) - \dim((1-\tau)\mathcal{H}^*(M,\mathbb{C})) \\ &= \dim((1+\tau)\mathcal{H}^m(M,\mathbb{C})) - \dim((1-\tau)\mathcal{H}^m(M,\mathbb{C})) \\ &= \mathrm{sig}(M). \end{aligned}$$

The last inequality follows from the fact that for $k \neq m$,

$$1 \pm \tau_k : \mathcal{H}^k(M,\mathbb{C}) \to \mathcal{H}^k(M,\mathbb{C}) \oplus \mathcal{H}^{2m-k}(M,\mathbb{C})$$

is clearly injective, and so for $k \neq m$,

$$\dim\left(1+\tau\right)\mathcal{H}^k\left(M,\mathbb{C}\right) = \dim\mathcal{H}^k\left(M,\mathbb{C}\right) = \dim\left(1-\tau\right)\mathcal{H}^k\left(M,\mathbb{C}\right).$$

Now Theorem 18.55 yields

$$\begin{aligned}\operatorname{sig}\left(M\right) = \operatorname{Index}\left(\mathcal{D}^{W,+}\right) &= \left(\mathbf{ch}\left(W\left(M\right)\right) \smile \widetilde{\mathbf{A}}\left(M\right)\right)\left[M\right] \\ &= \left(\mathbf{ch}\left(\Lambda^*\left(T^*M\right)\right) \smile \widetilde{\mathbf{A}}\left(M\right)\right)\left[M\right].\end{aligned}$$

Using (16.118), p.426, and (18.123), we obtain

$$\begin{aligned}&\mathbf{ch}\left(\Lambda^*\left(\left(T_{\mathbb{C}}M\right)^*\right)\right) \smile \widetilde{\mathbf{A}}\left(M\right) \\ &\qquad = \mathbf{MC}\left(4\cosh^2\left(y/2\right), TM\right) \smile \mathbf{MC}\left(\frac{y/2}{\sinh y}, TM\right) \\ &\qquad = \mathbf{MC}\left(4\cosh^2\left(y/2\right)\frac{y/2}{\sinh y}, TM\right) \\ &\qquad = \mathbf{MC}\left(4\cosh^2\left(y/2\right)\frac{y/2}{2\sinh\left(y/2\right)\cosh\left(y/2\right)}, TM\right) \\ &\qquad = \mathbf{MC}\left(\frac{y}{\tanh\left(y/2\right)}, TM\right).\end{aligned} \tag{18.132}$$

Let $\deg_m$ denote the m-th degree part of a power series in $y_1, \ldots, y_m$. Then

$$\begin{aligned}&\deg_m\left(\prod_{k=1}^m \frac{y_k}{\tanh\left(y_k/2\right)}\right) = 2^m \deg_m\left(\prod_{k=1}^m \frac{y_k/2}{\tanh\left(y_k/2\right)}\right) \\ &= 2^m 2^{-m}\deg_m\left(\prod_{k=1}^m \frac{y_k}{\tanh y_k}\right) = \deg_m\left(\prod_{k=1}^m \frac{y_k}{\tanh y_k}\right).\end{aligned}$$

Since $\mathbf{L}\left(M\right) := \mathbf{L}\left(TM\right) = \mathbf{MC}\left(\frac{y}{\tanh y}, TM\right)$, we then obtain

THEOREM 18.59 (Hirzebruch Signature Theorem). *Let M be a compact, oriented Riemannian $2m$-manifold, where m is even. Then*

$$\operatorname{sig}\left(M\right) = \mathbf{L}\left(M\right)\left[M\right] =: L\left(M\right) = L\text{-genus of } M.$$

In terms of the Pontryagin classes $p_k = p_k\left(TM\right)$, we have

$$\begin{aligned}2m = 4 \quad &\Rightarrow \quad \operatorname{sig}\left(M\right) = \tfrac{1}{3}p_1\left[M\right] \\ 2m = 8 \quad &\Rightarrow \quad \operatorname{sig}\left(M\right) = \tfrac{1}{45}\left(7p_2 - p_1^2\right)\left[M\right],\end{aligned}$$

in particular, and one may extend this using (16.109).

There is a twisted version of this theorem which we describe as follows. Let $E \to M$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}\left(E\right)$. Recall from Section 16.16.3, that there is an exterior covariant derivative operator $D^\varepsilon : \Omega^k\left(M,E\right) \to \Omega^{k+1}\left(M,E\right)$ which generalizes $d : \Omega^k\left(M\right) \to \Omega^{k+1}\left(M\right)$, but $D^\varepsilon \circ D^\varepsilon \neq 0$ in general. Moreover, D^ε has a formal adjoint $\delta^\varepsilon : \Omega^{k+1}\left(M,E\right) \to \Omega^k\left(M,E\right)$. Using the local formulas (16.22) and (16.23), we see that for $W = \Lambda^*\left(\left(\mathbb{C}^{2m}\right)^*\right)$, the twisted generalized Dirac operator is

$$\mathcal{D}^{E,W} = \left(d+\delta\right)^E := D^\varepsilon + \delta^\varepsilon : \Omega^*\left(M,E\right) \to \Omega^*\left(M,E\right). \tag{18.133}$$

which is called the **twisted DeRham-Dirac operator**. The twisted version of the Hirzebruch Signature Theorem has the additional "twist" that perhaps unexpectedly, $\operatorname{Index}\left((d+\delta)^{E,+}\right) \neq (\mathbf{ch}(E) \smile \mathbf{L}(M))[M]$ in general.

THEOREM 18.60 (Twisted Hirzebruch Signature Theorem). *Let M be a compact, oriented Riemannian $2m$-manifold, where m is even, and let $E \to M$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then the index of the twisted signature operator*

$$(d+\delta)^{E,+} = \mathcal{D}^{E,W+} : \Omega^{+}(M,E) \to \Omega^{-}(M,E)$$

(where $\Omega^{\pm}(M,E) := C^{\infty}(E \otimes \Lambda^{\pm}(M))$) is given by

$$\operatorname{Index}\left((d+\delta)^{E,+}\right) = \left(\mathbf{ch}(E) \smile \mathbf{MC}\left(\frac{y}{\tanh(y/2)}, TM\right)\right)[M]$$
$$= \left(\mathbf{ch}_2(E) \smile \mathbf{MC}\left(\frac{y}{\tanh y}, TM\right)\right)[M] = (\mathbf{ch}_2(E) \smile \mathbf{L}(M))[M],$$

where $\mathbf{ch}_2(E) := \bigoplus_{j=0}^{m} 2^j ch_j(E)$.

PROOF. Using Theorem 18.56 and (18.132),

$$\begin{aligned}
&\operatorname{Index}\left((d+\delta)^{E,+}\right)\\
&= \left(\mathbf{ch}(E) \smile \left(\mathbf{ch}\left(\Lambda^{*}\left((T_{\mathbb{C}}M)^{*}\right)\right) \smile \widetilde{\mathbf{A}}(M)\right)\right)[M]\\
&= \left(\mathbf{ch}(E) \smile \mathbf{MC}\left(\frac{y}{\tanh y/2}, TM\right)\right)[M]
\end{aligned}$$

Now

$$\begin{aligned}
\mathbf{MC}\left(\frac{y}{\tanh y/2}, TM\right) &= 2^m \mathbf{MC}\left(\frac{y/2}{\tanh y/2}, TM\right)\\
&= 2^m \bigoplus_{k=1}^{m} 2^{-2k} L_k(y).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbf{ch}(E) \smile \mathbf{MC}\left(\frac{y}{\tanh y/2}, TM\right)[M]\\
&\quad= \left(\bigoplus_{j=1}^{m} ch_j(E) \smile 2^m \bigoplus_{k=1}^{m} 2^{-2k} L_k(y)\right)[M]\\
&\quad= \left(\bigoplus_{\{(j,k):j+2k=m\}} 2^{m-2k} ch_j(E) \smile L_k(y)\right)[M]\\
&\quad= \left(\bigoplus_{\{(j,k):j+2k=m\}} 2^{j} ch_j(E) \smile L_k(y)\right)[M]\\
&\quad= (\mathbf{ch}_2(E) \smile \mathbf{L}(M))[M].
\end{aligned}$$

□

Recall that $J \in \operatorname{End}(TM)$ is an almost complex structure of a manifold M, if $J^2 = -I$. If J exists, then TM becomes a complex vector bundle by defining $(a+ib)X = aX + bJX$. For a complex structure to exist, $\dim M$ must be even.

EXERCISE 18.61. Here we use Theorem 18.60 to show if the sphere S^n $(n = 2m > 0)$ admits an almost complex structure, then $n = 1, 2$ or 3.
(a) Show that the total Pontryagin class $\mathbf{p}(S^n)$ is $1 \in H^0(S^n)$. [Hint. If N is the (trivial) normal bundle of $S^n \subseteq \mathbb{R}^{n+1}$, then $TS^n \oplus N$ is trivial.]
(b) In the identity $(x + x_1) \cdots (x + x_m) = x^m + \sigma_1 x^{m-1} + \cdots + \sigma_m$, where σ_i is the i-th elementary symmetric polynomial in the x_i, sucessively subtitute $x = x_1, \ldots, x = x_m$ and add the results to get Newton's formula

$$0 = (-1)^m \sum_{i=1}^{m} x_i^m + (-1)^{m-1} \sigma_1 \sum_{i=1}^{m} x_i^{m-1} + \cdots + m\sigma_m. \tag{18.134}$$

(b) Assuming that S^{2m} admits an almost complex structure, use (18.134) and the fact $c_m\left(TS^{2m}\right) = \chi\left(S^{2m}\right)$ (see (16.115) p. 424) to deduce that

$$2^m ch_m\left(TS^{2m}\right)\left[S^{2m}\right] = \frac{(-1)^{m+1} 2^{m+1}}{(m-1)!},$$

(c) Finally show that (b) and Theorem 18.60 imply that m must be 1, 2, or 3.

REMARK 18.62. Of course S^2 has a complex (and hence almost-complex) structure. S^4 does not have a complex structure, since otherwise,

$$\begin{aligned} 1 = 1 - p_1\left(TS^4\right) &= \mathbf{c}\left(\mathbb{C} \otimes TS^4\right) = \mathbf{c}\left(TS^4 \oplus \overline{TS^4}\right) \\ &= \mathbf{c}\left(TS^4\right) \smile \mathbf{c}\left(\overline{TS^4}\right) = \left(1 + c_2\left(TS^4\right)\right)^2 = 1 + 2c_2\left(TS^4\right), \end{aligned}$$

but $c_2\left(TS^4\right) = \chi\left(S^{2m}\right) \neq 0$; the same argument works for S^{4k}, using $1 + (-1)^k p_k\left(TS^4\right) = \mathbf{c}\left(\mathbb{C} \otimes TS^{4k}\right)$. We show that S^6 has a complex structure as follows. Recall that the Cayley numbers $\mathbb{O} := \mathbb{H} \times \mathbb{H}$, with multiplication

$$pq = (p_1, p_2)(q_1, q_2) = \left(p_1 q_1 - \overline{q_2} p_2, q_2 p_1 + p_2 \overline{q_1}\right),$$

are neither associative nor commutative, but $(1, 0)$ is a multiplicative identity. Let $\overline{p} := (\overline{p_1}, -p_2)$ and $\operatorname{Re}(p) := \frac{1}{2}(p + \overline{p}) = \operatorname{Re}(p_1)$. Then

$$\begin{aligned} \langle p, q \rangle &:= \operatorname{Re}(p\overline{q}) = \operatorname{Re}\left(p_1 \overline{q_1} + \overline{q_2} p_2\right) \text{ and} \\ |p|^2 &= \langle p, p \rangle = \operatorname{Re}\left(p_1 \overline{p_1} + \overline{p_2} p_2\right) = |p_1|^2 + |p_2|^2. \end{aligned}$$

Thus, if we identify p and q with vectors in $\mathbb{R}^8$, then $\langle p, q \rangle$ is just the usual dot product, and

$$\Sigma^6 := \{p \in \mathbb{O} : |p| = 1 \text{ and } \langle p, (1, 0) \rangle = 0\}$$

is a 6-sphere. Note that

$$\begin{aligned} p \in \Sigma^6 \Rightarrow p^2 &= \left(p_1 p_1 - \overline{p_2} p_2, p_2 p_1 + p_2 \overline{p_1}\right) \\ &= \left(-p_1 \overline{p_1} - \overline{p_2} p_2, p_2 p_1 - p_2 p_1\right) = -|p|^2 (1, 0) = -(1, 0). \end{aligned}$$

One can check that $|pq| = |p|\,|q|$ and that (in spite of nonassociativity), $p(pq) = p^2 q$. These are rather involved computations, where, e.g., one eventually needs to use $\operatorname{Re}(xy) = \operatorname{Re}(yx)$ for $x, y \in \mathbb{H}$, and

$$\begin{aligned} &p_2 \overline{q_1} p_1 + p_2 \overline{q_1 p_1} - \left(p_2 p_1 \overline{q_1} + p_2 \overline{p_1 q_1}\right) \\ &= \left(p_2 \overline{q_1}\right)\left(p_1 + \overline{p_1}\right) - p_2\left(p_1 + \overline{p_1}\right)\overline{q_1} = 0. \end{aligned}$$

For $p \in \Sigma^6$, we have $\langle pq, pq' \rangle = \langle q, q' \rangle$ for all $q, q' \in \mathbb{O}$, since $|pq| = |p|\,|q| = |q|$. Suppose that $\langle q, (1,0) \rangle = 0$ and $\langle q, p \rangle = 0$, so that $q \in T_p\Sigma^6$. Then

$$\begin{aligned} \langle pq, (1,0) \rangle &= \langle pq, -pp \rangle = -\langle q, p \rangle = 0 \text{ and} \\ \langle pq, p \rangle &= \langle pq, p\,(1,0) \rangle = \langle q, (1,0) \rangle = 0. \end{aligned}$$

Thus, we have a well-defined linear map $J_p : T_p\Sigma^6 \to T_p\Sigma^6$ given by $J_p q = pq$. Moreover, J_p defines a complex structure since

$$J_p^2\,(q) = p\,(pq) = p^2 q = -\,|p|^2\, q = -q.$$

The Gauss-Bonnet-Chern Formula.

As noted before, the representation

$$Q : \mathbb{C}l_{2m} \to \operatorname{End}\left(\Lambda^*\left(\mathbb{C}^{2m}\right)\right),$$

determined by $Q\,(w) = w \wedge \alpha - w \llcorner \alpha$ for $w \in \mathbb{C}^{2m}$, is highly reducible. Indeed for purely dimensional reasons, Q is the sum of 2^m copies of the unique irreducible representation $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right)$. We now describe a courser (if $m > 1$) decomposition of $\Lambda^*\left(\mathbb{C}^{2m}\right)$ into just two $\mathbb{C}l_{2m}$-modules, which leads to the Gauss-Bonnet-Chern Formula. Let

$$\Lambda^{\pm} := \left(1 \pm \omega_{\mathbb{C}}\,(Q)\right)\left(\Lambda^*\left(\mathbb{C}^m\right)\right) \text{ and } \Lambda^{even(odd),\pm} := \Lambda^{even(odd)}\left(\mathbb{C}^{2m}\right) \cap \Lambda^{\pm}.$$

In view of the fact that for $w \in \mathbb{C}^{2m}$,

$$\begin{aligned} & Q\,(w)\left(\Lambda^{\pm}\right) = \Lambda^{\mp} \text{ and } Q\,(w)\left(\Lambda^{even(odd)}\left(\mathbb{C}^{2m}\right)\right) = \Lambda^{odd(even)}\left(\mathbb{C}^{2m}\right) \\ & \Rightarrow\ Q\,(w)\left(\Lambda^{even(odd),\pm}\right) = \Lambda^{odd(even),\mp}, \end{aligned}$$

the $\mathbb{C}l_{2m}$-module $\Lambda^*\left(\mathbb{C}^{2m}\right)$ splits into two submodules:

$$\Lambda^*\left(\mathbb{C}^{2m}\right) = \left(\Lambda^{even,+} \oplus \Lambda^{odd,-}\right) \oplus \left(\Lambda^{odd,+} \oplus \Lambda^{even,-}\right) =: W^e \oplus W^o.$$

Consequently, we have two generalized Dirac operators

$$\begin{aligned} \mathcal{D}^{W^e} &= d + \delta \in \operatorname{End}\left(\Omega^{even,+}\,(M) \oplus \Omega^{odd,-}\,(M)\right) \\ \mathcal{D}^{W^o} &= d + \delta \in \operatorname{End}\left(\Omega^{odd,+}\,(M) \oplus \Omega^{even,-}\,(M)\right), \end{aligned} \tag{18.135}$$

where $\Omega^{even(odd),\pm}\,(M) := C^\infty\left(\Lambda^{even(odd),\pm}\,(M)\right)$. Note that

$$\begin{aligned} \mathcal{D}^{W^e+} &: \Omega^{even,+}\,(M) \to \Omega^{odd,-}\,(M) \text{ and} \\ \mathcal{D}^{W^o+} &: \Omega^{odd,+}\,(M) \to \Omega^{even,-}\,(M),\ \text{with} \end{aligned}$$

$$\begin{aligned} \operatorname{Index}\left(\mathcal{D}^{W^e+}\right) &= \left(\mathbf{ch}\left(\Lambda^{even,+}\,(M) \oplus \Lambda^{odd,-}\,(M)\right) \smile \widetilde{\mathbf{A}}\,(M)\right)[M] \\ \operatorname{Index}\left(\mathcal{D}^{W^o+}\right) &= \left(\mathbf{ch}\left(\Lambda^{odd,+}\,(M) \oplus \Lambda^{even,-}\,(M)\right) \smile \widetilde{\mathbf{A}}\,(M)\right)[M]. \end{aligned}$$

The signature operator can be written as

$$(d + \delta)^+ = \mathcal{D}^{W^e+} \oplus \mathcal{D}^{W^o+} : \Omega^+\,(M) \to \Omega^-\,(M).$$

The **Euler operator** is

$$(d + \delta)^\chi := d + \delta : \Omega^{even}\,(M) \to \Omega^{odd}\,(M), \tag{18.136}$$

whose index (according to Hodge Theory) is $\chi(M)$. Indeed,

$$\begin{aligned}\chi(M) &:= \dim\left(\bigoplus_{k=0}^m H^{2k}(M)\right) - \dim\left(\bigoplus_{k=1}^m H^{2k-1}(M)\right)\\ &= \dim\left(\bigoplus_{k=0}^m \mathcal{H}^{2k}(M)\right) - \dim\left(\bigoplus_{k=1}^m \mathcal{H}^{2k-1}(M)\right)\\ &= \dim\left(\operatorname{Ker}(d+\delta)^\chi\right) - \dim\left(\operatorname{Ker}\left((d+\delta)^\chi\right)^*\right)\\ &= \operatorname{Index}\left((d+\delta)^\chi\right). \qquad (18.137)\end{aligned}$$

In terms of $\mathcal{D}^{W^e+}$ and $\mathcal{D}^{W^o+}$, we have that $(d+\delta)^\chi$ is

$$\mathcal{D}^{W^e+} \oplus \left(\mathcal{D}^{W^o+}\right)^* : \Omega^{even,+}(M) \oplus \Omega^{even,-}(M) \to \Omega^{odd,-}(M) \oplus \Omega^{odd,+}(M).$$

Hence,

$$\begin{aligned}\operatorname{Index}\left((d+\delta)^\chi\right) &= \operatorname{Index}\left(\mathcal{D}^{W^e,+}\right) + \operatorname{Index}\left(\left(\mathcal{D}^{W^o,+}\right)^*\right)\\ &= \operatorname{Index}\left(\mathcal{D}^{W^e,+}\right) - \operatorname{Index}\left(\mathcal{D}^{W^o,+}\right)\\ &= \left(\mathbf{ch}\left(\Lambda^{even,+}(M) \oplus \Lambda^{odd,-}(M)\right) \smile \widetilde{\mathbf{A}}(M)\right)[M]\\ &\quad - \left(\mathbf{ch}\left(\Lambda^{odd,+}(M) \oplus \Lambda^{even,-}(M)\right) \smile \widetilde{\mathbf{A}}(M)\right)[M]\\ &= \begin{pmatrix} \mathbf{ch}\left(\Lambda^{even,+}(M)\right) - \mathbf{ch}\left(\Lambda^{even,-}(M)\right) \\ +\mathbf{ch}\left(\Lambda^{odd,-}(M)\right) - \mathbf{ch}\left(\Lambda^{odd,+}(M)\right) \end{pmatrix} \smile \widetilde{\mathbf{A}}(M)[M] \qquad (18.138)\end{aligned}$$

The individual results for the $\mathbf{ch}\left(\Lambda^{even(odd),\pm}(M)\right)$ are not all that simple or interesting (as far as we know), but the two combinations

$$\mathbf{ch}\left(\Lambda^{even,+}(M)\right) - \mathbf{ch}\left(\Lambda^{even,-}(M)\right) \text{ and } \mathbf{ch}\left(\Lambda^{odd,-}(M)\right) - \mathbf{ch}\left(\Lambda^{odd,+}(M)\right)$$

admit dramatic simplification. Indeed, we will find that when $m := \frac{1}{2}\dim M$ is odd the first of these is 0 and the second, when cupped with $\widetilde{\mathbf{A}}(M)$, is the Euler class. When m is even, the opposite is the case. Thus, whether m is even or odd, we will obtain the Gauss-Bonnet-Chern Theorem.

Let $\Lambda : \mathrm{SO}(n) \to \mathrm{U}\left(\Lambda^*\left((\mathbb{C}^n)^*\right)\right)$ be the representation determined by

$$\Lambda(A)(\alpha) = \left(A^T\right)^* \alpha = \alpha \circ A^T \text{ for } \alpha \in (\mathbb{C}^n)^* = \Lambda^1\left((\mathbb{C}^n)^*\right),\ A \in \mathrm{SO}(n).$$

The various $\mathbf{ch}\left(\Lambda^{even(odd),\pm}(M)\right)$ can be computed by finding the eigenvalues of of $\Lambda'(B) \in \operatorname{End}\left(\Lambda^*\left((\mathbb{C}^n)^*\right)\right)$ for $B \in \mathfrak{so}(n)$, where $\Lambda' : \mathfrak{so}(n) \to \mathfrak{u}\left(\Lambda^*\left((\mathbb{C}^n)^*\right)\right)$ is the Lie algebra representation. There is an orthonormal basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ for which B has the matrix $\bigoplus_{k=1}^m \begin{bmatrix} 0 & -y_k \\ y_k & 0 \end{bmatrix}$. One purpose of the rather lengthy digression in the next paragraph is to show that

$$\begin{aligned}&\mathbf{ch}\left(\Lambda'(B)^{even,+}\right) - \mathbf{ch}\left(\Lambda'(B)^{even,-}\right) + \mathbf{ch}\left(\Lambda'(B)^{odd,-}\right) - \mathbf{ch}\left(\Lambda'(B)^{odd,+}\right)\\ &= \prod_{k=1}^m \left(-2\sinh y_k\right),\end{aligned}$$

from which the Gauss-Bonnet-Chern Theorem will follow. The results in digression are also used in proving Hirzebruch-Riemann-Roch Theorem.

Let $\varphi_1, \ldots, \varphi_n$ be the basis of $(\mathbb{R}^n)^*$ dual to $e_1, \ldots, e_n$. Then

$$\begin{aligned}\Lambda'(B)(\varphi_1 + i\varphi_2)(e_1) &= (\varphi_1 + i\varphi_2)\left(B^T e_1\right) = (\varphi_1 + i\varphi_2)(-Be_1) \\ &= (\varphi_1 + i\varphi_2)(-y_1 e_2) = -iy_1 \\ \Lambda'(B)(\varphi_1 + i\varphi_2)(e_2) &= (\varphi_1 + i\varphi_2)(-Be_2) = (\varphi_1 + i\varphi_2)(y_1 e_1) = y_1 \\ \Lambda'(B)(\varphi_1 + i\varphi_2) &= -iy_1\varphi_1 + y_1\varphi_2 = -iy_1(\varphi_1 + i\varphi_2)\end{aligned}$$

Thus,

$$\begin{aligned}\Lambda'(B)(\varphi_1 + i\varphi_2) &= -iy_1(\varphi_1 + i\varphi_2) \\ \Lambda'(B)(\varphi_1 - i\varphi_2) &= iy_1(\varphi_1 - i\varphi_2)\end{aligned}$$

For $j = 1, \ldots, m$, let $\xi_j = \varphi_{2j-1} + i\varphi_{2j}$ and $\bar{\xi}_j = \varphi_{2j-1} - i\varphi_{2j}$. For a multi-index $(i)_k := (i_1, \ldots, i_h)$ where $1 \le i_1 < \cdots < i_h \le m$, we let

$$\xi_{(i)_h} := \xi_{i_1} \wedge \cdots \wedge \xi_{i_h} \text{ and } \bar{\xi}_{(i)_h} = \bar{\xi}_{i_1} \wedge \cdots \wedge \bar{\xi}_{i_h}.$$

Note that $\xi_{(i)_m} = \xi_1 \wedge \cdots \wedge \xi_m$ and $\bar{\xi}_{(i)_m} = \bar{\xi}_1 \wedge \cdots \wedge \bar{\xi}_m$. The $2^m \cdot 2^m = 2^n$ forms $\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}$ make a basis of eigenvectors of $\Lambda'(B)$ on $\Lambda^*\left((\mathbb{C}^n)^*\right)$ since

$$\Lambda'(B)\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = -i\left((y_{i_1} + \cdots + y_{i_h}) - (y_{j_1} + \cdots + y_{j_k})\right)\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}.$$

However, we need a basis in $\Lambda^{even(odd),\pm}$. For this, we will determine $*\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right)$ where $*$ is the complex-linear extension of usual star operator on $\Lambda^*(\mathbb{R}^n)$. If $b(\alpha, \beta)$ is the symmetric, bilinear (not Hermitian) extension of the usual inner product on $\Lambda^*(\mathbb{R}^n)$, we have

$$\alpha \wedge *\beta = b(\alpha, \beta)\,\varphi_1 \wedge \cdots \wedge \varphi_n.$$

Note that

$$\begin{aligned}b(\xi_j, \xi_j) &= b(\varphi_{2j-1} + i\varphi_{2j}, \varphi_{2j-1} + i\varphi_{2j}) = |\varphi_{2j-1}|^2 - |\varphi_{2j}|^2 = 0, \text{ and} \\ b\left(\xi_j, \bar{\xi}_j\right) &= b(\varphi_{2j-1} + i\varphi_{2j}, \varphi_{2j-1} - i\varphi_{2j}) = |\varphi_{2j-1}|^2 + |\varphi_{2j}|^2 = 2.\end{aligned}$$

Hence,

$$b\left(\bar{\xi}_{(i')_{h'}} \wedge \xi_{(j')_{k'}}, \xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = \begin{cases} 2^{h+k} & \text{if } (i')_{h'} = (i)_h\,,\ \ (j')_{k'} = (j)_k \\ 0 & \text{otherwise.}\end{cases}$$

Moreover, using $\xi_1 \wedge \bar{\xi}_1 = (\varphi_1 + i\varphi_2) \wedge (\varphi_1 - i\varphi_2) = -2i\varphi_1 \wedge \varphi_2$, we get

$$\begin{aligned}(-2i)^m\,\varphi_1 \wedge \cdots \wedge \varphi_n &= \xi_1 \wedge \bar{\xi}_1 \wedge \cdots \wedge \xi_m \wedge \bar{\xi}_m \\ &= (-1)^{m(m-1)/2}\,\xi_1 \wedge \cdots \wedge \xi_m \wedge \bar{\xi}_1 \wedge \cdots \wedge \bar{\xi}_m \\ &= i^{m(m-1)}\xi_1 \wedge \cdots \wedge \xi_m \wedge \bar{\xi}_1 \wedge \cdots \wedge \bar{\xi}_m, \text{ or}\end{aligned}$$

$$\begin{aligned}\nu_n := \varphi_1 \wedge \cdots \wedge \varphi_n &= (-2i)^{-m}\, i^{m(m-1)}\xi_1 \wedge \cdots \wedge \xi_m \wedge \bar{\xi}_1 \wedge \cdots \wedge \bar{\xi}_m \\ &= 2^{-m} i^{m^2}\xi_1 \wedge \cdots \wedge \xi_m \wedge \bar{\xi}_1 \wedge \cdots \wedge \bar{\xi}_m = 2^{-m} i^{m^2}\xi_{(i)_m} \wedge \bar{\xi}_{(i)_m}. \qquad (18.139)\end{aligned}$$

From

$$\left(\bar{\xi}_{(i')_{h'}} \wedge \xi_{(j')_{k'}}\right) \wedge *\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = b\left(\bar{\xi}_{(i')_{h'}} \wedge \xi_{(j')_{k'}}, \xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right)\nu_n,$$

we deduce that for some scalar $C = C\left((i)_h, (j)_k\right) \in \mathbb{C}$,

$$*\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = C\xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}}, \qquad (18.140)$$

where $(j^c)_{m-k}$ is the multi-index complementary $(j)_k$ in the sense that

$$\{j_1, \ldots, j_k\} \cup \left\{j_1^c, \ldots, j_{m-k}^c\right\} = \{1, \ldots, m\}.$$

We find below that

$$C = C\left((i)_h, (j)_k\right) = 2^{h+k-m} (-1)^h \, i^{m^2} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}}. \tag{18.141}$$

Indeed if $(i')_{h'} \neq (i)_h$ or $(j')_{k'} \neq (j)_k$, then

$$\left(\bar{\xi}_{(i')_{h'}} \wedge \xi_{(j')_{k'}}\right) \wedge \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}} = b\left(\bar{\xi}_{(i')_{h'}} \wedge \xi_{(j')_{k'}}, \xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) \nu_n = 0,$$

while if $(i')_{h'} = (i)_h$ and $(j')_{k'} = (j)_k$, then

$$\begin{aligned}
&\left(\bar{\xi}_{(i)_h} \wedge \xi_{(j)_k}\right) \wedge C\xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}} = C(-1)^{hm} \xi_{(j)_k} \wedge \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i)_h} \wedge \bar{\xi}_{(i^c)_{m-h}} \\
&= C(-1)^{hm} \varepsilon_{(j)_k (j^c)_{m-k}} \varepsilon_{(i)_h (i^c)_{m-h}} \xi_{(j)_m} \wedge \bar{\xi}_{(i)_m} \\
&= C(-1)^{hm} \varepsilon_{(i)_h (i^c)_{m-h}} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j)_m} \wedge \bar{\xi}_{(i)_m} \\
&= C(-1)^{hm} (-1)^{h(m-h)} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j)_m} \wedge \bar{\xi}_{(i)_m} \\
&= C(-1)^h \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j)_m} \wedge \bar{\xi}_{(i)_m} \\
&= C(-1)^h \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \frac{1}{2^{-m} i^{m^2}} \nu_n \\
&= b\left(\bar{\xi}_{(i)_h} \wedge \xi_{(j)_k}, \xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) \nu_n = 2^{h+k} \nu_n,
\end{aligned}$$

which yields the value of C in (18.141). We have

$$\begin{aligned}
&\tau_{h+k}\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = i^{m+(h+k)(h+k-1)} * \left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) \\
&= i^{m+(h+k)(h+k-1)} C\left((i)_h, (j)_k\right) \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}} \\
&= i^{m+(h+k)(h+k-1)} 2^{h+k-m} (-1)^h \, i^{m^2} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}} \\
&= i^{m(m+1)} i^{(h+k)(h+k-1)+2h} 2^{h+k-m} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}} \\
&= (-1)^{\frac{1}{2}m(m+1)+\frac{1}{2}(h-k)(h-k+1)} 2^{h+k-m} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}},
\end{aligned}$$

where we have used

$$(h+k)(h+k-1) + 2h = ((h-k)(h-k+1)) + 4hk.$$

In the case $h+k=m$, we have

$$\begin{aligned}
&\frac{1}{2} m(m+1) + \frac{1}{2}(h-k)(h-k+1) \\
&= \frac{1}{2} m(m+1) + \frac{1}{2}(2h-m)(2h-m+1) \\
&= m^2 + 2h^2 - 2mh + h \equiv m + h \mod 2, \text{ and so}
\end{aligned}$$

$$h+k=m \Rightarrow \tau_m\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = (-1)^{m+h} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}}.$$

Note that if $h+k=m$, then $(1 \pm \tau_m)\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right)$ are both eigenvectors of τ_m (with the same eigenvalue for $\Lambda'(B)$), unless one is 0. To determine when $(1 \pm \tau_m)\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = 0$, note that (18.140) yields

$$\tau_m\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = \pm \xi_{(i)_h} \wedge \bar{\xi}_{(j)_k} \Leftrightarrow (j)_k = (i^c)_{m-h} \ \text{(or } (i)_h = (j^c)_{m-k}\text{)}.$$

In this case, $\varepsilon_{(i^c)_{m-h}(i)_h}\varepsilon_{(j)_k(j^c)_{m-k}} = \left(\varepsilon_{(i^c)_{m-h}(i)_h}\right)^2 = 1$, and so

$$\tau_m\left(\xi_{(i)_h}\wedge\bar{\xi}_{(i^c)_{m-h}}\right) = (-1)^{m+h}\,\xi_{(i)_h}\wedge\bar{\xi}_{(i^c)_{m-h}},\ \text{for } 0\le h\le m, \tag{18.142}$$

while $(1\pm\tau_m)\left(\xi_{(i)_h}\wedge\bar{\xi}_{(j)_k}\right)\neq 0$ for $(i)_h\neq(j^c)_{m-k}$. Consequently, if $(j)_k\neq(i^c)_{m-h}$, the eigenvector $(1+\tau_{h+k})\left(\xi_{(i)_h}\wedge\bar{\xi}_{(j)_k}\right)$ of $\Lambda'(B)\,|_{\Lambda^{even(odd),+}}\left((\mathbb{C}^m)^*\right)$ corresponds to an eigenvector $(1-\tau_{h+k})\left(\xi_{(i)_h}\wedge\bar{\xi}_{(j)_k}\right)$ of $\Lambda'(B)\,|_{\Lambda^{even(odd),-}}\left((\mathbb{C}^m)^*\right)$ with the same eigenvalue. In the remaining case $(j)_k=(i^c)_{m-h}$, we have

$$\Lambda'(B)\left(\xi_{(i)_h}\wedge\bar{\xi}_{(i^c)_{m-h}}\right) = -i\left((y_{i_1}+\cdots+y_{i_h})-\left(y_{i_1^c}+\cdots+y_{i_{m-h}^c}\right)\right)$$

Thus taking (18.142) into account, we have

$$\begin{aligned}
&\mathbf{ch}\left(\Lambda'(B)^{even,+}\right)-\mathbf{ch}\left(\Lambda'(B)^{even,-}\right)+\mathbf{ch}\left(\Lambda'(B)^{odd,-}\right)-\mathbf{ch}\left(\Lambda'(B)^{odd,+}\right)\\
&=\sum\nolimits_{h=0}^{m}\sum\nolimits_{(i)_h}(-1)^{m+h}\exp\left(-\left((y_{i_1}+\cdots+y_{i_h})-\left(y_{i_1^c}+\cdots+y_{i_{m-h}^c}\right)\right)\right)\\
&=(-1)^m\sum\nolimits_{h=0}^{m}\sum\nolimits_{(i)_h}(-1)^h\exp\left(\left(y_{i_1^c}+\cdots+y_{i_{m-h}^c}\right)-(y_{i_1}+\cdots+y_{i_h})\right)\\
&=(-1)^m\prod\nolimits_{k=1}^{m}\left(e^{y_k}-e^{-y_k}\right)=\prod\nolimits_{k=1}^{m}\left(-2\sinh y_k\right),\ \text{and so}
\end{aligned}$$

$$\begin{aligned}
&\begin{pmatrix}\mathbf{ch}\left(\Lambda'(B)^{even,+}\right)-\mathbf{ch}\left(\Lambda'(B)^{even,-}\right)\\ +\mathbf{ch}\left(\Lambda'(B)^{odd,-}\right)-\mathbf{ch}\left(\Lambda'(B)^{odd,+}\right)\end{pmatrix}\widetilde{A}(B)\\
&=\prod\nolimits_{k=1}^{m}\left(-2\sinh y_k\right)\prod\nolimits_{k=1}^{m}\frac{y_k/2}{\sinh y_k}=(-1)^m y_1\ldots y_m=(-1)^m\,\mathrm{Pf}(B).
\end{aligned}\tag{18.143}$$

THEOREM 18.63 (Gauss-Bonnet-Chern Theorem). *Let M be a compact, orientable, Riemannian manifold of even dimension $n=2m$, and let $\mathcal{H}^k(M)$ be the space of harmonic k-forms on M. Then*

$$\begin{aligned}
\chi(M)=\mathrm{Index}\left((d+\delta)^\chi\right)&=\sum_{k=0}^{n}(-1)^k\dim\left(\mathcal{H}^k(M)\right)\\
&=\mathrm{GB}(TM)[M]=\int_M\mathrm{GB}\left(\Omega^\theta\right),
\end{aligned}$$

where $\mathrm{GB}\left(\Omega^\theta\right)\in\Omega^n(M)$ is the Gauss-Bonnet form, determined by

$$\begin{aligned}
\pi^*\,\mathrm{GB}\left(\Omega^\theta\right)&=(-1)^m\,\mathrm{Pf}\left(\tfrac{1}{2\pi}\Omega^\theta\right)\\
&=\frac{1}{2^{2m}\pi^m m!}\sum\nolimits_{(i)}\varepsilon_{i_1\cdots i_{2m}}\Omega^\theta_{i_1i_2}\wedge\cdots\wedge\Omega^\theta_{i_{2m-1}i_{2m}},
\end{aligned}$$

where Ω^θ is the curvature of the Levi-Civita connection θ (or indeed any connection) on the bundle $\pi: FM\to M$ of oriented, orthonormal frames.

PROOF. Using (18.137), (18.143) and (16.113), p.423,

$$\begin{aligned}
&\sum\nolimits_{k=0}^{n} (-1)^k \dim \left(\mathcal{H}^k (M)\right) = \operatorname{Index}\left((d+\delta)^{\chi}\right) \\
&= \begin{pmatrix} \mathbf{ch}\left(\Lambda^{even,+}(M)\right) - \mathbf{ch}\left(\Lambda^{even,-}(M)\right) \\ +\mathbf{ch}\left(\Lambda^{odd,-}(M)\right) - \mathbf{ch}\left(\Lambda^{odd,+}(M)\right) \end{pmatrix} \smile \widetilde{\mathbf{A}}(M)[M] \\
&= \mathrm{GB}(TM)[M] = \int_M \mathrm{GB}\left(\Omega^{\theta}\right).
\end{aligned}$$

□

There is also a twisted version of the Gauss-Bonnet-Chern Theorem, which has the additional "twist" (or rather "untwist") that the index of the twisted Euler operator is only affected by the dimension of the twisting bundle E, rather than any actual twisting (i.e., nontriviality) of E. We define the twisted Euler operator

$$(d+\delta)^{E,\chi} : \Omega^{even}(M,E) \to \Omega^{odd}(M,E) := (d+\delta)^E \mid_{\Omega^{even}(M,E)}$$

as the restriction $(d+\delta)^E \mid_{\Omega^{even}(M,E)}$ of the twisted DeRham-Dirac operator $(d+\delta)^E$ in (18.133).

THEOREM 18.64 (Twisted Gauss-Bonnet-Chern Theorem). *Let M be a compact, oriented Riemannian $2m$-manifold, and let $E \to M$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then the index of the twisted Euler operator is given by*

$$\operatorname{Index}\left((d+\delta)^{E,\chi}\right) = \dim E \cdot \mathrm{GB}[TM][M] = \dim E \cdot \chi(M).$$

PROOF. We compute

$$\begin{aligned}
&\operatorname{Index}\left((d+\delta)^{E,\chi}\right) = \operatorname{Index}\left(\mathcal{D}^{W^e,E+}\right) + \operatorname{Index}\left(\left(\mathcal{D}^{W^o,E+}\right)^*\right) \\
&= \operatorname{Index}\left(\mathcal{D}^{W^e,E+}\right) - \operatorname{Index}\left(\mathcal{D}^{W^o,E+}\right) \\
&= \left(\mathbf{ch}(E) \smile \mathbf{ch}\left(\Lambda^{even,+}(M) \oplus \Lambda^{odd,-}(M)\right) \smile \widetilde{\mathbf{A}}(M)\right)[M] \\
&- \left(\mathbf{ch}(E) \smile \mathbf{ch}\left(\Lambda^{odd,+}(M) \oplus \Lambda^{even,-}(M)\right) \smile \widetilde{\mathbf{A}}(M)\right)[M] \\
&= \mathbf{ch}(E) \smile \begin{pmatrix} \mathbf{ch}\left(\Lambda^{even,+}(M)\right) - \mathbf{ch}\left(\Lambda^{even,-}(M)\right) \\ +\mathbf{ch}\left(\Lambda^{odd,-}(M)\right) - \mathbf{ch}\left(\Lambda^{odd,+}(M)\right) \end{pmatrix} \smile \widetilde{\mathbf{A}}(M)[M] \\
&= \left(\mathbf{ch}(E) \smile \mathrm{GB}(TM)\right)[M] = \left(ch_0(E) \smile \mathrm{GB}(TM)\right)[M] \\
&= \dim E \cdot \mathrm{GB}(TM)[M].
\end{aligned}$$

□

The Generalized Yang-Mills Index Theorem

Recall that in (18.135) we introduced the two generalized Dirac operators $\mathcal{D}^{W^e}$ and $\mathcal{D}^{W^o}$ and we used their twisted versions

$$\mathcal{D}^{W^e,E} = (d+\delta)^E \in \operatorname{End}\left(\Omega^{even,+}(E) \oplus \Omega^{odd,-}(E)\right)$$
$$\mathcal{D}^{W^o,E} = (d+\delta)^E \in \operatorname{End}\left(\Omega^{odd,+}(E) \oplus \Omega^{even,-}(E)\right) \tag{18.144}$$

in the proof of the Twisted Gauss-Bonnet Theorem. For reasons that will be explained below, we call $\mathcal{D}^{W^e,E}$ and $\mathcal{D}^{W^o,E}$ **Yang-Mills-Dirac operators**.

THEOREM 18.65 (The Yang-Mills-Dirac Index Theorem). *Let M be a compact, oriented Riemannian $2m$-manifold, where m is even, and let $E \to M$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then*

$$\operatorname{Index}\left(\mathcal{D}^{W^e,E+}\right) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(M)\right)[M] + \tfrac{1}{2}\dim E \cdot \chi(M) \quad \textit{and}$$
$$\operatorname{Index}\left(\mathcal{D}^{W^o,E+}\right) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(M)\right)[M] - \tfrac{1}{2}\dim E \cdot \chi(M).$$

PROOF. Using $(d+\delta)^{E,+} = \mathcal{D}^{W^e,+} \oplus \mathcal{D}^{W^o,+}$ and the above proof of Theorem 18.64, we have

$$\operatorname{Index}\left((d+\delta)^{E,+}\right) = \operatorname{Index}\left(\mathcal{D}^{W^e,+}\right) + \operatorname{Index}\left(\mathcal{D}^{W^o,+}\right)$$
$$\operatorname{Index}\left((d+\delta)^{E,\chi}\right) = \operatorname{Index}\left(\mathcal{D}^{W^e,+}\right) - \operatorname{Index}\left(\mathcal{D}^{W^o,+}\right).$$

The results follow from adding and subtracting, since

$$\operatorname{Index}\left((d+\delta)^{E,+}\right) = \left(\mathbf{ch}_2(E) \smile \mathbf{L}(M)\right)[M]$$
$$\operatorname{Index}\left((d+\delta)^{E,\chi}\right) = \dim E \cdot \chi(M),$$

by Theorems 18.60 (p.574) and 18.64. □

We now explain the "Yang-Mills-Dirac" nomenclature. In the proof of Theorem 17.15 (p.462), we computed the index of the operator

$$\mathcal{T} : \Omega^1(E) \to \Omega^0(E) \oplus \Omega^2_-(E), \text{ given by}$$
$$\mathcal{T}(\alpha) := \left(\delta^\omega \alpha, \tfrac{1}{2}(1-*)\, D^\omega \alpha\right),$$

where $E = P \times_G \mathfrak{g}_{\mathbb{C}}$ for some principal G-bundle over a compact Riemannian 4-manifold M and where $\mathfrak{g}_{\mathbb{C}}$ is the complexification of the Lie algebra of G. The kernel of $\mathcal{T}$ can be regarded as the formal dimension of tangent space (at ω) of the manifold of moduli of connections on P with self-dual curvature. The operator $\mathcal{T}$ bears a strong resemblence to the operator

$$\mathcal{D}^{W^o,E+} : \Omega^{odd,+}(E) \to \Omega^{even,-}(E),$$

which is the restriction of $(d+\delta)^E := D^\omega + \delta^\omega \in \operatorname{End}(\Omega^*(E))$. Indeed, for

$$\pi_\pm := \tfrac{1}{2}(1 \pm \tau) : \Omega^*(E) \to \Omega^{*,\pm}(E),$$

we have $\pi_- \circ \mathcal{T} = \mathcal{D}^{W^o,E} \circ \pi_+$. We have isomorphisms

$$\pi_+|_{\Omega^1(E)} : \Omega^1(E) \cong \Omega^{odd,+}(E) \subset \Omega^1(E) \oplus \Omega^3(E), \text{ and}$$
$$\pi_-|_{\Omega^0(E)\oplus\Omega^2_-(E)} : \Omega^0(E) \oplus \Omega^2_-(E) \cong \Omega^{even,-}(E).$$

Thus, using $\mathbf{ch}_2(E) = \dim E + 2ch_1(E) + 4ch_2(E)$, it follows that

$$\begin{aligned}
\operatorname{Index}(\mathcal{T}) &= \operatorname{Index}\left(\pi_-|_{\Omega^0(E)}\right)^{-1} \circ \mathcal{D}^{W^o,E} \circ \left(\pi_-|_{\Omega^0(E)} \oplus \Omega^2_-(E)\right) \\
&= \operatorname{Index}\left(\mathcal{D}^{W^o,E+}\right) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(M)\right)[M] - \tfrac{1}{2}\dim E \cdot \chi(M) \\
&= \tfrac{1}{2}\left(\left(\dim E + 2ch_1(E) + 4ch_2(E)\right) \smile \mathbf{L}(M)\right)[M] - \tfrac{1}{2}\dim E \cdot \chi(M) \\
&= \tfrac{1}{2}\left(4ch_2(E)[M] + \dim E \cdot \mathbf{L}(M)[M]\right) - \tfrac{1}{2}\dim E \cdot \chi(M) \\
&= 2ch_2(E)[M] - \tfrac{1}{2}\dim E \cdot \left(\chi(M) - \operatorname{sig}(M)\right), \qquad (18.145)
\end{aligned}$$

in agreement with computation in the proof of Theorem 17.15 (p.462).

The Hirzebruch-Riemann-Roch Formula

Let M be a *complex manifold* with $n = \dim M = 2\dim_{\mathbb{C}} M = 2m$. In other words, M is a smooth n-manifold, and there is a covering $\{U\}$ of M and a collection $\{\varphi_U\}$ of coordinate charts $\varphi_U : U \to \mathbb{C}^m$, such that $\varphi_V \circ \varphi_U^{-1} : \varphi_U(U \cap V) \to \varphi_V(U \cap V)$ is holomorphic (i.e., $\left(\varphi_V \circ \varphi_U^{-1}\right)_* : T_{\varphi_U(x)}\mathbb{C}^m \to T_{\varphi_V(p)}\mathbb{C}^m$ is complex linear for each $x \in U \cap V$). The tangent spaces T_xM then possess a well-defined map $J_x \in T_xM$ (with $J_x^2 = -\operatorname{Id}_x$) which corresponds to multiplication by $i = \sqrt{-1}$ under $(\varphi_U)_* : T_xM \to T_{\varphi_U(x)}\mathbb{C}^m \cong \mathbb{C}^m$ (i.e., $J_x(X) = (\varphi_U)_*^{-1}(i\varphi_U(X))$). The bundle automorphism $J \in \operatorname{End}(TM)$ is known as the *complex structure* of the complex manifold M. While it is tempting to explicitly make T_xM a complex vector space by defining iX to be JX for $X \in T_xM$, this ultimately leads to profound confusion, since it is customary (and of great utility) to consider the complexification of the real vector space T_xM, namely

$$(T_{\mathbb{C}}M)_x := \mathbb{C} \otimes T_xM = \{V + iW : V, W \in T_xM\}.$$

of complex dimension $2m$. The problem is that multiplication by i in $T_{\mathbb{C}}M$ is not the same as the complex linear extension of J to $T_{\mathbb{C}}M$. In particular while J preserves the real subspace $T_xM \subset (T_{\mathbb{C}}M)_x$, multiplication by i does not. Thus, to avoid confusion, we use J_x instead of i for the complex structure on T_xM. Let

$$\varphi_U(x) = \left(z^1(x), \ldots, z^m(x)\right) = \left(x^1(x) + iy^1(x), \ldots, x^m(x) + iy^m(x)\right).$$

We define $\mathbb{C}$-valued, $\mathbb{R}$-linear functionals dz^j and $d\bar{z}^j$ on T_xM via

$$\begin{aligned}
dz^j &:= dx^j + idy^j : T_xM \to \mathbb{C} \text{ and} \\
d\bar{z}^j &:= dx^j - idy^j : T_xM \to \mathbb{C}
\end{aligned}$$

Any $\mathbb{R}$-linear functional on T_xM, such as dz^k or $d\bar{z}^k$, extends uniquely to a $\mathbb{C}$-linear functional on $(T_{\mathbb{C}}M)_x$, and we use the same symbols to denote these extensions; i.e., dz^j, $d\bar{z}^j \in (T_{\mathbb{C}}M)_x^*$. The local complex vector fields (local sections of $T_{\mathbb{C}}M$)

$$\partial_{z^k} := \tfrac{1}{2}\left(\partial_{x^k} - i\partial_{y^k}\right), \ \partial_{\bar{z}^k} := \tfrac{1}{2}\left(\partial_{x^k} + i\partial_{y^k}\right)$$

are dual to dz^j and $d\bar{z}^j \in (T_{\mathbb{C}}M)^*$, in the sense that

$$\begin{aligned}
dz^j(\partial_{z^k}) &= \tfrac{1}{2}\left(dx^j + idy^j\right)\left(\partial_{x^k} - i\partial_{y^k}\right) = \delta^j_k \\
d\bar{z}^j(\partial_{\bar{z}^k}) &= \tfrac{1}{2}\left(dx^j - idy^j\right)\left(\partial_{x^k} + i\partial_{y^k}\right) = \delta^j_k \\
d\bar{z}^j(\partial_{z^k}) &= dz^j(\partial_{\bar{z}^k}) = 0.
\end{aligned}$$

There is complex-linear extension of J_x to $(T_{\mathbb{C}}M)_x$. We denote this extension by the same symbol J_x. Since $J_x^2 = -\operatorname{Id}$, the eigenvalues of J_x are i and $-i$, and the eigenspaces of $J_x \in \operatorname{End}((T_{\mathbb{C}}M)_x)$ are

$$T_x^{1,0}M := \{V - iJV : V \in T_xM\} \text{ and } T_x^{0,1}M := \{V + iJV : V \in T_xM\},$$

respectively. Note that $\{\partial_{z^1}, \ldots, \partial_{z^m}\}$ and $\{\partial_{\bar{z}^1}, \ldots, \partial_{\bar{z}^m}\}$ are local framings of $C^\infty\left(T_x^{1,0}M\right)$ and $C^\infty\left(T_x^{0,1}M\right)$, respectively. We set

$$\Lambda^{p,0}\left(T_{\mathbb{C}}M^*\right)_x := \text{the vector space of all anti-symmetric multi-complex-linear functionals defined on } T_x^{1,0}M \times \overset{p}{\cdots} \times T_x^{1,0}M.$$

The $\Lambda^{p,0}\left(T_{\mathbb{C}}M^*\right)_x$ are the fibers of a complex vector bundle $\Lambda^{p,0}\left(T_{\mathbb{C}}M^*\right) \to M$. Let $\Omega^{p,0}(M)$ be the space of C^∞ sections of $\Lambda^{p,0}\left(T_{\mathbb{C}}M^*\right)$; i.e.,

$$\Omega^{p,0}(M) := C^\infty\left(\Lambda^{p,0}\left(T_{\mathbb{C}}M^*\right)\right).$$

On a coordinate neighborhood U, such a section is of the form

$$\frac{1}{p!}\sum_{(j)} f_{j_1\cdots j_p} dz^{j_1} \wedge \cdots \wedge dz^{j_p},$$

where the $f_{j_1\cdots j_p} \in C^\infty(U, \mathbb{C})$ are antisymmetric in $j_1 \cdots j_p$. Similarly, we may define $\Lambda^{0,q}\left(T_{\mathbb{C}}M^*\right)_x$ and the bundles $\Lambda^{0,q}\left(T_{\mathbb{C}}M^*\right)$, and the space $\Omega^{0,q}(M) := C^\infty\left(\Lambda^{0,q}\left(T_{\mathbb{C}}M^*\right)\right)$ of sections which locally are of the form

$$\frac{1}{q!}\sum_{(k)} f_{k_1\cdots k_q} d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q}.$$

More generally, one has the bundles $\Lambda^{p,q}\left(T_{\mathbb{C}}M^*\right)$ whose sections in $\Omega^{p,q}(M) := C^\infty\left(\Lambda^{p,q}\left(T_{\mathbb{C}}M^*\right)\right)$ are locally of the form

$$\frac{1}{p!q!}\sum_{(j)(k)} f_{j_1\cdots j_p;k_1\cdots k_q} dz^{j_1} \wedge \cdots \wedge dz^{j_p} \wedge d\bar{z}^{k_1} \cdots \wedge d\bar{z}^{k_q}$$

and are called forms of bidegree (p,q). By writing $dz^j = dx^j + idy^j$ and $d\bar{z}^j = dx^j - idy^j$, we can regard such forms as ordinary forms in $\Omega^{p+q}(M, \mathbb{C})$. Conversely, writing $dx^j = \frac{1}{2}\left(dz^j + d\bar{z}^j\right)$ and $dy^j = \frac{1}{2i}\left(dz^j - d\bar{z}^j\right)$, we see that

$$\Lambda^l\left(TM^*, \mathbb{C}\right) := \mathbb{C} \otimes \Lambda^l\left(TM^*, \mathbb{R}\right) \cong \bigoplus\nolimits_{p+q=l} \Lambda^{p,q}\left(T_{\mathbb{C}}M^*\right) \text{ and}$$
$$\Omega^l(M, \mathbb{C}) := \mathbb{C} \otimes \Omega^l(M, \mathbb{R}) \cong \sum_{p+q=l} \Omega^{p,q}(M).$$

There is an operator $\bar{\partial} : \Omega^{0,q}(M) \to \Omega^{0,q+1}(M)$ given locally by

$$\bar{\partial}\left(\frac{1}{q!}\sum\nolimits_{(k)} f_{k_1\cdots k_q} d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q}\right)$$
$$:= \frac{1}{q!}\sum\nolimits_{(k)}\sum\nolimits_{k_0=1}^{m} \partial_{\bar{z}^{k_0}}\left(f_{k_1\cdots k_q}\right) d\bar{z}^{k_0} \wedge d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q}.$$

More generally, one analogously defines $\bar{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$, as well as $\partial : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M)$. While we have given these operators locally, they are independent of local holomorphic coordinates. The operator $\partial + \bar{\partial}$ is the restriction of the usual exterior derivative on $\Omega^{p,q}(M) \subset \Omega^{p+q}(M, \mathbb{C})$, namely

$$\partial + \bar{\partial} = d|_{\Omega^{p,q}(M)} : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M) + \Omega^{p,q+1}(M).$$

This is a consequence of the fact that, for $f \in C^{\infty}(M, \mathbb{C})$, locally we have

$$\begin{aligned}
\left(\partial + \bar{\partial}\right) f &= \sum_{k=1}^{m} \partial_{z^k}(f)\, dz^j + \partial_{\bar{z}^k}(f)\, d\bar{z}^j \\
&= \sum_{j=1}^{m} \tfrac{1}{2}\left(\partial_{x^k} f - i\partial_{y^k} f\right)\left(dx^k + i dy^k\right) + \tfrac{1}{2}\left(\partial_{x^k} f + i\partial_{y^k} f\right)\left(dx^k - i dy^k\right) \\
&= \sum_{j=1}^{m} \partial_{x^k}(f)\, dx^k + \partial_{y^k}(f)\, dy^k = df.
\end{aligned}$$

We have $\partial^2 = 0$, $\partial\bar{\partial} + \bar{\partial}\partial = 0$ and $\bar{\partial}^2 = 0$, since

$$0 = d^2 = \left(\partial + \bar{\partial}\right)^2 = \partial^2 \oplus \left(\partial\bar{\partial} + \bar{\partial}\partial\right) \oplus \bar{\partial}^2.$$

In particular, since $\bar{\partial}^2 = 0$, we have a chain complex

$$0 \to \Omega^{0,0}(M) \xrightarrow{\bar{\partial}} \Omega^{0,1}(M) \xrightarrow{\bar{\partial}} \cdots \xrightarrow{\bar{\partial}} \Omega^{0,m}(M)$$

and **Dolbeault cohomology spaces**

$$H^{0,q}(M) := \frac{\operatorname{Ker}\left(\bar{\partial}|_{\Omega^{0,q}(M)}\right)}{\bar{\partial}\left(\Omega^{0,q-1}(M)\right)}.$$

In order to define harmonic representatives of Dolbeault cohomology classes, we need a Hermitian metric to define an adjoint

$$\bar{\partial}^* : \Omega^{0,q+1}(M) \to \Omega^{0,q}(M),$$

for $\bar{\partial} : \Omega^{0,q}(M) \to \Omega^{0,q+1}(M)$. Suppose that we are given a Riemannian metric h on M, so that for all all $x \in M$ and $V, W \in T_x M$, we have

$$h(JV, W) = -h(V, JW), \text{ or equivalently } h(JV, JW) = h(V, W).$$

Such can always found by setting $h(V, W) = h_0(JV, JW) + h_0(V, W)$ for an arbitrary metric h_0. Then h_x uniquely extends to a complex bilinear form $h_{\mathbb{C}}$ on $\mathbb{C} \otimes_{\mathbb{R}} T_x M$, so that for $V_1, V_2, W_1, W_2 \in T_x M$

$$\begin{aligned}
&h_{\mathbb{C}}(V_1 + iV_2, W_1 + iW_2) \\
&:= h(V_1, W_1) - h(V_2, W_2) + i\left(h(V_1, W_2) + h(V_2, W_1)\right).
\end{aligned}$$

We have

$$\begin{aligned}
&h_{\mathbb{C}}(V \pm iJV, W \pm iJW) \\
&= h(V_1, W_1) - h(JV, JW) \pm i\left(h(V, JW) + h(JV, W)\right) = 0.
\end{aligned}$$

Thus, the restrictions of $h_{\mathbb{C}}$ to $T^{1,0}M$ and $T^{0,1}M$ are 0. For

$$\left(h_{\mathbb{C}}\right)_{j\bar{k}} := h_{\mathbb{C}}\left(\partial_{z^j}, \partial_{\bar{z}^k}\right) = h_{\mathbb{C}}\left(\partial_{\bar{z}^k}, \partial_{z^j}\right) =: \left(h_{\mathbb{C}}\right)_{\bar{k}j},$$

we have $\left(h_{\mathbb{C}}\right)_{\bar{j}k} = \overline{\left(h_{\mathbb{C}}\right)_{j\bar{k}}}$ and (where $\otimes_s$denotes symmetric tensor product)

$$\begin{aligned}
h_{\mathbb{C}} &= \sum\nolimits_{j,k} \left(h_{\mathbb{C}}\right)_{j\bar{k}}\, dz^j \otimes d\bar{z}^k + \left(h_{\mathbb{C}}\right)_{\bar{k}j}\, d\bar{z}^k \otimes dz^j \\
&= 2\sum\nolimits_{j,k} \left(h_{\mathbb{C}}\right)_{j\bar{k}}\, \tfrac{1}{2}\left(dz^j \otimes d\bar{z}^k + d\bar{z}^k \otimes dz^j\right) = 2\sum\nolimits_{j,k} \left(h_{\mathbb{C}}\right)_{j\bar{k}}\, dz^j \otimes_s d\bar{z}^k.
\end{aligned}$$

DEFINITION 18.66. The **Kähler 2-form** $\kappa \in \Omega^2(M,\mathbb{C})$ on a complex manifold with Riemannian metric h with $h(JV,JW) = h(V,W)$ is given by

$$\kappa(X,Y) := h_{\mathbb{C}}(JX,Y).$$

Note that $\kappa \in \Omega^{1,1}(M)$, since locally

$$\kappa(\partial_{z^j},\partial_{\bar z^k}) = h_{\mathbb{C}}(J\partial_{z^j},\partial_{\bar z^k}) = h_{\mathbb{C}}(i\partial_{z^j},\partial_{\bar z^k}) = i(h_{\mathbb{C}})_{j\bar k}, \text{ and}$$
$$\kappa(\partial_{\bar z^j},\partial_{\bar z^k}) = \kappa(\partial_{z^j},\partial_{z^k}) = 0 \Rightarrow \kappa = i\sum\nolimits_{j,k}(h_{\mathbb{C}})_{j\bar k}\,dz^j\wedge d\bar z^k.$$

If $\left(\partial_{x^1},\partial_{y^1},\ldots,\partial_{x^m},\partial_{y^m}\right)$ is orthonormal at $x\in M$, then

$$\begin{aligned}(h_{\mathbb{C}})_{j\bar k} &= h_{\mathbb{C}}(\partial_{z^j},\partial_{\bar z^k}) = h_{\mathbb{C}}\left(\tfrac12\left(\partial_{x^j}-i\partial_{y^j}\right),\tfrac12\left(\partial_{x^k}+i\partial_{y^k}\right)\right)\\ &= \tfrac14 h(\partial_{x^j},\partial_{x^k}) + \tfrac14 h\left(\partial_{y^j},\partial_{y^k}\right) = \tfrac12\delta_{jk}.\end{aligned}$$

Thus, for example, $(h_{\mathbb{C}})_{k\bar k} = \frac12 h(\partial_{x^k},\partial_{x^k}) = \frac12 h\left(\partial_{y^k},\partial_{y^k}\right) = \frac12$, which is why we use the notation $(h_{\mathbb{C}})_{j\bar k}$, instead of simply $h_{j\bar k}$. Perhaps the bar over one of the indices suffices to avoid any confusion in practice, but there remain oddities, such as $0 = (h_{\mathbb{C}})_{11} \neq h_{11} = h(\partial_{x^1},\partial_{x^1}) = 1$, which would yield contradictions if we were to denote $(h_{\mathbb{C}})_{11}$ simply by h_{11}.

If $\left(\partial_{x^1},\partial_{y^1},\ldots,\partial_{x^m},\partial_{y^m}\right)$ is orthonormal at $x\in M$, then at x

$$\begin{aligned}\kappa &= i\sum\nolimits_{j,k=1}^m (h_{\mathbb{C}})_{j\bar k}\,dz^j\wedge d\bar z^k = i\sum\nolimits_{k=1}^m (h_{\mathbb{C}})_{k\bar k}\,dz^k\wedge d\bar z^k\\ &= \frac{i}{2}\sum\nolimits_{k=1}^m dz^k\wedge d\bar z^k = \frac{i}{2}\sum\nolimits_{k=1}^m\left(dx^k+i\,dy^k\right)\wedge\left(dx^k-i\,dy^k\right)\\ &= \sum\nolimits_{k=1}^m dx^k\wedge dy^k, \text{ and}\end{aligned}$$

$$\begin{aligned}\wedge^m\kappa &= \left(\sum\nolimits_{k_1=1}^m dx^{k_1}\wedge dy^{k_1}\right)\wedge\cdots\wedge\left(\sum\nolimits_{k_m=1}^m dx^{k_m}\wedge dy^{k_m}\right)\\ &= m!\,dx^1\wedge dy^1\wedge\cdots\wedge dx^m\wedge dy^m\end{aligned}$$

Moreover, according to (18.139), we have

$$\nu_h = dx^1\wedge dy^1\wedge\cdots\wedge dx^m\wedge dy^m = 2^{-m}i^{m^2}dz^1\wedge\cdots\wedge dz^m\wedge d\bar z^1\wedge\cdots\wedge d\bar z^m \tag{18.146}$$

If $\left(\partial_{x^1},\partial_{y^1},\ldots,\partial_{x^m},\partial_{y^m}\right)$ is not necessarily orthonormal, we still have

$$\nu_h = \tfrac{1}{m!}\wedge^m\kappa = \tfrac{i^m}{m!}\wedge^m\left(\sum\nolimits_{j,k}(h_{\mathbb{C}})_{j\bar k}\,dz^j\wedge d\bar z^k\right).$$

Define a Hermitian metric H_x (complex linear in first slot and conjugate linear in the second slot) on $\mathbb{C}\otimes_{\mathbb{R}} T_xM$ by

$$H(V,W) := (h_{\mathbb{C}})\left(V,\overline{W}\right). \tag{18.147}$$

Note that $H|_{T_x^{1,0}M}$ and $H|_{T_x^{0,1}M}$ are nondegenerate and relative to H, $T_x^{1,0}M\perp T_x^{0,1}M$. Indeed, for $V,W\in T_xM$, we have

$$\begin{aligned}H(V\pm iJV,W\pm iJW) &= H(V,W)+H(V,\pm iJW)+H(\pm iJV,W)+H(iJV,iJW)\\ &= H(V,W)+H(JV,JW) = 2h(V,W)\\ H(V\pm iJV,W\mp iJW) &= H(V,W)+H(V,\mp iJW)+H(\pm iJV,W)-H(iJV,iJW)\\ &= \pm i\left(h(V,JW)+h(JV,W)\right) = 0.\end{aligned}$$

We set

$$\begin{aligned}
H_{jk} &:= H\left(\partial_{z^j}, \partial_{z^k}\right) = \left(h_{\mathbb{C}}\right)\left(\partial_{z^j}, \partial_{\bar{z}^k}\right) = \left(h_{\mathbb{C}}\right)_{j\bar{k}} \\
H_{j\bar{k}} &:= H\left(\partial_{z^j}, \partial_{\bar{z}^k}\right) = \left(h_{\mathbb{C}}\right)\left(\partial_{z^j}, \partial_{z^k}\right) = \left(h_{\mathbb{C}}\right)_{jk} = 0 \\
H_{\bar{k}j} &:= H\left(\partial_{\bar{z}^k}, \partial_{z^j}\right) = \left(h_{\mathbb{C}}\right)\left(\partial_{\bar{z}^k}, \partial_{\bar{z}^j}\right) = \left(h_{\mathbb{C}}\right)_{\bar{k}\bar{j}} = 0 \text{ and} \\
H_{\bar{j}\bar{k}} &:= H\left(\partial_{\bar{z}^j}, \partial_{\bar{z}^k}\right) = \left(h_{\mathbb{C}}\right)\left(\partial_{\bar{z}^j}, \partial_{z^k}\right) = \left(h_{\mathbb{C}}\right)_{\bar{j}k}.
\end{aligned}$$

In particular, if $\left(\partial_{x^1}, \partial_{y^1}, \ldots, \partial_{x^m}, \partial_{y^m}\right)$ is orthonormal at $x \in M$, then

$$\begin{aligned}
H_{jk} &= H\left(\partial_{z^j}, \partial_{z^k}\right) = \left(h_{\mathbb{C}}\right)_{j\bar{k}} = \tfrac{1}{2}\delta_{jk} = \left(h_{\mathbb{C}}\right)_{\bar{j}k} = H\left(\partial_{\bar{z}^j}, \partial_{\bar{z}^k}\right) = H_{\bar{j}\bar{k}} \\
&\Rightarrow \left(\sqrt{2}\partial_{z^1}, \ldots, \sqrt{2}\partial_{z^m}, \sqrt{2}\partial_{\bar{z}^1}, \ldots, \sqrt{2}\partial_{\bar{z}^m}\right)
\end{aligned}$$

is orthonormal for H.

There is a conjugate-linear bundle isomorphism

$$\flat : \left(T_{\mathbb{C}}M\right)_x \to \left(T_{\mathbb{C}}M\right)_x^*$$

given, for $V, W \in \left(T_{\mathbb{C}}M\right)_x$, by

$$\flat_x(V)(W) = H(W, V).$$

We have $\flat\left(T^{1,0}M\right) = \Lambda^{1,0}\left(TM^*\right)$ and $\flat\left(T^{0,1}M\right) = \Lambda^{0,1}\left(TM^*\right)$. Indeed,

$$\flat\left(\partial_{z^k}\right) = \sum\nolimits_{l=1}^{m} H_{lk}dz^l \text{ and } \flat\left(\partial_{\bar{z}^k}\right) = \sum\nolimits_{l=1}^{m} H_{\bar{l}\bar{k}}d\bar{z}^l, \text{ since}$$

$$\begin{aligned}
\flat\left(\partial_{z^k}\right)\left(\partial_{z^j}\right) &= H\left(\partial_{z^j}, \partial_{z^k}\right) = H_{jk} = \sum\nolimits_{l=1}^{m} H_{lk}dz^l\left(\partial_{z^j}\right) \\
\flat\left(\partial_{z^k}\right)\left(\partial_{\bar{z}^j}\right) &= H\left(\partial_{\bar{z}^j}, \partial_{z^k}\right) = H_{\bar{j}k} = 0 \\
\flat\left(\partial_{\bar{z}^k}\right)\left(\partial_{\bar{z}^j}\right) &= H\left(\partial_{\bar{z}^j}, \partial_{\bar{z}^k}\right) = H_{\bar{j}\bar{k}} = \sum\nolimits_{\bar{l}=1}^{m} H_{\bar{l}\bar{k}}d\bar{z}^l\left(\partial_{\bar{z}^j}\right) \\
\flat\left(\partial_{\bar{z}^k}\right)\left(\partial_{z^j}\right) &= H\left(\partial_{z^j}, \partial_{\bar{z}^k}\right) = H_{j\bar{k}} = 0.
\end{aligned}$$

In particular, if $\left(\partial_{x^1}, \partial_{y^1}, \ldots, \partial_{x^m}, \partial_{y^m}\right)$ is orthonormal for h at $x \in M$, then $\flat\left(\partial_{z^k}\right) = \frac{1}{2}dz^k$ and $\flat\left(\partial_{\bar{z}^k}\right) = \frac{1}{2}d\bar{z}^k$. There is a conjugate-linear inverse to $\flat$, denoted by

$$\# : \left(T_{\mathbb{C}}M\right)_x^* \to \left(T_{\mathbb{C}}M\right)_x$$

Now H on $\left(T_{\mathbb{C}}M\right)_x$ induces a Hermitian inner product (still denoted by H) on $\Lambda^1\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$, given for $\alpha, \beta \in \Lambda^1\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$, by

$$H(\alpha, \beta) = H(\#\beta, \#\alpha),$$

where the switch in the order is due to the conjugate linearity of $\#$. We then have an induced Hermitian inner product on $\Lambda^k\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$ for $1 \leq k \leq m$, where $\left\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 < \cdots < i_k\right\}$ is an orthonormal basis for $\Lambda^k\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$ if $e_1, \ldots, e_m$ is an orthonormal basis of $\Lambda^1\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$. By restriction, we have Hermitian inner products on $\Lambda^{p,q}\left(\left(T_{\mathbb{C}}M\right)_x^*\right)$ as well.

Let $(\cdot, \cdot)$ be the Hermitian L^2 inner product on $\Omega^{0,q}(M)$ induced by H, namely

$$(\alpha, \beta) := \int_M H(\alpha, \beta)\, v_h, \text{ with } \|\alpha\| := \sqrt{(\alpha, \alpha)}.$$

We may then speak of the formal adjoint $\bar{\partial}^* : \Omega^{p,q+1}(M) \to \Omega^{p,q}(M)$ of $\bar{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$, having the property

$$\left(\bar{\partial}\alpha, \beta\right) = \left(\alpha, \bar{\partial}^*\beta\right).$$

In order to exhibit a formula for $\bar{\partial}^*$, we have a $\mathbb{C}$-linear star operator

$$* : \Omega^k(M,\mathbb{C}) \to \Omega^{n-k}(M,\mathbb{C}) \text{ determined by}$$
$$\alpha \wedge *\beta := h_{\mathbb{C}}(\alpha,\beta)\, v_h \text{ for } \alpha,\beta \in \Omega^k(M,\mathbb{C}).$$

Some authors (e.g., [**GH**]) and use H instead of $h_{\mathbb{C}}$ in the definition of $*$, which makes $*$ conjugate linear, while we, as others (e.g., [**Go**]), use a $\mathbb{C}$-linear star operator. Thus, here

$$\alpha \wedge *\bar{\beta} := g_{\mathbb{C}}\left(\alpha,\bar{\beta}\right) v_h = H(\alpha,\beta)\, v_h.$$

Since $\dim_{\mathbb{R}} M = 2m$ is even, $*^2|_{\Omega^k(M,\mathbb{C})} = (-1)^k \,\mathrm{Id}$ and the formal adjoint of $d : \Omega^k(M,\mathbb{R}) \to \Omega^{k+1}(M,\mathbb{R})$ is $\delta := -*d*$. By the $\mathbb{C}$-linearity of $*$ and the fact that d is the $\mathbb{C}$-linear extension of $\partial + \bar{\partial}$, δ has the $\mathbb{C}$-linear extension

$$\delta = -*d* = -*\left(\partial + \bar{\partial}\right)* = (-*\partial*) \oplus \left(-*\bar{\partial}*\right)$$

By (18.140), we have $*\Omega^{p,q}(M) = \Omega^{m-q,m-p}(M)$, and so

$$\begin{aligned}
\left(-*\bar{\partial}*\right)\left(\Omega^{p+1,q}(M)\right) &\subseteq \left(-*\bar{\partial}\right)\left(\Omega^{m-q,m-(p+1)}(M)\right)\\
&\subseteq (-*)\left(\Omega^{m-q,m-p}(M)\right) \subseteq \Omega^{p,q}(M), \text{ while}\\
(-*\partial*)\left(\Omega^{p+1,q}(M)\right) &\subseteq (-*\partial)\left(\Omega^{m-q,m-p-1}(M)\right)\\
&\subseteq (-*)\left(\Omega^{m-q+1,m-p-1}(M)\right) \subseteq \Omega^{p+1,q-1}(M).
\end{aligned}$$

Thus, for $\alpha \in \Omega^{p,q}(M)$ and $\beta \in \Omega^{p+1,q}(M)$, we have

$$\begin{aligned}
&H(d\alpha,\beta) = H\left(\partial\alpha + \bar{\partial}\alpha,\beta\right) = H(\partial\alpha,\beta)\\
&H(\alpha,\delta\beta) = H\left(\alpha, -*\partial*\beta - *\bar{\partial}*\beta\right) = H\left(\alpha,-*\bar{\partial}*\beta\right), \text{ and so}
\end{aligned}$$

$$\begin{aligned}
\int_M H(\partial\alpha,\beta)\, v_h &= \int_M H(d\alpha,\beta)\, v_h = \int_M h_{\mathbb{C}}\left(d\alpha,\bar{\beta}\right) v_h\\
&= \int_M h_{\mathbb{C}}\left(\alpha,\delta\bar{\beta}\right) v_h = \int_M h_{\mathbb{C}}\left(\alpha,\overline{\delta\beta}\right) v_h\\
&= \int_M H(\alpha,\delta\beta)\, v_h = \int_M H\left(\alpha,\left(-*\bar{\partial}*\right)\beta\right) v_h.
\end{aligned}$$

Similarly,

$$\int_M H\left(\bar{\partial}\alpha,\beta\right) v_h = \int_M H(\alpha,(-*\partial*)\beta)\, v_h.$$

Thus, the formal adjoints of ∂ and $\bar{\partial}$ are given by

$$\partial^* = -*\bar{\partial}* \text{ and } \bar{\partial}^* = -*\partial*.$$

We define the space of **harmonic** $(0,q)$-forms by

$$\mathcal{H}^{0,q}(M) := \left\{\alpha \in \Omega^{0,q}(M) : \left(\bar{\partial} + \bar{\partial}^*\right)\alpha = 0\right\}.$$

THEOREM 18.67 (Hodge Decomposition). *There is an orthogonal decomposition*

$$\begin{aligned}
\Omega^{0,q}(M) &= \mathcal{H}^{0,q}(M) \oplus \bar{\partial}\left(\Omega^{0,q-1}(M)\right) \oplus \bar{\partial}^*\left(\Omega^{0,q+1}(M)\right)\\
&= \mathcal{H}^{0,q}(M) \oplus \left(\bar{\partial}\bar{\partial}^*\right)\left(\Omega^{0,q}(M)\right) \oplus \left(\bar{\partial}^*\bar{\partial}\right)\left(\Omega^{0,q}(M)\right).
\end{aligned} \tag{18.148}$$

PROOF. The operator

$$\Delta_{\bar{\partial}} := \left(\bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}\right) = \left(\bar{\partial} + \bar{\partial}^*\right)^2 : \Omega^{0,q}(M) \to \Omega^{0,q}(M)$$

is elliptic, and hence we have the orthogonal decomposition

$$\Omega^{0,q}(M) = \operatorname{Ker}\Delta_{\bar{\partial}} \oplus \Delta_{\bar{\partial}}\left(\Omega^{0,q}(M)\right).$$

Moreover, $\operatorname{Ker}\Delta_{\bar{\partial}} = \mathcal{H}^{0,q}(M)$. Indeed,

$$\begin{aligned}\alpha \in \mathcal{H}^{0,q}(M) &\Rightarrow \left(\bar{\partial} + \bar{\partial}^*\right)\alpha = 0 \\ &\Rightarrow \bar{\partial}\alpha = 0 \text{ and } \bar{\partial}^*\alpha = 0 \Rightarrow \left(\bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}\right)\alpha = 0,\end{aligned}$$

and conversely if $\alpha \in \operatorname{Ker}\Delta_{\bar{\partial}}$, then

$$0 = \left(\left(\bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}\right)\alpha, \alpha\right) = \left\|\bar{\partial}\alpha\right\|^2 + \left\|\bar{\partial}\alpha\right\|^2 \Rightarrow \bar{\partial}\alpha + \bar{\partial}^*\alpha = 0.$$

It remains to prove that we have an orthogonal decompostion

$$\Delta_{\bar{\partial}}\left(\Omega^{0,q}(M)\right) = \bar{\partial}\left(\Omega^{0,q-1}(M)\right) \oplus \bar{\partial}^*\left(\Omega^{0,q+1}(M)\right). \tag{18.149}$$

The summands are orthogonal since

$$\beta \in \Omega^{0,q-1}(M) \text{ and } \gamma \in \Omega^{0,q+1}(M) \Rightarrow \left(\bar{\partial}\beta, \bar{\partial}^*\gamma\right) = \left(\bar{\partial}^2\beta, \gamma\right) = 0.$$

Moreover, for any $\alpha \in \Omega^{0,q}(M)$, we have

$$\Delta_{\bar{\partial}}\alpha = \bar{\partial}\left(\bar{\partial}^*\alpha\right) + \bar{\partial}^*\left(\bar{\partial}\alpha\right) \in \bar{\partial}\left(\Omega^{0,q-1}(M)\right) \oplus \bar{\partial}^*\left(\Omega^{0,q+1}(M)\right), \tag{18.150}$$

$$\text{and so } \Delta_{\bar{\partial}}\left(\Omega^{0,q}(M)\right) \subseteq \bar{\partial}\left(\Omega^{0,q-1}(M)\right) \oplus \bar{\partial}^*\left(\Omega^{0,q+1}(M)\right).$$

For the reverse inclusion, note that $\bar{\partial}\left(\Omega^{0,q-1}(M)\right)$ and $\bar{\partial}^*\left(\Omega^{0,q-1}(M)\right)$ are both in $\mathcal{H}^{0,q}(M)^{\perp}$, since

$$\alpha \in \mathcal{H}^{0,q}(M) \Rightarrow \left(\bar{\partial}\beta, \alpha\right) = \left(\beta, \bar{\partial}^*\alpha\right) = 0 = \left(\gamma, \bar{\partial}\alpha\right) = \left(\bar{\partial}^*\gamma, \alpha\right),$$

$$\text{and so } \bar{\partial}\left(\Omega^{0,q-1}(M)\right) \oplus \bar{\partial}^*\left(\Omega^{0,q+1}(M)\right) \subseteq \mathcal{H}^{0,q}(M)^{\perp} = \Delta_{\bar{\partial}}\left(\Omega^{0,q}(M)\right).$$

Note that (18.150) and (18.149) then give the second equality of (18.148). □

COROLLARY 18.68. *Suppose that $\bar{\partial}\gamma = 0$ for some $\gamma \in \Omega^{0,q}(M)$. There is a unique $\alpha \in \mathcal{H}^{0,q}(M)$, such that for some $\beta \in \Omega^{0,q-1}(M)$, $\gamma = \alpha + \bar{\partial}\beta$. In other words, every cohomology class in $H^{0,q}(M)$ has a unique harmonic representative.*

PROOF. Theorem 18.67 yields a unique $\alpha \in \mathcal{H}^{0,q}(M)$ such that

$$\gamma = \alpha + \bar{\partial}\beta + \bar{\partial}^*\beta'$$

for some $\beta \in \Omega^{0,q-1}(M)$ and $\beta' \in \Omega^{0,q+1}(M)$. Now

$$\begin{aligned}0 = \bar{\partial}\gamma &= \bar{\partial}\alpha + \bar{\partial}^2\beta + \bar{\partial}\bar{\partial}^*\beta' = \bar{\partial}\bar{\partial}^*\beta' \\ &\Rightarrow \left(\bar{\partial}\bar{\partial}^*\beta', \beta'\right) = 0 \Rightarrow \left\|\bar{\partial}^*\beta'\right\|^2 = 0 \Rightarrow \bar{\partial}^*\beta' = 0.\end{aligned}$$

□

Suppose that ∇ is the covariant derivative for the Levi-Civita connection of the Riemannian metric h. We have seen that it is always possible to choose coordinates about a point $x \in M$ such that the coordinate vector fields have vanishing ∇-derivatives at x. However, it is not necessarily the case that such coordinates can be chosen of the form $\left(x^1, y^1, \ldots, x^m, y^m\right)$, where $\left(z^1, \ldots, z^m\right) = \left(x^1 + iy^1, \ldots, x^m + iy^m\right)$ is a complex-analytic coordinate chart. If for each $x \in M$ such coordinates can be found, then the complex manifold M with Riemannian metric h is a Kähler manifold. While one can take this to be the definition of a Kähler manifold, usually one of the other equivalent conditions in the following theorem is taken to be the definition.

THEOREM 18.69. *Let M be a complex manifold with complex structure J, and Riemannian metric h, with Levi-Civita covariant derivative ∇. The following are equivalent.*

I. About each $x \in M$, there is a complex chart $\left(x^1 + iy^1, \ldots, x^m + iy^m\right)$, such that $\nabla\left(\partial_{x^i}\right) = \nabla\left(\partial_{y^i}\right) = 0$ at x.

II. $\nabla J = 0$ (i.e., $(\nabla J)(X) = J(\nabla X) - \nabla(J(X)) = 0$).

III. The 2-form Kähler 2-form $\kappa \in \Omega^2(M, \mathbb{R})$, given by $\kappa(X, Y) := h(X, JY)$, is closed (i.e., $d\kappa = 0$).

PROOF. Assuming I, at any point x we have II, since

$$
\begin{aligned}
&(\nabla J)\left(\partial_{x^k}\right) = J\left(\nabla \partial_{x^k}\right) - \nabla\left(J\left(\partial_{x^k}\right)\right) = 0 - \nabla\left(\partial_{y^k}\right) = 0 \text{ and} \\
&(\nabla J)\left(\partial_{y^k}\right) = J\left(\nabla \partial_{y^k}\right) - \nabla\left(J\left(\partial_{y^k}\right)\right) = 0 + \nabla\left(\partial_{x^k}\right) = 0.
\end{aligned}
$$

that induced Hermitian metric H on $\mathbb{C} \otimes_{\mathbb{R}} TM$ defined by (18.147), and Levi-Civita covariant derivative ∇. We now show that II $\Rightarrow$ III. For vector fields X, Y, Z, we have

$$
\begin{aligned}
3(d\kappa)(X, Y, Z) &= X[\kappa(Y, Z)] + Y[\kappa(Z, X)] + Z[\kappa(X, Y)] \\
&- \kappa([X, Y], Z) - \kappa([Y, Z], X) - \kappa([Z, X], Y).
\end{aligned}
$$

If X, J, Z are coordinate vector fields, then (since ∇ is torion-free)

$$
\nabla_X Y - \nabla_Y X = [X, Y] = 0, \text{ and so}
$$

$$
\begin{aligned}
&3(d\kappa)(X, Y, Z) = X[\kappa(Y, Z)] + Y[\kappa(Z, X)] + Z[\kappa(X, Y)] \\
&= X[h(Y, JZ)] + Y[h(Z, JX)] + Z[h(X, JY)] \\
&= h(\nabla_X Y, JZ) + h(Y, \nabla_X(JZ)) + h(\nabla_Y Z, JX) + h(Z, \nabla_Y(JX)) \\
&+ h(\nabla_Z X, JY) + h(X, \nabla_Z(JY)) \\
&= h(\nabla_X Y, JZ) - h(\nabla_X Z, JY) + h(\nabla_Y Z, JX) - h(\nabla_Y X, JZ) \\
&+ h(\nabla_Z X, JY) - h(\nabla_Z Y, JX) \\
&= h(\nabla_X Y - \nabla_Y X, JZ) + h(\nabla_Z X - \nabla_X Z, JY) + h(\nabla_Y Z - \nabla_Z Y, JX) = 0.
\end{aligned}
$$

To show that III $\Rightarrow$ I, we note that by a $\mathbb{C}$-linear change of complex coordinates about x, we may assume (since h is positive definite) that

$$
\begin{aligned}
h_{\mathbb{C}} &= 2\sum\nolimits_{j,k} \left(h_{\mathbb{C}}\right)_{j\bar{k}}\, dz^j \otimes_s d\bar{z}^k \\
&= \sum\nolimits_{j,k,l} \left(\delta_{jk} + \sum\nolimits_l \left(a_{j\bar{k}l} z^l + a_{j\bar{k}\bar{l}} \bar{z}^l\right) + \mathrm{O}\left(\|z\|^2\right)\right) dz^j \otimes_s d\bar{z}^k,
\end{aligned}
$$

where $\|z\|^2 = \sum_{j=1}^m z^j \bar{z}^j$ and

$$
\begin{aligned}
&\overline{\left(h_{\mathbb{C}}\right)_{j\bar{k}}} = \left(h_{\mathbb{C}}\right)_{\bar{j}k} = \left(h_{\mathbb{C}}\right)_{k\bar{j}} \\
&\Rightarrow \sum\nolimits_l \overline{a_{j\bar{k}l}} \bar{z}^l + \overline{a_{j\bar{k}\bar{l}}} z^l = \sum\nolimits_l \overline{a_{j\bar{k}l} z^l + a_{j\bar{k}\bar{l}} \bar{z}^l} = \sum\nolimits_l \left(a_{k\bar{j}l} z^l + a_{k\bar{j}\bar{l}} \bar{z}^l\right)
\end{aligned}
$$

$$
\Rightarrow \overline{a_{j\bar{k}\bar{l}}} = a_{k\bar{j}l} \text{ and } \overline{a_{j\bar{k}\bar{l}}} = a_{k\bar{j}l}. \tag{18.151}
$$

Since

$$\begin{aligned}\kappa\left(\partial_{z^j},\partial_{\bar z^k}\right)&=h_{\mathbb{C}}\left(\partial_{z^j},J\partial_{\bar z^k}\right)=h_{\mathbb{C}}\left(\partial_{z^j},-i\partial_{\bar z^k}\right)\\&=-ih_{\mathbb{C}}\left(\partial_{z^j},\partial_{\bar z^k}\right)=-i\left(h_{\mathbb{C}}\right)_{j\bar k},\text{ and}\end{aligned}$$

similarly $\kappa\left(\partial_{z^j},\partial_{z^k}\right)=\kappa\left(\partial_{\bar z^j},\partial_{\bar z^k}\right)=0$, we have

$$\kappa=\tfrac{1}{2}\sum\nolimits_{j,k=1}^{m}\kappa\left(\partial_{z^j},\partial_{\bar z^k}\right)dz^j\wedge d\bar z^k=-\tfrac{i}{2}\sum\nolimits_{j,k=1}^{m}\left(h_{\mathbb{C}}\right)_{j\bar k}dz^j\wedge d\bar z^k.$$

Thus,

$$\begin{aligned}0=d\kappa&=-\tfrac{i}{2}\sum\nolimits_{j,k=1}^{m}\left(\partial_{z^l}\left(h_{\mathbb{C}}\right)_{j\bar k}dz^l+\partial_{\bar z^l}\left(h_{\mathbb{C}}\right)_{j\bar k}d\bar z^l\right)\wedge dz^j\wedge d\bar z^k\\&=-\tfrac{i}{2}\sum\nolimits_{j,k=1}^{m}\left(a_{j\bar kl}dz^l+a_{j\bar k\bar l}d\bar z^l\right)\wedge dz^j\wedge d\bar z^k\\&\Rightarrow a_{j\bar kl}=a_{l\bar kj}\text{ and }a_{j\bar k\bar l}=a_{j\bar l\bar k}.\end{aligned}$$

For $b^k_{pq}=b^k_{qp}\in\mathbb{C}$, let

$$\begin{aligned}z^k&=w^k+\tfrac{1}{2}\sum\nolimits_{p,q=1}^{m}b^k_{pq}w^pw^q\\dz^k&=dw^k+\sum\nolimits_{p,q=1}^{m}b^k_{pq}w^qdw^p\end{aligned}$$

Then, modulo $\mathrm{O}\left(\|w\|^2\right)$ or $\mathrm{O}\left(\|z\|^2\right)$ terms,

$$\begin{aligned}h_{\mathbb{C}}&=\sum\nolimits_{j,k}\left(\delta_{jk}+a_{j\bar kl}z^l+a_{j\bar k\bar l}\bar z^l\right)dz^j\otimes_s d\bar z^k\\&=\sum\nolimits_{j,k}\left(\delta_{jk}+\sum\nolimits_l\left(a_{j\bar kl}w^l+a_{j\bar k\bar l}\bar w^l\right)\right)\\&\left(dw^j+\sum\nolimits_{p,q=1}^{m}b^j_{pq}w^qdw^p\right)\otimes_s\left(d\bar w^k+\sum\nolimits_{p,q=1}^{m}\overline{b^k_{p'q'}}\bar w^{q'}d\bar w^{p'}\right)\\&=\sum\nolimits_{j,k}\left(\delta_{jk}+\sum\nolimits_l\left(a_{j\bar kl}w^l+a_{j\bar k\bar l}\bar w^l\right)\right)\cdot\\&\left(\begin{array}{c}dw^j\otimes_s d\bar w^k+\sum_{p,q=1}^{m}b^j_{pq}w^qdw^p\otimes_s d\bar w^k\\+\sum_{p,q=1}^{m}b^k_{p'q'}\bar w^{q'}dw^j\otimes_s d\bar w^{p'}\end{array}\right)\\&=\sum\nolimits_{j,k}\left(\delta_{jk}+\sum\nolimits_l\left(a_{j\bar kl}w^l+a_{j\bar k\bar l}\bar w^l\right)\right)dw^j\otimes_s d\bar w^k\\&+\sum\nolimits_{j,k}\delta_{jk}\sum\nolimits_{p,q=1}^{m}b^j_{pq}w^qdw^p\otimes_s d\bar w^k\\&+\sum\nolimits_{j,k}\delta_{jk}\sum\nolimits_{p,q=1}^{m}\overline{b^k_{p'q'}}\bar w^{q'}dw^j\otimes_s d\bar w^{p'}\\&=\sum\nolimits_{j,k}\left(\delta_{jk}+\sum\nolimits_l\left(a_{j\bar kl}w^l+a_{j\bar k\bar l}\bar w^l\right)\right)dw^j\otimes_s d\bar w^k\\&+\sum\nolimits_{j,k,l=1}^{m}b^k_{jl}w^ldw^j\otimes_s d\bar w^k+\sum\nolimits_{j,k,l=1}^{m}\overline{b^j_{kl}}\bar w^ldw^j\otimes_s d\bar w^k\\&=\sum\nolimits_{j,k}\left(\delta_{jk}+\sum\nolimits_l\left(\left(a_{j\bar kl}+b^k_{jl}\right)w^l+\left(a_{j\bar k\bar l}+\overline{b^j_{kl}}\right)\bar w^l\right)\right)dw^j\otimes_s d\bar w^k.\end{aligned}$$

Thus, we should (if possible) choose

$$b^k_{jl}=-a_{j\bar kl}\text{ and }\overline{b^j_{kl}}=-a_{j\bar k\bar l}\text{ (i.e., }\overline{b^k_{jl}}=-a_{k\bar j\bar l})$$

We can do this provided that

1. $a_{j\bar{k}l} = a_{l\bar{k}j}$, so that b^k_{jl} is symmetrical j and l, and
2. $a_{j\bar{k}l} = \overline{a_{k\bar{j}\bar{l}}}$, so that $b^k_{jl} = -a_{j\bar{k}l} \Rightarrow \overline{b^j_{kl}} = -a_{j\bar{k}\bar{l}}$.

but this is the case, since we have seen that $d\kappa = 0 \Rightarrow a_{j\bar{k}l} = a_{l\bar{k}j}$ and $\overline{a_{k\bar{j}\bar{l}}} = a_{\bar{k}jl} = a_{j\bar{k}l}$ by (18.151). □

Until further notice, we assume that M (with Riemannian metric h) is a Kähler manifold, we let $\left(x^1 + iy^1, \ldots, x^m + iy^m\right)$ be complex coordinates about some point $x \in M$ such that $\left(\partial_{x^1}, \partial_{y^1}, \ldots, \partial_{x^m}, \partial_{y^m}\right)$ are orthonormal at x relative to h_x, and $\nabla\left(\partial_{x^i}\right) = \nabla\left(\partial_{y^i}\right) = 0$ at x. We find a formula for $\bar{\partial}^* : \Omega^{0,q}(M) \to \Omega^{0,q}(M)$ which will allow us to exhibit $\sqrt{2}\left(\bar{\partial} + \bar{\partial}^*\right)$ locally as a twisted Dirac operator. In view of Theorem 18.69, for $\alpha \in \Omega^{0,q}(M)$, we have

$$\bar{\partial}\alpha = \sum_{j=1}^m d\bar{z}^j \wedge \nabla_{\partial_{\bar{z}^j}} \alpha \quad \text{at } x \in M,$$

since $\nabla_{\partial_{\bar{z}^j}}\left(d\bar{z}^k\right)\left(\partial_{\bar{z}^l}\right) = \partial_{\bar{z}^j}\left(d\bar{z}^k\left(\partial_{\bar{z}^l}\right)\right) - d\bar{z}^k\left(\nabla_{\partial_{\bar{z}^j}}\partial_{\bar{z}^l}\right) = 0$ at x implies that

$$\nabla_{\partial_{\bar{z}^j}}\left(\alpha_{k_1\cdots k_q} d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q}\right) = \partial_{\bar{z}^j}\left(\alpha_{k_1\cdots k_q}\right) d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q} \text{ at } x.$$

We derive a formula for $\bar{\partial}^*$, namely for $\alpha \in \Omega^{0,q}(M)$,

$$\bar{\partial}^*\alpha = -\sum_{j=1}^m \left(d\bar{z}^j\right) \llcorner \nabla_{\partial_{\bar{z}^j}} \alpha \text{ at } x. \tag{18.152}$$

We use the result (18.140), namely

(18.153)
$$*\left(\xi_{(i)_h} \wedge \bar{\xi}_{(j)_k}\right) = 2^{h+k-m}(-1)^h i^{m^2} \varepsilon_{(i^c)_{m-h}(i)_h} \varepsilon_{(j)_k (j^c)_{m-k}} \xi_{(j^c)_{m-k}} \wedge \bar{\xi}_{(i^c)_{m-h}},$$

where $(j)_k$ is the multi-index $(j_1, \ldots, j_k)$ with $1 \le j_1 < \cdots < j_k \le m$, and $(j^c)_{m-k}$ is the multi-index complementary $(j)_k$ in the sense that

$$\{j_1, \ldots, j_k\} \cup \left\{j^c_1, \ldots, j^c_{m-k}\right\} = \{1, \ldots, m\}.$$

and $1 \le j^c_1 < \cdots < j^c_{m-k} < m$. Note that (18.153) implies

$$\begin{aligned} *\left(dz^h * \left(d\bar{z}^{(j)_k}\right)\right) &= *\left(dz^q \wedge 2^{k-m} i^{m^2} \varepsilon_{(j)_k (j^c)_{m-k}} dz^{(j^c)_{m-k}} \wedge d\bar{z}^{(i^c)_m}\right) \\ &= 2^{k-m} i^{m^2} \varepsilon_{(j)_k (j^c)_{m-k}} * \left(dz^q \wedge dz^{(j^c)_{m-k}} \wedge d\bar{z}^{(i^c)_m}\right), \end{aligned}$$

which is zero unless h is j_p for some j_p in $(j)_k$. In the following, $\left(\widehat{j_p}, j\right)_{k-1}$ denotes $(j)_k$, but with j_p removed, and $(j_p, j^c)_{m-k+1}$ denotes the multi-index $(j^c)_{m-k}$ with j_p inserted in the "correct" position which produces an increasing order. Note that

$$\varepsilon_{\left(\widehat{j_p}, j\right)_{k-1} (j_p, j^c)_{m-k+1}} = (-1)^{k-p} \delta^{j_p j^c}_{(j_p, j^c)} \varepsilon_{(j)_k (j^c)_{m-k}}, \tag{18.154}$$

since j_p is moved through $k-p$ indices in $(j)_k (j^c)_{m-k}$ to produce $\left(\widehat{j_p}, j\right)_{k-1} j_p (j^c)_{m-k+1}$, and j_p is moved through $\delta^{j_p j^c}_{(j_p, j^c)}$ indices in $j_p (j^c)_{m-k+1}$ to produce $(j_p, j^c)_{m-k+1}$.

We use (18.154) in the following computation

$$
\begin{aligned}
& * \left(dz^{j_p} * \left(d\bar{z}^{(j)_k} \right) \right) \\
&= 2^{k-m} i^{m^2} \varepsilon_{(j)_k (j^c)_{m-k}} * \left(dz^{j_p} \wedge dz^{(j^c)_{m-k}} \wedge d\bar{z}^{(i^c)_m} \right) \\
&= 2^{k-m} i^{m^2} \varepsilon_{(j)_k (j^c)_{m-k}} \delta^{j_p j^c}_{(j_p, j^c)} * \left(dz^{(j_p, j^c)_{m-k+1}} \wedge d\bar{z}^{(i^c)_m} \right) \\
&= 2^{k-m} i^{m^2} \varepsilon_{(j)_k (j^c)_{m-k}} \delta^{j_p j^c}_{(j_p, j^c)} 2^{(m-k+1)+m-m} (-1)^{m-k+1} i^{m^2} \varepsilon_{\left(\widehat{j_p}, j\right)_{k-1} (j_p, j^c)_{m-k+1}} d\bar{z}^{\left(\widehat{j_p}, j\right)_{k-1}} \\
&= 2 (-1)^{m^2} (-1)^{m-k+1} \varepsilon_{(j)_k (j^c)_{m-k}} \delta^{j_p j^c}_{(j_p, j^c)} \varepsilon_{\left(\widehat{j_p}, j\right)_{k-1} (j_p, j^c)_{m-k+1}} d\bar{z}^{\left(\widehat{j_p}, j\right)_{k-1}} \\
&= 2 (-1)^{k-1} \varepsilon_{(j)_k (j^c)_{m-k}} \delta^{j_p j^c}_{(j_p, j^c)} \varepsilon_{\left(\widehat{j_p}, j\right)_{k-1} (j_p, j^c)_{m-k+1}} d\bar{z}^{\left(\widehat{j_p}, j\right)_{k-1}} \\
&= 2 (-1)^{k-1} (-1)^{k-p} d\bar{z}^{\left(\widehat{j_p}, j\right)_{k-1}} = 2 (-1)^{p-1} d\bar{z}^{\left(\widehat{j_p}, j\right)_{k-1}} = d\bar{z}^{j_p} \llcorner d\bar{z}^{(j)_k}.
\end{aligned}
$$

It then follows that for $\beta \in \Omega^{1,0}(M)$ and $\alpha \in \Omega^{0,q}(M)$,

$$* (\beta \wedge *\alpha) = \bar{\beta} \llcorner \alpha.$$

Thus, with $\alpha = \frac{1}{q!} \sum_{(k)} \alpha_{k_1 \cdots k_q} d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q}$ and formally substituting $\partial = \sum_{k_0} \partial_{z^{k_0}} dz^{k_0} \wedge$ for $\beta \wedge$, we have

$$
\begin{aligned}
\bar{\partial}^* \alpha &= - * (\partial * \alpha) = -\bar{\partial} \llcorner \alpha \\
&= - \left(\sum\nolimits_{k_0} \left(d\bar{z}^{k_0} \right) \partial_{\bar{z}^{k_0}} \right) \llcorner \frac{1}{q!} \sum\nolimits_{(k)} \alpha_{k_1 \cdots k_q} d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q} \\
&= - \frac{1}{q!} \sum\nolimits_{k_0, (k)} \partial_{\bar{z}^{k_0}} \left(\alpha_{k_1 \cdots k_q} \right) \left(d\bar{z}^{k_0} \right) \llcorner \left(d\bar{z}^{k_1} \wedge \cdots \wedge d\bar{z}^{k_q} \right).
\end{aligned}
$$

At $x \in M$ (where the $\partial_{\bar{z}^k}$ are parallel), we then have (18.152).

Ideally, we would like to show that $\Omega^{0,*}(M)$ is a Clifford module bundle $W(M)$ and that $\mathcal{D}^W = \sqrt{2} \left(\bar{\partial} + \bar{\partial}^* \right)$. We could then obtain the Hirzebruch-Riemann-Roch Theorem as a consequence of the Index Theorem for Generalized Dirac Operators (Theorem 18.55, p.567). However, there are technical details of great consequence which stand in the way of doing this, as we now explain.

Note that

$$
\begin{aligned}
& \mathbb{C}^m \to (\mathbb{C} \otimes \mathbb{C}^m)^{1,0} \subset \mathbb{C} \otimes \mathbb{C}^m \text{ given by} \\
& v \mapsto 2^{\frac{1}{2}} v^{0,1} := 2^{\frac{1}{2}} \tfrac{1}{2} (v + iJv) = 2^{-\frac{1}{2}} (v + iJv)
\end{aligned}
$$

is an isometry. Indeed, if $\langle \cdot, \cdot \rangle_m$ is the standard Hermitian inner product on $\mathbb{C}^m$

$$
\begin{aligned}
H \left(2^{-\frac{1}{2}} (v + iJv), 2^{-\frac{1}{2}} (v + iJv) \right) &= \left\langle 2^{-\frac{1}{2}} (v + iJv), 2^{-\frac{1}{2}} (v - iJv) \right\rangle \\
&= \tfrac{1}{2} \left(\langle v, v \rangle + \langle Jv, Jv \rangle \right) = \langle v, v \rangle.
\end{aligned}
$$

We take $W := \Lambda^{0,*} \left((\mathbb{C} \otimes \mathbb{C}^m)^* \right) \cong \Lambda^* (\mathbb{C}^m)$ and let $Q : \mathbb{C}l_n \to \operatorname{End}(W)$ be determined by

$$Q_1(v) \alpha = 2^{\frac{1}{2}} \left(\flat \left(v^{0,1} \right) \wedge \alpha - \flat \left(v^{0,1} \right) \llcorner \alpha \right),$$

for $v \in \mathbb{R}^n = \mathbb{C}^m$. Let $q : U(m) \to U(W) = \Lambda^{0,*} \left((\mathbb{C}^m)^* \right)$ be the usual representation induced by $q_1 : U(m) \to U \left(\Lambda^{0,1} \left((\mathbb{C}^m)^* \right) \right)$ which we describe as follows. For

$A \in U(m) \subset \mathrm{SO}(2m)$, we have the $\mathbb{C}$-linear extension $A_c \in U(2m)$ which preserves the summands of the decomposition

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m = (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^{1,0} \oplus (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^{0,1}$$

We also $\widehat{A} \in U\left((\mathbb{C}^m)^*\right)$ given by $\widehat{A}(\varphi) = \varphi \circ A^{-1}$, and its $\mathbb{C}$-linear extension $\widehat{A}_c \in U\left(\left(\mathbb{C}^{2m}\right)^*\right) = U\left((\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^*\right)$ which preserves the summands of

$$(\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^* = (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^{*1,0} \oplus (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^{*0,1}$$

For $A \in U(m)$, we take

$$q_1(A) = (\widehat{A}_c)^{0,1} \in \mathrm{End}\left((\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^{*0,1}\right) = \mathrm{End}\left(\Lambda^{0,1}\left((\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^*\right)\right),$$

and let $q(A)$ be the natural extension of $q_1(A)$ to all of $\Lambda^{0,*}\left((\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^*\right)$. For $A \in U(m)$ and $v \in \mathbb{C}^m$, we have

$$q(A) \circ Q(v) \circ q\left(A^{-1}\right) = Q(Av).$$

Indeed for $\alpha \in W = \Lambda^{0,*}\left((\mathbb{C}^m)^*\right)\Lambda^{0,*}\left((\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}^m)^*\right) \cong \mathbb{C}l_n$,

$$\begin{aligned}
\left(q(A) \circ Q(v) \circ q\left(A^{-1}\right)\right)\alpha &= q(A)\left(Q(v)\left(q\left(A^{-1}\right)\alpha\right)\right) \\
&= 2^{\frac{1}{2}} q(A)\left(\flat\left(v^{0,1}\right) \wedge \left(q\left(A^{-1}\right)(\alpha)\right) - \flat\left(v^{0,1}\right) \llcorner \left(q\left(A^{-1}\right)(\alpha)\right)\right) \\
&= 2^{\frac{1}{2}} q(A)\left(\flat\left(v^{0,1}\right)\right) \wedge (\alpha) - q(A)\left(\flat\left(v^{0,1}\right)\right) \llcorner \alpha \\
&= 2^{\frac{1}{2}} \flat\left((Av)^{0,1}\right) \wedge \alpha - \flat\left((Av)^{0,1}\right) \llcorner \alpha \\
&= Q(Av)(\alpha).
\end{aligned}$$

Hence for $A \in U(m)$ and $\alpha \in \mathbb{C}l_n$, we have

$$q(A) Q(\alpha) q\left(A^{-1}\right) = Q(r(A)\alpha).$$

In view of this, there is a well-defined

$$\begin{aligned}
cl &: \mathbb{C}l(TM) \otimes W(M) \to W(M) \\
cl &: \mathbb{C}l(TM) \otimes \Lambda^{0,*}(T_{\mathbb{C}}M^*) \to \Lambda^{0,*}(T_{\mathbb{C}}M^*)
\end{aligned}$$

Moreover, $cl : \mathbb{C}l(TM) \otimes W(M) \to W(M)$ induces map on sections,

$$\begin{aligned}
cl &: C^\infty(\mathbb{C}l(TM)) \otimes C^\infty(W(M)) \to C^\infty(W(M)) \\
cl &: C^\infty(\mathbb{C}l(TM)) \otimes \Omega^{0,*}(M) \to \Omega^{0,*}(M).
\end{aligned}$$

For $\alpha \in C^\infty(\mathbb{C}l(TM))$ and $\psi \in C^\infty\left(\Omega^{0,*}(M)\right)$, we have

$$\nabla^{0,1}(cl(\alpha \otimes \psi)) = cl(\nabla\alpha \otimes \psi) + cl\left(\alpha \otimes \nabla^{0,1}\psi\right) \tag{18.155}$$

where

$$\nabla^{0,1} : \Omega^{0,*}(M) \overset{\nabla^{0,1}}{\to} \Omega^1(M) \otimes \Omega^{0,*}(M)$$

is just the complex extention of Levi-Civita covariant derivative which has the property that

$$\psi \in \Omega^{0,k}(M) \Rightarrow \nabla_X \psi \in \Omega^{0,k}(M)$$

and "generalized Dirac operator"

$$\mathcal{D}^W : \Omega^{0,*}(M) \overset{\nabla^{0,1}}{\to} \Omega^1(M) \otimes \Omega^{0,*}(M) \overset{cl}{\to} \Omega^{0,*}(M).$$

which is found to be $2^{\frac{1}{2}}\left(\bar{\partial}+\bar{\partial}^*\right)$. However, the problem is that

$$W(M)=\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)=UM\times_{U(m)}\Lambda^{0,*}\left(\left(\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m\right)^*\right)$$

is an associated bundle of UM instead of FM or P for a spin structure $P\to FM$, as is required in the proof of Theorem 18.55, p. 567; otherwise, it is not clear that $\mathcal{D}^W$ is locally a twisted Dirac operator. Let us assume, for the moment, that this does not matter. Then Theorem 18.55 would yield

$$\operatorname{Index}\left(\bar{\partial}+\bar{\partial}^*\right)=\operatorname{Index}\left(\mathcal{D}^{W+}\right)=\left(\mathbf{ch}\left(\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)\right)\smile\widetilde{\mathbf{A}}(M)\right)[M].$$

Now, identifying $\mathbf{c}_k(TM)$ with $\sigma_k(y_1,\dots,y_m)$, we have

$$\mathbf{ch}\left(\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)\right)=\sum_{h=0}^m\sum_{(j)_h}e^{y_{j_1}+\cdots+y_{j_h}}=\prod_{k=1}^m\left(1+e^{y_k}\right),\text{ and}$$

$$\begin{aligned}\mathbf{ch}\left(\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)\right)\smile\widetilde{\mathbf{A}}(M)&=\prod_{k=1}^m\left(1+e^{y_k}\right)\prod_{k=1}^m\tfrac{y_k/2}{\sinh(y_k)}\\&=\prod_{k=1}^m\left(1+e^{y_k}\right)\tfrac{y_k/2}{\sinh(y_k)}=\prod_{k=1}^m\frac{y_k}{1-e^{-y_k}}\\&=\mathbf{Td}\left(T^{1,0}(M)\right)=\mathbf{Td}(TM),\text{ and so}\end{aligned}\tag{18.156}$$

$$\operatorname{Index}\left(\bar{\partial}+\bar{\partial}^*\right)=\mathbf{Td}(TM)[M].$$

This is correct and indeed it is the Hirzebruch-Riemann-Roch formula, but since $W(M)=\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)$ is associated to $U(M)$ instead of FM or a spin structure bundle P, Theorem 18.55 does not apply. This suggests that the notion of Clifford module bundle can be generalized in such a way that Theorem 18.55 still holds. In fact this is the case, and we will examine the case at hand to determine a suitable generalization.

To fix the fact that $\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)$ is associated to $U(M)$, instead of FM or P, ideally one would like to find a homomorphism

$$h:U(m)\to\mathrm{SO}(2m),$$

a representation $r:\mathrm{SO}(2m)\to\mathrm{GL}(V)$, and an equivariant isomorphism

$$\phi:\Lambda^{0,*}\left(\left(\mathbb{C}\otimes\mathbb{C}^m\right)^*\right)\to V$$

Then we would have $\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)\cong FM\times_{\mathrm{SO}(2m)}V$, exhibiting $\Lambda^{0,*}\left(T_{\mathbb{C}}M^*\right)$ as a bundle associated with FM. It is natural to take h to be inclusion and ϕ to be the identity. However, $\Lambda^{0,*}\left(\left(\mathbb{C}\otimes\mathbb{C}^m\right)^*\right)$ is not an invariant subspace of the representation $\mathrm{SO}(2m)\to\Lambda^*\left(\left(\mathbb{C}\otimes\mathbb{C}^m\right)^*\right)$, but rather of its restriction to $U(m)$. While we will not prove the fact that there is no suitable choice r and ϕ, at least the most obvious attempt fails. The somewhat involved resolution of this difficulty is doubly worthwhile, since it also provides is a nice way of motivating Seiberg-Witten theory. In a nutshell, we will exploit the fact that while

$$\Lambda^{0,*}:U(m)\to\operatorname{End}\left(\Lambda^{0,*}\left(\left(\mathbb{C}\otimes\mathbb{C}^m\right)^*\right)\right)$$

does not extend to a representation of $\mathrm{SO}(2m)$, we have linear isomorphisms

$$\Lambda^{0,*}\left(\left(\mathbb{C}\otimes\mathbb{C}^m\right)^*\right)\cong\Lambda^*\left(\mathbb{C}^m\right)\cong\Sigma_{2m},$$

and there is the spinor representation $\rho:\mathrm{Spin}(2m)\to\operatorname{End}(\Sigma_{2m})$. However on the Lie algebra level, the restriction of $\rho':\mathfrak{spin}(2m)\to\operatorname{End}(\Sigma_{2m})$ to $\mathfrak{u}(m)\subset\mathfrak{so}(2m)\cong$

$\mathfrak{spin}(2m)$ does not coincide with $\left(\Lambda^{0,*}\right)' : \mathfrak{u}(m) \to \operatorname{End}\left(\Lambda^{0,*}\left((\mathbb{C}\otimes\mathbb{C}^m)^*\right)\right)$. Indeed, there is a crucial 1-dimensional twist which is needed to make a correct identification. We proceed with the details.

Since we will be considering $\mathbb{C}\otimes\mathbb{C}^m$, we use J to denote multiplication by $\sqrt{-1}$ in $\mathbb{C}^m$ and i to denote multiplication by $\sqrt{-1}$ in $\mathbb{C}\otimes\mathbb{C}^m$. Note that the map $\flat^{0,1} : \mathbb{C}^m \to \left((\mathbb{C}\otimes\mathbb{C}^m)^*\right)^{0,1}$ given by $\flat^{0,1}(v) = 2^{-\frac{1}{2}}\flat(v+iJv)$ is a $\mathbb{C}$-linear isometry, since it is clearly an $\mathbb{R}$-linear isometry and

$$\flat^{0,1}(Jv) = 2^{-\frac{1}{2}}\flat(Jv - iv) = 2^{-\frac{1}{2}}\flat(-i(v-iJv)) = i2^{-\frac{1}{2}}\flat(v-iJv) = i\flat^{0,1}(v)$$

Thus, $\flat^{0,1}$ extends to a $\mathbb{C}$-linear isometry

$$\Lambda\left(\flat^{0,1}\right) : \Lambda^*(\mathbb{C}^m) \cong \Lambda^{0,*}\left((\mathbb{C}\otimes\mathbb{C}^m)^*\right).$$

Recall that $\Lambda^*(\mathbb{C}^m)$ provides a representation space for an irreducible $\mathbb{C}l_{2m}$-module

$$\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right).$$

Here $\rho_{\mathbb{C}}$ is the $\mathbb{C}$-linear extension of

$$\rho : Cl_{2m} \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right),$$

which was described as follows. Let $\mathbb{R}^{2m} \cong \mathbb{C}^m$ be the usual isomorphism

$$(x_1, y_1, \ldots, x_m, y_m) = (x_1 + iy_1, \ldots, x_m + iy_m).$$

For $w \in \mathbb{C}^m$, define

$$\rho_1 : \mathbb{R}^{2m} \cong \mathbb{C}^m \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right) \text{ by}$$
$$\rho_1(w)(\alpha) := (w\wedge - w\llcorner)(\alpha) = w\wedge\alpha - w\llcorner\alpha.$$

We verified that ρ_1 extends to $\rho : Cl_{2m} \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right)$. In this way one may view the space Σ_{2m} of spinors concretely as $\Lambda^*(\mathbb{C}^m)$. There is then the representation

$$\rho|_{\operatorname{Spin}(n)} : \operatorname{Spin}(n) \to \operatorname{SU}(\Sigma_{2m}).$$

Since $\rho : Cl_{2m} \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right)$ is an *algebra* homomorphism, for $\alpha \in \mathfrak{spin}(n) = \mathcal{L}\left(\Lambda^2(\mathbb{C}^m)\right) \subset Cl_{2m}$, we have $\rho\left(e^{t\alpha}\right) = e^{t\rho(\alpha)}$, and so

$$\left(\rho|_{\operatorname{Spin}(n)}\right)'(\alpha) = \tfrac{d}{dt}\rho\left(e^{t\alpha}\right) = \tfrac{d}{dt}\left(e^{t\rho(\alpha)}\right) = \rho(\alpha).$$

Thus,

$$\left(\rho|_{\operatorname{Spin}(n)}\right)' = \rho|_{\mathfrak{spin}(n)} = \rho|_{\mathcal{L}(\Lambda^2(\mathbb{C}^m))}.$$

Recall that we also have the 2-fold covering $c : \operatorname{Spin}(n) \to \operatorname{SO}(n)$. Since there is a natural injection

$$\iota : \operatorname{U}(m) \to \operatorname{SO}(n),$$

at first one might guess that on the level of Lie algebras, the action of $\mathfrak{u}(m) \subset \mathfrak{so}(2m) \cong \mathfrak{spin}(2m)$ on $\Lambda^*(\mathbb{C}^m)$ given by

$$\left(\rho|_{\operatorname{Spin}(n)}\right)' : \mathfrak{u}(m) \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right)$$

coincides with the usual $(\Lambda^*)' : \mathfrak{u}(m) \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right)$ where $\Lambda^* : \operatorname{U}(m) \to \operatorname{End}\left(\Lambda^*(\mathbb{C}^m)\right)$ is the usual representation induced by the defining (identitity) representation $\operatorname{U}(m) \to \operatorname{U}(\mathbb{C}^m)$. Explicitly, for $A \in \operatorname{U}(m)$ and $0 \le l \le m$, we have $\Lambda^l(A) \in \operatorname{End}\left(\Lambda^l(\mathbb{C}^m)\right)$ given by

$$\Lambda^l(A)\left(f_{j_1}\wedge f_{j_2}\wedge\cdots\wedge f_{j_l}\right) := Af_{j_1}\wedge Af_{j_2}\wedge\cdots\wedge Af_{j_l},$$

and this extends to $\Lambda^* : \mathrm{U}(m) \to \mathrm{End}(\Lambda^*(\mathbb{C}^m))$. Although $\left(\rho|_{\mathrm{Spin}(n)}\right)' \neq (\Lambda^*)'$, the two representations are related in a way which will ultimately exhibit $\sqrt{2}\left(\bar{\partial} + \bar{\partial}^*\right)$ at least locally as a twisted generalized Dirac operator, thereby justifying the computation (18.156) via forms. We now describe how $\left(\rho|_{\mathrm{Spin}(n)}\right)'$ and $(\Lambda^*)'$ are related.

For $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, let $B_\lambda \in \mathfrak{u}(m)$ be the diagonal matrix

$$B_\lambda := diag\,(i\lambda_1, \dots, i\lambda_m).$$

The endomorphism $(\Lambda^*)'(B_\lambda) \in \mathrm{End}(\Lambda^*(\mathbb{C}^m))$ leaves each of the $\Lambda^l(\mathbb{C}^m)$ $(l = 1, \cdots, m)$ fixed. Explicitly, $\left(\Lambda^l\right)'(B_\lambda) \in \mathrm{End}\left(\Lambda^l(\mathbb{C}^m)\right)$ (induced by $B_\lambda \in \mathrm{End}(\mathbb{C}^m)$) is given simply by

$$\Lambda^l(B_\lambda)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) = i\,(\lambda_{j_1} + \cdots + \lambda_{j_l})(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}).$$

We wish to compare $(\Lambda^*)'(B_\lambda)$ with

$$\left(\rho|_{\mathrm{Spin}(n)}\right)'\left(c'^{-1}(\iota(B_\lambda))\right) \in \mathrm{End}(\Sigma_{2m}) = \mathrm{End}(\Lambda^*(\mathbb{C}^m)).$$

Recall that for $A = [a_{ij}] \in \mathfrak{so}(n)$, the Lie algebra isomorphism $c' : \mathfrak{spin}(2m) = \mathcal{L}\left(\Lambda^2\left(\mathbb{R}^{2m}\right)\right) \cong \mathfrak{so}(n)$ for the covering $c : \mathrm{Spin}(2m) \to \mathrm{SO}(2m)$ is given by

$$c'(-\tfrac{1}{4}\sum_{i,j} a_{ij} e_i e_j) := A.$$

For $\iota : U(m) \to \mathrm{SO}(n)$, the associated Lie algebra map $\iota' : \mathfrak{u}(m) \to \mathfrak{so}(n)$ applied to B_λ is given by

$$\iota'(B_\lambda)(e_{2j-1}) = \lambda_j e_{2j} \text{ and } \iota'(B_\lambda)(e_{2j}) = -\lambda_j e_{2j-1},\ j = 1, \dots, m, \text{ so that}$$
$$\iota'(B_\lambda)_{2j,2j-1} = \lambda_j \text{ and } \iota'(B_\lambda)_{2j-1,2j} = -\lambda_j.$$

Then

$$\begin{aligned}\iota'(B_\lambda) &= c'(-\tfrac{1}{4}\textstyle\sum_{i,j=1}^{n} \iota(B_\lambda)_{ij}\, e_i e_j) = c'(-\tfrac{1}{2}\textstyle\sum_{j=1}^{m} \iota(B_\lambda)_{2j-1,2j}\, e_{2j-1} e_{2j}) \\ &= c'(\tfrac{1}{2}\textstyle\sum_{j=1}^{m} \lambda_j e_{2j-1} e_{2j}).\end{aligned}$$

We have shown (see (18.14), p. 490) that the action of $\rho(e_{2j-1}e_{2j})$ on $f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}$ is given by

$$\begin{aligned}&\rho(e_{2j-1}e_{2j})(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) \\ &= \begin{cases} i\,(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) & \text{if } j_k = j \text{ for some } k \\ -i\,(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) & \text{if } j_k \neq j \text{ for all } k \end{cases} \\ &= 2i\left(\textstyle\sum_{k=1}^{l} \delta^j_{j_k}(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l})\right) - i\,(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}), \text{ and so}\end{aligned}$$

$$\begin{aligned}&\left(\rho|_{\mathrm{Spin}(n)}\right)'\left(c'^{-1}(\iota'(B_\lambda))\right)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) \\ &= \rho(\tfrac{1}{2}\sum_j \lambda_j e_{2j-1}e_{2j})(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) \\ &= i\textstyle\sum_{j=1}^{m}\sum_{k=1}^{l} \lambda_j \delta^j_{j_k}(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) - i\left(\tfrac{1}{2}\textstyle\sum_{j=1}^{m}\lambda_j\right)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) \\ &= \left(\Lambda^l\right)'(B_\lambda)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) - i\left(\tfrac{1}{2}\textstyle\sum_{j=1}^{m}\lambda_j\right)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}) \\ &= \left(\left(\Lambda^l\right)'(B_\lambda) - i\left(\tfrac{1}{2}\textstyle\sum_{j=1}^{m}\lambda_j\right)\right)(f_{j_1} \wedge f_{j_2} \wedge \cdots \wedge f_{j_l}).\end{aligned}$$

Since this holds for arbitrary l,

$$\begin{aligned}\rho\left(c'^{-1}\left(\iota'\left(B_\lambda\right)\right)\right) &= -\tfrac{1}{2}\left(i\sum\nolimits_{j=1}^{m}\lambda_j\right) + \left(\Lambda^*\right)'\left(B_\lambda\right)\\ &= -\tfrac{1}{2}\operatorname{Tr}\left(B_\lambda\right)\operatorname{Id} + \left(\Lambda^*\right)'\left(B_\lambda\right),\ \text{ or}\end{aligned}$$

$$\left(\Lambda^*\right)'\left(B_\lambda\right) = \tfrac{1}{2}\operatorname{Tr}\left(B_\lambda\right)\operatorname{Id} + \rho\left(c'^{-1}\left(\iota'\left(B_\lambda\right)\right)\right). \tag{18.157}$$

Although $\left(\Lambda^*\right)'\left(B_\lambda\right)$ and $\rho\left(c'^{-1}\left(\iota'\left(B_\lambda\right)\right)\right)$ are different, this formula provides a definite comparison. While $\left(\Lambda^*\right)'$ does not have values in $\rho\left(\mathfrak{spin}\left(2m\right)\right) = \rho\left(\mathcal{L}\left(\Lambda^2\left(\mathbb{R}^{2m}\right)\right)\right)$, it does have values in the slightly larger space $\rho_{\mathbb{C}}\left(i\mathbb{R}\oplus\mathfrak{spin}\left(2m\right)\right)$. Note that $i\mathbb{R}\oplus\mathfrak{spin}\left(2m\right)$ is the Lie algebra of

$$U\left(1\right)\operatorname{Spin}\left(2m\right) := \left\{zg : g\in\operatorname{Spin}\left(2m\right),\ z\in\mathbb{C},\ |z|=1\right\}\subset\mathbb{C}l\left(2m\right).$$

This suggests (as we will show) that possibly there is a homomorphism

$$\begin{aligned}j : U\left(m\right) &\to U\left(1\right)\operatorname{Spin}\left(2m\right)\ \text{ such that}\\ \rho_{\mathbb{C}}\circ j' = \left(\Lambda^*\right)' : \mathfrak{u}\left(m\right) &\to \operatorname{End}\left(\Lambda^*\left(\mathbb{C}^m\right)\right).\end{aligned}$$

If so, then by (18.157),

$$j'\left(B_\lambda\right) = \rho_{\mathbb{C}}^{-1}\circ\Lambda^{*\prime}\left(B_\lambda\right) = \tfrac{1}{2}\operatorname{Tr}\left(B_\lambda\right) + c'^{-1}\left(\iota'\left(B_\lambda\right)\right)\subseteq i\mathbb{R}\oplus\mathfrak{spin}\left(2m\right).$$

From this, it can be conjectured that for $A\in U\left(m\right)$ with an orthonormal basis of eigenvectors $v_1,\ldots,v_m$ and eigenvalues $e^{i\theta_1},\ldots,e^{i\theta_m}$, one can define

$$j\left(A\right) = \exp\left(\tfrac{i}{2}\sum\nolimits_{k=1}^{m}\theta_k\right)\prod\nolimits_{k=1}^{m}\left(\cos\left(\tfrac{1}{2}\theta_k\right)+\sin\left(\tfrac{1}{2}\theta_k\right)\right)v_k J v_k.$$

However, there is a problem (not impossibly difficult) with showing that this yields a well defined, continuous homomorphism, independent of the choice of the v_k. Alternatively, we obtain j as follows. Note that the homomorphism

$$\chi : \mathrm{U}\left(1\right)\times\operatorname{Spin}\left(2m\right)\to\mathrm{U}\left(1\right)\operatorname{Spin}\left(2m\right)\ \text{ given by }\chi\left(z,\sigma\right) := z\sigma \tag{18.158}$$

is onto with kernel $\left\{\pm\left(1,1\right)\right\}$, since

$$zg = 1\Rightarrow g = z^{-1}\Rightarrow g\in\mathbb{R}\cap U\left(1\right)\Rightarrow g = \pm1\Rightarrow\left(z,g\right) = \pm\left(1,1\right).$$

Thus, we are motivated to define

$$\operatorname{Spin}^c\left(n\right) := \mathrm{U}\left(1\right)\operatorname{Spin}\left(2m\right)\cong\frac{\mathrm{U}\left(1\right)\times\operatorname{Spin}\left(n\right)}{\left\{\pm\left(1,1\right)\right\}}. \tag{18.159}$$

Often $\operatorname{Spin}^c\left(n\right)$ is defined to be the group on the right. This has the advantage of making it clear that $\operatorname{Spin}^c\left(n\right)$ is not $U\left(1\right)\times\operatorname{Spin}\left(n\right)$, and our definition exhibits $\operatorname{Spin}^c\left(n\right)$ as a subgroup of $\mathbb{C}l\left(2m\right)$, with Lie algebra $i\mathbb{R}\oplus\mathfrak{spin}\left(2m\right)$. There is a double-covering homomorphism (with kernel $\left\{1,-1\right\}$)

$$\begin{aligned}Sq\times c : \operatorname{Spin}^c\left(n\right) &\to U\left(1\right)\times\operatorname{SO}\left(n\right)\ \text{ given by}\\ \left(Sq\times c\right)\left(zg\right) &:= \left(z^2, c\left(g\right)\right),\end{aligned} \tag{18.160}$$

which is well-defined since $\left(\left(-z\right)^2, c\left(-g\right)\right) = \left(z^2, c\left(g\right)\right)$.

PROPOSITION 18.70. *There exists a lifting homomorphism* $j : \mathrm{U}(m) \to \mathrm{Spin}^c(n)$, *such that* $\det \times \iota = j \circ (Sq \times c)$; *i.e., the diagram*

$$\begin{array}{ccccc} & & \mathrm{Spin}^c(n) & \overset{\rho_{\mathbb{C}}}{\to} & U(\Sigma_{2m}) = U(\Lambda^*(\mathbb{C}^m)) \\ & \overset{j}{\nearrow} & \downarrow Sq \times c & & \\ \mathrm{U}(m) & \overset{\det \times \iota}{\to} & \mathrm{U}(1) \times \mathrm{SO}(n) & & \end{array} \tag{18.161}$$

commutes. Moreover,

$$\rho_{\mathbb{C}} \circ j = \Lambda^* : \mathrm{U}(m) \to \mathrm{U}(\Lambda^*(\mathbb{C}^m)).$$

PROOF. The existence of a continuous map $j : \mathrm{U}(m) \to \mathrm{Spin}^c(n)$ with $\det \times \iota = j \circ (Sq \times c)$ will be deduced from covering space theory by showing

$$(\det \times \iota)_{\#}(\pi_1(\mathrm{U}(m), I)) \subseteq (Sq \times c)_{\#}(\pi_1(\mathrm{Spin}^c(n), 1)).$$

We have

$$\begin{aligned} &\pi_1(\mathrm{U}(m), I) \cong \mathbb{Z},\ \pi_1(\mathrm{Spin}^c(n), 1) \cong \mathbb{Z}, \text{ and} \\ &\pi_1(\mathrm{U}(1) \times \mathrm{SO}(n), (1, I)) \cong \mathbb{Z} \times \mathbb{Z}_2. \end{aligned}$$

The generator, say $g_{U(m)}$, of $\pi_1(\mathrm{U}(m), I)$ is represented by $\mathrm{U}(1) \to \mathrm{U}(m)$, given by $e^{i\theta} \mapsto diag(e^{i\theta}, 1, \ldots, 1)$, and the generator $g_{\mathrm{SO}(n)}$ of $\pi_1(\mathrm{SO}(n), I)$ is represented by

$$e^{i\theta} \mapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \oplus I_{2n-2}.$$

By means of $g_{U(m)}$ and $g_{\mathrm{SO}(n)}$, we identify $\pi_1(\mathrm{U}(1) \times \mathrm{SO}(n), (1, I)) = \mathbb{Z} \times \mathbb{Z}_2$. We have $(\det \times \iota)_{\#}(g_{U(m)}) = (g_{U(1)}, g_{\mathrm{SO}(n)})$, since

$$(\det \times \iota)\left(diag\left(e^{i\theta}, 1, \ldots, 1\right)\right) = \left(e^{i\theta}, \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \oplus I_{2n-2}\right),$$

The group G of covering transformations of the covering $Sq \times c : \mathrm{Spin}^c(n) \to \mathrm{U}(1) \times \mathrm{SO}(n)$ consists of Id and the map $zg \mapsto -zg$. The image, say N, of $\pi_1(\mathrm{Spin}^c(n), 1)$ under the monomorphism $(Sq \times c)_{\#}$ is then a subgroup of $\pi_1(\mathrm{U}(1) \times \mathrm{SO}(n), (1, I)) = \mathbb{Z} \times \mathbb{Z}_2$. Note that N is normal since $\mathbb{Z} \times \mathbb{Z}_2$ is abelian. From standard covering space theory, we then have $\mathbb{Z}_2 \cong G \cong (\mathbb{Z} \times \mathbb{Z}_2)/N$. There are several subgroups of $\mathbb{Z} \times \mathbb{Z}_2$ of index 2, namely $\mathbb{Z} \times \{0\}$, $2\mathbb{Z} \times \mathbb{Z}_2$, and

$$N = \{(p, p \bmod 2) : p \in \mathbb{Z}\} \tag{18.162}$$

which is generated by $(1, 1) = (g_{U(1)}, g_{\mathrm{SO}(n)})$. To prove (18.162), it suffices to observe that

$$e^{i\theta} \mapsto e^{i\theta/2}\left(\cos\left(\tfrac{1}{2}\theta\right) + \sin\left(\tfrac{1}{2}\theta\right) e_1 e_2\right), \quad 0 \le \theta \le 2\pi$$

is a loop in $\mathrm{Spin}^c(n)$ projecting via $Sq \times c$ to a representative of $(1, 1)$. This loop then represents the generator, say g_c, of $\pi_1(\mathrm{Spin}^c(n), 1) \cong N \cong \mathbb{Z}$. Now

$$\begin{aligned} &(Sq \times c)\left[e^{i\theta/2}, \cos\left(\tfrac{1}{2}\theta\right) + \sin\left(\tfrac{1}{2}\theta\right) e_1 e_2\right] = \left(e^{i\theta}, c\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2} e_1 e_2\right)\right) \\ &= \left(e^{i\theta}, \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \oplus I_{2n-2}\right) = (\det \times \iota)\left(diag\left(e^{i\theta}, 1, \ldots, 1\right)\right) \\ &\Rightarrow (\det \times \iota)_{\#}\left(g_{U(m)}\right) = (Sq \times c)_{\#}(g_c) \\ &\Rightarrow (\det \times \iota)_{\#}(\pi_1(\mathrm{U}(m))) = (Sq \times c)_{\#}(\pi_1(\mathrm{Spin}^c(n))). \end{aligned}$$

Covering space theory then provides us with a topological lifting j in the diagram (18.161) which sends $I \in \mathrm{U}(m)$ to $1 \in \mathrm{Spin}^c(n)$. We prove that j is a homomorphism as follows. Since $\det \times \iota$ is a homomorphism and $Sq \times c$ is a two-fold cover, for all $A_1, A_2 \in \mathrm{U}(m)$, we have

$$j(A_1A_2)\, j(A_2)^{-1}\, j(A_1)^{-1} = 1 \text{ or } -1,$$

but the left side is a continuous function of (A_1, A_2) which is 1 at $(A_1, A_2) = (I, I)$, whence $j(A_1A_2)\, j(A_2)^{-1}\, j(A_1)^{-1} = 1$ for all (A_1, A_2) in the connected space $\mathrm{U}(m)\times\mathrm{U}(m)$. Of course, the lift j of the C^∞ map $\det\times\iota$ is smooth, since the covering $Sq \times c$ is a local diffeomorphism. That $\rho_{\mathbb{C}} \circ j = \Lambda^* : \mathrm{U}(m) \to \mathrm{U}(\Lambda^*(\mathbb{C}^m))$ can be deduced as follows. Since

$$\begin{aligned}(\rho_{\mathbb{C}} \circ j)'(B_\lambda) &= \left(\rho_{\mathbb{C}} \circ (Sq\times c)'^{-1} \circ (\det \times \iota)'\right)(B_\lambda)\\ &= \rho_{\mathbb{C}} \circ (Sq \times c)'^{-1}(\mathrm{Tr}(B_\lambda), \iota'(B_\lambda))\\ &= \rho_{\mathbb{C}}\left(\tfrac{1}{2}\mathrm{Tr}(B_\lambda) + c'^{-1}(\iota'(B_\lambda))\right)\\ &= \tfrac{1}{2}\mathrm{Tr}(B_\lambda) + \rho\left(c'^{-1}(\iota'(B_\lambda))\right)\\ &= (\Lambda^*)'(B_\lambda),\end{aligned}$$

by (18.157), $(\rho_{\mathbb{C}} \circ j)'$ and $\Lambda^{*\prime}$ agree on the (Cartan) subalgebra $\mathfrak{t} := \{B_\lambda : \lambda \in \mathbb{R}^m\} \subset \mathfrak{u}(m)$ which is the Lie algebra of the (maximal) torus

$$T := \left\{diag\left(e^{i\lambda_1}, \ldots, e^{i\lambda_m}\right) : \lambda \in \mathbb{R}^m\right\} \subset \mathrm{U}(m).$$

Thus, $\Lambda^*|_T = (\rho_{\mathbb{C}} \circ j)|_T$. Those sufficiently familiar with the theory of representations of compact Lie groups will then conclude that $\Lambda^* = \rho_{\mathbb{C}} \circ j$, but for those less familiar with the theory, we offer the following. Since for every $A \in \mathrm{U}(m)$ there is $B \in \mathrm{U}(m)$ such that $BAB^{-1} \in T$, we have

$$\begin{aligned}\mathrm{Tr}(\Lambda^*(A)) &= \mathrm{Tr}\left(\Lambda^*(B)^{-1}\Lambda^*(A)\Lambda^*(B)\right) = \mathrm{Tr}\left(\Lambda^*\left(BAB^{-1}\right)\right)\\ &= \mathrm{Tr}\left((\rho_{\mathbb{C}} \circ j)\left(BAB^{-1}\right)\right) = \mathrm{Tr}\left((\rho_{\mathbb{C}} \circ j)(A)\right).\end{aligned}$$

Hence the characters $\mathrm{Tr}\circ(\rho_{\mathbb{C}} \circ j)$ and $\mathrm{Tr}\circ\Lambda^*$ of the representations $\rho_{\mathbb{C}} \circ j$ and Λ^* are the same, which implies (see 4.10.2, p. 107 of [**Wal**]) that $\rho_{\mathbb{C}} \circ j$ and Λ^* are equivalent. Thus $\rho_{\mathbb{C}} \circ j$ and Λ^* differ by a constant multiple on each of the irreducible subspaces $\Lambda^l(\mathbb{C}^m)$. But the constant multipliers are all 1, since $\rho \circ j$ and Λ^* agree on T. Hence, $\Lambda^* = \rho_{\mathbb{C}} \circ j$; i.e., we have

$$\begin{array}{ccc}\mathrm{Spin}^c(n) & \overset{\rho_{\mathbb{C}}}{\to} & \mathrm{U}(\Lambda^*(\mathbb{C}^m)) = \mathrm{U}(\Sigma_{2m})\\ \uparrow j & & \downarrow \mathrm{Id}\\ \mathrm{U}(m) & \overset{\Lambda^*}{\to} & \mathrm{U}(\Lambda^*(\mathbb{C}^m)).\end{array} \tag{18.163}$$

□

Let M be a complex manifold of complex dimension m with a Riemannian metric h which makes M a Kähler manifold. The homomorphism $j : \mathrm{U}(m) \to \mathrm{Spin}^c(n)$ yields a $\mathrm{Spin}^c(n)$-bundle $P_{\mathrm{Spin}^c(n)} \to M$, namely

$$\begin{aligned}&P_{\mathrm{Spin}^c(n)} := \mathrm{U}(M) \times_{\mathrm{U}(m)} \mathrm{Spin}^c(n), \text{ and a bundle morphism}\\ &f_j : \mathrm{U}(M) \to P_{\mathrm{Spin}(n)} \text{ (namely, } u \mapsto [u, 1]\text{)},\end{aligned}$$

which is equivariant in the sense that $f_j(uA) = f_j(u)\, j(A)$. The bundle $\Lambda^{0,n}(T_{\mathbb{C}}M^*)$ of complex dimension 1 is known as the *canonical line bundle* for M. Let $P_{\mathrm{U}(1)}$ be the unitary frame bundle for $\Lambda^{n,0}(T_{\mathbb{C}}M)$. Then

$$\Lambda^{0,n}(T_{\mathbb{C}}M^*) \cong \Lambda^{n,0}(T_{\mathbb{C}}M) = P_{\mathrm{U}(1)} \times_{\mathrm{U}(1)} \mathbb{C},$$

There is an equivariant bundle morphism $\wedge^n : \mathrm{U}(M) \to P_{\mathrm{U}(1)}$ given by $\wedge^n(u) \mapsto u(e_1)\wedge\cdots\wedge u(e_m)$, where $e_1, \ldots, e_m$ is the standard basis for $\mathbb{C}^m$ and $u : \mathbb{C}^m \to TM$ is a unitary frame. Since we will find that it is only necessary to exhibit $\sqrt{2}\left(\partial + \bar{\partial}\right)$ *locally* as a twisted Dirac operator, until further notice, we now hypothesize (as is always locally the case)

(18.164)

Assumption: There are

1. a spin structure $P_{\mathrm{Spin}(n)} \to FM$ and principal $\mathrm{U}(1)$-bundle $P'_{\mathrm{U}(1)}$, and
2. an equivariant bundle map $\pi_\chi : P'_{\mathrm{U}(1)} \times_f P_{\mathrm{Spin}(n)} \to P_{\mathrm{Spin}^c(n)}$,

such that $\pi_\chi((p', p)\, g) = \pi_\chi((p', p))\, \chi(g)$ for all $g \in \mathrm{U}(1) \times \mathrm{Spin}(n)$.

We then have a commutative diagram of bundle morphisms

$$\begin{array}{ccc} & & P'_{\mathrm{U}(1)} \times_f P_{\mathrm{Spin}(n)} \\ & & \pi_\chi \downarrow \\ \mathrm{U}(M) & \overset{f_j}{\to} & P_{\mathrm{Spin}^c(n)} \\ \downarrow \wedge^n & \searrow^{f_{\det \times \iota}} & \downarrow \pi_{Sq \times c} \\ P_{\mathrm{U}(1)} & \overset{\mathrm{Id} \times I_n}{\to} & P_{\mathrm{U}(1)} \times_f FM \end{array}$$

equivariant with respect to the following diagram of group homomorphisms:

$$\begin{array}{ccc} & & \mathrm{U}(1) \times \mathrm{Spin}(n) \\ & & \downarrow \chi \\ \mathrm{U}(m) & \overset{j}{\to} & \mathrm{Spin}^c(n) \\ \downarrow \det & \searrow^{\det \times \iota} & \downarrow Sq \times c \\ \mathrm{U}(1) & \overset{\mathrm{Id} \times I}{\to} & \mathrm{U}(1) \times \mathrm{SO}(n). \end{array}$$

Since $((Sq \times c) \circ \chi)(z, g) = (Sq \times c)(zg) = \left(z^2, c(g)\right)$, under the composition $\pi_{Sq\times c} \circ \pi_\chi$, a point $(p', p) \in P'_{\mathrm{U}(1)} \times_f P_{\mathrm{Spin}(n)}$ is sent to a point $(q', q) \in P_{\mathrm{U}(1)} \times_f FM$ where q' is independent of the choice of p. Thus, $\pi_{Sq\times c} \circ \pi_\chi$ determines a bundle map, say

$$\pi_s : P'_{\mathrm{U}(1)} \to P_{\mathrm{U}(1)}.$$

Moreover, since $\pi_s(p'z) = \pi_s(p')\, z^2$, π_s is a double cover. Since π_s and $\mathbb{C} \otimes \mathbb{C} \cong \mathbb{C}$ are equivariant relative to the homomorphism $z \mapsto z^2$ of $\mathrm{U}(1)$, it follows that $L := P'_{\mathrm{U}(1)} \times_{\mathrm{U}(1)} \mathbb{C}$ is a line bundle, such that

$$L \otimes L \cong P_{\mathrm{U}(1)} \times_{\mathrm{U}(1)} \mathbb{C} = \Lambda^{n,0}(T_{\mathbb{C}}M) \cong \Lambda^{0,n}(T_{\mathbb{C}}M^*).$$

In other words, under the assumption (18.164), L is a square-root of the canonical line bundle for M. In general, L exists locally. As usual, let

$$\Sigma^{\pm}(M) := P_{\mathrm{Spin}(n)} \times_{\mathrm{Spin}(n)} \Sigma^{\pm}_{2m} \text{ and } \Sigma(M) := \Sigma^+(M) \oplus \Sigma^-(M).$$

We claim that there is a natural isomorphism

$$L \otimes \Sigma(M) \cong \Lambda^{0,*}(T_{\mathbb{C}}M^*). \tag{18.165}$$

Let $\mu : \mathrm{U}(1) \to \mathrm{U}(\mathbb{C})$ be given by $\mu(z)\, w = zw$. We then have the representation

$$\mu \otimes \rho : \mathrm{U}(1) \times \mathrm{Spin}(n) \to \mathrm{U}(\mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m})$$

There is also an isomorphism $\phi : \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m} \to \Sigma_{2m}$ given by $\phi(z \otimes \sigma) := z\sigma$. Note that ϕ equivariant in the sense that for each $(\zeta, g) \in U(1) \times \mathrm{Spin}(n)$,

$$\begin{array}{ccc} \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m} & \overset{(\mu\otimes\rho)(\zeta,g)}{\to} & \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m} \\ \phi \downarrow & & \downarrow \phi \\ \Sigma_{2m} & \overset{\rho_{\mathbb{C}}(\chi(\zeta,g))}{\to} & \Sigma_{2m} \end{array}$$

commutes. Indeed, for $z \otimes \sigma \in \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m}$,

$$\begin{aligned} \rho_{\mathbb{C}}(\chi(\zeta, g))(\phi(z \otimes \sigma)) &= \rho_{\mathbb{C}}(\zeta g)(z\sigma) = \zeta\rho(g)(z\sigma) = \zeta z \rho(g)(\sigma) \\ &= \phi(\zeta z \otimes \rho(g)(\sigma)) = \phi((\iota \otimes \rho)(\zeta, g)\, z \otimes \sigma). \end{aligned}$$

Consequently, the diagram (where $\Phi(T) := \phi \circ T \circ \phi^{-1}$)

$$\begin{array}{ccc} \mathrm{U}(1) \times \mathrm{Spin}(n) & \overset{\mu\otimes\rho}{\to} & \mathrm{U}(\mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m}) \\ \chi \downarrow & & \downarrow \Phi \\ \mathrm{Spin}^c(n) & \overset{\rho_{\mathbb{C}}}{\to} & \mathrm{U}(\Sigma_{2m}), \\ \uparrow j & & \downarrow \mathrm{Id} \\ \mathrm{U}(m) & \overset{\Lambda^*}{\to} & \mathrm{U}(\Lambda^*(\mathbb{C}^m)) \end{array}$$

commutes and extends (18.163). Since $\Lambda^* = \rho_{\mathbb{C}} \circ j : \mathrm{U}(m) \to \mathrm{U}(\Sigma_{2m}) = \mathrm{U}(\Lambda^*(\mathbb{C}^m))$, it follows (see Proposition 16.22, p.380) that

$$\mathrm{U}(M) \times_{\mathrm{U}(m)} \Lambda^*(\mathbb{C}^m) \cong P_{\mathrm{Spin}^c(n)} \times_{\mathrm{Spin}^c(n)} \Sigma_{2m}. \tag{18.166}$$

Moreover, since the diagrams

$$\begin{array}{ccc} \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m} & \overset{(\mu\otimes\rho)(\zeta,g)}{\to} & \mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m} \\ \phi \downarrow & & \downarrow \phi \\ \Sigma_{2m} & \overset{\rho_{\mathbb{C}}(\chi(\zeta,g))}{\to} & \Sigma_{2m} \end{array} \quad \text{and} \quad \begin{array}{ccc} U(1) \times \mathrm{Spin}(n) & \overset{\mu\otimes\rho}{\to} & U(\mathbb{C} \otimes_{\mathbb{R}} \Sigma_{2m}) \\ \chi \downarrow & & \downarrow \Phi \\ \mathrm{Spin}^c(n) & \overset{\rho_{\mathbb{C}}}{\to} & U(\Sigma_{2m}), \end{array}$$

commute, the isomorphism $\phi : \mathbb{C} \otimes \Sigma_{2m} \to \Sigma_{2m}$ is equivariant. Then (see Proposition 16.22, p.380)

$$\begin{aligned} L \otimes \Sigma(M) &= P_{\mathrm{U}(1)\times\mathrm{Spin}(n)} \times_{\mathrm{U}(1)\times\mathrm{Spin}(n)} (\mathbb{C} \otimes \Sigma_{2m}) \\ &\cong P_{\mathrm{Spin}^c(n)} \times_{\mathrm{Spin}^c(n)} \Sigma_{2m} \cong U(M) \times_{\mathrm{U}(m)} \Lambda^*(\mathbb{C}^m) \\ &\cong \Lambda^{0,*}(T_{\mathbb{C}} M^*), \end{aligned} \tag{18.167}$$

where we have used (18.166). This is the desired natural isomorphism (18.165). Note that via (18.167), $\Lambda^{0,*}(T_{\mathbb{C}} M^*)$ may be regarded as an associated bundle of $P_{\mathrm{U}(1)\times\mathrm{Spin}(n)}$. However, $\Lambda^{0,*}(T_{\mathbb{C}} M^*)$ is not associated to FM or $P_{\mathrm{Spin}(n)}$ and hence is not a Clifford module bundle in the strict sense of Definition (18.54). Nevertheless, under the Assumption (18.164), $\Lambda^{0,*}(T_{\mathbb{C}} M^*)$ is a *twisted* Clifford module bundle, since $\Sigma(M)$ is a Clifford module bundle and $\Lambda^{0,*}(T_{\mathbb{C}} M^*)$ is obtained by twisting $\Sigma(M)$ with L. To prove the Hirzebruch-Riemann-Roch Theorem, it remains to check the twisted Dirac operator $\mathcal{D}^{L,\Sigma_{2m}}$ corresponds to $\sqrt{2}\left(\bar\partial + \bar\partial^*\right)$. More precisely, the above isomorphism in (18.167), say

$$\phi_M : L \otimes \Sigma(M) \cong \Lambda^{0,*}(T_{\mathbb{C}} M^*),$$

provides us with a linear isomorphism

$$\begin{aligned}&\Gamma(\phi_M) : C^\infty(L\otimes\Sigma(M)) \to \Omega^{0,*}(M) \text{ given by}\\&\Gamma(\phi_M)[\psi](y) = \phi_M(\psi(y)).\end{aligned}$$

Of course the isomorphism ϕ_M of vector bundles also induces a corresponding isomorphism, still denoted by $\Gamma(\phi_M)$, between the spaces of exterior forms with values in these vector bundles. Now we need to show that, under $\Gamma(\phi_M)$, the twisted Dirac operator $\mathcal{D}^{L,\Sigma_{2m}}$ on $C^\infty(L\otimes\Sigma(M))$ corresponds to $\sqrt{2}\left(\bar\partial+\bar\partial^*\right)$ on $\Omega^{0,*}(M)$. In other words,

$$\Gamma(\phi_M)\left[\mathcal{D}^{L,\Sigma_{2m}}\psi\right] = \sqrt{2}\left(\bar\partial+\bar\partial^*\right)(\Gamma(\phi_M)[\psi])$$

Recall that

$$\mathcal{D}^{L,\Sigma_{2m}} := (1\otimes cl)\circ\nabla : C^\infty(L\otimes\Sigma(M)) \to C^\infty(L\otimes\Sigma(M)),$$

while

$$\sqrt{2}\left(\bar\partial+\bar\partial^*\right) := \Omega^{0,*}(M)\to\Omega^{0,*}(M).$$

We first need to show that $\Gamma(\phi_M)$ respects covariant differentiation

$$\Gamma(\phi_M)[\nabla\psi] = \nabla(\Gamma(\phi_M)[\psi]),$$

but this is a consequence of the facts that ϕ is equivariant and that the connection forms on $P_{\mathrm{U}(1)\times\mathrm{Spin}(n)}$ and $P_{\mathrm{Spin}^c(n)}$ are both pull-backs via

$$P_{\mathrm{U}(1)\times\mathrm{Spin}(n)} \overset{\pi_\chi}{\to} P_{\mathrm{Spin}^c(n)} \overset{\pi_{Sq\times c}}{\to} P_{\mathrm{U}(1)\times\mathrm{SO}(n)}$$

of the connection 1-form $\theta^{0,n}\oplus\theta$ on $P_{\mathrm{U}(1)\times\mathrm{SO}(n)}$, where θ is the Levi-Civita on FM and $\theta^{0,n}$ is the connection induced by θ on unitrary frame bundle for $\Lambda^{0,n}(T_{\mathbb{C}}M^*)$, namely $P_{\mathrm{U}(1)}$. We show

$$\Gamma(\phi_M)[\mathcal{D}\psi] = \sqrt{2}\left(\bar\partial+\bar\partial^*\right)(\Gamma(\phi_M)[\psi]) \tag{18.168}$$

as follows. It is enough to show that (18.168) holds at an arbitrary point $x\in M$. We continue to let $\left(x^1+iy^1,\ldots,x^m+iy^m\right)$ be complex coordinates about $x\in M$ such that $\left(\partial_{x^1},\partial_{y^1},\ldots,\partial_{x^m},\partial_{y^m}\right)$ are orthonormal at x relative to h_x, and $\nabla(\partial_{x^i}) = \nabla\left(\partial_{y^i}\right) = 0$ at x. It is convenient to replace our model $\Lambda^*(\mathbb{C}^m)$ for Σ_{2m} by $\Lambda^{0,*}\left((\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m)^*\right)$ which is unitarially equivalent to $\Lambda^*(\mathbb{C}^m)$ via the the $\mathbb{C}$-linear map

$$\begin{aligned}&\Lambda^*(\mathbb{C}^m)\cong\Lambda^{0,*}\left((\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m)^*\right) \text{ induced by}\\&\sqrt{2}\flat\circ\pi^{0,1} : \mathbb{C}^m\cong\Lambda^{0,1}\left((\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m)^*\right), \text{ i.e.,}\\&\Lambda^1(\mathbb{C}^m)\ni v\mapsto\sqrt{2}\flat\left(\tfrac{1}{2}(v+iJv)\right)\in\Lambda^{0,1}\left((\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m)^*\right)\end{aligned}$$

Note that $\left|\left(\flat\circ\pi^{0,1}\right)(v)\right| = \left|\pi^{0,1}(v)\right| = \frac{1}{\sqrt{2}}|v|$, so that $\left|\sqrt{2}\left(\flat\circ\pi^{0,1}\right)(v)\right| = |v|$. Clifford multiplication of $v\in\mathbb{R}^{2m}\cong\mathbb{C}^m$ on $\Lambda^{0,*}\left((\mathbb{C}\otimes_{\mathbb{R}}\mathbb{C}^m)^*\right)$ is then

$$\sqrt{2}\left(\left(\flat\circ\pi^{0,1}\right)(v)\wedge - \llcorner\left(\flat\circ\pi^{0,1}\right)(v)\right).$$

Any section of $\psi\in C^\infty(L\otimes\Sigma(M))$ is locally of the form

$$\psi = \sum_{k=1}^m\sum_{(j)_k}\psi_{j_1\ldots j_k}(\phi_M)^{-1}\left(d\bar z^{j_1}\wedge\cdots\wedge d\bar z^{j_k}\right)$$

and $\Gamma(\phi_M)[\psi] \in \Omega^{0,*}(M,\mathbb{C})$ is given by

$$\Gamma(\phi_M)[\psi] = \sum_{k=1}^{m}\sum_{(j)_k} \psi_{j_1\dots j_k} d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k},\ \psi_{j_1\dots j_k} \in C^\infty(M,\mathbb{C}).$$

Let

$$\mu_{\mathbb{C}} = \frac{dz^1 \wedge \cdots \wedge dz^m}{|dz^1 \wedge \cdots \wedge dz^m|} \text{ and } \bar{\mu}_{\mathbb{C}} = \frac{d\bar{z}^1 \wedge \cdots \wedge d\bar{z}^m}{|d\bar{z}^1 \wedge \cdots \wedge d\bar{z}^m|}$$

Then $\bar{\mu}_{\mathbb{C}}$ is a local frame for the canonical line bundle $\Lambda^{0,n}(T_{\mathbb{C}}M^*)$; i.e., $\bar{\mu}_{\mathbb{C}}$ is a local section of $P_{U(1)}$ and we let $\sqrt{\bar{\mu}_{\mathbb{C}}}$ denote the corresponding local section of $P'_{U(1)}$ so that

$$(\phi_M)^{-1}\left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right) = \sqrt{\bar{\mu}_{\mathbb{C}}} \otimes \left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right).$$

Since dz^k and $d\bar{z}^k$ are parallel at x, $\mu_{\mathbb{C}}$ and $\bar{\mu}_{\mathbb{C}}$ are also parallel at x. Moreover,

$$\left(\flat \circ \pi^{0,1}\right)(\partial_{x^k}) = \flat \tfrac{1}{2}\left(\partial_{x^k} + i\partial_{y^k}\right) = \flat\left(\partial_{\bar{z}^k}\right) = \tfrac{1}{2} d\bar{z}^k \text{ and}$$
$$\left(\flat \circ \pi^{0,1}\right)(\partial_{y^k}) = \flat \tfrac{1}{2}\left(\partial_{y^k} - i\partial_{x^k}\right) = \flat\left(-i\partial_{\bar{z}^k}\right) = i\flat\left(\partial_{\bar{z}^k}\right) = \tfrac{i}{2} d\bar{z}^k.$$

Thus at x,

$$\begin{aligned}\mathcal{D}^{L,\Sigma_{2m}}\psi &= \sum_{k=1}^{m}\sum_{(j)_k} \mathcal{D}^{L,\Sigma_{2m}}\left(\psi_{j_1\dots j_k}(\phi_M)^{-1}\left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right)\right) \\ &= \sum_{k=1}^{m}\sum_{(j)_k} \mathcal{D}^{L,\Sigma_{2m}}\left(\psi_{j_1\dots j_k}\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes \left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right)\right) \text{ and}\end{aligned}$$

$$\begin{aligned}&\tfrac{1}{\sqrt{2}}\mathcal{D}^{L,\Sigma_{2m}}\left(\psi_{j_1\dots j_k}\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes \left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right)\right) \\ &= \left(\flat\left(\partial^{0,1}_{x^l}\right)\wedge\ - \flat\left(\partial^{0,1}_{x^l}\right)\llcorner\right)\partial_{x^l}\left(\psi_{j_1\dots j_k}\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &+ \left(\flat\left(\partial^{0,1}_{y^l}\right)\wedge\ - \flat\left(\partial^{0,1}_{y^l}\right)\llcorner\right)\partial_{y^l}\left(\psi_{j_1\dots j_k}\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &= \left(\tfrac{1}{2}d\bar{z}^l \wedge\ - \tfrac{1}{2}d\bar{z}^l\llcorner\right)\partial_{x^l}\left(\psi_{j_1\dots j_k}\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &+ \left(\tfrac{i}{2}d\bar{z}^l \wedge\ - \tfrac{i}{2}d\bar{z}^l\llcorner\right)\partial_{y^l}\left(\psi_{j_1\dots j_k}\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &= \left(d\bar{z}^l \wedge\ - d\bar{z}^l\llcorner\right)\tfrac{1}{2}\left(\partial_{x^l} + i\partial_{y^l}\right)\left(\psi_{j_1\dots j_k}\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &= \partial_{\bar{z}^l}\left(\psi_{j_1\dots j_k}\right)\left(d\bar{z}^l \wedge\ - d\bar{z}^l\llcorner\right)\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}.\end{aligned}$$

Hence,

$$\begin{aligned}&\tfrac{1}{\sqrt{2}}\Gamma(\phi_M)\left[\mathcal{D}^{L,\Sigma_{2m}}\left(\psi_{j_1\dots j_k}\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes \left(d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right)\right)\right] \\ &= \partial_{\bar{z}^l}\left(\psi_{j_1\dots j_k}\right)\left(d\bar{z}^l \wedge\ - d\bar{z}^l\llcorner\right)\Gamma(\phi)\left(\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right) \\ &= \partial_{\bar{z}^l}\left(\psi_{j_1\dots j_k}\right)\left(d\bar{z}^l \wedge\ - d\bar{z}^l\llcorner\right)d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k} \\ &= \left(\bar{\partial} + \bar{\partial}^*\right)\left(\psi_{j_1\dots j_k} d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right) \\ &= \left(\bar{\partial} + \bar{\partial}^*\right)\left(\Gamma(\phi_M)\left[\psi_{j_1\dots j_k}\sqrt{\bar{\mu}_{\mathbb{C}}} \otimes d\bar{z}^{j_1} \wedge \cdots \wedge d\bar{z}^{j_k}\right]\right)\end{aligned}$$

and (18.168) holds by linearity. Since

$$\Sigma^+_{2m} = \Lambda^{\text{even}}(\mathbb{C}^m) := \bigoplus\nolimits_{l \text{ even}} \Lambda^l(\mathbb{C}^m), \text{ while}$$
$$\Sigma^-_{2m} = \Lambda^{\text{odd}}(\mathbb{C}^m) := \bigoplus\nolimits_{l \text{ odd}} \Lambda^l(\mathbb{C}^m),$$

we have

$$L\otimes\Sigma^{+}(M)\overset{\phi_M}{\cong}\Lambda^{0,\mathrm{even}}(T_{\mathbb{C}}M^{*})\ \text{ and }\ L\otimes\Sigma^{-}(M)\overset{\phi_M}{\cong}\Lambda^{0,\mathrm{odd}}(T_{\mathbb{C}}M^{*}),\ \text{ and}$$
$$\Gamma(\phi_M)\circ\mathcal{D}^{L,\Sigma_{2m}}\circ\Gamma(\phi_M)^{-1}=\sqrt{2}\left(\bar\partial+\bar\partial^{*}\right):\Omega^{0,\mathrm{even}}(M)\to\Omega^{0,\mathrm{odd}}(M).$$

Hence under Assumption (18.164), we have exhibited $\sqrt{2}\left(\bar\partial+\bar\partial^{*}\right)$ as a twisted Dirac operator, and locally so if (18.164) fails globally. Thus, on a suitable neighborhood B about any point (e.g., a normal ball), the local index density for $\sqrt{2}\left(\bar\partial+\bar\partial^{*}\right)$ coincides with that for $\mathcal{D}^{L,\Sigma_{2m}}$ on B, where $P'_{\mathrm{U}(1)}$, $P_{\mathrm{Spin}(n)}$, L and $\Sigma(B)$ may only exist over B. Let θ be the Levi-Civita connection 1-form on FM, let $\widetilde\theta$ be the connection 1-form $c'^{-1}\left(C^{*}\left(\theta|_{FB}\right)\right)$ on $P_{\mathrm{Spin}(n)}$, let $\theta^{0,m}$ be the connection form on $P_{\mathrm{U}(1)}=\mathrm{U}\left(\Lambda^{0,m}(T_{\mathbb{C}}M^{*})\right)$ induced by $\theta|_{\mathrm{U}(M)}$, and let $\widetilde\theta^{0,m}=\frac{1}{2}\pi_s^{*}\theta^{0,m}$ be the related connection on $P'_{\mathrm{U}(1)}$ where $\pi_s:P'_{\mathrm{U}(1)}\to P_{\mathrm{U}(1)}$ is the double cover. The local index density form for $\sqrt{2}\left(\bar\partial+\bar\partial^{*}\right)|_{\Omega^{0,\mathrm{even}}(M)}$ is the same as that for $\mathcal{D}^{L,\Sigma_{2m}}$ on B, namely

$$\begin{aligned}
&\mathbf{ch}\left(L,\widetilde\theta^{0,m}\right)\wedge\widehat{\mathbf{A}}(M,\theta)\\
&=\mathbf{ch}\left(L,\widetilde\theta^{0,m}\right)\wedge\mathbf{ch}\left(P_{\mathrm{Spin}(n)},\widetilde\theta,\rho'\right)\wedge\widetilde{\mathbf{A}}(M,\theta)\\
&=\left(\mathbf{ch}\left(P'_{\mathrm{U}(1)},\widetilde\theta^{0,m},\mu'\right)\wedge\mathbf{ch}\left(P_{\mathrm{Spin}(n)},\widetilde\theta,\rho'\right)\right)\wedge\widetilde{\mathbf{A}}(M,\theta)\\
&=\mathbf{ch}\left(P'_{\mathrm{U}(1)}\times P_{\mathrm{Spin}(n)},\widetilde\theta^{0,m}\oplus\widetilde\theta,(\mu\otimes\rho)'\right)\wedge\widetilde{\mathbf{A}}(M,\theta)\\
&=\mathbf{ch}\left(\mathrm{U}(M),\theta|_{\mathrm{U}(M)},\Lambda^{0,*\prime}\right)\wedge\widetilde{\mathbf{A}}(M,\theta)\\
&=\mathbf{ch}\left(\Lambda^{0,*}\left((T_{\mathbb{C}}M)^{*}\right)\right)\wedge\widetilde{\mathbf{A}}(M,\theta)=\mathbf{Td}(TM,\theta).
\end{aligned}\tag{18.169}$$

Thus, we obtain

THEOREM 18.71 (Hirzebruch-Riemann-Roch Theorem). *Let M be a compact Kähler manifold. Then*

$$\mathrm{Index}\left(\left(\bar\partial+\bar\partial^{*}\right):\Omega^{0,even}(M)\to\Omega^{0,odd}(M)\right)=\mathbf{Td}(TM)[M].$$

There is also a twisted version of the Hirzebruch-Riemann-Roch Theorem. Let $E\to M$ be a Hermitian vector bundle with a covariant derivative ∇^{E} arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. We have the spaces $\Omega^{p,q}(M,E):=C^{\infty}\left(E\otimes\Lambda^{p,q}(T_{\mathbb{C}}M^{*})\right)$ of E-valued forms of bidegree (p,q), which are locally of the form (in multi-index notation)

$$\phi=\frac{1}{p!q!}\sum_{(j),(k)}f_{(j)_p;(k)_q}\otimes dz^{(j)_p}\wedge d\bar z^{(k)_q},$$

where $f_{(j)_p;(k)_q}$ is a local section of E. Moreover, there are global operators

$$\partial^{E}:\Omega^{p,q}(M,E)\to\Omega^{p+1,q}(M,E)\ \text{ and }\ \bar\partial^{E}:\Omega^{p,q}(M,E)\to\Omega^{p,q+1}(M,E)$$

determined locally by

$$\partial^E \phi = \frac{1}{p!q!} \sum_{h,(j),(k)} \nabla^E_{\partial_{z^h}} \left(f_{(j)_p;(k)_q} \right) dz^h \wedge dz^{(j)_p} \wedge d\bar{z}^{(k)_q}, \text{ and}$$

$$\bar{\partial}^E \phi = \frac{1}{p!q!} \sum_{h,(j),(k)} \nabla^E_{\partial_{\bar{z}^h}} \left(f_{(j)_p;(k)_q} \right) d\bar{z}^h \wedge dz^{(j)_p} \wedge d\bar{z}^{(k)_q}. \tag{18.170}$$

Note that just as $\bar{\partial} + \bar{\partial}^*$ was shown to be locally a twisted Dirac operator (twisted by L), $\bar{\partial}^E + \bar{\partial}^{E*}$ is also locally a twisted Dirac operator (twisted by $E \otimes L$).

THEOREM 18.72 (Twisted Hirzebruch-Riemann-Roch Theorem). *Let M be a compact Kähler manifold and let $E \to M$ be a Hermitian vector bundle. Then*

$$\text{Index}\left(\bar{\partial}^E + \bar{\partial}^{E*} : \Omega^{0,even}(M,E) \to \Omega^{0,odd}(M,E)\right) = \left(\mathbf{ch}(E) \smile \mathbf{Td}(M)\right)[M].$$

PROOF. Using the local computation (18.169), we get

$$\begin{aligned}
&\mathbf{ch}\left(E \otimes L, (\varepsilon, \widetilde{\theta}^{0,m})\right) \wedge \widehat{\mathbf{A}}(M,\theta) \\
&= \mathbf{ch}(E,\varepsilon) \wedge \left(\mathbf{ch}\left(L, \widetilde{\theta}^{0,m}\right) \wedge \widehat{\mathbf{A}}(M,\theta)\right) \\
&= \mathbf{ch}(E,\varepsilon) \wedge \mathbf{Td}(TM,\theta).
\end{aligned}$$

□

There is a perhaps more familiar version of Theorem 18.72 which applies when $E \to M$ is holomorphic. We say that the complex vector bundle of complex fiber dimension N $\pi_E : E \to M$ over the complex manifold M is *holomorphic* if E has the structure of a complex manifold, and about each point $x \in M$ there is a neighborhood B and a biholomorphic vector bundle map

$$\phi_B : \pi_E^{-1}(B) \to B \times \mathbb{C}^N.$$

which we call a *holomorphic trivialization.* If E has complex fiber dimension N, then the $\mathbb{C}$-linear frames of E (linear maps u from $\mathbb{C}^N$ to E_x for $x \in M$) form a principal $\mathrm{GL}(N,\mathbb{C})$-bundle, say $\widetilde{\pi}_E : \mathrm{GL}(E) \to M$, and if E has a Hermitian structure, $\mathrm{U}(E) \to M$ is a principal subbundle. Although $\mathrm{U}(E)$ does not generally have a complex structure, we show that $\mathrm{GL}(E)$ does. Each holomorphic trivialization ϕ_B determines a section $u_\phi : B \to \mathrm{GL}(E)$ given (for $\mathbf{z} \in \mathbb{C}^N$) by

$$u_\phi(y)(\mathbf{z}) = \phi_B^{-1}(y,\mathbf{z}).$$

An arbitrary frame $u \in \mathrm{GL}(E)$ at $y \in B$ is of the form $u_\phi(y) \circ A_\phi(u)$ for some $A_\phi(u) \in \mathrm{GL}(N,\mathbb{C})$. We then obtain a map

$$\tilde{\phi}_B : \mathrm{GL}(E)|_B \to B \times \mathrm{GL}(N,\mathbb{C}) \text{ given by } \tilde{\phi}_B(u) = (y, A_\phi(u))$$

If $\phi'_{B'} : \pi_E^{-1}(B') \to B' \times \mathbb{C}^N$ is another holomorphic trivialization, then

$$u_\phi(y) \circ A_\phi(u) = u = u_{\phi'}(y) \circ A_{\phi'}(u), \text{ and so}$$

$$\left(\tilde{\phi}'_{B'} \circ \tilde{\phi}_B^{-1}\right)(y, A_\phi(u)) = \tilde{\phi}'_{B'}(u) = \left(y, u_{\phi'}(y)^{-1} \circ u_\phi(y) \circ A_\phi(u)\right).$$

Now $\tilde{\phi}'_{B'} \circ \tilde{\phi}_B^{-1} : (B \cap B') \times \mathrm{GL}(N, \mathbb{C}) \to (B \cap B') \times \mathrm{GL}(N, \mathbb{C})$ is holomorphic, since for $(y, C) \in (B \cap B') \times \mathrm{GL}(N, \mathbb{C})$,

$$\begin{aligned}\left(\tilde{\phi}'_{B'} \circ \tilde{\phi}_B^{-1}\right)(y, C) &= \left(y, u_{\phi'}(y)^{-1} \circ u_\phi(y) \circ C\right) \\ &= \left(y, \left(\phi'^{-1}_{B'}(y, \cdot)\right)^{-1} \circ \phi_B^{-1}(y, \cdot) \circ C\right) \end{aligned} \tag{18.171}$$

is a holomorphic function of (y, C). We now define an operator

$$\bar{\partial}_E : \Omega^{p,q}(M, E) \to \Omega^{p,q+1}(M, E)$$

that generalizes $\bar{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$ which is the special case $E = M \times \mathbb{C}$. If $f_1, \ldots, f_N$ is the standard basis for $\mathbb{C}^N$ and $s_j(x) := \phi_B^{-1}(x, f_j)$, then locally, $\bar{\partial}_E$ is given by

$$\begin{aligned}&\bar{\partial}_E\left(\sum\nolimits_{k,(i)^p,(j)^q} \alpha_{k,(i)(j)} s_k \otimes dz^{(i)_p} \wedge d\bar{z}^{(j)_q}\right) \\ &:= \sum\nolimits_{k,(i),(j)} \bar{\partial}\left(\alpha_{k,(i)(j)}\right) s_k \otimes dz^{(i)} \wedge d\bar{z}^{(j)}. \end{aligned} \tag{18.172}$$

This is well-defined (independent of local holomorphic trivialization and coordinates), since $\bar{\partial}$ kills holomorphic functions. By definition, the kernel of $\bar{\partial}_E$ consists of the *holomorphic sections of* E. Note that $\bar{\partial}_E \circ \bar{\partial}_E = 0$ and so we have an exact sequence

$$\Omega^{0,0}(M, E) \overset{\bar{\partial}_E}{\to} \cdots \overset{\bar{\partial}_E}{\to} \Omega^{0,q}(M, E) \overset{\bar{\partial}_E}{\to} \Omega^{0,q+1}(M, E) \overset{\bar{\partial}_E}{\to} \cdots,$$

along with cohomology groups

$$H^q(\mathcal{O}_E) := \frac{\mathrm{Ker}\left(\bar{\partial}_E|_{\Omega^{0,q}(M,E)}\right)}{\bar{\partial}_E\left(\Omega^{0,q-1}(M, E)\right)}.$$

Since $\bar{\partial}_E \circ \bar{\partial}_E = 0$, $\bar{\partial}_E$ is like a covariant derivative operator for a connection 1-form on $\mathrm{GL}(E)$ with zero curvature. Indeed, on $B \times \mathrm{GL}(N, \mathbb{C})$ there is a standard flat connection 1-form given in terms of coordinates (y, C) by $C^{-1}dC$. On $B' \times \mathrm{GL}(N, \mathbb{C})$, with coordinates (y, C'), we have $C'^{-1}dC'$. In view of (18.171),

$$\begin{aligned}C' &= \left(\phi'^{-1}_{B'}(y, \cdot)\right)^{-1} \circ \phi_B^{-1}(y, \cdot) \circ C \text{ and so} \\ dC' &= d\left(\left(\phi'^{-1}_{B'}(y, \cdot)\right)^{-1} \circ \phi_B^{-1}(y, \cdot)\right) \circ C' + \left(\phi'^{-1}_{B'}(y, \cdot)\right)^{-1} \circ \phi_B^{-1}(y, \cdot) \circ dC\end{aligned}$$

Because of the first term, it is not true that $C'^{-1}dC' = C^{-1}dC$. However, the $(0,1)$-component of the first term is 0 since $\left(\phi'^{-1}_{B'}(y, \cdot)\right)^{-1} \circ \phi_B^{-1}(y, \cdot)$ is a holomorphic function of y, and so

$$C'^{-1}\bar{\partial}C' = C^{-1}\bar{\partial}C \text{ on } (B \cap B') \times \mathrm{GL}(N, \mathbb{C}).$$

In particular, the equations $\bar{\partial}C' = 0$ and $\bar{\partial}C = 0$ define the same horizontal distribution of subspaces of $T^{0,1}\left(\widetilde{\pi}_E^{-1}(B \cap B')\right) \subseteq T^{0,1}(\mathrm{GL}(E))$, of $\mathbb{C}$-dimension m, and so we obtain a well-defined horizontal distribution of $T^{0,1}(\mathrm{GL}(E))$. A genuine connection on $\mathrm{GL}(E)$ determines a horizontal distribution of real subspaces $T(\mathrm{GL}(E))$ of $\mathbb{R}$-dimension $2m$, and hence a horizontal distribution $T_{\mathbb{C}}(\mathrm{GL}(E))$ of $\mathbb{C}$-dimension $2m$. In essence, $C^{-1}\bar{\partial}C$ is the well-defined $(0,1)$-component of an ill-defined flat connection $C^{-1}dC$, and $\bar{\partial}_E$ is the well-defined $(0,1)$-component of the "ill-defined covariant derivative" for $C^{-1}dC$. To avoid additional notation, we still use $C^{-1}\bar{\partial}C$ to denote the global $(0,1)$-form given locally as $C^{-1}\bar{\partial}C$. Proposition

18.76 below asserts that the Hermitian metric on E determines a standard well-defined $(1,0)$-component $\omega^{1,0}$ whose sum with $C^{-1}\bar{\partial}C$ is a "complex connection 1-form"

$$\omega = \omega^{1,0} \oplus C^{-1}\bar{\partial}C \in \Omega^1_{\mathbb{C}}\left(\mathrm{GL}\left(E\right), \mathfrak{gl}_{\mathbb{C}}\left(n, \mathbb{C}\right)\right),$$

where $\mathfrak{gl}_{\mathbb{C}}(n,\mathbb{C}) = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{gl}(n,\mathbb{C})$. By "complex connection 1-form", we mean that although ω is not necessarily the complex extension of a (real) connection 1-form in $\Omega^1(\mathrm{GL}(E), \mathfrak{gl}(n,\mathbb{C}))$, at each point $u \in \mathrm{GL}(E)$ the form ω is defined on $T_{\mathbb{C}}\,\mathrm{GL}(E)_u := \mathbb{C} \otimes_{\mathbb{R}} T_u\,\mathrm{GL}(E)$ and satisfies the usual relations

$$\omega_u\left(\tfrac{d}{dt}\left(u \exp tA\right)\right) = A \text{ and } R_g^*\omega = g^{-1}\omega g,$$

for $A \in \mathfrak{gl}(N,\mathbb{C})$ and $g \in \mathrm{GL}(N,\mathbb{C})$.

Definition 18.73. Let ω be a complex connection 1-form on $\mathrm{GL}(E)$ for a holomorphic vector bundle E. Then ω is **compatible with the complex structure on** E if $\omega^{0,1} = C^{-1}\bar{\partial}C$.

Let $Herm(N)$ be the space of all Hermitian $N \times N$ matrices and let r : $\mathrm{GL}(N,\mathbb{C}) \to \mathrm{End}(Herm(N))$ be the representation given by

$$r(g)(\eta) = \left(g^{-1}\right)^T \eta \bar{g}^{-1}, \text{ for } g \in \mathrm{GL}(N,\mathbb{C}) \text{ and } \eta \in Herm(N).$$

Note that $r(g)$ is just pull-back by g^{-1}, since

$$X^T\left(r\left(g\right)\left(\eta\right)\right)\bar{Y} = X^T\left(g^{-1}\right)^T \eta \bar{g}^{-1}\bar{Y} = \left(g^{-1}X\right)^T \eta \left(\overline{g^{-1}Y}\right).$$

Definition 18.74. If E has a Hermitian structure H, then ω is a **Hermitian connection for** H if $D^\omega \widetilde{H} = 0$, where

$$\begin{aligned}\widetilde{H} &\in \bar{\Omega}^0\left(\mathrm{GL}\left(E\right), Herm\left(N\right)\right)\\ &= \left\{f \in C^\infty\left(\mathrm{GL}\left(E\right), Herm\left(N\right)\right) : f\left(ug\right) = r\left(g^{-1}\right)f\left(u\right)\right\}.\end{aligned}$$

is the equivariant function corresponding to H.

Remark 18.75. If $\nabla^E : C^\infty(M,E) \to \Omega^1(M,E)$ is the covariant derivative arising from ω, then ω is compatible with the complex structure on E iff

$$\pi^{0,1} \circ \nabla^E = \bar{\partial}_E \text{ or } \nabla^E = \bar{\partial}_E + \left(\nabla^E\right)^{1,0}.$$

Moreover, ω is a Hermitian connection for H if

$$d\left(H\left(V,W\right)\right) = H\left(\nabla^E V, W\right) + H\left(V, \nabla^E W\right).$$

The connection in the following Proposition is variously known as the *Chern connection*, the $(1,0)$-*connection*, the *metric connection*, and the *canonical connection* for the Hermitian, holomorphic vector bundle E.

Proposition 18.76. *Let M be a complex manifold and let $E \to M$ be a holomorphic vector bundle with Hermitian metric H. Then there is a unique complex connection 1-form ω on $\mathrm{GL}(E)$ which is both compatible with the complex structure on E and a Hermitian connection for H.*

Proof. First we prove uniqueness. Suppose that ω exists. Then

$$0 = D^\omega \widetilde{H} = d\widetilde{H} + r'(\omega)\widetilde{H} = d\widetilde{H} - \omega\widetilde{H} - \widetilde{H}\omega. \tag{18.173}$$

Note also that

$$\widetilde{H}(ug) = r\left(g^{-1}\right)\left(\widetilde{H}(u)\right) = g^T \widetilde{H}(u)\,\bar{g}$$
$$\Rightarrow d\widetilde{H}\left(A^*_u\right) = A^T \widetilde{H}(u) + \widetilde{H}(u)\,\bar{A}. \tag{18.174}$$

Using the complex structure on $\mathrm{GL}(E)$, ω decomposes as $\omega = \omega^{1,0} \oplus \omega^{0,1}$. Observe that $\omega^{1,0}$ and $\omega^{0,1}$ have values in $\mathbb{C} \otimes_{\mathbb{R}} \mathfrak{gl}(n,\mathbb{C})$, and any $A \in \mathfrak{gl}(n,\mathbb{C})$ splits as $A^{1,0} + A^{0,1}$, where $\overline{A^{1,0}} = A^{0,1} \in \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{gl}(n,\mathbb{C})$. Since ω is compatible with the complex structure on E, $\omega^{0,1} = C^{-1}\bar{\partial}C$ in terms of the "coordinates" (y, C) in $B \times \mathrm{GL}(N,\mathbb{C})$. Then $\omega^{1,0}$ is uniquely determined by taking the $(1,0)$-part of (18.173), namely

$$0 = \partial\widetilde{H} - \left(\omega^{1,0}\right)^T \widetilde{H} - \widetilde{H}\overline{\omega^{0,1}} \Rightarrow \left(\omega^{1,0}\right)^T \widetilde{H} = \partial\widetilde{H} - \widetilde{H}\overline{\omega^{0,1}}$$
$$\Rightarrow \omega^{1,0} = \left(\left(\partial\widetilde{H} - \widetilde{H}\overline{\omega^{0,1}}\right)\left(\widetilde{H}\right)^{-1}\right)^T = \widetilde{H}^{-1T}\left(\partial\widetilde{H}^T - \overline{\omega^{0,1}}^T \widetilde{H}^T\right).$$

Since $\overline{\omega^{0,1}}^T = \overline{\left(C^{-1}\bar{\partial}C\right)}^T = \partial\left(\bar{C}^T\right)\left(\bar{C}^T\right)^{-1}$, we get

$$\omega^{1,0} = \widetilde{H}^{-1T}\left(\partial\widetilde{H}^T - \partial\left(\bar{C}^T\right)\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right). \tag{18.175}$$

Thus, ω is unique. We mention that taking the $(0,1)$-part of (18.173) and using $\overline{\widetilde{H}} = \widetilde{H}^T$ gives the same result. Now if $\omega^{1,0}$ is *defined* by (18.175), then the sum

$$\begin{aligned}\omega &:= \widetilde{H}^{-1T}\left(\partial\widetilde{H}^T - \partial\left(\bar{C}^T\right)\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right) \oplus C^{-1}\bar{\partial}C \\ &= \widetilde{H}^{-1T}\left(\partial\left(\bar{C}^T\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right) - \partial\left(\bar{C}^T\right)\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right) \oplus C^{-1}\bar{\partial}C \\ &= \widetilde{H}^{-1T}\bar{C}^T\partial\left(\left(\bar{C}^T\right)^{-1}\right)\widetilde{H}^T \oplus C^{-1}\bar{\partial}C\end{aligned}$$

is a complex connection on $\mathrm{GL}(E)$. Indeed, we have $R_g^*\omega = g^{-1}\omega g$, since

$$(Cg)^{-1}\,\bar{\partial}(Cg) = g^{-1}\left(C^{-1}\bar{\partial}C\right)g, \text{ and}$$

$$\begin{aligned}&\left(r\left(g^{-1}\right)\widetilde{H}\right)^{-1T}\left(\partial\left(r\left(g^{-1}\right)\widetilde{H}\right)^T - \partial\left(\overline{Cg}^T\right)\left(\overline{Cg}^T\right)^{-1}\left(r\left(g^{-1}\right)\widetilde{H}\right)^T\right) \\ &= \left(g^T\widetilde{H}\bar{g}\right)^{-1T}\left(\partial\left(g^T\widetilde{H}\bar{g}\right)^T - \partial\left(\overline{Cg}^T\right)\left(\overline{Cg}^T\right)^{-1}\left(g^T\widetilde{H}\bar{g}\right)^T\right) \\ &= \left(g^{-1}\widetilde{H}^{-1T}\bar{g}^{-1T}\right)\left(\bar{g}^T\partial\left(\widetilde{H}^T\right)g - \bar{g}^T\partial\left(\overline{C}^T\right)\left(\overline{C}^{T-1}\bar{g}^{T-1}\right)\left(\bar{g}^T\widetilde{H}^T g\right)\right) \\ &= g^{-1}\widetilde{H}^{-1T}\left(\partial\widetilde{H}^T - \bar{g}^T\partial\left(\bar{C}^T\right)\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right)g.\end{aligned}$$

To show $\omega(A^*) = A$ for $A \in \mathfrak{gl}(n,\mathbb{C})$, note that (18.174) implies

$$\partial\widetilde{H}^T\left(A^*_u\right) = \widetilde{H}^T A^{1,0} + \left(\bar{A}^T\right)^{1,0}\widetilde{H}^T, \text{ and then}$$

$$\begin{aligned}\widetilde{\varepsilon}(A^*) &= \widetilde{H}^{-1T}\left(\partial\widetilde{H}^T(A^*) - \partial\left(\bar{C}^T\right)(A^*)\left(\bar{C}^T\right)^{-1}\widetilde{H}^T\right) \oplus C^{-1}\bar{\partial}C(A^*) \\ &= \widetilde{H}^{-1T}\left(\left(\widetilde{H}^T A^{1,0} + \left(\bar{A}^T\right)^{1,0}\widetilde{H}^T\right) - \left(\bar{A}^T\right)^{1,0}\widetilde{H}^T\right) \oplus A^{0,1} \\ &= A^{1,0} \oplus A^{0,1} = A.\end{aligned}$$

□

If E is a holomorphic Hermitian bundle over a Kähler manifold and ∇^E is the covariant derivative of a complex connection 1-form on $\mathrm{GL}(E)$ which is a Hermitian connection and compatible with the complex structure on E, then $\bar{\partial}^E$ defined in (18.170) agrees with $\bar{\partial}_E$ defined in (18.172). Hence, by a Hodge theoretic proof strictly analogous to that of Theorem 18.67, when E is Hermitian and holomorphic we have

$$H^q(\mathcal{O}_E) \cong \mathcal{H}^q(E) := \mathrm{Ker}\left(\left(\bar{\partial}^E + \bar{\partial}^{E\,*}\right)|_{\Omega^{0,q}(M,E)}\right).$$

The following theorem is then immediate.

THEOREM 18.77 (Holomorphic Hirzebruch-Riemann-Roch Theorem). *Let M be a compact Kähler manifold and let $E \to M$ be a Hermitian holomorphic vector bundle. Then the holomorphic Euler characteristic of E is given by*

$$\begin{aligned}\chi(E) &:= \sum\nolimits_{q=0}^{m} (-1)^q \dim H^q(\mathcal{O}_E) \\ &= \mathrm{Index}\left(\bar{\partial}^E + \bar{\partial}^{E*} : \Omega^{0,even}(M,E) \to \Omega^{0,odd}(M,E)\right) \\ &= \left(\mathbf{ch}(E) \smile \mathbf{Td}(M)\right)[M].\end{aligned}$$

In particular, $\chi(E)$, which seems to depend on the holomorphic structure of E, actually only depends on the topology of E.

CHAPTER 19

Seiberg-Witten Theory

1. Background and Survey

Since the early 1980s, S. K. Donaldson and others have been proving results about smooth 4-manifolds using moduli spaces of instantons (self-dual SU(2)*-connections modulo gauge transformations). The approach inspired by Seiberg and Witten (see* [**SW**]*) uses monopole moduli spaces (twisted spinor fields paired with abelian* U(1)*-connections, modulo gauge transformations). Seiberg-Witten theory not only provides simpler proofs of most of Donaldson's results, but has generated many new results, notably Taubes' work concerning symplectic manifolds (e.g., see* [**Tau94**] *and* [**Tau95**]*). We first introduce some notation, and review the major results.*

For a compact, simply-connected (and connected), oriented 4-manifold X, let $[X] \in H^4(X;\mathbb{Z})$ be the fundamental class determined by the orientation. For the standard definitions and results from algebraic topology in what follows, we refer the reader to [**Gr**] and [**Sp**]. Let

$$Q^X : H^2(X;\mathbb{Z}) \times H^2(X;\mathbb{Z}) \to \mathbb{Z}$$

be the symmetric bilinear form (the so-called *intersection form* of X) given by

$$Q^X(\alpha,\beta) := [X] \frown (\alpha \smile \beta) = \langle \alpha \smile \beta, [X]\rangle\,, \tag{19.1}$$

where $\smile$ and $\frown$ are the cup and cap products, and $\langle\cdot,\cdot\rangle$ denotes the pairing between homology and cohomology, namely the Kronecker product

$$\langle\cdot,\cdot\rangle : H^q(X;\mathbb{Z}) \times H_q(X;\mathbb{Z}) \to \mathbb{Z}.$$

Note that if the orientation of X is switched, then Q^X changes to $-Q^X$. It is known (see [**Gr**, p.180]) that $\alpha \mapsto \langle\alpha,\cdot\rangle$ defines a surjection

$$h : H^q(X;\mathbb{Z}) \to \mathrm{Hom}\,(H_q(X;\mathbb{Z});\mathbb{Z})\,,$$

but in general there may be a kernel. The Universal Coefficient Theorem (see [**Gr**, p.189]) yields a split exact sequence

$$0 \to \mathrm{Ext}\,(H_{q-1}(X;\mathbb{Z});\mathbb{Z}) \to H^q(X;\mathbb{Z}) \xrightarrow{h} \mathrm{Hom}\,(H_q(X;\mathbb{Z}),\mathbb{Z}) \to 0.$$

Since we assumed that X is simply-connected, $H_1(X;\mathbb{Z}) = 0$, and so

$$h : H^2(X;\mathbb{Z}) \cong \mathrm{Hom}(H_2(X;\mathbb{Z}),\mathbb{Z}), \tag{19.2}$$

which is a free, abelian $\mathbb{Z}$-module (i.e., a *lattice*). In the smooth category, $Q^X(\alpha,\beta) = \int_X \tilde\alpha\wedge\tilde\beta$, where $\tilde\alpha$ and $\tilde\beta$ are closed 2-forms representing (non-torsion) integral classes α and β, and the integral is defined via the given orientation on X. Poincaré duality states that we have an isomorphism

$$P := [X] \frown (\cdot)\ : H^i(X;\mathbb{Z}) \cong H_{4-i}(X;\mathbb{Z}) \tag{19.3}$$

Thus, $H_2(X;\mathbb{Z}) \cong H^2(X;\mathbb{Z})$ is also a lattice. Hence, we may alternatively use

$$Q_X : H_2(X;\mathbb{Z}) \times H_2(X;\mathbb{Z}) \to \mathbb{Z}.$$

The term "intersection form" follows from the fact that if two classes in $H_2(X;\mathbb{Z})$ are represented by a compact, oriented embedded surfaces which intersect transversely, then Q_X on this pair is the intersection number of the two surfaces (i.e., the algebraic number of signed intersections, where the sign is ± 1, depending on whether the combined orientation of surfaces at an intersection point agrees with that of X).

Proposition 19.1. *If X is a compact, simply-connected, oriented 4-manifold, then any class in $H_2(X;\mathbb{Z})$ is a homological sum of compact, oriented embedded surfaces.*

Proof. (sketch) Since X is simply-connected, the Hurewicz Isomorphism Theorem (see [**Sp**, p.398]) yields an isomorphism $\pi_2(X, x_0) \cong H_2(X;\mathbb{Z})$. This implies that the generators of $H_2(X;\mathbb{Z})$ can be represented by maps $S^2 \to X$. Such a map can be approximated (within the same homotopy class) by an immersion f with transverse self-intersections. We can find a small 4-ball B_i about each self-intersection i (say $i \in \{1, \ldots, g\}$), such that $f\left(S^2\right) \cap B_i$ consists of two disks D_{i_1} and D_{i_2} intersecting at a point. Note that $f\left(S^2\right) \cap \partial B_i$ is the union of the disjoint circles ∂D_1 and ∂D_2 which can be joined by an embedded tube $T_i \subseteq B_i$ (e.g., $\left((1-u)\, e^{i2\pi t}, u e^{i2\pi t}\right)$, $t, u \in [0,1]$). Replacing $D_{i_1} \cup D_{i_2}$ by T_i for each i, we obtain an embedded surface Σ. Note that homology classes of Σ and $f\left(S^2\right)$ in $H_2(X;\mathbb{Z})$ are the same in $H_2(X;\mathbb{Z})$, since these classes trivially determine the same element in $H_2(X - \cup_i B;\mathbb{Z})$ which is isomorphic to $H_2(X;\mathbb{Z})$ because attaching 4-balls to $X - \cup_i B$ does not alter H_2 (see [**Gr**, p.118]). Moreover, since Σ is obtained by attaching g handles to $S^2 - \cup_{i=1}^{g} (D_{i_1} \cup D_{i_2})$, Σ is an oriented surface of genus g. □

Exercise 19.2. The **connected sum** $M_1 \# M_2$ of two n-manifolds M_1 and M_2 is the manifold obtained by removing a ball from each and gluing the exposed spherical boundaries together. Show that for two compact, simply-connected, oriented 4-manifolds X_1 and X_2, we have

$$Q_{X_1 \# X_2} \cong Q_{X_1} \oplus Q_{X_2}. \tag{19.4}$$

If needed, see [**MilH**, p.103-5].

Since X is simply-connected,

$$H^1(X;\mathbb{Z}) \cong \operatorname{Hom}(H_1(X;\mathbb{Z}),\mathbb{Z}) \oplus \operatorname{Tor}(H_0(X;\mathbb{Z}) = \{0\}$$

$$\{0\} = H^1(X;\mathbb{Z}) \overset{P}{\cong} H_3(X;\mathbb{Z}) \text{ and } H^3(X;\mathbb{Z}) \overset{P}{\cong} H_1(X;\mathbb{Z}) = \{0\}.$$

Thus, Q_X provides all of the homological/cohomological information about X. Indeed, Milnor (in [**Mil58**]) announced that the oriented homotopy type of a simply-connected, compact, oriented four-manifold X is completely determined by Q_X. A proof is given in [**MilH**, p.103-5]. The form Q_X is nondegenerate over $\mathbb{Z}$, meaning that

$$w \mapsto Q_X(w, \cdot) \text{ defines an isomorphim } H_2(X;\mathbb{Z}) \cong \operatorname{Hom}(H_2(X;\mathbb{Z}),\mathbb{Z}).$$

This isomorphism is just

$$h \circ P^{-1} : H_2(X;\mathbb{Z}) \overset{P^{-1}}{\cong} H^2(X;\mathbb{Z}) \overset{h}{\cong} \operatorname{Hom}(H_2(X;\mathbb{Z}),\mathbb{Z}).$$

Indeed, we can write $w = [X] \frown \alpha$ for a unique $\alpha = P^{-1}(w) \in H^2(X;\mathbb{Z})$. Then for $z := [X] \frown \beta \in H_2(X;\mathbb{Z})$,

$$\begin{aligned} Q_X(w,z) &= Q_X([X] \frown \alpha, [X] \frown \beta) = Q^X(\alpha,\beta) = [X] \frown (\alpha \smile \beta) \\ &= [X] \frown (\beta \smile \alpha) = ([X] \frown \beta) \frown \alpha \\ &= z \frown \alpha = \langle \alpha, z \rangle = \left[\left(h \circ P^{-1}\right)(w)\right](z)\,. \end{aligned}$$

Thus, the matrix of Q_X relative to a basis (i.e., a set of independent *generators*) of $H_2(X;\mathbb{Z})$ not only has integer entries, but also has determinant ± 1 (i.e., is *unimodular*). Indeed, if $\{e_1,...,e_n\}$ is a basis of $H_2(X;\mathbb{Z})$ and $\{\delta_1,...,\delta_n\}$ is the dual basis, then nondegeneracy implies that there are $\alpha_1,...,\alpha_n \in H_2(X,\mathbb{Z})$, such that (for $\alpha_i = \sum_{k=1}^n \alpha_{ik} e_k$, $\alpha_{ik} \in \mathbb{Z}$)

$$\delta_i = Q_X(\alpha_i,\cdot) \quad \text{or} \quad \delta_{ij} = \delta_i(e_j) = Q_X(\alpha_i,e_j) = \sum_{k=1}^n \alpha_{ik}\,[Q_X]_{kj}\,. \tag{19.5}$$

Thus, $[\alpha_{ik}]$ is the inverse matrix of $[Q_X]$ with integer entries, and $\det(Q_X)\det([\alpha_{ik}]) = 1 \Rightarrow \det(Q_X) = \pm 1$. In [**Fr**], Michael Freedman showed that any unimodular symmetric bilinear form over $\mathbb{Z}^n$ $(0 < n \in \mathbb{Z})$ is isomorphic to Q_X for some *topological*, simply-connected, compact, oriented 4-manifold X. Moreover, up to homeomorphism there are at most two such X having a given intersection form. In order to state this result more precisely, we introduce some terminology.

A unimodular symmetric bilinear form Q on a lattice, say $\mathbb{Z}^n$ (n is the *rank* of Q), is called *even* if $Q(\alpha,\alpha)$ is even for all $\alpha \in \mathbb{Z}^n$. Otherwise, Q is called *odd*. Note Q is even iff all of the diagonal elements of a matrix representing Q are even, since

$$Q(\alpha,\alpha) = \sum_{i=1}^n \alpha_i^2 Q(e_i,e_i) + 2\sum_{i<j}^n \alpha_i\alpha_j Q(e_i,e_j) \equiv \sum_{i=1}^n \alpha_i^2 Q(e_i,e_i) \bmod 2.$$

The *signature* $\tau(Q)$ of Q is the number of positive minus the number of negative entries when Q is diagonalized over $\mathbb{R}$. If $\tau(Q) = \pm\mathrm{rank}(Q)$, then Q is *definite* and otherwise Q is *indefinite* . The forms which indefinite are easy to describe (see [**MilH**]):

1. If Q is indefinite and *odd* of rank n, then, relative to some basis, the matrix $[Q]$ of Q is given by the standard forms

$$[Q] = diag(1, \overset{p}{\ldots}, 1, -1, \overset{q}{\ldots}, -1),\;\; p+q=n. \tag{19.6}$$

2. If Q is indefinite and *even*, then relative to some basis, $[Q]$ is block-diagonal, with

$$[Q] = H \oplus ... \oplus H \oplus \pm(E_8 \oplus ... \oplus E_8)\,, \text{ (at least one } H \text{ summand) where} \tag{19.7}$$

$$H := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad E_8 := \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix}. \tag{19.8}$$

REMARK 19.3. We show that (19.6) is the intersection form for the connected sum

$$\left(\#_p \mathbb{CP}^2\right) \# \left(\#_q \overline{\mathbb{CP}^2}\right) := \left(\mathbb{CP}^2 \# \overset{p}{\cdots} \# \mathbb{CP}^2\right) \# \left(\overline{\mathbb{CP}^2} \# \overset{q}{\cdots} \# \overline{\mathbb{CP}^2}\right). \tag{19.9}$$

Here $\overline{\mathbb{CP}^2}$ is $\mathbb{CP}^2$ but with the orientation reversed. Recall (see [**Gr**, p.119]) that $H_2\left(\mathbb{CP}^2\right) \cong \mathbb{Z}$ with generator $\alpha := \left[\left(\mathbb{CP}^1\right)_1\right]$, where $\left(\mathbb{CP}^1\right)_1 := \{[1, z_1, 0]\} \subset \mathbb{CP}^2$. To compute $Q(\alpha, \alpha)$, we note that $\left(\mathbb{CP}^1\right)_1$ meets the homologous $\left(\mathbb{CP}^1\right)_2 := \{[1, 0, z_2]\} \subset \mathbb{CP}^2$ (with $\left[\left(\mathbb{CP}^1\right)_2\right] = \alpha = \left[\left(\mathbb{CP}^1\right)_1\right]$) at the single point $[1,0,0]$ with intersection number $+1$, whereas in $\overline{\mathbb{CP}^2}$ the intersection number is -1.

REMARK 19.4. Over the reals, H (where "H" stands for hyperbolic) is equivalent to $diag(1,-1)$, but H is not equivalent to $\operatorname{diag}(1,-1)$ over $\mathbb{Z}$, since $diag(1,-1)$ is odd but H is even. We can see that H is the intersection form for $S^2 \times S^2$, as follows. Each of the generators $\left[S^2 \times \{y\}\right]$ or $\left[\{x\} \times S^2\right]$ in $H_2\left(S^2 \times S^2\right)$ has self-intersection number 0, since (for $y \neq y'$, $\left[S^2 \times \{y\}\right] = \left[S^2 \times \{y'\}\right]$ but $S^2 \times \{y\} \cap S^2 \times \{y'\} = \phi$. The intersection number of $S^2 \times \{y\}$ with $\{x\} \times S^2$ is 1; here $S^2 \times S^2$ inherits its orientation from the S^2 factors. Note that $S^2 \times S^2$ is not homotopic to $\mathbb{CP}(2) \# \overline{\mathbb{CP}(2)}$, since $H \ncong \operatorname{diag}(1,-1)$.

REMARK 19.5. Note that the number of E_8 summands is $\frac{1}{8}$ of the signature of the even, indefinite form Q in (19.7). The form E_8 gets its name from the fact that it is isomorphic to the Cartan matrix for $\mathfrak{e}_8$, one of the exceptional Lie algebras. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors of $\mathbb{R}^n$. In general, for any integer $k > 0$, there is a lattice $\mathcal{E}_{4k}$ in $\mathbb{R}^{4k}$ generated by the vectors $\mathbf{e}_i \pm \mathbf{e}_j$, $-\mathbf{e}_i \pm \mathbf{e}_j$ $(i, j \in \{1, \ldots, 4k\}, i \neq j)$, and $\frac{1}{2}(\pm \mathbf{e}_1 \pm \cdots \pm \mathbf{e}_{4k})$, where the numbers of "+ signs" and "− signs" in the last expression are equal mod 2. In the case $k = 2$, there are $4\binom{8}{2} + 2^7 = 240$ of these vectors and these can be identified with the roots of $\mathfrak{e}_8$ (note $\dim \mathfrak{e}_8 = 8 + 240 = 248$); see [**Hum**, p.65]. Moreover, there is a transitive group of isometries on this set (see [**MilH**, p.28]). By definition, the form E_{4k} is the restriction to $\mathcal{E}_{4k}$ of the usual dot product on $\mathbb{R}^{4k}$. To see that this definition is consistent with the form E_8 defined in (19.8), one may do the following exercise.

EXERCISE 19.6. Find generators $\mathbf{f}_1, \ldots, \mathbf{f}_8$ of the lattice $\mathcal{E}_8$, such that the matrix $[\mathbf{f}_i \cdot \mathbf{f}_j]$ is the one in (19.8). [Hint. One can choose $\mathbf{f}_1 = \mathbf{e}_1 - \mathbf{e}_2$, $\mathbf{f}_2 = \mathbf{e}_3 - \mathbf{e}_2$ and $\mathbf{f}_8 = \frac{1}{2}((\mathbf{e}_1 + \cdots + \mathbf{e}_6) - (\mathbf{e}_7 + \mathbf{e}_8))$. Find $\mathbf{f}_3, \ldots, \mathbf{f}_7$.]

If Q is *definite* and *even*, the rank of Q is a multiple of 8. When the rank $n = 8$, up to isomorphism there is only E_8. When $n = 16$, there are $E_8 \oplus E_8$ and E_{16}. When $n = 24$, we have 24, including $E_8 \oplus E_8 \oplus E_8$, $E_8 \oplus E_{16}$, and the so-called

Leech lattice. However, the possibilities increase very rapidly. When $n = 32$, there are over 10 million different classes. When $n = 40$, there are over 10^{51} different classes. There are also definite, non-diagonalizable *odd* Q; e.g, the direct sum of a definite, even form and $\mathrm{diag}(1, ..., 1)$. Moreover, there are *indecomposable*, odd, definite Q. These results are stated in [**MilH**], but the primary source is [**Ser**].

THEOREM 19.7. *(M. Freedman, [**Fr**]; see also [**FQ**]). Any unimodular quadratic form Q over $\mathbb{Z}^n$ (n arbitrary) is isomorphic to Q_X for some topological, simply-connected, compact, oriented 4-manifold X. Moreover, up to homeomorphism, X is unique if Q is even. If Q is odd there are two such X. For one of these $X \times S^1$ is smoothable, and for the other $X \times S^1$ is not smoothable (i.e., its Kirby-Siebenmann invariant is nonzero).*

COROLLARY 19.8. *(Poincaré's conjecture in dimension 4) If a topological manifold is X is homotopic to S^4 (e.g., X is simply-connected and has trivial even intersection form $Q_X = 0$), then X is homeomorphic to S^4.*

COROLLARY 19.9. *There are over* 10 *million topologically distinct, simply-connected topological 4-manifolds with even, definite intersection forms of rank* 32.

By the following result of Donaldson, none of the manifolds in Corollary 19.9 admits a differentiable structure.

THEOREM 19.10. *(S. K. Donaldson [**Don83**]). If X is a compact, simply-connected, smooth (i.e., C^∞) 4-manifold with a definite intersection form Q_X, then Q_X is a standard diagonalizable form (i.e., according to Theorem 19.7, topologically X must be a connected sum of $\mathbb{CP}^2$s or $\overline{\mathbb{CP}^2}$s).*

Donaldson also proved some results about the possible *indefinite, even* forms of smooth, compact, simply connected, 4-manifolds.

THEOREM 19.11. *([**Don90**]). If the form*

$$(\oplus_M H) \oplus (\pm (\oplus_N E_8)) \tag{19.10}$$

is realized by a smooth, compact, simply connected 4-manifold and $N > 0$, then we must have $M \geq 3$.

A $K3$ *surface* is a compact, simply-connected, complex surface X ($\dim_{\mathbb{C}} X = 2$) with trivial canonical bundle $\Lambda^{2,0}(T_{\mathbb{C}}M^*)$ (see p.584 and 601). An example is the complex variety ${z_1}^4 + {z_2}^4 + {z_3}^4 + {z_4}^4 = 0$ in $\mathbb{CP}^3$. It is a deep fact that all $K3$ surfaces are diffeomorphic. However, in the holomorphic category the family of $K3$ surfaces is 20-dimensional [**GH**, p. 593]. For a $K3$ surface, the intersection form is known to be

$$Q^{K3} = (\oplus_3 H) \oplus (- \oplus_2 E_8).$$

By Proposition 18.18 (p. 18.18) we know that a compact, orientable manifold X admits a spin structure if and only if $w_2(X) = 0$. There is also the following characterization in terms of the intersection form Q_X.

PROPOSITION 19.12. *Let X be a compact, simply-connected 4-manifold. Then X has a spin structure (i.e., $w_2(X) = 0$) $\Leftrightarrow Q^X$ is even.*

PROOF. Since X is simply-connected, $H^2(X;\mathbb{Z})$ is free and so reduction mod 2 $r_* : H^2(X;\mathbb{Z}) \to H^2(X;\mathbb{Z}_2)$ is onto. We first establish a special case of Wu's formula

$$Q_2^X(w_2(X), r_*(\alpha)) = Q_2^X(r_*(\alpha), r_*(\alpha)) \text{ for all } \alpha \in H^2(X;\mathbb{Z}), \tag{19.11}$$

where $Q_2^X : H^2(X;\mathbb{Z}_2) \times H^2(X;\mathbb{Z}_2) \to \mathbb{Z}_2$ is simply given by

$$Q_2^X(r_*(\alpha), r_*(\alpha)) := Q^X(\alpha,\alpha) \bmod 2 \in \mathbb{Z}_2.$$

By Proposition 19.1, we may assume that the Poincaré dual of α is $[\Sigma] \in H_2(X;\mathbb{Z})$ for some compact orientable surface Σ embedded in X. Using (18.16), p. 496,

$$\begin{aligned} Q_2^X(w_2(X), r_*(\alpha)) &= \langle w_2(TX|\Sigma_g), r_*([\Sigma_g])\rangle \\ &= \langle w_2(T\Sigma_g \oplus N\Sigma_g), r_*([\Sigma_g])\rangle \\ &= \langle w_2(T\Sigma_g), r_*([\Sigma_g])\rangle + \langle w_2(N\Sigma_g), r_*([\Sigma_g])\rangle . \end{aligned}$$

By Proposition 18.19p.498, this is

$$\begin{aligned} &= (\langle c_1(T\Sigma_g), [\Sigma_g]\rangle + \langle c_1(N\Sigma_g), [\Sigma_g]\rangle) \bmod 2 = \langle c_1(N\Sigma_g), [\Sigma_g]\rangle \bmod 2 \\ &= Q_X([\Sigma_g], [\Sigma_g]) \bmod 2 = Q^X(\alpha,\alpha) \bmod 2 = Q_2^X(r_*(\alpha), r_*(\alpha)) . \end{aligned}$$

By Remark 16.66 (p.430) $\langle c_1(N\Sigma_g), [\Sigma_g]\rangle$ is the self-intersection number $Q_X([\Sigma_g], [\Sigma_g])$ of Σ_g. Thus,

$$\begin{aligned} Q_2^X(w_2(X), r_*(\alpha)) &= \langle c_1(N\Sigma_g), [\Sigma_g]\rangle \bmod 2 = Q_X([\Sigma_g], [\Sigma_g]) \bmod 2 \\ &= Q^X(\alpha,\alpha) \bmod 2 = Q_2^X(r_*(\alpha), r_*(\alpha)) , \end{aligned}$$

and so we have (19.11). □

Recall that X admits a spin structure if $w_2(X) = 0$. Thus, Rochlin's Theorem (see Corollary 18.52, p. 562) may be restated as

THEOREM 19.13. *(Rochlin's Theorem) The signature $\tau(X)$ of a simply connected, smooth, compact 4-manifold X with Q_X even must be a multiple of 16. Thus, N in 19.10 must be even.*

COROLLARY 19.14. *The topological manifold with intersection form E_8 guaranteed to exist by Freedman's Theorem 19.7 does not admit a smooth structure.*

In view of Theorems 19.10, 19.11 and 19.13,

THEOREM 19.15. *Among compact, simply connected, smooth 4-manifolds X with Q_X even (i.e., those which are spin), $K3$ surfaces (including those with opposite orientation) have simplest possible Q_X with an E_8 summand, namely $(\oplus_3 H) \oplus (\pm \oplus_2 E_8)$.*

PROOF. By Theorem 19.10, if Q_X is even, then it must be indefinite, and hence it must be of the form $(\oplus_M H) \oplus (\pm \oplus_N (E_8))$. By Theorem 19.11, we have $M \geq 3$, and by Theorem 19.13, $N \geq 2$. The $K3$ surface and its "opposite" realize the lower bounds. □

Note that for $X = \#_p(K3) \# (\#_q S^2)$, we have

$$Q_X = (\oplus_{3p} H) \oplus (- \oplus_{2p} E_8) \oplus (\oplus_q H) = (\oplus_{3p+q} H) \oplus (- \oplus_{2p} E_8)$$

For this X and all known examples of smooth, compact, simply connected 4-manifolds with even, indefinite Q_X as in (19.10), one has $M \geq \frac{3}{2}N$. For such Q_X, we have $b_2(X) = 2M + 8N$ and $|\tau(X)| = 8N$. Thus,

$$b_2(X) \geq \tfrac{11}{8}|\tau(X)| \Leftrightarrow 8(2M+8N) \geq 11 \cdot 8N$$
$$\Leftrightarrow 16M \geq (88-64)N = 24N \Leftrightarrow M \geq \tfrac{3}{2}N.$$

The conjecture that $M \geq \frac{3}{2}N$ is the same as

Conjecture 1. *(The $\frac{11}{8}$-conjecture)A simply connected, smooth, compact 4-manifold, X with Q_X even satisfies*

$$b_2(X) \geq \tfrac{11}{8}|\tau(X)|,$$

where $b_2(X)$ is the second Betti number (the rank of Q_X) and $\tau(X)$ is the signature of Q_X.

While the $\frac{11}{8}$-conjecture may still be open, M. Furuta [**Fu01**] used Seiberg-Witten Theory to prove a strict bound of $\frac{10}{8}$, namely

$$b_2(X) = 2M + 8N \geq 10N + 2 = \tfrac{10}{8}|\tau(X)| + 2.$$

For X a smooth, simply-connected, compact orientable 4-manifold with "self-dual betti number" $b^{2+}(X) > 1$ and odd, Donaldson also found some invariants

$$q_{d_{\bar{k}}}(X) : \times_{d_{\bar{k}}/2} H_2(X,\mathbb{Z}) \to \mathbb{Z} \tag{19.12}$$

which are symmetric polynomials (Donaldson's polynomial invariants) of degree $d_{\bar{k}}$, $\bar{k} = 1, 2, 3....$ Here

$$d_{\bar{k}} = 8\bar{k} - 3(1 - b^1(X) + b^{2+}(X)), \tag{19.13}$$

which reduces to the *even* integer $8\bar{k} - 3(b^{2+}(X) + 1)$ under the above assumptions on X. The number $d_{\bar{k}}$ in (19.13) is the "virtual dimension" of the moduli space $M_{\bar{k}}$ of *anti-self-dual* connections for a principal $\mathrm{SU}(2)$-bundle $P \to X$ with $\bar{k} := c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)[X]$.

Remark 19.16. Donaldson (and almost everyone else now) have found it convenient to work with moduli spaces of anti-self-dual (ASD) connections, instead of the self-dual connections which we have dealt with in this book. However, one can translate between the two by changing the orientation of X. In particular, we defined $k := -c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)[X]$, and replacing the orientation $[X]$ by $-[X]$, Donaldson's k (which we have denoted by $\bar{k}$) is obtained. Moreover, $b^{2+}(X)$ becomes $b^{2-}(X)$ under $[X] \to -[X]$.

As we will now explain, we implicitly obtained the analogous formula for "virtual dimension" d_k of the moduli space M_k of *self-dual* connections (where $k := -c_2\left(P \times_{\mathrm{SU}(2)} \mathbb{C}^2\right)[X]$), namely

$$d_k = 8k - 3(1 - b^1(X) + b^{2-}(X)).$$

Indeed, this is the index of the operator $\mathcal{T} : \Omega^1(E) \to \Omega^0(E) \oplus \Omega^2_-(E)$, where $E := P \times_{\mathrm{SU}(2)} \mathfrak{su}(2)$, and

$$\mathcal{T}(\tau) := \left(\delta^\omega \tau, \tfrac{1}{2}(1 - *) D^\omega \tau\right), \text{ for } \tau \in \Omega^k(E).$$

We found (see (17.39) and (17.40) p.463, or (18.145) p.583) that

$$\mathrm{Index}(\mathcal{T}) = 2ch_2(E)[M] - \tfrac{1}{2}\dim E \cdot (\chi(M) - \mathrm{sig}(M)).$$

To show that this is d_k, recall from Proposition 17.7 (p.447) that for $E' := P \times_{\mathrm{SU}(2)} \mathbb{C}^2$ and $E := P \times_{\mathrm{SU}(2)} \mathfrak{su}(2)_{\mathbb{C}}$, we have $-4k = 4c_2(E') = c_2(E) = -ch_2(E)$, and so

$$\begin{aligned}\operatorname{Index}(\mathcal{T}) &= 8k - \tfrac{3}{2}\left(\chi(M) - \operatorname{sig}(M)\right) \\ &= 8k - \tfrac{3}{2}\left(2 - 2b^1(X) + b^{2+}(X) + b^{2-}(X) - \left(b^{2+}(X) - b^{2-}(X)\right)\right) \\ &= 8k - 3\left(1 - b^1(X) + b^{2-}(X)\right).\end{aligned}$$

Donaldson's polynomial invariants can be used to distinguish smooth 4-manifolds having the same intersection form. For example, we have

PROPOSITION 19.17. *([**DK**, p.27]) For any simply connected, complex surface S with $b^{2+}(S) > 3$, there is a smooth 4-manifold which is homotopy equivalent to S, but not diffeomorphic to S, nor to any complex surface.*

In dimension 4, there is a radical difference between the smooth category and the topological category. Unlike in the topological category, the function

$$[X]_{\text{smooth}} \mapsto [Q_X]$$

is far from being 1-1 or onto. Moreover, differential topology in dimension 4 is unlike that in other dimensions. For one thing, there are uncountably many exotic $\mathbb{R}^4$'s (homeomorphic to $\mathbb{R}^4$ but not diffeomorphic to $\mathbb{R}^4$ or to each other); (see [**Gom**]). There is no exotic $\mathbb{R}^n$ for $n \neq 4$. In dimensions greater than 4, there can be only a finite number of nondiffeomorphic compact manifolds which have the same homotopy type and Pontrjagin classes. However, by generalizing examples of Donaldson, Robert Friedman and John Morgan proved

THEOREM 19.18. *([**FM**]). There are infinitely many nondiffeomorphic smooth structures on $\mathbb{CP}^2 \# \left(\#_q \overline{\mathbb{CP}^2}\right)$ for any $q \geq 9$.*

Recall that two C^∞ n-manifolds M_0 and M_1 are h-**cobordant** if there is a C^∞ $(n+1)$-manifold N, with $\partial\, N = M_0 \cup M_1$, such that M_0 and M_1 are strong deformation retracts of N. The triple is called (N, M_0, M_1) an h-**cobordism**.

THEOREM 19.19. *(h-cobordism theorem [**Sm61**]; see also [**Mil65**])). Let (N, M_0, M_1) be an h-cobordism with N (and hence M_0 and M_1) simply-connected. If $n = dim(M_i) \geq 5$, then N is diffeomorphic to $M_0 \times [0,1]$, and consequently M_0 and M_1 are diffeomorphic.*

We can use Theorem 19.18 this to show that the analog of Smale's h-cobordism theorem fails dramatically in dimension 4 in the *smooth* category. In contrast, Friedman showed that the h-cobordism theorem holds in the topological category in dimension 4. First, there is the result of C.T.C. Wall:

THEOREM 19.20. *([**Wall**, 1964])Two simply-connected, smooth, compact 4-manifolds with isomorphic intersection forms are h-cobordant.*

Thus, there is an h-cobordism N between $\mathbb{CP}^2\#(\#_9\overline{\mathbb{CP}^2})$ and any of its infinitely many exotic versions, say X. However, since $\mathbb{CP}^2\#(\#_9\overline{\mathbb{CP}^2})$ and X are not diffeomorphic, N cannot be diffeomorphic to a product.

Let X be a smooth, compact, oriented 4-manifold (not necessarily simply-connected). For any $L \in H^2(X,\mathbb{Z})$ (which can be interpreted as a complex line bundle) we will later define a Seiberg-Witten invariant $SW(L) \in \mathbb{Z}$, provided $b^{2+}(X) \geq 2$. Thus,

$$SW : H^2(X,\mathbb{Z}) \to \mathbb{Z}. \tag{19.14}$$

Moreover, we will see that (as with the Donaldson polynomials) the Seiberg-Witten invariants are all 0, unless $b^1(X) + b^{2+}(X)$ is odd. Reportedly, many (if not all) of the previous instanton results of Donaldson and his coworkers (e.g., Kronheimer, Friedman, Morgan, Mrowka) can be proved more easily and pushed further using Seiberg-Witten invariants rather than using of Donaldson's invariants. Moreover, Taubes has proven a number of new results concerning symplectic manifolds, which we describe following a few definitions.

DEFINITION 19.21. A **symplectic form** or **symplectic structure** on a smooth manifold X is a closed 2-form ω on X which is nondegenerate in the sense that $\omega_\flat(V) := \omega(\cdot, V)$ defines an isomorhism $T_xX \cong (T_xX)^*$ at each $x \in X$. The pair (X,ω) is called a **symplectic manifold**. A diffeomorphism $f : X_1 \to X_2$ between two symplectic manifolds (X_1,ω_1) and (X_2,ω_2) such that $f^*\omega_2 = \omega_1$ is called a **symplectomorphism**. If such **symplectomorphism** exists, then (X_1,ω_1) and (X_2,ω_2) are called **symplectomorphic**.

For a symplectic manifold (X,ω), one can choose (uniquely up to homotopy) an almost complex structure J which is **compatible** with ω in the sense that $h_J(V,W) := \omega(V,JW)$ is a Riemannian metric. Indeed, let $Met(M) \subset T^{0,2}(M)$ be the convex set of metric tensors and let $\mathcal{J}(M,\omega)$ be the set of complex structures. Following [**McS**, p.61], we will define a canonical surjective function $R : Met(M) \to \mathcal{J}(M,\omega)$. Then not only is $\mathcal{J}(M,\omega)$ nonvoid, but also given $J_0, J_1 \in \mathcal{J}(M,\omega)$ with $J_0 = R(g_0)$ and $J_1 = R(g_1)$, we have a homotopy $J(t) := R((1-t)g_0 + tg_1)$ connecting J_0 with J_1 in $\mathcal{J}(M,\omega)$. For $g \in Met(M)$, we define $R(g) \in \mathcal{J}(M,\omega)$ as follows. Let $A \in \mathrm{Aut}(TX)$ be determined by $g(AV,W) = \omega(V,W)$. Then the adjoint A^* of A relative to g is defined by

$$g(V,A^*W) = g(AV,W), \text{ and } A^* = -A, \text{ (i.e., } A \text{ is } g\text{-skew-symmetric) since}$$
$$g(V,A^*W) = g(AV,W) = \omega(V,W) = -\omega(W,V) = -g(AW,V) = g(V,-AW).$$

Now $-A^2 = A^*A \in \mathrm{Aut}(TX)$ is g-symmetric and positive-definite, and hence it has a unique g-symmetric, positive-definite square root, say $S \in \mathrm{Aut}(TX)$, such that $S^2 = A^*A = -A^2$. Let $R(g) := J := S^{-1}A$. Then $J^2 = S^{-2}A^2 = -A^{-2}A^2 = -\mathrm{Id}$. Moreover,

$$\omega(V,JW) = g(AV,JW) = g\left(AV,S^{-1}AW\right)$$

is symmetric in V and W, and $\omega(V,JV) = g\left(AV,S^{-1}AV\right) > 0$ for $V \neq 0$, since A is invertible and S^{-1} is positive-definite. Thus, $k_J(V,W) := g\left(AV,S^{-1}AV\right) = \omega(V,JW)$ is a positive-definite metric, and so $R(g) \in \mathcal{J}(M,\omega)$.

For a compatible $J \in \mathcal{J}(M,\omega)$, the associated **cannonical class** $K_X := -c_1(T_JX) = c_1(\Lambda_J^{2,0}(X))$ then only depends only on ω. In the case $\dim X = 4$, ω is self-dual relative to the metric g and orientation $\omega \wedge \omega$.

THEOREM 19.22. *([**Tau95**]) If (X, ω) is a compact, symplectic 4-manifold with $b^{2+}(X) \geq 2$, then the Seiberg-Witten invariant $SW(K_X)$ is ± 1 and $(K_X \smile [\omega])[X] \geq 0$. If $SW(k) \neq 0$ for any other class $k \in H^2(X, \mathbb{Z})$, then*

$$|k \smile [\omega]| \leq K_X \smile [\omega], \tag{19.15}$$

with equality only if $k = \pm K_X$.

COROLLARY 19.23. *If (X, ω) is a compact, symplectic 4-manifold, then $b^1(X) + b^{2+}(X)$ is odd; otherwise, $SW(K_X) = 0$, contrary to Theorem 19.22.*

COROLLARY 19.24. *If $X = X_1 \# X_2$, where neither X_1 nor X_2 have negative-definite intersection forms (i.e., $b^{2+}(X_1)$ and $b^{2+}(X_2)$ are strictly positive), then X does not have a symplectic structure compatible with the given orientation. For example, when $n \geq 2$ and $m \geq 0$,*

$$\left(\#_n \mathbb{CP}^2\right) \# \left(\#_m \overline{\mathbb{CP}^2}\right) = \mathbb{CP}^2 \# \left(\#_{n-1} \mathbb{CP}^2\right) \# \left(\#_m \overline{\mathbb{CP}^2}\right) \tag{19.16}$$

admits no symplectic structure.

PROOF. It is known that such X have Seiberg-Witten invariant 0. $\square$

The following proves a conjecture of Mark Gotay.

COROLLARY 19.25. *While $\#_3 \mathbb{CP}^2$ has an almost complex structure, it has no symplectic form.*

Proof. In order for $c \in H^2(X, \mathbb{Z})$ to be the first Chern class of an almost complex structure on the compact, orientable 4-manifold X, it is necessary and sufficient that $c = w_2(X) \bmod 2$, and $Q^X(c, c) = 2\chi(X) + 3\tau(X)$. For $X = \#_3 \mathbb{CP}^2$, $H^2(X, \mathbb{Z}) \cong \oplus_3 \mathbb{Z}$, $Q^X \cong \operatorname{diag}(1, 1, 1)$, $w_2(X) \longleftrightarrow (1, 1, 1)$ and $2\chi(X) + 3\tau(X) = 2 \cdot 5 + 3 \cdot 3 = 19$. If $c \longleftrightarrow (k_1, k_2, k_3)$, then $c = w_2(X) \bmod 2 \Leftrightarrow k_1, k_2$ and k_3 are odd, and $Q^X(c, c) = 2\chi(X) + 3\sigma(X) \Leftrightarrow \sum k_i^2 = 19$. Thus, $(k_1, k_2, k_3) = (1, 3, 3)$ will do. Of course, $\#_3 \mathbb{CP}^2$ has no symplectic form by Corollary 19.24.

Let $K \in H^2(\mathbb{CP}^2, \mathbb{Z})$ denote the Chern class of the canonical line bundle $\Lambda^{2,0}(\mathbb{CP}^2)$ of $\mathbb{CP}^2$. We have $K = -c_1(T\mathbb{CP}^2) = -3[\hat{\xi}] = -3[\kappa/\pi]$, where $\hat{\xi}$ is the dual of the tautological bundle $\xi := \left\{(p, v) \in \mathbb{CP}^2 \times \mathbb{C}^3 : v \in p\right\}$ and κ is the Kähler form of $\mathbb{CP}^2$ with the standard Fubini-Studi metric. Note that

$$(K \smile [\kappa])\left[\mathbb{CP}^2\right] = (-3[\kappa/\pi] \smile [\kappa])\left[\mathbb{CP}^2\right] = -3/\pi < 0.$$

More generally, using Seiberg-Witten theory Clifford Taubes proved

THEOREM 19.26. *([**Tau95**]) For any symplectic structure $\left(\mathbb{CP}^2, \omega\right)$, let K be the Chern class of the canonical line bundle of an almost complex structure J compatible with ω (i.e., $\omega(X, Y) = h(JX, Y)$ for a Riemannian metric h on $\mathbb{CP}^2$). Then $(K \smile [\omega])\left[\mathbb{CP}^2\right] < 0$.*

The Donaldson invariants and (more simply) SW invariants can also be used to prove the Thom Conjecture in "considerable generality" and specifically the classical case of $\mathbb{CP}^2$ (see [**KM94**]), namely

THEOREM 19.27. *(Thom Conjecture). In a compact Kähler surface X, any complex curve C (i.e., a holomorphic immersion of a Riemann surface) has the smallest genus among all surfaces immersed in X which are homologous to C.*

For a purely geometrical application, we have the result

Theorem 19.28. *([**Wit94**]) No compact 4-manifold X with $b_1(X) > 0$ and with a nonzero SW-invariant can have a metric with positive scalar curvature.*

We also have the following result of C. LeBrun

Theorem 19.29. *([**LeB**]). Any compact, Einstein 4-manifold X with a nonzero SW-invariant satisfies $\chi(X) \geq 3\tau(X)$.*

Remark 19.30. A result of Hitchin ([**?**, 1974]) says that an Einstein 4-manifold X satisfies $\chi(X) \geq \frac{3}{2}|\tau(X)|$, with equality implying that X is flat or is covered by a $K3$ surface (with $\chi = 24$ and $\tau = -16$). Thus, Theorem 19.29 provides a stronger result when $\tau(X) > 0$. In the case of a complex surface, $\chi(X) \geq 3\tau(X)$ is the same as $c_2 \geq {c_1}^2 - 2c_2$ or the Miyaoka-Yau inequality $3c_2(X) \geq c_1(X)^2$.

2. Spin^c Structures and the Seiberg-Witten Equations

Recall from (18.159), p.598, that

$$\mathrm{Spin}^c(n) := \mathrm{U}(1)\,\mathrm{Spin}(n) \cong \frac{\mathrm{U}(1) \times \mathrm{Spin}(n)}{\{\pm(1,1)\}}.$$

Moreover, recall that there is a homomorphism

$$r^c : \mathrm{Spin}^c(n) \to \mathrm{U}(1) \times \mathrm{SO}(n) \text{ given by}$$

$$r^c(zg) := \left(z^2, c(g)\right), \tag{19.17}$$

with kernel $\{[1, \mathbf{I}], [1, -\mathbf{I}]\} \cong \mathbb{Z}_2$, where $c : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ is the 2-fold cover.

Definition 19.31. A Spin^c **structure** for an oriented Riemannian n-manifold X consists of a principal Spin$^c(n)$-bundle $\pi^c : P_{\mathrm{Spin}^c(n)} \to X$, a principal $\mathrm{U}(1)$-bundle $\pi_1 : P_{\mathrm{U}(1)} \to X$, and an r^c-equivariant bundle map

$$\pi_{r^c} : P_{\mathrm{Spin}^c(n)} \to P_{\mathrm{U}(1)} \times FX, \tag{19.18}$$

where $FX \to X$ is the principal SO(n)-bundle of oriented orthonormal frames. Here, $P_{\mathrm{U}(1)} \times FX$ is the *fibered* (as opposed to Cartesian) product consisting of pairs (p_1, p_2) with $\pi_1(p_1) = \pi_F(p_2)$. Naturally, $P_{\mathrm{U}(1)} \times FX$ is a principal $\mathrm{U}(1) \times \mathrm{SO}(n)$-bundle. The Chern class in $H^2(X;\mathbb{Z})$ of the line bundle associated to $P_{\mathrm{U}(1)} \to X$ is called the **canonical class** of the Spin$^c(n)$-structure.

By Proposition 18.18 (p.497) that an oriented, Riemannian manifold X has a spin structure if and only if $w_2(X) = 0 \in H^2(X, \mathbb{Z}_2)$, where $w_2(X)$ is the second Stieffel-Whitney class defined in Definition 18.16, p.496. The condition under which X admits a Spin$^c(n)$-structure is much weaker. Indeed, we will show that any oriented, Riemannian 4-manifold X has a Spin$^c(n)$-structure. To describe the condition, observe that the short exact sequence $0 \to \mathbb{Z} \overset{m}{\to} \mathbb{Z} \overset{r}{\to} \mathbb{Z}/2\mathbb{Z} = \mathbb{Z}_2 \to 0$ (where m denotes multiplication by 2) induces a long exact sequence

$$... \to H^i(X;\mathbb{Z}) \overset{m_*}{\to} H^i(X;\mathbb{Z}) \overset{r_*}{\to} H^i(X;\mathbb{Z}_2) \overset{\delta}{\to} H^{i+1}(X;\mathbb{Z}) \to ..., \tag{19.19}$$

where δ is the Bochstein homomorphism.

THEOREM 19.32. *An oriented Riemannian n-manifold X has a* Spin^c *structure if and only if the Stiefel-Whitney class $w_2(X) \in H^2(X;\mathbb{Z}_2)$ is the mod 2 reduction $r_*([\tilde{w}])$ (see (19.19)) of an integral class $[\tilde{w}] \in H^2(X;\mathbb{Z})$. Then $[\tilde{w}]$ is the canonical class of some* $\mathrm{Spin}^c(n)$*-structure for X.*

PROOF. Up to smooth equivalence, complex line bundles (or equivalently principal U(1)-bundles) over X may be regarded (via their transition functions) as elements of $H^1(X;\mathrm{U}(1))$ (Čech cohomology with coefficients in the sheaf of germs of U(1)-valued functions); see the beginning of Section .18.2 where the construction was carried out for the frame bundle. Associated with the exact sequence

$$0 \to \mathbb{Z} \to \mathbb{R} \to \mathrm{U}(1) \to 0, \tag{19.20}$$

we have the long exact Čech cohomology sequence

$$... \to H^1(X;\widetilde{\mathbb{R}}) \to H^1(X;\mathrm{U}(1)) \overset{c_1}{\to} H^2(X;\mathbb{Z}) \to H^2(X;\widetilde{\mathbb{R}}) \to ..., \tag{19.21}$$

The homomorphism $c_1 : H^1(X;\mathrm{U}(1)) \to H^2(X;\mathbb{Z})$ is defined in the same way as the homomorphism $w_2 : H^1(X;\mathrm{SO}(n)) \to H^2(X;\mathbb{Z}_2)$ was defined in Definition 18.16, p.496, where the exact sequence in that case was $0 \to \mathbb{Z}_2 \to \mathrm{Spin}(n) \to \mathrm{SO}(n) \to 0$. By one definition, the first Chern class of a line bundle in $H^1(X;\mathrm{U}(1))$ is just its image under c_1 in (19.21). In (19.21), $\widetilde{\mathbb{R}}$ is the sheaf of germs of C^∞ $\mathbb{R}$-valued functions on X, as opposed to the constant sheaf $\mathbb{R}$ whose Čech cohomology coincides with the usual de Rham cohomology $H^*(X;\mathbb{R})$. It is not hard to show (via partitions of unity) that $H^i(X;\widetilde{\mathbb{R}}) = 0$ for $i > 0$. Thus, $c_1 : H^1(X;\mathrm{U}(1)) \cong H^2(X;\mathbb{Z})$; i.e., up to isomorphism a principal U(1)-bundle are determined by its first Chern class. Associated with the sequence

$$0 \to \mathbb{Z}_2 \overset{i}{\to} \mathrm{Spin}(n) \overset{c}{\to} \mathrm{SO}(n) \to 0,$$

we have exact Čech cohomology sequence

$$... \to H^1(X;\mathbb{Z}_2) \overset{i_*}{\to} H^1(X;\mathrm{Spin}(n)) \overset{c_*}{\to} H^1(X;\mathrm{SO}(n)) \overset{w_2}{\to} H^2(X;\mathbb{Z}_2), \tag{19.22}$$

where for nonabelian groups G, $H^1(X;G)$ is a pointed set instead of a group. By Definition 18.16, p.496, w_2 assigns to each equivalence class of SO(n)-bundles, its second Steifel-Whitney class. Thus, for a principal SO(n)-bundle ξ, $w_2(\xi)$ is the obstruction to finding a spin stucture covering ξ, and $w_2([FX])$ (or $w_2(X)$) the is obstruction in the special case of the bundle FX of oriented orthonormal frames. For the sequence

$$0 \to \mathbb{Z}_2 \overset{i}{\to} \mathrm{Spin}^c(n) \overset{r^c}{\to} \mathrm{U}(1) \times \mathrm{SO}(n) \to 0,$$

there is an exact seqence

$$\begin{aligned} &... \to H^1(X;\mathbb{Z}_2) \overset{i_*}{\to} H^1(X;\mathrm{Spin}^c(n)) \overset{r^c_*}{\to} \\ &\to H^1(X;\mathrm{U}(1)) \oplus H^1(X;\mathrm{SO}(n)) \overset{\tilde{c}_1 + w_2}{\to} H^2(X;\mathbb{Z}_2), \end{aligned} \tag{19.23}$$

where $\tilde{c}_1$ is the composition $H^1(X;\mathrm{U}(1)) \overset{c_1}{\to} H^2(X;\mathbb{Z}) \overset{r_*}{\to} H^2(X;\mathbb{Z}_2)$. Suppose $r_*([\tilde{w}]) = w_2(X)$ for some $[\tilde{w}] \in H^2(X;\mathbb{Z})$. Then $c_1^{-1}([\tilde{w}]) \in H^1(X;\mathrm{U}(1))$ defines (up to equivalence) a principal U(1)-bundle $P_{\mathrm{U}(1)}$. Thus, the obstruction to obtaining a Spin^c structure over $P_{U(1)} \times FX$ with canonical class $[\tilde{w}]$ is

$$\begin{aligned}(\tilde{c}_1 + w_2)\left(c_1^{-1}([\tilde{w}]) \oplus [FX]\right) &= \tilde{c}_1\left(c_1^{-1}([\tilde{w}])\right) + w_2([FX]) \\ &= r_*([\tilde{w}]) + w_2(X) = w_2(X) + w_2(X) = 0,\end{aligned} \tag{19.24}$$

since $w_2(X)$ is in the $\mathbb{Z}_2$-module $H^2(X;\mathbb{Z}_2)$. Conversely, if there is a Spin^c structure over $P_{U(1)} \times FX$ with canonical class $[\tilde{w}] \in H^2(X;\mathbb{Z})$, then $\tilde{c}_1\left(c_1^{-1}([\tilde{w}])\right) + w_2([FX]) = 0$, in which case

$$w_2([FX]) = \tilde{c}_1\left(c_1^{-1}([\tilde{w}])\right) = \left(r_* \circ c_1 \circ c_1^{-1}\right)[\tilde{w}] = r_*([\tilde{w}]).$$

□

REMARK 19.33. By the exact sequence (19.19), any two integral classes $[\tilde{w}] \in r_*^{-1}(w_2(X))$ differ by an element of

$$\operatorname{Ker}\left(H^2(X;\mathbb{Z}) \xrightarrow{r_*} H^2(X;\mathbb{Z}_2)\right) \cong m_* H^2(X;\mathbb{Z}) = 2H^2(X;\mathbb{Z}).$$

Thus, $2H^2(X;\mathbb{Z})$ acts on the set $r_*^{-1}(w_2(X))$ of choices for $[\tilde{w}]$. Once a choice of the canonical class $[\tilde{w}]$ is made, any two choices of isomorphism classes $[P_{\mathrm{Spin}^c(n)}]$ differ by the image of $i_*\left(H^1(X;\mathbb{Z}_2)\right)$ in (19.23). Assuming that X has a Spin^c structure, the set of Spin^c structures on X is then parametrized by $2H^2(X;\mathbb{Z}) \oplus i_*\left(H^1(X;\mathbb{Z}_2)\right)$.

We thank Bob Little for providing the references [**HiHo**] and [**Mas**] for following result.

THEOREM 19.34. *The Stiefel-Whitney class $w_2(X) \in H^2(X;\mathbb{Z}_2)$ of a compact, orientable Riemannian 4-manifold X is in fact the mod 2 reduction of some integral class $[\tilde{w}] \in H^2(X;\mathbb{Z})$. Consequently, any compact, orientable Riemannian 4-manifold admits a Spin^c-structure.*

PROOF. We need to show that $w_2(X) \in r_*\left(H^2(X;\mathbb{Z})\right)$, for $r_* : H^2(X;\mathbb{Z}) \to H^2(X;\mathbb{Z}_2)$ as in (19.19). Let $T^2(X)$ be the torsion subgroup of $H^2(X;\mathbb{Z})$. Suppose that we can show that

$$\begin{aligned} r_*\left(H^2(X;\mathbb{Z})\right) &= r_*\left(T^2(X)\right)^{\perp}, \text{ where} \\ r_*\left(T^2(X)\right)^{\perp} &:= \left\{z \in H^2(X;\mathbb{Z}_2) : z \smile r_*(t) = 0,\ \forall t \in T^2(X)\right\}.\end{aligned} \tag{19.25}$$

Then it would suffice to show that $w_2(X) \smile r_*(t) = 0$ for all $t \in T^2(X)$. According to Wu's formula

$$w_2(X) \smile r_*(t) = r_*(t) \smile r_*(t) = r_*(t \smile t) = 0,$$

where the last equality follows since $t \smile t$ is in the torsion subgroup of $H^4(X;\mathbb{Z}) \cong \mathbb{Z}$, whence $t \smile t = 0$. Thus it remains to show (19.25). Now $r_*\left(H^2(X;\mathbb{Z})\right) \subseteq r_*\left(T^2(X)\right)^{\perp}$, since $r_*(h) \smile r_*(t) = r_*(h \smile t) = 0$ because $h \smile t$ is a torsion element of the torsionless group $H^4(X;\mathbb{Z}) \cong \mathbb{Z}$. Thus, it suffices to prove that $\dim r_*\left(H^2(X;\mathbb{Z})\right) = \dim r_*\left(T^2(X)\right)^{\perp}$ as vector spaces over $\mathbb{Z}_2$. Since

$$\smile : H^2(X;\mathbb{Z}_2) \times H^2(X;\mathbb{Z}_2) \to H^4(X;\mathbb{Z}_2) \cong \mathbb{Z}_2$$

is nondegenerate, $\dim r_*\left(T^2(X)\right)^{\perp} = \dim H^2(X;\mathbb{Z}_2) - \dim r_*\left(T^2(X)\right)$. Let $b_i = \operatorname{rank} H^i(X;\mathbb{Z})$ and let c_i be the number of cyclic summands of $H^i(X;\mathbb{Z})$ of order equal to a power of 2 in the primary decomposition. We have $\dim r_*\left(T^2(X)\right) = c_2$, and $\dim r_*\left(H^2(X;\mathbb{Z})\right) = b_2+c_2$. Moreover, $\delta\left(H^2(X;\mathbb{Z}_2)\right) = \operatorname{Ker}\left(H^3(X;\mathbb{Z}) \overset{m_*}{\to} H^3(X;\mathbb{Z})\right) \cong \oplus_{c_3}\mathbb{Z}_2$. Thus, from the exact sequence

$$\ldots \to H^2(X;\mathbb{Z}) \overset{m_*}{\to} H^2(X;\mathbb{Z}) \overset{r_*}{\to} H^2(X;\mathbb{Z}_2) \overset{\delta}{\to} H^3(X;\mathbb{Z}) \overset{m_*}{\to} H^3(X;\mathbb{Z}) \to \ldots,$$

we obtain

$$\dim H^2(X;\mathbb{Z}_2) = \dim r_*\left(H^2(X;\mathbb{Z})\right) + \dim \delta\left(H^2(X;\mathbb{Z}_2)\right) = b_2 + c_2 + c_3.$$

Hence,

$$\begin{aligned}\dim r_*\left(T^2(X)\right)^{\perp} &= \dim H^2(X;\mathbb{Z}_2) - \dim r_*\left(T^2(X)\right) \\ &= b_2 + c_2 + c_3 - c_2 = b_2 + c_3.\end{aligned}$$

Now, $\dim r_*\left(H^2(X;\mathbb{Z})\right) = b_2 + c_2$. Thus, it remains to show that $c_2 = c_3$. Using the Universal Coefficient Theorem ([**Gr**, p.194] or [**Sp**, Corollary 4, p.244]),

$$H^3(X;\mathbb{Z}) \cong F_3(X) \oplus T_2(X), \tag{19.26}$$

where $T_i(X)$ is the torsion subgroup of $H_i(X;\mathbb{Z})$ and $F_i(X) = H_i(X;\mathbb{Z})/T_i(X)$. By Poincaré duality (which we can use, since X is orientable),

$$H^2(X;\mathbb{Z}) \cong H_2(X;\mathbb{Z}) \cong F_2(X) \oplus T_2(X). \tag{19.27}$$

Thus, the torsion subgroup of $H^3(X;\mathbb{Z})$ is the same as that of $H^2(X;\mathbb{Z})$, and hence $c_3 = c_2$. □

In the case of compact, almost complex $2m$-manifolds X, $w_2(X)$ is the mod 2 reduction of the integral class $c_1(\Lambda^{m,0}(X)) = -c_1(TX)$, and hence X admits a Spin^c-structure with canonical class $c_1(\Lambda^{m,0}(X))$. Since it might be the case that $H^1(X,\mathbb{Z}_2) \neq 0$, this does not uniquely determine a Spin^c structure for X. However, we will now show that there is a standard Spin^c structure for X, using the existence of the lift $h : \mathrm{U}(m) \to \mathrm{Spin}^c(2m)$ of $\det \times \iota : \mathrm{U}(m) \to \mathrm{U}(1) \times \mathrm{SO}(2m)$.

Proposition 19.35. *Any Riemannian $2m$-manifold X with almost complex structure J, with a metric h compatible with J (i.e., $h(JV,JV) = h(V,V)$) admits a standard* Spin^c*-structure, with canonical class being the Chern class of the canonical line bundle $\Lambda^{m,0}(X)$.*

Proof. Let $P_{\mathrm{U}(m)}$ be the unitary frame bundle of X. The homomorphism $j : \mathrm{U}(m) \to \mathrm{Spin}^c(2m)$ of Proposition 18.70 (p.599) provides a principal $\mathrm{Spin}^c(2m)$-bundle $P_{\mathrm{Spin}^c} \to X$ and j-equivariant map

$$P_{\mathrm{U}(m)} \to P_{\mathrm{U}(m)} \times_j \mathrm{Spin}^c(2m) =: P_{\mathrm{Spin}^c}$$

(see Proposition 16.19, p.378). The homomorphism $\det \times \iota : \mathrm{U}(m) \to \mathrm{U}(1) \times \mathrm{SO}(2m)$ yields a principal bundle $P_{\mathrm{U}(1)\times\mathrm{SO}(2m)} \to X$ which can be identified with $P_{\mathrm{U}(1)} \times FX$, where FX is the oriented orthonormal frame of X and $P_{\mathrm{U}(1)}$ is the frame bundle for the canonical line bundle $\Lambda^{m,0}(X)$; for this one may use Remark 16.20 (p. 379) and note that a frame in $P_{\mathrm{U}(m)}$ determines a frame in $P_{\mathrm{U}(1)}$ and a (oriented) frame in FX. The homomorphism $r^c : \mathrm{Spin}^c(2m) \to \mathrm{U}(1) \times \mathrm{SO}(2m)$ (that satisfies $r^c \circ j = \det \times \iota$) then yields an r^c-equivariant map

$$P_{\mathrm{Spin}^c} \to P_{\mathrm{U}(1)\times\mathrm{SO}(2m)} \cong P_{\mathrm{U}(1)} \times FX,$$

which provides the desired standard Spin^c-structure. □

For a Spin^c structure $P_{\mathrm{Spin}^c(n)} \overset{\pi_{r^c}}{\to} P_{\mathrm{U}(1)} \times FX$ we have a Spin$^c(n)$-principal bundle

$$\pi^c : P_{\mathrm{Spin}^c(n)} \overset{\pi_{r^c}}{\to} P_{\mathrm{U}(1)} \times FX \to X$$

Moreover, there is a representation

$$\mathrm{Spin}^c(n) \overset{\rho_{\mathbb{C}}}{\to} \mathrm{U}\left(\Sigma_{2m}\right)$$

which is just the restriction of $\rho_{\mathbb{C}} : \mathbb{C}l(n) \to \mathrm{End}\left(\Sigma_{2m}\right)$ to $\mathrm{Spin}^c(n) \subset \mathbb{C}l(n)$. As in (18.13), p. 18.13, there are also the Spin$^c(n)$-invariant eigenspaces $\Sigma_{2m}^{\pm}$ of $\nu_{\mathbb{C}} := i^m e_1 \cdots e_{2m}$. Thus, we may form the associated bundles

$$\Sigma_c^{\pm}(X) := P_{\mathrm{Spin}^c(n)} \times_{\rho_{\mathbb{C}}} \Sigma_{2m}^{\pm} \to X \text{ and } \Sigma_c(X) := \Sigma_c^{+}(X) \oplus \Sigma_c^{-}(X)$$

Let ω be a connection on $P_{\mathrm{U}(1)}$ and let θ be the Levi-Civita connection on FX. Then $\omega \oplus \theta$ is a connection on $P_{\mathrm{U}(1)} \times FX$ and $(\omega \oplus \theta)^c := \pi_{r^c}^{*}(\omega \oplus \theta)$ may be regarded as a connection on $P_{\mathrm{Spin}^c(n)}$ under the isomorphism of Lie algebras $\mathfrak{spin}^c(n) \cong \mathfrak{u}(1) \oplus \mathfrak{so}(n)$ induced by the double-covering $r^c : \mathrm{Spin}^c(n) \to \mathrm{U}(1) \times \mathrm{SO}(n)$. Hence, there is a covariant differentiation operator for $(\omega \oplus \theta)^c$, say

$$\nabla^{(\omega\oplus\theta)^c} : C^{\infty}\left(\Sigma_{2m}^c(X)\right) \to C^{\infty}\left(\Lambda^1(X) \otimes \Sigma_{2m}^c(X)\right).$$

Moreover, there is a well-defined Clifford multiplication

$$\Lambda^1(X) \otimes \Sigma_c(X) \overset{cl}{\to} \Sigma_c(X) \text{ with } cl\left(\Lambda^1(X) \otimes \Sigma_c^{\pm}(X)\right) \subset \Sigma_c^{\mp}(X).$$

induced by the representation $\rho_{\mathbb{C}} : \mathbb{C}l(n) \to \mathrm{End}\left(\Sigma_{2m}\right)$. Thus, we have a Spin^c-**Dirac operator**

$$\mathcal{D}_c : C^{\infty}\left(\Sigma_c(X)\right) \overset{cl\circ\nabla^{(\omega\oplus\theta)^c}}{\to} C^{\infty}\left(\Sigma_c(X)\right).$$

For index computations, as well as for a better understanding of this operator, we will show that $\mathcal{D}_c$ is locally a twisted Dirac operator. The twist is by a locally defined square root of the **canonical line bundle**

$$L := P_{\mathrm{U}(1)} \times_{\mu} \mathbb{C} \to X,$$

where $\mu : \mathrm{U}(1) \to \mathrm{U}(\mathbb{C})$ is simply given by $\mu(z)\, w = zw$; note that $c_1(L)$ is (by definition) the canonical class of the Spin^c structure $P_{\mathrm{Spin}^c(n)} \overset{\pi_{r^c}}{\to} P_{\mathrm{U}(1)} \times FX$. Let B_X be a coordinate ball of X. Since $P_{\mathrm{U}(1)}$ and FX are trivial over B_X, there is a μ^2-equivariant double cover $P'_{\mathrm{U}(1)} \to P_{\mathrm{U}(1)}|B_X$ and a trivial spin structure $P_{\mathrm{Spin}(n)} \to FX|B_X$. Define a 2-fold cover

$$\chi : \mathrm{U}(1) \times \mathrm{Spin}(2m) \to \mathrm{U}(1)\,\mathrm{Spin}(2m) = \mathrm{Spin}^c(2m) \text{ by } \chi(z,\sigma) := z\sigma.$$

Since $P_{\mathrm{Spin}^c(n)}|B_X$ is also trivial, we have a χ-equivariant covering

$$P'_{\mathrm{U}(1)} \times P_{\mathrm{Spin}(n)} \overset{\pi_{\chi}}{\to} P_{\mathrm{Spin}^c(n)}|B_X$$

and a 4-fold $r^c \circ \chi$-equivariant covering

$$\pi_{r^c} \circ \pi_{\chi} : P'_{\mathrm{U}(1)} \times P_{\mathrm{Spin}(n)} \overset{\pi_{\chi}}{\to} P_{\mathrm{Spin}^c(n)}|B_X \overset{\pi_{r^c}}{\to} \left(P_{\mathrm{U}(1)} \times FX\right)|B_X$$

Let $L'(B_X) := P'_{\mathrm{U}(1)} \times_{\mu} \mathbb{C}$. By Proposition 16.22 (p. 380) $L'(B_X) \otimes L'(B_X) \cong L|B_X$ naturally, since $\phi : \mathbb{C} \otimes \mathbb{C} \to \mathbb{C}$ (given by $\phi(w_1 \otimes w_2) := w_1 w_2$) satisfies

$$\begin{aligned} &\phi\left(\left((\mu \otimes \mu)(z)\right)(w_1 \otimes w_2)\right) = \phi\left(z w_1 \otimes z w_2\right) = z^2 w_1 w_2 \\ &= \mu\left(\mu^2(z)\right)\left(\phi\left(w \otimes w'\right)\right). \end{aligned}$$

Also using Proposition 16.22 again, we have

$$\Phi : L'(B_X) \otimes \Sigma_{2m}(B_X) \cong \Sigma^c_{2m}(X)\,|B_X.$$

Indeed in this case, let $\phi : \mathbb{C} \otimes \Sigma_{2m} \to \Sigma_{2m}$ be given by $\phi(w \otimes \psi) = w\psi$, and note that for $(z, \sigma) \in \mathrm{U}(1) \times \mathrm{Spin}(2m)$, we have χ-equivariance in the sense

$$\begin{aligned}\phi((z,\sigma)\cdot(w\otimes\psi)) &= \phi((zw\otimes\sigma\cdot\psi)) = zw(\sigma\cdot\psi)\\ &= z\sigma\cdot(w\psi) = \chi(z,\sigma)\cdot w\psi = \chi(z,\sigma)\cdot\phi(w\otimes\psi).\end{aligned}$$

Because of the naturality of the isomorphism Φ, the twisted Dirac operator $\mathcal{D}^{L'(B_X)} \in \mathrm{End}\,(C^\infty(L'(B_X) \otimes \Sigma_{2m}(B_X)))$ coresponds to $\mathcal{D}_c \in \mathrm{End}\,(C^\infty(\Sigma^c_{2m}(X)\,|B_X))$, in the sense that

$$\mathcal{D}^{L'(B_X)}(\psi) = \Phi^{-1}(\mathcal{D}_c(\Phi(\psi))).$$

Proposition 19.36. *For a* Spin^c*-structure* $\pi_{r^c} : P_{\mathrm{Spin}^c(n)} \to P_{\mathrm{U}(1)} \times FX$*, let* $\Omega^\omega \in \Omega^2(X, i\mathbb{R})$ *be the curvature of the connection* ω *on* $P_{\mathrm{U}(1)}$*, and let* S *be the scalar curvature of* X*. For* $\psi \in C^\infty(\Sigma_c(X))$*, and an orthonormal frame* $E_1, \ldots, E_n$ *at* $x \in X$*, let* $\mathfrak{R}^\omega_x \in \mathbb{C}l\,(TM)_x$ *be defined by*

$$\mathfrak{R}^\omega_x = \tfrac{1}{2}\sum_{j,k} \Omega^\omega_{jk} E_j E_k, \text{ and let } (\mathfrak{R}^\omega\psi)(x) := \mathfrak{R}^\omega_x \cdot \psi(x).$$

Then, we have

$$\mathcal{D}_c^2\psi = -\Delta\psi + \tfrac{1}{2}\mathfrak{R}^\omega\psi + \tfrac{1}{4}S\psi. \tag{19.28}$$

Proof. Since $\mathcal{D}_c$ is locally a twisted Dirac operator, this (19.28) is a consequence of Proposition 18.25 (p. 506). Note that we have the factor $\frac{1}{2}$ in $\frac{1}{2}\mathfrak{R}^\omega\psi$ since the local Dirac operator is twisted by $L' = P'_{\mathrm{U}(1)} \times_\mu \mathbb{C}$, rather than $L = P_{\mathrm{U}(1)} \times_\mu \mathbb{C} \cong L' \otimes L'$, and the curvature $\Omega^{\omega'} \in \Omega^2(X, i\mathbb{R})$ of the "lift" $\omega' := \frac{1}{2}\pi_1^*(\omega)$ of ω to $P'_{\mathrm{U}(1)}$ is $\frac{1}{2}\Omega^\omega$. Note since $\pi_1 : P'_{\mathrm{U}(1)} \to P_{\mathrm{U}(1)}|B_X$ is μ^2-equivariant, $\pi_{1*}(A^*) = 2A^*$ for $A \in \mathfrak{u}(1)$ so that (as required of a connection)

$$\omega'(A^*) = \tfrac{1}{2}\pi_1^*(\omega)(A^*) = \tfrac{1}{2}\omega(\pi_{1*}(A^*)) = \tfrac{1}{2}\omega(2A^*) = A.$$

□

One way to motivate the Seiberg-Witten equations without delving into supersymmetric QCD ([**SW**]) is to try to generalize to 4-manifolds with Spin^c structures, the fact that spin manifolds with positive scalar curvature have no nonzero harmonic spinors (see Corollary 18.26, p. 506). Mimicing the computation in the proof of Corollary 18.26, we have for $\psi \in C^\infty(\Sigma_c(X))$

$$\begin{aligned}\left\|\mathcal{D}_c^2\psi\right\|^2 &= \left(\mathcal{D}_c^2\psi, \psi\right) = \left(-\Delta\psi + \tfrac{1}{2}\mathfrak{R}^\omega\psi + \tfrac{1}{4}S\psi, \psi\right)\\ &= \left\|\nabla^{(\omega\oplus\theta)^c}\psi\right\|^2 + \tfrac{1}{2}(\mathfrak{R}^\omega\psi, \psi) + \tfrac{1}{4}(S\psi, \psi).\end{aligned} \tag{19.29}$$

Note that

$$\begin{aligned}\langle\mathfrak{R}^\omega\psi, \psi\rangle &= \left\langle \tfrac{1}{2}\sum\nolimits_{j,k}\Omega^\omega_{jk}E_jE_k\cdot\psi(x), \psi(x)\right\rangle\\ &= \tfrac{1}{2}\sum\nolimits_{j,k}\Omega^\omega_{jk}\langle E_jE_k\cdot\psi(x), \psi(x)\rangle\end{aligned}$$

Let $\Omega^\omega = \Omega^+ + \Omega^-$ be the decomposition of Ω^ω into its self-dual and anti-self-dual parts, and let $\mathfrak{R}^\omega = \mathfrak{R}^+ + \mathfrak{R}^-$ be the corresponding decomposition of $\mathfrak{R}^\omega$. We claim that

$$\left(\mathfrak{R}^\omega \psi, \psi\right) = \left(\mathfrak{R}^+ \psi^+, \psi^+\right) + \left(\mathfrak{R}^- \psi^-, \psi^-\right),$$

where $\psi = \psi^+ + \psi^-$ is the decomposition of

$$\psi \in C^\infty\left(\Sigma_c\left(X\right)\right) = C^\infty\left(\Sigma_c^+\left(X\right)\right) \oplus C^\infty\left(\Sigma_c^-\left(X\right)\right).$$

Indeed, for $\nu_{\mathbb{C}} := i^2 e_1 e_2 e_3 e_4 = -e_1 e_2 e_3 e_4$, we have $\nu_{\mathbb{C}} \mathfrak{R}^\pm = \pm \mathfrak{R}^\pm$, since $\nu_{\mathbb{C}}^2 = \mathrm{Id}$ and

$$\begin{aligned}
\nu_{\mathbb{C}} e_2 e_3 &= -e_1 e_2 e_3 e_4 e_2 e_3 = -e_1 e_2 e_3 e_2 e_3 e_4 = e_1 e_4 \\
\nu_{\mathbb{C}} e_3 e_1 &= -e_1 e_2 e_3 e_4 e_3 e_1 = e_2 e_3 e_4 e_3 = e_2 e_4 \\
\nu_{\mathbb{C}} e_1 e_2 &= -e_1 e_2 e_3 e_4 e_1 e_2 = -e_1 e_2 e_1 e_2 e_3 e_4 = e_3 e_4.
\end{aligned}$$

Thus, $\nu_{\mathbb{C}} \mathfrak{R}^\pm \psi^+ = \pm \mathfrak{R}^\pm \psi^+ \Rightarrow \mathfrak{R}^\pm \psi^+ \in C^\infty\left(\Sigma_c^\pm\left(X\right)\right)$. However, we know that $\mathfrak{R}^\pm \psi^+ \in C^\infty\left(\Sigma_c^+\left(X\right)\right)$, since $\mathfrak{R}_x = \frac{1}{2}\sum_{j,k} \Omega_{jk}^\omega E_j E_k$ is in the even part of $\mathbb{C}l\left(TM\right)_x$. Thus, $\mathfrak{R}^- \psi^+ = 0$ and $\mathfrak{R}^+ \psi^+ \in C^\infty\left(\Sigma_c^+\left(X\right)\right)$. Similarly, $\mathfrak{R}^+ \psi^- = 0$ and $\mathfrak{R}^- \psi^- \in C^\infty\left(\Sigma_c^-\left(X\right)\right)$. Thus, (19.29) becomes

$$\left\|\mathcal{D}_c^2 \psi\right\|^2 = \left\|\nabla^{(\omega\oplus\theta)^c} \psi\right\|^2 + \tfrac{1}{2}\left(\mathfrak{R}^+ \psi^+, \psi^+\right) + \tfrac{1}{2}\left(\mathfrak{R}^- \psi^-, \psi^-\right) + \tfrac{1}{4}\left(S\psi, \psi\right).$$

In particular,

$$\left\|\mathcal{D}_c^2 \psi^+\right\|^2 = \left\|\nabla^{(\omega\oplus\theta)^c} \psi^+\right\|^2 + \tfrac{1}{2}\left(\mathfrak{R}^+ \psi^+, \psi^+\right) + \tfrac{1}{4}\left(S\psi^+, \psi^+\right) \tag{19.30}$$

Note that at $x \in X$,

$$\begin{aligned}
\left\langle \mathfrak{R}^\omega\left(x\right)\psi^+\left(x\right), \psi^+\left(x\right)\right\rangle &= \left\langle \tfrac{1}{2}\sum\nolimits_{j,k} \Omega_{jk}^{\omega+} E_j E_k \cdot \psi^+\left(x\right), \psi^+\left(x\right)\right\rangle \\
&= \tfrac{1}{2}\sum\nolimits_{j,k} \Omega_{jk}^{\omega+} \left\langle E_j E_k \cdot \psi^+\left(x\right), \psi^+\left(x\right)\right\rangle \\
&= \left\langle \Omega_x^{\omega+}, q\left(\psi^+\left(x\right)\right)\right\rangle, \qquad (19.31)
\end{aligned}$$

where $q\left(\psi^+\left(x\right)\right)$ is the self-dual part of

$$Q\left(\psi^+\left(x\right)\right) := \tfrac{1}{2}\sum\nolimits_{j,k} \left\langle E_j E_k \cdot \psi^+\left(x\right), \psi^+\left(x\right)\right\rangle \varphi^j \wedge \varphi^k \in \Lambda_x^2\left(X, i\mathbb{R}\right). \tag{19.32}$$

Note that $Q\left(\psi^+\left(x\right)\right)$ has values in $i\mathbb{R}$, since

$$\begin{aligned}
&\left\langle E_j E_k \cdot \psi^+\left(x\right), \psi^+\left(x\right)\right\rangle \varphi^j \wedge \varphi^k = \left\langle \psi^+\left(x\right), E_k E_j \psi^+\left(x\right)\right\rangle \varphi^j \wedge \varphi^k \\
&= -\left\langle \psi^+\left(x\right), E_k E_j \psi^+\left(x\right)\right\rangle \varphi^k \wedge \varphi^j = -\overline{\left\langle E_k E_j \psi^+\left(x\right), \psi^+\left(x\right)\right\rangle} \varphi^k \wedge \varphi^j.
\end{aligned}$$

Since q appears throughout S-W theory, it is fitting to make a formal definition.

Definition 19.37. The quadratic map $q : C^\infty\left(\Sigma_c^+\left(X\right)\right) \to \Omega^{2+}\left(X, i\mathbb{R}\right)$ is defined by

$$q\left(\psi^+\right)\left(x\right) = \tfrac{1}{2}\left(1 + *\right) Q\left(\psi^+\left(x\right)\right),$$

where $Q\left(\psi^+\left(x\right)\right)$ is given by (19.32). See also Remarks 19.38 and 19.40 below.

From (19.30) and (19.31), we obtain

$$\left\|\mathcal{D}_c^2\psi^+\right\|^2 = \left\|\nabla^{(\omega\oplus\theta)^c}\psi^+\right\|^2 + \tfrac{1}{2}\left(\Omega^{\omega+}, q\left(\psi^+\right)\right) + \tfrac{1}{4}\left(S\psi^+, \psi^+\right),$$

and perhaps the simplest way to guarantee that

$$\mathcal{D}_c\psi^+ = 0 \text{ and } 0 \neq S \geq 0 \Rightarrow \psi^+ = 0$$

is to assume that $\Omega^{\omega+} = q\left(\psi^+\right)$ so that $\frac{1}{2}\left\langle\Omega^{\omega+}, q\left(\psi^+\right)\right\rangle = \frac{1}{2}\left|q\left(\psi^+\right)\right|^2 \geq 0$. We write $\mathcal{D}_c$ as $\mathcal{D}_c^\omega$ and Ω^+ as $\Omega^{\omega+}$, to indicate their dependence on ω. Based on the above motivation, the (unperturbed) **S-W equations** (S-W for Seiberg-Witten) for the pair (ω, ψ^+) are simply

$$\mathcal{D}_c^\omega\psi^+ = 0 \text{ and } \Omega^{\omega+} = q\left(\psi^+\right).$$

An immediate consequence of our derivation is that when $0 \neq S \geq 0$, (ω, ψ^+) is a solution of the S-W equations if and only if $\psi^+ = 0$ and $\Omega^{\omega+} = 0$ (i.e., ω has anti-self-dual curvature). Since the space of solutions of the S-W equations is not always nice, it is generally necessary to study the **perturbed S-W equations**

$$\mathcal{D}_c^\omega\psi^+ = 0 \text{ and } \Omega^{\omega+} = q\left(\psi^+\right) + \eta. \tag{19.33}$$

where $\eta \in \Omega^{2+}(X, i\mathbb{R})$ is a given self-dual 2-form which supplies the perturbation.

REMARK 19.38. The quadratic map $q : C^\infty\left(\Sigma_c^+(X)\right) \to \Omega^{2+}(X, i\mathbb{R})$ is induced by a pointwise map $q_x : \Sigma_c^+(X)_x \to \Lambda_x^{2+}(X, i\mathbb{R})$. This q_x is derived from a purely algebraic quadratic map $q_0 : \Sigma_2^+ \to i\Lambda^{2+}\left(\mathbb{R}^4\right)$, given by

$$\begin{aligned}
q_0(\psi) &= \left(\tfrac{1}{2}\sum\nolimits_{j,k}\left\langle e_je_k\psi^+, \psi^+\right\rangle e_j \wedge e_k\right)^+ \\
&= \left\langle e_2e_3\psi^+, \psi^+\right\rangle e_2 \wedge e_3 + \left\langle e_1e_4\psi^+, \psi^+\right\rangle e_1 \wedge e_4 + \cdots \\
&= \left\langle \tfrac{1}{2}\left(e_2e_3 + e_1e_4\right) + \tfrac{1}{2}\left(e_2e_3 - e_1e_4\right)\psi^+, \psi^+\right\rangle e_2 \wedge e_3 \\
&+ \left\langle \tfrac{1}{2}\left(e_2e_3 + e_1e_4\right) - \tfrac{1}{2}\left(e_2e_3 - e_1e_4\right)\psi^+, \psi^+\right\rangle e_1 \wedge e_4 + \cdots \\
&= \left\langle \tfrac{1}{2}\left(e_2e_3 + e_1e_4\right)\psi^+, \psi^+\right\rangle \left(e_2 \wedge e_3 + e_1 \wedge e_4\right) + \cdots \\
&= \tfrac{1}{2}\left\langle\left(e_2e_3 + e_1e_4\right)\cdot\psi^+, \psi^+\right\rangle\left(e_2 \wedge e_3 + e_1 \wedge e_4\right) + \\
&\tfrac{1}{2}\left\langle\left(e_3e_1 + e_2e_4\right)\cdot\psi^+, \psi^+\right\rangle\left(e_3 \wedge e_1 + e_2 \wedge e_4\right) + \\
&\tfrac{1}{2}\left\langle\left(e_1e_2 + e_3e_4\right)\cdot\psi^+, \psi^+\right\rangle\left(e_1 \wedge e_2 + e_3 \wedge e_4\right).
\end{aligned}$$

EXERCISE 19.39. We can make $q_0(\psi)$ more explicit by identifying Σ_2^+ with $\mathbb{C}^2$ and (in view of Example 18.8, p. 486) by using the fact that Clifford multiplication by $\frac{1}{2}\left(e_2e_3 + e_1e_4\right)$ corresponds to quaternionic multiplication by $\mathbf{i}$ on $\mathbb{H}$, or by $i\sigma_1$ on $\mathbb{C}^2$, where σ_1 is the first Pauli matrix (see 15.70, p. 352), etc.. Verify that
(a)

$$\begin{aligned}
q_0(\psi) = i\left(\psi_2\overline{\psi}_1 + \psi_1\overline{\psi}_2\right)&\left(e_2 \wedge e_3 + e_1 \wedge e_4\right) + \\
\left(\psi_2\overline{\psi}_1 - \psi_1\overline{\psi}_2\right)&\left(e_3 \wedge e_1 + e_2 \wedge e_4\right) + \\
\left(\psi_1\overline{\psi}_1 - \psi_2\overline{\psi}_2\right)&\left(e_1 \wedge e_2 + e_3 \wedge e_4\right) \in i\Lambda^{2+}\left(\mathbb{R}^4\right).
\end{aligned}$$

(b)

$$\tfrac{1}{2}\left|q_0(\psi)\right|^2 = |\psi|^4.$$

REMARK 19.40. Note that $q_0(\psi) = \widetilde{q}_0(\psi,\psi)$ where the real bilinear form $\widetilde{q}_0$ is given by

$$\begin{aligned}\widetilde{q}_0(\psi,\xi) = i\left(\psi_2\overline{\xi}_1 + \psi_1\overline{\xi}_2\right)(e_2\wedge e_3 + e_1\wedge e_4) + \\ \left(\psi_2\overline{\xi}_1 - \psi_1\overline{\xi}_2\right)(e_3\wedge e_1 + e_2\wedge e_4) + \\ i\left(\psi_1\overline{\xi}_1 - \psi_2\overline{\xi}_2\right)(e_1\wedge e_2 + e_3\wedge e_4) \in \mathbb{C}\otimes\Lambda^{2+}\left(\mathbb{R}^4\right),\end{aligned}$$

or equivalently

$$\begin{aligned}\widetilde{q}_0(\psi,\xi) &= \tfrac{1}{2}\left(q_0(\psi+\xi) - q_0(\psi) - q_0(\xi)\right) \\ &\quad + \tfrac{i}{2}\left(q_0(\psi+i\xi) - q_0(\psi) - q_0(\xi)\right).\end{aligned}$$

Also observe that $\widetilde{q}_0(\cdot,\cdot)$ is linear in the first slot, conjugate linear in the second slot, and $\widetilde{q}_0(\xi,\psi) = -\overline{\widetilde{q}_0(\psi,\xi)}$ so that $\widetilde{q}_0(\psi,\psi) \in i\Lambda^{2+}\left(\mathbb{R}^4\right)$. Similar statements hold for the form

$$\widetilde{q} : C^\infty\left(\Sigma_c^+(X)\right)\times C^\infty\left(\Sigma_c^+(X)\right) \to \Omega^{2+}(X,\mathbb{C}) \tag{19.34}$$

associated with $q : C^\infty(\Sigma_c^+(X)) \to \Omega^{2+}(X, i\mathbb{R})$.

3. The Manifold of Moduli

A certain subgroup $\mathrm{GA}_1\left(P_{\mathrm{Spin}^c(n)}\right)$ of the group of gauge transformations $\mathrm{GA}\left(P_{\mathrm{Spin}^c(n)}\right)$ acts on the space of solutions (ω,ψ^+) of the perturbed S-W equations. In this context, the quotient space is known as the *moduli space*. We will show that for a generic choice of the perturbation η, the moduli space is a compact manifold. The Seiberg-Witten invariant is obtained by integrating a certain form over this manifold. Much of the effort involved is directed toward showing that this Seiberg-Witten invariant is well-defined, independent of a suitable choice of perturbation η and Riemannian metric for X, so that it only depends on the differentiable structure on X and the choice of Spin^c structure.

The subgroup $\mathrm{GA}_1\left(P_{\mathrm{Spin}^c(n)}\right)$ of the group $\mathrm{GA}\left(P_{\mathrm{Spin}^c(n)}\right)$ of gauge transformations of $P_{\mathrm{Spin}^c(n)} \to X$ is defined as follows. For $s \in C^\infty(X, \mathrm{U}(1))$, note that

$$[s(x), 1] = \{\pm(s(x), 1)\} \in \frac{\mathrm{U}(1)\times\mathrm{Spin}(n)}{\{\pm(1,1)\}} = \mathrm{Spin}^c(n)$$

is in the center of $\mathrm{Spin}^c(n)$. Let $F_s \in C^\infty\left(P_{\mathrm{Spin}^c(n)}, P_{\mathrm{Spin}^c(n)}\right)$ be given by

$$F_s(p) := p\cdot[s(\pi^c(p)), 1],$$

where $\pi^c : P_{\mathrm{Spin}^c(n)} \to X$, is in $\mathrm{GA}\left(P_{\mathrm{Spin}^c(n)}\right)$. Then $F_s \in \mathrm{GA}\left(P_{\mathrm{Spin}^c(n)}\right)$, since

$$F_s(pg) = pg\cdot[s(\pi^c(p)), 1] = p\cdot[s(\pi^c(p)), 1]\,g = F_s(p)g, \tag{19.35}$$

for any $g \in \mathrm{Spin}^c(n)$. Thus, via $s \mapsto F_s$, we may regard $C^\infty(X, \mathrm{U}(1))$ as a subgroup, say $\mathrm{GA}_1\left(P_{\mathrm{Spin}^c(n)}\right)$, of $\mathrm{GA}\left(P_{\mathrm{Spin}^c(n)}\right)$. Note that for $\pi_{r^c} : P_{\mathrm{Spin}^c(n)} \to P_{\mathrm{U}(1)}\times FX$, $x = \pi^c(p)$ with $p \in P_{\mathrm{Spin}^c(n)}$, we have

$$\pi_{r^c}(F_s(p)) = \pi_{r^c}(p\cdot[s(x), 1]) = \pi_{r^c}(p)\cdot r^c([s(x), 1]) = \pi_{r^c}(p)\cdot\left(s(x)^2, \mathrm{Id}\right).$$

Thus, $F_s \in \mathrm{GA}_1\left(P_{\mathrm{Spin}^c(n)}\right)$ induces the gauge transformation $p_1 \mapsto p_1 s(x)^2$ on $P_{\mathrm{U}(1)}$. Hence, using Proposition 16.28 (p. 382) with $f_s \in C\left(P_{\mathrm{U}(1)}, \mathrm{U}(1)\right)$ given by

$$f_s(p_1) = p_1 s(\pi_1(p_1))^2 = p_1\left((\pi_1^* s)(p_1)\right)^2 = p_1\left(\pi_1^* s^2\right)(p_1)$$

(where $\pi_1 : P_{\mathrm{U}(1)} \to X$), we have that $C^\infty(X, \mathrm{U}(1))$ acts on the set $\mathcal{C}\left(P_{\mathrm{U}(1)}\right)$ of connections ω on $P_{\mathrm{U}(1)}$ via

$$\begin{aligned} s \cdot \omega &= f_s f_{s*}^{-1} + f_s \omega f_s^{-1} = f_s f_{s*}^{-1} + \omega = \left(\pi_1^* s^2\right) d\left(\pi_1^* s^{-2}\right) + \omega \\ &= \pi_1^*\left(s^2 ds^{-2}\right) + \omega = \pi_1^*\left(-s^2 2 s^{-3} ds\right) + \omega = \omega + \pi_1^*\left(-2s^{-1} ds\right). \end{aligned}$$

Moreover, Proposition 16.28 implies that $C^\infty(X, \mathrm{U}(1))$ ultimately acts on $C^\infty\left(X, \Sigma_c^+(X)\right)$ via the simple rule

$$(s \cdot \psi)(x) = s(x)\,\psi(x).$$

We show that if (ω, ψ) is solves the S-W equations, then

$$s \cdot (\omega, \psi) := (s \cdot \omega, s \cdot \psi) = \left(\omega + \pi_1^*\left(-2s^{-1} ds\right), s\psi\right). \tag{19.36}$$

(for all $s \in C^\infty(X, \mathrm{U}(1))$) is also a solution. By Corollary 16.29 (p. 383) and the fact the group $\mathrm{U}(1)$ of $P_{\mathrm{U}(1)}$ is abelian, $\Omega^{s\cdot\omega} = \pi_1^*\left(s^2\right)\Omega^\omega \pi_1^*\left(s^{-2}\right) = \Omega^\omega$. Alternatively,

$$\begin{aligned} \Omega^{s\cdot\omega} &= d(s\cdot\omega) = d\left(\omega + \pi_1^*\left(-2s^{-1}ds\right)\right) = d\omega - 2\pi_1^*\left(d\left(s^{-1}ds\right)\right) \\ &= d\omega - 2\pi_1^*\left(-s^{-2} ds \wedge ds + s^{-1} d^2 s\right) = \Omega^\omega, \end{aligned}$$

since $ds \wedge ds = 0$ and $d^2 s = 0$. Also, since the inner product on $\Sigma_c^+(X)$ is Hermitian, we have

$$q(s\cdot\psi) = q(s\psi) = s\bar{s}q(\psi) = |s|^2 q(\psi) = q(\psi).$$

Hence, as scalar multiplication does not affect duality, we have

$$\Omega^{\omega+} = q(\psi) + \eta \Rightarrow \Omega^{(s\cdot\omega)+} = \Omega^{\omega+} = q(\psi) + \eta = q(s\cdot\psi) + \eta. \tag{19.37}$$

Using Proposition 16.26 (p. 381) and the linearity of Clifford mutiplication, we also have

$$\mathcal{D}_c^{s\cdot\omega}(s\cdot\psi) = s\cdot\mathcal{D}_c^\omega\psi, \text{ so that } \mathcal{D}_c^\omega\psi = 0 \Rightarrow \mathcal{D}_c^{s\cdot\omega}(s\cdot\psi) = 0. \tag{19.38}$$

We denote the orbit of $(\omega,\psi) \in \mathcal{C}\left(P_{U(1)}\right) \times C^\infty\left(\Sigma_c^+(X)\right)$ under $C^\infty(X, \mathrm{U}(1))$ by

$$[\omega, \psi] := \left\{ s\cdot(\omega,\psi) : s \in C^\infty(X, \mathrm{U}(1)) \right\}$$

Definition 19.41. For a fixed self-dual $\eta \in \Omega^{2+}(X, i\mathbb{R})$, the set

$$M_\eta := \left\{ [\omega,\psi] : \mathcal{D}_c^\omega \psi^+ = 0 \text{ and } \Omega^{\omega+} = q\left(\psi^+\right) + \eta \right\}$$

of orbits of solutions (ω,ψ) of the perturbed S-W equations is known as the **moduli space for the** Spin^c **structure** $\left(P_{\mathrm{Spin}^c(n)}, P_{\mathrm{U}(1)}\right)$.

M_η is not always a manifold in a natural way. However, we will show that there are generic perturbations η such that M_η (if nonempty) is naturally a finite-dimensional manifold at points $[(\omega,\psi)]$ where $\psi \neq 0$.

We compute the dimension of M_η formally as follows. Any two connections on $P_{\mathrm{U}(1)}$ differ by a one-form which is the pull-back of a one-form on the base X. Thus, $\mathcal{C}\left(P_{\mathrm{U}(1)}\right)$ is an affine space associated with associated vector space

$$\overline{\Omega}^1\left(P_{U(1)}, i\mathbb{R}\right) \overset{\pi_1^*}{\cong} \Omega^1(X, i\mathbb{R}),$$

which then formally serves as the tangent space $T_\omega \mathcal{C}\left(P_{U(1)}\right)$. The formal tangent space of the orbit of a point $(\omega, \psi) \in \mathcal{C}\left(P_{U(1)}\right) \times C^\infty\left(\Sigma_c^+(X)\right)$ under the action of $C^\infty(X, \mathrm{U}(1))$ is the subspace of

$$T_{(\omega,\psi)}\left(\mathcal{C}\left(P_{U(1)}\right) \times C^\infty\left(\Sigma_c^+(X)\right)\right) := \Omega^1\left(P_{U(1)}, i\mathbb{R}\right) \oplus C^\infty\left(\Sigma_c^+(X)\right)$$

consisting of elements of the form (where $\theta \in C^\infty(X, \mathbb{R})$)

$$\begin{aligned}\frac{d}{dt}\left(e^{it\theta} \cdot (\omega, \psi)\right)\Big|_{t=0} &= \frac{d}{dt}\left(\omega + \pi_1^*\left(-2e^{-it\theta} d\left(e^{it\theta}\right)\right), e^{it\theta}\psi\right)\Big|_{t=0} \\ &= (-2id\theta, i\theta\psi) \in \Omega^1(X, i\mathbb{R}) \times C^\infty\left(\Sigma_c^+(X)\right),\end{aligned}$$

where we have (and will continue) to identify $\overline{\Omega}^1\left(P_{U(1)}, i\mathbb{R}\right)$ and $\Omega^1(X, i\mathbb{R})$. Now, $\Omega^1\left(P_{U(1)}, i\mathbb{R}\right) \oplus C^\infty\left(\Sigma_c^+(X)\right)$ has an inner product given by

$$\langle (i\alpha, \xi), (i\alpha', \xi') \rangle = \int_X h(\alpha, \alpha') + \mathrm{Re}\left(\langle \xi, \xi' \rangle\right) \nu_h,$$

where $h(\alpha, \alpha')$ is the usual inner product on covectors and ν_h is the volume form, each induced by the Riemannian metric h on X. A vector $(i\alpha, \xi)$ is normal to the orbit if we have (where $\delta = -*d*$ is the formal L^2-adjoint of d)

$$\begin{aligned}0 &= \langle (i\alpha, \xi), (-2id\theta, i\theta\psi) \rangle = \langle \alpha, -2d\theta \rangle + \tfrac{1}{2}\left(\langle \xi, i\theta\psi \rangle + \langle i\theta\psi, \xi \rangle\right) \\ &= \left\langle -2(\delta\alpha) - \tfrac{i}{2}\left(\langle \xi, \psi \rangle - \langle \psi, \xi \rangle\right), \theta \right\rangle \text{ for all } \theta \in C^\infty(X, \mathbb{R}), \text{ or}\end{aligned}$$

$$i\delta\alpha + \tfrac{1}{4}\left(\langle \psi, \xi \rangle - \langle \xi, \psi \rangle\right) = 0. \tag{19.39}$$

Consider the so-called Seiberg-Witten function

$$\Phi : \mathcal{C}\left(P_{\mathrm{U}(1)}\right) \times C^\infty\left(\Sigma_c^+(X)\right) \to \Omega^{2+}(X, i\mathbb{R}) \oplus C^\infty\left(\Sigma_c^+(X)\right) \tag{19.40}$$

given by

$$\Phi(\omega, \psi) = (\Omega^{\omega+} - q(\psi) - \eta, \mathcal{D}_c^\omega \psi). \tag{19.41}$$

The set of solutions of the S-W equations is $\Phi^{-1}(0,0)$. We formally compute the differential

$$\begin{aligned}\Phi_{*(\omega,\psi)}(i\alpha, \xi) &= \tfrac{d}{dt}(\Omega^{(\omega+ti\alpha)+} - q(\psi + t\xi), \mathcal{D}_c^{\omega+ti\alpha}(\psi + t\xi))\Big|_{t=0} \\ &= (i(d\alpha)^+ - \widetilde{q}(\xi, \psi) - \widetilde{q}(\psi, \xi), \mathcal{D}_c^\omega \xi + \tfrac{1}{2} i\alpha \cdot \psi),\end{aligned}$$

where $\alpha \cdot \psi$ is shorthand for $cl(\alpha^\# \otimes \psi)$ (i.e., Clifford multiplication of ψ by the vector field $\alpha^\#$ dual to the one-form α).

Since the tangent space of the moduli space M at $[\omega, \psi]$ can be formally identified with the intersection of $\operatorname{Ker} \Phi_{*(\omega,\psi)}$ with the normal space of the orbit of the action of $C^\infty(X, \mathrm{U}(1))$ through (ω, ψ), we see that formally the tangent space of M at $[(\omega, \psi)]$ is the kernel of the real operator

$$B : \Omega^1(X, i\mathbb{R}) \oplus C^\infty\left(\Sigma_c^+(X)\right) \to C^\infty(X, \mathbb{R}) \oplus \Omega^{2+}(X, i\mathbb{R}) \oplus C^\infty\left(\Sigma_c^-(X)\right) \tag{19.42}$$

given by

$$B(i\alpha, \xi) := \left(i\delta\alpha + \tfrac{1}{4}\left(\langle \psi, \xi \rangle - \langle \xi, \psi \rangle\right), \Phi_{*(\omega,\psi)}(i\alpha, \xi)\right).$$

We will now compute the Index B, which provides a lower bound on $\dim(\operatorname{Ker} B)$. In general, $\dim(\operatorname{Coker} B) \neq 0$, but later we will show that for a generic choice of self-dual 2-form η, $\operatorname{Coker}(B) = 0$ at any point in $\Phi_\eta^{-1}(0,0)$ for which $\psi \neq 0$. Hence the index of B will turn out to be the dimension of a (perturbed) moduli space.

One consequence of the Atiyah-Singer Index Theorem is that the index of an elliptic operator is determined by its highest order part. The top order part of differential operator B is the first-order differential operator

$$B_1(i\alpha,\xi) := \left(\delta\left(i\alpha\right), d\left(i\alpha\right)^{+}, \mathcal{D}_c^{\omega}\xi\right) = \left(\delta\oplus d^{+}\oplus\mathcal{D}_c^{\omega}\right)(i\alpha,\xi). \tag{19.43}$$

Each of the operators $\delta\oplus d^{+} : \Omega^1\left(X, i\mathbb{R}\right)\to\Omega^0\left(X,i\mathbb{R}\right)\oplus\Omega^{2+}\left(X,i\mathbb{R}\right)$ and $\mathcal{D}_c^{\omega} : C^{\infty}\left(\Sigma_c^{+}\left(X\right)\right)\to C^{\infty}\left(\Sigma_c^{-}\left(X\right)\right)$ is elliptic. Indeed, $\delta\oplus d^{+}$ is elliptic, since we have previously shown that $\delta\oplus d^{-}$ is elliptic (see the proof of Theorem 17.15, p. 462, where $\mathcal{T}_0 := \delta\oplus d^{-}$) and $\delta\oplus d^{+}$ results from $\delta\oplus d^{-}$ by changing the orientation of X, which does not affect ellipticity. The formal adjoint of $\left(\delta\oplus d^{+}\right)^{*}$ is $d\oplus\delta|\Omega^{2+}\left(X,i\mathbb{R}\right)$, and

$$\begin{aligned}\operatorname{Index}\left(\delta\oplus d^{+}\right) &= \dim\operatorname{Ker}\left(\delta\oplus d^{+}\right)-\dim\operatorname{Ker}\left(\left(\delta\oplus d^{+}\right)^{*}\right) = b_1-\left(1+b_2^{+}\right)\\ &= \tfrac{1}{2}\left(b_1+b_3\right)-\left(1+\tfrac{1}{2}\left(b_2^{+}-b_2^{-}\right)+\tfrac{1}{2}\left(b_2^{+}+b_2^{-}\right)\right)\\ &= -\tfrac{1}{2}\left(1-b_1+b_2-b_3+1\right)-\tfrac{1}{2}\operatorname{sig}\left(X\right)\\ &= -\tfrac{1}{2}\left(\chi\left(X\right)+\operatorname{sig}\left(X\right)\right).\end{aligned}$$

Now $\mathcal{D}_c^{\omega}$ is locally a Dirac operator twisted by L'. Thus, we may apply the Local Index Formula for twisted Dirac operators (Theorem 18.45, p. 544). Since locally $L'\otimes L' = L$, on the level of forms we have (where $\omega' := \frac{1}{2}\pi_1^{*}\left(\omega\right)$) $c_1\left(L',\omega'\right) = 2c_1\left(L,\omega\right)$ and

$$\begin{aligned}\mathbf{ch}\left(L',\omega'\right) &= 1+c_1\left(L',\omega'\right)+\tfrac{1}{2}c_1\left(L',\omega'\right)^2\\ &= 1+\tfrac{1}{2}c_1\left(L,\omega\right)+\tfrac{1}{8}c_1\left(L,\omega\right)^2.\end{aligned}$$

Then the Local Index Formula yields

$$\begin{aligned}\operatorname{Index}\mathcal{D}_c^{\omega} &= \left(1+\tfrac{1}{2}c_1\left(L\right)+\tfrac{1}{8}c_1\left(L\right)^2\smile\widehat{\mathbf{A}}\left(X\right)\right)[X]\\ &= \left(1+\tfrac{1}{2}c_1\left(L\right)+\tfrac{1}{8}c_1\left(L\right)^2\smile\left(1-\tfrac{1}{24}p_1\left(TX\right)\right)\right)[X]\\ &= \left(\tfrac{1}{8}c_1\left(L\right)^2-\tfrac{1}{24}p_1\left(TX\right)\right)[X]\\ &= \tfrac{1}{8}c_1\left(L\right)^2[X]-\tfrac{1}{24}3\operatorname{sig}\left(X\right) = \tfrac{1}{8}\left(c_1\left(L\right)^2[X]-\operatorname{sig}\left(X\right)\right),\end{aligned}$$

where $p_1\left(TX\right)[X] = 3\operatorname{sig}\left(X\right)$ by the Hirzebruch Signature Theorem (Theorem 18.59, p. 573). Since $\operatorname{Index}B$ will turn out to be the dimension of a (perturbed) moduli space, we denote $\operatorname{Index}B$ by $d\left(X,L\right)$. In summary, the index of the real operator B is then given by

$$\begin{aligned}d\left(X,L\right) &:= \operatorname{Index}B = \operatorname{Index}B_1 = \operatorname{Index}\left(\delta\oplus d^{+}\right)+\operatorname{Index}_{\mathbb{R}}\mathcal{D}_c^{\omega}\\ &= b_1-\left(1+b_2^{+}\right)+\operatorname{Index}_{\mathbb{R}}\mathcal{D}_c^{\omega}\\ &= -\tfrac{1}{2}\left(\chi\left(X\right)+\operatorname{sig}\left(X\right)\right)+2\cdot\tfrac{1}{8}\left(c_1\left(L\right)^2[X]-\operatorname{sig}\left(X\right)\right)\\ &= \tfrac{1}{4}\left(c_1(L)^2-2\chi(X)-3\operatorname{sig}(X)\right).\end{aligned}\tag{19.44}$$

REMARK 19.42. Note that as $\operatorname{index}_{\mathbb{R}}(\mathcal{D}_c^{\omega}) = 2\cdot\operatorname{index}_{\mathbb{C}}(\mathcal{D}_c^{\omega})$ is even, $\operatorname{Index}_{\mathbb{R}}(B)\equiv\operatorname{Index}\left(\delta\oplus d^{+}\right)\bmod 2$. Since $\operatorname{Index}\left(\delta\oplus d^{+}\right) = b^1-b^{2+}-1$, $\operatorname{index}_{\mathbb{R}}(B)$ is even if and only if b^1+b^{2+} is odd (i.e., $\operatorname{Index}_{\mathbb{R}}(B)$ and b^1+b^{2+} have opposite parity).

Having computed the formal dimension $d(X, L)$ of moduli space, one would like to verify that under suitable conditions that it is a manifold. This requires the introduction of function spaces and hard analysis. The set of solutions of the S-W equations is $\Phi^{-1}(0,0)$. In this chapter, we have assumed thus far that all objects are in the C^∞ category. Spaces of C^∞ sections can be made into Fréchet spaces, but without extra conditions, the implicit function theorem fails for maps between Fréchet spaces. Thus, even if we could prove that $\Phi_{*(\omega,\psi)}$ is onto, one could not deduce that $\Phi^{-1}(0,0)$ is a manifold. Instead, we enlarge the spaces involved to suitable Sobolev spaces which are Hilbert (or Banach) spaces. In that category, one has inverse and implicit function theorems, as we have seen in Sections 77.3 and/or 17.4. Moreover, we will need

THEOREM 19.43 (Unique Continuation (see [**Ar**])). *Let $E \to X$ be a vector bundle over a connected manifold X and let $D : C^\infty(E) \hookleftarrow$ be a second-order, elliptic differential operator whose symbol is a multiple of the identity at each nonzero covector (i.e., a "scalar" symbol). If $Du = 0$ and $u = 0$ on some open set, then $u = 0$ on X.*

The appropriate Sobolev extension of the S-W function (19.40) , an appropriate extension is (for sufficiently large k)

$$\Phi^{2,k+1} : \mathcal{C}^{2,k+1}\left(P_{U(1)}\right) \times W^{2,k+1}\left(\Sigma_c^+(X)\right) \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right).$$

Note that we need $k + 1 > \frac{4}{2}$ (i.e., $k \geq 2$) in order that q define a bounded bilinear form on $W^{2,k+1}(\Sigma_c^+(X)) \times W^{2,k+1}(\Sigma_c^+(X)) \to W^{2,k+1}\left(\Lambda^{2+}(X, i\mathbb{R})\right)$ according to Proposition 17.22, p. 469. We remark that while a connection 1-form $\omega \in \mathcal{C}\left(P_{\mathrm{U}(1)}\right)$ is not a 1-form defined on X, any two connection 1-forms ω_1 and ω_0 on $P_{\mathrm{U}(1)}$ differ by a 1-form $\omega_1 - \omega_0$ that uniquely projects to a 1-form on X which we also denote by $\omega_1 - \omega_0$. Thus, given a fixed arbitrary choice of a connection $\omega_0 \in \mathcal{C}\left(P_{\mathrm{U}(1)}\right)$, the space $\mathcal{C}\left(P_{\mathrm{U}(1)}\right)$ can be identified with $\Omega^1(X, i\mathbb{R})$, via $\omega_1 \leftrightarrow \omega_1 - \omega_0$. Then

$$\mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right) := \omega_0 + W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right),$$

which is really independent of the choice of ω_0. Assume that $k \geq 2$. Let

$$W^{2,k+1}\left(\Sigma_c^+(X)\right)^* := W^{2,k+1}\left(\Sigma_c^+(X)\right) - \{0\} \text{ and}$$

$$\mathcal{CWP}_k := \mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right) \times W^{2,k+1}\left(\Sigma_c^+(X)\right)^* \times W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right). \tag{19.45}$$

The tangent space of the Banach manifold $\mathcal{CWP}_k$ at (ω, ψ, η) is

$$T_{(\omega,\psi,\eta)}\mathcal{CWP}_k := W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right) \times W^{2,k+1}\left(\Sigma_c^+(X)\right)^* \times W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right)$$

Define

$$F : \mathcal{CWP}_k \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right) \text{ by}$$

$$F(\omega, \psi, \eta) := \Phi^{2,k+1}(\omega, \psi) = (\Omega^{\omega+} - q(\psi) - \eta, \mathcal{D}_c^\omega \psi). \tag{19.46}$$

THEOREM 19.44. *If for some triple $(\omega, \psi, \eta) \in \mathcal{CWP}_k$ we have $F(\omega, \psi, \eta) = 0$, then the differential*

$$F_* : T_{(\omega,\psi,\eta)}\mathcal{CWP}_k \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right)$$

at (ω,ψ,η) *is given by*

$$F_*(\omega',\psi',\eta') = \left((d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') - \eta',\ \mathcal{D}_c^\omega\psi' + \tfrac{1}{2}\omega'\cdot\psi\right),$$

and F_* *is onto.*

Proof. Note that for $\eta' \in W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right),\ F_*(0,0,-\eta') = (\eta',0)$, so that $\operatorname{Im}(F_*) \supseteq W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)\oplus 0$. It suffices to show that for any $\xi' \in W^{2,k}(\Sigma_c^-(X))$, there is some $\omega' \in W^{2,k+1}\left(\Lambda^1(X,i\mathbb{R})\right)$ and $\psi' \in W^{2,k+1}(\Sigma_c^+(X))$, such that

$$\xi' = \mathcal{D}_c^\omega\psi' + \tfrac{1}{2}\omega'\cdot\psi,$$

since then

$$F_*\left(\omega',\psi',(d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') - \eta'\right) = (\eta',\xi').$$

Thus, it remains to show that the map

$$D : W^{2,k+1}\left(\Lambda^1(X,i\mathbb{R})\right)\times W^{2,k+1}\left(\Sigma_c^+(X)\right) \to W^{2,k}\left(\Sigma_c^-(X)\right)$$

given by

$$D(\omega',\psi') = \mathcal{D}_c^\omega\psi' + \tfrac{1}{2}\omega'\cdot\psi \tag{19.47}$$

is onto. Since $\mathcal{D}_c^\omega$ is elliptic, the operator D is also elliptic. Thus, according to Proposition 17.23 (applied to the elliptic formal adjoint D^* of D), D is onto if $\operatorname{Ker}(D^*)$ is 0. We determine D^* as follows. For $\xi \in C^\infty(\Sigma_c^-(X))$, we have (where $(\cdot,\cdot)$ is the L^2 inner product on $C^\infty(\Sigma_c^-(X))$)

$$\left(\mathcal{D}_c^\omega\psi' + \tfrac{1}{2}\omega'\cdot\psi,\xi\right) = (\psi',\mathcal{D}_c^\omega\xi) + \int_X \left\langle \omega', i\operatorname{Im}\left(\left\langle \tfrac{1}{2}cl(\cdot)\psi,\xi\right\rangle\right)\right\rangle_g \nu_g,$$

using the fact that ω' is $i\mathbb{R}$-valued. Thus,

$$\begin{aligned} D^*\xi &= i\operatorname{Im}\left(\left\langle \tfrac{1}{2}cl(\cdot)\psi,\xi\right\rangle\right)\oplus\mathcal{D}_c^\omega\xi \\ &= \tfrac{i}{4}\left(\langle cl(\cdot)\psi,\xi\rangle - \langle\xi, cl(\cdot)\psi\rangle\right)\oplus\mathcal{D}_c^\omega\xi. \end{aligned}$$

Hence

$$D^*\xi = 0 \Leftrightarrow \mathcal{D}_c^\omega\xi = 0 \ \text{ and } \ \operatorname{Im}\langle cl(\cdot)\psi,\xi\rangle = 0.$$

We have assumed that $\psi \neq 0$ and $\mathcal{D}_c^\omega\psi = 0$. If $\mathcal{D}_c^\omega\xi = 0$ and $\xi \neq 0$, then by the Theorem 19.43 (p. 633), neither ψ nor ξ can vanish on an open set. (Note that we are actually applying Theorem 19.43 with $D = \mathcal{D}_c^{\omega *}\mathcal{D}_c^\omega$ which is a second-order operator with "scalar" symbol, and $\mathcal{D}_c^\omega\xi = 0 \Leftrightarrow \mathcal{D}_c^{\omega *}\mathcal{D}_c^\omega\xi = 0$.) Hence there must be a point $x \in X$ where both ψ and ξ are nonzero. At x the covector $\operatorname{Im}\left(\left\langle \tfrac{1}{2}cl(\cdot)\psi,\xi\right\rangle\right)$ is nonzero. Indeed, first recall that Clifford multiplication takes $\mathbb{R}^4$ onto $\mathbb{R}\cdot \mathrm{SU}(\Sigma_2^+,\Sigma_2^-)$. Since $\mathrm{SU}(2)$ acts transitively on the unit sphere in $\mathbb{C}^2$, we know there is some vector $V \in T_x(X)$, so that $cl(V)\psi(x) = i\xi(x)$ and $\operatorname{Im}\left\langle \tfrac{1}{2}cl(V)\psi(x),\xi(x)\right\rangle = \operatorname{Im}\left\langle \tfrac{1}{2}i\xi(x),\xi(x)\right\rangle = \tfrac{1}{2}\langle\xi(x),\xi(x)\rangle > 0$. Thus, necessarily $\xi = 0$, and D in (19.47) is onto as required. □

Theorem 19.44 and Theorem 17.24 (Implicit Function Theorem I, p. 470) yield

COROLLARY 19.45. *The parametrized solution space*

$$\mathcal{SWP}_k := F^{-1}(0,0)$$
$$= \left\{(\omega,\psi,\eta) \in \mathcal{CWP}_k : \Omega^{\omega+} - q\,(\psi) - \eta = 0,\, \mathcal{D}_c^{\omega}\psi = 0\right\} \tag{19.48}$$

(if nonvoid) is a Hilbert submanifold of $\mathcal{CWP}_k$ *with tangent space at* (ω,ψ,η) *being the kernel of* F_*.

We need to indicate the sense in which the group $C^\infty(X,U(1))$ of gauge transformations can be enlarged to a C^∞ Banach manifold $W^{2,k+2}(X,U(1))$ (where $k \geq 1$) and group for which the group operation $(r,s) \mapsto rs^{-1}$ is C^∞. We have (where the last inclusion is compact by Proposition 17.20, p.468, since $2\,(k+2) > 4$)

$$W^{2,k+2}(X,\mathrm{U}(1)) := \left\{s \in W^{2,k+2}(X,\mathbb{C}) : s\bar{s} = 1 \text{ (a.e.)}\right\}$$
$$\subseteq W^{2,k+2}(X,\mathbb{C}) \overset{\iota}{\subseteq} C_0(X,\mathbb{C}),$$

and so $W^{2,k+2}(X,\mathrm{U}(1))$ inherits a topological metric from $W^{2,k+2}(X,\mathbb{C})$. Moreover, since $C_0(X,\mathrm{U}(1))$ is closed in $C_0(X,\mathbb{C})$, $W^{2,k+2}(X,\mathrm{U}(1)) = \iota^{-1}\left(C_0(X,\mathrm{U}(1))\right)$ is closed in $W^{2,k+2}(X,\mathbb{C})$. Hence, $W^{2,k+2}(X,\mathrm{U}(1))$ is a complete metric space. For any $s \in W^{2,k+2}(X,\mathrm{U}(1))$, we define a function

$$Exp_s : W^{2,k+2}(X,i\mathbb{R}) \to W^{2,k+2}(X,\mathrm{U}(1)) \text{ by } Exp_s(i\theta)(x) = e^{i\theta(x)}s(x).$$

For $i\theta \in W^{2,k+2}(X,i\mathbb{R})$, the fact that $Exp_s(i\theta) \in W^{2,k+2}(X,\mathrm{U}(1))$ follows from Proposition 17.22 (p. 469) and

PROPOSITION 19.46. *If* $f : E \to F$ *is any* C^∞ *fiber-preserving map (not necessarily linear on fibers), then left-composition by* f *defines a* C^∞ *map from* $W^{p,k}(E)$ *to* $W^{p,k}(F)$ *if* $k - \frac{n}{p} > 0$.

PROOF. See [**Pal68**, p. ???]. □

PROPOSITION 19.47. *For* $k \geq 1$, $W^{2,k+2}(X,\mathrm{U}(1))$ *is a* C^∞ *Banach manifold and* $W^{2,k+2}(X,\mathrm{U}(1))^2 \ni (t,s) \mapsto ts^{-1} = t\bar{s} \in W^{2,k+2}(X,\mathrm{U}(1))$ *is* C^∞.

PROOF. Let $d(z_1,z_2)$ denote the angular distance between z_1 and z_2 in the unit circle U(1). For $\rho \in [0,\pi]$, define

$$U_\rho(s) := \left\{t \in W^{2,k+2}(X,\mathrm{U}(1)) : d(t(x),s(x)) < \rho \text{ for all } x \in X\right\},$$
$$\text{and } O_\rho := \left\{i\theta \in W^{2,k+2}(X,i\mathbb{R}) : |i\theta(x)| < \rho \text{ for all } x \in X\right\}.$$

Since the inclusion $W^{2,k+2}(X,i\mathbb{R}) \subseteq C^0(X,i\mathbb{R})$ is continuous, O_ρ is open. Moreover,

$$Exp_s|O_\rho : O_\rho \to U_\rho(s)$$

is bijective. We denote the inverse by $Log_s : U_\rho(s) \to O_\rho$, and it is given by

$$Log_s(t)(x) = i \arg\left(t(x)\overline{s(x)}\right),$$

where $\arg(e^{i\theta}) = \theta$ for $|\theta| < \pi$. In order that the set of charts Log_s, $s \in W^{2,k+2}(X,\mathrm{U}(1))$, define a differentiable structure on $W^{2,k+2}(X,\mathrm{U}(1))$, we need to show that

$$Log_s \circ Log_t^{-1} : Log_t\left(U_\rho(s) \cap U_\rho(t)\right) \to Log_s\left(U_\rho(s) \cap U_\rho(t)\right)$$

is C^∞. Note that if $\rho < \frac{\pi}{2}$, then

$$\begin{aligned}\left(Log_s \circ Log_t^{-1}\right)(i\theta)(x) &= Log_s\left(Log_t^{-1}(i\theta)\right)(x) = i\arg\left(Log_t^{-1}(i\theta)(x)\overline{s(x)}\right) \\ &= i\arg\left(Exp_t(i\theta)(x)\overline{s(x)}\right) = i\arg\left(e^{i\theta(x)}t(x)\overline{s(x)}\right) \\ &= i\,\theta(x) + i\,\arg\left(t(x)\overline{s(x)}\right) = i\,\theta(x) + Log_s(t)(x).\end{aligned}$$

Thus, $Log_s \circ Log_t^{-1}$ is a translation of $W^{2,k+2}(X, i\mathbb{R})$ by $Log_s(t) \in W^{2,k+2}(X, i\mathbb{R})$, which is C^∞. Hence, $W^{2,k+2}(X, \mathrm{U}(1))$ is a Banach manifold. The mapping from $W^{2,k+2}(X, U(1))^2$ to $W^{2,k+2}(X, \mathrm{U}(1))$ given by $(t, s) \mapsto ts^{-1} = t\bar{s}$ is C^∞, since in terms of charts about t_0, s_0 and $t_0\bar{s}_0$ it is given by

$$\begin{aligned}&\left(i\arg\left(t\,\overline{t_0}\right), i\arg\left(s\,\overline{s_0}\right)\right) = \left(Log_{t_0}(t), Log_{s_0}(s)\right) \\ &\mapsto Log_{t_0\bar{s}_0}(t\bar{s}) = i\arg\left(t\,\bar{s}\,\overline{t_0\bar{s}_0}\right) = i\arg\left(t\,\overline{t_0}\bar{s}\,s_0\right) = i\arg\left(t\,\overline{t_0}\right) - i\arg\left(s\,\overline{s_0}\right).\end{aligned}$$

(i.e., $(u, v) \mapsto u - v$). One could also show that $W^{2,k+2}(X, \mathrm{U}(1))$ is a closed (Hilbert) submanifold and subgroup of $W^{2,k+2}(X, \mathbb{C}^*)$. Since $W^{2,k+2}(X, \mathbb{C}^*)$ is covered by a single chart in which the group operations are certainly C^∞ (using Propositions 17.22 and 19.46), the restrictions of the group operations to $W^{2,k+2}(X, \mathrm{U}(1))$ are smooth. Thus, $W^{2,k+2}(X, \mathrm{U}(1))$ is a closed Lie subgroup of $W^{2,k+2}(X, \mathbb{C}^*)$. □

For $k \geq 2$, $W^{2,k+2}(X, \mathrm{U}(1))$ acts smoothly and freely on $\mathcal{CWP}_k$ via

$$s \cdot (\omega, \psi, \eta) = \left(\omega + \pi_1^*\left(-2s^{-1}ds\right), s\psi, \eta\right) \tag{19.49}$$

Note that $s^{-1}ds \in W^{2,k+1}(X, \Lambda^1(X, i\mathbb{R}))$ by Proposition 17.22 (p. 469), since $s^{-1} \in W^{2,k+2}(X, \mathrm{U}(1))$ and $ds \in W^{2,k+1}(X, \Lambda^1(X, \mathbb{C})))$ and $2(k+1) > 4$; also $s^{-1}ds$ is $i\mathbb{R}$-valued, since

$$\overline{s^{-1}ds} = \bar{s}^{-1}d\bar{s} = sd\left(s^{-1}\right) = s(-s^{-2}ds) = -s^{-1}ds).$$

For $k \geq 3$, $s^{-1}ds$ is C^1 since $k + 1 - \frac{4}{2} \geq 2 > 1$ and Proposition 17.20 (p. 468) then applies. Thus, for $k \geq 3$ we can compute $d\left(s^{-1}ds\right) = -s^{-2}ds \wedge ds = 0$ in the usual way. The computations leading to (19.37) and (19.38) then apply to show that $\mathcal{SWP}_k \subseteq \mathcal{CWP}_k$ is invariant under the action of $W^{2,k+2}(X, \mathrm{U}(1))$ for $k \geq 3$. We wish to construct quotient manifolds $\mathcal{CWP}_k / W^{2,k+2}(X, \mathrm{U}(1))$ and $\mathcal{SWP}_k / W^{2,k+2}(X, \mathrm{U}(1))$. In doing this, we will need to repeatedly use a family of elliptic operators parametrized by $(\omega, \psi, \eta) \in \mathcal{CWP}_k$ for $k \geq 2$. Using (19.42) for the motivation (and noting that η is just carried along for the ride for now), this family is

$$\begin{aligned}&B_{(\omega,\psi,\eta),k} : W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right) \oplus W^{2,k+1}\left(\Sigma_c^+(X)\right) \\ &\to W^{2,k}(X, i\mathbb{R}) \oplus W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right),\end{aligned}$$

given by

$$B_{(\omega,\psi,\eta),k}(\omega', \psi') := \begin{bmatrix} \delta\omega' - \frac{1}{4}\left(\langle\psi', \psi\rangle - \langle\psi, \psi'\rangle\right) \\ d\omega'^{+} - \widetilde{q}(\psi', \psi) - \widetilde{q}(\psi, \psi') \\ \mathcal{D}_\omega\psi' + \frac{1}{2}\omega' \cdot \psi \end{bmatrix} \tag{19.50}$$

Proposition 17.22 (p. 469) is used in showing that this is well-defined. For the time being, both $(\omega, \psi, \eta) \in \mathcal{CWP}_k$ and $k \geq 2$ will be fixed, and we write $B_{(\omega,\psi,\eta),k}$ simply as B. It is convenient to introduce the following notation.

NOTATION 19.48. Let

$$U_1 := W^{2,k+1}\left(\Lambda^1(X,i\mathbb{R})\right), \quad U_2 := W^{2,k+1}\left(\Sigma_c^+(X)\right)$$
$$V_1 := W^{2,k}(X,i\mathbb{R}), \quad V_2 := W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right), \quad V_3 := W^{2,k}\left(\Sigma_c^-(X)\right).$$

Then we write

$$B = \left(B^1, B^2, B^3\right) : U_1 \oplus U_2 \to V_1 \oplus V_2 \oplus V_3, \text{ where}$$

$$\begin{array}{ll} B^1 : U_1 \oplus U_2 \to V_1 & \text{with } B^1(\omega',\psi') = \delta\omega' - \frac{1}{4}\left(\langle\psi',\psi\rangle - \langle\psi,\psi'\rangle\right) \\ B^2 : U_1 \oplus U_2 \to V_2 & \text{with } B^2(\omega',\psi') = d\omega'^{+} - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') \\ B^3 : U_1 \oplus U_2 \to V_3 & \text{with } B^3(\omega',\psi') = \mathcal{D}_c^{\omega}\psi' + \frac{1}{2}\omega'\cdot\psi. \end{array}$$

Since B is elliptic, the symbols of B^1, B^2, and B^3 are surjective but not injective. We need to compute the appropriate Sobolev extensions of the L^2 formal adjoints of B^1, B^2, and B^3.

$$\begin{array}{l} \left(B^1\right)^* : W^{2,k+2}(X,i\mathbb{R}) \to U_1 \oplus U_2 \\ \left(B^2\right)^* : W^{2,k+2}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \to U_1 \oplus U_2 \\ \left(B^3\right)^* : W^{2,k+2}\left(\Sigma_c^-(X)\right) \to U_1 \oplus U_2 \end{array}$$

The final results are:

$$\begin{array}{l} \left(B^1\right)^*(i\theta) = \left(id^{k+2}\theta, -\frac{i}{2}\theta\psi\right) \\ \left(B^2\right)^*(\gamma) = \left(\delta^{k+2}\gamma, S_\psi^*(\gamma)\right) \\ \left(B^3\right)^*(\xi) = \left(T_\psi^*(\xi), \left(\mathcal{D}_c^{\omega}\right)^{k+2}\xi\right), \end{array}$$

where S_ψ^* and T_ψ^* are 0-th order operators, described below. Since $i\mathbb{R}$ is a real vector space with real inner product $\langle a,b\rangle := a\bar{b}$, the spaces $\Omega^j(X,i\mathbb{R})$ are also real vector spaces with real inner products, and so the formal adjoints need to be computed relative to real L^2 inner products. Hence, although $\Sigma_c^{\pm}(X)$ has a Hermitian inner product, say $\langle\psi_1,\psi_2\rangle$, we must use the real inner product $\operatorname{Re}(\langle\psi_1,\psi_2\rangle) = \frac{1}{2}\left(\langle\psi_1,\psi_2\rangle + \langle\psi_2,\psi_1\rangle\right)$. We first compute $\left(B^1\right)^*$:

$$\begin{aligned} \int_X \left\langle i\theta, B^1(\omega',\psi')\right\rangle \nu_g &= \int_X \left\langle i\theta, \delta\omega' + \tfrac{1}{4}\left(\langle\psi,\psi'\rangle - \langle\psi',\psi\rangle\right)\right\rangle \nu_g \\ &= \int_X \langle i\theta, \delta\omega'\rangle\, \nu_g - \int_X \tfrac{1}{4}\left(\langle i\theta\psi,\psi'\rangle + \langle\psi', i\theta\psi\rangle\right) \nu_g \\ &= \int_X \langle id\theta, \omega'\rangle\, \nu_g - \int_X \tfrac{1}{2}\operatorname{Re}\left(\langle i\theta\psi,\psi'\rangle\right) \nu_g \\ &= \int_X \left\langle \left(id\theta, -\tfrac{i}{2}\theta\psi\right), (\omega',\psi')\right\rangle \nu_g. \end{aligned}$$

For $\left(B^2\right)^*$, we first note that there is a 0-th order differential operator $S_\psi :$ $C^\infty\left(\Sigma_c^+(X)\right) \to \Omega^{2+}(X,i\mathbb{R})$ given by $S_\psi(\psi') = \widetilde{q}(\psi',\psi) + \widetilde{q}(\psi,\psi')$. The adjoint of S_ψ, say $S_\psi^* : \Omega^{2+}(X,i\mathbb{R}) \to C^\infty\left(\Sigma_c^+(X)\right)$, has the property that for $\gamma \in \Omega^{2+}(X,i\mathbb{R})$,

$$\langle\gamma, \widetilde{q}(\psi',\psi) + \widetilde{q}(\psi,\psi')\rangle = \langle\gamma, S_\psi(\psi')\rangle = \operatorname{Re}\left\langle S_\psi^*(\gamma), \psi'\right\rangle.$$

While one can find an explicit expression for $S_\psi^*(\gamma)$, the important point is that $S_\psi^* : \Omega^{2+}(X,i\mathbb{R}) \to C^\infty(\Sigma_c^+(X))$ is again a 0-th order differential operator. The formal adjoint of B^2 is computed via

$$\begin{aligned}\int_X \left\langle \gamma, B^2\left(\omega', \psi'\right)\right\rangle \nu_g &= \int_X \left\langle \gamma, d\omega' - \widetilde{q}\left(\psi', \psi\right) - \widetilde{q}\left(\psi, \psi'\right)\right\rangle \nu_g \\ &= \int_X \left\langle \delta\gamma, \omega'\right\rangle \nu_g - \int_X \left\langle \gamma, \widetilde{q}\left(\psi', \psi\right) + \widetilde{q}\left(\psi, \psi'\right)\right\rangle \nu_g \\ &= \int_X \left\langle \delta\gamma, \omega'\right\rangle \nu_g - \int_X \operatorname{Re}\left\langle S_\psi^*\left(\gamma\right), \psi'\right\rangle \nu_g \\ &= \int_X \left\langle \left(\delta\gamma, -S_\psi^*\left(\gamma\right)\right), \left(\omega', \psi'\right)\right\rangle \nu_g. \qquad (19.51)\end{aligned}$$

For $\left(B^3\right)^*$, let $\xi \in C^\infty(\Sigma_c^-(X))$, and note that

$$\begin{aligned}\int_X \operatorname{Re}\left\langle \xi, B^3\left(\omega', \psi'\right)\right\rangle \nu_g &= \int_X \operatorname{Re}\left\langle \xi, \mathcal{D}_c^\omega \psi' + \tfrac{1}{2}\omega' \cdot \psi\right\rangle \nu_g \\ &= \int_X \operatorname{Re}\left\langle \xi, \tfrac{1}{2}\omega' \cdot \psi\right\rangle + \operatorname{Re}\left\langle \mathcal{D}_c^\omega \xi, \psi'\right\rangle \nu_g \\ &= \int_X \left\langle T_\psi^*\left(\xi\right), \omega'\right\rangle + \operatorname{Re}\left\langle \mathcal{D}_c^\omega \xi, \psi'\right\rangle \nu_g \\ &= \int_X \left\langle \left(T_\psi^*\left(\xi\right), \mathcal{D}_c^\omega \xi\right), \left(\omega', \psi'\right)\right\rangle \nu_g. \qquad (19.52)\end{aligned}$$

Here T_ψ^* is the 0-th order operator which is the adjoint of the 0-th order operator $T_\psi : \Omega^1(X, i\mathbb{R}) \to C^\infty(\Sigma_c^-(X))$ given by $T_\psi(\omega') = \frac{1}{2}\omega' \cdot \psi$.

Theorem 19.49. *For $k \geq 2$, the quotient space of moduli*

$$\mathcal{MCWP}_k := \mathcal{CWP}_k / W^{2,k+2}(X, \mathrm{U}(1))$$

has the structure of a Hausdorff C^∞ Hilbert manifold.

Proof. We first will produce a "local slice" of the action of $W^{2,k+2}(X, \mathrm{U}(1))$ on $\mathcal{CWP}_k$. Previously (see (19.39)) we found that at a point $(\omega, \psi, \eta) \in \mathcal{CWP}_k$, a vector $(\omega', \psi', \eta') \in T_{(\omega,\psi,\eta)}\left(\mathcal{CWP}_k\right)$ is formally L^2-orthogonal to the orbit

$$W^{2,k+2}(X, \mathrm{U}(1)) \cdot (\omega, \psi, \eta)$$

iff

$$\delta\omega' - \tfrac{1}{4}\left(\langle \psi', \psi\rangle - \langle \psi, \psi'\rangle\right) = 0,$$

i.e., iff $(\omega', \psi') \in \operatorname{Ker}\left(B^1\right)$. Note that

$$\begin{gathered} -2B^{1*} \oplus 0 : W^{2,k+2}\left(X, i\mathbb{R}\right) \to U_1 \oplus U_2 \oplus V_2, \\ \text{given by } \left(-2B_k^{1*} \oplus 0\right)(i\theta) = \left(-2id\theta, i\theta\psi, 0\right), \end{gathered}$$

is the differential at $1 \in W^{2,k+2}\left(X, \mathrm{U}(1)\right)$ of the map

$$\begin{gathered} W^{2,k+2}\left(X, \mathrm{U}(1)\right) \to \mathcal{CWP}_k \text{ given by} \\ s \mapsto s \cdot (\omega, \psi, \eta) = \left(\omega - 2\pi_1^*\left(s^{-1}ds\right), s\psi, \eta\right). \end{gathered}$$

Since the symbol of B^{1*} is injective, we have the decomposition (see Proposition 17.23, p. 469)

$$U_1 \oplus U_2 = \operatorname{Ker}\left(B^1\right) \oplus \operatorname{Im}\left(B^{1*}\right) \qquad (19.53)$$

into closed subspaces which are L^2-orthogonal. For a fixed (ω,ψ,η), we define a C^∞ map

$$\begin{aligned}&J = J_{(\omega,\psi,\eta)} : \left(\mathrm{Ker}(B^1)\times V_2\right)\times W^{2,k+2}\left(X,U(1)\right)\to \mathcal{CWP}_k \text{ by}\\ &J(x,s) := s\cdot\left((\omega,\psi,\eta)+x\right)\ ;\text{ i.e.,}\end{aligned}$$

$$\begin{aligned}J((\omega',\psi',\eta')\,,\,s) &= s\cdot\left((\omega,\psi,\eta)+(\omega',\psi',\eta')\right)\\ &= \left(\omega+\omega'-2\pi_1^*\left(s^{-1}ds\right),\, s\left(\psi+\psi'\right),\,\eta+\eta'\right).\end{aligned}$$

At $(0,0,1)\in\mathrm{Ker}(B^1)\times V_2\times W^{2,k+2}\left(X,U(1)\right)$, the differential

$$J_{*(0,0,1)} : \mathrm{Ker}\left(B^1\right)\oplus V_2\oplus V_1\to T_{(\omega,\psi,\eta)}\left(\mathcal{CWP}_k\right)\cong U_1\oplus U_2\oplus V_2$$

is given by

$$\begin{aligned}J_{*(0,0,1)}((\omega',\psi')\,,\eta',\,i\theta) &= (\omega'-2id\theta, i\theta\psi+\psi',\eta')\\ &= (\omega',\psi',\eta')+(-2id\theta,i\theta\psi,0) = (\omega',\psi',\eta')-2B^{1*}\left(i\theta\right).\end{aligned}$$

Thus, on $\mathrm{Ker}\left(B^1\right)\oplus V_2\oplus 0$, $J_{*(0,0,1)}$ is the "inclusion" $\mathrm{Ker}\left(B^1\right)\oplus V_2\oplus 0\subseteq U_1\oplus U_2\oplus V_2$, and on $0\oplus 0\oplus V_1$, $J_{*(0,0,1)}$ is $-2B^{1*}: V_1\to U_1\oplus U_2\oplus 0$. Thus using (19.53), we have that $J_{*(0,0,1)}$ is onto. The kernel of $J_{*(0,0,1)}$ is trivial. Indeed, since $\mathrm{Ker}\left(B^1\right)\perp\mathrm{Im}\left(B^{1*}\right)$, we have

(19.54) $$(\omega',\psi',\eta')+B^{1*}\left(i\theta\right)=0\Rightarrow(\omega',\psi',\eta')=0\text{ and }B^{1*}\left(i\theta\right)=0,$$

and $i\theta=0$, since $\mathrm{Ker}(B^{1*})$ is trivial (recall $\psi\neq 0$). Thus, by the Inverse Function Theorem, Theorem 17.26 (p. 470), there are open neighborhoods U of $0\in\mathrm{Ker}\left(B^1\right)\times V_2$ and V of $1\in W^{2,k+2}\left(X,\mathrm{U}(1)\right)$, such that $J|\left(U\times V\right): U\times V\to J\left(U\times V\right)$ is a diffeomorphism onto a neighborhood $J\left(U\times V\right)$ of $(\omega,\psi,\eta)\in\mathcal{CWP}_k$. Thus, points of $\mathcal{CWP}_k$ near (ω,ψ,η) are uniquely of the form

$$J(x,s)=s\cdot\left((\omega,\psi,\eta)+x\right)\ \text{ for } x\in\mathrm{Ker}(B^1)\times V_2$$

(i.e., for x in the "normal space" to the orbit of (ω,ψ,η)). In other words, U is a "local slice of the action". On our way to a global slice, let

$$Q:\mathcal{CWP}_k\to\mathcal{CWP}_k/\,W^{2,k+2}\left(X,\mathrm{U}(1)\right)$$

be the quotient function and define a diffeomorphism

$$\begin{aligned}&j_{(\omega,\psi,\eta)}: U\to J_{(\omega,\psi,\eta)}\left(U\times\{1\}\right)\ \text{ by}\\ &j_{(\omega,\psi,\eta)}\left(x\right):=J_{(\omega,\psi,\eta)}\left(x,1\right)=(\omega,\psi,\eta)+x.\end{aligned}$$

We now show that the mapping

$$\varphi_{(\omega,\psi,\eta)}: U\to\mathcal{CWP}_k/\,W^{2,k+2}\left(X,\mathrm{U}(1)\right)$$

given by

(19.55) $$\varphi_{(\omega,\psi,\eta)}\left(K\right):=\left[(\omega,\psi,\eta)+K\right]=Q\circ j_{(\omega,\psi,\eta)}\left(K\right)$$

is 1-1 if U is chosen small enough. Suppose that for $K_1=(\omega_1,\psi_1,\eta_1)$ and $K_2=(\omega_2,\psi_2,\eta_2)\in U$, we have $\varphi_{(\omega,\psi,\eta)}\left(K_1\right)=\varphi_{(\omega,\psi,\eta)}\left(K_2\right)$. Then $s\cdot\left((\omega,\psi,\eta)+K_1\right)=$

$(\omega,\psi,\eta)+K_2$ for some $s\in W^{2,k+2}(X,\mathrm{U}(1))$, i.e.,

$$\begin{aligned} &s\cdot((\omega,\psi,\eta)+(\omega_1,\psi_1,\eta_1))\\ &=\left(\omega-2\pi_1^*\left(s^{-1}ds\right)+\omega_1,s\psi+s\psi_1,\eta+\eta_1\right)\\ &=(\omega+\omega_2,\psi+\psi_2,\eta+\eta_2). \end{aligned} \tag{19.56}$$

Thus, $\omega-2\pi_1^*\left(s^{-1}ds\right)+\omega_1=\omega+\omega_2$, $s\psi+s\psi_1=\psi+\psi_2$, and $\eta_1=\eta_2$. Hence,

$$2\pi_1^*\left(s^{-1}ds\right)=\omega_1-\omega_2 \quad\text{and}\quad (s-1)\,\psi=(\psi_2-\psi_1)-(s-1)\,\psi_1.$$

Now, let

$$U_\varepsilon:=\Big\{(\omega',\psi',\eta')\in U:\|\psi'\|_{2,k+1}\le\varepsilon \text{ and } \|\omega'\|_{2,k+1}\le\varepsilon\Big\}\subseteq U.$$

For K_1, $K_2\in U_\varepsilon$, we have

$$\left\|2s^{-1}ds\right\|_{2,k+1}=\|\omega_1-\omega_2\|_{2,k+1}\le 2\varepsilon, \tag{19.57}$$

and for some constant C,

$$\begin{aligned} &\|(s-1)\,\psi\|_{2,k+1}\le\|\psi_2-\psi_1\|+\|(s-1)\,\psi_1\|_{2,k+1}\\ &\le 2\varepsilon+C\,\|s-1\|_{2,k+1}\|\psi_1\|_{2,k+1}\le\Big(2+C\,\|s-1\|_{2,k+1}\Big)\,\varepsilon. \end{aligned}$$

Proposition 17.20 and (19.57) yield $\left\|s^{-1}ds\right\|_{C^0}\le C'\left\|s^{-1}ds\right\|_{2,k+1}\le C'\varepsilon$, for some constant C' and $k\ge 2$. Thus, for ε sufficiently small, we deduce that s is arbitrarily C_0-close to some constant, say s_0, with $\|s-s_0\|_{C^0}\le\varepsilon_0$. We show $s_0=1$. If $s_0\ne 1$, then for $\varepsilon_0<\frac12|1-s_0|$, $|s(x)-1|>\frac12|1-s_0|>0$ for all $x\in X$. For $k\ge 2$, there is a also a constant C'' such that

$$\|(s-1)\,\psi\|_{C^0}\le C''\,\|(s-1)\,\psi\|_{2,k+1}\le C''\Big(2+C\,\|(s-1)\|_{2,k+1}\Big)\,\varepsilon \tag{19.58}$$

can be made arbitrarily small with ε. However, since $\psi\ne 0$, there is some x_0 with $\psi(x_0)\ne 0$ and

$$|(s(x_0)-1)\,\psi(x_0)|=|(s(x_0)-1)|\,|\psi(x_0)|\ge\tfrac12|1-s_0|\,|\psi(x_0)|,$$

which contradicts (19.58). Thus, $s_0=1$, and (19.57) then implies that $s\in V$ for sufficiently small ε. Since $J|U_\varepsilon\times V$ is 1-1 and $s\in V$, (19.56) then yields

$$J(K_1,s)=J(K_2,1)\Rightarrow(K_1,s)=(K_2,1)\Rightarrow K_1=K_2.$$

Thus, by replacing U by U_ε for sufficiently small ε, the mapping

$$\varphi_{(\omega,\psi,\eta)}:U\to\varphi_{(\omega,\psi,\eta)}(U)=Q\circ j_{(\omega,\psi,\eta)}(U)\subseteq\mathcal{CWP}_k/\,W^{2,k+2}(X,\mathrm{U}(1))$$

is 1-1. We wish to show that the

$$\varphi^{-1}_{(\omega,\psi,\eta)}:\varphi_{(\omega,\psi,\eta)}(U)\to U \tag{19.59}$$

form a collection of coordinate charts for $\mathcal{CWP}_k/\,W^{2,k+2}(X,\mathrm{U}(1))$ with C^∞ coordinate transitions. Suppose that we have two such charts, say

$$\begin{aligned} &\varphi^{-1}_{(\omega_1,\psi_1,\eta_1)}:\varphi_{(\omega,\psi,\eta)}(U_1)\to U_1 \text{ and } \varphi^{-1}_{(\omega_2,\psi_2,\eta_2)}:\varphi_{(\omega,\psi,\eta)}(U_2)\to U_2\\ &\text{with } [(\omega_0,\psi_0,\eta_0)]\in\varphi_{(\omega_1,\psi_1,\eta_1)}(U_1)\cap\varphi_{(\omega_2,\psi_2,\eta_2)}(U_2). \end{aligned}$$

We may assume $(\omega_0,\psi_0,\eta_0)\in j_{(\omega_1,\psi_1,\eta_1)}(U_1)$. There is $s_0\in W^{2,k+2}(X,\mathrm{U}(1))$, such that $(\omega_0,\psi_0,\eta_0)\cdot s_0\in j_{(\omega_2,\psi_2,\eta_2)}(U_2)$. We show that there is a neighborhood U_1' of

$(\omega_0, \psi_0, \eta_0)$ in $j_{(\omega_1,\psi_1,\eta_1)}(U_1)$ and some $s = e^{i\theta} \in C^\infty\left(U_1', W^{2,k+2}(X, \mathrm{U}(1))\right)$, such that for all $x \in U_1'$

$$R_{s_0 s}(x) := x \cdot s_0 \cdot s(x) \in j_{(\omega_2,\psi_2,\eta_2)}(U_2),$$

where $R_{s_0 s} : U_1' \to \mathcal{CWP}_k$ is C^∞. The following computation will then yield $\varphi^{-1}_{(\omega_2,\psi_2,\eta_2)} \circ \varphi_{(\omega_1,\psi_1,\eta_1)} = \left(j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1} \circ R_{s_0 s} \circ j_{(\omega_1,\psi_1,\eta_1)}$ on U_1', which is a composition of C^∞ mappings, as required: For $x = j_{(\omega_1,\psi_1,\eta_1)}(y) \in U_1'$ (i.e., $y \in j^{-1}_{(\omega_1,\psi_1,\eta_1)}(U_1') \subseteq U_1$),

$$\begin{aligned}
&\left(\varphi^{-1}_{(\omega_2,\psi_2,\eta_2)} \circ \varphi_{(\omega_1,\psi_1,\eta_1)}\right)(y) = \left(\left(Q \circ j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1} \circ \left(Q \circ j_{(\omega_1,\psi_1,\eta_1)}\right)\right)(y) \\
&= \left(Q \circ j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1}([x]) = \left(Q \circ j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1}([x \cdot s_0 \cdot s(x)]) \\
&= \left(Q \circ j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1}([R_{s_0 s}(x)]) = \left(j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1}(R_{s_0 s}(x)) \\
&= \left(\left(j_{(\omega_2,\psi_2,\eta_2)}\right)^{-1} \circ R_{s_0 s} \circ j_{(\omega_1,\psi_1,\eta_1)}\right)(y).
\end{aligned}$$

We seek $\theta \in C^\infty\left(U_1', W^{2,k+2}(X, \mathbb{R})\right)$ such that, for

$$s = e^{i\theta} \in C^\infty\left(U_1', W^{2,k+2}(X, \mathrm{U}(1))\right),$$

we have

$$B^1_{(\omega_2,\psi_2,\eta_2)}\left((\omega,\psi,\eta) \cdot s_0 \cdot s(\omega,\psi,\eta) - (\omega_2,\psi_2,\eta_2)\right) = 0$$

for all (ω,ψ,η) in a neighborhood of (ω_0,ψ_0,η_0) in $j_{(\omega_1,\psi_1,\eta_1)}(U_1)$. Define

$$F : j_{(\omega_1,\psi_1,\eta_1)}(U_1) \times W^{2,k+2}(X,\mathbb{R}) \to W^{2,k}(X,\mathbb{R}) \text{ by}$$

$$\begin{aligned}
&F((\omega,\psi,\eta),\theta) = B^1_{(\omega_2,\psi_2,\eta_2)}\left((\omega,\psi,\eta) \cdot s_0 \cdot e^{i\theta} - (\omega_2,\psi_2,\eta_2)\right) \\
&= B^1_{(\omega_2,\psi_2,\eta_2)}\left(\omega - \omega_2 - 2s_0^{-1}ds_0 - 2id\theta,\ e^{i\theta}s_0\psi - \psi_2,\ \eta - \eta_2\right). \qquad (19.60)
\end{aligned}$$

Using the fact that $(\omega_0,\psi_0,\eta_0)\cdot s_0 \in J_{(\omega_2,\psi_2,\eta_2)}(U_2)$, we have $F((\omega_0,\psi_0,\eta_0),0) = 0$. The existence of θ will follow from the Implicit Function Theorem (Theorem 17.25, p. 470), provided that at $((\omega_0,\psi_0,\eta_0),0)$

$$\partial_\theta F : W^{2,k+2}(X,\mathbb{R}) \to W^{2,k}(X,\mathbb{R})$$

is onto. However, using

$$B^1_{(\omega_2,\psi_2,\eta_2)}(\omega',\psi',\eta') = \delta\omega' + \tfrac{1}{4}\left(\langle\psi,\psi'\rangle - \langle\psi',\psi\rangle\right), \text{ we have}$$

$$\begin{aligned}
\partial_\theta F(\theta') &= B^1_{(\omega_2,\psi_2,\eta_2)}\left(2id\theta', -i\theta' s_0^{-1}\psi_0, 0\right) \\
&= 2i\delta d\theta' + \tfrac{1}{4}\left(\left\langle\psi_2, -i\theta' s_0^{-1}\psi_0\right\rangle - \left\langle -i\theta' s_0^{-1}\psi_0, \psi_2\right\rangle\right) \\
&= 2i\delta d\theta' + \tfrac{1}{4}\left(i\theta'\left\langle\psi_2, s_0^{-1}\psi_0\right\rangle + i\theta'\left\langle s_0^{-1}\psi_0, \psi_2\right\rangle\right) \\
&= 2i\left(\delta d\theta' + \tfrac{1}{8}\theta'\left(\left\langle s_0^{-1}\psi_0, \psi_2\right\rangle + \left\langle\psi_2, s_0^{-1}\psi_0\right\rangle\right)\right).
\end{aligned}$$

Since the elliptic operator

$$\theta' \mapsto \delta d\theta' + \tfrac{1}{8}\theta'\left(\left\langle s_0^{-1}\psi_0, \psi_2\right\rangle + \left\langle\psi_2, s_0^{-1}\psi_0\right\rangle\right)$$

is self-adjoint, it suffices to prove that its kernel is trivial. If $\partial_\theta F(\theta') = 0$, then

$$0 = \int_X \left\langle \delta d\theta' + \tfrac{1}{8}\theta' \left(\left\langle s_0^{-1}\psi_0, \psi_2 \right\rangle + \left\langle \psi_2, s_0^{-1}\psi_0 \right\rangle \right), \theta' \right\rangle \nu_g$$

$$= \int_X \left| d\theta' \right|^2 + \tfrac{1}{8} \left(\left\langle s_0^{-1}\psi_0, \psi_2 \right\rangle + \left\langle \psi_2, s_0^{-1}\psi_0 \right\rangle \right) \theta'^2 \, \nu_g. \tag{19.61}$$

In the case $s_0^{-1}\psi_0 = \psi_2$, this last equation holds only if $\theta' = 0$. However, in general we must ensure that the neighborhoods U associated with points (ω, ψ, η) are chosen small enough so that for all $\left(\widetilde{\omega}, \widetilde{\psi}, \widetilde{\eta}\right) \in j_{(\omega,\psi,\eta)}(U)$, the operator

$$\Delta_{\widetilde{\psi}} : \theta \mapsto \delta d\theta + \tfrac{1}{4} \left(\left\langle \widetilde{\psi}, \psi \right\rangle + \left\langle \psi, \widetilde{\psi} \right\rangle \right) \theta$$

is invertible. This is certainly true for $\left(\widetilde{\omega}, \widetilde{\psi}, \widetilde{\eta}\right) = (\omega, \psi, \eta)$. The invertible operators from $W^{2,k+2}(X, \mathbb{R})$ to $W^{2,k}(X, \mathbb{R})$ form an open set in the Banach space $\operatorname{Hom}\left(W^{2,k+2}(X, \mathbb{R}), W^{2,k}(X, \mathbb{R})\right)$ of continuous linear maps. By Proposition 17.22 (p. 469), we have

$$\begin{aligned} \left\| \left(\Delta_{\widetilde{\psi}} - \Delta_\psi \right) \theta \right\|_{2,k} &= \left\| \tfrac{1}{4} \left(\left\langle \widetilde{\psi} - \psi, \psi \right\rangle + \left\langle \psi, \widetilde{\psi} - \psi \right\rangle \right) \theta \right\|_{2,k} \\ &\le C \left\| \psi \right\|_{2,k} \left\| \widetilde{\psi} - \psi \right\|_{2,k} \left\| \theta \right\|_{2,k} \\ &\le C \left\| \psi \right\|_{2,k} \left\| \widetilde{\psi} - \psi \right\|_{2,k+1} \left\| \theta \right\|_{2,k+2} \end{aligned}$$

Thus, $\left\| \Delta_{\widetilde{\psi}} - \Delta_\psi \right\| \le C \left\| \psi \right\|_{2,k} \left\| \widetilde{\psi} - \psi \right\|_{2,k+1}$ which shows that $\Delta_{\widetilde{\psi}}$ will be invertible for $\widetilde{\psi}$ sufficiently $W^{2,k+1}$-close (in fact $W^{2,k}$-close) to ψ. Thus, the neighborhoods U associated with points (ω, ψ, η) can be chosen small enough so that the coordinate transition functions will be smooth.

It remains to show that the topology of $\mathcal{CWP}_k / W^{2,k+2}(X, \mathrm{U}(1))$ induced by the charts is Hausdorff. Since all spaces involved are second countable, it suffices to show that if a sequence $[x_n]$ in $\mathcal{CWP}_k / W^{2,k+2}(X, \mathrm{U}(1))$ converges to both $[y]$ and $[z]$, then $[y] = [z]$. If $[x_n]$ converges to both $[y]$ and $[z]$, then there are sequences r_n and $t_n \in W^{2,k+2}(X, \mathrm{U}(1))$ such that $r_n \cdot x_n \to y$ and $t_n \cdot x_n \to z$. Let $y_n = r_n \cdot x_n$, then $y_n \to y$ and for $s_n := r_n^{-1} t_n$, we have $s_n \cdot y_n = t_n \cdot x_n \to z$. Thus, it suffices to prove that if $y_n \to y$ and $s_n \cdot y_n \to z$, then there is a convergent subsequence, say $s_{n_i} \to s$, in which case $s_{n_i} \cdot y_n \to y \cdot s$, and so $z = y \cdot s$ (since $\mathcal{CWP}_k$ is a metric space) and $[y] = [z]$. In other words, we need to show that the action of $W^{2,k+2}(X, \mathrm{U}(1))$ on $\mathcal{CWP}_k$ is *proper*. To show this, let

$$(\omega_n, \psi_n, \eta_n) \to (\omega, \psi, \eta) \in \mathcal{CWP}_k, \tag{19.62}$$

and suppose that

$$s_n \cdot (\omega_n, \psi_n, \eta_n) \to (\beta, \phi, \zeta) \quad \text{for } \; s_n \in W^{2,k+2}(X, \mathrm{U}(1)).$$

Then

$$\begin{aligned} &\omega_n - 2 s_n^{-1} ds_n = s_n \cdot \omega_n =: \beta_n \to \beta \in \mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right) \\ &\Rightarrow -2 s_n^{-1} ds_n = \beta_n - \omega_n \to \beta - \omega \in W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right). \end{aligned}$$

Since $(k+2) - \frac{4}{2} > 0$ for $k > 0$,

$$s_n \in W^{2,k+2}(X, \mathrm{U}(1)) \Rightarrow s_n \in C^0(X, \mathrm{U}(1)) \text{ and } \left| s_n(x) \right|^2 = 1.$$

Thus, s_n is bounded in $W^{p,0}(X, \mathrm{U}(1))$ for all $p \geq 1$. Also, $2s_n^{-1} ds_n = \beta_n - \omega_n \in W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right) \subseteq W^{2,2}\left(\Lambda^1(X, i\mathbb{R})\right)$. Thus, by Proposition 17.22 (p. 469) with

$$k_3 - \tfrac{n}{p_3} = 0 - \tfrac{4}{5} < -\tfrac{4}{6} = \left(0 - \tfrac{4}{6}\right) + \left(2 - \tfrac{4}{2}\right) = \left(k_1 - \tfrac{n}{p_1}\right) + \left(k_2 - \tfrac{n}{p_2}\right),$$

we have that

$$ds_n = \tfrac{1}{2} s_n (\beta_n - \omega_n) \in W^{5,0}\left(\Lambda^1(X, i\mathbb{R})\right), \text{ and}$$

$$\|ds_n\|_{5,0} \leq C \|s_n\|_{6,0} \ \|\beta_n - \omega_n\|_{2,2}. \tag{19.63}$$

Now $\beta_n - \omega_n$ is bounded in $W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right)$ and hence in $W^{2,q}\left(\Lambda^1(X, i\mathbb{R})\right)$ for $q \leq k+1$ (e.g., for $q \leq 3$). Thus, as both s_n and ds_n are bounded in $W^{5,0}(X, \mathrm{U}(1))$, we have that s_n bounded in $W^{5,1}(X, \mathrm{U}(1))$. Since $\beta_n - \omega_n$ is bounded in $W^{2,3}\left(X, \Lambda^1(i\mathbb{R})\right)$, it is also bounded in $W^{5,1}(X, \mathrm{U}(1))$, since $1 - \frac{4}{5} \leq 2 - \frac{4}{3}$. Hence, by (19.63) and the Proposition 17.22 (p. 469) with $k_i = 1$, $p_i = 5$, $k_i - \frac{4}{p_i} = \frac{1}{5} > 0$, $i = 1, 2, 3$,

$$\|ds_n\|_{5,1} \leq C \|s_n\|_{5,1} \ \|B_n - \omega_n\|_{5,1}$$

Thus, $\|ds_n\|_{5,1}$ is bounded, and hence $\|s_n\|_{5,2}$ is bounded. Now, as $k \geq 2$,

$$W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right) \subseteq W^{2,3}\left(\Lambda^1(X, i\mathbb{R})\right) \subseteq W^{3,2}\left(\Lambda^1(X, i\mathbb{R})\right),$$

since $3 - \frac{4}{2} > 2 - \frac{4}{3}$. Hence, as $\beta_n - \omega_n \in W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right)$ is bounded in $W^{3,2}\left(\Lambda^1(X, i\mathbb{R})\right)$, we have

$$\|ds_n\|_{3,2} \leq C \|s_n\|_{5,2} \ \|\beta_n - \omega_n\|_{3,2}.$$

Thus, $\|s_n\|_{3,3}$ is bounded, and consequently $\|s_n\|_{2,3}$ is bounded. Inductively, suppose that $\|s_n\|_{2,j+1}$ is bounded for some $2 \leq j \leq k$. Since $j + 1 - \frac{4}{2} = j - 1 > 0$, Proposition 17.22 (p. 469) yields

$$\|ds_n\|_{2,j+1} \leq C \|s_n\|_{2,j+1} \ \|B_n - \omega_n\|_{2,j+1},$$

so that $\|s_n\|_{2,j+2}$ is bounded. Hence, by induction, $\|s_n\|_{2,k+2}$ is bounded. By Theorem 7.15 (p. 198), there is a subsequence, say s_n', of s_n which converges in $W^{2,k+1}(X, \mathrm{U}(1))$. Now, in $W^{2,k+1}(X, \mathrm{U}(1))$,

$$\begin{aligned} ds_n' - ds_m' &= \tfrac{1}{2} s_n' (\beta_n' - \omega_n') - \tfrac{1}{2} s_m' (\beta_m' - \omega_m') \\ &= \tfrac{1}{2} (s_n' - s_m') (\beta_n' - \omega_n') + \tfrac{1}{2} s_m' (\beta_n' - \omega_n') - \tfrac{1}{2} s_m' (\beta_m' - \omega_m') \\ &= \tfrac{1}{2} (s_n' - s_m') (\beta_n' - \omega_n') + \tfrac{1}{2} s_m' \left((\beta_n' - \omega_n') - (\beta_m' - \omega_m')\right). \end{aligned}$$

Thus, by Proposition 17.22 (p. 469),

$$\begin{aligned} \|ds_n' - ds_m'\|_{2,k+1} \leq{}& C_1 \|(s_n' - s_m')\|_{2,k+1} \|\beta_n' - \omega_n'\|_{2,k+1} \\ &+ C_2 \|s_m'\|_{2,k+1} \|(\beta_n' - \omega_n') - (\beta_m' - \omega_m')\|_{2,k+1}. \end{aligned}$$

Hence, ds_n' converges in $W^{2,k+1}(X, \mathbb{C})$, and so s_n' converges in $W^{2,k+2}(X, \mathrm{U}(1))$, since s_n' is known to converge in $W^{2,0}(X, \mathrm{U}(1))$. □

We now turn our attention to defining the manifold structure for the set

$$\mathcal{MSWP}_k := \mathcal{SWP}_k / W^{2,k+2}(X, U(1))$$

called the parametrized moduli space of solutions of the perturbed SW equations, where the parameter is the perturbation η ranging over $W^{2,k}(X, \Lambda^{2+}(X))$.

THEOREM 19.50. *For $k \geq 3$, the parametrized moduli space $\mathcal{MSWP}_k$ (if nonempty) is a closed Hilbert submanifold of $\mathcal{MCWP}_k$. At a point $(\omega, \psi, \eta) \in \mathcal{SWP}_k$, the differential $Q_{*(\omega,\psi,\eta)}$ of the projection $Q : \mathcal{CWP}_k \to \mathcal{MCWP}_k$ restricts to an isomorphism*

$$Q_{*(\omega,\psi,\eta)} : T_{(\omega,\psi,\eta)} \left(\mathcal{SWP}_k\right) \cap \operatorname{Ker}\left(B^1\right) \to T_{[(\omega,\psi,\eta)]} \left(\mathcal{MSWP}_k\right).$$

In other words, the tangent space at $[(\omega, \psi, \eta)] \in \mathcal{MSWP}_k$ is the projection of the orthogonal complement of the tangent space of the orbit $W^{2,k+2}(X, \mathrm{U}(1)) \cdot (\omega, \psi, \eta)$ in $T_{(\omega,\psi,\eta)}(F^{-1}(0,0)) = \operatorname{Ker}\left(F_{*(\omega,\psi,\eta)}\right)$, namely

$$\begin{aligned}
&T_{[(\omega,\psi,\eta)]} \left(\mathcal{MSWP}_k\right)\\
&\cong \operatorname{Ker}\left(F_{*(\omega,\psi,\eta)}\right) \cap \left\{(\omega', \psi', \eta') : \delta\omega' - \tfrac{1}{2}\left(\langle\psi', \psi\rangle - \langle\psi, \psi'\rangle\right) = 0\right\}\\
&= \Big\{(\omega', \psi', \eta') : \left(d\omega'\right)^+ - \widetilde{q}\left(\psi', \psi\right) - \widetilde{q}\left(\psi, \psi'\right) - \eta' = 0,\\
&\mathcal{D}_c^\omega \psi' + \tfrac{1}{2}\omega' \cdot \psi = 0,\ \delta\omega' - \tfrac{1}{2}\left(\langle\psi', \psi\rangle - \langle\psi, \psi'\rangle\right) = 0\Big\}\\
&= \left\{(\omega', \psi', \eta') : B^1\left(\omega', \psi'\right) = 0, B^2\left(\omega', \psi'\right) - \eta' = 0, B^3\left(\omega', \psi'\right) = 0\right\}. \qquad (19.64)
\end{aligned}$$

PROOF. Recall that $\mathcal{SWP}_k$ was shown to be a submanifold of $\mathcal{CWP}_k$ which is invariant under $W^{2,k+2}(X, \mathrm{U}(1))$ for $k \geq 3$. We need to exhibit $\mathcal{MSWP}_k$ as a submanifold of $\mathcal{MCWP}_k$. This will be accomplished by showing, via the Implicit Function Theorem, that $\mathcal{MSWP}_k$ is a submanifold when viewed in a coordinate neighborhood of $\mathcal{MCWP}_k$. Recall (see (19.46)) that

$$\begin{aligned}
&F : \mathcal{CWP}_k \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right) \text{ is given by}\\
&F(\omega, \psi, \eta) = \left(\Omega^{\omega+} - q\left(\psi\right) - \eta, \mathcal{D}_c^\omega \psi\right).
\end{aligned}$$

A standard coordinate chart about $[\omega, \psi, \eta] \in \mathcal{MCWP}_k$ is $\varphi^{-1}_{(\omega,\psi,\eta)} : \varphi_{(\omega,\psi,\eta)}(U) \to U$ (see (19.55) and (19.59)), where

$$\varphi_{(\omega,\psi,\eta)} = Q \circ j_{(\omega,\psi,\eta)} : U \to \mathcal{MSWP}_k$$

and $U \subseteq \operatorname{Ker}(B^1) \times V_2$ and $\left(Q \circ j_{(\omega,\psi,\eta)}\right)(x) = [x + (\omega, \psi, \eta)]$. Since (ω, ψ, η) will be fixed, we drop the subscript (ω, ψ, η). It suffices to show that in terms of the coordinate chart about $(\omega, \psi, \eta) \in \mathcal{SWP}_k$, we have that $\varphi^{-1}\left(\mathcal{MSWP}_k \cap \varphi(U)\right)$ is a submanifold of U near $0 \in U$. We have

$$\begin{aligned}
&\varphi^{-1}\left(\mathcal{MSWP}_k \cap \varphi(U)\right) = (Q \circ j)^{-1}\left(\mathcal{MSWP}_k \cap \varphi(U)\right)\\
&= j^{-1} \circ (Q|j(U))^{-1}\left(\mathcal{MSWP}_k \cap \varphi(U)\right) = j^{-1}\left(\mathcal{SWP}_k \cap j(U)\right)\\
&= j^{-1}\left(F^{-1}(0,0) \cap j(U)\right) = (F \circ j)^{-1}(0,0).
\end{aligned}$$

By the Implicit Function Theorem I (Theorem 17.24, p. 470), we need to show that $(F \circ j)_{*0}$ is onto. Now

$$(F \circ j)_{*0} : \operatorname{Ker}(B^1) \oplus W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \to V_2 \oplus V_3, \qquad (19.65)$$

is given by

$$(F \circ j)_{*0}\left(\omega', \psi', \eta'\right) = F_{*(\omega,\psi,\eta)}\left(\omega', \psi', \eta'\right) = \left(B^2\left(\omega', \psi'\right) - \eta', B^3\left(\omega', \psi'\right)\right).$$

We assume that $(\omega, \psi, \eta) \in \mathcal{SWP}_k$, in which case $(F \circ j)(0) = (0, 0)$. Since $\operatorname{Im}(F \circ j)_{*0} \supseteq V_2 \oplus 0$, it suffices to prove that $B^3\left(\operatorname{Ker}(B^1)\right) = V_3$. The proof of Theorem 19.44 contains a proof that $B^3 : U_1 \oplus U_2 \to V_3$ is onto. Since $U_1 \oplus U_2 = \mathrm{Ker}\left(B^3\right) \oplus \mathrm{Im}\left(B^{3*}\right)$, we know that $B^3\left(\mathrm{Im}\left(B^{3*}\right)\right) = V_3$. Thus, it suffices to show that

$$\mathrm{Im}\left(B^{3*}\right) \subseteq \mathrm{Ker}\left(B^1\right) \quad \text{or} \quad \mathrm{Ker}\left(B^3\right) \supseteq \mathrm{Im}\left(B^{1*}\right).$$

Thus, we need $B^3 \circ B^{1*} = 0$. To this end

$$\begin{aligned} B^3\left(B^{1*}(i\theta)\right) &= B^3\left(id\theta, -\tfrac{i}{2}\theta\psi\right) = \mathcal{D}_c^\omega\left(-\tfrac{i}{2}\theta\psi\right) + \tfrac{i}{2}d\theta\cdot\psi \\ &= -\tfrac{i}{2}\theta\mathcal{D}_c^\omega\psi - \tfrac{i}{2}d\theta\cdot\psi + \tfrac{i}{2}d\theta\cdot\psi = -\tfrac{i}{2}\theta\mathcal{D}_c^\omega\psi = 0, \end{aligned}$$

since we have assumed $(\omega,\psi,\eta) \in \mathcal{SWP}_k$. Finally note that we have the identification

$$\begin{aligned} T_{(\omega,\psi,\eta)}\mathcal{MSWP}_k &\cong \mathrm{Ker}\left((F\circ j)_{*0}\right) \\ &= \left\{(\omega',\psi',\eta') : B^1(\omega',\psi') = 0, B^2(\omega',\psi') - \eta' = 0, B^3(\omega',\psi') = 0\right\}, \end{aligned} \tag{19.66}$$

as required. ∎

We have a smooth projection mapping

$$p : \mathcal{MSWP}_k \to W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right), \quad \text{where} \quad p([\omega,\psi,\eta]) = \eta. \tag{19.67}$$

Let

$$\mathcal{SW}_k(\eta) := p^{-1}(\eta) = \left\{[\omega,\psi,\eta] \in \mathcal{MSWP}_k : \Omega^{\omega+} - q(\psi) - \eta = 0, \mathcal{D}_c^\omega\psi = 0\right\}. \tag{19.68}$$

THEOREM 19.51. *Near a point $(\omega,\psi,\eta) \in \mathcal{SWP}_k$ where $p_{[\omega,\psi,\eta]*}$ is onto (i.e., η is an achieved regular value of p), we have that $\mathcal{SW}_k(\eta) := p^{-1}(\eta)$ is a smooth manifold of dimension*

$$d(X,L) := b^1 - \left(1 + b^{2+}\right) + \tfrac{1}{4}\left(c_1(L)^2 - \mathrm{sig}(X)\right).$$

PROOF. By the Implicit Function Theorem I (Theorem 17.24, p. 470), $\mathcal{SW}_k(\eta)$ is a smooth submanifold of $\mathcal{MSWP}_k$. Moreover, by (19.66) and (19.67),

$$p_{*[\omega,\psi,\eta]}(\omega',\psi',\eta') = \eta' = B^2(\omega',\psi') = (d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi'). \tag{19.69}$$

Thus,

$$\begin{aligned} &T_{[\omega,\psi,\eta]}\left(p^{-1}(\eta)\right) = \ker\left(p_{*[\omega,\psi,\eta]} : T_{[\omega,\psi,\eta]}\mathcal{MSWP}_k \to W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)\right) \\ &\cong \Big\{(\omega',\psi',0) \in T_{(\omega,\psi,\eta)}\mathcal{CWP}_k : (d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') = 0, \\ &\mathcal{D}_\omega\psi' + \tfrac{1}{2}\omega'\cdot\psi = 0,\ \delta\omega' - \tfrac{1}{2}\langle\psi,\psi'\rangle = 0\Big\} \end{aligned}$$

which can be identified with the kernel of the elliptic operator

$$\begin{aligned} B : &W^{2,k+1}\left(\Lambda^1(X,i\mathbb{R})\right) \oplus W^{2,k+1}\left(\Sigma_c^+(X)\right) \\ &\to W^{2,k}(X,i\mathbb{R}) \oplus W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right) \end{aligned}$$

of (19.50). We need to show that if $p_{[\omega,\psi,\eta]*}$ is onto, then B is onto, so that $\mathrm{Index}(B) = \dim\mathrm{Ker}(B)$. Assuming that $p_{[\omega,\psi,\eta]*}$ in (19.69) is onto, the map

$$(\omega',\psi') \mapsto (d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') \tag{19.70}$$

must be onto, even for (ω',ψ') constrained by $B^3(\omega',\psi') := \mathcal{D}_\omega\psi' + \frac{1}{2}\omega'\cdot\psi = 0$ (see (19.66)). Thus, the projection of the image of B to $W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right)$ contains $W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \oplus 0$. We already know from the proof of

Theorem 19.44 that for $F(\omega,\psi,\eta)=0$ and $\psi\neq 0$, the map (where (ψ',ω') is now unrestricted)

$$(\psi',\omega')\mapsto \mathcal{D}_\omega\psi'+\tfrac{1}{2}\omega'\cdot\psi$$

is onto. Thus, the projection of the image of B to $W^{2,k}\left(\Lambda^{2+}(X)\right)\oplus W^{2,k}\left(\Sigma_c^-(X)\right)$ is onto. It remains to show that the image B contains $W^{2,k}(M,i\mathbb{R})\oplus 0\oplus 0$. For an arbitrary function $\theta\in W^{2,k+2}(X,\mathbb{R})$, we have $(\omega,\psi,\eta)\cdot e^{it\theta}=\left(\omega+2itd\theta,e^{-it\theta}\psi,\eta\right)$. Since $F(\omega,\psi,\eta)=0\Rightarrow F\left((\omega,\psi,\eta)\cdot e^{it\theta}\right)=0$, we know that

$$\tfrac{d}{dt}\left(\omega+2itd\theta,e^{-it\theta}\psi,\eta\right)\big|_{t=0}=(2id\theta,-i\theta\psi,0)\in\ker\left(F_{*(\omega,\psi,\eta)}\right).$$

Thus,

$$\begin{aligned}B(2id\theta,-i\theta\psi)&=\begin{bmatrix}\delta(2d\theta)+\frac{i}{2}\left(\langle -i\theta\psi,\psi\rangle-\langle\psi,-i\theta\psi\rangle\right),\\ F_{*(\omega,\psi,\eta)}(2id\theta,-i\theta\psi,0)\end{bmatrix}\\ &=\left(2\delta(d\theta)+\tfrac{1}{2}\left(\langle\theta\psi,\psi\rangle+\langle\psi,\theta\psi\rangle\right),0,0\right)=\left(2\left(\delta d\theta+\tfrac{1}{4}\langle\psi,\psi\rangle\theta\right),0,0\right).\end{aligned}$$

Now $\delta d\theta=-\Delta\theta$, where Δ is the usual Laplace operator on $C^\infty(X,\mathbb{R})$, extended to $W^{2,k+2}(X,\mathbb{R})$. Since the operator $-\Delta+\frac{1}{4}\langle\psi,\psi\rangle$ is elliptic and formally self-adjoint, to show that

$$\theta\mapsto 2i\left(\delta d\theta+\tfrac{1}{4}\langle\psi,\psi\rangle\theta\right)$$

is onto $W^{2,k}(M,i\mathbb{R})$ it suffices to show that the kernel of $-\Delta+\frac{1}{4}\langle\psi,\psi\rangle$ is 0. However, $-\Delta\theta+\frac{1}{4}\langle\psi,\psi\rangle\theta=0$ implies

$$0=\int_X\left(-\Delta\theta+\tfrac{1}{4}\langle\psi,\psi\rangle\theta\right)\theta\,\nu_g=\int_X\|d\theta\|^2+\tfrac{1}{4}\langle\psi,\psi\rangle\theta^2\,\nu_g$$

Thus, $\theta=0$ since $\psi\neq 0$. Hence, B is onto, and its index is the dimension of its kernel. This index was found (see (19.44)) to be $b^1-(1+b^{2+})+\frac{1}{4}\left(c_1(L)^2-\mathrm{sig}(X)\right)$. □

We would like to use the Sard-Smale Theorem (see [**Sm68**]) stated below to conclude that the set of all $\eta\in W^{2,k}\left(\Lambda^{2+}(X)\right)$, for which $p^{-1}(\eta)$ is either void or a manifold of dimension $d(X,L)$, is residual. Recall that a **residual set** is a subset containing the intersection of a countable collection of open, dense subsets. A residual set of a complete metric space is dense by the Baire Category Theorem.

THEOREM 19.52 (Sard-Smale Theorem). *Let $f:B_1\to B_2$ be a C^r Fredholm mapping between Banach manifolds (i.e., f_{*x} is Fredholm for each $x\in B_1$). and let*

$$C:=\{f(x):x\in B_1\ \ \textit{and}\ \ f_{*x}\ \textit{is not onto}\}\tag{19.71}$$

(i.e., C is the set of critical values of f). If B_1 is separable and $r>\{\max(0,\mathit{index}(f_{*x})):x\in B_1\}$ (in particular, if f is C^∞), then $B_2-f(C)$ is residual.

REMARK 19.53. Often $B_2-f(C)$ is called the set of regular values, but if f is not onto, $B_2-f(C)$ will contain points which are not values of f at all (e.g., consider a constant map f).

In order to apply Theorem 19.52 to $p:\mathcal{MSWP}_k\to W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)$, we need

THEOREM 19.54. *For each* $(\omega, \psi, \eta) \in \mathcal{SWP}_k$,

$$p_{*[\omega,\psi,\eta]} : T_{[\omega,\psi,\eta]}\mathcal{MSWP}_k \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right)$$

is Fredholm. Moreover, the index of $p_{*[\omega,\psi,\eta]}$ *is the same as the index of*

$$\begin{aligned} B_{(\omega,\psi,\eta),k} &: W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right) \oplus W^{2,k+1}\left(\Sigma_c^+(X)\right) \\ &\to W^{2,k}(X, i\mathbb{R}) \oplus W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right). \end{aligned}$$

PROOF. We will use some of the facts we have shown about the related operator $B_{(\omega,\psi,\eta),k}$. We use Notation 19.48 (p. 637), namely

$$\begin{aligned} &U_1 := W^{2,k+1}\left(\Lambda^1(X, i\mathbb{R})\right), \quad U_2 := W^{2,k+1}\left(\Sigma_c^+(X)\right) \\ &V_1 := W^{2,k}(X, i\mathbb{R}), \quad V_2 := W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right), \quad V_3 := W^{2,k}\left(\Sigma_c^-(X)\right). \end{aligned}$$

and

$$B = \left(B^1, B^2, B^3\right) : U_1 \oplus U_2 \to V_1 \oplus V_2 \oplus V_3, \text{ where}$$

$$\begin{array}{ll} B^1 : U_1 \oplus U_2 \to V_1 & \text{with } B^1(\omega', \psi') = \delta\omega' - \frac{1}{4}\left(\langle\psi', \psi\rangle - \langle\psi, \psi'\rangle\right) \\ B^2 : U_1 \oplus U_2 \to V_2 & \text{with } B^2(\omega', \psi') = d\omega'^+ - \widetilde{q}(\psi', \psi) - \widetilde{q}(\psi, \psi') \\ B^3 : U_1 \oplus U_2 \to V_3 & \text{with } B^3(\omega', \psi') = \mathcal{D}_c^\omega\psi' + \frac{1}{2}\omega' \cdot \psi. \end{array}$$

From 19.64, we have

$$\begin{aligned} &T_{[\omega,\psi,\eta]}\mathcal{MSWP}_k \\ &\cong \Big\{(\omega', \psi', \eta') \in T_{(\omega,\psi,\eta)}\mathcal{CWP}_k : (d\omega')^+ - \widetilde{q}(\psi', \psi) - \widetilde{q}(\psi, \psi') = \eta', \\ &D_c^\omega\psi' + \tfrac{1}{2}\omega' \cdot \psi = 0,\ \delta\omega' - \tfrac{1}{2}\left(\langle\psi', \psi\rangle - \langle\psi, \psi'\rangle\right) = 0\Big\} \\ &= \left\{\begin{array}{l} (\omega', \psi', \eta') \in T_{(\omega,\psi,\eta)}\mathcal{CWP}_k : \\ B^1(\omega', \psi') = 0,\ B^2(\omega', \psi') = \eta',\ B^3(\omega', \psi') = 0 \end{array}\right\} \\ &\cong \operatorname{Ker}\left(B^1 \oplus B^3\right). \end{aligned}$$

Moreover,

$$p_{*[\omega,\psi,\eta]}(\omega', \psi', \eta') = \eta' = (d\omega')^+ - \widetilde{q}(\psi', \psi) - \widetilde{q}(\psi, \psi'),$$

or under the final isomorphism in 19.66,

$$p_{*[\omega,\psi,\eta]} = B^2|\operatorname{Ker}\left(B^1 \oplus B^3\right) : \operatorname{Ker}\left(B^1 \oplus B^3\right) \to V_2.$$

In order to show that $p_{*[\omega,\psi,\eta]}$ is Fredholm, we need the following

1.) $\operatorname{Ker}\left(p_{*[\omega,\psi,\eta]}\right) \cong \operatorname{Ker}\left(B^2|\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)$ is finite-dimensional
2.) $\operatorname{Im}\left(p_{*[\omega,\psi,\eta]}\right) = B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)$ is closed and of finite codimension in V_2.

Now (1.) holds, since

$$\operatorname{Ker}\left(p_{*[\omega,\psi,\eta]}\right) = \operatorname{Ker}\left(B^2|\ker\left(B^1 \oplus B^3\right)\right) = \operatorname{Ker}\left(B^1 \oplus B^2 \oplus B^3\right) = \operatorname{Ker}(B),$$

and B is elliptic. To prove (2.), we first show that it suffices to prove that

$$B(U_1 \oplus U_2)^\perp \cap \left(V_1 \oplus 0_{V_2} \oplus V_3\right) = 0. \tag{19.72}$$

Note that under the obvious imbedding $V_2 \to 0_{V_1} \oplus V_2 \oplus 0_{V_3}$,

$$B^2\left(\ker\left(B^1 \oplus B^3\right)\right) \cong B(U_1 \oplus U_2) \cap \left(0_{V_1} \oplus V_2 \oplus 0_{V_3}\right).$$

Thus, since $B(U_1 \oplus U_2)$ is closed in $V_1 \oplus V_2 \oplus V_3$, $B^2\left(\ker\left(B^1 \oplus B^3\right)\right)$ is closed in V_2. Assuming (19.72), we have a 1-1 projection

$$\pi : B(U_1 \oplus U_2)^{\perp} \to (V_1 \oplus 0_{V_2} \oplus V_3)^{\perp} = 0_{V_1} \oplus V_2 \oplus 0_{V_3} \cong V_2.$$

Note that from the yet to be proven fact (19.72), we also obtain

$$\begin{aligned} V_1 \oplus V_2 \oplus V_3 &= B(U_1 \oplus U_2) + (V_1 \oplus 0_{V_2} \oplus V_3)^{\perp} \\ &= B(U_1 \oplus U_2) + 0_{V_1} \oplus V_2 \oplus 0_{V_3}. \end{aligned}$$

Thus, $B^1 \oplus B^3 : U_1 \oplus U_2 \to V_1 \oplus V_3$ is onto. We will prove that

$$\operatorname{Im}(\pi) = \pi\left(B(U_1 \oplus U_2)^{\perp}\right) = B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp},$$

from which we will obtain

$$\pi : B(U_1 \oplus U_2)^{\perp} \cong B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp}. \tag{19.73}$$

Then, since B is Fredholm, $\dim B(U_1 \oplus U_2)^{\perp} < \infty$, and so $\dim B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp} < \infty$. Since

$$\begin{aligned} &0_{V_1} \oplus B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right) \oplus 0_{V_3} \\ &= B(U_1 \oplus U_2) \cap (0_{V_1} \oplus V_2 \oplus 0_{V_3}) \subseteq B(U_1 \oplus U_2), \end{aligned}$$

we have $\pi\left(B(U_1 \oplus U_2)^{\perp}\right) \subseteq B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp}$. To show the reverse inclusion, let $w \in B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp} \subseteq V_2$. For arbitrary $v \in V_1 \oplus V_3$, we have $v = \left(B^1 \oplus B^3\right)(u)$ for some $u \in U_1 \oplus U_2$, since $B^1 \oplus B^3 : U_1 \oplus U_2 \to V_1 \oplus V_3$ is onto. Define

$$L : V_1 \oplus V_3 \to \mathbb{R} \text{ by } \quad L(v) = -\left\langle B^2(u), w\right\rangle.$$

L is well defined since $w \in B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp}$. Thus, there is some $w' \in V_1 \oplus V_3$ such that

$$\left\langle w', \left(B^1 \oplus B^3\right)(u)\right\rangle = \langle w', v\rangle = L(v) = -\left\langle B^2(u), w\right\rangle$$

for all $u \in u \in U_1 \oplus U_2$. Then

$$\langle w' + w, B(u)\rangle = \left\langle w', \left(B^1 \oplus B^3\right)(u)\right\rangle + \left\langle B^2(u), w\right\rangle = 0,$$

so that $w' + w \in B(U_1 \oplus U_2)^{\perp}$. Since $w' \in V_1 \oplus V_3$ and $w \in B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp}$, we have $\pi(w' + w) = w$. Thus, $B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp} \subseteq \pi\left(B(U_1 \oplus U_2)^{\perp}\right)$, and we have the isomorphism 19.73. Thus, from the yet to be proven (19.72), we obtain

$$\begin{aligned} &\operatorname{Index}\left(p_{*[\omega,\psi,\eta]}\right) = \dim \operatorname{Ker} p_{*[\omega,\psi,\eta]} - \dim \operatorname{Coker} p_{*[\omega,\psi,\eta]} \\ &= \dim \operatorname{Ker} B - \dim\left(B^2\left(\operatorname{Ker}\left(B^1 \oplus B^3\right)\right)^{\perp}\right) = \dim \operatorname{Ker} B - \dim\left(B(U_1 \oplus U_2)^{\perp}\right) \\ &= \dim \operatorname{Ker} B - \dim \operatorname{Coker} B = \operatorname{Index} B. \end{aligned}$$

Hence, it only remains to show $B(U_1 \oplus U_2)^{\perp} \cap (V_1 \oplus 0_{V_2} \oplus V_3) = 0$. Suppose that $(f, 0, \xi) \in V_1 \oplus 0_{V_2} \oplus V_3$ and for all $(\omega', \psi') \in U_1 \oplus U_2$, we have (where the zero-th order operator T_ψ and its adjoint T_ψ^* were defined near 19.52)

$$\begin{aligned} 0 &= \langle B(\omega',\psi'),(f,0,\xi)\rangle = \left\langle B^1(\omega',\psi'),f\right\rangle + \left\langle B^3(\omega',\psi'),\xi\right\rangle \\ &= \left\langle (\omega',\psi'),\left(B^1\right)^* f\right\rangle + \left\langle (\omega',\psi'),\left(B^3\right)^*\xi\right\rangle \\ &= \left\langle (\omega',\psi'),\left(i\,df,-\tfrac{i}{2}f\psi\right)\right\rangle + \left\langle (\omega',\psi'),\left(T_\psi^*\xi, D_c^\omega\xi\right)\right\rangle \\ &= \left\langle \omega', i\,df + T_\psi^*\xi\right\rangle + \left\langle \psi', D_c^\omega\xi - \tfrac{i}{2}f\psi\right\rangle. \end{aligned} \tag{19.74}$$

Then

$$i\,df + T_\psi^*\xi = 0 \quad\text{and}\quad D_c^\omega\xi - \tfrac{i}{2}f\psi = 0. \tag{19.75}$$

Since $D_c^\omega\psi = 0$,

$$\begin{aligned} 0 &= D_c^\omega D_c^\omega\xi - \tfrac{i}{2}D_c^\omega(f\psi) = D_c^\omega D_c^\omega\xi - \tfrac{i}{2}df\cdot\psi - \tfrac{i}{2}fD_c^\omega\psi \\ &= D_c^\omega D_c^\omega\xi + \tfrac{1}{2}T_\psi^*\xi\cdot\psi. \end{aligned}$$

Thus,

$$\begin{aligned} 0 &= \int_X \left\langle D_c^\omega D_c^\omega\xi + \tfrac{1}{2}T_\psi^*\xi\cdot\psi,\xi\right\rangle v = \int_X \langle D_c^\omega D_c^\omega\xi,\xi\rangle + \tfrac{1}{2}\left\langle T_\psi\left(T_\psi^*\xi\right),\xi\right\rangle v \\ &= \int_X \langle D_c^\omega\xi, D_c^\omega\xi\rangle + \tfrac{1}{2}\left\langle T_\psi^*\xi, T_\psi^*\xi\right\rangle v \\ &\Rightarrow 0 = T_\psi^*\xi = -idf \quad\text{and}\quad \tfrac{i}{2}f\psi = D_c^\omega\xi = 0. \end{aligned}$$

Hence, f is constant, and $\frac{i}{2}f\psi = 0$ then implies $f = 0$, since $\psi \neq 0$. Also, at any point p where $\psi(p) \neq 0$, the map $\Lambda_p^1(X,i\mathbb{R}) \to W_p^-(X)$ given by $\alpha \mapsto \alpha\cdot\psi(p)$ is 1-1 since $\alpha\cdot(\alpha\cdot\psi(p)) = \psi(p) \neq 0$, and it is onto since $\dim_{\mathbb{R}}\Lambda_p^1(X,i\mathbb{R}) = \dim_{\mathbb{R}} W_p^-(X) = 4$. Hence T_ψ and T_ψ^* are isomorphisms at any point where $\psi(p) \neq 0$. Hence $0 = T_\psi^*\xi \Rightarrow \xi\psi = 0$, but then $\xi = 0$ on the nonvoid open set where $\psi \neq 0$. Since $D_c^\omega\xi = 0$, unique continuation (Theorem 19.43) then yields $\xi = 0$. Thus, $(f,0,\xi) = (0,0,0)$ and we have (19.72), as required. □

Theorem 19.55. *For a η in a residual subset of $W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)$, either $p^{-1}(\eta)$ (i.e., $\mathcal{SW}_k(\eta)$) is empty or $p^{-1}(\eta)$ is a submanifold of $\mathcal{MSWP}_k$ of dimension $d(X,L) = b^1 - \left(1+b^{2+}\right) + \frac{1}{4}\left(c_1(L)^2 - \operatorname{sig}(X)\right)$.*

Proof. According to Theorem 19.54, $p : \mathcal{MSWP}_k \to W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)$ has a Fredholm derivative throughout $\mathcal{MSWP}_k$. Thus, by Theorem 19.52, $W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) - p(C)$ is residual, where C is the critical set of p. For any point $\eta \in W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) - p(C)$, we have that p_* is onto at each point in $p^{-1}(\eta)$. If $p^{-1}(\eta)$ is not empty, then $p^{-1}(\eta)$ is a submanifold of $\mathcal{MSWP}_k$ of dimension $d(X,L)$ by Theorem 19.51. □

4. Compactness of Moduli Spaces and the Definition of S-W Invariants

Here we establish that moduli spaces of solutions of generically perturbed S-W equations are compact, and proceed to show that the S-W invariants are defined. First we examine the consequences of the S-W equations in conjunction with the "Spin$_c$ Bochner-Weitzenbock" formula (19.28), p. 626.

PROPOSITION 19.56. *If* $(\omega, \psi) \in \mathcal{C}\left(P_{\mathrm{U}(1)}\right) \times C^\infty\left(\Sigma_c^+(X)\right)$ *is a solution of the perturbed S-W equations* $\mathcal{D}_c^\omega \psi = 0$, $\Omega^{\omega+} = q(\psi) + \eta$, *for some* $\eta \in \Omega^{2+}(X, i\mathbb{R})$, *then we have*

$$\begin{aligned} 0 = \left\langle \mathcal{D}_c^2 \psi, \psi \right\rangle &= -\left\langle \Delta\psi, \psi\right\rangle + \tfrac{1}{2}\left\langle \mathfrak{R}^\omega \psi, \psi\right\rangle + \tfrac{1}{4} S\left|\psi\right|^2 \\ &= -\left\langle \Delta\psi, \psi\right\rangle + \tfrac{1}{2}\left|\psi\right|^4 + \tfrac{1}{2}\left\langle \eta, q(\psi)\right\rangle + \tfrac{1}{4} S\left|\psi\right|^2. \end{aligned} \tag{19.76}$$

at each point of X. *Integrating this, yields*

$$\int_X \left|\nabla\psi\right|^2 + \tfrac{1}{2}\left|\psi\right|^4 + \tfrac{1}{2}\left\langle \eta, q(\psi)\right\rangle + \tfrac{1}{4} S\left|\psi\right|^2 \,\nu_g = 0. \tag{19.77}$$

PROOF. Using (19.31) on p. 19.31, $\Omega^{\omega+} = q(\psi) + \eta$, and Exercise 19.39 (p. 628), we have

$$\begin{aligned} \left\langle \mathfrak{R}^\omega \psi, \psi\right\rangle = \left\langle \Omega^{\omega+}, q(\psi)\right\rangle &= \left\langle q(\psi) + \eta, q(\psi)\right\rangle \\ &= \left|q(\psi)\right|^2 + \left\langle \eta, q(\psi)\right\rangle = \left|\psi\right|^4 + \left\langle \eta, q(\psi)\right\rangle, \end{aligned}$$

from which (19.76) follows from $\mathcal{D}_c^\omega \psi = 0$ and Proposition 19.36, p. 626. Note that (19.77) follows upon integration, since $\Delta := -\nabla^*\nabla$. □

PROPOSITION 19.57. *For* (ω, ψ, η) *as in Proposition 19.56, we have*

$$\left|\psi(x)\right|^2 \le \max_{y\in X}\left(\sqrt{2}\left|\eta(y)\right| - \tfrac{1}{2} S(y), 0\right). \tag{19.78}$$

Hence, there is an *a priori* bound on $\|\psi\|_{C^0}$. In particular, if $\eta = 0$ and $S \ge 0$, then we must have $\psi = 0$ and $\Omega^{\omega+} = q(\psi) + \eta = 0$.

PROOF. Let $E_1, \dots, E_4$ be locally-defined orthonormal frame field, parallel at a point $x \in X$ (i.e., $\nabla_{E_i} E_j = 0$ at x). Then

$$\begin{aligned} &\Delta\left(\left|\psi\right|^2\right) = \sum\nolimits_{i=1}^4 E_i\left[E_i\left[\left\langle \psi, \psi\right\rangle\right]\right] \\ &= \sum_{i=1}^4 \left\langle \nabla_{E_i}\nabla_{E_i}\psi(x), \psi(x)\right\rangle + 2\left\langle \nabla_{E_i}\psi(x), \nabla_{E_i}\psi(x)\right\rangle + \left\langle \psi(x), \nabla_{E_i}\nabla_{E_i}\psi(x)\right\rangle \\ &= \left\langle -\nabla^*\nabla\psi(x), \psi(x)\right\rangle + 2\left|\nabla\psi(x)\right|^2 + \left\langle \psi(x), -\nabla^*\nabla\psi(x)\right\rangle. \end{aligned}$$

Hence,

$$\Delta\left(\left|\psi\right|^2\right) = 2\left|\nabla\psi\right|^2 - 2\,\mathrm{Re}\left\langle \nabla^*\nabla\psi, \psi\right\rangle.$$

If $\left|\psi\right|^2$ achieves its maximum at x',then $\Delta\left(\left|\psi\right|^2\right)(x') \le 0$, and so at x'

$$-2\,\mathrm{Re}\left\langle \Delta\psi, \psi\right\rangle = 2\,\mathrm{Re}\left\langle \nabla^*\nabla\psi, \psi\right\rangle = 2\left|\nabla\psi\right|^2 - \Delta\left(\left|\psi\right|^2\right) \ge 0.$$

From (19.76), we have

$$\tfrac{1}{2}|\psi|^4 + \tfrac{1}{2}\operatorname{Re}\langle \eta, q(\psi)\rangle + \tfrac{1}{4}S|\psi|^2 = \operatorname{Re}\langle \Delta\psi, \psi\rangle.$$

$$\operatorname{Re}\langle \eta, q(\psi)\rangle \le |\eta|\,|q(\psi)| \le \sqrt{2}\,|\eta|\,|\psi|^2,$$

using $\frac{1}{2}|q(\psi)|^2 = |\psi|^4$ from Exercise 19.39 b. Thus, at x',

$$\begin{aligned}\tfrac{1}{2}\left(|\psi|^2 - \sqrt{2}\,|\eta| + \tfrac{1}{2}S\right)|\psi|^2 &= \tfrac{1}{2}|\psi|^4 - \tfrac{1}{2}\sqrt{2}\,|\eta|\,|\psi|^2 + \tfrac{1}{4}S|\psi|^2\\ &\le \tfrac{1}{2}|\psi|^4 + \tfrac{1}{2}\operatorname{Re}\langle \eta, q(\psi)\rangle + \tfrac{1}{4}S|\psi|^2 = \operatorname{Re}\langle \Delta\psi,\psi\rangle \le 0\end{aligned}$$

Hence, either $\psi(x') = 0$ (and so $\psi = 0$), or for all $x \in X$,

$$|\psi(x)|^2 \le |\psi(x')|^2 \le \left(\sqrt{2}\,|\eta(x')| - \tfrac{1}{2}S(x')\right) \le \max_{y\in X}\left(\sqrt{2}\,|\eta(y)| - \tfrac{1}{2}S(y)\right).$$

□

COROLLARY 19.58. *For (ω, ψ, η) as in Proposition 19.56, we have*

$$\begin{aligned}\left|\Omega^{\omega+}\right| &= |q(\psi) + \eta| \le \sqrt{2}\,|\psi|^2 + |\eta|\\ &\le \max_{y\in X}\left(2\,|\eta(y)| - \tfrac{1}{2}\sqrt{2}S(y), 0\right) + \max_{y\in X}(|\eta(y)|).\end{aligned} \tag{19.79}$$

For a fixed oriented, Riemannian manifold (X, g) and form η, there is an upper bound on the number of Spin^c-structures for which the S-W moduli space has a nonnegative formal dimension.

PROOF. Using $\frac{1}{2}|q(\psi)|^2 = |\psi|^4$ from Exercise 19.39 b and Proposition 19.57, (19.79) is evident. For a Spin^c stucture $P_{\mathrm{Spin}^c} \to P_{\mathrm{U}(1)\times\mathrm{SO}(4)} \to X$, let L be the complex line bundle $P_{\mathrm{U}(1)} \times_{\mathrm{U}(1)} \mathbb{C}$. Recall that the formal dimension of moduli space is

$$\tfrac{1}{4}\left(c_1(L)^2[X] - 2\chi(X) - 3\tau(X)\right) \tag{19.80}$$

where $c_1(L)^2[X]$ is the evaluation of the the cup-square of the first Chern class $c_1(L)$ on $[X]$. If this is assumed to be nonnegative, then

$$2\chi(X) + 3\tau(X) \le c_1(L)^2[X]. \tag{19.81}$$

For any connection ω on $P_{\mathrm{U}(1)}$ with curvature form $\Omega^\omega \in \Omega^2(X, i\mathbb{R})$, we have

$$\begin{aligned}c_1(L)^2[X] &= \int_X \tfrac{i}{2\pi}\Omega^\omega \wedge \tfrac{i}{2\pi}\Omega^\omega = \tfrac{-1}{4\pi^2}\int_X \left(\Omega^{\omega+} + \Omega^{\omega-}\right)\wedge\left(\Omega^{\omega+} + \Omega^{\omega-}\right)\\ &= \tfrac{-1}{4\pi^2}\int_X \left(\Omega^{\omega+} + \Omega^{\omega-}\right)\wedge\left(*\Omega^{\omega+} - *\Omega^{\omega-}\right)\\ &= \tfrac{-1}{4\pi^2}\int_X \Omega^{\omega+}\wedge *\Omega^{\omega+} - \Omega^{\omega+}\wedge *\Omega^{\omega-} + \Omega^{\omega-}\wedge *\Omega^{\omega+} - \Omega^{\omega-}\wedge *\Omega^{\omega-}\\ &= \tfrac{1}{4\pi^2}\int_X \left(\left|\Omega^{\omega+}\right|^2 - \left\langle\Omega^{\omega+}, \Omega^{\omega-}\right\rangle + \left\langle\Omega^{\omega-}, \Omega^{\omega+}\right\rangle - \left|\Omega^{\omega-}\right|^2\right)\ \nu_g\\ &= \tfrac{1}{4\pi^2}\left(\left\|\Omega^{\omega+}\right\|^2 - \left\|\Omega^{\omega-}\right\|^2\right).\end{aligned} \tag{19.82}$$

By (19.79), we have an upper bound on $\|\Omega^{\omega+}\|^2$. Then (19.82) together with the lower bound (19.81) on $c_1(L)[X]$ gives us an upper bound on $\left\|F_A^-\right\|^2$. Hence, there is an upper bound on $\|\Omega^{\omega}\|^2$. By Hodge theory (see Theorem 18.58, p. 571), the closed form $\frac{i}{2\pi}\Omega^{\omega}$ has a unique representative, say β^{ω}, in the lattice of integral harmonic 2-forms. Since $\|\beta^{\omega}\| \leq \left\|\frac{i}{2\pi}\Omega^{\omega}\right\|$, we have an upper bound on $\|\beta^{\omega}\|$ and there are only a finite number of such β^{ω} in the lattice within a ball. Thus, there are only finitely many possibilities for the canonical class $c_1(L) = \left[\frac{i}{2\pi}\Omega^{\omega}\right]$. For each of these possibilities, there are only $\#H^1(X, \mathbb{Z}_2) < \infty$ distinct Spin^c structures (see 19.33, 623). □

LEMMA 19.59. *Let ω_0 be a C^{∞} connection on $P_{\mathrm{U}(1)} \to X$. For any $k \geq 0$, let $\omega \in \mathcal{C}^{2,k+1}(P_{\mathrm{U}(1)})$. Then there is $s \in W^{2,k+2}(X, \mathrm{U}(1))$, such that $s \cdot \omega := \omega_0 + \alpha$, where $\alpha \in W^{2,k+1}(\Lambda^1(X))$, $\delta\alpha = 0$, and*

$$\|\alpha\|_{2,k+1}^2 \leq C\left\|\Omega^{\omega+}\right\|_{2,k}^2 + K, \tag{19.83}$$

where C and K are independent of α.

PROOF. For some $\alpha_0 \in W^{2,k+1}(\Lambda^1(X))$ we have $\omega = \omega_0 + \alpha_0$. For $e^{i\theta} \in W^{2,k+2}(X, \mathrm{U}(1))$, $e^{i\theta} \cdot \omega = \omega_0 + \alpha_0 - 2id\theta$. Thus, we first find $\theta \in W^{2,k+2}(X, \mathbb{R})$, such that $\delta(\alpha_0 - 2id\theta) = 0$ or

$$\delta d\theta = -\tfrac{i}{2}\delta\alpha_0. \tag{19.84}$$

Since $\delta d = -\Delta$ is a formally self-adjoint elliptic operator, this can be solved for θ as long as $\delta\alpha_0$ is L^2 orthogonal to $\ker(\Delta)$ which consists of the constant functions. However, $\langle\delta\alpha_0, c\rangle_{2,0} = \langle\alpha_0, dc\rangle_{2,0} = 0$, for any constant function c. Hence, we can can solve for θ and take $\alpha = \alpha_0 + 2id\theta$ to obtain $\delta\alpha = 0$. However, further modifications are necessary to produce α satisfying 19.83. Thus, let $\alpha_1 = \alpha_0 + 2id\theta$. Now,

$$\Omega^{\omega} = \Omega^{e^{i\theta}\cdot\omega} = \Omega^{\omega_0+\alpha_1} = \Omega^{\omega_0} + d\alpha_1. \tag{19.85}$$

By the Hodge Decomposition Theorem (Theorem 18.57, p. 571), we can uniquely write $\alpha_1 = h + \beta$, where h is harmonic (i.e., $dh = 0$ and $\delta h = 0$) and $\beta \in W^{2,k}(\Lambda^1(X))$ is orthogonal to the subspace of harmonic forms. Note that $0 = \delta\alpha_1 = \delta(h + \beta) = \delta\beta$. We have

$$\Omega^{\omega+} = \Omega^{\omega_0+} + d\alpha_1^+ = \Omega^{\omega_0+} + dh^+ + d\beta^+ = \Omega^{\omega_0+} + d\beta^+. \tag{19.86}$$

The operator

$$\delta \oplus d^+ : \Omega^1(X) \to \Omega^0(X) \oplus \Omega^{2+}(X) \tag{19.87}$$

is elliptic and $\beta \perp \operatorname{Ker}(\delta \oplus d) = \operatorname{Ker}(\delta \oplus d^+)$, since

$$\begin{aligned}
(d\gamma)^+ = 0 \Rightarrow 0 &= \delta((1 + *)d\gamma) = \delta d\gamma + \delta * d\gamma \\
&= \delta d\gamma - *d**d\gamma = \delta d\gamma - *d^2\gamma = \delta d\gamma \\
&\Rightarrow 0 = (\delta d\gamma, \gamma) = (d\gamma, d\gamma) \Rightarrow d\gamma = 0.
\end{aligned}$$

Thus, there is a constant C (independent of β) such that

$$\begin{aligned}
\|\beta\|_{2,k+1}^2 &\leq C\left(\|\delta\beta\|_{2,k}^2 + \left\|d\beta^+\right\|_{2,k}^2\right) = C\left\|d\beta^+\right\|_{2,k}^2 = C\left\|\Omega^{\omega+} - \Omega^{\omega_0+}\right\|_{2,k}^2 \\
&\leq C\left\|\Omega^{\omega+}\right\|_{2,k}^2 + C\left\|\Omega^{\omega_0+}\right\|_{2,k}^2 = C\left\|F_{\omega}^+\right\|_{2,k}^2 + K'
\end{aligned}$$

However, $\alpha_1 = h+\beta$ and we cannot deduce that $\|h\|_{2,k+1}^2 \le K''$, and hence we need a further gauge transformation. The group $H^1(X,\mathbb{Z}) \subset H^1(X,\mathbb{R})$ can be regarded as a lattice in the b_1-dimensional vector space of harmonic 1-forms. For a harmonic 1-form $\xi \in H^1(X,\mathbb{Z})$, we have a well-defined function $s_0 \in C^\infty(X, \mathrm{U}(1))$, given by $s_0(x) := \exp\left(2\pi i \int_\gamma \xi\right)$ where γ is a path joining a fixed x_0 to x in X. We have

$$\left(e^{i\theta}s_0\right)\cdot\omega = s_0\cdot\left(e^{i\theta}\cdot\omega\right) = \omega_0 + \alpha_1 - 2{s_0}^{-1}ds_0 = \omega_0 + \beta + h - 4\pi i\xi$$

If d is the $\|\cdot\|_{2,k+1}^2$-diameter of a fundamental cell of the lattice $H^1(X,\mathbb{Z})$, then we can choose $\xi \in H^1(X,\mathbb{Z})$ so that $\|h - 4\pi i\xi\| \le 2\pi d$. Then for $s := e^{i\theta}s_0$ and $\alpha := \beta + h - 4\pi i\xi$, we have $s\cdot\omega := \omega_0 + \alpha$, where

$$\begin{aligned}\|\alpha\|_{2,k+1}^2 &= \|\beta + h - 4\pi i\xi\|_{2,k+1}^2 \le \|\beta\|_{2,k+1}^2 + \|h - 4\pi i\xi\|_{2,k+1}^2 \\ &\le C\left\|F_\omega^+\right\|_{2,k}^2 + K' + 2\pi d\end{aligned}$$

and C and $K' + 2\pi d$ are independent of α. □

THEOREM 19.60. *For $k \ge 4$, let $(\omega',\psi') \in \mathcal{C}^{2,k+1}(P_{\mathrm{U}(1)}) \times W^{2,k+1}(\Sigma_c^+(X))$ be a solution of*

$$\Omega^{\omega'+} - q(\psi') = \eta \quad \text{and} \quad \mathcal{D}_c^{\omega'}\psi' = 0, \tag{19.88}$$

where $\eta \in i\Omega^{2+}(X)$ is a C^∞ form and an achieved regular value of the projection $p : \mathcal{MSWP}_{k+1} \to W^{2,k}(\Lambda^{2+}(X, i\mathbb{R}))$. Let ω_0 be a fixed choice of C^∞ connection. In accordance with Lemma 19.59, for some $s \in W^{2,k+2}(X,\mathrm{U}(1))$, we have $s\cdot\omega' = \omega_0 + \alpha$, for $\alpha \in W^{2,k+1}(i\Lambda^1(X))$ with $\delta\alpha = 0$. Let $(\omega,\psi) := s\cdot(\omega',\psi')$. We assume that the harmonic component of $\frac{1}{2\pi i}\alpha$ lies in a fixed fundamental domain of the lattice of integral harmonic 1-forms (see the proof of Lemma 19.59). Then there are constants $C(k')$ depending only on g, η, ω_0 and $k' \ge 3$, such that

$$\|\alpha\|_{2,k'} + \|\psi\|_{2,k'} \le C(k'), \tag{19.89}$$

where $\|\psi\|_{2,k'}$ is computed using the Levi-Civita connection for X and the connection ω_0 on $P_{\mathrm{U}(1)}$. In particular, α and ψ are C^∞, as is $\omega = s\cdot\omega' = \omega_0 + \alpha$.

PROOF. Since $W^{2,5}(\Sigma_c^+(X)) \subseteq C^2(\Sigma_c^+(X))$, the argument of Proposition 19.57 applies to yield a uniform bound on $\|\psi\|_{C^0}$. Here and elsewhere, "uniform bound" means a bound which only depends on η and the metric g on X (in particular, not on ψ or ω). By (19.77), we have

$$\int_X |\nabla^\omega\psi|^2 \le \int_X \tfrac{1}{4}|S|\,|\psi|^2 + \tfrac{1}{2}|\psi|^4 + \tfrac{1}{2}\left|\langle\eta, q(\psi)\rangle\right|\,\nu_g. \tag{19.90}$$

However, this does not yet give us a uniform bound on $\|\psi\|_{2,1}$, since we are using ω_0 (not ω) to define $\|\psi\|_{2,1}$. Now

$$\nabla^\omega\psi = \nabla^{\omega_0}\psi + \tfrac{1}{2}\alpha\psi \quad \text{or} \quad \nabla^{\omega_0}\psi = \nabla^\omega\psi - \tfrac{1}{2}\alpha\psi. \tag{19.91}$$

Thus, we need a uniform bound on $\|\alpha\psi\|_{2,0}$ to produce a uniform bound on $\|\psi\|_{2,1}$. For this it suffices to produce a uniform bound on $\|\alpha\|_{2,0}$, since there is a uniform bound on $\|\psi\|_{C^0}$. We can in fact produce a uniform bound on $\|\alpha\|_{2,2}$ as follows. By Lemma 19.59, we have

$$\|\alpha\|_{2,k+1}^2 \le C\left\|\Omega^{\omega+}\right\|_{2,k}^2 + K, \tag{19.92}$$

whence it suffices to obtain a uniform bound on $\|\Omega^{\omega+}\|_{2,1}^2$ to get a uniform bound on $\|\alpha\|_{2,2}^2$. Since

$$\delta + d : \Omega^2(X) \to \Omega^1(X) \oplus \Omega^3(X)$$

has injective symbol, the Sobolev extension

$$(\delta + d)^{2,k+1} : W^{2,k+1}\left(\Lambda^2(X)\right) \to W^{2,k}\left(\Lambda^1(X)\right) \oplus W^{2,k}\left(\Lambda^3(X)\right)$$

has a finite-dimensional kernel $\mathcal{H}^2(X)$ consisting of C^∞ harmonic 2-forms and

$$W^{2,k+1}\left(\Lambda^2(X)\right) = \mathcal{H}^2(X) \oplus \mathcal{H}^2(X)^\perp.$$

Since $(\delta + d)^{2,k+1} |\mathcal{H}^2(X)^\perp$ has a continuous inverse on its image, there is a constant C, such that for all $\beta \in W^{2,k+1}\left(\Lambda^2(X)\right)$,

$$\left\|\beta^\perp\right\|_{2,k+1} \le C \left\|(\delta + d)^{2,k+1}\beta\right\|_{2,k} ,$$

where $\beta^\perp$ is the orthogonal projection of into $\mathcal{H}^2(X)^\perp$. Since $\delta = \pm * d*$, if β is self-dual, then $\delta\beta = \pm * d\beta$, and so $\|\delta\beta\|_{2,k} = \|d\beta\|_{2,k}$. Thus,

$$\left\|\beta^\perp\right\|_{2,k+1} \le \sqrt{2}C \left\|d^{2,k+1}\beta\right\|_{2,k} .$$

Applying this with $\beta = \Omega^{\omega+}$, we get

$$\left\|\left(\Omega^{\omega+}\right)^\perp\right\|_{2,k+1} \le \sqrt{2}C \left\|d^{2,k+1}\Omega^{\omega+}\right\|_{2,k} . \tag{19.93}$$

Since all norms are equivalent on the finite-dimensional kernel $\mathcal{H}^2(X)$, the fact that there is a uniform C^0 bound on $\Omega^{\omega+} = q(\psi) + \eta$ gives us a uniform $W^{2,k}$ bound on the harmonic part of $\Omega^{\omega+}$. To get a uniform bound on $\Omega^{\omega+}$ itself, first note (where θ is the Levi-Civita connection) that

$$\nabla^\theta\Omega^{\omega+} = \widetilde{q}\,(\nabla^\omega\psi, \psi) + \widetilde{q}\,(\psi, \nabla^\omega\psi) + \nabla^\theta\eta. \tag{19.94}$$

While we are in the process of getting a uniform bound on $\|\psi\|_{2,1}$, we already know that $\|\psi\|_{C^0}$ and $\|\nabla^\omega\psi\|_{2,0}$ are uniformly bounded by (19.78) and (19.90). Thus, $\left\|\nabla^\theta\Omega^{\omega+}\right\|_{2,0}$ is uniformly bounded by (19.94). This yields a uniform bound on $\left\|d^{2,k+1}\Omega^{\omega+}\right\|_{2,0}$ in (19.93), and hence on $\left\|\left(\Omega^{\omega+}\right)^\perp\right\|_{2,1}$. Thus, $\|\Omega^{\omega+}\|_{2,1}$ is uniformly bounded. By (19.92), $\|\alpha\|_{2,2}^2$ is then uniformly bounded, and $\|\psi\|_{2,1}$ is uniformly bounded via (19.91). Now we wish to show that $\|\psi\|_{2,3}$ is uniformly bounded. Note that

$$0 = \mathcal{D}_c^\omega\psi = \mathcal{D}_c^{\omega_0}\psi + \tfrac{1}{2}\alpha \cdot \psi \Rightarrow \mathcal{D}_c^{\omega_0}\psi = -\tfrac{1}{2}\alpha \cdot \psi.$$

Since $\|\alpha\|_{2,2}$ and $\|\psi\|_{p_2,0}$ are bounded for each $p_2 \ge 1$, Proposition 17.22 (p. 17.22) implies that $\|\alpha \cdot \psi\|_{p_3,k_3}$ is uniformly bounded, provided $k_3 - \frac{4}{p_3} \le \min\left(2 - \frac{4}{2}, 0 - \frac{4}{p_2}\right) = -\frac{4}{p_2}$. In particular, taking $p_2 \ge 4$, $\|\alpha \cdot \psi\|_{4,0}$ is uniformly bounded, and hence $\|\mathcal{D}_c^{\omega_0}\psi\|_{4,0}$ is uniformly bounded. Since $\mathcal{D}_c^{\omega_0}$ is an elliptic operator, we not only have a continuous map

$$\mathcal{D}_c^{\omega_0} : W^{4,1}(\Sigma_c^+(X)) \to W^{4,0}(\Sigma_c^-(X)),$$

but also by Proposition 17.21 (468),

$$\|\psi\|_{4,1} \le C\left(\|\mathcal{D}_c^{\omega_0}\psi\|_{4,0} + \|\psi\|_{4,0}\right)$$

for some constant C depending only on $\mathcal{D}_c^{\omega_0}$. Thus, there is a uniform bound on $\|\psi\|_{4,1}$. Since $\|\alpha\|_{2,2}$ and $\|\psi\|_{4,1}$ are bounded, Proposition 17.22 (p. 17.22) implies that $\|\alpha\cdot\psi\|_{p_3,k_3}$ is uniformly bounded, provided $k_3 - \frac{4}{p_3} < \min\left(2-\frac{4}{2}, 1-\frac{4}{4}\right) = 0$. Thus, with $k_3 = 1$ and $p_3 = 3$, $\|\mathcal{D}_c^{\omega_0}\psi\|_{3,1} = \|\alpha\cdot\psi\|_{3,1}$ is uniformly bounded. Using

$$\|\psi\|_{3,2} \leq C\left(\|\mathcal{D}_c^{\omega_0}\psi\|_{3,1} + \|\psi\|_{3,1}\right),$$

we deduce that $\|\psi\|_{3,2}$ is uniformly bounded. Then Proposition 17.22 (p. 17.22) implies that $\|\mathcal{D}_c^{\omega_0}\psi\|_{2,2} = \|\alpha\cdot\psi\|_{2,2}$ is uniformly bounded, since $\|\alpha\|_{2,2}$ and $\|\psi\|_{3,2}$ are uniformly bounded (note that $\max\left(2-\frac{4}{2}, 2-\frac{4}{3}\right) = \frac{2}{3} > 0$). Using

$$\|\psi\|_{2,3} \leq C\left(\left\|\left(\mathcal{D}_c^{\omega_0}\right)^{2,3}\psi\right\|_{2,2} + \|\psi\|_{2,2}\right),$$

$\|\psi\|_{2,3}$ is uniformly bounded. Moreover, from $\Omega_\omega^+ = q(\psi) + \eta$, we get that $\|\Omega_\omega^+\|_{3,2}$ is uniformly bounded since Proposition 17.22 (p. 17.22) implies that $\|q(\psi)\|_{3,2}$ is bounded (here $k - \frac{n}{p} = 2 - \frac{4}{3} > 0$). By (19.83), we then have a uniform bound on $\|\alpha\|_{2,4}$.

In summary, thus far we have shown $\|\alpha\|_{2,k'} + \|\psi\|_{2,k'} \leq C(k')$ for $k' = 2$ and $k' = 3$. Assume that we have this result for some $k' \geq 3$. By Proposition 17.22 (p. 17.22) $\|\mathcal{D}_c^{\omega_0}\psi\|_{2,k'} = \|\alpha\cdot\psi\|_{2,k'}$ is uniformly bounded since $k' - \frac{4}{2} > 0$. Then,

$$\|\psi\|_{2,k'+1} \leq C\left(\|\mathcal{D}_c^{\omega_0}\psi\|_{2,k'} + \|\psi\|_{2,k'}\right)$$

gives us a uniform bound on $\|\psi\|_{2,k'+1}$. From $\Omega^{\omega+} = q(\psi) + \eta$ and the fact that $\|q(\psi)\|_{2,k'}$ is uniformly bounded, we get a uniform bound for $\|\Omega^{\omega+}\|_{2,k'}$. Finally, (19.92) gives us a uniform bound on $\|\alpha\|_{2,k'+1}$. □

Corollary 19.61. *For $k \geq 4$, let $(\omega_n', \psi_n') \in \mathcal{C}^{2,k+1}(P_{\mathrm{U}(1)}) \times W^{2,k+1}(\Sigma_c^+(X))$ be a sequence of solutions of the S-W equations*

$$\Omega^{\omega'+} - q(\psi) = \eta \quad \text{and} \quad \mathcal{D}_c^{\omega'}\psi' = 0, \tag{19.95}$$

where $\eta \in i\Omega^{2+}(X)$ is a C^∞ form and an achieved regular value of the projection $p : \mathcal{MSWP}_{k+1} \to W^{2,k}(i\Lambda^{2+}(X))$. The sequence $(\omega_n, \psi_n) := s_n\cdot(\omega_n', \psi_n')$ of C^∞ solutions of the S-W equations produced in Theorem 19.60 has a subsequence which is convergent in the C^∞ topology to a C^∞ solution (ω, ψ) of the S-W equations 19.95.

Proof. The inclusion of $W^{2,k}$ spaces in the corresponding C^{k-3} spaces is compact for $k \geq 3$. This means that a sequence which is bounded in $W^{2,k}$ has a subsequence convergent in C^{k-3}. Because of the bound 19.89 for $k \geq 3$, we can choose a subsequence of (A_n, ψ_n) converging in C^0. Then we can choose a subsequence of the subsequence converging in C^1. Continuing, we obtain a sequence of subsequences, and the diagonal subsequence converges in C^k for all k, and hence in C^∞ to a C^∞ solution. □

Corollary 19.61 does not imply that the solution (ω, ψ) to which (ω_n, ψ_n) converges is in $p_k^{-1}(\eta)$, since we might have $\psi = 0$, and then $(\omega, 0) \notin \mathcal{CWP}_k$; see (19.45), p. 633). In order to prove that the manifold $p_k^{-1}(\eta)$ is compact, we need to avoid using those $\eta \in W^{2,k}\left(\Lambda^{2+}(X)\right)$ for which there are reducible solutions (i.e., solutions for which $\psi = 0$). If $b_2^+ > 0$, we now show how such η can be avoided. If (ω, ψ) is a reducible solution of the S-W equations, then

$$\Omega^{\omega+} = q(\psi) + \eta = q(0) + \eta = \eta. \tag{19.96}$$

We can always write

$$\Omega^{\omega} = \mathcal{H}(\Omega^{\omega}) + d\alpha$$

where $\mathcal{H}(\Omega^{\omega})$ is the harmonic part of Ω^{ω} and $\alpha \in W^{2,k+1}(\Lambda^1(X, i\mathbb{R}))$. The cohomology class of Ω^{ω} is determined by $P_{\mathrm{U}(1)}$, since $\left[\frac{i}{2\pi}\Omega^{\omega}\right] = c_1(P_{\mathrm{U}(1)})$. Thus, $\mathcal{H}(\Omega^{\omega})$ is uniquely determined by the metric and $P_{\mathrm{U}(1)}$ and is independent of the choice of $\omega \in \mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right)$. For a given metric g and $P_{\mathrm{U}(1)}$, we set $\gamma_g\left(P_{\mathrm{U}(1)}\right) := \mathcal{H}(\Omega^{\omega})$. Thus,

$$(\mathcal{H}(\Omega^{\omega}) + d\alpha)^+ = \mathcal{H}(\Omega^{\omega})^+ + d\alpha^+ = \gamma_g\left(P_{U(1)}\right)^+ + d\alpha^+$$

Let

$$\mathcal{H}^+ : W^{2,k+1}(\Lambda^2(X, i\mathbb{R})) \to \mathcal{H}^{2+}(X, i\mathbb{R})_g$$

denote the L^2-orthogonal projection (given by Hodge theory) onto the self-dual harmonic 2-forms relative to g. Then (19.96) will not hold for any $\omega \in \mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right)$, as long as

$$\mathcal{H}^+(\eta) \neq \gamma_g\left(P_{\mathrm{U}(1)}\right)^+, \tag{19.97}$$

since then $\mathcal{H}^+(\Omega^{\omega}) = \gamma_g\left(P_{\mathrm{U}(1)}\right)^+ \neq \mathcal{H}^+(\eta)$. The affine subspace

$$\mathcal{A}_g\left(P_{\mathrm{U}(1)}\right) := \left\{\eta \in W^{2,k+1}(\Lambda^{2+}(X, i\mathbb{R})) : \mathcal{H}^+(\eta) = \gamma_g\left(P_{U(1)}\right)^+\right\} \tag{19.98}$$

is closed and of codimension b_2^+ in $W^{2,k+1}(\Lambda^{2+}(X, i\mathbb{R}))$.

THEOREM 19.62. *(Compactness of $\mathcal{SW}_k(\eta)$). For any fixed metric g and* Spin^c *structure for X, let $p_k : \mathcal{MSWP}_k \to W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$ be the projection where $k \geq 4$. Assume that $b_2^+ > 0$. For generic $\eta \in W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$ (i.e., for η in a residual subset) and $\mathcal{SW}_k(\eta) := p_k^{-1}(\eta)$, either $\mathcal{SW}_k(\eta) = \phi$ or $\mathcal{SW}_k(\eta)$ is a compact C^∞ manifold of dimension*

$$d(X, L) = b^1 - \left(1 + b^{2+}\right) + \tfrac{1}{4}\left(c_1(L)^2 - \tau(X)\right), \tag{19.99}$$

where L is the canonical line bundle for the Spin^c*-structure. In the case $\eta \in \Omega^{2+}(X, i\mathbb{R}) := C^\infty(\Lambda^{2+}(X, i\mathbb{R}))$ is an achieved regular value of $p_k : \mathcal{MSWP}_k \to W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$, we let $\mathcal{SW}_\infty(\eta)$ be the set of C^∞ solutions of the S-W equations modulo $C^\infty(X, \mathrm{U}(1))$. The natural map $f_k : \mathcal{SW}_\infty(\eta) \to \mathcal{SW}_k(\eta)$ given by $f_k([\omega, \psi]) = [\omega, \psi]_k$ is a bijection.*

PROOF. By Theorem 19.55 (p. 649), the only new issue in the first claim is compactness. By intersecting the residual subset of Theorem 19.55 with the open, dense complement of $\mathcal{A}_g\left(P_{\mathrm{U}(1)}\right)$ (using $b_2^+ > 0$), we obtain residual subset, say $\mathcal{T}$, of $W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$ such for $\eta \in \mathcal{T}$, there are no $\omega \in \mathcal{C}^{2,k+1}\left(P_{\mathrm{U}(1)}\right)$ with $\Omega^{\omega+} = \eta$, and hence no solutions of the perturbed S-W equations with $\psi = 0$. Thus, for $\eta \in \mathcal{T}$, the limit (ω, ψ) in Corollary 19.61 is not reducible and hence it is in $p_k^{-1}(\eta)$ which is then compact. We now prove that $f_k : \mathcal{SW}_\infty(\eta) \to \mathcal{SW}_k(\eta)$ is a bijection. If $[\omega', \psi']_k \in \mathcal{SW}_k(\eta)$, then (by Theorem 19.60), $[\omega', \psi']_k = [\omega, \psi]_k$ for a C^∞ solution (ω, ψ) of the S-W equations perturbed by η. Thus, f_k is onto.

To show that f_k is 1-1, we argue as follows. If $[\omega_1, \psi_1]_k = [\omega_2, \psi_2]_k$ for some C^∞ $\omega_1, \psi_1, \omega_2, \psi_2$, then $(\omega_2, \psi_2) = s \cdot (\omega_1, \psi_1)$ for some $s \in W^{2,k+2}(X, \mathrm{U}(1))$, i.e.

$$2s^{-1}ds = \omega_2 - \omega_1, \text{ and } \psi_2 = s\psi_1. \tag{19.100}$$

Since locally $s = e^{i\theta}$, we have $2id\theta = 2s^{-1}ds = \omega_2 - \omega_1$ is C^∞ and hence θ and s are C^∞. Hence $[\omega_1, \psi_1] = [\omega_2, \psi_2]$ in $\mathcal{SW}_\infty(\eta)$. □

REMARK 19.63. If p_k has an achieved regular value say η_0, then all $\eta \in W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$ which are sufficiently close to η_0 will be achieved regular values (the compactness of $\mathcal{SW}_k(\eta_0)$ is used for this). Since $C^\infty(\Lambda^{2+}(X, i\mathbb{R}))$ is dense in $W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$, there will be C^∞ achieved regular values for which the bound 19.89 applies.

We wish to show how the moduli space $\mathcal{M}_{k+1}(\eta)$ is used to define an integer, known as the S-W-invariant for a given Spin^c structure for X. It turns out that the S-W-invariant is independent of a suitable choice of η and of the Riemannian metric on X, provided that $b_2^+ \geq 2$. We first give the definition and establish the independence later.

Let $\eta \in \Omega^{2+}(X, i\mathbb{R})^+ = C^\infty(\Lambda^{2+}(X, i\mathbb{R}))$ be a regular achieved value of $p_k : \mathcal{MSWP}_k \to W^{2,k}(\Lambda^2(X, i\mathbb{R})^+)$ so that Theorem 19.62 applies, and hence $\mathcal{SW}_\infty(\eta)$ is a compact, smooth manifold of dimension $d(X, L)$. We let $Sol\,(\eta)$ denote the set of all C^∞ solutions of the S-W equations (perturbed by η) so that

$$\mathcal{SW}_\infty(\eta) = Sol\,(\eta)\,/C^\infty(X, \mathrm{U}(1))$$

Choose $x_0 \in X$ and consider the subgroup say $G_0 \subseteq C^\infty(X, \mathrm{U}(1))$ of those s, such that $s(x_0) = 1$. We then have a mapping

$$\pi : Sol\,(\eta)\,/G_0 \to \mathcal{SW}(\eta) := \mathcal{SW}_\infty(\eta) \tag{19.101}$$

with $\pi([(\omega, \psi)]_0) = [(\omega, \psi)]$, where $[(\omega, \psi)]_0$ denotes the G_0-orbit of $(\omega, \psi) \in Sol\,(\eta)$. We claim that (19.101) is a principal U(1)-bundle. For $[(\omega, \psi)]_0 \in \mathcal{SW}_k/G_0$, let U(1) act on $Sol\,(\eta)\,/G_0$ from the right via

$$[(\omega, \psi)]_0 \cdot e^{i\theta_0} = \left[e^{-i\theta_0} \cdot (\omega, \psi)\right]_0 = \left[\left(\omega, e^{-i\theta_0}\psi\right)\right]_0.$$

To show that the action is free, suppose that in $\mathcal{SW}_k/G_0$ we have $\left[\left(\omega, e^{-i\theta_0}\psi\right)\right]_0 = [(\omega, \psi)]_0 \cdot e^{i\theta_0} = [(\omega, \psi)]_0$. Then there is $s \in G_0$, such $\omega = -2s^{-1}ds + \omega$ (i.e., s is constant and hence 1, since $s \in G_0$) and $e^{-i\theta_0}\psi = s^{-1}\psi = \psi$. Thus, since $\psi \neq 0$, $e^{-i\theta_0} = 1$, and hence the U(1) action is free. The fibers of π are the orbits of this U(1)-action. Indeed, if $\pi([(\omega_1, \psi_1)]_0) = \pi([(\omega_2, \psi_2)]_0)$, then $[(\omega_1, \psi_1)] = [(\omega_2, \psi_2)]$. Hence $(\omega_2, \psi_2) = s \cdot (\omega_1, \psi_1)$ for some $s \in C^\infty(X, \mathrm{U}(1))$. We can write

$$s = s\,s(x_0)^{-1}\,s(x_0) = s_0 s(x_0), \text{ where } s_0 := s\,s(x_0)^{-1} \in G_0.$$

Then

$$\begin{aligned}[(\omega_2, \psi_2)]_0 &= [s \cdot (\omega_1, \psi_1)]_0 = [s_0 s(x_0) \cdot (\omega_1, \psi_1)]_0 \\ &= [s_0 \cdot (\omega_1, \psi_1)]_0 \cdot s(x_0)^{-1} = [(\omega_1, \psi_1)]_0 \cdot s(x_0)^{-1}.\end{aligned}$$

Thus, U(1) acts transitively on the fibers of π, and it is clear that the U(1)-action preserves the fibers. The local triviality of 19.101 may be shown ultimately using constructions as in the proofs of Theorem 19.49 (p. 638) and Theorem 19.50 (p. 644), and so (19.101) defines a principal U(1)-bundle. Let $c_1(\eta) \in H^2(\mathcal{SW}(\eta), \mathbb{Z})$ be the first Chern class of (19.101). Note that $d(X,L)$ in (19.99) is even iff $b_1 + b_2^+$ is odd. Once we give $\mathcal{SW}(\eta)$ an orientation, if $d(X,L)$ is even, then the evaluation of the $\frac{1}{2}d(X,L)$-fold cup product $c_1(\eta)^{d(X,L)/2} := c_1(\eta) \smile \overset{d(X,L)/2}{\cdots} \smile c_1(\eta)$ on the fundamental class $[\mathcal{SW}(\eta)] \in H^{d(X,L)/2}(\mathcal{SW}(\eta), \mathbb{Z})$ yields the **Seiberg-Witten invariant**

$$\mathcal{SW}([P_{\mathrm{Spin}^c}]) := c_1(\eta)^{d(X,L)/2}[\mathcal{SW}(\eta)] \in \mathbb{Z}. \tag{19.102}$$

The orientation for $\mathcal{SW}(\eta)$ is defined as follows. The tangent space of $\mathcal{SW}(\eta)$ is (see Theorem 19.54, p. 647)

$$\begin{aligned} &T_{[\omega,\psi,\eta]}\left(p^{-1}(\eta)\right) \\ &= \operatorname{Ker}\left(p_{*[\omega,\psi,\eta]} : T_{[\omega,\psi,\eta]}\mathcal{MSWP}_k \to W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right)\right) \cong \operatorname{Ker} B_{(\omega,\psi,\eta),k} \\ &\cong \left\{ \begin{array}{l} (\omega',\psi',\eta') \in T_{(\omega,\psi,\eta)}\mathcal{CWP}_k : \\ 0 = B(\omega',\psi',\eta') := \left[\begin{array}{c} (d\omega')^+ - \widetilde{q}(\psi',\psi) - \widetilde{q}(\psi,\psi') \\ \mathcal{D}_c^\omega \psi' + \frac{1}{2}\omega' \cdot \psi \\ \delta\omega' - \frac{1}{2}\langle \psi, \psi' \rangle \end{array} \right] \end{array} \right\}, \end{aligned}$$

where

$$\begin{aligned} B = B_{(\omega,\psi,\eta),k} : W^{2,k+1}&\left(\Lambda^1(X,i\mathbb{R})\right) \oplus W^{2,k+1}\left(\Sigma_c^+(X)\right) \\ &\to W^{2,k}(X,i\mathbb{R}) \oplus W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \oplus W^{2,k}\left(\Sigma_c^-(X)\right). \end{aligned} \tag{19.103}$$

Thus, we need some way to orient $\operatorname{Ker}(B)$. Associated with any Fredholm operator $F : V \to W$, between Hilbert spaces, there is the 1-dimensional space

(19.104)

$$\det F := \Lambda^{\mathrm{top}}(\operatorname{Ker} F) \otimes \Lambda^{\mathrm{top}}(\operatorname{Ker} F^*) := \Lambda^{\dim(\ker F)}(\operatorname{Ker} F) \otimes \Lambda^{\dim(\ker F^*)}(\operatorname{Ker} F^*).$$

In the event that F is onto (i.e., $\operatorname{Ker} F^* = \{0\}$), orienting $\det F$ is equivalent to orienting $\operatorname{Ker} F$. Let

$$B(t)(\omega',\psi',\eta') := \left[\begin{array}{c} \delta\omega' - \frac{1-t}{2}\langle\psi,\psi'\rangle \\ (d\omega')^+ - (1-t)\left(\widetilde{q}(\psi',\psi) + \widetilde{q}(\psi,\psi')\right) \\ \mathcal{D}_c^\omega\psi' + \frac{1-t}{2}\omega'\cdot\psi = 0 \end{array} \right]. \tag{19.105}$$

This is a family of elliptic operators over $[0,1]$ and it can be shown that $\{\det B(t) : t \in [0,1]\}$ is a continuous line bundle over $[0,1]$. As $[0,1]$ is contractible, this line bundle is trivial and an orientation for $\det B(1)$ determines an orientation for $\det B(0)$ (i.e., for $\ker B$, since B is onto, by the assumption that η is an achieved regular value of p; see the proof of Theorem 19.51, p. 645). Now

$$B(1) = \left(\delta, d^+, \mathcal{D}_c^\omega\right),$$

and we have (see the computations following (19.43), p. 632)

$$\begin{aligned} \operatorname{Ker} B(1) &= \mathcal{H}^1(X,i\mathbb{R}) \oplus \operatorname{Ker}\mathcal{D}_c^\omega \ \text{ and} \\ \operatorname{Ker} B(1)^* &= i\mathbb{R} \oplus \mathcal{H}^{2+}(X,i\mathbb{R}) \oplus \operatorname{Ker}\left(\mathcal{D}_c^{\omega *}\right). \end{aligned}$$

Thus,

$$\det B(1) = \Lambda^{\mathrm{top}}\left(\mathcal{H}^1(X,i\mathbb{R}) \oplus \ker\mathcal{D}_c^\omega\right) \otimes \Lambda^{\mathrm{top}}\left(i\mathbb{R} \oplus \mathcal{H}^{2+}(X,i\mathbb{R}) \oplus \operatorname{Ker}\left(\mathcal{D}_c^{\omega*}\right)\right)$$

As $\ker\mathcal{D}_c^\omega$ and $\ker\mathcal{D}_c^{\omega*}$ are complex spaces, they have a natural orientation. Moreover, $i\mathbb{R}$ has the natural orientation i. Thus, we need only to choose a fixed orientation for $H^1(X,i\mathbb{R}) \cong \mathcal{H}^1(X,i\mathbb{R})$ and $H^{2+}(X,i\mathbb{R}) \cong \mathcal{H}^{2+}(X,i\mathbb{R})$ to determine an orientation for det $B(1)$. We assume that this has been done. Hence, the orientation class $[\mathcal{SW}(\eta)]$ in (19.102) is determined.

It remains to show that $\mathcal{SW}(P_{\mathrm{Spin}^c})$ is independent of the metric g on X and the perturbing form η. Let $FX(g)$ denote the bundle of oriented orthonormal frames for to the metric g. We first point out that when the metric is on X is changed from g_1 to g_2, the bundle of oriented orthonormal frames on X changes from $FX(g_1)$ to $FX(g_2)$. These are seen to be isomorphic as follows. Let $\mathrm{Aut}(TX) \subset \mathrm{End}(TX) = T^{1,1}X$ be the bundle of invertible linear transformations of the tangent spaces of X. First we show that there is a *canonical* automorphism $\varphi \in C^\infty(\mathrm{Aut}(TX))$ of the tangent bundle, such that

$$g_1(Y, Z) = (g_2 \cdot \varphi)(Y, Z) := g_2(\varphi(Y), \varphi(Z))$$

for all $Y, Z \in T_pX$ and $p \in X$. More precisely, we prove

PROPOSITION 19.64. *For Riemannian metrics g_1 and g_2 on X, there is a unique positive, g_1-symmetric $\varphi \in C^\infty(\mathrm{Aut}(TX))$ such that*

$$g_1 = g_2 \cdot \varphi \quad \text{and} \quad g_1(\varphi(Y), Z) = g_1(Y, \varphi(Z))$$

PROOF. At each point $p \in X$, $\alpha_p \in \mathrm{GL}(T_pX)$ exists, such that $g_2 \cdot \alpha_p = g_1$ at p. Just let α_p take g_1-orthonormal basis of T_pX to a g_2-orthonormal basis of T_pX. If $\beta_p \in \mathrm{GL}(T_pX)$ also satisfies $g_2 \cdot \beta_p = g_1$, then $\beta_p = \alpha_p \circ A$ for some orthogonal (relative to g_1) $A \in \mathrm{O}_{g_1}(T_pX)$. Indeed,

$$g_1 \cdot A = g_1 \cdot \left(\alpha_p^{-1} \circ \beta_p\right) = \left(g_1 \cdot \alpha_p^{-1}\right) \cdot \beta_p = g_2 \cdot \beta_p = g_1.$$

Using T to denote transpose relative to g_1, we have

$$\beta_p \circ \beta_p^T = (\alpha_p \circ A) \circ (\alpha_p \circ A)^T = \alpha_p \circ A \circ A^T \circ \alpha_p^T = \alpha_p \circ \alpha_p^T.$$

Thus, the positive g_1-symmetric map $\alpha_p \circ \alpha_p^T$ is independent of the choice of α_p, and it has a unique, g_1-symmetric, positive square-root $\sqrt{\alpha_p \circ \alpha_p^T}$. We take $\varphi_p \in \mathrm{Aut}(T_pX)$ to be $\sqrt{\alpha_p \circ \alpha_p^T}$, and $\varphi \in C^\infty(\mathrm{Aut}(TX))$ is defined. $\square$

NOTATION 19.65. Let $S^+_{g_1}X \to X$ be the bundle of positive g_1-symmetric operators.

Thus, we have a canonical $\varphi \in C^\infty\left(S^+_{g_1}X\right) \subseteq C^\infty(\mathrm{Aut}(TX))$ satisfying $g_2 \cdot \varphi = g_1$. We define

$$L_\varphi : FX(g_1) \to FX(g_2)$$

by $L_\varphi(u) := \varphi_p \circ u$ where $FX(g_1) \ni u : \mathbb{R}^4 \to T_pX$ is an isometry (i.e., an oriented orthonormal frame relative to g_1). Note that L_φ is equivariant ($L_\varphi(u \circ A) = \varphi \circ u \circ A = L_\varphi(u) \circ A$ for $A \in \mathrm{SO}(4)$). Consequently, the Levi-Civita connection for $FX(g_2)$ pulls back to a connection (but not necessarily the Levi-Civita connection) for $FX(g_1)$. We can extend L_φ to

$$I \times L_\varphi : P_{\mathrm{U}(1)} \times FX(g_1) \to P_{\mathrm{U}(1)} \times FX(g_2).$$

Then pull back by $I \times L_\varphi^{-1}$ of Spin^c-bundles $\pi_{r^c}(g_1) : P_{\mathrm{Spin}^c}(g_1) \to P_{\mathrm{U}(1)} \times FX(g_1)$ gives us a canonical one-to-one correspondence between the Spin^c structures for (X, g_1) and those for (X, g_2) :

$$\begin{array}{ccc} P_{\mathrm{Spin}^c}(g_1) & \overset{J_\varphi}{\to} & P_{\mathrm{Spin}^c}(g_2) := \left(I \times L_\varphi^{-1}\right)^! P_{\mathrm{Spin}^c}(g_1) \\ \downarrow \pi_{r^c}(g_1) & & \downarrow \pi_{r^c}(g_2) \\ P_{\mathrm{U}(1)} \times FX(g_1) & \overset{I \times L_\varphi}{\to} & P_{\mathrm{U}(1)} \times FX(g_2), \text{ where} \end{array} \tag{19.106}$$

$$\begin{aligned} &\left(I \times L_\varphi^{-1}\right)^! P_{\mathrm{Spin}^c}(g_1) \\ &:= \left\{ \left((p,u), \widetilde{p}\right) \in \left(P_{\mathrm{U}(1)} \times FX(g_2)\right) \times P_{\mathrm{Spin}^c}(g_1) : \pi_{r^c}(g_1)(\widetilde{p}) = \left(p, L_\varphi^{-1} u\right) \right\}. \end{aligned}$$

That is, the fiber of $P_{\mathrm{Spin}^c}(g_2)$ over $(p,u) \in P_{\mathrm{U}(1)} \times FX(g_2)$ is given by

$$\left(\left(I \times L_\varphi^{-1}\right)^! P_{\mathrm{Spin}^c}(g_1)\right)_{(p,u)} = \{(p,u)\} \times P_{\mathrm{Spin}^c}(g_1)_{\left(p, L_\varphi^{-1} u\right)}$$

The action of $\sigma \in \mathrm{Spin}^c(4)$ on $P_{\mathrm{Spin}^c}(g_2)$ is given by

$$\left((p,u), \widetilde{p}\right) \cdot \sigma = \left((p,u) \cdot r_c(\sigma), \widetilde{p} \cdot \sigma\right).$$

The map $\pi_{r^c}(g_2) : P_{\mathrm{Spin}^c}(g_2) \to P_{\mathrm{U}(1)} \times FX(g_2)$ is defined by

$$\pi_{r^c}(g_2)\left((p,u), \widetilde{p}\right) := (p,u),$$

and note that $\pi_{r^c}(g_2)$ is r^c-equivariant, since

$$\begin{aligned} \pi_{r^c}(g_2)\left(\left((p,u), \widetilde{p}\right) \cdot \sigma\right) &= \pi_{r^c}(g_2)\left(\left((p,u) \cdot r_c(\sigma), \widetilde{p} \cdot \sigma\right)\right) \\ &= (p,u) \cdot r_c(\sigma) = \pi_{r^c}(g_2)\left((p,u), \widetilde{p}\right) \cdot r_c(\sigma). \end{aligned}$$

Also, for $\widetilde{p} \in P_{\mathrm{Spin}^c}(g_1)$ with $\pi_{r^c}(g_1)(\widetilde{p}) = \left(p, L_\varphi^{-1} u\right)$, we have

$$\begin{aligned} J_\varphi(\widetilde{p}) &:= \left((p,u), \widetilde{p}\right) \in P_{\mathrm{Spin}^c}(g_2) \text{ and} \\ \left(\pi_{r^c}(g_2) \circ J_\varphi\right)(\widetilde{p}) &= \pi_{r^c}(g_2)\left((p,u), \widetilde{p}\right) = (p,u) \\ &= (I \times L_\varphi)\left(p, L_\varphi^{-1} u\right) = \left((I \times L_\varphi) \circ \pi_{r^c}(g_1)\right)(\widetilde{p}). \end{aligned}$$

Connections and equivariant functions (e.g., twisted spinor fields) on $P_{\mathrm{Spin}^c}(g_2)$ pullback via J_φ to connections and equivariant functions on $P_{\mathrm{Spin}^c}(g_1)$. Thus, the connections and twisted spinor fields on $P_{\mathrm{Spin}^c}(g_2)$ associated with the metric g_2 can all be identified with connections and twisted spinor fields on $P_{\mathrm{Spin}^c}(g_1)$ associated with a fixed metric g_1. However, the pull-back of the Dirac operator and the star operator for g_2 are not the same as the usual Dirac operator and star operator for g_1, and we must elaborate on this. Note that $\varphi \in C^\infty\left(S_{g_1}^+ X\right)$ induces $\varphi^* \in C^\infty\left(\mathrm{Aut}(\Lambda^*(X))\right)$, where $(\varphi^* \alpha)(Y_1, ..., Y_k) = \alpha\left(\varphi(Y_1), ..., \varphi(Y_k)\right)$. The star operator $*_{g_2}$ of $g_2 = g_1 \cdot \varphi^{-1}$ is given by

$$*_{g_2} = \varphi^{-1*} \circ *_{g_1} \circ \varphi^*.$$

Observe that

$$\begin{gathered} \eta \in \Omega^{2+_{g_1}}(X, i\mathbb{R}) \Leftrightarrow \varphi^{-1*}\eta \in \Omega^{2+_{g_2}}(X, i\mathbb{R}), \text{ since} \\ *_{g_2}\left(\varphi^{-1*}\eta\right) = \left(\varphi^{-1*} \circ *_{g_1} \circ \varphi^*\right)\left(\varphi^{-1*}\eta\right) = \varphi^{-1*}\left(*_{g_1}\eta\right). \end{gathered}$$

Note that

$$\begin{aligned}\Omega^{\omega+_{g_2}} = \varphi^{-1*}\eta &\Leftrightarrow \tfrac{1}{2}\left(1+\varphi^{-1*}\circ *_{g_1}\circ\varphi^*\right)\Omega^\omega = \varphi^{-1*}\eta \\ &\Leftrightarrow \tfrac{1}{2}\left(\varphi^* + *_{g_1}\circ\varphi^*\right)\Omega^\omega = \eta \\ &\Leftrightarrow \tfrac{1}{2}\left(1+*_{g_1}\right)\varphi^*\Omega^\omega = \eta \\ &\Leftrightarrow \left(\varphi^*\Omega^\omega\right)^{+_{g_1}} = \eta.\end{aligned}$$

Thus, in the case $\psi = 0$, the S-W equation $\Omega_\omega^{+_{g_2}} = \varphi^{-1*}\eta$ relative to g_2 is equivalent to $(\varphi^*\Omega_\omega)^{+_{g_1}} = \eta$. The other S-W equation $\mathcal{D}_c^\omega\psi = 0$ also depends on the metric. Let $\theta_2 \in \mathcal{C}(P_{\mathrm{SO}(4)}(g_2))$ be the Levi-Civita connection of g_2 and let $\omega \in \mathcal{C}(P_{\mathrm{U}(1)})$. Then $L_\varphi^*\theta_2 \in \mathcal{C}(P_{\mathrm{SO}(4)}(g_1))$, but $\theta_1 \neq L_\varphi^*\theta_2$ in general. Due to the commutativity of (19.106),

$$(\mathfrak{r}^c)^{-1}\left(P_{r^c}(g_1)^*\left(\omega\oplus L_\varphi^*\theta_2\right)\right) = J_\varphi^*\left((\mathfrak{r}^c)^{-1}\left(P_{r^c}(g_2)^*\left(\omega\oplus\theta_2\right)\right)\right).$$

If $\Sigma_{c,g}(X) = \Sigma_{c,g}^+(X)\oplus\Sigma_{c,g}^-(X)$ denotes the bundle of virtual twisted spinors relative to g, then

$$\Sigma_{c,g_1}(X) = P_{\mathrm{Spin}^c}(g_1)\times_{\mathrm{Spin}^c}\Sigma_4 \overset{J_\varphi\times_{\mathrm{Spin}^c}I_{\Sigma_4}}{\cong} P_{\mathrm{Spin}^c}(g_2)\times_{\mathrm{Spin}^c}\Sigma_4 = \Sigma_{c,g_2}(X)$$

where

$$\left(J_\varphi\times_{\mathrm{Spin}^c}I_{\Sigma_4}\right)([p,w]) := [J_\varphi(p), w].$$

If $\psi \in C^\infty(\Sigma_{c,g_2}(X))$ corresponds to the equivariant Σ_4-valued function $\widetilde{\psi} \in \overline{\Omega}^0(P_{\mathrm{Spin}^c}(g_2),\Sigma_4)$, then $\left(J_\varphi\times_{\mathrm{Spin}^c}I_{\Sigma_4}\right)^{-1}(\psi) \in C^\infty(\Sigma_{g_1,c}(X))$ corresponds to the pull-back

$$J_\varphi^*\widetilde{\psi} := \widetilde{\psi}\circ J_\varphi \in \overline{\Omega}^0(P_{\mathrm{Spin}^c}(g_1),\Sigma_4).$$

Consequently, we will denote the isomorphism $C^\infty(\Sigma_{c,g_2}(X)) \cong C^\infty(\Sigma_{c,g_1}(X))$ induced by $\left(J_\varphi\times_{\mathrm{Spin}^c}I_{\Sigma_4}\right)^{-1}$ simply by

$$J_\varphi^* : C^\infty(\Sigma_{c,g_2}(X)) \cong C^\infty(\Sigma_{c,g_1}(X))$$

In other words, there is a well-defined notion (i.e., J_φ^*) of pull-back of twisted spinors, induced by $\varphi \in C^\infty\left(S_{g_1}^+X\right)$. We let $\nabla^{(\omega,\theta_2)}$ denote covariant differentiation on $C^\infty(\Sigma_{c,g_2}(X))$ for the connection $(\mathfrak{r}^c)^{-1}\left(P_{r^c}(g_2)^*(\omega\oplus\theta_2)\right)$, and $\nabla^{\left(\omega,L_\varphi^*\theta_2\right)}$ denote covariant differentiation on $C^\infty(\Sigma_{c,g_1}(X))$ for the connection $(\mathfrak{r}^c)^{-1}\left(P_{r^c}(g_1)^*\left(\omega\oplus L_\varphi^*\theta_2\right)\right)$. One can check that for $\psi \in C^\infty(\Sigma_{c,g_2}(X))$,

$$\nabla_Y^{\left(\omega,L_\varphi^*\theta_2\right)}\left(J_\varphi^*\psi\right) = J_\varphi^*\nabla_{\varphi(Y)}^{(\omega,\theta_2)}\psi.$$

In other words, regarding $\nabla^{\left(\omega,L_\varphi^*\theta_2\right)}J_\varphi^*\psi \in C^\infty\left(\Lambda^1(X)\otimes\Sigma_{c,g_2}(X)\right)$ as a $\Sigma_g(X)$-valued 1-form,

$$\nabla^{\left(\omega,L_\varphi^*\theta_2\right)}\left(J_\varphi^*\psi\right) = \left(\varphi^*\otimes J_\varphi^*\right)\left(\nabla^{(\omega,\theta_2)}\psi\right).$$

We denote Clifford multiplication relative to g by

$$cl_g : \Lambda^*(X)\otimes\Sigma_{g,c}(X)\to\Sigma_{g,c}(X)$$

We have

$$cl_{g_1}\left(\left(\varphi^*\otimes J_\varphi^*\right)(\alpha\otimes\psi)\right) = cl_{g_1}\left((\varphi^*\alpha)\otimes J_\varphi^*\psi\right) = J_\varphi^*\left(cl_{g_2}(\alpha\otimes\psi)\right).$$

Thus,

$$J_\varphi^*\left(\mathcal{D}_{(\omega,\theta_2)}\psi\right) = J_\varphi^*\left(cl_{g_2}\left(\nabla^{(\omega,\theta_2)}(\psi)\right)\right)$$
$$= cl_{g_1}\left(\left(\varphi^*\otimes J_\varphi^*\right)\left(\nabla^{(\omega,\theta_2)}(\psi)\right)\right)$$
$$= cl_{g_1}\left(\nabla^{\left(\omega,L_\varphi^*\theta_2\right)}\left(J_\varphi^*\psi\right)\right).$$

Letting

$$\mathcal{D}_c^{\left(\omega,L_\varphi^*\theta_2\right)} := cl_{g_1}\circ\nabla^{\left(\omega,L_\varphi^*\theta_2\right)},$$

we then have

$$J_\varphi^*\left(\mathcal{D}_c^{(\omega,\theta_2)}\psi\right) = \mathcal{D}_c^{\left(\omega,L_\varphi^*\theta_2\right)}\left(J_\varphi^*\psi\right)$$

or

$$\mathcal{D}_c^{\left(\omega,L_\varphi^*\theta_2\right)} = J_\varphi^*\circ\mathcal{D}_c^{(\omega,\theta_2)}\circ J_{\varphi^{-1}}^* : C^\infty\left(\Sigma_{g_1}(X)\right)\to C^\infty\left(\Sigma_{g_1}(X)\right).$$

Thus,

$$\mathcal{D}_c^{(\omega,\theta_2)}J_{\varphi^{-1}}^*\psi = 0 \Leftrightarrow \mathcal{D}_c^{\left(\omega,L_\varphi^*\theta_2\right)}\psi = 0.$$

We can also view

$$cl_g : \Lambda^*(X)\otimes\Sigma_{c,g}(X)\to\Sigma_{c,g}(X) \quad \text{as} \quad cl_g : \Lambda^*(X)\to\operatorname{End}\left(\Sigma_{c,g}(X)\right).$$

Recall (see (19.34), p. 629) that there is a bilinear map

$$\widetilde{q}_g : C^\infty\left(\Sigma_{c,g}^+(X)\right)\times C^\infty\left(\Sigma_{c,g}^+(X)\right)\to\Omega^{2+_g}(X,\mathbb{C}).$$

We have

$$\widetilde{q}_{g_1}\left(J_\varphi^*\psi, J_\varphi^*\zeta\right) = \left(\varphi^*\widetilde{q}_{g_2}\right)(\psi,\zeta) \quad \text{or} \quad \widetilde{q}_{g_1}(\psi,\zeta) = \left(\varphi^*\widetilde{q}_{g_2}\right)\left(J_{\varphi^{-1}}^*\psi, J_{\varphi^{-1}}^*\zeta\right).$$

If q_g is the quadratic map corresspoonding to $\widetilde{q}_g$, we have

$$\Omega^{\omega+_{g_2}} - q_{g_2}\left(J_{\varphi^{-1}}^*\psi\right) = \varphi^{-1*}\eta$$
$$\Leftrightarrow \varphi^*\left(\Omega^{\omega+_{g_2}} - q_{g_2}\left(J_{\varphi^{-1}}^*\psi\right)\right) = \eta$$
$$\Leftrightarrow \left(\varphi^*\Omega^\omega\right)^{+_{g_1}} - \varphi^* q_{g_2}\left(J_{\varphi^{-1}}^*\psi\right) = \eta$$

$$\Leftrightarrow \left(\varphi^*\Omega_\omega\right)^{+_{g_1}} - q_{g_1}(\psi) = \eta. \tag{19.107}$$

In summary, we have the following equivalence between the S-W equations for the metric $g_2 := g_1\cdot\varphi^{-1}$ and the φ-perturbed S-W equations involving spinor fields which are associated with the fixed Spin^c structure for the metric g_1 $P_{\mathrm{Spin}^c}(g_1)\to P_{\mathrm{U}(1)}\times FX(g_1)$:

$$\boxed{\begin{array}{l}\mathcal{D}_c^{(\omega,\theta_2)}J_{\varphi^{-1}}^*\psi = 0\\ \Omega^{\omega+_{g_2}} - q_{g_2}\left(J_{\varphi^{-1}}^*\psi\right) = \varphi^{-1*}\eta\end{array}} \quad\Leftrightarrow\quad \boxed{\begin{array}{l}\mathcal{D}_c^{\left(\omega,L_\varphi^*\theta_2\right)}\psi = 0\\ \left(\varphi^*\Omega_\omega\right)^{+_{g_1}} - q_{g_1}(\psi) = \eta\end{array}} \tag{19.108}$$

We call the latter equations, the **g_1-based S-W equations perturbed by** $\varphi\in C^\infty(\operatorname{Aut}(TX))$ **and** $\eta\in\Omega^{2,+_{g_1}}(X,i\mathbb{R})$. In Section 19.3, the metric g was fixed and the S-W equations were only perturbed by $\eta\in\Omega^{2,+_g}(X,i\mathbb{R})$ (or $W^{2,k}\left(\Lambda^{2,+_g}(X,i\mathbb{R})\right)$. Here, we may still regard the metric g_1 as being fixed and C^∞. However, there is an additional space of perturbations, namely $C^\infty\left(S_{g_1}^+X\right)\subset C^\infty(\operatorname{End}(TX))$ or an appropriate Sobolev enlargement, say $W^{2,k'}\left(S_{g_1}^+X\right)\subset W^{2,k'}(\operatorname{End}(TX))$, where k' is sufficiently large depending on k. One complication which arises is that (in local coordinates) the coefficients of the Dirac operator $\mathcal{D}_{\left(\omega, L_{\varphi}^{*}\theta_2\right)}$ are not necessarily C^∞ if $\varphi \in W^{2,k'}\left(S_{g_1}^{+}X\right)$. Since the Levi-Civita connection for a metric may be locally expressed (via Christoffel symbols) in terms of the metric cofficients and their first derivatives, the coefficients of $\mathcal{D}_{\left(\omega, L_{\varphi}^{*}\theta_2\right)}$ will be in $W^{2,k'-1} \subseteq C^m$ for $m < k' - 3$. For elliptic operators with sufficiently regular coefficients, there are versions (see [**Can**]) of the Fundamental Elliptic Estimate (Proposition 17.21, p. 17.21), Elliptic Decomposition (Proposition 17.23) and Unique Continuation (Theorem 19.43) with the same conclusions. We always choose k' large enough so that these conclusions hold. We mention that the Sobolev norms in these theorems are still all defined in terms of the fixed C^∞metric g_1, as opposed to $g_1 \cdot \varphi^{-1}$. We need to address how the theorems of Section 19.3 are modified when the additional pertubation space $W^{2,k'}\left(S_{g_1}^{+}X\right)$ is included.

We set

$$\mathcal{CWPS}_k := \mathcal{CWP}_k \times W^{2,k'}\left(S_{g_1}^{+}X\right)\right),$$

where the "$\mathcal{S}$" stands for "symmetric". We have the extended map (see (19.46))

$$FS : \mathcal{CWPS}_k \to W^{2,k}\left(\Lambda^{2+}\left(X, i\mathbb{R}\right)\right) \oplus W^{2,k}\left(\Sigma_c^{-}\left(X\right)\right)$$

given by

$$FS\left(\omega, \psi, \eta, \varphi\right) := \left(\left(\varphi^{*}\Omega^{\omega}\right)^{+g_1} - q_{g_1}\left(\psi\right) - \eta,\ \mathcal{D}_c^{\left(\omega, L_{\varphi}^{*}\theta_2\right)}\psi\right)$$

If $FS\left(\omega, \psi, \eta, \varphi\right) = 0$, then the differential $FS_{*(\omega,\psi,\eta,\varphi)}$ is onto. Indeed, the proof of Theorem 19.44 shows that even the partial derivative of $FS\left(A, \psi, \eta, \varphi\right)$ with respect to $\left(A, \psi, \eta\right)$ is onto, where φ is held fixed. Essentially all we need to do is apply Theorem 19.44 (p. 633) in the case where the underlying metric is $g_1 \cdot \varphi^{-1}$. Even though $g_1 \cdot \varphi^{-1}$ need not be C^∞, the same argument goes through if k' is chosen large enough. Thus,

$$\mathcal{SWPS}_k := \left(FS\right)^{-1}\left(0, 0\right)$$

is a Hilbert submanifold of $\mathcal{CWPS}_k$ by Theorem 17.24 (Implicit Function Theorem I, p. 470). The group $W^{2,k+2}(X, \mathrm{U}(1))$ of gauge transformations acts trivially on the extra parameter space $W^{2,k'}\left(S_{g_1}^{+}X\right)$. Consequently, Theorem 19.49 (p. 638) implies that

$$\mathcal{MCWPS}_k := \mathcal{CWPS}_k / W^{2,k+2}(X, \mathrm{U}(1)) \cong \mathcal{CWP}_k / W^{2,k+2}(X, \mathrm{U}(1)) \times \mathcal{S}_k$$

is a Hausdorff C^∞ Hilbert manifold. Since $J_{\varphi^{-1}}^{*}\left(s \cdot \psi\right) = J_{\varphi^{-1}}^{*}\left(s^{-1}\psi\right) = s^{-1}J_{\varphi^{-1}}^{*}\left(\psi\right)$, the group $W^{2,k+2}(X, \mathrm{U}(1))$ leaves the space $\mathcal{SWPS}_k$ of parametrized g_1-based solutions of equations (19.108) invariant. We may then form the quotient

$$\mathcal{MSWPS}_k := \mathcal{SWPS}_k / W^{2,k+2}(X, \mathrm{U}(1)),$$

and prove that this is a closed Hilbert submanifold of $\mathcal{MCWPS}_k$ as in Theorem 19.50, p. 644. There is a projection map

$$ps : \mathcal{MSWPS}_k \to W^{2,k}\left(\Lambda^{2+}\left(X, i\mathbb{R}\right)\right) \times W^{2,k'}\left(S_{g_1}^{+}X\right)\right)$$

given by

$$ps\left(\left[\left(A, \psi, \eta, \varphi\right)\right]\right) = \left(\eta, \varphi\right),$$

and the g_1-based moduli space $\mathcal{SW}_{g_1,k}(\eta, \varphi)$ at a fixed (η, φ) is defined by

$$\mathcal{SW}_{g_1,k}(\eta, \varphi) := \left(ps\right)^{-1}\left(\eta, \varphi\right).$$

For the moduli spaces $\mathcal{SW}_k(\eta)$ (defined in (19.68), p. 645), the underlying metric g was fixed and hence it was not included in the notation. If we do include it, we write $\mathcal{SW}_k(\eta, g)$. Then we have

$$\mathcal{SW}_{g_1,k}(\eta,\varphi) \cong \mathcal{SW}_k(\varphi^{-1*}\eta, g_1 \cdot \varphi^{-1}).$$

In the definition $\mathcal{SW}_k(\eta) := p_k^{-1}(\eta)$ (see Theorem 19.62) the right side depends on the choice of η, but it could also depend on the chosen metric on X. Thus, in place of (19.102) we should really write

$$\mathcal{SW}([P_{Spin^c}], \eta, g) := c_1(\eta, g)^{d(L)/2}[\mathcal{SW}(\eta, g)],$$

and show that this is independent of the choice of a suitable pair (η, g), in the case $b_2^+ \geq 2$. Here "suitable" means that η is in the residual subset of Theorem 19.62, where the underlying metric is g. In terms of the g_1-based framework, we are to show that

$$\mathcal{SW}_{g_1}([P_{Spin^c}], \eta, \varphi) := c_1(\eta,\varphi)^{d(L)/2}[\mathcal{SW}_{g_1}(\eta,\varphi)]$$

is independent of a suitable pair (η,φ), namely a pair for which $(\varphi^{-1*}\eta, g_1 \cdot \varphi^{-1})$ is suitable. One necessary condition for the suitability of (η,φ) is that $(\varphi^*\Omega^\omega)^{+_{g_1}} \neq \eta$ for all connections ω, so that there will be no solutions of the S-W$_{g_1}$ equations for which $\psi = 0$. We call the set of such pairs in $W^{2,k}\left(\Lambda^{2+}(X, i\mathbb{R})\right) \times W^{2,k'}(S_{g_1}X)$ meeting this condition $\mathbf{FR}_k$, since the condition guarantees that $W^{2,k+2}(X, U(1))$ acts *freely* on $\mathcal{CWPS}_k$, and we call such (η,φ) an $\mathbf{FR}_k$ **pair**. As in the case where the metric is fixed, we will show that the complement of $\mathbf{FR}_k$ lies in a submanifold of codimension b_2^+. Then as a consequence of the Fredholm Transversality Theorem below, for $b_2^+ \geq 2$, two $\mathbf{FR}_k$ pairs (η_0,φ_0) and (η_1,φ_1) can be joined by a path say $(\eta(t), \varphi(t))$ of $\mathbf{FR}_k$ pairs. We will eventually show that generically such a path yields an oriented cobordism between $\mathcal{SW}_{g_1}(\eta_0,\varphi_0)$ and $\mathcal{SW}_{g_1}(\eta_1,\varphi_1)$. The Chern classes $c_1(\eta,\varphi) \in H^2(\mathcal{SW}_{g_1}([P_{Spin^c}],\eta,\varphi),\mathbb{Z})$ are restrictions of the Chern class, say $c_1 \in H^2(\mathcal{MCWPS},\mathbb{Z})$, of the U(1) bundle

$$\mathcal{CWPS}/G_0 \to \mathcal{MCWPS} := \mathcal{CWPS}/C^\infty(X, U(1)).$$

Thus, the evaluation of $c_1^{d(L)/2}$on the two cobordent (and hence homologous) submanifolds $\mathcal{SW}_{g_1}(\eta_0,\varphi_0)$ and $\mathcal{SW}_{g_1}(\eta_1,\varphi_1)$ yields the same integer. Hence,

$$\begin{aligned}
&\mathcal{SW}_{g_1}([P_{Spin^c}],\eta_0,\varphi_0) := c_1(\eta_0,\varphi_0)^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_0,\varphi_0)] \\
&= c_1^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_0,\varphi_0)] = c_1^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_1,\varphi_1)] \\
&= c_1(\eta_1,\varphi_1)^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_1,\varphi_1)] =: \mathcal{SW}_{g_1}([P_{Spin^c}],\eta_1,\varphi_1),
\end{aligned}$$

or equivalently,

$$\mathcal{SW}([P_{Spin^c}], \varphi_0^{-1*}\eta, g_1\cdot\varphi_0) = \mathcal{SW}([P_{Spin^c}], \varphi_1^{-1*}\eta, g_1\cdot\varphi_1).$$

Since any suitable pair (η_0, g_0) can be written in the form $\left(\varphi_0^{-1*}\eta, g_1\cdot\varphi_0\right)$ for some suitable (η,φ_0), the independence of $\mathcal{SW}([P_{Spin^c}],\eta,g)$ on (η,g) will then be established.

We now show that the set

$$(\mathbf{FR}_k)^c = \left\{ \begin{array}{c} (\eta,\varphi) \in W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \times W^{2,k'}(P_{g_1}(TX)) : \\ (\varphi^* F_A)^{+_{g_1}} = \eta \text{ for some } A \in \mathcal{C}^{2,k+1}(P_{U(1)}) \end{array} \right\} \tag{19.109}$$

of *unsuitable* pairs is contained in a submanifold of codimension b_2^+. Recall that for any metric g on X, the g-harmonic representative $\mathcal{H}_g(\Omega^\omega) = \gamma_g(P_{\mathrm{U}(1)})$ of $[\Omega^\omega] =$

$-2\pi i\, c_1(P_{\mathrm{U}(1)}) \in 2\pi i\, H^2(X,\mathbb{Z})$ is independent of the choice of ω. If $(\eta,\varphi) \in (\mathbf{FR}_k)^c$, then $(\varphi^*\Omega^\omega)^{+g_1} = \eta$ for some ω, and

$$\begin{aligned} 0 = (\varphi^*\Omega^\omega)^{+g_1} - \eta \Rightarrow 0 &= \mathcal{H}_{g_2}\left(\varphi^{-1*}\left((\varphi^*\Omega^\omega)^{+g_1} - \eta\right)\right) \\ &= \mathcal{H}_{g_2}\left(\Omega^{\omega+g_2} - \varphi^{-1*}\eta\right) = \mathcal{H}_{g_2}\left(\Omega^\omega\right)^{+g_2} - \mathcal{H}_{g_2}\left(\varphi^{-1*}\eta\right) \\ &= \gamma_{g_2}(P_{\mathrm{U}(1)})^{+g_2} - \mathcal{H}_{g_2}\left(\varphi^{-1*}\eta\right) \end{aligned}$$

or

$$\mathcal{H}_{g_2}\left(\varphi^{-1*}\eta\right) = \gamma_{g_2}(P_{\mathrm{U}(1)})^{+g_2} \in \mathcal{H}^{2+}_{g_2}(X, i\mathbb{R}).$$

Thus, it suffices to prove that the map

$$\begin{aligned} &J : W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right) \times W^{2,k'}\left(S_{g_1}X\right) \to \mathcal{H}^{2+}_{g_2}(X,i\mathbb{R}), \text{ given by} \\ &J(\eta,\varphi) = \mathcal{H}_{g_2}\left(\varphi^{-1*}\eta\right), \end{aligned}$$

has a surjective differential at (η,φ), since $J^{-1}\left(\gamma_{g_2}(P_{U(1)})^{+g_2}\right)$ will then be a submanifold of codimension $\dim\left(\mathcal{H}^{2+}_{g_2}(X,i\mathbb{R})\right) = b_2^+$ containing $(\mathbf{FR}_k)^c$. We have

$$\begin{aligned} J_{(\eta,\varphi)*}(\eta',\varphi') &= \mathcal{H}_{g_2}\left(\left(\varphi^{-1*}\varphi'^*\varphi^{-1*}\right)(\eta)\right)^{+g_2} + \mathcal{H}_{g_2}\left(\varphi^{-1*}\eta'\right)^{+g_2} \\ &= \mathcal{H}_{g_2}\left(\left(\varphi^{-1}\circ\varphi'\circ\varphi^{-1}\right)^*(\eta)\right)^{+g_2} + \mathcal{H}_{g_2}\left(\varphi^{-1*}\eta'\right)^{+g_2}. \end{aligned}$$

Given any $\xi \in \mathcal{H}^{2+}_{g_2}(X,i\mathbb{R})$, we have $\varphi^*\xi \in W^{2,k}\left(\Lambda^{2+}(X,i\mathbb{R})\right)$ and $J_{(\eta,\varphi)*}(\varphi^*\xi,0) = \mathcal{H}_{g_2}\left(\varphi^{-1*}\varphi^*\xi\right) = \mathcal{H}_{g_2}(\xi) = \xi$, as required. In order to produce an oriented cobordism between $\mathcal{SW}_{g_1}(\eta_0,\varphi_0)$ and $\mathcal{SW}_{g_1}(\eta_1,\varphi_1)$, we need to recall some definitions and results from transversality theory.

Definition 19.66 (transversality). Given C^∞ maps $f_1 : M_1 \to N$ and $f_2 : M_2 \to N$ for Banach manifolds M_1, M_2 and N, we say that f_1 and f_2 are transverse to each other if for all $(x_1,x_2) \in M_1 \times M_2$ with $f_1(x_1) = f_2(x_2)$, we have $f_{1*}(T_{x_1}M_1) + f_{2*}(T_{x_2}M_2) = T_yN$, where $y = f_1(x_1) = f_2(x_2)$. We write $f_1 \pitchfork f_2$.

We can show that if f_1 and f_2 are transverse to each other, then (with $\Delta_N := \mathit{diag}(N)$)

$$(f_1 \times f_2)^{-1}(\Delta_N) := \{(x_1,x_2) \in M_1 \times M_2 : f_1(x_1) = f_2(x_2)\}$$

is a submanifold of $M_1 \times M_2$. Indeed, in terms of a coordinate ball $\varphi : U \to B$ about $y = f_1(x_1) = f_2(x_2) \in U$ (where B is a ball about 0 in some Banach space), we have

$$\begin{aligned} &(f_1 \times f_2)^{-1}\left(\Delta_{\varphi^{-1}\left(\frac{1}{2}B\right)}\right) \\ &= \left\{ \begin{array}{l} (x_1,x_2) \in (\varphi\circ f_1)^{-1}\left(\tfrac{1}{2}B\right) \times (\varphi\circ f_2)^{-1}\left(\tfrac{1}{2}B\right) : \\ (\varphi\circ f_1)(x_1) - (\varphi\circ f_2)(x_2) = 0 \end{array} \right\} \\ &= (\varphi\circ f_1 - \varphi\circ f_2)^{-1}(0), \end{aligned}$$

where

$$\varphi\circ f_1 - \varphi\circ f_2 : (\varphi\circ f_1)^{-1}\left(\tfrac{1}{2}B\right) \times (\varphi\circ f_2)^{-1}\left(\tfrac{1}{2}B\right) \to B$$

is given by

$$(\varphi\circ f_1 - \varphi\circ f_2)(x_1,x_2) = (\varphi\circ f_1)(x_1) - (\varphi\circ f_2)(x_2).$$

Since f_1 and f_2 are transverse, $(\varphi \circ f_1 - \varphi \circ f_2)_*$ is onto at each point of $(\varphi \circ f_1 - \varphi \circ f_2)^{-1}(0)$. Thus, by the Implicit Function Theorem I (Theorem 17.24, p. 470), $(f_1 \times f_2)^{-1}(\Delta_N)$ is a submanifold of $M_1 \times M_2$. We also have

THEOREM 19.67 (Fredholm Transversality). *Let $f_1 : M_1 \to N$ and $f_2 : M_2 \to N$ be C^∞ maps for Banach manifolds M_1, M_2 and N with $\dim M_2 < \infty$ and f_2 Fredholm. Then there is a map $\tilde{f}_2 : M_2 \to N$ arbitrarily close to f_2 in the topology of C^∞ convergenge on compact sets), such that $f_1 \pitchfork \tilde{f}_2$. Moreover, if $f_1 \pitchfork f_2$ for points in a closed subset C of M_2, then we may assume that $\tilde{f}_2 = f_2$ on C.*

One consequence of the FTT (Fredholm Transversality Theorem) is that any two points, say p and q, in the complement of a codimension 2 submanifold M_1 of a pathwise connected Banach manifold N can be joined by a curve which lies outside of M_1. Indeed, let $M_2 = [0,1]$, let $f_2 : [0,1] \to N$ be a curve joining p to q, and let $f_1 : M_1 \to N$ be inclusion. By the FTT, f_2 can be purturbed to $\tilde{f}_2 : [0,1] \to N$ (still joining p to q), with $f_1 \pitchfork \tilde{f}_2$. It must be the case that $\tilde{f}_2([0,1]) \cap M_1 = \phi$, since $f_{1*}(T_{x_1}M_1) + f_{2*}(T_{x_2}M_2) = T_yN$ is impossible as M_1 has codimension 2 and $\dim f_{2*}(T_{x_2}M_2) \le 1$. As a corollary, we have

THEOREM 19.68. *For $b_2^+ \ge 2$, two $\boldsymbol{FR}_k$ pairs (η_0, φ_0) and (η_1, φ_1) can be joined by a path say $(\eta(t), \varphi(t))$ of $\boldsymbol{FR}_k$ pairs.*

We now apply the FTT to the case where $N = W^{2,k+1}\left(\Lambda^{2,+}(X, i\mathbb{R})\right) \times W^{2,k'}(S_{g_1}X)$, $M_1 = \mathcal{MSWPS}_k$, $M_2 = [0,1]$,

$$f_1 = ps : \mathcal{MSWPS}_k \to W^{2,k+1}\left(\Lambda^{2,+}(X, i\mathbb{R})\right) \times W^{2,k'}(S_{g_1}X)$$

and

$$f_2 : [0,1] \to W^{2,k+1}\left(\Lambda^{2,+}(X, i\mathbb{R})\right) \times W^{2,k'}(S_{g_1}X)$$

is a curve joining two $\mathbf{FR}_k$ pairs (η_0, φ_0) and (η_1, φ_1) which are regular values of ps. Since (η_0, φ_0) and (η_1, φ_1) are regular values of ps,we have that $ps \pitchfork f_2$ at the endpoints 0 and 1 of $[0,1]$. By the FTT, f_2 can be purturbed to $\tilde{f}_2$ (still joining (η_0, φ_0) and (η_1, φ_1)), so that $ps \pitchfork \tilde{f}_2$. Then $\left(f_1 \times \tilde{f}_2\right)^{-1}(\Delta_N)$ is a submanifold of $\mathcal{MSWPS}_k \times [0,1]$. Moreover,

$$\begin{aligned} C &:= \left(f_1 \times \tilde{f}_2\right)^{-1}(\Delta_N) \\ &= \left\{([\omega, \psi, \eta, \varphi], t) \in \mathcal{MSWPS}_k \times [0,1] : (\eta, \varphi) = ps([A, \psi, \eta, \varphi]) = \tilde{f}_2(t)\right\} \\ &= \cup_{0 \le t \le 1} \mathcal{SW}_{g_1,k}(\tilde{f}_2(t)) \times \{t\}. \end{aligned}$$

We have that

$$\begin{aligned} \partial\left(\left(f_1 \times \tilde{f}_2\right)^{-1}(\Delta_N)\right) &= \left(\mathcal{SW}_{g_1,k}(\tilde{f}_2(0)) \times \{0\}\right) \cup \left(\mathcal{SW}_{g_1,k}(\tilde{f}_2(1)) \times \{1\}\right) \\ &= \left(\mathcal{SW}_{g_1,k}(\eta_0, \varphi_0) \times \{0\}\right) \cup \left(\mathcal{SW}_{g_1,k}(\eta_1, \varphi_1) \times \{1\}\right). \end{aligned}$$

We still must show that C is compact and orientable. Once this is done, we proceed as follows. The mapping $\pi : C \to \mathcal{MSWPS}_k$ given by $([\omega, \psi, \eta, \varphi], t) \mapsto [\omega, \psi, \eta, \varphi]$ exhibits the equality of homology classes (we drop the index k)

$$[\mathcal{SW}_{g_1}(\eta_0, \varphi_0)] = [\mathcal{SW}_{g_1}(\eta_1, \varphi_1)] \in H^{d(L)}(\mathcal{MCWPS}, \mathbb{Z})$$

The invariance of the S-W invariant then follows:

$$\begin{aligned}
&\mathcal{SW}_{g_1\cdot\varphi_0}([P_{\mathrm{Spin}^c}],\eta_0)\\
&=\mathcal{SW}_{g_1}([P_{\mathrm{Spin}^c}],\eta_0,\varphi_0)=c_1(\eta_0,\varphi_0)^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_0,\varphi_0)]\\
&=c_1^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_0,\varphi_0)]=c_1^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_1,\varphi_1)]\\
&=c_1(\eta_1,\varphi_1)^{d(L)/2}[\mathcal{SW}_{g_1}(\eta_1,\varphi_1)]=\mathcal{SW}_{g_1}([P_{\mathrm{Spin}^c}],\eta_1,\varphi_1)\\
&=\mathcal{SW}_{g_1\cdot\varphi_1}([P_{\mathrm{Spin}^c}],\eta_1).
\end{aligned}$$

We now show that C is compact. The compactness results (see Corollary 19.61 and Theorem 19.62) for the individual moduli spaces depend on the estimate (19.89), namely

$$\|\alpha\|_{2,k'}+\|\psi\|_{2,k'}\leq C(k') \tag{19.110}$$

where the constants $C(k')$ only depend on X, η, ω_0 and $k'\geq 3$. There, X had a fixed Riemannian metric g and the dependence of $C(k')$ on X implies a possible dependence of $C(k')$ on g (or equivalently, on $\varphi\in W^{2,k'}(S_{g_1}X)$). However, in retracing the steps involved in producing the constant $C(k')$, one finds as long as η and φ vary within a compact subset of $W^{2,k+1}\left(\Lambda^{2,+}(X,i\mathbb{R})\right)\times W^{2,k''}(S_{g_1}X)$ for sufficiently large k and k'', there is an upper bound on the constants $C(k')$ for a fixed k'. For example, the relevant inequalities on which (19.110) is based are (**??**), (**??**), and various inequalities which come from applying the Fundamental Elliptic Estimate (Proposition 17.21, p. 468) and Sobolev Multiplication (Proposition 17.22, p. 469). In all cases the constants involved depend on the Sobolev norms (relative to g_1) of η and φ of sufficiently large order. Thus, we can find constants $C(k')$ such that 19.110 (or a g_1-based version) holds for $(\eta,\varphi)\in\tilde{f}_2([0,1])$. The compactness of C then follows.

We can produce an orientation on C as follows. Let $z:=(z_1,1):=([A,\psi,\eta,\varphi],t)\in C$. The tangent space of C at z is

$$T_zC=\left\{\left(v,a\tfrac{\partial}{\partial t}\right)\in T_{z_1}(\mathcal{MSWPS}_k)\times\mathbb{R}:(ps)_*(v)=a\tilde{f}_2'(t)\right\}.$$

We wish to identify T_zC as the kernel of a Fredholm map which onto, and proceed as in the discussion following (19.104). Let

$$FS(\omega,\psi,\eta,\varphi):=\left((\varphi^*\Omega^\omega)^{+_{g_1}}-q_{g_1}(\psi)-\eta,\ \mathcal{D}_{(\omega,L_\varphi^*\theta_2)}\psi\right),$$

so that for $z_1=[\omega,\psi,\eta,\varphi]$, $FS(\omega,\psi,\eta,\varphi)=0$. Now $(\omega',\psi',\eta',\varphi')\in T_{z_1}(\mathcal{MSWPS}_k)$ means that $FS_{*(\omega,\psi,\eta,\varphi)}(\omega',\psi',\eta',\varphi')=0$, *and*

$$B^1(\omega',\psi',\eta',\varphi')+L_1(\varphi')=0,$$

where

$$B^1(\omega',\psi',\eta',\varphi'):=\delta\omega'-\tfrac{1}{4}(\langle\psi',\psi\rangle-\langle\psi,\psi'\rangle)$$

and $L_1=L_1(\omega,\psi,\eta,\varphi)$ is a zero-th order operator. We also have

$$FS_{*(\omega,\psi,\eta,\varphi)}(\omega',\psi',\eta',\varphi')=\begin{pmatrix}B^2(\omega',\psi')-\eta'+L_2(\varphi'),\\ B^3(\omega',\psi')+L_3(\varphi')\end{pmatrix},$$

where

$$B^2(\omega',\psi'):=(\varphi^*d\omega')^+-\widetilde{q}_{g_1}(\psi',\psi)-\widetilde{q}_{g_1}(\psi,\psi')$$

$$B^3(\omega',\psi'):=\mathcal{D}_{(\omega,L_\varphi^*\theta_2)}\psi'+\tfrac{1}{2}\omega'\cdot\psi \tag{19.111}$$

and $L_i = L_i(\omega, \psi, \eta, \varphi)$ is a zero-th order operator ($i = 2, 3$). We have

$$v := \left((\omega', \psi', \eta', \varphi'), a\tfrac{\partial}{\partial t}\right) \in T_z C(\omega', \psi', \eta', \varphi')$$
$$\Leftrightarrow \begin{cases} B^1(\omega', \psi') + L_1(\varphi') = 0 \\ FS_{*(\omega,\psi,\eta,\varphi)}(\omega', \psi', \eta', \varphi') = 0 \\ (\eta', \varphi') = (ps)_*(\omega', \psi', \eta', \varphi') = a\tilde{f}_2'(t) \end{cases}$$
$$\Leftrightarrow \begin{cases} B^1(\omega', \psi') + L_1(\varphi') = 0 \\ B^2(\omega', \psi') - \eta' + L_2(\varphi') = 0 \\ B^3(\omega', \psi') + L_3(\varphi') = 0 \\ (\eta', \varphi') - a\tilde{f}_2'(t) = 0. \end{cases}$$
$$\Leftrightarrow \begin{cases} H^1(v) := B^1(\omega', \psi') + L_1(\varphi') = 0 \\ H^2(v) := B^2(\omega', \psi') - \frac{1}{2}\left(\eta' + a\tilde{f}_{22}'(t)\right) + L_2(\varphi') = 0 \\ H^3(v) := B^3(\omega', \psi') + L_3(\varphi') = 0 \\ H^4(v) := (\eta', \varphi') - a\tilde{f}_2'(t) = 0. \end{cases}$$

Since $f_1 = ps$ and $\tilde{f}_2$ are transverse, any $\xi \in L^{2,k}\left(\Lambda^{2,+}(X, i\mathbb{R})\right) \oplus L^{2,k'}(S_{g_1}X)$ is of the form $\xi = (\eta', \varphi') + \tilde{f}_{2*}\left(c\frac{\partial}{\partial t}\right)$ for some $c \in \mathbb{R}$ and $(\omega', \psi', \eta', \varphi') \in T_{q_1}(\mathcal{MSWPS}_k)$. Hence, the Fredholm map $H = \left(H^1, H^2, H^3, H^4\right)$ is onto and has kernel $T_{z_1}(\mathcal{MSWPS}_k)$. If $FS(\omega, \psi, \eta, \varphi) = 0$, then the differential $FS_{*(\omega,\psi,\eta,\varphi)}$ is onto. For $s \in [0, 1]$, define the deformation H_s by

$$\begin{aligned} H_s^1(v) &:= B^1(s)(\omega', \psi') + (1-s)L_1(\varphi') \\ H_s^2(v) &:= B^2(s)(\omega', \psi') - \tfrac{1}{2}(1-s)\left(\eta' + a\tilde{f}_{22}'(t)\right) + (1-s)L_2(\varphi') \\ H_s^3(v) &:= B^3(s)(\omega', \psi') + (1-s)L_3(\varphi') \\ H_s^4(v) &:= (\eta', \varphi') - (1-s)a\tilde{f}_2'(t), \end{aligned}$$

where $B^i(s)$ was the deformation defined in (19.105). When $s = 0$, we arrive at the operator with block matrix

$$\begin{bmatrix} B^1(1) & \mathbf{0} & 0 \\ B^2(1) & \mathbf{0} & 0 \\ B^3(1) & \mathbf{0} & 0 \\ 0 & \mathrm{Id} & 0 \end{bmatrix}$$

where the boldface entries are defined on $L^{2,k}\left(\Lambda^{2,+}(X, i\mathbb{R})\right) \oplus L^{2,k'}(S_{g_1}(X))$. The kernel of H_1 is $\ker\left(B^1(1) \oplus B^2(1) \oplus B^3(1)\right) \oplus (0,0) \oplus \mathbb{R}$, and the cokernel of H_1 is $\operatorname{coker}\left(B^1(1) \oplus B^2(1) \oplus B^3(1)\right) \oplus (0,0) \oplus 0$. Each of these has a natural orientation (see the discussion following (19.105)).

APPENDIX A

Fourier Series and Integrals - Fundamental Principles

Synopsis. Fourier Series: The Fundamental Function Spaces on S^1; Density; Orthonormal Basis; Fourier Coefficients; Plancherel's Identity; Product and Convolution. The Fourier Integral: Different Integral Conventions; Duality Between Local and Global - Point and Neighborhood - Multiplication and Differentiation - Bounded and Continuous; Fourier Inversion Formula; Plancherel and Poisson Summation Formulae; Parseval's Equality; Higher Dimensional Fourier Integrals

The following fundamentals and elementary facts are standard mathematical knowledge today, and can be found in a great number of text books in analysis. As a general reference, we mention [**DM**, 1972].

1. Fourier Series

We use the notation $S^1 := \{z \in \mathbb{C} : |z| = 1\}$, and define the function spaces

$$C^0(S^1) := \begin{cases} \text{the Banach space of continuous } \mathbb{C}\text{-valued functions on } S^1 \text{ with} \\ \text{the norm } \|f\|_\infty := \sup\left\{|f(z)| : z \in S^1\right\}, \end{cases}$$

$$L^1(S^1) := \begin{cases} \text{the Banach space of } \mathbb{C}\text{-valued, integrable functions on } S^1 \text{ with} \\ \text{the norm } \|f\|_1 := \int_{S^1} |f|, \end{cases}$$

$$L^2(S^1) := \begin{cases} \text{the Hilbert space of square-integrable } \mathbb{C}\text{-valued functions on } S^1 \\ \text{with inner product } \langle f, g\rangle := \int_{S^1} f\bar{g} \text{ and norm } \|f\|_2 := \sqrt{\langle f, f\rangle}. \end{cases}$$

Warning: Functions in $L^1(S^1)$ or $L^2(S^1)$ are identified if they agree outside a set of measure zero. In particular, f is identified with the zero function, if f is zero *almost everywhere*; i.e., f is nonzero only on a set of measure zero. In this way, we have $\|f\| = 0$ precisely when $f = 0$. Thus, strictly speaking, the elements of $L^1(S^1)$ or $L^2(S^1)$ are not functions, but rather equivalence classes of functions. While this is true, in practice it is much simpler and generally harmless to disregard this fine distinction, and we will do this in what follows. Moreover, it is convenient to regard $L^1(S^1)$ as $L^1([0,1])$, and $L^2(S^1)$ as $L^2([0,1])$, and we will do so often without comment. Then $\|f\|_1 := \int_0^1 |f(x)|\,dx$ and $\langle f, g\rangle := \int_0^1 f(x)\overline{g(x)}\,dx$. Perhaps $f(0) \neq f(1)$, but this does not matter in L^1 or L^2 since $\{0, 1\}$ is of measure 0. However, $C^0(S^1)$ and $C^0([0,1])$ are not naturally identified.

Exercise A.1. Show that $L^2(S^1)$ with $\langle \cdot, \cdot\rangle$ is indeed a Hilbert space. We need to show:
a) $\langle \cdot, \cdot\rangle : L^2(S^1) \times L^2(S^1) \to \mathbb{C}$ is well defined. [Hint: The pointwise estimate

$$2|f(z)\overline{g(z)}| = 2\,|f(z)|\,|g(z)| \leq |f(z)|^2 + |g(z)|^2$$

shows that $f\bar{g} \in L^1(S^1)$ for $f, g \in L^2(S^1)$.]
b) $\langle \cdot, \cdot \rangle$ is sesquilinear, namely $\langle f, g \rangle$ is $\mathbb{C}$-linear in f, and $\langle f, g \rangle = \overline{\langle g, f \rangle}$ (i.e., $\langle f, g \rangle$ is conjugate linear in g). Also $\langle \cdot, \cdot \rangle$ is positive; i.e., $\langle f, f \rangle \geq 0$ and $\langle f, f \rangle = 0$ only for $f = 0$. (All of this is trivial.)
c) $L^2(S^1)$ is a complex vector space. [Hint: For closure under addition, prove Hermann Minkowski's inequality $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$ (the Triangle Inequality).]
d) $L^2(S^1)$ is complete. [Hint: In order to prove that a Cauchy sequence $\{f_n\}$ in $L^2(S^1)$ (i.e., $\|f_n - f_m\|_2 \to 0$ as $n, m \to \infty$) possesses a limit $f \in L^2(S^1)$ (i.e., $\|f_n - f\|_2 \to 0$ as $n \to \infty$), one applies the fundamental convergence theorems which distinguish the Lebesgue integral from the Riemann integral. The rather technical proof can be found in [**DM**, p.16-20].]
e) $L^2(S^1)$ is separable. [Hint: Show that the family of piecewise constant functions, having rational real and imaginary parts and with jumps at finitely many rational points, is dense in $L^2(S^1)$.]

Approximation: With the help of the smoothing functions of the kind

$$g(x) := \begin{cases} c \exp\left((x-a)^{-1} (x-b)^{-1} \right) & \text{for } a < x < b \\ 0 & \text{for } x \leq a \text{ or } x \geq b \end{cases}$$

it follows that $C^\infty(S^1)$ is dense in $L^2(S^1)$.

Convolution: In $L^1(S^1)$, there is a commutative and associative product (known as *convolution*) given by

$$(f * g)(x) := \int_0^1 f(x-y) g(y) \, dy, \ x \in [0,1],$$

where we assume that f is extended periodically of period 1 so that $f(x-y)$ makes sense when $x - y \notin [0,1]$. This makes $L^1(S^1)$ an algebra (without identity). By applying the theorem of Guido Fubini on iterated integrals, we obtain

$$\|f * g\|_1 \leq \|f\|_1 \ \|g\|_1 .$$

Moreover, one can show that, relative to $*$, $L^2(S^1)$ is an ideal in $L^1(S^1)$ hence, $f * g$ is in $L^2(S)$, whenever one of the factors lies in $L^2(S)$. See [**DM**, p.41].

Orthonormal systems: The family $\{z^n : n \in \mathbb{Z}\}$, where $z^n : S^1 \to \mathbb{C}$ is the function that assigns to each $z \in S^1$ the value z^n, is a complete orthonormal system in $L^2(S^1)$. Regarding $L^2(S^1)$ as $L^2([0,1])$, the corresponding functions have the form $e^{2\pi i n x}$.

Fourier series: This orthonormal system is complete; i.e., its linear span is dense in $L^2(S^1)$. Because of this, each function $f \in L^2(S^1)$ can be expanded in a Fourier series

$$f = \sum\nolimits_{n=-\infty}^{\infty} \widehat{f}(n) z^n \ \ (\text{i.e., } \lim_{k \to \infty} \| f - \sum\nolimits_{n=-k}^{k} \widehat{f}(n) z^n \|_2 = 0).$$

with the Fourier coefficients

$$\widehat{f}(n) := \langle f, z^n \rangle = \int_0^1 f(x) e^{-2\pi i n x} \, dx. \tag{A.1}$$

Note that f equals its infinite Fourier series, in the sense that the partial sums $\sum_{|n| \leq k} \widehat{f}(n) z^n$ converge to f in the $L^2(S^1)$-norm as $k \to \infty$, but not necessarily pointwise. The function

$$f \mapsto \left\{ \widehat{f}(n) \right\}_{n=-\infty}^{\infty} = \ldots, \widehat{f}(-1), \widehat{f}(0), \widehat{f}(1), \ldots$$

is an isomorphism from $L^2(S^1)$ to the space $L^2(\mathbb{Z})$ of absolute square-summable sequences of complex numbers. The isomorphism is an isometry, namely

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 \quad \text{(Plancherel's Identity).}$$

Details are in [**DM**]. The Fourier coefficients are also defined for $f \in L^1(S^1)$, and by Fubini's Theorem, it then follows [**DM**, p.42] that

$$(f * g)\widehat{\ }(n) = \widehat{f}(n)\widehat{g}(n).$$

Incidentally, the algebra A of those sequences which appear as the Fourier coefficients of integrable functions has been barely investigated: "The best information available to date indicates that A has no decent description at all." [**DM**, p.43].

On the other hand, when the product fg (in the sense of pointwise multiplication) is integrable (e.g., when $f, g \in L^2(S)$, by Exercise A.1a), then the Fourier coefficients satisfy

$$(fg)\widehat{\ }(n) = (\widehat{f} * \widehat{g})(n) := \sum_{k=-\infty}^{\infty} \widehat{f}(n-k)\,\widehat{g}(k).$$

For the proof, we do not need Fubini's Theorem as above, but rather we insert the Fourier series of g in the formula for $(fg)\widehat{\ }(n)$, and then use the usual limit theorems for the Lebesgue integral, to interchange the integral and sum.

2. The Fourier Integral

One can proceed from the standard representation of functions on a circle as functions of period 1 on the real line, to the more general case of period T, and then let T go to ∞. This leads to the concept of the Fourier transform. Let $L^1(\mathbb{R})$ denote the Banach space of integrable functions with

$$\|f\|_1 := \int_{-\infty}^{\infty} |f(x)|\, dx < \infty.$$

Different Integral Conventions. In [**DM**, p.86f.], the Fourier transform of $f \in L^1(\mathbb{R})$ is defined as

$$\widehat{f}_{DM}(\xi) := \int_{-\infty}^{\infty} f(x)e^{-i2\pi\xi x}dx, \tag{A.2}$$

which is a natural extension of (A.1), namely $\widehat{f}(n) := \int_0^1 f(x)e^{-i2\pi nx}dx$.
In the previous edition [**BlBo85**, p.82] of this book, the Fourier transform was given as

$$\widehat{f}_B(\xi) := \int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx = \widehat{f}_{DM}(\xi/2\pi).$$

This agrees with the main stream in analysis and information theory. In [**Rud**, p.167] (and [**BlCs**, p.423]), we find the definition

$$\widehat{f}_R(\xi) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx = \frac{1}{\sqrt{2\pi}}\widehat{f}_B(\xi) = \frac{1}{\sqrt{2\pi}}\widehat{f}_{DM}(\xi/2\pi).$$

In Remark A.3 (below, p.674), there are cogent reasons for adopting any one of the above definitions. Since our emphasis in this book is on the application of Fourier transforms to differential equations and we wish the Fourier transform to be an

L^2 isometry, we use $\widehat{f}_R$ as Remark A.3 suggests. Thus, we define the **Fourier transform** of $f \in L^1(\mathbb{R})$ via

$$\widehat{f}(\xi) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx. \tag{A.3}$$

Since it is an irritating annoyance to have to include the factor $1/\sqrt{2\pi}$, we adopt the notation

$$đx := dx/\sqrt{2\pi}, \text{ so that } \widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} đx.$$

Convolution, Multiplication, Differentiation, and Inversion. We consider the Fourier transformation on the spaces $C_{\downarrow}^{\infty}(\mathbb{R})$, $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$. Here, $C_{\downarrow}^{\infty}(\mathbb{R})$ is the space of *rapidly decreasing* C^{∞} functions on $\mathbb{R}$ (with complex values). *Rapidly decreasing* means that these functions and all their derivatives tend to 0 at infinity, even when they are multiplied by arbitrary polynomials. $L^1(\mathbb{R})$ is an algebra (without identity) under convolution

$$(f * g)(x) := \int_{-\infty}^{\infty} f(x-y) g(y) dy,$$

and we have [**DM**, p.87f.]:

$$\|f * g\|_1 \le \|f\|_1 \ \|g\|_1 .$$

$L^2(\mathbb{R})$ is the Hilbert space (proved as in Exercise A.1) of square integrable functions with

$$\langle f, g \rangle := \int_{-\infty}^{\infty} f(x) \overline{g(x)}\, dx \text{ and } \|f\|_2 := \sqrt{\langle f, f \rangle}$$

By the argument of Exercise A.1a, it follows that $C_{\downarrow}^{\infty}(\mathbb{R})$ is dense in $L^1(\mathbb{R})$, as well as in $L^2(\mathbb{R})$. Naturally one cannot expect, as with functions on the (compact) circle S^1, that $C^{\infty}(\mathbb{R})$ or $C^0(\mathbb{R})$ will be contained in the Lebesgue spaces.

A further complication arises since we no longer have $L^2 \subseteq L^1$, or even $L^1 \subseteq L^2$. For example, if $f(x) = 1$ for $|x| < 1$ and $f(x) = 0$ for $|x| \ge 1$, then

$$f(x)\, |x|^{-2/3} \in L^1(\mathbb{R}) \setminus L^2(\mathbb{R}) \quad \text{and} \quad (1 - f(x))\, |x|^{-2/3} \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R}).$$

Exercise A.2. The following statements are easy to prove:
(a) For each $f \in L^1(\mathbb{R})$ and each $\xi \in \mathbb{R}$, $\widehat{f}(\xi)$ is well defined.
(b) If $f \in C^1(\mathbb{R})$, $f' \in L^1(\mathbb{R})$ and $ixf(x)$ stands for the function $x \mapsto ixf(x)$ which is assumed to be in $L^1(\mathbb{R})$, then we have

$$(f')\widehat{\ }(\xi) = i\xi \widehat{f}(\xi) \text{ and} \tag{A.4}$$

$$(\widehat{f})'(\xi) = -(ixf(x))\widehat{\ }(\xi). \tag{A.5}$$

The proof of each is via integration by parts. Using the following notation for the various operations above

$$(Df)(x) := \frac{1}{i} f'(x), \ \ (Mf)(x) := xf(x), \text{ and } \mathcal{F}f := \widehat{f},$$

we have

$$\mathcal{F}Df = M\mathcal{F}f \text{ and } D\mathcal{F}f = -\mathcal{F}Mf.$$

More generally, by induction, we have (for $f \in C^\infty_\downarrow(\mathbb{R})$ and $p, q \in \mathbb{N}$)

$$M^p D^q \mathcal{F} f = (-1)^q \, \mathcal{F} M^q D^p f.$$

This formula is of fundamental importance for the treatment of differential operators (with constant coefficients) which are converted into simple multipliers. See Chapter 8 on pseudo-differential operators. From the topological viewpoint these formulas are most remarkable, because they express a deep duality between local and global properties: Thus (A.4) relates the smoothness of f with the rate of decay (asymptotic behavior) of $\widehat{f}$, and (A.5) relates the smoothness of $\widehat{f}$ with the decay of f. In fact, $\widehat{f}$ is differentiable (a local property) when f decreases so fast that the Fourier integral of $-ixf(x)$ converges. This local–global duality is also a feature of the index formula for elliptic operators, and we will deal with it further in that context.
(c) The **Fourier Inversion Formula**

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{i\xi x} d\xi$$

holds for $f \in C^\infty_\downarrow(\mathbb{R})$. In direct analogy with the role of Fourier coefficients in Fourier series, $\widehat{f}(\xi)$ is the density of the frequency ξ in the harmonic decomposition of f.
[Hint: Prove the formula first for functions with compact support (i.e., vanishing outside a compact subset of $\mathbb{R}$). In this case there are no difficulties with the limit process which reduces to functions of period T and then let T go to infinity. Use smoothing functions as following Exercise A.1e in order to approximate rapidly decreasing C^∞ functions by functions of compact support. The $L^1(\mathbb{R})$ estimates needed next are somewhat tricky, but can be looked up in [**DM**, p.89f]. A shorter direct proof can be found in [**Hö63**, 1963, p.18f].]
(d) As a corollary to the proof of (c), one obtains the **Plancherel formula**

$$\|\widehat{f}\|_2 = \|f\|_2$$

and that

$$\mathcal{F} : C^\infty_\downarrow(\mathbb{R}) \to C^\infty_\downarrow(\mathbb{R})$$

is linear and bijective, where $\mathcal{F}$ again denotes the Fourier transformation. By the Fourier Inversion Formula, we obtain the inverse transformation

$$\left(\mathcal{F}^{-1} f\right)(x) = (\mathcal{F} f)(-x).$$

(e) Extend $\mathcal{F}$ from $C^\infty_\downarrow(\mathbb{R})$ to $L^2(\mathbb{R})$! [Hint: Approximate $f \in L^2(\mathbb{R})$ in the $L^2(\mathbb{R})$-norm by a sequence $\{f_n\}$ with $f_n \in C^\infty_\downarrow(\mathbb{R})$. Using the additivity of $\mathcal{F}$ and the Plancherel Identity, show that $\{\widehat{f_n}\}$ is a Cauchy sequence in $L^2(\mathbb{R})$, whence $\widehat{f} := \lim \widehat{f_n}$ defines an element of the Hilbert space $L^2(\mathbb{R})$. Finally check that $\widehat{f}$ indeed depends only on f and not on the choice of the sequence. In this way, one obtains an isomorphism from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$, which we denote by $\mathcal{F}$ again.]
(f) The spaces $C^\infty_\downarrow(\mathbb{R})$ and $L^2(\mathbb{R})$ share the property that they are mapped into themselves by $\mathcal{F}$. This is not true for $L^1(\mathbb{R})$. Still, one can easily show [**DM**, p.102]

that for $f \in L^1(\mathbb{R})$,

$$(i)\ \widehat{f} \in C^0(\mathbb{R})$$
$$(ii)\ \lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$$
$$(iii)\ (f * g)\widehat{\ }(\xi) = \sqrt{2\pi}\widehat{f}(\xi)\,\widehat{g}(\xi)\,.$$

REMARK A.3. Here we consider the merits of the three definitions

$$\widehat{f}_{DM}(\xi) := \int_{-\infty}^{\infty} f(x)e^{-i2\pi\xi x}dx,$$
$$\widehat{f}_B(\xi) := \int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx = \widehat{f}_{DM}(\xi/2\pi) \text{ and}$$
$$\widehat{f}_R(\xi) := \int_{-\infty}^{\infty} f(x)e^{-i\xi x}d'x := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx = \frac{1}{\sqrt{2\pi}}\widehat{f}_{DM}(\xi/2\pi),$$

found in [**DM**, p.87], [**BlBo85**, p.82], and [**Rud**, p.167]. The main advantage of $\widehat{f}_B$ is that there is no 2π or $\sqrt{2\pi}$. However, in terms of $\widehat{f}_B$, the Plancherel formula becomes $\|\widehat{f}_B\|_2 = \sqrt{2\pi}\,\|f\|_2$ so that $f \mapsto \widehat{f}_B$ is not an isometry, which is a drawback. We do have $\|\widehat{f}_{DM}\|_2 = \|f\|_2$. Moreover, $\widehat{f}_{DM}$ is good for expressing the remarkable **Poisson summation formula**

$$\sum_{k=-\infty}^{\infty} f(kL) = \frac{1}{L}\sum_{k=-\infty}^{\infty} \widehat{f}_{DM}(k/L)\,, \tag{A.6}$$

which holds for $f \in C_{\downarrow}^{\infty}(\mathbb{R})$ and any $L > 0$. This relates the sum of f over the lattice $\{kL : k \in \mathbb{Z}\}$ to the sum of $\widehat{f}$ over the reciprocal lattice $\{k/L : k \in \mathbb{Z}\}$. Using $\widehat{f}_R$ or $\widehat{f}_B$, this becomes

$$\sum_{k=-\infty}^{\infty} f(kL) = \frac{\sqrt{2\pi}}{L}\sum_{k=-\infty}^{\infty} \widehat{f}_R(2\pi k/L) = \frac{1}{L}\sum_{k=-\infty}^{\infty} \widehat{f}_B(2\pi k/L)$$

which is less esthetic and harder to recall. The convolution theorem gives $(f*g)\widehat{\ }_B = \widehat{f}_B\widehat{g}_B$ and $(f*g)\widehat{\ }_{DM} = \widehat{f}_{DM}\widehat{g}_{DM}$, both of which look better than $(f*g)\widehat{\ }_R = \sqrt{2\pi}\widehat{f}_R\widehat{g}_R$. So far, $\widehat{f}_{DM}$ seems to be the best choice. However,

$$(f')\widehat{\ }_{DM}\ (\xi) = 2\pi i\xi\widehat{f}_{DM}(\xi),$$

and the excess baggage of the 2π makes $\widehat{f}_{DM}$ a bit cumbersome for applications to differential equations. Thus, we have adopted $\widehat{f}_R$, but not passionately and not exclusively; indeed in almost all of Chapter 4, we use $\widehat{f}_B$ since the factor $1/\sqrt{2\pi}$ only serves as a needless distraction.

Higher Dimensional Fourier Integrals. By the theorem of Guido Fubini, the closed linear span of the n-fold products

$$f_1(x_1)\cdots f_n(x_n)$$

of functions in $L^2(\mathbb{R})$ is in $L^2(\mathbb{R}^n)$ (Prove!). Thus, the preceding concepts and results carry over directly to the case of several variables:

DEFINITION A.4. (a) As above, define the spaces $C_{\downarrow}^{\infty}(\mathbb{R}^n)$, $L^1(\mathbb{R}^n)$, and $L^2(\mathbb{R}^n)$; (b) the **Fourier transform**

$$\widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)e^{-i\langle x,\xi\rangle}đx \quad \text{with } đx = (2\pi)^{-n/2}dx_1\cdots dx_n,$$

where $x = (x_1, ..., x_n)$, $\xi = (\xi_1, ..., \xi_n)$, and $\langle x, \xi\rangle = x_1\xi_1 + ... + x_n\xi_n$; and
(c) the **convolution**

$$\begin{array}{ccc} L^1(\mathbb{R}^n) \times L^1(\mathbb{R}^n) & \longrightarrow & L^1(\mathbb{R}^n) \\ (f, g) & \mapsto & (f * g)(x) := \int_{\mathbb{R}^n} f(x-y)g(y)dy,\ x \in \mathbb{R}^n. \end{array}$$

We leave it to the reader to check that $f * g$ is well defined, i.e., for almost all $x \in \mathbb{R}^n$, the integrand $y \mapsto f(x-y)g(y)dy$ belongs to L^1.

EXERCISE A.5. Show:
(a) The **Fourier Inversion Formula**

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{i\langle x,\xi\rangle} đ\xi, \text{ where } f \in C^\infty_\downarrow(\mathbb{R}^n).$$

(b) The **Differentiation-Multiplication Conversion** for multi-indices p, q,

$$M^p D^q \mathcal{F} = (-1)^{|q|} \mathcal{F} D^p M^q, \qquad \text{with } D^q := (-1)^{|q|} \frac{\partial^{|q|}}{\partial x_1^{q_1} \cdots \partial x_n^{q_n}},$$

where $\mathcal{F} f := \widehat{f}$ for $f \in C^\infty_\downarrow(\mathbb{R}^n)$ and $|q| := q_1 + \cdots + q_n$.
(c) The **Integrable-Continuous Conversion** $f \in L^1(\mathbb{R}^n) \Rightarrow \widehat{f} \in C^0(\mathbb{R}^n)$.
(d) **Plancherel Formula** $\|\widehat{f}\|_2 = \|f\|_2$ and more general **Parseval's Equality** $\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi$, for $f, g \in L^2(\mathbb{R}^n)$.
(e) The **Convolution Theorems** $\widehat{fg} = (2\pi)^{n/2}\left(\widehat{f} * \widehat{g}\right)$ and $\widehat{f}\widehat{g} = (2\pi)^{-n/2}(f * g)\widehat{\ }$, if $f, g, \widehat{f}, \widehat{g} \in L^1(\mathbb{R}^n)$.
Details are in [**DM**, p.132f] or [**Hö63**, 1963, p.17-19].

APPENDIX B

Vector Bundles

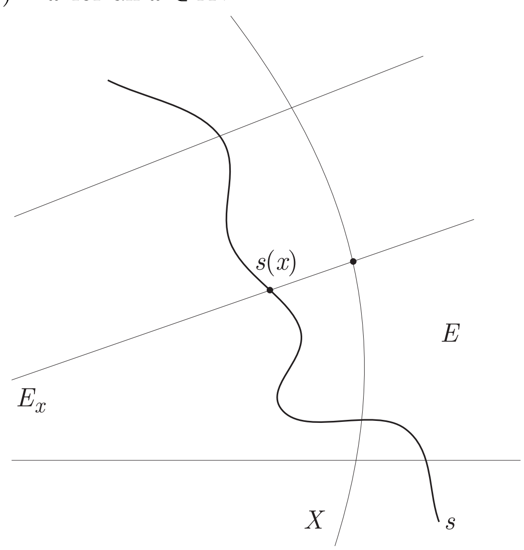
1. Basic Definitions and First Examples

Let X be a topological space. A **family of vector spaces** over X is a topological space E together with

(i) a continuous surjective map $p : E \to X$ and

(ii) a vector space structure of finite dimension in each $E_x := p^{-1}(x)$, which carries the topology induced by E.

By "vector spaces", we mean complex vector spaces, unless explicitly indicated otherwise. The mapping p is called the **projection**;; E is called the **total space** of the family; X is the **parameter space** or **base space** of the family; for $x \in X$, E_x is the **fiber** over x. A section of a family $p : E \to X$ is a continuous map $s : X \to E$ such that $(p \circ s)(x) = x$ for all $x \in X$.



A **homomorphism** (*bundle map*) from one family $p : E \to X$ to another $q : F \to X$ is a continuous map $\phi : E \to F$ such that

(i) $q \circ \phi = p$ and

(ii) $\phi_x : E_x \to F_x$ is a linear map for each $x \in X$. We write $\phi \in \mathrm{Hom}(E, F)$.

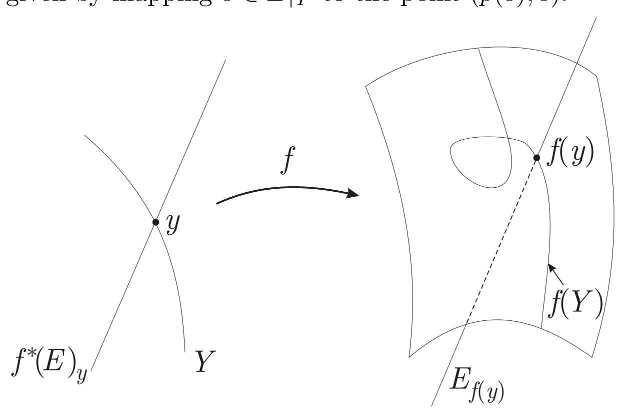
We say that such a ϕ is an **isomorphism** when ϕ is bijective and ϕ^{-1} is continuous. E and F are called **isomorphic** when there is an isomorphism between them. We write $\phi \in \mathrm{Iso}(E, F)$ and $E \cong F$.

EXERCISE B.1. a) Let V be a finite-dimensional vector space; e.g., $V = \mathbb{C}^N$. Show that a family of vector spaces over X is obtained by taking $E := X \times V$ with $p : E \to X$ being the projection on the first factor. This is the product family V_X

with fiber V.
b) If F is a family which is isomorphic to a product family, then one calls F **trivial**. Show that a trivial family of finite dimensional (real) vector spaces is obtained, if

$$E := \{(x, y, -\lambda y, \lambda x) : x, y, \lambda \in \mathbb{R} \text{ and } x^2 + y^2 = 1\} \text{ and}$$
$$p(x, y, \cdot, \cdot) := (x, y).$$

In general, prove that a bundle F is isomorphic to a product family V_X, if and only if one can find N sections $s_i : X \to P$ such that $s_1(x), ..., s_N(x)$ forms a basis for F_x for each $x \in X$; here $N = \dim V$.
c) Let Y be a subset of X and E a family of vector spaces over X with projection p. Show that $p^{-1}(Y) \to Y$ is a family over Y. We call this the **restriction** of E to Y, and write $E|_Y$ for this family.
d) More generally: Let Y be an arbitrary topological space and $f : Y \to X$ a continuous map. As follows, define the **induced family** $f^*(p) : f^*(E) \to Y$: Take $f^*(E)$ to be the subspace of $Y \times E$ consisting of points (y, e) with $f(y) = p(e)$; the projection and the vector space structure on the fibers are self-evident. Show: For each further map $g : Z \to Y$, there is a natural isomorphism $(fg)^*(E) \stackrel{\cong}{\to} g^* f^*(E)$ which one obtains by mapping each point of the form (z, e) with $z \in Z$ and $e \in E$ to the point $(z, g(z), e)$. If $f : Y \to X$ is the inclusion, then there is an isomorphism $E|_Y \stackrel{\cong}{\to} f^*(F)$ given by mapping $e \in E|_Y$ to the point $(p(e), e)$.



A family of vector spaces is called **locally trivial**, if each $x \in X$ possesses a neighborhood U such that $E|U$ is trivial. A locally trivial family is called a **vector bundle**; trivial families are called **trivial bundles**. If $f : Y \to X$ and E is a vector bundle over X, clearly $f^*(E)$ is a vector bundle over Y which we call the **induced bundle (lifted bundle, pull-back)**.

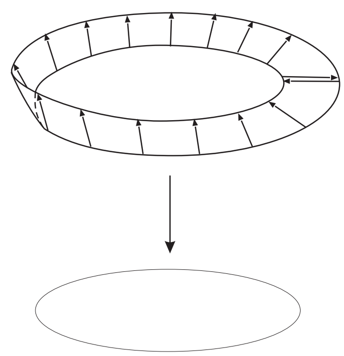
Note: If E is a vector bundle over X, then $x \mapsto \dim(E_x)$ is a locally constant function on X, and hence is constant on each connected component of X. If $\dim(E_x)$ is constant on all of X, then one says that E has **(fiber-)dimension** equal to the common fiber dimension $\dim(E_x)$. If X is a manifold, then the real dimension of E (regarded as a topological space) equals $\dim(X) + 2\dim(E_x)$. Vector bundles of fiber dimension 1 are also called **line bundles**.

Since a vector bundle is locally trivial, each section can be written locally as a vector-valued function on the base space. For a vector bundle E, we denote the space of sections of E by $C^0(E)$; $C^0(E)$ is a vector space in a natural way via pointwise addition, etc.

Exercise B.2. a) Let V be a (complex) vector space and $\mathbb{P}V$ its associated "projective space" of all one-dimensional linear subspaces of V. We can write $\mathbb{P}V = (V \setminus \{0\})/\sim$, where $\sim$ is the equivalence relation $v \sim w \Leftrightarrow \lambda v = w$ for some $\lambda \in \mathbb{C}$. We define $H_V \subseteq \mathbb{P}V \times V$ as the set of all (x, v) such that $x \in \mathbb{P}V$, $v \in V$, and v belongs to the complex line x. Show that H is a vector bundle in a natural way. (The construction goes back to Heinz Hopf.)
b) Go through the corresponding construction in the (more intuitive) category of real vector bundles when V is real (e.g., $\mathbb{R}^n$), and show that H_V is a real subbundle of $\mathbb{P}V \times V$ of fiber dimension 1, and that H_V is nontrivial if $\dim V \geq 2$.
c) For the moment, we remain in the real category and consider the following family (parametrized by $\theta \in [0, 2\pi]$) of integro-differential equations for C^∞ functions f on the unit interval which satisfy the boundary condition $f(0) = f(1)$:

$$\cos\theta\ f(x) + \sin\theta\ \frac{df}{dx} = \cos\theta \int_0^1 f(x)\,dx,\ \ \theta \in [0, \pi]\,,$$
$$\cos\theta\ f(x) + \sin\theta\ Lf(x)) = \cos\theta \int_0^1 f(x)\,dx + \sin\theta \int_0^1 Lf(x)\,dx,\ \ \theta \in [\pi, 2\pi]\,.$$

Here, $L : C^0(S^1) \to C^0(S^1)$ is a fixed operator with $L^2 = -\,\mathrm{Id}$. Show that the solutions of the family of equations form a real vector bundle over the circle $S^1 := \mathbb{R}/2\pi\mathbb{Z}$ that is nontrivial and isomorphic to the bundle $H_{\mathbb{R}^2}$. (Actually, every real line bundle over S^1 is either trivial or isomorphic to $H_{\mathbb{R}^2}$ see also [**BJ**, 3.23.9].)



[Hint for b): In contrast to the complex numbers, -1 cannot be deformed into 1 without going through 0. Thus, a real bundle is nontrivial, if it remains connected after the zero section is removed.
For c): First show that for $0 \leq \theta \leq \pi$, the solutions are the constant functions $c\mathbf{1}$, and for $\pi \leq \theta \leq 2\pi$ the solutions are the functions $c\cos\theta\ \mathbf{1} - \sin\theta\ L\,(\mathbf{1})\,,\ c \in \mathbb{R}$. With the topology of $S \times C^0(I)$ or of $S^1 \times \mathbb{R}$ (since every solution can be written in the form $c_1 + c_2 L(\mathbf{1})$), construct a family of 1-dimensional (real) vector spaces over S^1 and show the local triviality. For this use the initial value map $f \mapsto f(0)$. Note that this map can vanish for a solution $f \neq 0$ at a parameter θ_0, namely if $\cos\theta_0 = \sin\theta_0\ L\,(\mathbf{1})\,(0)$. How can one proceed in a neighborhood of θ_0? Distinguish the cases where $L\,(\mathbf{1})\,(0)$ is positive, negative or zero. Incidentally, how does one obtain an L with $L^2 = -\,\mathrm{Id}$? Start with the Fourier series $f(x) = a_0 + \sum_{\nu=1}^{\infty} (a_\nu \sin\nu x + b_\nu \cos\nu x)$, and replace a_ν by $b_{\nu+1}$ and b_ν by $-a_{\nu-1}$; see also [**Si70**].]

REMARK B.3. a) and b) describe the origin of the bundle concept in analytic and projective geometry. Part c) is characteristic for many functional analytic situations with "jumps" where the passage from one side to another (from one solution curve to another of the same equation) cannot be understood within the given space but requires an extension of the system (e.g., by parametrization). A basic model for such a process is present in the geometry of number fields (see Hint for b)). Many classical results of analysis – especially concerning the dependence of the solutions of a functional equation on the variation of its coefficients and on the zeros and poles of its solutions – can be aptly formulated in the language of vector bundles. Conversely, the theorem of Grothendieck, for example, that every holomorphic vector bundle E on the Riemannian number sphere $S^2 = \mathbb{P}(\mathbb{C}^2)$ can be represented as a Whitney sum $E_1 \oplus \cdots \oplus E_n$ of line bundles (Am. J. Math. **79** (1957), 121-138) was known to analysts at the beginning of the century: See G. Birkhoff, Math. Ann. **54** (1913), 122-139, where Grothendieck's theorem appears as a theorem about matrices of analytic functions. Birkhoff was led to this theorem through his investigation of the singular points of ordinary differential equations; further see D. Hilbert, Gött. Nachr. (1905), 307-358, who gave a proof of Grothendieck's theorem for $N = 2$ in his "Fundamentals of a General Theory of Integral Equations" in connection with the "Riemannian Problem" (Contributed by M. Schneider).

EXERCISE B.4. Show that the usual operations for vector spaces in linear algebra also make sense for vector bundles. In particular, for vector bundles E and F over the same base, investigate the **direct sum** $E \oplus F$, the **tensor product** $E \otimes F$, the **homomorphism bundle** $\mathrm{Hom}(E, F)$, the **isomorphism bundle** $\mathrm{Iso}(E, F)$, and the **dual bundle** $E^* := Hom(E, \mathbb{C}_X)$. Show that the bundles $E^* \otimes F$ and $\mathrm{Hom}(E, F)$ are isomorphic. Also, carry over the concepts of subspace and quotient space from linear algebra to the corresponding concepts of a **subbundle** F of E and a **quotient bundle** E/F.
[Hint: Make use of the fact that the corresponding operations in the "structure group" $GL(N, \mathbb{C})$ are continuous! Example: recall $E^* = \mathrm{Hom}\,(E, \mathbb{C}_X)$ and introduce a topology on $E^*|_U$ which makes $U \times \mathbb{C}^N \to E^*|_U$ a homeomorphism, where $\phi : E|_U \to U \times \mathbb{C}^N$ is a local trivialization for E over the open subset $U \subseteq X$. Let $\psi : E|_V \to V \times \mathbb{C}^N$ be another trivialization. Do ϕ and ψ define the same topology on $E^*|_{U \cap V}$? Does the continuity of $(\psi \circ \phi^{-1})^* : (U \cap V) \times \mathbb{C}^N \to (U \cap V) \times \mathbb{C}^N$ follow from that of $(U \cap V) \times \mathbb{C}^N \to (U \cap V) \times \mathbb{C}^N$? For this, write the two chart changes in the form $U \cap V \to GL(N, \mathbb{C})$ and prove (!) that $GL(N, \mathbb{C}) \xrightarrow{*} GL(N, \mathbb{C})$ is continuous.]

2. Homotopy Equivalence and Isomorphy

We denote the set of isomorphism classes of vector bundles over X by $\mathrm{Vect}(X)$, and let $\mathrm{Vect}_N(X)$ denote the subset of $\mathrm{Vect}(X)$ consisting of the classes of bundles of dimension N. $\mathrm{Vect}(X)$ is an abelian semigroup under the operation $\oplus$. In $\mathrm{Vect}(X)$, there is a naturally distinguished element, namely the class of the trivial bundle of dimension N. A vector bundle over a point is a vector space, and hence $\mathrm{Vect}(X)$ can be identified with the semigroup $\mathbb{Z}^+$ of non-negative integers, in this case. However, in the general ease, when there are nontrivial bundles (see Exercise

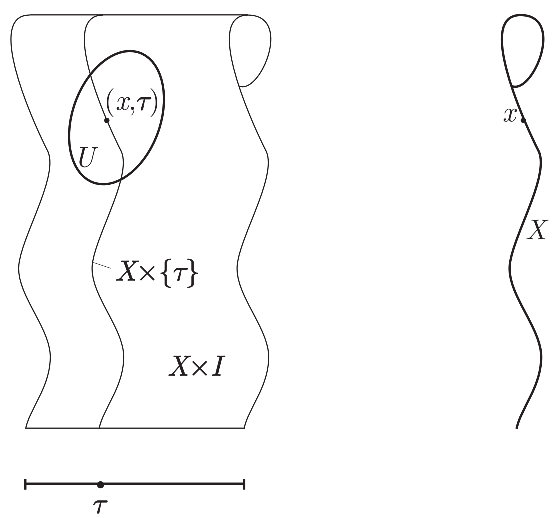
B.2 above), the isomorphism classes of vector bundles are not determined by their dimensions.

Two continuous mappings $f, h : X \to Y$ are **homotopic**, if there is a continuous map $F : X \times I \to Y$ ($I := [0,1]$) such that $F_0 := F(\cdot, 0) = f$ and $F_1 := F(\cdot, 1) = h$. The map $f : X \to Y$ is a **homotopy-equivalence**, if there is a continuous map $g : Y \to X$ such that $g \circ f \sim \mathrm{Id}_X$ and $f \circ g \sim \mathrm{Id}_Y$ ("$\sim$" means "homotopic"). X and Y are then called **homotopy equivalent**. The set of homotopy classes of maps $X \to Y$ is denoted by $[X, Y]$. X is called **contractible**, if X is homotopy equivalent to a point.

Theorem B.5. *(i) If $f : X \to Y$ is a homotopy equivalence, then the transformation $f^* : \mathrm{Vect}(Y) \to \mathrm{Vect}(X)$ of Exercise B.1d is bijective. (Assume X and Y compact.)*
(ii) If X is contractible, then every bundle over X is trivial, and $\mathrm{Vect}(X)$ is isomorphic to the non-negative integers.

Proof. (ii) follows easily from (i), since $\mathrm{Vect}_N(P)$ consists only of the isomorphism class of the trivial bundle of dimension N, in the case where P is a point. (i) follows from the fact that $F_0^* E \cong F_1^* E$, if $F : X \times I \to Y$ is a homotopy of f and E is a vector bundle over Y. We give a proof of this in three steps:

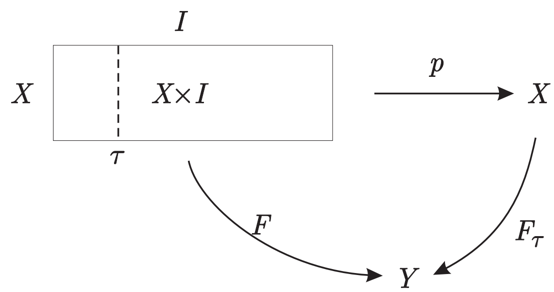
Step 1: Let H be a vector bundle over $X \times I$, fix a $\tau \in I$, and let $s \in C^0(H|_{X \times \{\tau\}})$. We show that s can be extended to a section $S \in C^0(H)$ with $S(\cdot, \tau) = s$. Since a section of a vector bundle can be regarded locally as a graph of a continuous vector-valued function, one can locally apply the Tietze Extension Theorem [**Du**]: We can find a neighborhood U about each (x, τ), and a section $t \in C^0(H|_U)$ such that t and s coincide on $(X \times \{\tau\}) \cap U$.



By the compactness of X, we can obtain a finite system $\{U_j\}$, $\{t_j\}$ with $X \times \{\tau\} \subset \cup U_j$. If $\{\phi_j\}$ is a C^0 partition of unity subordinate to $\{U_j\}$ (see also Theorem 6.4, p. 157), then we set

$$s_j := \begin{cases} t_j \phi_j & \text{on } U_j \\ 0 & \text{on } (X \times I) \setminus U_j. \end{cases}$$

By construction $s_j \in C^0(H)$; hence, $\sum s_j \in C^0(H)$ is a well-defined extension of s.
Step 2: From Step 1, we conclude that two vector bundles G and G' over $X \times I$ which are isomorphic over $X \times \{\tau\}$, are also isomorphic in a neighborhood of $X \times \{\tau\}$: Each $s \in \operatorname{Iso}(G|_{X\times\{\tau\}}, G'|_{X\times\{\tau\}})$ be regarded as a section of $H|_{X\times\{\tau\}}$, where H denotes the bundle $\operatorname{Hom}(G, G')$ of linear maps from fibers of G to corresponding fibers of G'. Let $S \in C^0(H)$ be an extension of s. Then the set $W := \{z \in X \times I : S_z \in \operatorname{Hom}(G, G') \text{ is bijective}\}$ is open in $X \times I$ (by the classical zero determinant argument) and contains all of $X \times \{\tau\}$ by construction. Since the inverse map of $\mathrm{GL}(N, \mathbb{C})$ is continuous, it follows that the mapping $z \to S_z^{-1}$ is continuous, and hence a bundle isomorphism is defined on W.
Step 3. We now set $G := F^*E$ and $G' := p^*(F_\tau)^*E$, where $F_\tau(x) := F(x, \tau)$ and $p : X \times I \to X$ is the projection.



By Exercise B.1d, G and G' are isomorphic over $X \times \{\tau\}$, and by Step 2, they are also isomorphic in a whole neighborhood, which we can take to be a strip $X \times \delta(\tau)$, by the compactness of X. For all $\rho \in \delta(\tau)$, we then have $(F_\rho)^*E \cong (F_\tau)^*E$. Since the unit interval I is compact and connected, we obtain that the isomorphism classes of $(F_\tau)^*E$ do not depend on τ. □

Remark B.6. The statements proved in the first two steps of the proof, also apply to more general situations, and are occasionally formulated as independent theorems; e.g., see [**AB64b**, p.233f] or [**Ati67a**, p.16]. One can also express the result of our proof somewhat more generally (we write $X \times Z$ instead of $X \times I$): Each vector bundle E over the topological space $X \times Z$ (X and Z compact) can be regarded as a "continuous" family of vector bundles E over X where the parameter z is in Z, and the isomorphism classes of E in $\operatorname{Vect}(X)$ are locally constant.

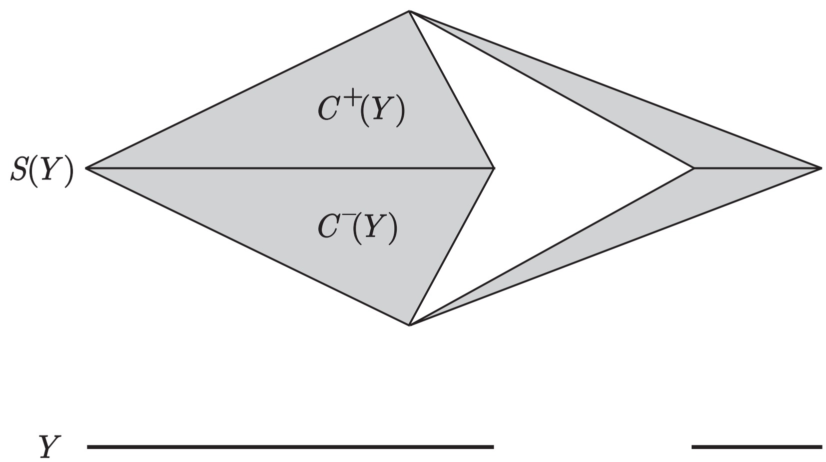
3. Clutching Construction and Suspension

Vector bundles are often given via a **clutching** or **gluing construction**: Let $X = X_1 \cup X_2$ and $A = X_1 \cap X_2$, where all spaces are compact. Let E_i be a vector bundle over X_i and $\phi : E_1|A \to E_2|A$ an isomorphism. Then we define the vector bundle $E_1 \cup_\phi E_2$ over X as follows. As a topological space $E_1 \cup_\phi E_2$ is the quotient space of the disjoint union $E_1 + E_2$ under the equivalence relation which identifies $e_1 \in E_1|A$ with $\phi(e_1) \in E_2|A$. If we regard X as the corresponding quotient space of $X_1 + X_2$, then we obtain a natural projection $p : E_1 \cup_\phi E_2 \to X$ and $p^{-1}(x)$ has a natural vector space structure for each $x \in X$.

Exercise B.7. a) Show that $E_1 \cup_\phi E_2$ is a vector bundle.
b) Show that the isomorphism class of $E_1 \cup_\phi E_2$ depends solely on the homotopy class of the isomorphism $\phi : E_1|A \to E_2|A$.

[Hint for a): It remains only to show that $E_1 \cup_\phi E_2$ is locally trivial. Outside of A, this is clear. In order to extend a trivialization of E_1 on a neighborhood $V_1 \subseteq X$ of a point $a \in A$ to a trivialization of E_2 over V_2 with $a \in V_2 \subseteq X_2$, argue as in Step 2 of Theorem B.5. See also [**AB64b**, p.235] or [**Ati67a**, p.21]. For b): Reduce to Theorem B.5. See also the given sources.]

We now give $\mathrm{Vect}_N(X)$ a homotopy-theoretic interpretation, when X can be represented as the suspension $S(Y)$ of another space Y.[1] Here the **suspension** $S(Y)$ is the union of two cones over Y. Thus, we write $S(Y) := C^+(Y) \cup C^-(Y)$, where $C^+(Y) := Y \times [0, \frac{1}{2}]/Y \times \{0\}$ and $C^-(Y) := Y \times [\frac{1}{2}, 1]/Y \times \{1\}$. Then $Y = C^+(Y) \cap C^-(Y)$. We note that the suspension $S(S^n)$ of the n-sphere is homeomorphic to the $(n+1)$-sphere S^{n+1}.



THEOREM B.8. *The clutching of trivial bundles over $C^+(Y)$ and $C^-(Y)$ defines a natural isomorphism* $[Y, \mathrm{GL}(N,\mathbb{C})] \overset{\cong}{\to} \mathrm{Vect}_N(S(Y))$.

PROOF. (i) Each $l : Y \to GL(N,\mathbb{C})$ yields a bundle over SY via the clutching of the N-dimensional trivial bundles over the two cones, and homotopic maps l_0 and l_1 yield isomorphic bundles; see the proof of Theorem B.5(i). (ii) Conversely, we have the composition

$$\mathrm{Vect}_N(SY) \to \mathrm{Vect}_N(C^-Y) \oplus \mathrm{Vect}_N(C^+Y) \to [Y, \mathrm{GL}(N,\mathbb{C})],$$

the left arrow is given by the restrictions of the bundle, where one obtains trivial bundles since $C^\pm Y$ are contractible (see Theorem B.5(ii).) If $\alpha^\pm$ are such trivializations, then the right arrow is defined by taking the homotopy class of $(\alpha^+|_Y)\,(\alpha^-|_Y)^{-1} : Y \to \mathrm{GL}(N,\mathbb{C})$, which actually only depends on the homotopy classes of $\alpha^\pm$ and hence only on the isomorphism class in $\mathrm{Vect}_N(SY)$. (iii) By construction the functions given in (i) and (ii) are inverse to each other. □

EXERCISE B.9. Show $H_{\mathbb{C}^2} \cong \mathbb{C}_{B_0} \cup_a \mathbb{C}_{B_\infty}$, where $H_{\mathbb{C}^2}$ is the complex line bundle over $\mathbb{PC}^2 = \mathbb{C} \cup \{\infty\} = S^2 = B_0 \cup B_\infty$ defined in Exercise B.2a; here (z_0, z_1) are homogeneous coordinates for $\mathbb{PC}^2$ with $(0,1) = \infty$, and $z = z_1/z_0$ is the coordinate for $\mathbb{C}$, $B_0 := \{z \in \mathbb{C} : |z| \le 1\}$, and $B_\infty := \{z \in \mathbb{C} : |z| \ge 1\} \cup \{\infty\}$

[1] With the concept of Grassmann manifolds, one can give a homotopy-theoretic definition of $\mathrm{Vect}_N(X)$ for *arbitrary* X; see [**Ati67a**, p.24-30].

(the two canonical hemispheres of S^2). Finally, $a : z \mapsto z$, is the standard map $S^1 \to \mathbb{C}^\times := \{\mathbb{C} \setminus \{0\}\} = \mathrm{GL}(1, \mathbb{C})$.

Exercise B.10. Show that for each bundle E, there is a bundle F such $E \oplus F$ is trivial. [Hint: Show first with the help of a finite open cover of the compact parameter space X and a suitable partition of unity, that $C^0(E)$ contains an "ample" subspace; i.e., a subspace $V \subseteq C^0(E)$ such that each point of E is in the image of a section $s \in V$. If $\dim V = N$, then we have an epimorphism $\phi : X \times \mathbb{C}^N \to E$, and consequently there is an isomorphism $E \oplus F \to X \times \mathbb{C}^N$ where F is the kernel bundle of ϕ; see also [**Ati67a**, p.26 f] or the related technique in the proof of our Embedding Theorem 6.7c, p.162.]

Exercise B.11. Let X be a topological space which possesses in addition the structure of a C^∞ manifold of dimension n (see Chapter 6).
a) Show that the tangent bundle TX is a (real) vector bundle over X; do the same for the "normal bundle" NX, when X is a submanifold of a Riemannian manifold Y.
b) When may one call a continuous vector bundle over X, whose total space is a C^∞ manifold, a C^∞ vector bundle? Show that each (continuous) vector bundle over X is isomorphic to a C^∞ vector bundle.
[Hint for a): First investigate the case $X = S^1$ and show that TS^1 is isomorphic to the real line bundle defined in Exercise B.1b. In the general case, this "direct" method is also possible, since one can (see Theorem 6.7c, p. 160) embed X into a high dimensional Euclidean space, and thus realize TX as a (real) subbundle of a higher dimensional trivial bundle. It is easier (especially since TX is in general not trivial; e.g., for $X = S^2$, see Exercise 11.23, p. 284) to carry out a local analysis, where each chart u from the C^∞ atlas yields a local trivialization of TX via the differential forms $(du_1, ..., du_n)$; see Chapter 6. See also the discussion below.
For b): For the definition of C^∞ vector bundles, see also [**BJ**, Ch.3], and for the topological equivalence of the categories of C^0 and C^∞ vector bundles, see the Whitney Approximation Theorem in [**BJ**, p.66]. Details of the argument are in [**Hir**, p.101], where it is shown that one can make E itself into a C^∞ vector bundle.]

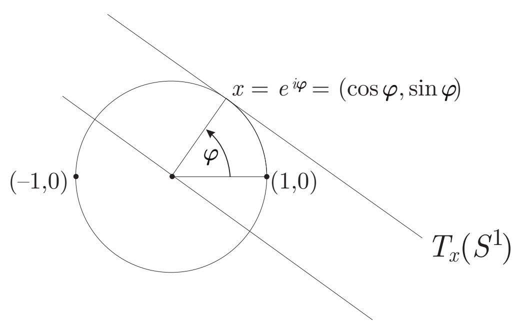
Remark B.12. From Exercise B.11b, it follows (without loss of generality) that we only need to investigate C^∞ vector bundles, if the base is C^∞. Exercise B.10 does not say that we always encounter trivial bundles (compare with the analogous - but deeper embedding theorem for manifolds, Theorem 6.7c, p. 160): Roughly, the carrying along of additional "irrelevant" parameters is not only redundant, but can also produce so much "noise" that this noise destroys the structure of the problem or renders it unrecognizable. This is the case with the index problem for elliptic operators, whose solution consists exactly in distinguishing certain vector bundles generated by the symbol of the operator; see Part III.

The idea of a vector bundle originates in the analysis of non-Euclidean manifolds. Otherwise, according to Theorem B.5(ii), there are only trivial bundles, since Euclidean space is contractible. As an example, let $c : I \to X$ be a differentiable path in a manifold X. Classical mechanics considers the velocity vector $\dot{c}(t)$ for $t \in I$. For the physicist it was always clear (by reasons of physics) how $c(t)$ is multiplied by a scalar, how $\dot{c}_1(t_1)$ and $\dot{c}_2(t_2)$ are added or how the equality of $\dot{c}_1(t_1)$

and $\dot{c}_2(t_2)$ is checked when $c_1, c_2 : I \to X$. are two paths in X with $c_1(t_1) = c_2(t_2)$. From the point of view of physics no confusion between a velocity vector and a position vector was conceivable, but the earliest mathematical abstractions could not express the difference: If the position vectors $c(t)$ are represented as triples of real numbers $(c_1(t), c_2(t), c_3(t))$, then the velocity vector is $(\dot{c}_1(t), \dot{c}_2(t), \dot{c}_3(t))$ where $\dot{c}_i(t)$ is the derivative of $\dot{c}_i$ at t. Thus, $c(t)$ and $\dot{c}(t)$ are both elements of the single vector space $\mathbb{R}^3$. This low level of abstraction was fully sufficient as long as X was Euclidean space (actually an affine space – but choose a base point and call it 0). In this case there is indeed a natural interpretation of the velocity vector $\dot{c}(t)$ at the space point $c(t)$ as a velocity vector $\dot{\widetilde{c}}(t)$ at the point $\widetilde{c}(t) = 0$. In fact, consider the translated path $\widetilde{c} : I \to X$ with $\widetilde{c}(\tau) := c(\tau) - c(t)$, $\tau \in I$. This means that the tangent spaces at the various points of X can be identified canonically (via the retraction $r : X \to \{0\}$) with the tangent space at the point 0. In the language of vector bundles, we could say $TX \cong r^*(TX|_{\{0\}})$. Here TX is the, totality of all velocity vectors, the fiber T_xX is $\mathbb{R}^3$, the restricted bundle $TX|_{\{0\}}$ is precisely the space $\{0\} \times \mathbb{R}^3 \cong \mathbb{R}^3$ and the induced bundle $r^*(\{0\} \times \mathbb{R}^3)$ is the trivial bundle $X \times \mathbb{R}^3$.

Let us next consider the case that X is a submanifold of a Euclidean space Y, say the 2-sphere in 3-space. Every velocity vector $\dot{c}\,(t)$ can be considered as an element of the tangent space of Y at the point $c(t)$ and hence an element of the tangent space of Y at 0 (due to the translation described above), i.e., as n-tuple of derivatives of the coordinate functions of c. Already by reason of dimensions it is clear that, in general, the tangent space T_xX of all velocity vectors X of paths through x cannot be identified with the full tangent space T_0Y but only with some subspace and it may be different for each x

Example B.13. $X = S^1$ and $Y = \mathbb{R}^2$



If we follow the translation with a rotation through the angle ϕ, then we have identified all tangent spaces T_xS^1 with the same subspace of $T_0\mathbb{R}^2$ or with the tangent space $T_{(0,1)}S^1$. We write $TS^1 = q^*(TS^1|_{\{(1,0)\}})$, where $q : S^1 \to (1,0)$ is the retraction and $TS^1|\,\{(1,0)\}$ can be identified with $\{(1,0)\} \times \mathbb{R}$. Consequently $TS^1 \cong S^1 \times \mathbb{R}$, i.e., the tangent bundle of S^1 is trivial.

In analytic terms this circumstance is expressed by saying that there exists a nowhere vanishing tangential vector field on S^1; i.e., at each point a non-vanishing velocity vector can be chosen and in a continuous fashion. For example choose at $x = (\cos\theta, \sin\theta)$ the unit velocity vector $\dot{c}(\phi/2\pi)$, where $c : I \to X$ is given by $c(t) = (\cos 2\pi t, \sin 2\pi t)$. Usually we write $\frac{d}{d\theta}\big|_x$ instead of $\dot{c}(\phi/2\pi)$. Then $\frac{d}{d\theta}$ is the nowhere vanishing vector field and it defines an isomorphism $TS^1 \cong S^1 \times \mathbb{R}$. The isomorphism is, for all $x \in S^1$, given by the map $T_xS^1 \to \mathbb{R}$ which assigns to a velocity vector the value λ, if it is equal to $\lambda \cdot \frac{d}{d\theta}\big|_x$.

EXAMPLE B.14. $X = S^2$ and $Y = \mathbb{R}^3$. Here the situation is different. There is no canonical way of identifying all tangent spaces. In other words, the tangent bundle TS^2 is non-trivial. If not there would exist two tangential vector fields on S^2 which are linearly independent at each point of S^2. But on S^2 there is not even a single vector field that vanishes nowhere (Exercise 11.23, p. 284). This example shows that it might be useful (indeed necessary) to distinguish tangent spaces at different points. As one goes on to consider manifolds without an explicitly given imbedding in Euclidean spaces, as in physics with the theory of relativity, the notion of a bundle becomes indispensable.

The modern concept of a bundle evolved from the topology and geometry of manifolds as practiced by Heinz Hopf and others since the 1920's. In the 1950's the notion was precisely formulated, and the classification of bundles and their systematic employment in deep problems of geometry and analysis started. (For the theory of characteristic classes, see e.g., the concise [**GIN**] *or the more elaborated* [**Hi66a**, p.49 f], [**KN69**, Chapter XII], *and Section 16.16.7.) Crudely expressed, the success of these methods derives from their utilizing given or manufactured "classical" structures (such as the tangent bundles or bundles of differential forms) on the manifolds under discussion to the greatest extent possible and thus shifting the plane of study from manifolds – which are conceptually simpler but harder to understand – to vector bundles which are more easily analyzed.*

Differently put, while for topological manifolds and other "triangulable" spaces, one depends at first on the combinatorial methods of the analysis of "cell decompositions", and while for differentiable manifolds only the group of diffeomorphisms is a priori *available for investigation, vector bundles offer more opportunity for manipulations because of their richer structure. Principally, large parts of linear algebra can be used directly. (Incidentally, these play an important role also for the other method, albeit under the surface.) For example (see above Exercise B.4) one can perform linear constructions with vector bundles such as forming direct sums and quotients, which is impossible to do with manifolds. Also, "clutching functions" which can be used to build complicated manifolds from simpler ones (see e.g., Exercise 6.46, p. 188), i.e., the diffeomorphisms, become linear only in the first derivative ("functional matrix" or "Jacobian"). In contrast, Theorem B.5 shows how much (via linear clutching functions) the topology of vector bundles can be reduced to the geometry of the matrix spaces of linear algebra.*

At the start of the 1960's linear algebra had "matured" enough with the Periodicity Theorem (for the (stable) homotopy groups of invertible matrices) discovered by Raoul Bott just before. Michael Francis Atiyah and Friedrich Hirzebruch extracted from these methods an abstract formalism – K-theory – which they developed as a generalized cohomology theory, using stability classes of vector bundles; see Chapter 11.

Bibliography

[AM] ABRAHAM, R. AND J. MARSDEN, *Foundations of Mechanics*, Benjamin-Cummings, Reading MA, 1978.

[Ad] ADAMS, R.A., *Sobolev Spaces*, Academic Press, New York, 1975.

[ADN] AGMON, S., A. DOUGLIS AND L. NIRENBERG, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II, *Comm. Pure Appl. Math.* **17** (1964), 35-92.

[Ag] AGRANOVICH, M.S., Elliptic singular integro-differential operators (translation from *Uspehi Mat. Nauk* **20** (1965), no. 5 (125), 3-120), *Russian Math. Surveys* **20** (1965), no. 5/6, 1-121.

[Ah] AHLFORS, L., *Complex Analysis*, McGraw-Hill, New York, 1953.

[AkGl] AKHIEZER, N.I. AND I.M. GLAZMAN, *Theory of Linear Operators in Hilbert Space*, two volumes bound as one, Dover Publications Inc., New York, 1993 (translation from Gosudarstv. Izdat. Techn.-Teor. Lit., Moscow–Leningrad, 1950).

[Alb] ALBERT, A., *Introduction to Algebraic Theories*, Univ. of Chicago Press, Chicago, 1941.

[AH] Alexandroff, P. and H. Hopf, *Topologie* I, Springer-Verlag, Berlin, 1935.

[Ar] ARONSZAJN, N., A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of the second order, *J. Math. Pures. Appl.* (9) **36** (1957), 235-249.

[Ati] ATIYAH, M.F., *Collected Works*, 5 vol., The Clarendon Press, Oxford University Press, Oxford, New York, 1988.

[Ati62] —, Harmonic spinors and elliptic operators. Workshop lecture, Bonn, 1962.

[Ati63a] —, The Grothendieck ring in geometry and topology, *Proc. Int. Congr. Math.* (Stockholm 1962), Uppsala (1963), 442-446; in [**Ati**], vol. 2, pp. 293-297.

[Ati63b] —, The index of elliptic operators on compact manifolds, *Sém. Bourbaki, Exp.* **253**, Paris (1963); in [**Ati**], vol. 3, pp. 13-24.

[Ati66a] —, The Euler characteristic theorem I. Arbeitstagung lectures, Vervielfältigt, Bonn, 1966.

[Ati66b] —, The role of algebraic topology in mathematics, *J. Lond. Math. Soc.* **41** (1966), 63-69; in [**Ati**], vol. 1, pp. 189-196.

[Ati67a] —, *K-Theory*, Benjamin, New York, (Russ. transl. Moscow 1967); in [**Ati**], vol. 2, pp. 431-603.

[Ati67b] —, Algebraic topology and elliptic operators, *Comm. Pure Appl. Math.* **20** (1967), 237-249; in [**Ati**], vol. 3, pp. 57-71.

[Ati68a] —, Bott periodicity and the index of elliptic operators, *Quart. J. Math., Oxford* (2) **19** (1968), 113-140; in [**Ati**], vol. 2, pp. 605-633.

[Ati68b] —, Global aspects of the theory of elliptic differential operators, *Reports of the Internat. Congress of Mathematics* (Moscow 1966), Verlag Mir, Moscow, 1968, 57-64; in [**Ati**], vol. 3, pp. 73-81.

[Ati69] —, Algebraic topology and operators in Hilbert space, In: C. T. Tamm (ed.), *Lectures in Modern Analysis and Applications I*, Lecture Notes in Mathematics 103, Springer-Verlag, Berlin-Heidelberg-New York, 1969, 101-121.

[Ati70a] —, Topology of elliptic operators, *Amer. Math. Soc. Symp. Pure Math.* **16** (1970), 101-119; in [**Ati**], vol. 3, pp. 387-407.

[Ati70b] —, Global theory of elliptic operators, *Proc. Int. Conf. Functional Analysis*, Tokyo, 1970, 21-50; in [**Ati**], vol. 3, pp. 337-347.

[Ati70c] —, Vector fields on manifolds, *Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen* **200** (1970); in [**Ati**], vol. 2, pp. 727-749.

[Ati70d] —, Elliptic operators and singularities of vector fields, *Actes, Congres Intern. Math.* (Nice 1970), Tome 2, 1970, 207-209; in [**Ati**], vol. 2, pp. 803-807.

[Ati71] —, Riemann surfaces and spin structures, *Ann. Sci. École Norm Sup.* (4) **4** (1971), 47-62; in [**Ati**], vol. 3, pp. 431-447.

[Ati74] —, *Elliptic operators and compact groups*, Lecture Notes in Mathematics 401, Springer-Verlag, Berlin-Heidelberg-New York, 1974; in [**Ati**], vol. 3, pp. 499-593.

[Ati75a] —, Eigenvalues and Riemannian geometry. In: *Proc. Internat. Conf. on Manifolds and Related Topics in Topology* (Tokyo 1973), Univ. Tokyo Press, Tokyo, 1975, 5-9; in [**Ati**], vol. 4, pp. 333-339.

[Ati75b] —, *Classical groups and classical differential operators on manifolds*, Centro Internazionale Matematico Estivo (Varenna 1975), Edizioni Cremonese, Rome, 1975, 5-48; in [**Ati**], vol. 4, pp. 341-385.

[Ati76a] —, Elliptic operators, discrete groups and von Neumann algebras, *Soc. Math. de France Astérisque* **32/33** (1976), 43-72; in [**Ati**], vol. 4, pp. 387-417.

[Ati76b] —, Trends in pure mathematics. In: *Proc. of the 3rd Internat. Congress on Mathematical Education*, Karlsruhe, 1976, 61-74; in [**Ati**], vol. 1, pp. 261-265.

[Ati79] —, *Geometry of Yang-Mills Fields*, Lezioni Fermaine, Accademia Nazionale dei Lincei & Scuola Normale Superiore, Pisa (1979); in [**Ati**], vol. 5, pp. 75-173.

[AB64a] Atiyah, M.F. and R. Bott, The index problem for manifolds with boundary, *Coll. Differential Analysis*, Tata Institute, Bombay, Oxford University Press, Oxford, 1964, 175-186; in [**Ati**], vol. 3, pp. 25-37.

[AB64b] —, —, On the periodicity theorem for complex vector bundles, *Acta Math.* **112** (1964), 229-247; in [**Ati**], vol. 2, pp. 337-357.

[AB65] —,—, Notes on the Lefschetz fixed point theorem for elliptic complexes (Russ. transl. Moscow, 1966), Notes, Harvard University, 1965.

[AB66] —,—, A Lefschetz fixed point formula for elliptic differential operators, *Bull. Amer. Math. Soc.* **72** (1966), 245-250; in [**Ati**], vol. 3, pp. 83-89.

[AB67] —,—, The Lefschetz fixed point theorem for elliptic complexes I, *Ann. of Math.* **86** (1967), 374-407; in [**Ati**], vol. 3, pp. 91-125.

[AB68] —,—, The Lefschetz fixed point theorem for elliptic complexes II, *Ann. of Math.* **88** (1968), 451-491; in [**Ati**], vol. 3, pp. 127-169.

[ABP] Atiyah, M.F., R. Bott and V.K. Patodi, On the heat equation and the index theorem (Russ. transl., Moscow, 1973), *Invent. Math.* **19** (1973), 279-330 (and Errata **28** (1975), 277-280); in [**Ati**], vol. 4, pp. 11-63.

[ADHM] Atiyah, M.F., V.G. Drinfield, N.J. Hitchin and Y.I. Manin, Construction of instantons, *Phys. Lett.* **65A** (1978), 185-187; in [**Ati**], vol. 5, pp. 21-25.

[AD] Atiyah, M.F. and J.L. Dupont, Vector fields with finite singularities, *Acta Math.* **128** (1972), 1-40; in [**Ati**], vol. 2, pp. 761-802.

[AH59] Atiyah, M.F. and F. Hirzebruch, Riemann-Roch theorems for differentiable manifolds, *Bull. Amer. Math. Soc.* **65** (1959), 276-281; in [**Ati**], vol. 2, pp. 7-13.

[AH61] —,—, Vector bundles and homogeneous spaces, *Amer. Math. Soc. Symp. Pure Math.* **3** (1961), 7-38; in [**Ati**], vol. 2, pp. 51-83.

[AH62] —,—, Analytic cycles on complex manifolds, *Topology* **1** (1962), 25-46; in [**Ati**], vol. 2, pp. 249-271.

[AHS1] Atiyah, M.F., N.J. Hitchin and I.M. Singer, Deformation of instantons, *Proc. Nat. Acad. Sci.* **74** (1977), 2662-2663; in [**Ati**], vol. 5, pp. 7-9.

[AHS2] —,—,—, Self-duality in four dimensional Riemannian geometry, *Proc. Roy. Soc. London* **A362** (1978), 425-461; in [**Ati**], vol. 5, pp. 27-65.

[APS73] Atiyah, M.F., V.K. Patodi and I.M. Singer, Spectral asymmetry and Riemannian geometry, *Bull. London Math. Soc.* **5** (1973), 229-234; in [**Ati**], vol. 4, pp. 71-79.

[APS75] —,—,—, Spectral asymmetry and Riemannian geometry I/II/III. *Math. Proc. Camb. Phil. Soc.* **77** (1975a); 43-69/**78** (1975b); 405-432/**79** (1976), 71-99; in [**Ati**], vol. 4, pp. 81-169.

[ASe] Atiyah, M.F. and G.B. Segal, The index of elliptic operators II, *Ann. of Math.* **87** (1968), 531-545; in [**Ati**], vol. 3, pp. 223-237.

[AS63] Atiyah, M.F. and I.M. Singer, The index of elliptic operators on compact manifolds, *Bull. Amer. Math. Soc.* **69** (1963), 422-433; in [**Ati**], vol. 3, pp. 11-24.

[AS68a] —,—, The index of elliptic operators I, *Ann. of Math.* **87** (1968), 484-530; in [**Ati**], vol. 3, pp. 171-219.

[AS68b] —,—, The index of elliptic operators III, *Ann. of Math.* **87** (1968), 546-604; in [**Ati**], vol. 3, pp. 239-299.

[AS69] —,—, Index theory for skew-adjoint Fredholm operators, *Publ. Math. Inst. Hautes Etudes Sci.* **37** (1969), 305-325; in [**Ati**], vol. 3, pp. 349-372.

[AS71a] —,—, The index of elliptic operators IV, *Ann. of Math.* **93** (1971), 119-138; in [**Ati**], vol. 3, pp. 301-322.

[AS71b] —,—, The index of elliptic operators V, *Ann. of Math.* **93** (1971), 139-149; in [**Ati**], vol. 3, pp. 323-335.

[AW] Atiyah, M.F. and R.S. Ward, Instantons and algebraic geometry, *Commun. Math. Phys.* **55** (1977), 117-124; in [**Ati**], vol. 5, pp. 11-20.

[Atk] Atkinson, F.W., The normal solubility of linear equations in normed spaces, *Mat. Sb.* **28** (70) (1951), 3-14 (Russian).

[Avr] Avramidi, I.G., *Heat Kernel and Quantum Gravity*, Lecture Notes in Physics. New Series m: Monographs, 64. Springer-Verlag, Berlin, 2000. x+149 pp.

[BFI] Bauer, W., K. Furutani and C. Iwasaki, Spectral analysis and geometry of a sub-Laplacian and related Grushin type operators. In: *Partial Differential Equations and Spectral Theory*, Operator Theory: Advances and Applications **211**, Springer 2011, pp. 183-287.

[BPST] Belavin, A., A.M Polykov, A.S. Schwarz and Y.S. Tyupkin, Pseudo-particle solutions of the Yang-Mills equations, *Phys. Lett.* **59B** (1975), 85-87.

[BeCo] Bellman, R. and K.L. Cooke, *Modern Elementary Differential Equations*, Addison-Wesley, Reading, 1971.

[BK] Bellman, R. and R. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems*, American Elsevier, New York, 1965.

[Ben] Benedetto, J.J., *Spectral synthesis*, Teubner, Stuttgart, 1975.

[Berg] Berger, M., *A Panoramic View of Riemannian Geometry*, Springer-Verlag, Berlin-Heidelberg-New York, 2003.

[BeGeVe] Berline, N., E. Getzler and M. Vergne, *Heat Kernels and Dirac Operators*, Springer-Verlag, Berlin-Heidelberg-New York, 1992.

[Ber] Bernard, C.W., in *Geometrical and Topological Methods in Gauge Theories*, Springer-Verlag, New York, 1980.

[BCGW] Bernard, C.W., N.H. Christ, A.H. Guth and E.J. Weinberg, Instanton parameters for arbitrary gauge groups, *Phys. Rev.* **D16** (1977), 2967-2977.

[BJS] Bers, L., F. John, and M. Schechter, *Partial differential equations* (Russ. translation, Moscow, 1966), Interscience Publishers, New York-London-Sidney, 1964.

[Bl81] Bleecker, D., *Gauge Theory and Variational Principles*, Addison-Wesley, Reading Mass., 1981.

[Bl08] —, Index theory, in: D. Krupka and D. Saudners (eds.), *Handbook of Global Analysis*, Elsevier, 2008, pp. 75-145.

[BlBo85] Bleecker, D. and B. Booss–Bavnbek, *Topology and Analysis*, Springer-Verlag, Berlin-Heidelberg-New York, 1985.

[BlBo03] —,—, Spectral invariants of operators of Dirac type on partitioned manifolds, (in *Aspects of Boundary Problems in Analysis and Geometry* [Edited by J. B. Gil, T. Krainer and I. Witt]), *Operator Theory: Advances and Applications* **143**, Birkhäuser Verlag, Basel (2003), 1-125.

[BlCs] Bleecker, D. and G. Csordas, *Basic Partial Differential Equations*, International Press, Boston, 1996.

[Boc] Bochner, S., Vector fields and Ricci curvature, *Bull. Amer. Math. Soc.* **52**, 1946, 776-797.

[Boh] Bohn, M., On Rho invariants of fiber bundles, 263 pages, 8 figures, doctoral thesis, 2009, arXiv:0907.3530v1 [math.GT]. ***

[Boj] Bojarski, B.W., On the index problem for systems of singular integral equations, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **11** (1963), 653-655.

[Bok] Bokobza-Haggiag, J., Opérateurs pseudo-différentiels sur une variété différentiable, *Ann. Inst. Fourier* (Genoble) **19** (1969), 125-177.

[Bo72] Booss, B., *Elliptische Topologie von Transmissionsproblemen*, Bonner Mathematische Schriften no. 58, Bonn, 1972.

[Bo02] Booss–Bavnbek, B., *The Determinant of Elliptic Boundary Problems for Dirac Operators. Reader on the Scott–Wojciechowski Theorem*, http://milne.ruc.dk/~Booss/ell/determinant.pdf.

[BCLZ] Booss–Bavnbek, B., G. Chen, M. Lesch and C. Zhu, Perturbation of sectorial projections of elliptic pseudo-differential operators, *J. Pseudo-Differ. Oper. Appl.*. To appear. Published online: 14 October 2011, DOI: 10.1007/s11868-011-0042-5, 31 pages, arXiv:1101.0067 [math.SP].

[BoFu1] Booss–Bavnbek, B. and K. Furutani, The Maslov index – a functional analytical definition and the spectral flow formula, *Tokyo J. Math.* **21** (1998), 1–34.

[BoFu2] —, —, Symplectic functional analysis and spectral invariants, In: B. Booß–Bavnbek, K.P. Wojciechowski (eds.), *Geometric Aspects of Partial Differential Equations*, Amer. Math. Soc. Series *Contemporary Mathematics*, vol. 242, Providence, R.I., 1999, pp. 53–83.

[BLP] Booss–Bavnbek, B., M. Lesch and J. Phillips, Unbounded Fredholm operators and spectral flow, *Canad. J. Math.* **57**, no. 2 (2005), 225–250, arXiv:0108014 [math.FA].

[BoWo] Booss–Bavnbek, B. and K.P. Wojciechowski, *Elliptic Boundary Problems for Dirac Operators*, Birkhäuser, Boston, 1993.

[Bott] Bott, R., Homogeneous differential operators. In: *Differential and Combinatorial Topology* (S. S. Cairns, ed.), Princeton Univ. Press, Princeton 1965, 167-186.

[Bou] Boutet de Monvel, L., Boundary problems for pseudo-differential operators, *Acta Math.* **126** (1971), 11-51.

[BrGi] Branson, Th.P. and P.B. Gilkey Residues of the eta function for an operator of Dirac type, *J. Funct. Anal.* **108** (1992), 47-87.

[Brd] Bredon, G. E., *Topology and Geometry*, Springer–Verlag, New York, 1993.

[Bre] Breuer, M., Fredholm theories in von Neumann algebras I/II., *Math. Ann.* **178** (1968), 243-254, and **180** (1969), 313-325.

[Bri63] Brieskorn, E. et al., Die Atiyah-Singer-Indexformel, Seminarvorträge, Bonn, 1963.

[Bri74] —, Über die Dialektik in der Mathematik. In: M. Otte (ed.) Mathematiker über die Mathematik, Springer-Verlag, Berlin-Heidelberg-New York, 1974, 221-285.

[Bri76] —, The development of geometry and topology, Notes of introductory lectures given at the University of La Habana in 1973. M. z. Berufspraxis Math. Heft 17, Bielefeld 1976, 109-203.

[BN] Brill, A. and M. Noether, Die Entwicklung der Theorie der algebraischen Funktionen in älterer und neuerer Zeit. *Jber. Deutsch. Math.-Verein.* **3** (1892/93), 107-566.

[BJ] Bröcker, Th. and K. Jänich, *Introduction to Differential Topology* (transl. from German), Cambridge Univ. Press, New York, 1982.

[BDF] Brown, L.G., R.G. Douglas and P.A. Fillmore, Unitary equivalence modulo the compact operators and extensions of C^*-algebras, In Fillmore, P. A. (ed.), *Proceedings of a Conference on Operator Theory*, Lecture Notes in Mathematics 345, Springer-Verlag, Berlin-Heidelberg-New York, 1973.

[BruLes] Brüning, J. and M. Lesch On boundary value problems for Dirac type operators. I. Regularity and self-adjointness, *J. Funct. Anal.* **185**/1 (2001), 162.

[Cald63] Calderon, A.P., Boundary value problems for elliptic equations. In: *Outlines of the Joint Soviet-Amer. Symp. Partial Diff. Equations* (Novosibirsk, 1963), Acad. Sci. USSR, Siberian Section, Moscow, 1963, pp. 303-304.

[Cald67] —, The analytic calculation of the index of elliptic equations, *Proc. Nat. Acad. Sci.* **57** (1967), 1193-1194.

[CZ] Calderon, A.P. and A. Zygmund, On singular integrals, *Amer. J. Math.* **78** (1956), 289-309.

[Calk] Calkin, J.W., Two-sided ideals and congruence in the ring of bounded operators in Hilbert space, *Ann. of Math.* **42** (1941), 839-873.

[Can] Cantor, M., Elliptic operators and the decomposition of tensor fields, *Bull. Amer. Math. Soc.* **5** (1981), 235-262.

[CaSc] Cartan, H. and L. Schwartz et al., Théorème d'Atiyah-Singer sur l'indice d'un opérateur différentiel elliptique. Séminaire Henri Cartan 16e année 1963/64, École Normale Supérieure, Paris 1965.

[ChSi] Chern, S.S and J. Simons, Characteristic forms and geometric invariants, *Ann. of Math.* **99** (1974), 48-69.

[CDSS] Coburn, L.A., R.G. Douglas, D.G. Schaeffer and I.M. Singer, C*-algebra of operators on a half space II - Index theory, Publ. M*ath. Inst. Hautes Études Sci.* **40** (1971), 69-79.

[CodLev] Coddington, E.A. and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill Book Company, New York, 1955.

[CorLab] Cordes, H.O. and J.P. Labrousse, The invariance of the index in the metric space of closed operators, *J. Math. Mech.* **12**, (1963), 693–719.

[CF] Corrigan, E.F. and D.B. Farley, Scalar field theory and exact solutions to a classical SU (2) Gauge Theory, *Phys. Lett.* **67B** (1977), 69-71.

[CG] Corrigan, E. and Goddard, P., In: *Geometrical and Topological Methods in Gauge Theories*, Springer-Verlag, New York, 1980.

[CD] Coughlan, G.D. and Dodd, J.E., *The Ideas of Particle Physics*, Second Ed., Cambridge Univ. Press, Cambridge, 1991.

[CH] Courant, R. and D. Hilbert, *Methods of Mathematical Physics I and II* (trans. from German), Interscience Publishers, New York, 1953/62.

[De] Devinatz, A., On Wiener-Hopf operators, In: B. Gelbaum (ed.), *Functional Analysis*, Thompson, Washington, 1967, 81-118.

[DeWi] DeWitt, B.S., Quantum gravity: yesterday and today, *General Relativity and Gravitation* **41** (2009), 413–419.

[Di43] Dieudonné, J., Sur les homomorphismes d'espace normés, *Bull. Sci. Math.* (2) **67** (1943), 72-84.

[Di] Dieudonné, J., *Cours de Géométrie Algébrique, Vol. I*: *Aperçu Historique sur le Développement de la Géometrie Algébrique*, Presses Univ. de France, Paris, 1974.

[Dix] Dixmier, J., *C*–Algebras* (transl. from French), North-Holland, Amsterdam, 1977.

[Don83] Donaldson, S.K., An application of gauge theory to the topology of 4-manifolds, *J. Diff. Geom.* **18** (1983), 269-316.

[Don90] —, Polynomial invariants for smooth 4-manifolds, *Topology* **29** (1990), 257-315.

[Don96] —, The Seiberg-Witten equations and 4-manifold topology, *Bull. Amer. Math. Soc.* **33** (1996), 45-70.

[DK] Donaldson, S.K., and P.B. Kronheimer, *The Geometry of Four-Manifolds*, Clarendon Press, Oxford, 1990.

[Do] Douglas, R.G., *Banach Algebra Techniques in Operator Theory*, Academic Press, New York, 1972.

[DM] Drechsler, W. and M.E. Mayer, *Fiber Bundle Techniques in Gauge Theories*, Lecture Notes in Physics, Vol. **67**, Springer-Verlag, New York, 1977.

[Du] Dugundji, J., *Topology*, Allyn and Bacon, Boston, 1966.

[DG] Duistermaat, J.J. and V.W. Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics, *Invent. Math.* **29** (1975), 39-79.

[DunSch] Dunford, N. and J.T. Schwartz, *Linear Operators*, 3 volumes, Interscience Pub., 1963, Reprinted Wiley Classics Library, 1988.

[DM] Dym, H. and H.P. McKean, *Fourier Series and Integrals*, Academic Press, New York, 1972.

[DY] Dynkin, E.B. and A.A. Yushkevich, Sätze und Aufgaben über Markoffsche Prozesse (translation from Russian), Heidelberger Taschenbücher 51, Springer-Verlag, Berlin-Heidelberg-New York, 1969.

[EGH] Eguchi, T., P.B. Gilkey and A.J. Hanson, Gravitation, gauge theories, and differential geometry, *Phys. Reports* **66** (1980), 213-393.

[ES] Eilenberg, S. and N. Steenrod, *Foundations of Algebraic Topology*, Princeton Univ. Press, Princeton, 1952.

[Ei] Eisenhart, L.P., *Riemannian Geometry*, Princeton Univ. Press, Princeton, 1966.

[El] Elizalde, E., *Ten Physical Applications of Spectral Zeta Functions*, Springer, Berlin, 1995.

[En] Enflo, P, A counterexample to the approximation problem in Banach spaces, *Acta Math.* **130** (1973), 309-317.

[Es73] Eskin, G.I., Boundary value problems and the parametrix for systems of elliptic pseudo-differential equations (translation from *Uspehi Mat. Nauk* **28** (1973), 75-116), *Trans. Moscow Math. Soc.* **28** (1973), 74-115.

[Es81] Eskin, G.I., Boundary value problems for elliptic pseudo-differential equations (translation from Russian), *Amer. Math. Soc. Transl.* **52**, 1981.

[Esp98] Esposito, G., *Dirac Operators and Spectral Geometry*, Cambridge Lecture Notes in Physics vol. 12, Cambridge University Press, Cambridge, 1998.

[Esp11] —, *An introduction to quantum gravity*, arXiv:1108.3269v1 [hep-th], 58pp, 2011.

[Fe] Fedosov, B.W., Analytic formulae for the index of elliptic operators, *Trans. Moscow Math. Soc.* **30** (1974), 159-240 (translation from Trudy Moskov Mat. Obshch. **30** (1974), 159-241).

[Fi] Fillmore, P.A. (ed.), *Proceedings of a Conference on Operator Theory*, Lecture Notes in Mathematics 345, Springer-Verlag, Berlin-Heidelberg-New York, 1973.

[Fred] Fredholm, E.I., Sur une classe d'équations fonctionelles, *Acta Math.* **27** (1903), 365-390.

[FU] Freed, D.S. and K.K. Uhlenbeck, *Instantons and Four-Manifolds*, MSRI Pub., Vol. 1, 2nd Ed., Springer-Verlag, New York, 1991.

[Fr] Freedman, M.H., The topology of four-dimensional manifolds, *J. Diff. Geom.* **17** (1982), 357-453.

[FK] Freedman, M.H. and A. Kirby, Geometric proof of Rohlin's theorem: In: *Proc. Symp. Pure Math.*, Vol. **32**, Amer. Math. Soc., Providence, R.I. (1978), pp. 85-97.

[FQ] Freedman, M.H. and F. Quinn, *Topology of 4-Manifolds*, Princeton Univ. Press, Princeton, 1990.

[FM] Friedman, R. and J. Morgan, Algebraic surfaces and four-manifolds: some conjectures and speculations, *Bull. Amer. Math. Soc.* **18** (1988), 1-19.

[FV] Fursaev, D. and D. Vassilevich, *Operators, Geometry, and Quanta - Methods of Spectral Geometry in Quantum Field Theory*, Springer, Berlin, 2011.

[Fu99] Furuta, M., *Index theorem. 1*, Translated from the 1999 Japanese original by Kauro Ono. Translations of Mathematical Monographs, 235. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 2007. xviii+205 pp.

[Fu01] —, Monopole equation and the $\frac{11}{8}$-conjecture, *Math. Res. Lett.* **8**/3 (2001), 279-291.

[Fur] Furutani, K., On the Quillen determinant, *J. Geom. Phys.* **49** (2004), 366-375.

[Ge] Gelfand, I.M., On elliptic equations (translation from *Uspehi Mat. Nauk* **15**, 3 (1960), 121-132), *Russ. Math. Surveys* **15**, no. 3 (1960), 113-123.

[GRS] Gelfand, I.M., D.A. Raikov and G.E. Shilov, *Kommutative normierte Algebren*, VEB Deutscher Verlag der Wissenshaften (translation from Russian), Berlin, 1964.

[Get] Getzler, E., Pseudodifferential operators on supermanifolds and the Atiyah-Singer index theorem, *Comm. Math. Phys.* **92** (1983), 163-178.

[GML] Ghorbani, H., D. Musso and A. Lerda, Stringy instanton effects in $\mathcal{N} = 2$ gauge theories, arXiv:1012.1122v1 [hep-th], Dec. 2010, 31 pp.

[Gi73] P. B. Gilkey, Curvature and eigenvalues of the Laplace operator for elliptic complexes, *Advances in Math.* **10** (1973), 344-382.

[Gi75] Gilkey, P.B., The boundary integrand in the formula for the signature and Euler characteristic of a Riemannian manifold with boundary, *Advances in Math.* **15** (1975), 334-360.

[Gi95] —, *Invariance Theory, the Index Theorem and the Heat Equation*, CRC Press, Boca Ratan-Ann Arbor-London-Tokyo, 1995.

[Gi04] —, *Asymptotic Formulae in Spectral Geometry*, CRC Press, Boca Ratan-Ann Arbor-London-Tokyo, 2004.

[Gi08] —, The spectral geometry of operators of Dirac and Laplace type, in: D. Krupka and D. Saudners (eds.), *Handbook of Global Analysis*, Elsevier, 2008, pp. 287-324, also http://pages.uoregon.edu/gilkey/ .

[GIKP] Gilkey, P., I. Ivanova, K. Kirsten and J.H. Park, The index theorem, in: J.-P. Francoise, G.L. Naber and Tsou S.T. (eds.), *Encyclopedia of Mathematical Physics*, (Elsevier Academic) Oxford: Elsevier, 2006, volume 3, pp. 23-31, also http://pages.uoregon.edu/gilkey/ .

[GIN] Gilkey, P., I. Ivanova and S. Nikčević, Characteristic classes, in: J.-P. Francoise, G.L. Naber and Tsou S.T. (eds.), *Encyclopedia of Mathematical Physics*, (Elsevier Academic) Oxford: Elsevier, 2006, volume 3, pp. 488-496, also http://pages.uoregon.edu/gilkey/ .

[GJ] Glimm, J. and Jaffe, A., *Quantum Physics*, Second Ed., Springer-Verlag, New York, 1987.

[GF] Gohberg, I.Z. and I.A. Feldman, *Faltungsgleichungen und Projektionsverfahren zu ihrer Lösung* (tranlation from Russian), Akademie-Vlg., Berlin, 1974.

[GG] Gohberg, I.Z. and S. Goldberg, *Basic Operator Theory*, Birhäuser, Boston, 1981.

[GK57] Gohberg, I.Z. and M.G. Krein, The basic propositions on defect numbers, root numbers, and indices of linear operators (translation from Upehi Mat. Nauk **12**, no. 2 (1957), 44-118), *Amer. Math. Soc. Transl.* (2) **13** (1960), 185-264.

[GK58] Gohberg, I.Z. and M.G. Krein, Systems of integal equations on a half-line with kernels depending on the difference of the arguments (translation from Upehi Mat. Nauk **13**, no. 2 (1958), 3-72), *Amer. Math. Soc. Transl.* (2) **14**, (1960), 217-287.

[Go] Goldberg, S.I., *Curvature and Homology*, Dover Edition, Dover, New York, 1982.

[Gom] Gomph, R., An infinite set of exotic $\mathbb{R}^4$s, *J. Diff. Geom.* **21** (1985), 283-300.

[GWW] Gordon, C.S., D. Webb and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, *Invent. Math.* **110** (1995), 1-22.

[Gra] Gracia-Bondía, J., Notes on quantum gravity and noncommutative geometry. In: *New Paths Towards Quantum Gravity*, eds. B. Booß–Bavnbek, G. Esposito and M. Lesch, Springer Lecture Notes in Physics Vol. 807, Springer, Heidelberg-Berlin-New York, 2010, pp. 3-58.

[GH] Griffiths, P. and J. Harris, *Principles of Algebraic Geometry*, Wiley, New York, 1978.

[Gr] Greenberg, M.J., *Lectures on Algebraic Topology*, Benjamin, New York, 1967.

[GS] Grigis, A. and J. Sjöstrand, *Microlocal Analysis for Differential Operators. An Introduction*, London Mathematical Society Lecture Note Series, 196, Cambridge University Press, Cambridge, 1994, iv+151 pp.

[Gro] Grothendieck, A., Produits tensoriels topologiques et espaces nucléaires. *Mem. Amer. Math. Soc.* **1955** (1955), no. 16, 140 pp.

[Gru] Grubb, G., *Distributions and Operators*, Graduate Texts in Mathematics, 252, Springer, New York, 2009, xii+461 pp.

[GruG] Grubb, G. and G. Geymonat, The essential spectrum of elliptic systems of mixed order, *Math. Ann.* **227** (1977), 247-276.

[Gu93] Guentner, E., K-homology and the Index Theorem, in: *Index Theory and Operator Algebras*, *Contemp. Math.* **148**, (1993), pp. 47–66.

[GP] Guillemin, V. and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, 1974.

[HH] Haefliger, A. and M. Hirsch, On the existence and classification of differentiable embeddings, *Topology* **2** (1963) 129-135.

[Haa] Haack, W., Probleme höherer Charakteristik für ein System von zwei elliptischen Differentialgleichungen, *Math. Nachr.* ***8*** (1952), 123–132.

[Ha] Hartmann, P., *Ordinary Differential Equations*, Wiley, New York, 1964.

[Hel] Hellwig, G., Das Randwertproblem eines linearen elliptischen Systems, *Math. Zeitschrift* **56** (1952), 388–408.

[HT] Hellinger, E. and O. Toeplitz, Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten, In *Enzyklopädie der mathematischen Wissenschaften* II, C 13, Teubner, Leipzig, 1927, 1335-1601.

[He] Hermann, R., *Vector Bundles in Mathematical Physics* I/II, Benjamin, New York, 1970.

[Hir] Hirsch, M. W., *Differential topology*, Springer-Verlag, Berlin-Heidelberg-New York, 1976.

[Hi66a] Hirzebruch, F., *Topolgical Methods in Algebraic Geometry* (1st German ed., 1956), Third Ed., Springer-Verlag, New York, 1966.

[Hi66b] —, Elliptische Differentialoperatoren auf Mannigfaltigkeiten, *Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen* **33** (1966), 583-608.

[Hi73] —, Hilbert modular surfaces, *Enseignement Math.* **19** (1973), 183-281.

[HiHo] Hirzebruch, F. and H. Hopf, Felder von Flächenelementen in 4-dimensionalen Mannigfaltigkeiten, *Math. Ann.* **136** (1958), 156-172

[HiSch] Hirzebruch, F. and W. Scharlau, *Einführung in die Funktionalanalysis*, B.I.-Wissenschaftsverlag, Mannheim, 1971.***

[HiZa] Hirzebruch, F. and D. Zagier, *The Atiyah-Singer theorem and elementary number theory*, Mathematics Lecture Series 3, Publish or Perish Inc., Boston, 1974.

[Hit] Hitchin, N., On compact four-dimensional Einstein manifolds, *J. Diff. Geom.* **9** (1974), 435-442.

[Hö63] Hörmander, L., *Linear Partial Differential Equations*, Springer-Verlag, Berlin-Heidelberg-New York, 1963/1969.

[Hö71] —, On the index of pseudodifferential operators. In: G. Anger (ed.), *Elliptische Differentialgleichungen*, Band II, Akademie-Berlag, Berlin 1971, 127-146.

[Hö66a] —, Pseudo-differential operators and non-elliptic boundary problems, *Ann. of Math.* **83** (1966), 129-209.

[Hö66b] —, Pseudo-differential operators and hypoelliptic equations, *Math. Soc. Symp. X on Singular Integral Operators*, 1966, 138-183.

[Hö71a] —, The calculus of Fourier integral operators. In: *Prospects in Mathematics* (F. Hirzebruch et al.), Princeton University Press, Princeton, N.J., 1971, pp. 33-57.

[Hö71b] —, Fourier integral operators I, *Acta Math.* **127** (1971), 79-183.

[Hö71c] —, On the existence and the regularity of solutions of linear pseudo-differential equations, *Enseignement Math.* **17** (1971), 99-163.

[Hö2] —, *The Analysis of Linear Partial Differential Operators. II. Differential Operators with Constant Coefficients*, Springer-Verlag, Berlin, 1983, pp. viii+399.

[Hö3] —, *The Analysis of Linear Partial Differential Operators. III. Pseudo-Differential Operators*, Springer-Verlag, Berlin, 1985, pp. viii+525.

[Hö4] —, *The Analysis of Linear Partial Differential Operators. IV. Fourier Integral Operators*, Springer-Verlag, Berlin, 1985, pp. vii+352.

[Hum] Humphreys, J.E., *Introduction to Lie Algebras and Representation Theory*, (Third printing, revised) Graduate Texts in Mathematics, Springer-Verlag, Berlin-Heidelberg-New York, 1980.

[Hu] Husemoller, D., *Fibre Bundles*, McGraw-Hill, New York, 1966 (Russ. transl., Moscow, 1970); 2nd Edition, Springer-Verlag, New York, 1975.

[Il] Illusie, L., Contractibilité du groupe linéaire des espaces de Hilbert de dimension infinie, *Séminaire Bourbaki* **17**, No. 284 (1964/65).

[IL] Ivachnenko, A.G. and V.G. Lapa, *Cybernetics and Forcasting Techniques*, American Elsevier, New York, 1967 (translation from Russian).

[Ize] Ize, J., Bifurcation theory for Fredholm operators, Mem. Am. Math. Soc. **7**/174 (Sept. 1976).

[JNR] Jackiw, R., C. Nohl and C. Rebbi, Conformal properties of pseudoparticle configurations, *Phys. Rev.* **D15** (1976), 1642-1646.

[Ja] Jänich, K., Vektorraumbündel und der Raum der Fredholm-Operatoren, *Math. Ann.* **161** (1965), 129-142.

[Jö] Jörgens, K., *Lineare Integraloperatoren*, Teubner, Stuttgart, 1970; Eng. transl. Pitman, San Francisco, 1982.

[Jo] Jost, J. *Riemannian Geometry and Geometric Analysis*, Sixth edition. Universitext. Springer, Heidelberg, 2011. xiv+611 pp.

[Kac] Kac, M., On some connections between probability and differential equations. In: *Proceedings of the Second Berkeley Symposium on Probability and Statistics*, J. Neyman, ed., University of California Press, Berkeley (1950), pp. 189-215.

[Kac66] —, Can one hear the shape of a drum?, *Amer. Math. Monthly* **73** (1966), 1-23.

[Kak] Kaku, M., *Quantum Field Theory*, Oxford University Press, New York, 1993.

[Kal] Kaluza, T., *Sitzber. Preuss Akad. Wiss.* (1921), 966.

[Kar] Karoubi, M., *K-theory – An Introduction*, Springer-Verlag, Berlin-Heidelberg-New York, 1978.

[KR] Karoui, N. and H. Reinhard, Processus de diffusion dans $\mathbb{R}^n$. In: *Sém. Probabilités* VII, Univ. de Strasbourg, Springer-Verlag LN 321, Berlin-Heidelberg-New York, 1973, pp. 95ff.

[Kat] Kato, T., *Perturbation Theory for Linear Operators*, Springer–Verlag, Berlin, 1966, 2d ed. 1976.

[Kato] —, Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups. In: *Topics in Functional Analysis* (I. Gohberg and M. Kac, eds.), Academic Press, New York, 1978, pp. ???-???.

[Kir] Kirsten, K., *Spectral Functions in Mathematics and Physics*, Chapman & Hall/CRC, Boca Raton, FL, 2001, 382 pp.

[Kle26] Klein, O., Quantentheorie und fünfdimensionale Relativitätstheorie, *Z. Phys.* **37** (1926), 895.

[Kle38] —, On the theory of charged fields (In: *New Theories in Physics*, Proc. of a Conf. held in Warsaw, 1938), Institute for Intellectual Collaboration, Paris.

[Kli] Kline, M., *Mathematical Thought from Ancient to Modern Times*, Oxford Univ. Press, New York, 1972.

[KN63] Kobayashi, S. and K. Nomizu, *Foundations of Differential Geometry, Vol. I,* Wiley, New York, 1963.

[KN69] —, —, *Foundations of Differential Geometry, Vol. II,* Wiley, New York, 1969.

[Kod] Kodaira, K., On the structure of campact complex analytic surfaces I, *Am. J. Math.* **84** (1964), 751-798.

[KMS] Kolár, I., P. Michor and J. Slovák, *Natural Operations in Differential Geometry*, Springer-Verlag, 1993; free copy from `http://www.emis.de/monographs/KSM/index.html`.

[Kol] Kolmogorov, A. N., Interpolation and extrapolation of stationary stochastic sequences, *Ivz. Akad. Nauk SSSR Ser. Mat.* **5** (1941), 3-14 (Russian).

[Kos] KOSCHORKE, U., Frame fields and nondegenerate singularities, *Bull. Amer. Math. Soc.* **81** (1975), 157-160.

[Kr] KREIN, M.G., Integal equations on a half-line with kernel depending on the difference of the arguments, (translation from Yonexn Mathem. Hayk **13**, no. 5 (1958), 3-120), *Amer. Math. Soc. Transl.* (2) **22** (1962), 163-288.

[KM94] KRONHEIMER, P.B. AND T.S. MROWKA, The genus of embedded surfaces in the projective plane, *Math. Res. Letters* **1** (1994), 797-808.

[KM08] —, —, Knot homology groups from instantons, arXiv:0806.1053v2 [math.GT], Aug. 2008, 119 pp.

[Kui] KUIPER, N.H., The homotopy type of the unitary group of Hilbert space, *Topology* **3** (1965), 19-30.

[KG] KUMANO-GO, H., *Pseudo-Differential Operators*, MIT Press, Cambridge, Mass., 1981.

[Lag] LAGRANGE, J.L., Letter to Euler 12 August, 1755 (Latin), in: J.A. Serret and G. Darboux (eds.), *Œuvres de Lagrange* **14**, Paris, 1892, pp. 366–375.

[Lan] LANG, S., *Differentiable Manifolds*, Addison-Wesley, Reading, Mass., 1972.

[LaMi] LAWSON, H.B. AND M. MICHELSOHN, *Spin Geometry*, Princeton Univ. Press, Princeton, 1989.

[LeB] LEBRUN, C., Einstein metrics and Mostow rigidity, *Math. Res. Letters* **2** (1995), 1-8.

[Les] LESCH, M., *Operators of Fuchs Type, Conical Singularities, and Asymptotic Methods*, Teubner Verlag, Stuttgart, 1997.

[Les04] LESCH, M.,The uniqueness of the spectral flow on spaces of unbounded self-adjoint Fredholm operators, In: B. Booß–Bavnbek, G. Grubb and K.P. Wojciechowski (eds.), *Spectral Geometry of Manifolds with Boundary and Decomposition of Manifolds*, Amer. Math. Soc. Series *Contemporary Mathematics*, vol. 366, Providence, R.I., 2005, pp. 193-224.

[LiMa] LIONS, J.L. AND E. MAGENES, *Problèmes aux limites non homogènes et applications*, I, Dunod, Paris, 1968.

[Lo] LOGAN, D.J., *Applied Partial Differential Equations*, Springer–Verlag, New York, 1998.

[Mac] MACLANE, S., *Homology*, Springer-Verlag, Berlin-Heidelberg-New York, 1963.

[Mad] MADORE, J., Geometric methods in classical field theory, *Phys. Reports* **75**, (1981), 125-204.

[Mag] MAGENES, E., Spazi di interpolazione ed equazioni a derivate parziali, *Atti del VII Congresso dell' Unione Mat. Ital.* (Genova 1963), Ediz. Cremonese, Rom 1964.***

[Mas] MASSEY, W.S., On the Stieffel-Whitney classes of a manifold. II, *Proc. Amer. Math. Soc.* **13** (1962), 938-942

[May] MAYER, K.H., Elliptische Differentialoperatoren und Ganzzahligkeitssätze für charakteristische Zahlen, *Topology* **4** (1965), 295-313.

[McS] MCDUFF, D. AND D. SALAMON, *Introduction to Symplectic Topology*, Clarendon Press, Oxford, 1995.

[MS] MCKEAN, H.P., JR. AND I.M. SINGER, Curvature and the eigenvalues of the Laplacian, *J. Diff. Geom.* **1** (1967), 43-69.

[Mel] MELROSE, R.B., *The Atiyah-Patodi-Singer Index Theorem*, A. K. Peters, Boston, 1993.

[Me] MEYER, P.A., *Probabilités et potentiels*, Hermann, Paris, 1966.

[Mik] MIKHLIN [MIHLIN], S.G., *Multidimensional Singular Integrals and Integral Equations*, Pergamon Press, New York, 1965 (translation from Fizmatgiz, Moscow, 1962).

[Mil58] MILNOR, J.W., On simply connected 4-manifolds. In: *Proc. Internat. Sympos. Algebraic Topology, Mexico* held in 1956 (1958), pp. 122-128.

[Mil] —, *Morse Theory*, Princeton Univ. Press, Princeton, N. J., 1963.

[Mil64] —, Eigenvalues of the Laplace operator on certain manifolds, *Proc. Nat. Acad. of Sciences U.S.A* **51** (1964), 542.

[Mil65] —, *Lectures on the h-cobordism*, Princeton Mathematical Notes, 1965.

[MilH] MILNOR, J. AND D. HUSEMOLLER, *Symmetric Bilinear Forms*, Springer-Verlag, New York, 1973.

[MilS] MILNOR, J. AND J. STASHEFF, *Characteristic Classes*, Ann. Math. Studies 76, Princeton Univ. Press, Princeton, 1974.

[MiPl] MINAKSHISUNDARAM, S. AND Å. PLEIJEL, Some properties of eigenfunctions of the Laplace operator on Riemannian manifolds, *Can. J. Math.* **1** (1949), 242-256.

[MV] MITTER, P.K. AND C.M. VIALLET, On the bundle of connections and gauge orbit manifold in Yang-Mills theory, *Commun. Math. Phys.* **79** (1981), 457-472.

[Miz] Mizohata, S., *The Theory of Partial Differential Equations*, Cambridge University Press, Cambridge, 1973 (translation from Iwanami Shoten Publ., Tokyo, 1965).

[Mo] Morrey, C., *Multiple integrals and the calculus of variations*, Springer-Verlag, Berlin-Heidelberg-New York, 1966.

[Mu] Mumford, D., *Introduction to Algebraic Geometry*, Vol. 1, Springer-Verlag, Berlin-Heidelberg-New York, 1976.

[Na] Narasimhan, R., *Analysis on real and complex manifolds*, Masson, Paris, 1973.

[Nas] Nash, C., *Differential Topology and Quantum Field Theory*, Academic Press, London, 1991.

[NS] Nemitsky, V.V. and V.V. Stepanov, *Qualitative theory of differential equations* (transl. from Russian), Princeton Univ. Press, Princeton, N.J., 1960.

[Neu] von Neumann, J., Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren, *Math. Ann.* **102** (1929), 49-131.

[Nic] Nicolaescu, L.I., *Lectures on the Geometry of Manifolds*, World Scientific, Singapore, 1996.***

[Ni] Nirenberg, L., Pseudo-differential operators, *Amer. Math. Soc. Symp. Pure Math.* **16** (1970), 149-167.

[No] Noether, F., Über eine Klasse singulärer Integralgleichungen, *Math. Ann.* **82** (1921), 42-63.

[OPS] Osgood, B., R. Phillips and P. Sarnak, Extremals of determinants of Laplacians, *J. Funct. Anal.* **80** (1988), 148-211.

[Pal65] Palais, R.S. (ed.), *Seminar on the Atiyah-Singer index theorem*, Ann. of Math. Studies **57** (Russ. transl. Moscow, 1970), Princeton Univ. Press, Princeton, 1965.

[Pal68] —, *Foundations of Global Non-Linear Functional Analysis*, Benjamin, New York, 1968.

[PaWo02] J. Park and K.P. Wojciechowski, Scattering theory and adiabatic decomposition of the ζ-determinant of the Dirac Laplacian, *Math. Res. Lett.* **9** (2002), 17–25.

[Par] Parker, T.H., Gauge theories on four-dimensional Riemannian manifolds, *Commun. Math. Phys.* **85** (1982), 563-602.

[Pat] Patodi, V.K., Curvature and the eigenforms of the Laplace operator, *J. Diff. Geom.* **5** (1971), 233-249.

[Ped] Pedersen, G.K., *Analysis Now*, Springer-Verlag, New York, 1989.

[Per02] Perelman, G., The entropy formula for the Ricci flow and its geometric applications. arXiv:math.DG/0211159 [math.DG] (November 11, 2002).

[Per03a] — Ricci flow with surgery on three-manifolds. arXiv:math.DG/0303109 [math.DG] (March 10, 2003).

[Per03b] — Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. arXiv:math.DG/0307245 [math.DG] (July 17, 2003).

[PS] Peskin, M.E. and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley, Reading Mass., 1995.

[Pf] Pflaum, M.J., The normal symbol on Riemannian manifolds, *New York J. Math.* **4** (1998), 97-125 (electronic).

[Po] Pontryagin, L.S., *Ordinary Differential equations* (transl. from Russian), Pergamon Press, London, 1962.

[Prö72] Prössdorf, D., Uber eine Algebra von Pseudodifferentialoperatoren im Halbraum, *Math. Nachr.* **52** (1972), 113-139.

[Prö74] —, *Einige Klassen singulärer Gleichungen*, Akademie-Verlag, Berlin, 1974 (Engl. translation North-Holland /Elsevier, New York, 1978).

[PRR] Przevorska-Rolevicz, D. and S. Rolewicz, *Equations in Linear Spaces*, Panstvove Vydavnistvo Naukove, Warsaw, 1968.

[Q85] Quillen, D., Determinants of Cauchy-Riemann operators on Riemann surfaces, *Funct. Anal. and its Appl.* **19**, no. 1 (1985), 37-41.

[RaSi71] Ray, D. and I.M. Singer, R-torsion and the Laplacian on Riemannian manifolds, *Advances in Math.* **7** (1971), 145-210.

[RaSi73] —,—, Analytic torsion for complex manifolds, *Ann. of Math.*, **98** (1973), 154-177.

[ReSi72] Reed, M. and B. Simon, *Methods of Modern Mathematical Physics, I. Functional Analysis,* Academic Press, New York, 1972.

[ReSi75] —,—, *Methods of Modern Mathematical Physics, II. Fourier Analysis,* Academic Press, New York, 1975.

[Ro] Rochlin, V.A., New results in the theory of four-manifolds (in Russian), *Dokl. Acad. Nauk USSR* **84** (1952), 221-224.

[Rud] Rudin, W., *Functional Analysis*, McGraw-Hill, New York, 1973.

[Sa] Sannino, F., Conformal dynamics for TeV physics and cosmology, *Acta Phys. Polon. B* **40** (2009), 3533–3743, arXiv:0911.0931 [hep-ph].

[Schi] Schiff, L.I., *Quantum Mechanics*, McGraw-Hill, New York, 1968.

[Sche] Schechter, M., *Principles of Functional Analysis*, Second Edition, Graduate Studies in Mathematics v. **36**, Amer. Math. Soc., Providence, Rhode Island, 2002.

[Schu76a] Schulze, B.W., On the set of all elliptic boundary value problems for an elliptic pseudodifferential operator on a manifold, *Math. Nachr.* **75** (1976), 271-282.

[Schu76b] —, Elliptic operators on manifolds with boundary, *Proc.* "*Globale Analysis*", Ludwigsfelde, Berlin, 1976 (notes).

[Schw50] Schwartz, L., *Théorie des distributions* I, II, Hermann, Paris, 1950/5l.

[Schw64] Schwarz, A.S., The homotopy-topology of Banach spaces (Russian), *Dokl. Akad. Nauk SSSR* **154** (1964), 61-63. English translation in *Sov. Math. Dokl.* **5** (1964), 57-59.

[Schw79] —, Instantons and fermions in the field of an instanton, *Commun. Math. Phys.* **64** (1979), 233-268.

[Schw66] Schwarzenberger, R.L.E., Appendix One. In: F. Hirzebruch, *Topological Methods in Algebraic Geometry*, Springer-Verlag, Berlin-Heidelberg-New York, 1966, 159-201.

[Sco02] Scott, S.G., Zeta determinants on manifolds with boundary, *J. Funct. Anal.* **192** (2002), 112–185.

[ScWo00] Scott, S.G. and K.P. Wojciechowski, The ζ–determinant and Quillen determinant for a Dirac operator on a manifold with boundary, *GAFA – Geom. Funct. Anal.* **10** (2000), 1202–1236.

[See65] Seeley, R.T., Integro-differential operators on vector bundles, *Trans. Amer. Math. Soc.* **117** (1965), 167-204.

[See67] —, Complex powers of an elliptic operator (Russ. transl., Moscow, 1968), *Amer. Math. Soc. Proc. Sympos. Pure Math.* **10** (1967), 288-307.

[See68] —, Topics in pseudo-differential operators. In: *C.I.M.E. Conference on Pseudo-Differential Operators*, Stresa 1968, Cremonese, Rome, 1969, pp. 167-305.

[See69] —, The resolvent of an elliptic boundary problem, *Amer. J. Math.* **91** (1969), 889-920.

[Seg] Segal, G.B., Equivariant K-theory, *Publ. Math. Inst. Hautes Etudes Sci.* **34** (1968), 105-151.

[Seg90] —, The definition of conformal field theory, Oxford preprint, 1990.

[SW] Seiberg, N. and E. Witten, Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD, *Nucl. Phys. B* **431** (1994), 581-640.

[ST] Seifert, H. and W. Threlfall, *A Textbook of Topology* (1st German ed. 1934), Academic Press, New York, 1980.

[Ser] Serre, J.-P., *A Course in Arithmetic,* Springer-Verlag 1973.

[Sha] Shanahan, P., *The Atiyah–Singer Index Theorem: An Introduction*, Lecture Notes in Mathematics 638, Springer–Verlag, Berlin, 1978.

[Shu] Shubin, M.A., *Pseudodifferential Operators and Spectral Theory*, Springer–Verlag, Berlin, 1987 (translation from Nauka, Moscow, 1978).

[Sie] Siegel, C.L., *Topics in Complex Function Theory* I-III, Wiley-Interscience New York, 1969/1971/1973.

[Sim] Simon, B., Notes on infinite determinants of Hilbert space operaors, *Advances in Math.* **24** (1977), 244-273.

[Si63] Singer, I.M., On the index of elliptic operators. In: *Outlines Joint Soviet Amer. Sympos. Partial Diff. Equations* (Nowosibirsk 1963), Moscow, 1963, pp. ???-???.

[Si69] —, Elliptic operators on manifolds. In: *Pseudo-Differential Operators*, Centro Internazionale Matematico Estivo Stresa 1968, Edizioni Cremonese, Rom, 1969, 333-375.

[Si70] —, Operator theory and K-theory, Lecture presented in the Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, 1970.

[Si71a] —, Future extensions of index theory and elliptic operators. In: *Prospects in Mathematics* by F. Hirzebruch et al., Princeton University Press, Princeton, 1971, pp. 171-185.

[Si71b] —, Operator theory and periodicity, *Indiana Univ. Math. J.* **20** (1971), 949-951.

[Si73] —, Recent applications of index theory for elliptic operators, *Amer. Math. Soc. Symp. Pure Math.* **23** (1973), 11-31.

[Si75] —, Eigenvalues of the Laplacian and invariants of manifolds, *Proc. Int. Congr. Math.* (Vancouver 1974), Vancouver, 1975, 187-200.

[ST] Singer, I.M. and J.A. Thorpe, *Lecture Notes on Elementary Topology and Geometry*, Scott Foresmann, Glenview, Ill., 1967; reprinted, Springer-Verlag, New York.

[Sm61] Smale, S., Generalized Poincaré's conjecture in dimensions greater than four, *Am. J. Math.* **84** (1961), 391-406.

[Sm62] —, On the structure of Manifolds, *Am. J. Math.* **84** (1962), 387-399.

[Sm68] —, An infinite-dimensional version of Sard's theorem, *Am. J. Math.* **87** (1968), 861-866.

[Sp] Spanier, E.H., *Algebraic Topology*, McGraw-Hill, New York, 1966.

[Ste] Steenrod, N., *The Topology of Fibre Bundles*, Princeton Univ. Press, Princeton, 1951.

[Sto] Stong, R.E., *Notes on Cobordism Theory*, Princeton Univ. Press, Princeton, 1958.

[Sza] Szankowski, A., $B(\mathcal{H})$ does not have the approximation property, *Acta Math.* **147** (1981), no. 1-2, 89-108.

[Tau82] Taubes, C.H., Self-dual Yang-Mills connections on non-self-dual 4-manifolds, *J. Diff. Geom.* **17** (1982), 139-170.

[Tau94] —, The Seiberg-Witten invariants and symplectic forms, *Math. Res. Letters* **1** (1994), 809-822.

[Tau95] —, More constraints on symplectic forms from Seiberg-Witten invariants, *Math. Res. Letters* **2** (1995), 9-14.

[Tay74] Taylor, M.E., *Pseudo Differential Operators*, Lecture Notes in Mathematics: 416, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

[Tay81] —, *Pseudo-Differential Operators*, Princeton Univ. Press Princeton, 1981.

[Tay1] —, *Partial Differential Equations I. Basic Theory*, Springer–Verlag, New York, 1996.

[Tay2] —, *Partial Differential Equations II. Qualitative Studies of Linear Equations*, Springer–Verlag, New York, 1996.

[Th] Thomas, E., Vector fields on manifolds, *Bull. Amer. Math. Soc.* **75** (1969), 643-683.

[Ti] Titchmarsh, E.C., *Introduction to the Theory of Fourier Integrals I & II*, Oxford University Press, Oxford, 1937 and 1948.

[TT] Toledo, D. and L.L. Tong, A parametrix for $\bar{\partial}$ and Riemann-Roch in Cech theory, *Topology* **15** (1976), 273-301.

[Tra] Trautman, A., Fiber bundles, gauge fields, and gravitation. In: *General Relativity and Gravitation* I, edited by A. Held, Plenum Press, New York, 1980, pp. 287-307.

[Tri] Triebel, H., *Interpolation Theory-Function Spaces, Differential Operators* (transl. from German), North-Holland/Elsevier, New York, 1978.

[Tre] Treves, F., *Introduction to Pseudo-Differential and Fourier Integral Operators I and II*, Plenum Press, New York, 1980.

[Tro] Trotter, H.F., Approximation of semi-groups of operators, *Pacific J. Math.* **8** (1958), 887–919.

[Uh82a] Uhlenbeck, K.K., Removable singularities in Yang-Mills fields, *Commun. Math. Phys.* **83** (1982), 11-29.

[Uh82b] —, Connections with L^p bounds on curvature, *Commun. Math. Phys.* **83** (1982), 31-42.

[Ul] Ulam, S.M., *A Collection of Mathematical Problems*, Interscience, New York-London, 1960.

[UN] UNESCO (ed.) *Mathematics Applied To Physics*, with contributions from G. A. Deschamps and others, Edited by E. Roubine for UNESCO, Springer-Verlag, Berlin-Heidelberg-New York, 1960.

[vdW] van der Waerden, B.L., *Algebra* I/II, Springer-Verlag, Berlin-Heidelberg-New York, 1970.

[Va] Vassilevich, D.V, Spectral problems from quantum field theory, In: B. Booß–Bavnbek, G. Grubb and K.P. Wojciechowski (eds.), *Spectral Geometry of Manifolds with Boundary and Decomposition of Manifolds*, Amer. Math. Soc. Series *Contemporary Mathematics*, vol. 366, Providence, R.I., 2005, pp. 3–21.

[VT] Vaynberg, M.M. and V.A. Trenogyn, *Theorie der Lösungsverzweigungen bei nichtlinearen Gleichungen* (transl. from Russian), Akademie-Vlg., Berlin, 1973.

[Ve56] Vekua, I.N., *Systeme von Differentialgleichungen erster Ordnung vom elliptischen Typus und Randwertaufgaben* (transl. from Russian), VEB Deutscher Vlg. der Wissenschaften, Berlin, 1956.

[Ve62] —, *Generalized Analytic Functions* (transl. from Russian), Pergamon Press, Oxford, 1962.

[VE] Vishik, M.I. and G.I. Eskin, Normally solvable problems for elliptic systems of equations in convolutions (transl. from Natem. Co. **74** (1967), no. 3, 326-356), *Math. USSR-Sbornik* **3** (1967), 303-332.

[Vor] Voronov, T., Quantization of forms on the cotangent bundle, *Comm. Math. Phys.* **205** (1999), 315-336.

[Wall] Wall, C.T.C., On simply-connected 4-manifolds, *J. London Math. Soc.* **39** (1964), 141-149.

[Wa] Wallace, A.H., *Differential Topology - First Steps*, Benjamin, New York, 1968.

[Wal] Wallach, N.R., *Harmonic Analysis on Homogeneous Spaces*, Marcel Dekker, New York, 1973.

[Wells] Wells, R.O., *Differential Analysis on Complex Manifolds*, Springer-Verlag, Berlin-Heidelberg-New York, 1980.

[We46] Weyl, H., *The Classical Groups*, Princeton Univ. Press, Princeton, 1946.

[We64] —, *The Concept of a Riemann Surface*, 3rd ed. (1st German ed., 1913), Addison-Wesley, Reading, Mass., 1964.

[Wh] Whitney, H., The self-intersections of a smooth n-manifold in $2n$-space, *Ann. of Math.* **45** (1944), 220-246.

[Wid] Widom, H., A complete symbolic calculus for pseudodifferential operators, *Bull. Sci. Math.* (2) **104** (1980), 19-63.

[Wie33] Wiener, N., *The Fourier Integral and Certain of its Applications,* Cambridge University Press, Cambridge, 1933.

[Wie49] —, *Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications*, M.I.T. Press, Cambridge, 1949.

[Wie61] —, *Cybernetics - Control and communication in the animal and the Machine*, 2nd ed., M.I.T. Press, New York, 1961.

[Wil] Wilczek, F., ??? In: *Quark Confinement and Field Theory*, eds. D. Stump and D. Weingarten, J. Wiley and Sons, New York, 1977, pp. ???-???.

[Wit] Witten, E., Some exact multipseudoparticle solutions of classical Yang-Mills theory, *Phys. Rev. Lett.* **38** (1977), 121-124.

[Wit94] —, Monopoles and four-manifolds, *Math. Res. Letters* **1** (1994), 769-796.

[Yoo] Yood, B., Properties of linear transformations preserved under addition of completely continuous transformations, *Duke Math. J.* **18** (1951), 599-612.

[Yosh] Yoshida, T., *Index Theorem for Dirac Operators*, Kyoritsu Shuppan Publishers, 1998 (in Japanese).

[Yo] Yosida, K., *Functional Analysis*, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

[Zy] Zypkin, J.S., *Adaption und Lernen in Kybernetischen Systemen.*, Oldenbourg Blg., München-Wien, 1970 (translation from Russian).