Smooth 4-manifolds and the Seiberg–Witten equations

T. PERUTZ
UNIVERSITY OF TEXAS AT AUSTIN, SPRING SEMESTER, 2018

Contents

1 Classification problems in differential topology **8**
1.1 Smooth manifolds and their classification 8
1.2 Dimensions 2 and 3 8
1.3 Higher dimensions: limitations 9
1.4 Higher dimensions: revised goals 9
1.5 4-manifolds 10

2 Review: The algebraic topology of manifolds **12**
2.1 Cup products 12
2.1.1 Čech cohomology 12
2.2 De Rham cohomology 13
2.3 Poincaré duality 13
2.3.1 The fundamental homology class 13
2.3.2 Cap product and Poincaré duality 14
2.3.3 Intersection of submanifolds 14

3 The intersection form **16**

4 The cup product in middle dimensional cohomology **16**
4.1 Symmetric forms over $\mathbb{R}$ 17
4.2 Characteristic vectors 18
4.3 The E_8 lattice 19
4.4 Topological examples 20

5 The intersection form and characteristic classes **21**
5.1 Cohomology of 4-manifolds 21
5.2 Characteristic classes 21
5.2.1 Stiefel–Whitney classes 21
5.2.2 Chern classes 23
5.2.3 Pontryagin classes 25

6 Tangent bundles of 4-manifolds 27
6.1 Vector bundles and obstruction theory . . . 27
6.2 Obstruction theory . . . 27
6.2.1 The Stiefel–Whitney classes as primary obstructions . . . 28
6.3 Vector bundles over a 4-manifold . . . 28

7 Rokhlin's theorem and homotopy theory 30
7.1 Rokhlin's theorem . . . 30
7.2 The Pontryagin–Thom construction . . . 30
7.3 The J-homomorphism . . . 31
7.4 Framed 3-manifolds and their bounding 4-manifolds . . . 32
7.5 Background: Relative homotopy groups . . . 33
7.6 Appendix: Effect on π_3 of Lie group homomorphisms . . . 33

8 Realization of unimodular forms by 4-manifolds 36
8.1 The theorem of Whitehead on 4-dimensional homotopy types . . . 36
8.2 Topological 4-manifolds . . . 37
8.3 Realizing unimodular forms by smooth 4-manifolds . . . 38

8 Self-duality: linear-algebraic aspects 42
8.1 Self-duality . . . 42
8.1.1 The Hodge star . . . 42
8.1.2 2-forms in 4 dimensions . . . 42
8.1.3 Equivalence of conformal structures with maximal positive-definite subspaces . 43
8.1.4 Conformal structures as maps $\Lambda^- \to \Lambda^+$ . . . 44
8.2 2-planes . . . 45

9 Self-duality in 4 dimensions: Hodge-theoretic aspects 46
9.1 The Hodge theorem . . . 46
9.1.1 The co-differential . . . 46
9.1.2 The co-differential as formal adjoint . . . 46
9.1.3 Harmonic forms . . . 46
9.1.4 Variational characterization. . . . 47
9.1.5 The theorem . . . 47
9.2 Self-dual and anti-self-dual harmonic forms . . . 48
9.3 The self-duality complex . . . 48
9.4 The derivative of the period map . . . 49

10 Covariant derivatives **51**
10.1 Covariant derivatives in vector bundles 51
10.2 Curvature 52
10.3 Gauge transformations 54
10.4 Flat connections 54
10.5 Flat connections are local systems 55

11 $\mathrm{U}(1)$-connections **56**
11.1 Connections and gauge transformations in line bundles 56
11.1.1 Chern–Weil theory 56
11.1.2 Structure of $\mathcal{B}_L$ 57
11.2 $\mathrm{U}(1)$-instantons 58

12 Instantons in $\mathrm{U}(1)$-bundles **59**
12.1 Instantons in $\mathrm{U}(1)$-bundles 59
12.1.1 Generic non-existence 60

13 Differential operators **62**
13.1 First-order differential operators 62
13.2 Higher-order operators 64
13.3 Examples of symbols 64
13.4 Elliptic operators 65

14 Analysis of elliptic operators **66**
14.1 Fredholm operators 66
14.2 Sobolev spaces and elliptic estimates 67
14.2.1 Sobolev spaces 67
14.3 Elliptic estimates 69
14.4 L^p bounds 70

15 Clifford algebras, spinors and spin groups **71**
15.1 Clifford algebras 71
15.1.1 Orthogonal sums 72
15.2 Spinors 72
15.2.1 A quick note on spinors in odd dimension 74
15.3 Projective actions 74
15.3.1 Projective action of the orthogonal group 74
15.3.2 Projective action of the orthogonal Lie algebra 75
15.4 Spin groups 76
15.4.1 Clifford groups 76
15.4.2 Spin groups 77
15.4.3 Representations of spin 77

16 Spin groups and spin structures in low dimensions **78**
16.1 The compact Lie groups Spin(n) 78
16.1.1 The story so far 78
16.1.2 Exponentials 78
16.2 Spinors 79
16.3 Low-dimensional cases 80
16.3.1 Spin(2) 80
16.3.2 Spin(3) 80
16.4 Spin(4) 81

17 Spin and Spin^c-structures: topology **83**
17.1 Spin structures on vector bundles 83
17.1.1 Uniqueness 84
17.2 Existence and uniqueness in full 84
17.3 Spin^c-structures 85
17.4 Existence and uniqueness for Spin^c-structures 86
17.4.1 The case of 4-manifolds 86

18 Dirac operators **88**
18.1 The Levi-Civita connection 88
18.2 Clifford connections 88
18.3 The Dirac operator 90
18.4 The formal adjoint to a covariant derivative 90
18.5 The Lichnérowicz formula 91

19 The Seiberg–Witten equations **95**
19.1 Spin^c-structures in 4 dimensions 95
19.2 Spin^c-structures and self-duality 95
19.3 The configuration space 96
19.4 The Seiberg–Witten equations 96
19.4.1 The Dirac equation 96
19.4.2 The curvature equation 97
19.5 The index 98
19.5.1 A reinterpretation of the Seiberg–Witten index 99

20 The Seiberg–Witten equations: bounds **101**
20.0.2 An inequality for Laplacians 101
20.1 *A priori* bounds for solutions to the SW equations 102
20.2 Finiteness 103

21 The compactness theorem **105**
21.1 Statement of the theorem . . . 105
21.2 Sobolev multiplication . . . 105
21.3 The Seiberg–Witten equations in Sobolev spaces . . . 106
21.4 Elliptic estimates and the proof of compactness . . . 107
21.4.1 The positive feedback loop . . . 107
21.5 L^2_1 bounds . . . 108
21.5.1 From L^2_1 to L^2_3 . . . 109
21.5.2 Weak and strong L^2_1 convergence . . . 110

22 Transversality **112**
22.1 Reducible solutions . . . 112
22.1.1 Reducible configurations . . . 112
22.1.2 Reducible solutions and abelian instantons . . . 112
22.2 Transversality for irreducible solutions . . . 113

23 Transversality, continued **115**
23.1 Generic transversality for irreducibles . . . 115
23.1.1 Previously . . . 115
23.1.2 The schematic argument . . . 115
23.2 Fredholm maps and the Sard–Smale theorem . . . 117

24 The diagonalization theorem **120**
24.1 Statement and a preliminary reduction . . . 120
24.2 Seiberg–Witten moduli spaces . . . 121
24.3 Generic regularity of the reducible solution . . . 122
24.4 Proof of Elkies' theorem . . . 123
24.4.1 θ-functions of lattices . . . 124
24.4.2 Modularity . . . 124
24.4.3 The lattice $\mathbb{Z}^N$ . . . 125
24.4.4 The ratio of θ-functions . . . 125

25 Seiberg–Witten invariants **127**
25.1 Preliminaries . . . 127
25.1.1 Homology orientations . . . 127
25.2 Conjugation of Spin^c-structures . . . 127
25.3 Formulation of the invariants . . . 127
25.4 Configuration spaces . . . 128
25.5 The construction . . . 130
25.5.1 Properties . . . 130
25.6 Orientations . . . 131
25.6.1 Virtual vector bundles . . . 131
25.6.2 The virtual index bundle . . . 131
25.6.3 The determinant index bundle . . . 132
25.6.4 A family of Fredholm operators . . . 132

26 Taubes's constraints on symplectic 4-manifolds **134**
26.1 The canonical Spin^c-structure . . . 134
26.2 Statement of the constraints . . . 135
26.3 Geometry of almost complex manifolds . . . 136
26.4 Almost Kähler manifolds . . . 137
26.5 Symplectic 4-manifolds . . . 139

27 Taubes's constraints, continued **141**
27.1 The canonical solution to the Dirac equation . . . 141
27.2 The SW equations . . . 143
27.3 Proof of Taubes's constraints . . . 143

28 The symplectic Thom conjecture **145**
28.1 The minimal genus problem . . . 145
28.2 Proving the adjunction inequality . . . 146
28.2.1 Blowing up . . . 147
28.3 The symplectic Thom conjecture . . . 149

29 Wish-list **150**

Bibliography **151**

Part I. Algebraic and differential topology in dimension 4

1 Classification problems in differential topology

1.1 Smooth manifolds and their classification

The central problem addressed by differential topology is the classification, up to diffeomorphism, of smooth manifolds. An ideal solution would look like this:

(1) We can write down a collection $\{X_i\}_{i\in I}$ of smooth connected manifolds of a particular dimension (perhaps satisfying some other constraints, such as compactness or orientability), which represents all diffeomorphism types without any redundancy.

(2) When someone hands us a description of a manifold M, we can decide to which of the standard manifolds X_i it is diffeomorphic by computing certain invariants. If M is described by a finite set of data—such as a set of polynomial equations with rational coefficients—we can ask for an algorithm for this determination.

(3) When someone hands us two manifolds, M and M', we can compute invariants that decide whether or not they are diffeomorphic (again, there is an algorithmic version of this statement).

(4) We understand what kinds of *families* of manifolds, all diffeomorphic to a fixed manifold M, are possible. For instance, we understand the homotopy type of the topological group $\operatorname{Diff} M$ of self-diffeomorphisms. (This problem is less precisely defined than the others.)

1.2 Dimensions 2 and 3

For compact smooth surfaces, complete solutions are available. The first invariant is orientability, and we state the solution only in the oriented (and connected) case:

(1) For each integer $g \geq 0$, there is a standard surface Σ_g, which can be described—for instance—as the connected sum of the sphere S^2 and g copies of the 2-torus T^2.

(2, 3) The Euler characteristic $\chi(M)$ is a complete invariant of M. It can be computed algorithmically from any reasonable description—for instance a polyhedron underlying the topological space M; or from an atlas such that all multiple intersections of charts are contractible or empty; or from the monodromy data describing M as a Riemann surface with a holomorphic map to S^2. One has $\chi(M) = 2 - 2g$.

(4) [7] We describe the groups $\operatorname{Diff}^+(\Sigma_g)$ of orientation-preserving self-diffeomorphisms. The identity component is denoted by $\operatorname{Diff}^+(\Sigma_g)$, and the group of components (the *mapping class group*) is $\pi_0 \operatorname{Diff}^+(\Sigma_g)$.

  (a) The inclusion $\mathrm{SO}(3) \to \operatorname{Diff}^+(S^2)$ is a homotopy equivalence.

  (b) Writing $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, the inclusion $T^2 \to \operatorname{Diff}^+(T^2)_0$ (where T^2 acts on itself by translations) is a homotopy equivalence, while $\pi_0 \operatorname{Diff}^+(T^2) \cong SL_2(\mathbb{Z})$ (via the action of the mapping class group on $H_1(T^2;\mathbb{Z}) = \mathbb{Z}^2$).

  (c) For $g > 1$, $\operatorname{Diff}^+(\Sigma_g)_0$ is contractible. The mapping class group is an infinite group which acts with finite stabilizers on a certain contractible space (Teichmüller space).

A solution which is almost as complete is known for the far more intricate case of compact 3-manifolds, thanks to the vision of W. Thurston in his *geometrization conjecture*, and its realization via Ricci flow by R. Hamilton and G. Perelman, with contributions from others. The fundamental group is, very nearly, a complete invariant.

1.3 Higher dimensions: limitations

In higher dimension, and in the non-compact case, the desiderata (1–4) are in general overambitious:

- A basic invariant is the fundamental group $\pi_1(M)$; when one presents M as an n-dimensional *handlebody* (closely related to a presentation as a CW complex) one obtains a group presentation for $\pi_1(M)$ (finite when M is compact). For compact manifolds of dimension $n \geq 4$, *any* finite presentation of a group can arise in this way. Deciding whether a finite presentation presents the trivial group is an algorithmically unsolvable problem; so it is not algorithmically possible to decide simple connectivity of arbitrary compact n-manifolds encoded as handlebodies.
- The success with manifolds of dimension 2 and 3 stems from the existence (with some provisos), and uniqueness (ditto), of 'optimal' Riemannian metrics (metrics whose isometry group is transitive, say). In high dimension, there is no known class of metrics for which one can expect simultaneous existence and uniqueness, and there are strong senses in which no such class of metrics can exist [21].
- Any attempt to handle non-compact manifolds, even simply connected ones, must deal with the fact that there are uncountably many diffeomorphism-types. (There are only countably many compact diffeomorphism types.)

1.4 Higher dimensions: revised goals

Focus on the compact case. Assume there is an isomorphism $\pi_1(M) \cong G$ with some standard group G. Most basic is the simply connected case, where one assumes $\pi_1(M)$ trivial, and we shall do that henceforth.

In some cases, such as that of simply connected compact 5-manifolds [1], or homotopy n-spheres for (conservatively) $5 \leq n \leq 18$ [11], solutions to (1–3) are available.

In a much wider range of cases, the framework of *surgery theory* (e.g. [13, 14]) gives conceptual answers to the following questions:

(a) Given the homotopy type of a finite CW complex (simply connected, say), when is it realizable as the homotopy type of a compact manifold of prescribed dimension $n \geq 5$?

(b) Given a compact, simply connected n-manifold M, what are the diffeomorphism types of manifolds homotopy-equivalent to M?

Surgery theory also has something to say about problem (4), but I will not discuss that.

The solution to (a) takes the following form:

To be realizable by a compact n-manifold, a simply connected homotopy type must be a Poincaré space *of dimension n (that is, its (co)homology satisfies Poincaré duality); it must admit a* tangent bundle *(that is, a rank n vector bundle T related in a certain way to the underlying homotopy type); and when n is a multiple of 4, T must obey the Hirzebruch index theorem (this amounts to the vanishing of a certain integer invariant; there is a related vanishing condition in $\mathbb{Z}/2$ when n is 2 mod 4).*

The solution to (b) is closely related to that of (a): one enumerates possible tangent bundles T; for each of them, there is a finite set of possible manifolds, which can be understood via 'surgery obstructions'.

1.5 4-manifolds

The greatest mystery in differential topology is the following broad question:

Question 1.1 *What is the classification of simply connected, compact, 4-dimensional smooth manifolds?*

It turns out that a 4-dimensional simply connected Poincaré complex, up to homotopy equivalence, can be neatly encoded in a *unimodular matrix*, that is, a $b \times b$ symmetric matrix Q (for some $b \geq 0$) with integer coefficients and determinant ± 1. These matrices are to be considered up to *integral equivalence*, meaning replacement by $M^T QM$ for $M \in GL_b(\mathbb{Z})$. The integral equivalence class of matrices is called the *intersection form*, since—in the case of a smooth 4-manifold X—it encodes algebraic intersection numbers of oriented surfaces embedded in X.

The more precise version of the question is this:

Question 1.2 *(i) Which unimodular matrices Q arise as intersection forms of compact, 4-dimensional smooth manifolds?*

(ii) How can we enumerate the diffeomorphism-types of 4-manifolds representing a given Q?

Freedman [8] showed, in epochal work from the early 80s, that the classification of simply connected compact *topological* 4-manifolds, up to homeomorphism, is precisely the same as that of the underlying homotopy types, save for a subtlety concerning a $\mathbb{Z}/2$-valued Kirby–Siebenmann invariant (see Theorem 8.3).

In the smooth category, the existence of a tangent bundle imposes a mild constraint on Q:

Theorem 1.3 (Rokhlin) *Suppose X is a simply connected 4-manifold whose intersection form is represented by a matrix Q with even diagonal entries. Then the signature of Q (the number of positive eigenvalues minus the number of negative eigenvalues) must be divisible by 16.*

Divisibility by 8 is true as a matter of algebra.

In the early 1980s, S. Donaldson, then a graduate student, discovered a more drastic constraint [5]:

Theorem 1.4 (Donaldson's diagonalizability theorem) *Suppose X is a 4-manifold whose intersection form is positive-definite. Then the intersection form is represented by the identity matrix.*

The proof used *gauge theory*, namely, analysis of a moduli space of solutions to a non-linear PDE with *gauge symmetry*, the *instanton* equation.

In subsequent years, 4-manifold invariants (*Donaldson* or *instanton* invariants) were developed which showed that there can be infinitely many diffeomorphism-classes of smooth 4-manifolds within one simply connected homotopy type (or equally by Freedman, homeomorphism type).

In 1994, Witten [22], building on his joint work with Seiberg in string theory, introduced a new pair of equations, the *Seiberg–Witten (SW) equations*, with 4-dimensional gauge symmetry. These equations are a little harder to grasp than the instanton equations, but are in fact much more convenient—in particular, their spaces of solutions have stronger compactness properties. In a flurry of activity, the mathematicians who had been working on instantons switched their attention to the SW equations.

Not only could they rapidly re-prove the known results (such as the diagonalizability theorem), they obtained many new ones. Over the subsequent decade, the SW equations led to huge progress on 4-manifolds, on the geometry of complex surfaces, and on symplectic 4-manifolds. This course will be an introduction to those ideas.

Around the year 2000, *Floer theoretic* invariants for 3-manifolds were introduced, based on the SW equations (*monopole Floer theory* [12]) or on an equivalent formulation in symplectic topology (*Heegaard Floer theory* [**?**]). Such methods have led to progress in contact geometry, most spectacularly the Weinstein conjecture in dimension 3 [**?**], in the relations of knot theory to 3-manifold topology via Dehn surgery (e.g. [**?**]), and even on an old problem in high-dimensional topological manifold theory [**?**]. There is a whole constellation of related theories (monopole, Heegaard, instanton, etc.), but for several major results the *only* known proofs go via SW theory itself.

Known results give an answer to (i) which is nearly, but not quite, complete. As for (ii), SW invariants distinguish many diffeomorphism-types, but we do not know when they are complete invariants, and we do not have a way to capture all diffeomorphism-types within a homotopy type. A solution to (ii) will need truly new ideas.

This course

For the first few weeks, we will study background on 4-manifolds and in differential geometry. We will then carry out the analysis of the SW equations that forms the foundation for the theory. We then turn to applications to 4-manifold topology and to symplectic geometry.

2 Review: The algebraic topology of manifolds

2.1 Cup products

If X and Y are CW complexes, their product $X \times Y$ is again a CW complex (its cells are products of those from X and from Y). The cellular cochain complexes (with $\mathbb{Z}$ coefficients) are related by a *Künneth isomorphism*

$$C^*(X \times Y) \xrightarrow{\cong} C^*(X) \otimes C^*(Y)$$

(map of cochain complexes). The diagonal map $\Delta\colon X \to X \times X$, $x \mapsto (x,x)$ is not cellular (it does not map the k-skeleton to the k-skeleton), but it can be homotoped to a cellular map δ, which induces a map of cochain complexes

$$\delta^*\colon C^*(X \times X) \to C^*(X).$$

Composing, one gets a cochain map

$$\cup\colon C^*(X) \otimes C^*(X) \cong C^*(X \times X) \xrightarrow{\delta^*} C^*(X).$$

On cohomology, this defines the cup product

$$\cup\colon H^*(X) \otimes H^*(X) \to H^*(X),$$

which is in fact an invariant of X (since it can be realized invariantly in the singular cohomology), and is associative and graded-commutative (i.e. $x \cup y = (-1)^{|x||y|} y \cup x$). This makes $H^*(X)$ a graded, graded, unital, commutative ring (e.g. [15]).

The snag with this construction is that the homotopy from Δ to δ makes the cup product non-explicit. One can instead work with singular cochains, in which case there is no need for such a homotopy, but the Künneth isomorphism becomes more cumbersome.

For geometric and computational purposes, it is useful to have more transparent models of the cup product. In the case of smooth manifolds, several such models are available:

- Čech cohomology (in which cup product is given by an explicit formula, and which in the case of compact manifolds is finite and built from essentially combinatorial data);
- De Rham cohomology (in which cup product is realized as wedge product of forms);
- Oriented submanifolds (when available) and Poincaré duality, for which cup product corresponds to intersections of transverse representatives.

We shall discuss the first two now, and the third after reviewing Poincaré duality.

2.1.1 Čech cohomology

Any open covering of a manifold M admits a refinement which is a *good covering* $\mathfrak{U} = \{U_i\}_{i \in I}$, that is, a locally finite covering by open sets U_i such that for every $J \subset I$, the intersection $U_J := \bigcap_{j \in J} U_j$ is either empty or contractible. Any two good coverings admit a common refinement, hence a common good refinement. One can demonstrate the existence of good coverings refining given coverings by either of the following methods: (i) embed M in $\mathbb{R}^N$, and consider intersections with M of very small balls in $\mathbb{R}^N$, centered at points of M; or (ii) via geodesic balls for a Riemannian metric.

Take an open covering $\mathfrak{U} = \{U_i\}_{i \in I}$. Define a k*-simplex* to be an injection $\sigma\colon \{0,1,\dots,k\} \to I$ such that $U_{\sigma(0),\dots,\sigma(k)}$ is non-empty; let S_k be the set of k-simplices. For $0 \le i \le k$, define $\partial_i\colon S_k \to S_{k-1}$

by saying that $\partial_i\sigma$ is the result of deleting $\sigma(i)$ (i.e., $\partial_i\sigma(j) = \sigma(j)$ for $j \leq i$ while $\partial_i\sigma(j) = \sigma(j+1)$ for $j \geq i$). The *Čech complex*, with coefficients in the ring A, is the sum $\check{C}^*(X,\mathcal{U};A) = \bigoplus_{k\geq 0}\check{C}^k(X,\mathcal{U};A)$ where

$$\check{C}^k(X,\mathcal{U};A) = \prod_{S_k} A.$$

It comes with the differential $\delta\colon \check{C}^k \to \check{C}^{k+1}$, given by

$$(\delta\eta)(\sigma) = \sum_{i=0}^{k}(-1)^{i+1}\eta(\partial_i\sigma).$$

One has $\delta\circ\delta = 0$; we write $\check{H}^*(M,\mathcal{U},A)$ for the resulting cohomology groups. A refinement of $\mathcal{U}$ to another covering $\mathcal{V}$ results in a chain map $\check{C}^*(M,\mathcal{V};A) \to \check{C}^*(M,\mathcal{U};A)$, which is a quasi-isomorphism provided that both coverings are good. It follows that $\check{H}^*(M;A) = \check{H}^*(M,\mathcal{U};A)$ is independent of the good covering $\mathcal{U}$. In fact, $\check{C}^*(X,\mathcal{U};A)$ is quasi-isomorphic to the singular cochain complex $S^*(X;A)$ (as one can see the fact that both are acyclic resolutions of the constant sheaf [20]).

The formula for the cup product on the Čech complex is as follows: for $\alpha \in \check{C}^a$ and $\beta \in \check{C}^b$, one has $\alpha\cup\beta \in \check{C}^{a+b}$; in a notation which I hope is self-explanatory,

$$(\alpha\cup\beta)(\sigma) = \alpha(\mathsf{beginning}(\sigma))\cdot\beta(\mathsf{end}(\sigma)).$$

(Graded commutativity on cohomology is not manifest in this model!)

2.2 De Rham cohomology

The de Rham complex $\Omega^*(M)$ is the differential graded algebra (DGA) formed as the direct sum of the vector spaces $\Omega^k(M)$ of k-forms. It has the differential d, the exterior derivative. The wedge product of forms $\Omega^a(M)\times\Omega^b(M) \to \Omega^{a+b}(M)$ is associative and graded-commutative, and d is a derivation.[1]

The relation between Čech and de Rham can be understood through the total complex of a double complex which contains both $\check{C}^*(M,\mathcal{U};\mathbb{R})$ and $\Omega^\bullet(M)$, namely, the Čech–de Rham complex $D^* = \mathsf{Tot}\,\check{C}^*(M;\Omega^\bullet(M))$ [3]. The outcome is that there is a complex D^* and a pair of canonical chain maps

$$\check{C}^*(M,\mathcal{U};\mathbb{R}) \longrightarrow D^* \longleftarrow \Omega^*(M),$$

both quasi-isomorphisms; hence $\check{H}^*(M;\mathbb{R}) \cong H^*_{DR}(M)$ canonically (and in fact, naturally with respect to smooth maps). There is a product on D^* which is respected by the two quasi-isomorphisms, so $\check{H}^*(M;\mathbb{R}) \cong H^*_{DR}(M)$ as $\mathbb{R}$-algebras.

2.3 Poincaré duality

2.3.1 The fundamental homology class

Two basic features of the algebraic topology of topological manifolds X (of dimension n) are the following theorems:

Theorem 2.1 *(see e.g [15])*

[1] Over $\mathbb{Z}$, or indeed $\mathbb{Z}/p$, no such functorial commutative DGA is available for computing cohomology; the Steenrod squares obstruct it.

(i) Vanishing: *One has $H_k(X) = 0$ for $k > n$, and also for $k = n$ when X is connected but not compact.*

(ii) Fundamental class: *One has $H_n(X) \cong \mathbb{Z}$ when X is compact, connected and orientable.*

In the closed, orientable case, a choice of orientation for X determines a generator $[X]$ for $H_n(X)$, called a *fundamental class*. For an orientation-preserving homeomorphism $f\colon X \to Y$, one has $f_*[X] = [Y]$; when f reverses orientation, $f_*[X] = -[Y]$.

If X is given as a CW complex with cells of dimension $\leq n$ and one cell e_n of degree n, one necessarily has $\partial e_n = 0$ in the cellular complex $C_*(X)$, with $[e_n]$ is a fundamental class.

The fundamental class defines a map $H^n(X;R) \to R$, given by evaluating cohomology classes on $[X]$. In the case of smooth manifolds and de Rham cohomology, it is a fact that *the evaluation map $H^n_{DR}(X) \to \mathbb{R}$ is the integration map $\int_X$.*[2]

2.3.2 Cap product and Poincaré duality

Degree-reversed homology $H_{-*}(X)$ is a graded module over the cohomology ring, via the cap product

$$\cap\colon H^k(X) \otimes H_j(X) \to H_{j-k}(X).$$

In cellular (co)homology, this is defined via the cellular approximation δ to the diagonal $\Delta\colon X \to X \times X$:

$$\cap\colon C^*(X) \otimes C_*(X) \xrightarrow{\mathrm{id}\otimes\delta_*} C^*(X) \otimes C_*(X) \otimes C_*(X) \xrightarrow{\mathrm{ev}\otimes\mathrm{id}} C_*(X).$$

If $f\colon X^n \to Y^m$ is a smooth map of manifolds, and X is closed and oriented, then the resulting map $\cdot \cap f_*[X]\colon H^n_{DR}(Y) \to H^0_{DR}(Y) = \mathbb{R}$ is given by

$$[\eta] \cap f_*[X] = \int_X f^*\eta$$

for $\eta \in \Omega^n(Y)$ with $d\eta = 0$.

The Poincaré duality theorem [15] says that for X a closed and oriented topological manifold of dimension n, the map

$$D_X = \cdot \cap [X]\colon H^*(X) \to H_{n-*}(X)$$

is an isomorphism.

We shall write D^X for the inverse isomorphism.

2.3.3 Intersection of submanifolds

Suppose that X^n, Y^{n-p} and Z^{n-q} are closed, oriented manifolds, and $f\colon Y \to X$ and $g\colon Z \to X$ a pair of maps. We then have cohomology classes

$$c_Y = D^X(f_*[Y]) \in H^p(X), \quad c_Z = D^X(g_*[Z]) \in H^q(X),$$

and a cup product

$$c_Y \cup c_Z \in H^{p+q}(X).$$

This has the following geometric interpretations: let f' be a map homotopic to f and transverse to g, meaning that whenever $f(y) = g(z) = x$, say, one has $T_xX = Df'(T_yY) + Dg(T_zZ)$. Such an f' always

[2] To prove assertions such as this one, it is helpful to generalize them to the non-compact case. For that purpose one must work with compactly supported cohomology and compactly supported differential forms.

exists, according to standard transversality theory. Then the fiber product $P = Y_{f'} \times_g Z$ is naturally a closed, oriented manifold of dimension $n-(p+q)$, equipped with a map $\phi = (f', g)\colon P \to X$. This defines a class

$$c_P = D^X(\phi_*[P]) \in H^{p+q}(X),$$

and one has

$$c_P = c_Y \cup c_Z.$$

In particular, when f and g are embeddings, P is their intersection, and ϕ its embedding in X.

Intersection of submanifolds gives a geometric realization of the cup product, but it has a limitation: not every homology class is representable by a submanifold, or even by a smooth map from another manifold. Codimension 0 classes, $c \in H^n(X)$, are representable by maps from a manifold (several disjoint copies of X), but not embedded submanifolds. Two important classes of homology classes that *are* realizable by embedded submanifolds are the following:

- Codimension 1 classes, elements of $H_{n-1}(X) \cong H^1(X)$. Indeed, the map $[X, S^1] \to H^1(X;\mathbb{Z})$ sending the homotopy class of $f\colon X \to S^1$ to $f^*\omega$ (where $\omega \in H^1(S^1)$ is a generator) is bijective. The Poincaré dual to $f^*\omega$ is represented by a hypersurface $H_t = f^{-1}(t)$, where t is a regular value (the normal bundle to H_t is identified with T_tS^1, hence is oriented; this, with the chosen orientation of TX, determines and orientation for H_t).
- Codimension 2 classes, elements of $H_{n-2}(X) \cong H^2(X)$. Indeed, the map $[X, \mathbb{C}P^\infty] \to H^2(X)$ sending $[f]$ to f^*c is bijective: here $c \in H^2(\mathbb{C}P^\infty) = \mathbb{Z}$ is the generator which restricts to the point class in $H^2(\mathbb{C}P^1) = H_0(\mathbb{C}P^1)$). One can represents the homotopy class f by a map $g\colon X \to \mathbb{C}P^N$ for a finite N; then f^*c is Poincaré dual to $g^{-1}(D)$, where D is a hyperplane in $\mathbb{C}P^N$ to which g is transverse.

In both these cases, the bijectivity of the relevant map is an instance of the bijection $H^n(X) \cong [X, K(\mathbb{Z}, n)]$, where the Eilenberg–MacLane space $K(\mathbb{Z}, n)$ is characterized as having vanishing homotopy groups $\pi_k K(\mathbb{Z}, n)$ $(k > 0)$ except that $\pi_n K(\mathbb{Z}, n) \cong \mathbb{Z}$. The circle S^1 is a $K(\mathbb{Z}, 1)$, since its fundamental group is $\mathbb{Z}$ and its universal cover is contractible, while $\mathbb{C}P^\infty$ is a $K(\mathbb{Z}, 2)$ (it is simply connected, so $\pi_2 = H_2 = \mathbb{Z}$ by Hurewicz; higher homotopy groups vanish by the exact sequence of the fibration $S^1 \to S^{2N+1} \to \mathbb{C}P^N$).

Codimension 3 classes are not realizable, in general, as discovered by R. Thom [19] For instance, the 10-dimensional compact Lie group $\mathsf{Sp}(2)$ has cohomology $H^*(\mathsf{Sp}(2)) \cong \Lambda[x_3, x_7]$, the exterior algebra on generators of degrees 3 and 7. The class x_3 is not realizable by an embedded submanifold [2].

3 The intersection form

4 The cup product in middle dimensional cohomology

Suppose that M is a closed, oriented manifold of even dimension $2n$. Its middle-degree cohomology group $H^n(M)$ then carries a bilinear form, the *cup-product form*,

$$H^n(M) \times H^n(M) \to \mathbb{Z}, \quad (x,y) \mapsto x \cdot y := \mathsf{eval}\,(x \cup y, [M]).$$

The cup product $x \cup y$ lies in $H^{2n}(M)$, and eval denotes the evaluation of a cohomology class on a homology class.

The cup-product form is skew-symmetric when n is odd, and symmetric when n is even.

Lemma 4.1 *$x \cdot y$ is equal to the evaluation of x on $D_X y$.*

Proof Under the isomorphism $H_0(X) \cong \mathbb{Z}$ sending the homology class of a point to 1, one has $x \cdot y = (x \cup y) \cap [X] = x \cap (y \cap [X]) = \langle x, D_X y \rangle$. □

As we saw in Lecture 2, the cup product, when interpreted as a pairing on homology $H_n(M) \times H_n(M) \to \mathbb{Z}$ by applying Poincaré duality to both factors, amounts to an intersection product. Concretely, if S and S' are closed, oriented submanifolds of M, of dimension n and intersecting transversely, and $s = D^M[S]$, $s' = D^M[S']$, then

$$s \cdot s' = \sum_{x \in S \cap S'} \varepsilon_x,$$

where $\varepsilon_x = 1$ if, given oriented bases $(e_1, \dots, e_n)$ of T_xS and $(e'_1, \dots, e'_n)$ of T_xS', the basis $(e_1, \dots, e_n, e'_1, \dots, e'_n)$ for T_xM is also oriented; otherwise, $\varepsilon_x = -1$.

Notation: For an abelian group A, we write

$$A' := A/A_{\mathsf{tors}} \tag{1}$$

for the largest torsion-free quotient of A.

The cup product form necessarily descends to a form on the free abelian group $H^n(M)'$. We shall denote the latter form by Q_M.

Proposition 4.2 *The cup-product form Q_M is non-degenerate, i.e., the group homomorphism*

$$H^n(M)' \to \mathsf{Hom}(H^n(M'), \mathbb{Z}), \quad x \mapsto (y \mapsto x \cdot y)$$

is an isomorphism.

Proof By the lemma, an equivalent assertion is that evaluation defines a non-degenerate pairing of the torsion-free quotients $H^n(X)'$ and $H_n(X)'$. This is true as a matter of homological algebra: it is a weak form of the cohomological universal coefficients theorem. □

If we choose an integral basis $(e_1, \dots, e_b)$ of $H^n(M)'$, we obtain a square matrix Q of size $b \times b$, where $b = b_n(M)$, with entries $Q_{ij} = e_i \cdot e_j$. It is symmetric or skew symmetric depending on the parity of n. Non-degeneracy of the form Q_M is equivalent to the unimodularity condition $\det Q = \pm 1$.

Proposition 4.3 *Suppose that N is a compact, oriented manifold with boundary M, and $i\colon M \to N$ the inclusion. Let $L = \operatorname{im} i^* \subset H^n(M;\mathbb{R})$. Then (i) L is isotropic, i.e., $x \cdot y = 0$ for $x, y \in L$; and (ii) $\dim L = \frac{1}{2}\dim H^n(M;\mathbb{R})$.*

Proof (i) We have $i^*u \cdot i^*v = \mathsf{eval}(i^*(u \cup v), [M]) = \mathsf{eval}(u \cup v, i_*[M])$. But $i_*[M] = 0$, since the fundamental cycle of M is bounded by that of N.

(ii) (I follow the proof in [15].) There is a commutative diagram with exact rows as follows:

$$\begin{tikzcd}
\cdots \arrow[r] & H^n(N;\mathbb{R}) \arrow[r, "i^*"] \arrow[d, "D_N"] & H^n(M;\mathbb{R}) \arrow[r, "\delta"] \arrow[d, "D_M"] & H^{n+1}(N,M;\mathbb{R}) \arrow[r, "q"] \arrow[d, "D_{N,M}"'] & H^{n+1}(N;\mathbb{R}) \arrow[r] \arrow[d, "D_N"] & \cdots \\
\cdots \arrow[r] & H_{n+1}(N,M;\mathbb{R}) \arrow[r, "\partial"] & H_n(M;\mathbb{R}) \arrow[r, "i_*"] & H_n(N;\mathbb{R}) \arrow[r, "p"] & H_n(M,N;\mathbb{R}) \arrow[r] & \cdots
\end{tikzcd}$$

The top row is the cohomology exact sequence of the pair (N,M), the bottom row the homology exact sequence of the same pair; and the vertical maps are duality isomorphisms: D_M is Poincaré duality, the remaining vertical maps Poincaré–Lefschetz duality (which we have not reviewed). Fix a complement K to L in $H^n(M;\mathbb{R})$. We shall show that $\dim K = \dim L$.

From exactness of the top row, we see that $L = \ker\delta$, so $K \cong \operatorname{im}\delta \cong \ker q$. But $\ker q \cong \ker p \cong \operatorname{im} i_*$, so $K \cong \operatorname{im} i_*$. Real cohomology is dual to real homology, and i^* is dual to i_*. Thus $\operatorname{im} i^*$ is the annihilator of $\ker i_*$, and $\dim \operatorname{im} i^* = \dim \operatorname{im} i_*$, i.e. $\dim L = \dim K$. □

4.1 Symmetric forms over $\mathbb{R}$

We concentrate now on the case where n is even, so Q_M is symmetric.

Definition 4.4 A *unimodular lattice* (Λ, σ) is a free abelian group Λ of finite rank, together with a non-degenerate, symmetric bilinear form $\sigma\colon \Lambda \times \Lambda \to \mathbb{Z}$.

Q_M is a unimodular lattice.

Recall that given a symmetric bilinear form σ on a real finite-dimensional real vector space V, there is an orthogonal decomposition

$$V = R \oplus V^+ \oplus V^-,$$

where $R = \{v \in V : \sigma(v,\cdot) = 0\}$ is the radical, and where σ is positive-definite on V^+ and negative-definite on V^-. The dimensions $\dim V^\pm$ are invariants of (V,σ), and together with that of R they are *complete* invariants.

We define the signature $\tau(\Lambda)$ of a unimodular lattice (Λ,σ) to be that of $\Lambda \otimes \mathbb{R}$, and the signature of M to be that of Q_M.

The fact that this $\tau(M)$ is an invariant of a closed oriented manifolds (of dimension divisible by 4) immediately gives the

Proposition 4.5 *A $4k$-dimensional closed oriented manifold M admits an orientation-reversing self-diffeomorphism only if its signature vanishes.*

Theorem 4.6 *(a) Let Y be a an oriented cobordism between 4-manifolds X_1 and X_2 (i.e., Y is a compact oriented 5-manifold with boundary ∂Y, together with an oriented diffeomorphism $\partial Y \cong -X_1 \amalg X_2$). Then $\tau_{X_1} = \tau_{X_2}$.*

(b) Conversely, if $\tau_{X_1} = \tau_{X_2}$, an oriented cobordism exists.

Proof (a) By the proposition above, the cup-product form of $-X_1 \amalg X_2$ admits a middle-dimensional isotropic subspace. It follows, as a matter of algebra, that $\tau(-X_1 \amalg X_2) = 0$. But the cup-product form $Q_{-X_1 \amalg X_2}$ is the orthogonal sum of $Q_{-X_1} = -Q_{X_1}$ and Q_{X_2}, so $\tau(-X_1 \amalg X_2) = \tau(X_2) - \tau(X_1)$.

(b) [Sketch.] It follows from Thom's cobordism theory that the group Ω_d of cobordism classes of closed oriented d-manifolds, under disjoint union, is isomorphic to the homotopy group $\pi_{d+k} M\,\mathrm{SO}(k)$ in the 'stable range' where k is reasonably large. Here $M\,\mathrm{SO}(k)$ is the Thom space of the universal vector bundle $E\,\mathrm{SO}(k) \to B\,\mathrm{SO}(k)$ over the classifying space for the Lie group $\mathrm{SO}(k)$. Note that the homology group $H_{d+k}(M\,\mathrm{SO}(k))$ is isomorphic (by the Thom isomorphism) to $H_d(B\,\mathrm{SO}(k))$; so there are Hurewicz maps $\Omega_d \to H_d B\,\mathrm{SO}(k)$, and in particular map $\Omega_4 \to H_4(B\,\mathrm{SO}(k)) \cong \mathbb{Z}$. Thom proves that $\Omega_d = 0$ for $d \leq 3$ and that $\Omega_4 \to \mathbb{Z}$ is an isomorphism. The signature homomorphism $\tau \colon \Omega_4 \to \mathbb{Z}$ is surjective, since $\tau(\mathbb{C}P^2) = 1$, and therefore an isomorphism. □

4.2 Characteristic vectors

Having examined Q_M over $\mathbb{R}$, we turn next to an aspect of its mod 2 arithmetic.

Definition 4.7 A *characteristic vector* c for a unimodular lattice is an element $c \in \Lambda$ such that $c \cdot x \equiv x \cdot x \mod 2$ for all $x \in \lambda$.

Lemma 4.8 *The characteristic vectors form a coset of 2Λ in Λ.*

Proof Let $\lambda = \Lambda \otimes_{\mathbb{Z}} (\mathbb{Z}/2)$. It is a $\mathbb{Z}/2$-vector space of dimension d, with a symmetric pairing $(\cdot,\cdot)$, still non-degenerate. The map $\lambda \to \lambda$ given by $z \mapsto (z,z)$ is $\mathbb{Z}/2$-linear, and so by non-degeneracy can be represented as $(z,z) = (\bar{c}, z)$ for a unique element $\bar{c} \in \lambda$. The characteristic vectors c are precisely the lifts of $\bar{c}$ to Λ. □

Definition 4.9 A unimodular lattice (Λ, σ) is called *even* if 0 is characteristic, i.e., if (x,x) is always even; otherwise the lattice is called *odd*. The property of being even or odd is called the *type* of the lattice.

Lemma 4.10 *For any two characteristic vectors c and c', one has $\sigma(c,c) \equiv \sigma(c',c')$ modulo 8.*

Proof Write $c' = c + 2x$. Then

$$\sigma(c',c') = \sigma(c,c) + 4(\sigma(c,x) + \sigma(x,x)),$$

and $\sigma(c,x) + \sigma(x,x)$ is even. □

Theorem 4.11 (Hasse–Minkowski) *A unimodular form σ on a lattice $\Lambda \cong \mathbb{Z}^r$, which is indefinite (i.e. neither positive- nor negative-definite) is determined, up to isomorphism, by its rank r, signature $\tau \in \mathbb{Z}$, and type $t \in \mathbb{Z}/2$.*

This is a deep and powerful result which we will not prove; see [18]. The key point is to find an isotropic vector, i.e. a vector $x \neq 0$ such that $\sigma(x,x) = 0$. It suffices to find an isotropic vector x in $\Lambda \otimes \mathbb{Q}$; and according Hasse–Minkowski's local-to-global principle for quadratic forms over $\mathbb{Q}$, for existence of such an isotropic vector it is necessary and sufficient that there are isotropic vectors in $\Lambda \otimes \mathbb{R}$ (to which indefiniteness is clearly the only obstruction) and in $\Lambda \otimes \mathbb{Q}_p$ for each prime p. Quadratic forms over the p-adics $\mathbb{Q}_p$ can be concretely understood, and it turns out that (when the rank is at least 5) there is a p-adic isotropic vector as soon as the form is indefinite (additional arguments are needed for low rank).

Let I_+ denote the unimodular lattice $\mathbb{Z}$ with form $(x,y) \mapsto xy$; let $I_- = -I_+$. Part of the statement of Hasse–Minkowski is that, if Λ is odd and indefinite, it is isomorphic to a direct sum

$$rI_+ \oplus sI_-$$

for suitable r and s. To prove this, one uses an isotropic vector to find an orthogonal direct sum decomposition $\Lambda = I_+ \oplus I_- \oplus \Lambda'$. Then $I_+ \oplus \Lambda'$ and $I_- \oplus \Lambda'$ have lower rank than Λ, and both are odd. One of them is indefinite, so one can proceed by induction on the rank.

The classification of odd indefinite unimodular forms has the following

Corollary 4.12 *In any unimodular lattice, any characteristic vector c has $\sigma(c,c) \equiv \tau$ mod 8. In particular, the signature of an even unimodular lattice is divisible by 8.*

Proof The form $rI_+ \oplus sI_-$ has characteristic vector $c = (1, \dots, 1)$, for which one has $c^2 = \tau$. Thus for any characteristic vector one has $c^2 \equiv \tau$ modulo 8. By the classification, the corollary holds for odd, indefinite unimodular forms. We can make any unimodular form odd and indefinite by adding I_+ or I_-, which has the effect of adding or subtracting 1 to the signature. If c is characteristic for Λ then $c \oplus 1$ is characteristic for $\Lambda \oplus I_\pm$, with $(c \oplus 1)^2 = c^2 \pm 1$, so we deduce the corollary for Λ. □

The basic example of an *even* unimodular form is the lattice $U = \mathbb{Z}^2$ with $(a,b)^2 = 2ab$. Its matrix is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

To classify even indefinite unimodular forms one proceeds as follows. Suppose Λ_1 and Λ_2 are indefinite, unimodular and even, of the same rank and signature. One uses the existence of an isotropic vector to prove that $\Lambda_i \cong U \oplus \Lambda_i'$ for even unimodular lattices Λ_i'. From what has been proved about the odd case, one knows that $\Lambda_1' \oplus I_+ \oplus I_- \cong \Lambda_2' \oplus I_+ \oplus I_-$, and with some work one deduces that $\Lambda_1' \oplus U \cong \Lambda_2' \oplus U$, i.e., that $\Lambda_1 \cong \Lambda_2$.

4.3 The E_8 lattice

There is an important example of a positive-definite even unimodular form of rank 8. This is the form E_8 arising from the E_8 root system (or Dynkin diagram). Start with the lattice $\mathbb{Z}^8$ (standard inner product). Let $\Gamma \subset \mathbb{Z}^8$ be the sub-lattice formed by $x \in \mathbb{Z}^8$ with $x \cdot x$ even. Then E_8 is formed from Γ by adjoining the vector $\frac{1}{2}(e_1 + \cdots + e_8)$. Since this vector has length-squared 2, E_8 is even.

Exercise 4.13 (1) Let Λ be a lattice in $\mathbb{R}^n$ (with inner product inherited from $\mathbb{R}^n$) and $\Lambda' \subset \Lambda$ a sub-lattice of finite index $[\Lambda : \Lambda']$. Show that the determinants of the matrices representing these lattices are related by

$$\det \Lambda = [\Lambda : \Lambda'] \det \Lambda'.$$

(2) Show that $[\mathbb{Z}^8 : \Gamma] = 2$ and $[E_8 : \Gamma] = 2$.

(3) Deduce that $\det E_8 = 1$.

E_8 has basis $(v_1, \dots, v_8)$ where

$$v_i = e_{i+1} - e_i \quad (1 \leq i \leq 6), \quad v_7 = \tfrac{1}{2}(e_1 + e_8) - \tfrac{1}{2}(e_2 + \cdots + e_7), \quad v_8 = e_1 + e_2.$$

One has $v_i \cdot v_i = 2$; $v_1 \cdot v_2 = v_2 \cdot v_3 = \cdots = v_5 \cdot v_6 = -1$; $v_7 \cdot v_2 = -1$; $v_8 \cdot v_7 = 0$. All the other pairs are orthogonal. (One usually depicts this situation via the E_8 Dynkin graph.)

We typically prefer to use the negative-definite version $-E_8$. This has basis $(v_1, \dots, v_8)$ and matrix

$$-E_8 = \begin{bmatrix} -2 & 1 & & & & & 1 & \\ 1 & -2 & 1 & & & & & \\ & 1 & -2 & 1 & & & & \\ & & 1 & -2 & 1 & & & \\ & & & 1 & -2 & 1 & & \\ & & & & 1 & -2 & & \\ & 1 & & & & & -2 & 1 \\ & & & & & & 1 & -2 \end{bmatrix}.$$

The direct sum

$$rU \oplus s(\pm E_8)$$

is even unimodular of rank $2r \pm 8s$ and signature $\pm 8s$. By Hasse–Minkowski and the fact that the signature of an even unimodular form is divisible by 8, we see that every indefinite even unimodular form takes this shape.

4.4 Topological examples

It is straightforward to write down an example of a $4k$-manifold with cup-product form H: one can simply take $S^{2k} \times S^{2k}$. In particular, in 4 dimensions we have $S^2 \times S^2$.

In 4 dimensions, it is also easy to come up with an example with cup-product for I_+: one can take $\mathbb{C}P^2$, with its orientation as a complex surface. One has $H^2(\mathbb{C}P^2) = \mathbb{Z}$, the generator ℓ being the Poincaré dual to any projective line $L \subset \mathbb{C}P^2$. Any two such lines, L and L', if distinct, intersect positively at a single point, so $\ell \cdot \ell = 1$. We can get I_- as intersection form by taking $-\mathbb{C}P^2$ (i.e. reversing orientation).

Next time, we shall use characteristic classes of the tangent bundle to prove the following

Proposition 4.14 *Let X be a smooth quartic complex surface in $\mathbb{C}P^3$. Then X has even intersection form of rank 22 and signature -16.*

Thus from Hasse–Minkowski, we deduce that X has intersection form

$$3U \oplus 2(-E_8).$$

It is not a simple task to write down an integral basis for $H_2(X)$, let alone to calculate the intersection form explicitly, so Hasse–Minkowski is a convenient shortcut.

5 The intersection form and characteristic classes

5.1 Cohomology of 4-manifolds

The homology and cohomology of a closed oriented 4-manifold look like this—the two columns are related by Poincaré duality isomorphisms:

$$\begin{array}{ll}
H^0(X) = \mathbb{Z} = \mathbb{Z}\cdot 1 & H_4(X) = \mathbb{Z} = \mathbb{Z}\cdot [X] \\
H^1(X) = \mathsf{Hom}(\pi_1(X), \mathbb{Z}) & H_3(X) \\
H^2(X) \cong \mathsf{Hom}(H_2(X), \mathbb{Z}) \oplus H_1(X)_{tors.} & H_2(X) \\
H^3(X) & H_1(X) = \pi_1(X)^{ab} \\
H^4(X) & H_0(X) = \mathbb{Z} = \mathbb{Z}\cdot [\mathsf{point}]
\end{array}$$

All these groups are determined by π_1 and H^2. The cohomology ring includes the cup-product form Q_X.

However, a moment's reflection makes clear that there is further structure to consider—for instance, the form of the cup product $H^1 \otimes H^2 \to H^3$, the relation of integer to mod p cohomology, and how the Hurewicz map $\pi_2(X) \to H_2(X)$ fits into the picture (according to a theorem of Hopf, its cokernel is isomorphic to $H^2(B\pi_1(X))$.

In the simply connected case, the situation simplifies to the following:

$$\begin{array}{ll}
H^0(X) = \mathbb{Z} = \mathbb{Z}\cdot 1 & H_4(X) = \mathbb{Z} = \mathbb{Z}\cdot [X] \\
H^1(X) = 0 & H_3(X) = 0 \\
H^2(X) \cong \mathbb{Z}^d & H_2(X) \cong \mathbb{Z}^d \\
H^3(X) = 0 & H_1(X) = 0 \\
H^4(X) = \mathbb{Z} & H_0(X) = \mathbb{Z} = \mathbb{Z}\cdot [\mathsf{point}].
\end{array}$$

The the cup-product form Q_X on $H^2(X)$ fully determines the ring $H^*(X)$ and the module $H_*(X)$. All mod p or rational cohomology classes are reductions of integral ones, and the Hurewicz map $\pi_2 \to H_2$ is an isomorphism.

Conclusion: When X is simply connected, Q_X is the only (co)homological information we can find.

5.2 Characteristic classes

The next 4-manifold invariant we will study is the tangent bundle $TX \to X$ viewed as a distinguished rank 4 vector bundle. We look especially at its characteristic classes. While they will prove disappointing as tools for distinguishing 4-manifolds, they are very helpful both in in computing the intersection form Q_X.

5.2.1 Stiefel–Whitney classes

(See e.g. [9].) For any finite-rank vector bundle $V \to X$ over an arbitrary space X, there are *Stiefel–Whitney classes* $w_i(V) \in H^i(X;\mathbb{Z}/2)$, for $i \geq 0$, with $w_0 = 1$, vanishing for $i \gg 0$. The total Stiefel–Whitney class is

$$w(V) = w_0(V) + w_1(V) + w_2(V) + \cdots \in H^*(X;\mathbb{Z}/2).$$

Among such assignments, they are uniquely characterized by the following properties:

- For a map $f\colon Y \to X$, one has $w_i(f^*V) = f^*w_i(V)$.
- $w_i(V) = 0$ for $i > \operatorname{rank} V$.
- $w(U \oplus V) = w(U) \cup w(V)$.
- For the tautological line bundle $L \to \mathbb{R}P^1$, whose fiber over the line $\lambda \subset \mathbb{R}^2$ is λ, one has $w_1(L) \neq 0 \in H^1(\mathbb{R}P^1; \mathbb{Z}/2) = \mathbb{Z}/2$.

Short exact sequences of topological (or smooth) vector bundles over paracompact Hausdorff spaces necessarily split, e.g. by taking orthogonal complements with respect to a Euclidean metric. Thus the formula for $W = U \oplus V$ (the 'Whitney sum formula') is applicable as soon as one has a short exact sequence $0 \to U \to W \to V \to 0$.

The Stiefel–Whitney classes have the following two properties:

(1) For path-connected spaces X, under the standard isomorphism $H^1(X; \mathbb{Z}/2) = \mathrm{Hom}(\pi_1(X), \mathbb{Z}/2)$, $w_1(V)$ maps to the orientation character of V. For an orientable vector bundle one therefore has $w_1(V) = 0$—in particular, $w_1(TM) = 0$ for orientable manifolds M.

(2) When M is a closed, smooth n-manifold, and V has rank r, the *top* Stiefel–Whitney class $w_r(V) \in H^r(M; \mathbb{Z}/2)$ is Poincaré dual to the class in $H_{n-r}(M; \mathbb{Z}/2)$ of the zero-locus of a transverse-to-zero section $s\colon M \to V$. (Hence, in the oriented case, w_r is the mod 2 reduction of the Euler class.)

Example 5.1 We compute $w(T(\mathbb{R}P^n))$. To do so, let $V = \mathbb{R}^{n+1}$, so $\mathbb{R}P^n = \mathbb{P}V$. A point of $\mathbb{R}P^n$ is a line $\lambda \subset V$, and

$$T_\lambda \mathbb{P}V \cong \mathrm{Hom}(\lambda, V)/\mathrm{Hom}(\lambda, \lambda)$$

canonically. Let $L \to \mathbb{P}V$ be the tautological line bundle, and $\mathbf{V} \to \mathbb{P}V$ the trivial bundle with fiber V. Since $\mathrm{Hom}(\lambda, \lambda) = \mathbb{R} \cdot \mathrm{id}_\lambda$, there is a short exact sequence

$$0 \to \mathbb{R} \to \mathrm{Hom}(L, \mathbf{V}) \to T(\mathbb{P}V) \to 0.$$

We have $\mathrm{Hom}(L, \mathbf{V}) \cong L^* \oplus \cdots \oplus L^*$ ($n+1$ summands).

Let $H \in H^1(\mathbb{P}V; \mathbb{Z}/2) = \mathbb{Z}/2$ be the non-trivial element. We find

$$w(T\mathbb{P}V) = (1+H)^{n+1}.$$

Tangent bundles of 4-manifolds

Theorem 5.2 (Wu) *For X a closed 4-manifold, $(w_1^2 + w_2)(TX)$ is the characteristic element of $H^2(X; \mathbb{Z}/2)$, i.e. $u \cup u = (w_1^2 + w_2)(TX) \cup u$ for all $u \in H^2(X; \mathbb{Z}/2)$.*

Proof This is an instance of Wu's theorem about Stiefel–Whitney classes $w(TX)$ of closed n-manifolds, which says that $\mathrm{Sq}\, v = w(TX)$ where $v = \sum v_i$ is the class which represents the Steenrod square Sq^i with respect to the cup-product form: $\mathrm{Sq}^i x_{n-i} = v_i \cup x_{n-i}$ [16]. When $n = 4$, these conditions says that $v = 1 + v_1 + v_2$; that $v_1 = w_1$; and hat v_2 is the characteristic element. Wu's formula then gives $w_1^2 + w_2$ is characteristic. See J. Milnor and J. Stasheff, *Characteristic classes*. □

Corollary 5.3 *On a simply connected, closed 4-manifold, (i) $w_2(TX)$ is fully determined by the intersection form Q_X; and (ii) $w_2(TX) = 0$ if and only if Q_X is even.*

The top Stiefel–Whitney class $w_4(TX)$, when evaluated on $[X]$, counts the zeros of a vector field mod 2. Hence $w_4(TX)[X]$ is the mod 2 Euler characteristic $\bar{\chi}(X) \in \mathbb{Z}/2$: no new information beyond Q_X.

Theorem 5.4 (Hirzebruch–Hopf) *On a closed, oriented 4-manifold one has* $w_3(TX) = 0$.

In the simply connected case, this theorem is trivially true: $H^3(X;\mathbb{Z}/2) \cong H_1(X;\mathbb{Z}/2) = 0$. In general, according to Wu's formula and the fact that $w_2 = v_2$, one has

$$w_3 = \mathsf{Sq}^1 w_2,$$

and $\mathsf{Sq}^1 x$ is the Bockstein operation β measuring whether x has a lift to $\mathbb{Z}$-coefficients. Thus the theorem says w_2 has an integral lift. This form will be critical to us, since existence of an integral lift of w_2 is the obstruction to a $\mathrm{Spin^c}$-structure.

5.2.2 Chern classes

For any finite-rank complex vector bundle $E \to X$ there are *Chern classes* $c_i(E) \in H^{2i}(X;\mathbb{Z})$, $i \geq 0$, $c_0 = 1$, vanishing for $i \gg 0$. The total Chern class is $c(E) = 1 + c_1(E) + c_2(E) + \cdots \in H^{even}(X)$. They are uniquely characterized by the following properties:

- For a map $f\colon Y \to X$, one has $c_i(f^*E) = f^*c_i(E)$.
- $c_i(E) = 0$ for $i > \mathsf{rank}\, E$.
- $c(E \oplus F) = c(E) \cup c(F)$.
- For the tautological line bundle $L \to \mathbb{C}P^1$ (whose fiber over the line $\lambda \subset \mathbb{C}^2$ is λ), one has $\mathsf{eval}(c_1(L), [\mathbb{C}P^1]) = -1$ (here we use the complex orientation of $\mathbb{C}P^1$).

The top Chern class $c_r(E)$ of a rank r vector bundle over a manifold X is Poincar'e dual to the $[Z]$ class of the zero-locus of a section $s\colon X \to E$. In other words, $c_r(E)$ is the *Euler class* $e(E_{\mathbb{R}})$ of the underlying real oriented vector bundle.

Example 5.5 The Chern classes of $T(\mathbb{C}P^n)$ can be computed by a formally identical argument to the one we used to compute $w(T(\mathbb{R}P^n))$. The result is

$$c(T(\mathbb{C}P^n)) = (1 + H)^{n+1},$$

where $H = -c_1(L) \in H^2(\mathbb{C}P^n)$ is the hyperplane class.

The topological line bundles over X form a group $\mathrm{Pic}(X)$ under tensor product $\otimes$—the (topological) Picard group. The first Chern class defines a homomorphism

$$c_1\colon\ \mathrm{Pic}(X) \to H^2(X;\mathbb{Z}).$$

That

$$c_1(L_1 \otimes L_2) = c_1(L_1) + c_1(L_2)$$

is clear (when X is a manifold, which includes the 'universal' case $\mathbb{C}P^N$) from the zero-locus interpretation.

Theorem 5.6 $c_1\colon\ \mathrm{Pic}(X) \to H^2(X;\mathbb{Z})$ *is an isomorphism.*

Proof Both $\mathrm{Pic}(X)$ and $H^2(X;\mathbb{Z})$ are homotopy classes of maps to $\mathbb{C}P^\infty$, a space which is simultaneously the classifying space $B\,\mathrm{U}(1)$ for line bundles and the Eilenberg–MacLane space $K(\mathbb{Z},2)$. The universal line bundle $L \to B\,\mathrm{U}(1)$ is the tautological bundle $L \to \mathbb{C}P^\infty$, whose Chern class $c_1(L) = -H$ is also the universal degree 2 cohomology class. □

Theorem 5.7 *For a complex vector bundle $E \to X$ regarded as a real vector bundle $E_\mathbb{R}$, $w_{2i}(E_\mathbb{R})$ is the mod 2 reduction of $c_i(E)$ while $w_{2i+1}(E_\mathbb{R}) = 0$.*

Proof We have $w(E_\mathbb{R} \oplus F_\mathbb{R}) = w(E_\mathbb{R})w(F_\mathbb{R})$, and $c(E \oplus F) = c(E)c(F)$; thus the two assertions are compatible with direct sums. They are also compatible with pullbacks. By the splitting principle [16], any complex vector bundle is the pullback of a sum of line bundles. Thus it suffices to treat the case of a complex line bundle. But then it suffices to treat the case of the universal line bundle, i.e. the tautological line bundle $\lambda \to \mathbb{C}P^\infty$. One has $w_1(\lambda) = 0$, so it suffices to show that $w_2(\lambda) \neq 0$. But the restriction of λ to $\mathbb{R}P^2 \subset \mathbb{C}P^2 \subset \mathbb{C}P^\infty$ is the complexification $\ell \otimes \mathbb{C} = \ell \oplus i\ell$ of the tautological line bundle $\ell \to \mathbb{R}P^2$. Thus $w_2(\lambda)|_{\mathbb{R}P^2} = w_1(\ell)^2 \neq 0$. □

An *almost complex structure* J on a manifold X^{2n} is an endomorphism of TX such that $J^2 = -\mathrm{id}$. It makes TX a complex vector bundle, with i acting via J. Complex manifolds are, of course, almost complex.

An almost complex manifold (X,J) has Chern classes

$$c_j(TX,J) \in H^{2j}(X).$$

The top Chern class $c_n(TX,J) \in H^{2n}(X)$, evaluated on the fundamental class $[X]$ to give an integer, is (up to sign) the Euler characteristic $\chi(X)$.

Example: complex hypersurfaces. Let $X \subset \mathbb{C}P^n$ be the complex hypersurface cut out as $F = 0$ for a homogeneous polynomial F of degree d (smoothness is guaranteed if we assume that there is no point where $F = 0$ and $\partial_i F = 0$ for all i).

We recall two principles from complex analytic geometry: first, a hypersurface D in a complex manifold M defines an invertible sheaf $\mathcal{O}_M(D)$, the sheaf of meromorphic functions with simple poles along D. We can equally view this sheaf as the section of a holomorphic line bundle $\mathcal{L}_D$. Second, the restriction $\mathcal{L}_D|_D$ is identified with the holomorphic normal bundle $N_{D/M}$.

In the case of $\mathbb{P}^n$, a holomorphic line bundle is determined by its degree (or first Chern class), since $H^1(\mathcal{O}) = 0$. Thus $\mathcal{O}_{\mathbb{P}^n}(X) \cong \mathcal{O}(1)^{\otimes d}$, where $\mathcal{O}(1)$ is the dual to the tautological line bundle $\mathcal{O}(-1)$. So we have $N_{X/\mathbb{P}^n} \cong (\mathcal{O}(1)|_X)^{\otimes d}$. Let $H = c_1(\mathcal{O}(1)) \in H^2(\mathbb{P}^n)$, and let $h = H|_X$. Then $c(N_{X/\mathbb{P}^n}) = 1 + dh$.

From the short exact sequence $0 \to TX \to T(\mathbb{P}^n)|_X \to N_{X/\mathbb{P}^n} \to 0$ and the formula for $c(T\mathbb{P}^n)$, we deduce

$$c(TX)(1 + dh) = (1 + h)^{n+1},$$

i.e., for $1 \leq j < n$,

$$c_j(TX) + dc_{j-1}(TX)h = \binom{n+1}{j} h^j$$

Thus

$$c_1(TX) = (n + 1 - d)h,$$

and our formula recursively determines the higher Chern classes. For the case of complex hypersurfaces in $\mathbb{P}^3$,

$$c_1(TX) = (4-d)h, \quad c_2(TX) = (d^2-4d+6)h^2.$$

Now, $D^{\mathbb{P}^3}[X] = dH \in H^2(\mathbb{P}^3)$, so

$$c_2(TX)[X] = d(d^2-4d+6).$$

This is just the Euler characteristic $\chi(X)$, so

$$\chi(X) = d(d^2-4d+6).$$

It follows from the Lefschetz hyperplane that hypersurfaces in a projective space $\mathbb{C}P^n$ for $n \geq 3$ are *simply connected*. So $b_1(X) = 0$. With $n = 3$, we then get

$$b_2(X) = \chi(X) - 2 = d^3 - 4d^2 + 6d - 2.$$

5.2.3 Pontryagin classes

The Pontryagin classes of a real vector bundle $V \to X$ are defined as follows:

$$p_i(V) = (-1)^i c_{2i}(V \otimes \mathbb{C}) \in H^{4i}(X).$$

They satisfy the naturality property $p_i(f^*V) = f^*p_i(V)$.

In the case of an oriented 4-manifold X, the only non-trivial Pontryagin class of TX is $p_1(TX) \in H^4(X) = \mathbb{Z}$.

Lemma 5.8 *For a closed oriented 4-manifold X, the integer $p_1(TX)[X] \in \mathbb{Z}$ depends only on the oriented cobordism class of X.*

Proof If Y is a cobordism from X_1 to X_2, one has $TY|_{X_i} \cong TX_I \oplus \underline{\mathbb{R}}$. Consequently $p_1(TY|_{X_i}) = p_1(TX_i)$. Thus $p_1(TX_1)[X_1] = p_1(TY)[X_1] = p_1(TY)[X_2] = p_2(TX_2)[X_2]$. □

Lemma 5.9 *For an almost complex manifold X, one has $p_1(TX) = (c_1^2 - 2c_2)(TX, J)$.*

Proof Let V be a complex vector space, $\bar{V}$ its conjugate (the same space, with i now acting by the old action of $-i$), and $V_{\mathbb{R}}$ the underlying real vector space. Then there is a $\mathbb{C}$-linear isomorphism

$$V_{\mathbb{R}} \otimes \mathbb{C} \cong V \oplus \bar{V}.$$

This generalizes to complex vector bundles $V \to X$:

$$V_{\mathbb{R}} \otimes \mathbb{C} \cong V \oplus \bar{V}.$$

Thus

$$p_1(V_{\mathbb{R}}) = -c_2(V_{\mathbb{R}} \otimes \mathbb{C}) = -c_2(V \oplus \bar{V}) = -c_2(V) - c_2(\bar{V}) - c_1(V)c_1(\bar{V}).$$

In the case where V is a direct sum $\lambda_1 \oplus \cdots \oplus \lambda_r$ of line bundles, one has $\bar{V} \cong \lambda_1^* \oplus \cdots \oplus \lambda_r^*$. With $\ell_i = c_1(\lambda_i)$, one then has $c(V) = \prod(1+\ell_i)$ and $c(\bar{V}) = \prod(1-\ell_i)$, so that $c_j(\bar{V}) = (-1)^j c_j(V)$. The *splitting principle* [] implies that the identity $c_j(\bar{V}) = (-1)^j c_j(V)$ remains true even when V is not such a direct sum. Hence

$$p_1(V_{\mathbb{R}}) = c_1(V)^2 - 2c_2(V).$$

The lemma follows by applying this formula to $V = TX$. □

Example 5.10 When $X = \mathbb{C}P^n$, one has $c_1(TX) = (n+1)h$ and $c_2(TX) = \frac{1}{2}n(n+1)h^2$, so

$$p_1(TX) = -(n+1)h^2.$$

In particular,

$$p_1(T(\mathbb{C}P^2)) = -3h^2.$$

Theorem 5.11 (Hirzebruch signature theorem) *For a closed oriented 4-manifold X, one has $p_1(TX)[X] = 3\tau(X)$.*

Proof We have seen that the signature τ defines an isomorphism $\Omega_4 \to \mathbb{Z}$ of the cobordism group with $\mathbb{Z}$. By the lemma, $X \mapsto p_1(TX)[X]$ is another such homomorphism, so we need only compare them for one example with non-zero signature. In the case of $\mathbb{C}P^2$, $\tau = 1$ and $p_1(TX)[X] = 3$. □

Example: Complex hypersurfaces We are now in a position to compute the intersection form for a degree d complex hypersurface X_d in $\mathbb{C}P^3$:

Theorem 5.12 *For $d \geq 2$, the intersection form of X_d is indefinite, and is characterized up to equivalence by the facts that*

$$\begin{aligned} b_2(X_d) &= d^3 - 4d^2 + 6d - 2, \\ \tau(X_d) &= -\frac{1}{3}(d+2)d(d-2), \\ \mathsf{type}\,(X_d) &= d \mod 2. \end{aligned}$$

We already computed b_2. We have $p_1(TX_d) = (c_1^2 - 2c_2)(TX_d)[X_d] = (4-d^2)dh^2$, and so $\tau(X) = \frac{1}{3}p_1(TX)[X] = \frac{1}{3}(4-d^2)d$. The type is determined by w_2, which is the mod 2 reduction of $c_1 = (4-d)h$.

Example 5.13 A *K3 surface* is a compact complex surface with $b_1 = 0$ and $c_1(TX) = 0$. Quartic hypersurfaces $X \subset \mathbb{P}^3$ are examples of K3 surfaces. From our formulas, we have

$$b_2(X) = 24, \quad \tau(X) = -16, \quad \mathsf{type}(X) = \mathsf{even}.$$

By Hasse–Minkowski, then,

$$H^2(X) \cong 3U \oplus 2(-E_8).$$

6 Tangent bundles of 4-manifolds

The aim of this lecture is to show how, for a closed, simply connected 4-manifold X, $w_2(TX)$ measures non-triviality of the tangent bundle over 2-skeleton, while $\tau(X)$ and $\chi(X)$ then measure non-triviality over the 4-skeleton.

6.1 Vector bundles and obstruction theory

A good reference for this section is A. Hatcher's *Vector bundles and K-theory*.

6.2 Obstruction theory

Obstruction theory addresses the following problem:

Suppose $E \to X$ is a topological fiber bundle. When does it admit a section $s\colon X \to E$?

We shall assume X is a simply connected CW complex, with k-skeleta X^k, starting from $X^0 = \{x\}$, and that F is the fiber of E over x. The strategy is to construct s over X^k by induction on k. To this end, it addresses the following question:

Suppose s^k is a given section over X^k. When can we extend it to X^{k+1}?

Suppose $\Phi\colon (D^{k+1}, S^k) \to (X^{k+1}, X^k)$ is the inclusion of an $(k+1)$-cell, and $\phi = \Phi|_{S^k}\colon S^k \to X^k$ its attaching map. We have a section $\phi^* s^k$ of $\phi^* E \to X^k$. Now, $\phi^* E \to S^k$ is a trivial bundle, with a canonical-up-to-homotopy trivialization, because it extends to $\Phi^* E \to D^{k+1}$. So one can think of $\phi^* s^k$ as a map $X^k \to E_{\Phi(0)}$ to the fiber over $\Phi(0)$, defined up to homotopy.

A homotopy class of paths in X a basepoint x to $\Phi(0)$ defines an isomorphism $\pi_k E_{\Phi(0)} \cong \pi_k F$, and since X is simply connected, this isomorphism is canonical. Thus s^k determines an element of $\pi_k F$.

Running through the $(k+1)$-cells, we get a map

$$\{(k+1)\text{-cells}\} \to \pi_k(F),$$

i..e., a cellular cochain

$$o^{k+1} \in C^{k+1}(X; \pi_k(F)).$$

The key results of obstruction theory are as follows

- o^{k+1} is a cocycle.
- Its cohomology class $\mathfrak{o}^{k+1} = [o^{k+1}] \in H^{k+1}(X; \pi_k F)$ depends only on the homotopy class of s^k.
- If $\mathfrak{o}^{k+1} = 0$, one can extend s^k to s^{k+1}.
- If $\pi_i(F) = 0$ for $i < k$, then the *primary obstruction* $\mathfrak{o}^{k+1} \in H^{k+1}(X; \pi_k F)$ is an invariant $\mathfrak{o}^{k+1}(E)$ of the fiber bundle.
- The above assertions are valid also when $\pi_1(X)$ is non-trivial but acts trivially in $\pi_i(F)$ for $i \leq k$.

6.2.1 The Stiefel–Whitney classes as primary obstructions

Let $E \to X$ be a rank n euclidean vector bundle, and let $V_k(E) \to X$ be the associated fiber bundle whose fiber is $V_k(E_x)$, the Stiefel manifold of orthonormal k-frames in E_x. The typical fiber is $V_k(\mathbb{R}^n)$.

Using the homotopy exact sequence of the fibration $V_{k-1}(\mathbb{R}^n) \to V_k(\mathbb{R}^n) \to V_1(\mathbb{R}^n) = S^{n-1}$, one shows that the first non-vanishing homotopy group of $V_k(\mathbb{R}^n)$ is π_{n-k}. Moreover,

$$\pi_{n-k} V_k(\mathbb{R}^n) \cong \begin{cases} \mathbb{Z} & \text{if } n-k \text{ is even or } k=1; \\ \mathbb{Z}/2 & \text{else.} \end{cases}$$

Thus the primary obstruction to a section of $V_k(E) \to X$ is a class

$$\mathfrak{o}^{n-k+1} \in H^{n-k+1}(X; \pi_{n-k} V_k(\mathbb{R}^n)).$$

These amount to characteristic classes for rank n bundles,

$$\mathfrak{o}_n^k(E) \in \begin{cases} H^{n-k+1}(X;\mathbb{Z}), & n-k \text{ even or } k=1; \\ H^{n-k+1}(X;\mathbb{Z}/2), & n-k \text{ odd and } k>1, \end{cases}$$

except that in the $\mathbb{Z}$-cases one should think through what data are needed to pin down the sign of the isomorphism $\pi_{n-k}(\mathbb{R}^n) \to \mathbb{Z}$. For the case $k=1$, the necessary datum is an orientation for E; the resulting class $\mathfrak{o}_n^1 \in H^n(X;\mathbb{Z})$ is the *Euler class* $e(E)$ (one can take this as the definition of $e(E)$).

In all cases, there is a class $\bar{\mathfrak{o}}_n^k \in H^{n-k+1}(X;\mathbb{Z}/2)$, by reducing mod 2.

Theorem 6.1 $\bar{\mathfrak{o}}_n^k(E) = w_{n-k+1}(E)$.

The reduced class $\bar{\mathfrak{o}}_n^k(E)$ can be defined regardless of $\pi_1(X)$, and the conclusion still holds.

6.3 Vector bundles over a 4-manifold

The following is a case of Theorem 6.1:

When the vector bundle $E \to X$ has rank 4, $w_2(E)$ is the obstruction $\mathfrak{o}_4^3(E)$ to finding 3 linearly independent sections of E over X^2.

Corollary 6.2 *Over a closed, oriented 4-manifold X, $w_2(TX)$ is the obstruction to trivializing TX over the complement of a point $x \in X$.*

Proof Fix a Riemannian metric g in TX. Choose a CW decomposition for $Y = X \setminus \{x\}$.

We can find 3 orthonormal vector fields v_1, v_2 and v_3 over the 2-skeleton Y^2. The line bundle $\ell = (v_1, v_2, v_3)^\perp \to Y^2$ is trivial, since $w_1(TY) = 0$, so it is spanned by a fourth unit vector field v_4. We can extend this trivialization over Y^3, because the obstruction lies in $H^3(Y; \pi_2 \mathsf{SO}(4)) = 0$. We use here the fact that $\pi_2 \mathsf{SO}(4) = 0$, which is true because the universal cover of $\mathsf{SO}(4)$ is $S^3 \times S^3$, or as an instance of the theorem that $\pi_2(G) = 0$ for any Lie group G (see J. Milnor, *Morse theory*).

If there are any higher-dimensional cells, we can extend over those skeleta since $H^i(Y; \pi_{i-1} \mathsf{SO}(4)) = 0$ for $i > 3$. □

Theorem 6.3 *Suppose X is a closed, oriented 4-manifold, and that T and T' are oriented rank 4 real vector bundles with $w_2(T) = w_2(T') = 0$. Then $T \oplus \underline{\mathbb{R}} \cong T' \oplus \underline{\mathbb{R}}$ if and only if $p_1(T) = p_1(T')$. Moreover, $T \cong T'$ as oriented bundles if and only if, in addition, $e(T) = e(T') \in H^4(X;\mathbb{Z})$.*

Lemma 6.4 *There is a rank 4 vector bundle $E \to S^4$ with $p_1(E)[S^4] = -2$ and $e(E)[S^4] = 1$.*

Proof S^4 is diffeomorphic to quaternion projective space $\mathbb{H}P^1$, which carries the tautological quaternionic line bundle $\Lambda \to \mathbb{H}P^1$ (whose fiber over an $\mathbb{H}$-line λ is λ. Let E be the dual $\mathsf{Hom}_{\mathbb{H}}(\Lambda, \mathbb{H})$, viewed as a rank 2 complex vector bundle. When we write homogeneous coordinates $[X : Y]$, X and Y are sections of E. Thus the locus $X = 0$ (a point) is Poincaré dual to $e(E) = c_2(E)$. Hence $e(E)[S^4] = 1$, while $p_1(E)[S^4] = (c_1^2 - 2c_2)(E)[S^4] = -2$. □

Proof of the theorem A 'Pontryagin–Thom collapse map' is a map $f\colon X \to S^4$, of degree $+1$, with the property that, for some $x \in X$, f is smooth near x with $D_x f$ an isomorphism and $f^{-1}(f(x)) = \{x\}$. These are easy to construct: fix a closed neighborhood B of a chosen x and a diffeomorphism $\phi\colon B \to D^4$. Compose with the projection $D^4 \to D^4/\partial D^4$ followed by a suitable homeomorphism $D^4 \partial D^4 \to S^4$. Fix such an f.

Since $w_2(T) = 0$, T can be trivialized away from X. Consequently, there is a vector bundle $U \to S^4$ and an isomorphism $f^*U \cong T$. Thus $f^*p_1(U) = p_1(T)$ and $f^*e(U) = e(T)$. Thus it will suffice to show that $p_1(U)$ determines $U \oplus \underline{\mathbb{R}}$, and that $(p_1(U), e(U))$ determines U.

The stabilized bundle $U \oplus \underline{\mathbb{R}} \to S^4$ is trivial over the two hemispheres $D_\pm$ of S^4, and over the equator S^3 these two trivializations differ by a map $g\colon S^3 \to \mathsf{SO}(5)$, the 'clutching map', whose class $[\gamma] \in \pi_3\,\mathsf{SO}(5)$ is well-defined up to homotopy. Conversely, any $\gamma\colon S^3 \to \mathsf{SO}(5)$ determines a rank 5 bundle $V_\gamma \to S^4$.

The Pontryagin class determines a homomorphism $p_1\colon \pi_3(\mathsf{SO}(5)) \to \mathbb{Z}$, $[\gamma] \mapsto p_1(V_\gamma)[S^4]$. Taking E as in Lemma 6.4, one has $p_1(E \oplus \underline{\mathbb{R}})[S^4] = -2$, so p_1 is a non-trivial homomorphism. Now, $\mathsf{SO}(5)$ is a connected, simple Lie group, and as such, $\pi_3\,\mathsf{SO}(5) \cong \mathbb{Z}$. Hence the homomorphism p_1 is injective. This proves that $p_1(U)$ determines $U \oplus \underline{\mathbb{R}}$.

The clutching function for U itself lies in $\pi_3\,\mathsf{SO}(4)$. Now, $\mathsf{SO}(4)$ has universal cover $\mathsf{SU}(2) \times \mathsf{SU}(2) \cong S^3 \times S^3$ (as we shall discuss in depth later), so $\pi_3\,\mathsf{SO}(4) \cong \mathbb{Z}^2$. We have a homomorphism $(p_1, e)\colon \pi_3\,\mathsf{SO}(4) \to \mathbb{Z}^2$. For the bundle E constructed above one has $(p_1, e) = (-2, 1)$. And $p_1(TS^4)[S^4] = 3\tau(S^4) = 0$, while $e(TS^4)[S^4] = \chi(S^4) = 2$. Since $(-2, 1)$ and $(0, 2)$ are linearly independent in $\mathbb{R}^2$, (p_1, e) is an injective homomorphism. □

Note: We shall not prove it, but the theorem is true under the assumption $w_2(T) = w_2(T')$, even if these classes are non-zero. This sharper result has the

Corollary 6.5 *Suppose that X and X' are simply connected, closed oriented 4-manifolds, and $f\colon X' \to X$ a degree $+1$ homotopy equivalence. Then $f^*TX \cong TX'$.*

Proof f defines an isometry of $Q_{X'}$ with Q_X, and these intersection forms determine w_2, p_1 and e. □

7 Rokhlin's theorem and homotopy theory

7.1 Rokhlin's theorem

Theorem 7.1 (Rokhlin) *If X is a closed oriented 4-manifold with $w_2(TX) = 0$ then 16 divides $\tau(X)$.*

We have seen that the signature of an even unimodular lattice is divisible by 8. For instance, E_8 has signature 8. So the content is that the integer $\tau(X)/8$ is even. The statement is sharp, since we have seen that a quartic surface $X_4 \subset \mathbb{C}P^3$ has signature -16.

By the Hirzebruch signature theorem, Rokhlin's theorem is equivalent to the assertion that 48 divides $p_1(X) := p_1(TX)[X]$.

Our present purpose is to sketch a proof of the logical equivalence of Rokhlin's theorem with a statement in homotopy theory:

Theorem 7.2 *The following assertions are equivalent:*

(i) *If X is a closed oriented 4-manifold with $w_2(TX) = 0$ then 48 divides $p_1(X)$.*

(ii) $\pi_8(S^5) \cong \mathbb{Z}/24$.

Note that $\pi_8(S^5)$ is a *stable* homotopy group: for $k \geq 5$, the groups $\pi_{3+k}(S^k)$ are isomorphic via the suspension homomorphisms.

The theorem is also due to Rokhlin. There is a précis in M. Kervaire and J. Milnor *Bernoulli numbers, homotopy groups and a theorem of Rohlin*, 1960.

That $\pi_8(S^5) \cong \mathbb{Z}/24$ can be proven by methods internal to homotopy theory; Adams's work on the image of the J-homomorphism is particularly relevant. For instance, in A. Hatcher's book draft *Vector bundles and K-theory*, one finds a K-theoretic proof that $|\pi_8(S^5)| \geq 24$.

On the other hand, one can also prove Rokhlin's theorem using differential geometry, and deduce the fact about $\pi_{3+k}(S^k)$. Indeed, we will see later that $-\tau(X)/16$ has an interpretation as the quaternionic index of the Dirac operator.

7.2 The Pontryagin–Thom construction

References: R. Thom, *Quelques propriétés des variétés différentiables*; J. Milnor, *Topology from the differentiable viewpoint*.

Consider closed k-dimensional manifolds M^k embedded in $\mathbb{R}^{k+m}$, where $m > k+1$.

Definition 7.3

- A *normal framing for M* is a trivialization $\phi \colon N_M \to \underline{\mathbb{R}^m}$ of the normal bundle (usually it is the homotopy class of ϕ that matters).
- A *framed cobordism* from a normally framed manifold (M_0, ϕ_0) to a normally framed manifold (M_1, ϕ_1) is a compact manifold P^{k+1} with boundary $\partial P = M_0 \amalg M_1$, with an embedding $j \colon P \to \mathbb{R}^{k+m} \times [0,1]$, transverse to the boundary $\{0,1\} \times \mathbb{R}^{k+m}$, and such that $M_i = j^{-1}(\mathbb{R}^{k+m} \times \{i\})$ for $i \in \{0,1\}$; together with a framing Φ of N_P which, on M_i, agrees with ϕ_i.

Existence of a framed cobordism defines an equivalence relation (it is transitive since one can concatenate cobordisms, and symmetric since one can flip the direction of $[0,1]$). The equivalence classes form an abelian group Ω_k^{framed} under the operation of disjoint union. The identity element is represented by the empty framed manifold, or by any boundary of a framed manifold.

There is a homomorphism

$$\mathsf{PT}\colon \pi_{k+m}(S^m) \to \Omega_k^{framed}.$$

Given a smooth representative f of $[f] \in \pi_{k+m}(S^m)$, choose any regular value $x \in \mathbb{R}^m \subset S^m = \mathbb{R}^m \cup \{\infty\}$. Then $M = f^{-1}(x) \subset \mathbb{R}^{k+m}$ is a closed k-manifold, and the standard basis for $\mathbb{R}^m$ pulls back via f to define a normal framing ϕ for M. The class $\mathsf{PT}[f] = [M, \phi]$ is well defined in Ω^k_{framed}, since given two regular values x_0 and x_1, one can connect them via a regular path in S^m.

Theorem 7.4 PT *is an isomorphism.*

(Properly speaking, the Pontryagin–Thom construction is really the construction of the inverse map.)

7.3 The J-homomorphism

The J*-homomorphism*

$$J_k^m\colon \pi_k\,\mathrm{SO}(m) \to \pi_{k+m}(S^m),$$

is defined as follows. Regard S^{k+m} as $\partial(D^{k+1}\times D^m) = (S^k \times D^m) \cup (D^{k+1} \times S^{m-1})$ (the two parts meet along $S^k \times S^{m-1}$). Given $\theta\colon (S^k, *) \to (\mathrm{SO}(m), I)$, we want to define

$$J_m^k(\theta)\colon (S^k \times D^m) \cup (D^{k+1} \times S^{m-1}) \to D^m/\partial D^m.$$

To do so, extract from $\theta\colon S^k \to \mathrm{SO}(m)$ the family of linear isometries $\theta_x\colon (D^m, \partial D^m) \to (D^m, \partial D^m)$ parameterized by $x \in S^k$. Collectively these define a map $S^k \times D^m \to D^m/\partial D^m$. It sends $S^k \times S^{m-1}$ to the basepoint $*$. Together with the constant map $(D^{k+1} \times S^{m-1}) \to *$, this prescription defines $J_k^m(\theta)$.

We consider the stable range $k > m+1$, in which the inclusion $\mathrm{SO}(m) \to \mathrm{SO}(m+1)$ induces an isomorphism on π_k, and in which suspension $\pi_{k+m}(S^m) \to \pi_{k+m+1}(S^{m+1})$ is an isomorphism. These isomorphisms intertwine J_k^m with J_k^{m+1}, so we really just have one 'stable J-homomorphism J_k for each k.

The stable J-homomorphism has a geometric meaning, as follows.

The framings for S^k inside $\mathbb{R}^{k+m}$ form a group under an operation of connected sum (facilitated by insisting that the framing is standard over a chosen disc inside S^k). The identity element comes from is the framing of $S^k \subset \mathbb{R}^{k+1}$. On the other hand, the framings also form a torsor for $\pi_k\,\mathrm{SO}(m)$; and indeed, these decscriptions are compatible, so that one can simply say

$$\{\text{framings of } S^k \subset \mathbb{R}^{k+m}\} = \pi_k\,\mathrm{SO}(m).$$

Via the Pontryagin–Thom construction, J_k^m then amounts to a map

$$J_k^m\colon \{\text{framings of } S^k \subset \mathbb{R}^{k+m}\} \to \Omega_k^{framed}.$$

Proposition 7.5 J_k^m *maps a normal framing* ϕ *of* S^k *to the framed cobordism class* $[S^k, \phi]$.

Exercise 7.6 Prove the proposition.

7.4 Framed 3-manifolds and their bounding 4-manifolds

We now take $k = 3$ and $m = 5$. The homomorphism

$$J_3 = J_3^5 \colon \pi_3\,\mathrm{SO}(5) \to \Omega_3^{framed}$$

sends normal framings of $S^3 \subset \mathbb{R}^8$ to framed cobordism classes of 3-manifolds in $\mathbb{R}^8$.

Theorem 7.7 *J_3 is surjective.*

sketch In this proof we freely use the notion of spin cobordism (i.e. cobordism of manifolds with a 'trivialization of w_2'). We invoke the fact that the spin cobordism group Ω_3^{spin} is trivial. Thus, given a normally framed 3-manifold (M,ϕ), one can find a compact oriented 4-manifold $P \subset \mathbb{R}^8 \times [0,1]$ with oriented boundary $\partial P = -S^3 \amalg M$, such that $w_2(TP) = 0$. The obstructions to extending ϕ to a normal framing Φ for P lie in $H^{i+1}(P,M;\pi_i\,\mathrm{SO}(5))$. When $i = 1$, the obstruction is exactly $w_2(TP)$. When$i > 1$, the obstruction group vanishes. Thus we can find a framing Φ, and it defines a representative (S^3,ψ) for $[M,\phi]$. □

Noting that $\pi_3\,\mathrm{SO}(5) \cong \mathbb{Z}$, J_3 is a surjection

$$J_3 \colon \mathbb{Z} \cong \pi_3\,\mathrm{SO}(5) \to \pi_8(S^5) \to 0.$$

Now we come to the relation with Rokhlin's theorem.

Proof of Theorem 7.2 (ii) $\Rightarrow$ (i): Here we assume that $\pi_8(S^5) \cong \mathbb{Z}/24$.
Then $\ker J_3 = 24\mathbb{Z}$. Suppose $X \subset \mathbb{R}^9$ is a closed oriented 4-manifold with $w_2(TX) = 0$. We saw in Lecture 5 that we can trivialize $T(X-B)$, where B is a ball. Hence we can find a trivialization Φ of the normal bundle ν to $X-B$ inside $\mathbb{R}^9$. There is an obstruction $\mathfrak{o}^4(\nu,\Phi)$ to extending Φ to X, lying in $H^4(X;\pi_3\,\mathrm{SO}(5)) \cong H^4(X;\mathbb{Z}) \cong \mathbb{Z}$. The integer $\mathfrak{o}^4(\nu,\Phi)$ is exactly the homotopy class of $\Phi|_{\partial B}$ when viewed as an element of $\pi_3\,\mathrm{SO}(5) = \mathbb{Z}$. Thus $\mathfrak{o}(\nu,\Phi)$ is an element of $\ker J_3 = 24\mathbb{Z}$.
In a lemma from Lecture 5, we constructed a complex vector bundle $E \to S^4$ with $c_2(E) = 1$. This vector bundle arose as the tautological line bundle over $\mathbb{H}P^1$. Thus E, being a quaternionic line bundle, has a clutching function $\gamma_E \in \pi_3(\mathrm{SU}(2))$. Since $c_2(E) = 1$, the clutching function of E is a generator for the infinite cyclic group $\pi_3\mathrm{SU}(2)$.

Lemma 7.8 *The clutching function for $E \oplus \mathbb{R}$ generates $\pi_3\,\mathrm{SO}(5)$.*

Proof The clutching function for $E \oplus \mathbb{R}^2$ is deduced from γ_E by means of the composite of inclusions $\mathrm{SU}(2) \to \mathrm{SO}(4) \to \mathrm{SO}(6)$. But there is a commutative diagram of inclusions of Lie groups

$$\begin{tikzcd} \mathrm{SU}(2) \arrow[r] \arrow[d] & \mathrm{SO}(4) \arrow[d] \\ \mathrm{SU}(3) \arrow[r] & \mathrm{SO}(6) \end{tikzcd},$$

so it also deduced by means of the other composite $\mathrm{SU}(2) \to \mathrm{SU}(3) \to \mathrm{SO}(6)$. Since γ_E generates $\pi_3\mathrm{SU}(2)$, it suffices to show that $\mathrm{SU}(2) \to \mathrm{SU}(3)$ and $\mathrm{SU}(3) \to \mathrm{SO}(6)$ induce isomorphisms on π_3. The first of these assertions is easy to deduce from the homotopy exact sequence of the fibration $\mathrm{SU}(2) \to \mathrm{SU}(3) \to S^5$. In general, in a simply-laced, simple Lie group G, the map $\mathrm{SU}(2) \to G$, integrating a map of Lie algebra $\mathfrak{su}(2) \to \mathfrak{g}$ corresponding to a simple root of G, induces an isomorphism on π_3.[3] The inclusion of $\mathrm{SU}(n)$ in $\mathrm{SO}(2n)$ (for $n \geq 3$) respects simple roots—it corresponds to the evident inclusion of Dynkin diagrams—and so induces an isomorphism on π_3. □

[3] At present, I don't have a convenient reference for this point.

Since the clutching function for $E \oplus \underline{\mathbb{R}}$ generates $\pi_3 \mathsf{SO}(5))$, and $p_1(E \oplus \underline{\mathbb{R}}) = -2c_2(E) = -2$, we deduce that

$$p_1(X) = \pm 2\mathfrak{o}^4(\nu, t).$$

But then $48 | p_1(X)$, as required.

(i) $\Rightarrow$ (ii): Now suppose we know Rokhlin's theorem. Take a normally framed 3-sphere (S^3, ϕ) representing a class $a \in \ker J_3 = \mathbb{Z}$. There is ta normally framed 4-manifold (P, Φ) bounding (S^3, ϕ). Let $X = P \cup_{S^3} D^4$. Then $w_2(X) = 0$—adding a 4-ball does not affect degree 2 cohomology—so by Rokhlin, $p_1(TX)[X] = 48r$ for some $r \in \mathbb{Z}$. On the other hand $p_1(TX)[T] = \pm 2a$, so $a = 24r$, and $\ker J_3 \subset 24\mathbb{Z}$. The example of the quartic surface, with $p_1 = -48$, shows that $24 \in \ker J_3$. So J_3 induces an isomorphism

$$J_3 \colon \mathbb{Z}/24\mathbb{Z} \to \pi_3(S^8).$$

□

7.5 Background: Relative homotopy groups

This material, which is preparatory for the next lecture, is covered in Hatcher's *Algebraic Topology*.

If A is a subspace of X, $x \in A$ is a basepoint, and $n \geq 1$, $\pi_n(X, A, x)$ (or just $\pi_n(X, A)$) denotes the set of homotopy classes of maps $(D^n, \partial D^n, *) \to (X, A, x)$. It includes a distinguished element e, the unique homotopy class of a map $D^n \to A$. For $n \geq 2$, there is a 'collapsing map' $c \colon D^n \to D^n \vee D^n$, and this makes $\pi_n(X, A, x)$ a group, with identity e, under the operation $[f] * [g] = (f \vee g) \circ c$. For $n \geq 3$ it is abelian.

The relative homotopy groups are functorial, and have the following properties:

- The inclusions $A \to X$ and $(X, \emptyset) \to (X, A)$ define maps in a long exact sequence
$$\cdots \to \pi_n(A, x) \to \pi_n(X, x) \to \pi_n(X, A, x) \xrightarrow{\partial} \pi_{n-1}(A) \to \dots$$
in which ∂ is defined by restriction from D^n to S^{n-1}. The sequence becomes non-abelian $(\cdots \to \pi_2(X) \to \pi_2(X, A) \to \pi_1(A) \to \pi_1(X) \to \dots)$ and then ends with an 'exact sequence of pointed sets' $\pi_1(X, x) \to \pi_1(X, A, x) \to \pi_0(A, x) \to \pi_0(X, x)$.
- There are Hurewicz maps $h \colon \pi_n(X, A, x) \to H_n(X, A)$, mapping $[f]$ to $f_*[D^n, \partial D^n]$ (the image under f_* of the fundamental class of $(D^n, \partial D^n)$). Hurewicz maps intertwine the homotopy and homology exact sequences of the pair (X, A).
- There is a notion of (X, A) being an *n-connected pair*. It means that for all $i \leq n$, every map $(D^i, \partial D^i) \to (X, A)$ is homotopic (as a pair) to a map $(D^i, \partial D^i) \to (A, A)$. Equivalently, for all $x_0 \in$, one has $\pi_i(X, A, x_0) = 0$ for $1 \leq i \leq n$ and $\pi_0(A, x) \to \pi_0(X, x_0)$ is onto.
- There is a relative Hurewicz theorem: *If (X, A) is $(n-1)$-connected*, $n \geq 2$, and A non-empty and simply connected, then $H_i(X, A) = 0$ for $i < n$ and $h \colon \pi_n(X, A) \xrightarrow{\simeq} H_n(X, A)$.

7.6 Appendix: Effect on π_3 of Lie group homomorphisms

Proposition 7.9 *Let G be a connected, simple Lie group. Then $\pi_3(G)' \cong \mathbb{Z}$.*

Here $\pi_3(G)'$ is the torsion-free quotient. In fact, $\pi_3(G) \cong \mathbb{Z}$; as uniform (rather than case-checking) proof of the latter fact is due to Bott, and uses Morse theory on loop-spaces (cf. [**?**]). We will prove the weaker version using Chern–Weil theory, as follows:

Proof We may assume G compact, since every Lie group deformation-retracts to a maximal compact subgroup. We may assume G simply connected, since the universal cover is again a compact, simple Lie group with the same homotopy groups as G.

Recall that $\pi_2(G) = 0$ (this can be proved using Bott's method, see [**?**], or via Whitehead's lemma that $H^2(\mathfrak{g};\mathbb{R}) = 0$). So G is 2-connected; its classifying space BG is therefore 3-connected. Thus, using Hurewicz,

$$\pi_3(G) \cong \pi_4(G) \cong H_4(BG),$$

and by universal coefficients,

$$H^4(BG) \cong H_4(BG)^\vee \cong \pi_3(BG)^\vee.$$

It suffices, then, to show that $H^4(BG)' \cong \mathbb{Z}$, or equivalently that $H^4(BG;\mathbb{C}) \cong \mathbb{C}$.

The Chern–Weil homomorphism

$$\mathsf{cw}_G \colon \mathbb{C}[\mathfrak{g}]^G \to H^*(BG;\mathbb{C})$$

maps a G-invariant polynomial p on $\mathfrak{g}$ to the cohomology class of $p(F_\nabla)$, where F_∇ is the curvature of a connection ∇ in the universal principal G-bundle $EG \to BG$. We assign a degree d polynomial degree $2d$ in $\mathbb{C}[\mathfrak{g}]^G$, whereupon cw_G preserves degree. We quote the fact that it is an isomorphism of graded algebras.

Note that $(\mathbb{C}(\mathfrak{g})^G)_4$, the degree 4 part, consists of G-invariant quadratic forms on $\mathfrak{g}$. Since G is simple, all such forms are multiples of the Killing form. Therefore $H^4(BG) \cong \mathbb{C}$. $\square$

Suppose now that $\phi\colon \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra homomorphism between simple Lie algebras. We define the *Dynkin index* $\operatorname{ind}\phi$ by

$$(\phi(x), \phi(x))_{\mathfrak{g}} = (\operatorname{ind}\phi)\cdot (x,x)_{\mathfrak{h}}, \quad x \in \mathfrak{h}.$$

(the ratio is well-defined since the space of invariant symmetric bilinear forms on a simple Lie algebra is 1-dimensional).

Theorem 7.10 *Let H and G be connected, simple Lie groups, and $\Phi\colon H \to G$ a homomorphism. Let $\phi\colon \mathfrak{h} \to \mathfrak{g}$ be the corresponding Lie algebra homomorphism. The map*

$$\phi_*\colon \pi_3(H)' \to \pi_3(G)'$$

amounts to a homomorphism $\mathbb{Z} \to \mathbb{Z}$, and so is given by multiplication by d for some $d \in \mathbb{Z}$ (determined up to sign). One has

$$d = \operatorname{ind}(\phi).$$

I believe this theorem is due to E. B. Dynkin. A reference for a closely related discussion is Atiyah–Hitchin–Singer, *Self-duality in four-dimensional Riemannian geometry*.

The integer d describing the effect on π_3' can be characterized in two additional ways:

- $(B\phi)^*\colon H^4(BG) \to H^4(BH)$, which is the map $\mathbb{Z} \to \mathbb{Z}$ dual to the one on π_3', is multiplication by $\pm d$.

- $\phi^*\colon (\mathbb{C}[\mathfrak{g}]^G)_4 \to (\mathbb{C}[\mathfrak{h}]^H)_4$ maps the integer lattice in $(\mathbb{C}[\mathfrak{g}]^G)_4$ to d times the integer lattice in $(\mathbb{C}[\mathfrak{h}]^H)_4$.

(Here the integer lattices are those defined by $H^4(BG)$ and $H^4(BH)$.) The latter holds because the Chern–Weil homomorphism is natural: there is a commutative diagram

$$\begin{tikzcd} \mathbb{C}[\mathfrak{g}]^G \arrow[r, "\phi^*"] \arrow[d, "\mathsf{cw}_G"'] & \mathbb{C}[\mathfrak{h}]^H \arrow[d, "\mathsf{cw}_H"] \\ H^*(BG;\mathbb{C}) \arrow[r, "(B\phi)^*"] & H^*(BH;\mathbb{C}) \end{tikzcd} .$$

Now, the bases for $(\mathbb{C}[\mathfrak{g}]^G)_4$ and $(\mathbb{C}[\mathfrak{h}]^H)_4$ given by $(\cdot,\cdot)_{\mathfrak{g}}$ and $(\cdot,\cdot)_{\mathfrak{h}}$ have the property that $\phi^*(\cdot,\cdot)_{\mathfrak{g}} = (\operatorname{ind}\phi)(\cdot,\cdot)_{\mathfrak{h}}$. Thus to assert that $d = \operatorname{ind}\phi$ for all ϕ is equivalent to asserting that there is a universal constant c for which $(\cdot,\cdot)_{\mathfrak{g}}$ represents c times a generator of the integer cohomology $H^4(BG)$.

Lemma 7.11 *When G is simply connected and simple, $H^4(BG)$ is generated by the class*

$$\mathsf{cw}_G\left(\frac{1}{h^\vee}\kappa_{\mathfrak{g}}\right).$$

Here $\kappa_{\mathfrak{g}}$ is the Killing form (i.e., $\kappa_{\mathfrak{g}}(\xi) = \operatorname{tr}(\mathsf{ad}\,\xi \circ \mathsf{ad}\,\xi)$), and $h^\vee$ the dual Coxeter number; that is, $h^\vee = 1/\kappa_{\mathfrak{g}}(\theta)$, where θ is the sum of the simple roots of $\mathfrak{g}$.

8 Realization of unimodular forms by 4-manifolds

8.1 The theorem of Whitehead on 4-dimensional homotopy types

J. H. C. Whitehead's 1949 paper *On simply connected 4-dimensional polyhedra* classifies the homotopy types of simply connected 4-dimensional CW complexes in terms of cohomological data. There are subtleties related to the torsion which can, in general, appear in H^3, and which Whitehead handles via Pontryagin's refinement of the mod n cup-square operation. But in the case of simply connected 4-manifolds, as we have seen, the intersection form governs every aspect of the cohomology, and the result is particularly simple to state:

Theorem 8.1 *Suppose that X and X' are closed, oriented, simply connected 4-manifolds. Then any isometry $H_2(X) \to H_2(X')$ of their intersection forms is induced by a degree 1 homotopy equivalence $X \to X'$.*

The proof we shall present is from J. Milnor and D. Husemoller's book *Symmetric bilinear forms*.

We approach the theorem via a construction of simply connected, 4-dimensional CW complexes. Take a wedge sum $\bigvee_{i=1}^n S^2$ of 2-spheres. It is a 2-complex with one 0-cell and n 2-cells. Now let X_f be the complex obtained by attaching a 4-cell via $f\colon S^3 \to \bigvee_{i=1}^n S^2$:

$$X_f = \left(\bigvee^n S^2\right) \cup_f D^4.$$

The homotopy type of f is determined by the class of f in $\pi_3(\bigvee S^2)$.

Since the cellular cochain complex is purely even, one has $H^2(X_f) = \mathbb{Z}^n$, $H^4(X_f) = \mathbb{Z}$, $H^{odd}(X_f) = 0$. The cup product $H^2 \times H^2 \to H^4$ therefore takes the form of an $n \times n$ symmetric matrix Q_f.

Lemma 8.2 *The map $f \mapsto Q_f$ defines an isomorphism*

$$\pi_3(\bigvee^n S^2) \to \{n \times n \text{ symmetric matrices over } \mathbb{Z}\}.$$

When $n = 1$, we recover the well-known fact that $\pi_3(S^2) \cong \mathbb{Z}$ via the Hopf invariant, and it is instructive to recall one of the proofs of that fact:

Embed $S^2 = \mathbb{P}^1$ into $\mathbb{P}^2$ in the usual way. Then $H_k(\mathbb{P}^2, \mathbb{P}^1) = 0$ for $k \leq 3$, while $H_4(\mathbb{P}^2, \mathbb{P}^1) \cong \mathbb{Z}$. By relative Hurewicz, the first non-trivial relative homotopy group coincides with the first relative homology group. Thus $(\mathbb{P}^2, \mathbb{P}^1)$ is 3-connected, and $\pi_4(\mathbb{P}^2, \mathbb{P}^1) \xrightarrow{\simeq} H_4(\mathbb{P}^2, \mathbb{P}^1) \cong \mathbb{Z}$.

The Hopf fibration $S^1 \to S^5 \to \mathbb{P}^5$ gives a long exact sequence of homotopy groups, showing that $\pi_4(\mathbb{P}^2) = 0$ and $\pi_3(\mathbb{P}^2) = 0$; by the homotopy long exact sequence of the pair $(\mathbb{P}^2, \mathbb{P}^1)$, one has $\pi_4(\mathbb{P}^2, \mathbb{P}^1) \xrightarrow{\simeq} \pi_3(\mathbb{P}^1)$.

Proof of the lemma. Let $P = (\mathbb{P}^2)^{\times n}$. Embed $\bigvee^n \mathbb{P}^1$ into $(\mathbb{P}^1)^{\times n}$ as the subspaces where all but one coordinate is a basepoint $* \in \mathbb{P}^1$. Embed $(\mathbb{P}^1)^{\times n}$ into P as the product of the standard embeddings $\mathbb{P}^1 \to \mathbb{P}^2$. It is easy to check that $H_k(P, \bigvee^n \mathbb{P}^1) = 0$ for $k \leq 3$, and that $H_4(P) \xrightarrow{\simeq} H_4(P, \bigvee^n \mathbb{P}^1)$.

By the Künneth formula in cellular homology, $H^4(P)$ is free abelian with basis $\{e^i \cup e^j\}_{1 \leq i,j \leq n}$, where $e^i \in H^2(P)$ is the pullback of the generator of $H^2(\mathbb{P}^2)$ via the ith projection. And $H^4(P) = \mathsf{Hom}(H_4(P), \mathbb{Z})$.

Now, $\pi_4(P) = \prod^n \pi_4(\mathbb{P}^2) = 0$, and likewise $\pi_3(P) = 0$, so $\pi_4(P, \bigvee \mathbb{P}^1) \xrightarrow{\simeq} \pi_3(P)$. And by relative Hurewicz, $\pi_4(P, \bigvee \mathbb{P}^1) \xrightarrow{\simeq} H_4(P, \bigvee \mathbb{P}^1)$. Hence we have an isomorphism $\pi_3(P) \cong H_4(P)$, $[f] \mapsto h_f$.

We get an isomorphism $\pi_3(P) \to \{n \times n \text{ symmetric matrices}\}$ by mapping f to the matrix $\tilde{Q}_f$ with entries $(\tilde{Q}_f)_{ij} = \langle e^i \cup e^j, h_f \rangle$. What remains is to show that $Q_f = \tilde{Q}_f$.

The composite $S^3 \xrightarrow{f} \bigvee \mathbb{P}^1 \hookrightarrow P$ represents a class in $\pi_3 P = 0$; being nullhomotopic, it extends to a map $D^4 \to P$. Thus the inclusion $\bigvee \mathbb{P}^1 \hookrightarrow P$ extends to a map $\eta \colon X_f \to P$. The fundamental class $[X_f] \in H_4(X_f)$ maps under η to $\eta_*[X_f] = h_f$, while the classes $\eta^* e^i$ represent the evident cellular basis for $H^2(X_f)$. We have

$$(\tilde{Q}_f)_{ij} = \langle e^i \cup e^j, h_f \rangle = \langle \eta^*(e^i \cup e^j), [X_f] \rangle = \langle \eta^* e^i, \eta^* e^j, [X_f] \rangle = (Q_f)_{ij}.$$

□

We can now prove Whitehead's theorem. It applies even to *topological* 4-manifolds, where the conclusion we shall prove is that the isometry of forms is induced by a weak homotopy equivalence. In fact, compact topological manifolds have the homotopy type of finite CW complexes—a much deeper fact than its specialization to smooth manifolds—so these weak equivalences are homotopy equivalences.

Fix a basepoint $x \in X$ and a closed 4-ball B not containing x. Let $X_0 = X \setminus B$. Notice that X_0 is simply connected with $H_2(X_0) = H_2(X)$. By simple connectivity, the Hurewicz map $\pi_2(X_0, x) \to H_2(X_0)$ is an isomorphism. Choose a basis $(e_1, \dots, e_n)$, for $H_2(X_0)$, and represent e_i as $(s_i)_*[S^2]$ for a map $s_i \colon (S^2, *) \to (X, x)$. Assemble the s_i into a map $s \colon \bigvee^n S^2 \to X_0$.

Observe next that $s \colon \bigvee^n S^2 \to X_0$ induces an isomorphism on H_*—after all, both spaces have homology only in degrees 0 and 2. By Whitehead's theorem, s is a weak homotopy equivalence.

Now, $X = X_0 \cup D^4$—the resulting of attaching a 4-cell to X_0. Replacing X_0 by its homotopy-equivalent space $\bigvee^n S^2$, we see that $X \simeq X_f$ for some $f \in \pi_3 \bigvee^n S^2$. We can take the homotopy equivalence to map the basis vector $e_i \in H_2(X)$ to the ith cell $e_i \in H_2(X_f)$. Thus $Q_X = Q_f$. So Q_X determines Q_f and therefore, by the lemma, f. Hence Q_X determines the homotopy type of X_f, and therefore of X.

8.2 Topological 4-manifolds

Any topological manifold M has a *Kirby–Siebenmann invariant* $\kappa(M) \in H^4(M; \mathbb{Z}/2)$. It is an obstruction to finding a piecewise linear (PL) structure on M; in dimensions > 4, it is the only obstruction (see R. Kirby, L. Sibenmann, *Foundational Essays on Topological Manifolds, Smoothings, and Triangulation* or Y. Rudyak's survey *Piecewise linear structures on topological manifolds*). Since smooth manifolds admit PL structures, it also obstructs smoothing M.

In the case of a closed 4-manifold, one has $\kappa(M) \in \mathbb{Z}/2$. It is known that PL structures on 4-manifolds admit unique smoothings, but that vanishing of $\kappa(M)$ is not sufficient for a PL structure to exist.

Theorem 8.3 (M. Freedman) *Let Q be a unimodular form.*

- *Q arises as the intersection form Q_X of a simply connected, closed, oriented topological 4-manifold X.*
- *Any automorphism of Q is realized by an oriented self-homeomorphism of X.*
- *If Q is even, then $\kappa(X) = 0$ and X is unique up to homeomorphism.*

- *If Q is odd, there are exactly two non-homeomorphic realizations of Q, say X_0 and X_1, where the subscript is $i = \kappa(X_i) \in \mathbb{Z}/2$. Thus the manifold X_1 is not smoothable.*

The proof of the theorem is a highly intricate development of the notion of a 'Casson handle', and the only complete reference seems to be Freedman's original paper *The topology of four-dimensional manifolds* (J. Diff. Geom., 1982).

Recall that any odd, indefinite unimodular forms are of shape $mI_+ \oplus nI_-$. The latter form is realized by a connected sum $m(\mathbb{C}P^2)\#n(\overline{\mathbb{C}P^2})$ of m copies of $\mathbb{C}P^2$ with its complex orientation, and n copies of $\mathbb{C}P^2$ with its other orientation. Any even, indefinite unimodular form is (up to an overall change of sign) of shape $rU \oplus s(-E_8)$. According to Freedman's theorem, there is a topological 4-manifold M with intersection form E_8, and we can then realize $rU \oplus s(-E_8)$ by the connected sum

$$r(S^2 \times S^2)\# s\overline{M}.$$

There is also a multitude of definite unimodular forms, all represented by topological 4-manifolds.

8.3 Realizing unimodular forms by smooth 4-manifolds

In what follows, X is always a closed, oriented, smooth 4-manifold.

The following questions are key:

(1) Which *indefinite* unimodular forms are represented as intersection forms Q_X?

(2) Which *definite* unimodular forms are represented as intersection forms Q_X?

For question (1), the answer is immediate in the odd case, so the question is really about even forms.

Rokhlin's theorem in combination with Freedman's shows that there are topological 4-manifolds which admit no smooth structure: the E_8-manifold M, for instance. An even indefinite intersection form must (up to a sign) take the shape

$$rU + s(-E_8), \quad r > 0,\ s \geq 0;$$

Rokhlin says that s is *even* when the form comes from a smooth manifold.

Question 2 was settled using gauge theory.

Theorem 8.4 (S. Donaldson, 1981) *If Q_X is definite and $H_1(X)$ has no 2-torsion then Q_X is diagonalizable over $\mathbb{Z}$.*

Being diagonalizable, Q_X is then represented by $\pm I$, where I is the identity matrix.

Donaldson proved this theorem using instanton gauge theory, but Kronheimer reproved it using the Seiberg–Witten equations, and in doing so sharpened it:

Theorem 8.5 (P. Kronheimer, 1994) *If Q_X is definite then Q_X is diagonalizable over $\mathbb{Z}$.*

The spin case. We now return to realization of even forms. For this is natural to assume $w_2(TX) = 0$ (in the simply connected case, evenness is equivalent to vanishing w_2).
The diagonalization theorem has the following

Corollary 8.6 *If $w_2(TX) = 0$ and Q_X is definite then $b_2(X) = 0$.*

We turn now to the indefinite case. Set

$$Q_{r,s} = rU + sE_8, \quad r \geq 0,\, s \in \mathbb{Z}.$$

We can allow s negative. By Rokhlin, s must be even to be realizable.

Theorem 8.7 (Donaldson) *Assume $H_1(X)$ has no 2-torsion and $w_2(TX) = 0$. If $Q_X \cong Q_{r,2t}$ with $t \neq 0$, then $r \geq 3$.*

Again, Kronheimer was able to remove the assumption on $H_1(X)$.
It follows that if a K3 surface X is smoothly a connected sum $X_1 \# X_2$, then one of the summands has $b_2 = 0$. By contrast, X is topologically the sum of 3 copies of $S^2 \times S^2$ and 2 copies of the $-E_8$-manifold $\bar{M}$.
Exactly which forms $Q_{r,2t}$ are realizable? Notice that if $t < 0$, we have $r = b^+(X)$, the dimension of a maximal positive-definite subspace in $H^2(X;\mathbb{R})$.
By taking a connected sum of t K3 surfaces and u copies of $S^2 \times S^2$, we can realize $Q_{3|t|+u,2t}$ for any $u \geq 0$. Notice that $Q_{3|t|+u,2t}$ has rank $b_2 = 2(3|t| + u) + 16|t| = 2(11t + u)$, and signature $\tau = -16t$. So

$$b_2 = \left(\frac{11}{8} + \frac{u}{8}\right)|\tau| \geq \frac{11}{8}|\tau|.$$

Conjecture 8.8 (The 11/8 conjecture) *If $w_2(TX) = 0$ then*

$$b_2(X) \geq \frac{11}{8}|\tau(X)|.$$

Equivalently, if $Q_X \cong Q_{r,2t}$ then $r \geq 3|t|$.

It is known that to prove 11/8 it suffices to do so in the simply connected case.
The conjecture is open, but there has been substantial progress towards it. The theorem of Donaldson (plus Kronheimer) quoted above proves it when $|t| = 1$. It is also known when $|t| = 2$ [M. Furuta, Y. Kametani, H. Matsue, 2001], but not (to my knowledge) when $|t| = 3$. However, one has Furuta's 5/4 theorem:

Theorem 8.9 (M. Furuta, 1997) *If $w_2(TX) = 0$ then*

$$b_2(X) \geq \frac{5}{4}|\tau(X)| + 2.$$

Equivalently, if $Q_X \cong Q_{r,2t}$ then $r \geq 2|t| + 1$.

Several proofs are known, but all have at their core the same strategy, which is to utilize refinements of Seiberg–Witten theory. SW theory produces a closed, oriented manifold $\mathcal{M}$ of solutions to the SW equations inside an ambient space $\mathcal{B}^*$. 'Standard' SW theory involves studying this manifold, and produces invariants from its fundamental class $[\mathcal{M}] \in H_*(\mathcal{B})$. Refined SW theory produces says, roughly speaking, that as a regular level set of a smooth map, $\mathcal{M}$ should comes with a normal framing (as in the Pontryagin–Thom construction), and uses this idea to set up invariants of a homotopical nature. Furuta's theorem specifically depends on 'Seiberg–Witten KO-theory'.[4]

[4] No proof is known using Donaldson theory, nor Heegaard Floer theory.

Conclusion: Gauge theory offers a solution to the problem of realizing unimodular forms by smooth manifolds, *except* for a question mark regarding the forms $Q_{r,2t}$ with $2|t|+2 \leq r < 3|t|$ and $|t| \geq 3$.

For the complementary uniqueness question—of enumerating diffeomorphism-types of simply connected 4-manifolds Y in the homotopy type of a fixed manifold X—gauge theory can frequently distinguish infinite collections of diffeomorphism-types, but for no X has it supplied a complete answer.

Part II. Differential geometry

8 Self-duality: linear-algebraic aspects

This lecture is based on [6, chapter 1]. It is about oriented inner product spaces, chiefly of dimension 4. These will arise as cotangent spaces of manifolds. Today's material will, almost in its entirety, generalize to vector bundles, and in particular to cotangent bundles of manifolds.

8.1 Self-duality

8.1.1 The Hodge star

In general, an inner product g on a finite-dimensional vector space V, with an orthonormal basis $(v_1, \dots, v_n)$, induces $\mathsf{O}(V)$-invariant inner products on the exterior products $\Lambda^k V$ (with orthonormal basis $\{v_{i_1} \wedge \cdots \wedge v_{i_k}\}_{i_1 < \cdots < i_k}$).

Fix an orientation for V, and assume that $(v_1, \dots, v_n)$ is an oriented basis. Let $\mathsf{vol} = v_1 \wedge \cdots \wedge v_n \in \det V$. The bilinear form $\wedge \colon \Lambda^k V \times \Lambda^{n-k} V \to \Lambda^n V = \det V$ is non-degenerate. Hence it is valid to define the *Hodge star*

$$\star \colon \Lambda^k V \to \Lambda^{n-k} V$$

as the unique linear map such that

$$\alpha \wedge \star\beta = \langle \alpha, \beta \rangle \, \mathsf{vol}, \quad \alpha, \beta \in \Lambda^k V.$$

It has the following properties:

- The symmetry of the inner product implies graded self-adjointness,
$$\alpha \wedge \star\beta = (-1)^{k(n-k)} \star \alpha \wedge \beta.$$
- For $I \in \binom{n}{k}$ a k-element subset of $\{1, \dots, n\}$, think of I as an ordered k-tuple $(i_1, \dots, i_k)$ by prescribing $i_1 < \cdots < i_k$. Set $v_I = v_{i_1} \wedge \cdots \wedge v_{i_k}$. Then
$$\star\, v_I = \sigma(I, I^c) v_{I^c}, \tag{2}$$
where $I^c = \{1, \dots, n\} \setminus I$, and $\sigma(I, I^c)$ is the sign of the permutation (I, I^c) of $\{1, \dots, n\}$.
- From (2) we see that $\star$ is an isometry.
- It is clear from (2) and the $\mathrm{SO}(V)$-equivariance of $\star$ that $\star\star = \pm I$, where the sign depends only on n and k. But $\star \star (v_1 \wedge \cdots \wedge v_k) = \star(v_{k+1} \wedge \cdots \wedge v_n) = (-1)^{k(n-k)} v_1 \wedge \cdots \wedge v_k$, so
$$\star\star = (-1)^{k(n-k)} I.$$
- When one conformally rescales the inner product, replacing $\langle \cdot, \cdot \rangle$ by $\lambda \langle \cdot, \cdot \rangle$ for $\lambda > 0$, then $(\lambda^{-1/2} v_1, \dots, \lambda^{-1/2} v_n)$ is an orthonormal basis for the new product. From this we see that
$$\star_{\mathsf{new}} = \lambda^{k-(n/2)} \star_{\mathsf{old}} .$$

8.1.2 2-forms in 4 dimensions

When $n = 4$ and $k = 2$, one has an *isometry of order 2*

$$\star \in \mathsf{O}(\Lambda^2 V), \quad \star\star = \mathrm{id},$$

which *depends only on the conformal class of the inner product.* The ± 1 eigenspaces of $\star$ define an orthogonal splitting into 3-dimensional vector spaces of self-dual and anti-self-dual forms:

$$\Lambda^2 V = \Lambda^+ V \oplus \Lambda^- V \quad \omega = \omega^+ + \omega^- = \frac{1}{2}(I + *)\omega + \frac{1}{2}(I - *)\omega.$$

The subspaces $\Lambda^\pm = \Lambda^\pm V$ depend only on the conformal class of the inner product. $\Lambda^\pm$ has basis

$$\begin{aligned}(\omega_1^\pm, \omega_2^\pm, \omega_2^\pm) &:= ((v_{12})^\pm, (v_{13})^\pm, (v_{14})^\pm) \\ &= (\tfrac{1}{2}(v_{12} \pm v_{34}),\ \tfrac{1}{2}(v_{13} \pm v_{42}),\ \tfrac{1}{2}(v_{14} \pm v_{23}).\end{aligned}$$

If we permute the v_i, the $\omega_j^\pm$ are permuted via the familiar epimomorphism $\theta\colon S_4 \to S_3$—the action of S_4 on the three splittings of $\{1, 2, 3, 4\}$ into 2 sets of 2—whose kernel is the Klein 4-group. Note that θ sends transpositions to transpositions, and so preserves the sign of permutations. Thus Λ^+ and Λ^- inherit orientations from V. Using the metric inherited from $\Lambda^2 V$, one sees that the action of $\mathsf{O}(V)$ on V induces actions on Λ^2 permuting or preserving $\Lambda^\pm$ according to the sign of the determinant, and therefore defining

$$\lambda^\pm\colon\ \mathsf{SO}(V) \to \mathsf{SO}(\Lambda^\pm)$$

Lemma 8.1 *Unit-length, decomposable 2-forms $\eta = u \wedge v \in \Lambda^2 V$ correspond bijectively with oriented 2-planes $P \subset V$, by mapping P to its oriented volume form vol_P.*

Proof The inverse map sends η to the plane $P \subset V$ formed by vectors v with $v \wedge \eta = 0$, with the orientation such that $\eta > 0$. □

Lemma 8.2 *There is a short exact sequence of Lie groups*

$$1 \to \{\pm I\} \to \mathsf{SO}(V) \xrightarrow{(\lambda^+, \lambda^-)} \mathsf{SO}(\Lambda^+) \times \mathsf{SO}(\Lambda^-) \to 1.$$

Proof If $A \in \ker(\lambda^+, \lambda^-)$, so A acts trivially in $\Lambda^\pm$, then it also acts trivially in $\Lambda^2 V$. Hence, by the previous lemma, A preserves every 2-plane. It therefore preserves the intersection of any two 2-planes, hence preserves every line, hence is scalar; so $A = \pm I$.

We have $\dim \mathsf{SO}(4) = 2 \dim \mathsf{SO}(3) = 6$. The corresponding map of Lie algebras (i.e., the derivative $D_I(\lambda^+, \lambda^-)$) is injective, since $\ker(\lambda^+, \lambda^-)$ is discrete, hence surjective because of the equality of dimensions. Thus the image of (λ^+, λ^-) contains the identity component of $\mathsf{SO}(3) \times \mathsf{SO}(3)$, which is the whole group. □

Hence

$$\mathrm{SO}(4)/\{\pm I\} \cong \mathrm{SO}(3) \times \mathrm{SO}(3).$$

8.1.3 Equivalence of conformal structures with maximal positive-definite subspaces

Let V be 4-dimensional. Then $\Lambda^2 V$ carries the intrinsic symmetric pairing

$$q\colon \Lambda^2 V \times \Lambda^2 V \to \det V$$

defined by wedge product: $q(\eta) = \eta \wedge \eta$. An element of $\mathsf{GL}(V)$ of determinant -1 induces an isometry $(\Lambda^2 V, q) \to (\Lambda^2 V, -q)$, so q has neutral signature. Once one orients V, q has a well-defined sign, and then has signature zero. The choice of a conformal structure $[g] = \mathbb{R}_{>0} \cdot g$ defines the maximal positive subspace Λ^+, and the maximal negative subspace Λ^- which is exactly $(\Lambda^+)^\perp$.

We consider the map

$$\sigma\colon \mathsf{conf}(V) \to \mathsf{Gr}_3(\Lambda^2 V)^+, \quad \sigma[g] = \Lambda^+_g$$

from the space of conformal structures (an open set in $(\mathrm{Sym}^2 V^*)/\mathbb{R}_+$) to the Grassmannian $\mathsf{Gr}_3(\Lambda^2 V)$ of 3-planes, or more precisely, to its open set $\mathsf{Gr}_3(\Lambda^2 V)^+$ of positive-definite 3-planes.

Proposition 8.3 *σ is bijective.*

Proof $\mathsf{SL}(V)$ acts transitively on $\mathsf{conf}(V)$, so a choice of inner product g gives us an identification

$$\mathsf{conf}(V) = \frac{\mathsf{SL}(V)}{\mathsf{SO}(V,g)}.$$

The Lie group $\mathsf{SO}(\Lambda^2 V, q)$ has two components: given $h \in \mathsf{SO}(\Lambda^2 V, q)$, one has a map $\Lambda_g^+ \to \Lambda_g^+$ given by $h\colon \Lambda_g^+ \to \Lambda^2 V$ followed by orthogonal projection $\Lambda^2 V \to \Lambda_g^+$; this map may preserve or reverse orientation. The identity component $\mathsf{SO}(\Lambda^2 V, q)_0$ acts transitively on $\mathsf{Gr}_3(\Lambda^2 V)^+$ with stabilizer $\mathsf{SO}(\Lambda_g^+) \times \mathsf{SO}(\Lambda_g^-)$, so

$$\mathsf{Gr}_3(\Lambda^2 V)^+ = \frac{\mathsf{SO}(\Lambda^2 V, q)_0}{\mathsf{SO}(\Lambda_g^+) \times \mathsf{SO}(\Lambda_g^-)}.$$

The action of $\mathsf{SL}(V)$ on 2-forms defines a representation

$$\rho\colon\ \mathsf{SL}(V) \to \mathsf{SO}(\Lambda^2 V, q)_0.$$

As we saw in the proof of Lemma 8.2, one has $\ker \rho = \pm I$. Now, $\mathsf{SL}(V)$ has dimension $4^2 - 1 = 15$, while $\mathsf{SO}(\Lambda^2 V, q)$ has complexification $\mathrm{SO}(6,\mathbb{C})$ and so has dimension $\binom{6}{2} = 15$. Therefore $D\rho$ is an isomorphism of Lie algebras, and it follows that ρ defines a short exact sequence

$$1 \to \pm I \to \mathsf{SL}(V) \xrightarrow{\rho} \mathsf{SO}(\Lambda^2 V, q)_0 \to 1.$$

ρ maps $\mathsf{SO}(V,g)$ to $\mathsf{SO}(\Lambda_g^+) \times \mathsf{SO}(\Lambda_g^-)$ by the 2-fold cover (λ^+, λ^-), and therefore induces a diffeomorphism

$$\frac{\mathsf{SL}(V)}{\mathsf{SO}(V,g)} = \frac{\mathsf{SL}(V)/\pm I}{\mathsf{SO}(V,g)/\pm I} \xrightarrow{\cong} \frac{\mathsf{SO}(\Lambda^2 V, q)_0}{\mathsf{SO}(\Lambda_g^+) \times \mathsf{SO}(\Lambda_g^-)}.$$

Since it preserves the natural basepoints, and is $\mathsf{SL}(V)$-equivariant, this diffeomorphism coincides with our map $\mathsf{conf}(V) \to \mathsf{Gr}_3(\Lambda^2 V)^+$. □

8.1.4 Conformal structures as maps $\Lambda^- \to \Lambda^+$

Fix a reference metric g_0, defining the splitting $\Lambda^2 V = \Lambda_0^+ \oplus \Lambda_0^-$. Let Λ^- be another negative-definite 3-dimensional subspace of $\Lambda^2 V$. Its intersection with Λ_0^+ is trivial, so it projects isomorphically to Λ_0^-. Therefore

$$\Lambda^- = \Gamma_m,$$

the graph of a uniquely determined linear map $m \in \mathsf{Hom}(\Lambda_0^-, \Lambda_0^+)$. For $\alpha \neq 0 \in \Lambda_0^-$, one has

$$0 > q(\alpha + m(\alpha)) = \alpha^2 + m(\alpha)^2 = |\alpha|^2 - |m(\alpha)|^2$$

so the graph Γ_m is negative-definite if and only if m has operator norm $|m| < 1$. Thus

Proposition 8.4 *The map $m \mapsto \Gamma_m$ identifies maps $\Lambda_0^- \to \Lambda_0^+$, of operator norm < 1, with $\mathsf{Gr}_3(\Lambda^2 V)^-$.*

Lemma 8.5 *If $\Lambda^- = \Gamma_m$ then its orthogonal complement Λ^+ is Γ_{m^*}, the graph of the adjoint map m^*.*

Proof For $\alpha_\pm \in \Lambda_0^\pm$, one has

$$(\alpha_+ + m^*\alpha_+) \wedge (\alpha_- + m\alpha_-) = \alpha_+ \wedge m\alpha_- + m^*\alpha_+ \wedge \alpha_- = g(\alpha_+, m\alpha_-) - g(m^*\alpha_+, \alpha_-) = 0.$$

□

This gives a representation of conformal structures as linear maps.

We can compute the components of $\alpha = \alpha_+ + \alpha_-$ with respect to the new decomposition $\Lambda^2 V = \Gamma_m \oplus \Gamma^*_m$. Write

$$\alpha = (\gamma + m\gamma) + (\beta + m^*\beta) \in \Gamma_m \oplus \Gamma_{m^*},$$

so $\beta \in \Lambda^+_0$ and $\gamma \in \Lambda^-_0$. Then

$$\alpha_+ = \beta + m\gamma, \quad \alpha_- = \gamma + m^*\beta,$$

so

$$m^*\alpha_+ - \alpha_- = (m^*m - I)\gamma, \quad \alpha_+ - m\alpha_- = (I - mm^*)\beta.$$

Now, mm^* and m^*m have operator norm < 1, so $m^*m - I$ and $I - mm^*$ have vanishing kernel and so are invertible. Thus

$$\gamma = (m^*m - I)^{-1}(m^*\alpha_+ - \alpha_-),$$
$$\beta = (I - mm^*)^{-1}(\alpha_+ - m\alpha_-).$$

8.2 2-planes

We consider the null-cone $N_V \subset \Lambda^2 V$ of the wedge-square form q, i.e. the space of 2-forms η with $q(\eta) = 0$.

Lemma 8.6 *N_V is the space of decomposable 2-forms $u \wedge v$.*

Proof Certainly $q(u \wedge v) = 0$. Any $\omega \in \Lambda^2 V$ can be thought of as a skew form on V^*, and as such its rank is 0, 2 or 4. If $\omega \wedge \omega = 0$ but $\omega \neq 0$ then the rank is 2. In this case, there is a 2-dimensional space $K_\omega \subset V$ of vectors x such that $\omega \wedge x = 0$. Then ω is necessarily a volume form for K_ω. □

Lemma 8.7 *Fix an inner product g in V. Then a 2-form $\omega \in \Lambda^2 V$ with $|\omega|^2 = 2$ is decomposable if and only if $|\omega^+| = 1 = |\omega^-|$.*

Proof One has

$$\omega^2 = |\omega^+|^2 - |\omega^-|^2, \quad |\omega|^2 = |\omega^+|^2 + |\omega^-|^2.$$

Since ω is decomposable if and only if $\omega \wedge \omega = 0$, the lemma follows. □

There is a map

$$\widetilde{\mathsf{Gr}}_2(V) \to (N_V \setminus \{0\})/\mathbb{R}_+,$$

from the Grassmannian of oriented 2-planes in V to rays in the null-cone, sending K to an oriented volume form ω_K for K. We already observed that oriented 2-planes correspond bijectively with decomposable 2-forms, so by the first of the two foregoing lemmas, this map is bijective. Thus by the second lemma of the pair we deduce

Proposition 8.8 *There is a diffeomorphism $\widetilde{\mathsf{Gr}}_2(V) \to S^2 \times S^2$, mapping K to the self-dual and anti-self-dual parts of ω_K.*

9 Self-duality in 4 dimensions: Hodge-theoretic aspects

9.1 The Hodge theorem

9.1.1 The co-differential

For an oriented Riemannian n-manifold (M, g), and $x \in M$, the inner product on T_x^*M determines the Hodge star operator $\star\colon \Lambda^k T_x^* X \to \Lambda^{n-k} T_x^* X$. It amounts to a map of vector bundles $\star\colon \Lambda^k T^*X \to \Lambda^{n-k} T^*X$.

We define the co-differential $d^*\colon \Omega_M^{k+1} \to \Omega_M^k$ by

$$d^* = (-1)^{k+1} \star^{-1} d \star = (-1)^{k+1}(-1)^{k(n-k)} \star d \star = (-1)^{nk+1} \star d \star .$$

Here d is the exterior derivative. From the facts that $d^2 = 0$ and $\star^2 = \pm I$ it follows that $(d^*)^2 = 0$.

9.1.2 The co-differential as formal adjoint

Assume now that M is compact. One then has the L^2-inner product on Ω_M^k:

$$\langle \alpha_1, \alpha_2 \rangle_{L^2} = \int_M g(\alpha_1, \alpha_2) \,\mathsf{vol}_g.$$

Stokes's theorem implies the 'integration by parts' formula $\int_M d\alpha \wedge \beta = -(-1)^k \int_M \alpha \wedge d\beta$ for $\alpha \in \Omega_M^k$ and $\beta \in \Omega_M^{n-k-1}$. Taking $\beta = \star\gamma$, where $\gamma \in \Omega_M^{k+1}$, this implies that

$$\int_M g(d\alpha, \gamma) \mathsf{vol}_g = (-1)^{k+1} \int_M g(\alpha, \star^{-1} d \star \gamma) \mathsf{vol}_g.$$

That is d^* is the formal adjoint to d:

$$\langle d\alpha, \beta \rangle_{L^2} = \langle \alpha, d^*\beta \rangle_{L^2}.$$

9.1.3 Harmonic forms

On any oriented Riemannian manifold (M, g), the Hodge Laplacian on k-forms is

$$\Delta = dd^* + d^*d = (d + d^*)^2\colon \Omega_M^k \to \Omega_M^k$$

and the space of *g-harmonic k-forms* is

$$\mathcal{H}_g^k := \ker \Delta.$$

Plainly $\ker(d + d^*) \subset \mathcal{H}_g^k$. Reverting now to the assumption that M is compact, the identity

$$\langle \alpha, \Delta\alpha \rangle_{L^2} = \|d^*\alpha\|_{L^2}^2 + \|d\alpha\|_{L^2}^2$$

shows that $\mathcal{H}_g^k \subset \ker(d + d^*)$, i.e, that

$$\mathcal{H}_g^k = \ker(d + d^*).$$

9.1.4 Variational characterization.

If $\eta \in \mathcal{H}^k_g$ then $d\eta = 0$, so η represents a cohomology class $[\eta]$.

Lemma 9.1 *A harmonic form η strictly minimizes L^2-norm among representatives of its cohomology class.*

Proof We have
$$\|\eta + d\gamma\|^2_{L^2} - \|\eta\|^2_{L^2} = \|d\gamma\|^2_{L^2} + 2\langle \eta, d\gamma\rangle = \|d\gamma\|^2_{L^2},$$
so $\eta + d\gamma$ has strictly larger L^2-norm than η unless $d\gamma = 0$. □

Consequently, *there is at most one harmonic representative of a fixed cohomology class.*

Lemma 9.2 *If η minimizes L^2 norm among forms $\eta' = \eta + d\gamma$ then $d^*\eta = 0$. Thus if η is also closed then it is harmonic.*

Proof The minimization property implies that $\frac{d}{dt}\big|_{t=0} \|\eta + td\gamma\|^2_{L^2} = 0$. That is, $0 = 2\langle \eta, d\gamma\rangle_{L^2} = 2\langle d^*\eta, \gamma\rangle_{L^2}$. Taking $\gamma = d^*\eta$, we see that $d^*\eta = 0$. □

9.1.5 The theorem

Notice that $\ker d$ is precisely the L^2-orthogonal complement to $\operatorname{im} d^*$ in Ω^k_M, since if α is orthogonal to $\operatorname{im} d^*$ then $\langle d\alpha, \beta\rangle_{L^2} = \langle \alpha, d^*\beta\rangle_{L^2} = 0$ for any β, whence $\|d\alpha\|^2_{L^2} = 0$, so $d\alpha = 0$.

Theorem 9.3 (The Hodge theorem) *There are L^2-orthogonal decompositions*
$$\Omega^k_M = \ker d \oplus \operatorname{im} d^*.$$
and
$$\ker d = \mathcal{H}^k_g \oplus \operatorname{im} d.$$
Hence the quotient map $\ker d \to H^k_{DR}(M) = \ker d / \operatorname{im} d$ defines an isomorphism
$$\mathcal{H}^k_g \xrightarrow{\simeq} H^k_{DR}(M).$$

These assertions are proved by working with suitable Hilbert spaces of forms (the disadvantage of the L^2 inner product on Ω^k_M being that its norm is incomplete). The relevant analysis will be described later in the course; but schematically, the argument is as follows:

- One works with Hilbert spaces of forms of Sobolev class L^2_ℓ, for integers $\ell > 0$; one then has a bounded linear map $\delta = d \oplus d^* \colon L^2_{\ell+1}(\Lambda^*) \to L^2_\ell(\Lambda^*)$.
- δ is formally self-adjoint, meaning that $\langle \delta\alpha, \beta\rangle_{L^2} = \langle \alpha, \delta\beta\rangle_{L^2}$.
- The step is to prove an *elliptic estimate* (Garding's inequality),
$$\|\eta\|_{L^2_{\ell+1}} \leq C\left(\|\delta\eta\|_{L^2_\ell} + \|\eta\|_{L^2_\ell}\right)$$
which implies that δ is Fredholm—in particular, its image is closed—and that its cokernel is the kernel of the formal adjoint δ.

These points will be stated precisely later in the course, though we will not prove the elliptic estimate. Thus $L^2_\ell(\Lambda^*) = \mathcal{H}^*_g \oplus \operatorname{im}\delta$. Since $\operatorname{im}\delta \subset \operatorname{im} d + \operatorname{im} d^*$, one has $L^2_\ell(\Lambda^*) = \mathcal{H}^*_g \oplus \operatorname{im} d \oplus \operatorname{im} d^*$. The elliptic estimate also implies that $\mathcal{H}^*_g = \ker\delta$ consists of forms of class $\bigcap_\ell L^2_\ell = C^\infty$. One takes the intersection over ℓ of the spaces $L^2_\ell(\Lambda^*)$, with their decompositions $\mathcal{H}^*_g \oplus \operatorname{im} d \oplus \operatorname{im} d^*$, to obtain the theorem.

A sample reference is [10].

9.2 Self-dual and anti-self-dual harmonic forms

Now let (X,g) be a closed, oriented, Riemannian 4-manifold. The bundle of 2-forms $\Lambda^2 = \Lambda^2 T^*X$ decomposes under the Hodge star:

$$\Lambda^2 = \Lambda_g^+ \oplus \Lambda_g^-.$$

We write $\Omega_g^\pm$ for sections of $\Lambda_g^\pm$. The whole of the previous lecture applies to $\Lambda_g^\pm$; for instance, conformal structures on X are equivalent to maximal positive-definite sub-bundles of Λ^2, by mapping $[g]$ to Λ_g^+. However, our focus now is on harmonic forms.

One has

$$d^* = -\star d \star \colon \Omega_X^2 \to \Omega_X^1.$$

Hence, for a 2-form η,

$$\eta \in \ker(d+d^*) \;\Leftrightarrow\; \star\eta \in \ker(d+d^*).$$

Hence, *if one takes a harmonic 2-form $\eta \in \mathcal{H}_g^2(X)$, its components $\eta^\pm \in \Omega^\pm(X)$ are again harmonic.*

Thus we have a decomposition

$$\mathcal{H}_g^2 = \mathcal{H}_g^+ \oplus \mathcal{H}_g^-$$

into self-dual and anti-self-dual harmonic 2-forms. Notice that $\int_X \eta\wedge\eta = \|\eta\|_{L^2}^2 > 0$ when $\eta \neq 0 \in \mathcal{H}_g^+$, while $\int_X \omega\wedge\omega = -\|\omega|_{L^2}^2 < 0$ when $\omega \neq 0 \in \mathcal{H}_g^-$. Thus $\mathcal{H}_g^+$ and $\mathcal{H}_g^-$ are positive- and negative-definite subspaces for the wedge product quadratic form. They are maximal such subspaces, since the complement one another. Thus

$$\dim \mathcal{H}_g^\pm = b^\pm(X),$$

where $b^\pm(X)$ denotes the dimension of a maximal positive- or negative-definite subspace in $H^2(X;\mathbb{R})$.

9.3 The self-duality complex

Consider the sequence of operators

$$0 \to \Omega_X^0 \xrightarrow{d} \Omega_X^1 \xrightarrow{d^+} \Omega_g^+ \to 0.$$

Here $d^+\alpha = (d\alpha)^+ = \frac{1}{2}(1+\star)d\alpha$. We can regard this sequence as a cochain complex $(\mathcal{E}^*,\delta)$, where $\mathcal{E}^0 = \Omega_X^0$, $\mathcal{E}^1 = \Omega_X^1$ and $\mathcal{E}_X^2 = \Omega_g^+$.

Theorem 9.4 *The cohomology spaces of this complex are as follows:*

$$H^0(\mathcal{E}) = H_{DR}^0(X);$$
$$H^1(\mathcal{E}) = H_{DR}^1(X);$$
$$H^2(\mathcal{E}) \xleftarrow{\cong} \mathcal{H}_g^+(X).$$

Proof $H^0(\mathcal{E}) = \Omega_X^0$ by definition. If $\alpha \in \Omega_X^1$ then we can write $d\alpha = d^+\alpha + d^-\alpha$, and $\int_X d\alpha\wedge d\alpha = \|d^+\alpha\|_{L^2}^2 - \|d^-\alpha\|_{L^2}^2$. But $\int_X d\alpha\wedge d\alpha = \int_X d(\alpha\wedge d\alpha) = 0$, so

$$\|d^+\alpha\|_{L^2}^2 = \|d^-\alpha\|_{L^2}^2.$$

Hence $\ker d^+ = \ker d^- = \ker d$, and we see that $H^1(\mathcal{E}) = H_{DR}^1(X)$.

As for $H^2(\mathcal{E})$, it is easy to check that $\mathcal{H}^+_g$ is L^2-orthogonal to $\operatorname{im} d^+$. Hence the composite

$$\mathcal{H}^+_g \hookrightarrow \Omega^+_g \to \Omega^+_g / \operatorname{im} d^+$$

is injective. The point is to show that it is also surjective. Take $\omega \in \Omega^+$, and use the Hodge theorem to write it as

$$\omega = \omega_{harm} + d\alpha + \star d\beta,$$

the three components being respectively in $\mathcal{H}^2_g$, $\operatorname{im} d$ and $\operatorname{im} d^*$. These Hodge components are unique. Since $\star\omega = \omega$, we see that $\star\omega_{harm} = \omega_{harm}$ and $d\alpha = d\beta$ so

$$\omega = \omega_{harm} + 2d^+\alpha \in \mathcal{H}^+_g \oplus \operatorname{im} d^+.$$

□

9.4 The derivative of the period map

Let $\mathcal{C}_X$ be the space of conformal structures on our 4-manifold X. Precisely, fix an integer $r \geq 3$, and let $\mathcal{C}_X$ be the space of Riemannian metric of class C^r, modulo positive functions $f > 0$ also of class C^r.

Once a reference conformal structure $[g_0] \in \mathcal{C}_X$ is chosen, $\mathcal{C}_X$ can be identified with an open set in a Banach space. Namely, it is is the space of C^r maps $m\colon \Lambda^- \to \Lambda^+$ with $|m_x| < 1$ for all $x \in X$. The resulting smooth Banach manifold structure is independent of $[g_0]$. The tangent space to $\mathcal{C}_X$ at $[g_0]$ is the space of all bundle maps $m\colon \Lambda^- \to \Lambda^+$.

Lemma 9.5 *Consider the map* $\mathsf{star}\colon \mathcal{C}_X \to C^r(X, \mathsf{Hom}(\Lambda^2_X, \Lambda^2_X))$ *sending* $[g]$ *to* $\star_g$. *Its derivative at* $[g_0]$ *is given by*

$$D_{[g_0]}\mathsf{star}(m) = 2\begin{bmatrix} 0 & m^* \\ -m & 0 \end{bmatrix}$$

as a map $\Lambda^+ \oplus \Lambda^- \to \Lambda^+ \oplus \Lambda^-$.

Proof Take $\alpha_- \in \Lambda^-$. Let $[g_t]$ be the conformal structure with bundle map tm, and $\star_t = \star_{[g_t]}$. Then $\Lambda^-_t = \Gamma_{tm}$, so

$$\star_t(\alpha_- + tm\alpha_-) = -\alpha_- - tm\alpha_-.$$

Differentiating with respect to t, at $t = 0$, we get

$$D\mathsf{star}(m)(\alpha_-) + \star_0 m(\alpha_0) = -m\alpha_-,$$

i.e.,

$$D\mathsf{star}(m)(\alpha_-) = -2m\alpha_-.$$

This handles the first column of the matrix; the second is obtained similarly by writing $\Lambda^+_t = \Gamma_{tm^*}$. □

In this context, the *period map* is the map

$$P\colon \mathcal{C}_X \to \mathsf{Gr}_{b^-(X)} H^2_{DR}(X), \quad [g] \mapsto \mathcal{H}^-_g$$

mapping a conformal structure to its space of ASD harmonic forms, viewed as a subspace of the fixed space $H^2_{DR}(X)$.

The reference conformal structure $[g_0]$ in $\mathcal{C}_X$ has a neighborhood identified with an open neighborhood of the origin in $C^r(X, \mathsf{Hom}(\Lambda^+, \Lambda^-)$. The Grassmannian has a neighborhood of $\mathcal{H}^- = \mathcal{H}^-_{g_0}$ identified with $\mathsf{Hom}(\mathcal{H}^-, \mathcal{H}^+)$ (by taking graphs), so near $[g_0]$ we think of P as a map

$$P\colon C^r(X, \mathsf{Hom}(\Lambda^+, \Lambda^-)) \to \mathsf{Hom}(\mathcal{H}^-, \mathcal{H}^+),$$

defined near the origin.

In particular, the tangent space to the Grassmannian at $P[g_0]$ is $\mathsf{Hom}(\mathcal{H}^-, \mathcal{H}^+)$, so the derivative of P is a map

$$D_{[g_0]}P\colon C^r(X, \mathsf{Hom}(\Lambda^-, \Lambda^+)) \to \mathsf{Hom}(\mathcal{H}^-, \mathcal{H}^+).$$

Proposition 9.6 $(D_{[g_0]}P)(m)\colon \mathcal{H}^- \to \mathcal{H}^+$ *is the map*

$$\alpha^- \mapsto m(\alpha^-)_{harm},$$

where $_{harm}\colon \Omega^+ \to \mathcal{H}^+$ *is the projection arising from the* L^2*-orthogonal decomposition* $\Omega^+ = \operatorname{im} d^+ \oplus \mathcal{H}^+$.

Proof Let g_t be the conformal class whose corresponding bundle map is tm, $|t| < 1$, and $\star_t$ its Hodge star. Think of P as a map $C^r(X, \mathsf{Hom}(\Lambda^-, \Lambda^+) \to \mathsf{Hom}(\mathcal{H}^-, \mathcal{H}^+)$, defined near the origin. Then, for $\alpha^- \in \mathcal{H}^-$, we have a closed 2-form $\alpha_t^- = \alpha^- + P(g_t)(\alpha^-) \in \Gamma_{P(g_t)}$. This form is cohomologous to a closed, g_t-ASD form; say

$$\alpha_t^- + d\phi_t \in \mathcal{H}_t^-.$$

We have

$$(D_{[g_0]}P)(m)(\alpha^-) = \left.\frac{d\alpha_t^-}{dt}\right|_{t=0} =: \xi \in \mathcal{H}^+.$$

Differentiating the relation

$$(1 + \star_t)(\alpha_t^- + d\phi_t) = 0$$

at $t = 0$, using our computation of the derivative of star, we find

$$-2m(\alpha^-) + (1 + \star_0)(\xi + \dot{\phi}) = 0,$$

where $\dot{\phi} = (d/dt)|_{t=0}\phi_t$. That is,

$$\xi = m(\alpha) - d^+\dot{\phi}.$$

□

10 Covariant derivatives

10.1 Covariant derivatives in vector bundles

Let $E \to M$ be a complex vector bundle over the manifold M, equipped with a hermitian product $(\cdot,\cdot)$. We write Γ for the vector space of C^∞ sections of a vector bundle—so $\Gamma(M,E)$ means the space of sections of $E \to M$.

Definition 10.1 A *covariant derivative*, or *connection*, in E is a $\mathbb{C}$-linear map

$$\nabla\colon \Gamma(M,E) \to \Gamma(M, T^*M \otimes E),$$

obeying the Leibniz rule

$$\nabla(fs) = df \otimes s + f\nabla s, \quad f \in C^\infty(M,\mathbb{C}), \quad s \in C^\infty(M,E).$$

It is called *unitary* if

$$\nabla(s_1,s_2) = (\nabla s_1, s_2) + (s_1, \nabla s_2).$$

Lemma 10.2 *Covariant derivatives are local operators:* $(\nabla s)(x)$ *depends only on the germ of* s *near* x.

Proof If s_1 and s_2 agree on a neighborhood U of x, take a function $\chi \in C^\infty(M)$, supported in U, $\chi = 1$ near x. Then $\chi s_1 = \chi s_2$. The Leibniz rule then shows that $\nabla(\chi s_i)(x) = (\nabla s_i)(x)$, so $(\nabla s_0)(x) = (\nabla s_1)(x)$. □

Example 10.3 In a trivialized line bundle $\underline{\mathbb{C}} \to M$, a connection amounts to a map $C^\infty(M) \to C^\infty(M, T^*M)$ satisfying the Leibniz rule. The exterior derivative d is an example of a connection, called the *trivial connection*. If the trivialization is unitary, i.e. $(\cdot,\cdot)_x$ is independent of $x \in M$, then d is a unitary connection.

Example 10.4 In a higher rank trivial line bundle $\underline{V} \to M$, with fiber the vector space V, we still have a trivial connection d:

$$d = d \otimes 1_V\colon C^\infty(M, \mathbb{C} \otimes V) \to C^\infty(M, T^*M \otimes V).$$

The difference $\nabla - \nabla'$ between covariant derivatives in E is $C^\infty(M)$-linear; indeed, it is any section of the vector bundle $T^*M \otimes \operatorname{End} E$. Thus the space of covariant derivatives $\mathcal{C}(E)$ is an affine space:

$$\mathcal{C}(E) = \nabla + \Omega^1_M(\operatorname{End} E).$$

If ∇ and ∇' are unitary, $\nabla - \nabla'$ is a section of $T^*M \otimes \mathfrak{u}(E)$, where $\mathfrak{u}(E)$ is the Lie algebra bundle of skew-hermitian endomorphisms of the fibers E_x. Thus the space of unitary covariant derivatives is

$$\mathcal{C}(E, \langle\cdot,\cdot\rangle) = \nabla + \Omega^1_M(\mathfrak{u}(E)).$$

Example 10.5 In the case of a trivialized bundle $\underline{V} \to M$, covariant derivatives take the form

$$d + A, \quad A \in C^\infty(M, T^*M \otimes \operatorname{End} V).$$

Covariant derivatives are first-order operators. Since every vector bundle is locally trivial, and ∇ is a local operator, we can always represent ∇ locally in the form $\nabla = d + A$. Hence covariant derivatives are *first-order operators*, in the sense that $(\nabla s)(x)$ depends only on the 1-jet of s at x, i.e. on $s(x) \in E_x$ and on $ds(x)\colon T_xM \to T_{s(x)}E$.

Pullback. Connections pull back: if $f\colon N \to M$ is a smooth map, and ∇ a connection in $E \to M$, then there is a connection $f^*\nabla$ in $f^*E \to N$. A section of f^*E takes the form f^*s, where s is a section of E. We put $(f^*\nabla)(f^*s) = (Df^* \otimes f^*)(\nabla s)$.
One has $(g \circ f)^* = f^* \circ g^*$, as usual.

Existence. We can now verify that (unitary) connections exist in any (hermitian) vector bundle $E \to M$. Let $\{U_i\}$ be a locally finite open cover such that $E|_{U_i}$ is trivial. Then a connection ∇_i exists in $E|_{U_i}$ (since it exists locally). Let $\{\rho_i\}$ be a partition of unity subordinate to the cover. Let $\nabla s = \sum_i \rho_i \nabla_i s$. Then

$$\nabla(f \cdot s) = \sum_i (\rho_i f \nabla_i s + \rho_i df \otimes s) = f\nabla s + df \otimes s,$$

so ∇ is a connection. The same construction works for unitary connections.

10.2 Curvature

A connection ∇ in E defines a 'coupled exterior derivative'

$$d_\nabla \colon \Omega^k(E) \to \Omega^{k+1}(E),$$

which for $k = 0$ is just given by $d_\nabla = \nabla$. The rule is

$$d_\nabla(\eta \otimes s) = (-1)^k \eta \wedge \nabla s + d\eta \otimes s, \quad s \in C^\infty(M,E), \quad \eta \in \Omega^k_M.$$

This again obeys the Leibniz rule, since for functions f, one has

$$d_\nabla(f\eta \otimes s) = (-1)^k f\eta \wedge \nabla s + f\, d\eta \otimes s + (df \wedge \eta) \otimes s = f\, d_\nabla(\eta \otimes s) + df \wedge (\eta \otimes s).$$

Lemma 10.6 *$d_\nabla \circ d_\nabla \colon \Omega^*_M(E) \to \Omega^{*+2}_M(E)$ is linear over the algebra Ω^*_M.*

Proof For $\eta \in \Omega^k_M$,

$$\begin{aligned} d_\nabla \circ d_\nabla(\eta \otimes s) &= d_\nabla((-1)^k \eta \wedge \nabla s + d\eta \otimes s) \\ &= \eta \wedge d_\nabla \circ d_\nabla s + (-1)^k d\eta \wedge \nabla s + (-1)^{k+1} d\eta \wedge \nabla s \\ &= \eta \wedge d_\nabla \circ d_\nabla s. \end{aligned}$$

□

We define the *curvature*

$$F_\nabla \in \Omega^2_M(\operatorname{End} E)$$

to be the endomorphism-valued 2-form such that $d_\nabla \circ d_\nabla s = F_\nabla \otimes s$. In the case of unitary connections, $F_\nabla \in \Omega^2_M(\mathfrak{u}(E))$. By construction,

$$F_{f^*\nabla} = f^* F_\nabla.$$

Of course, the trivial connection d, in the trivial bundle, is *flat*, $F_d = 0$, because $d^2 = 0$.
To repeat: In this formulation, *curvature is defined as the obstruction to $d_\nabla^2 = 0$.*
A covariant derivative in E induces one in $\operatorname{End} E$; its coupled exterior derivative is

$$d_\nabla \colon \Omega^k_M(\operatorname{End} E) \to \Omega^{k+1}_M(\operatorname{End} E), \quad (d_\nabla \alpha) = [d_\nabla, \alpha].$$

Taking note of the sign that appears when we apply d_∇ to As, we see that

$$F_{\nabla + A} = F_\nabla + d_\nabla A + A \wedge A.$$

Here $A \wedge A$ combines the composition of endomorphisms with the wedge product of forms. In particular, in a trivial bundle, one has

$$F_{d+A} = dA + A \wedge A.$$

Example 10.7 In the case of a line bundle $L \to M$, the 1-form A is valued in $\operatorname{End}\mathbb{C} = \mathbb{C}$. Similarly F_∇ is a complex 2-form. In this case, the antisymmetry of the wedge product shows that

$$F_{\nabla+A} = F_\nabla + dA.$$

For unitary connections, A and F_A are $i\mathbb{R}$-valued.

It is instructive to write the formula $F_{d+A} = dA + A \wedge A$ in local coordinates $(x_1, \ldots, x_n)$ on M. Then $A = \sum A_k dx_k$, where A_k is an $r \times r$ matrix, and the curvature is $F_{d+A} = \sum_{i,j} F_{ij} dx_i \wedge dx_j$, where

$$F_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} + [A_i, A_j].$$

For a vector field v, write $\nabla_v \colon \Gamma(M,E) \to \Gamma(M,E)$ for the contraction of ∇ with v. Note that $\nabla_{fv} = f\nabla_v$ for functions f.

In the case of the coordinate vector field $\partial_i = \partial/\partial x_i$, we will abbreviate the notation to $\nabla_i s = \nabla_{\partial_i}$. Then, in our local trivialization, $\nabla_i = \partial_i + A_i$, and

$$[\nabla_i, \nabla_j] = [\partial_i + A_i, \partial_j + A_j] = F_{ij},$$

so that we arrive at a second interpretation of the curvature: *its components measure failure of commutativity of covariant partial derivatives*.

From this formula follows a more general one:

Lemma 10.8 *For vector fields u and v, one has*

$$F_\nabla(u,v) = [\nabla_u, \nabla_v] - \nabla_{[u,v]}.$$

Proof The right-hand side is evidently $\mathbb{R}$-bilinear in u and v. It is actually $C^\infty(M)$-bilinear, since for functions f, one has

$$[\nabla_{fu}, \nabla_v] = f[\nabla_u, \nabla_v] + df(v)\nabla_u$$

while $[f, uv] = f[u,v] - df(v)u$, so that

$$\nabla_{[fu,v]} = f\nabla_{[u,v]} - df(v)\nabla_u.$$

Once one knows that both sides are $C^\infty(M)$-bilinear, it suffices to see that they agree for coordinate vector fields, which we have already established. □

So: *curvature measures the failure of ∇ to preserve brackets.*

Lemma 10.9 (Bianchi identity) $d_\nabla F_\nabla = 0$.

Proof In local coordinates,

$$\begin{aligned} d_\nabla F_\nabla &= \sum_i [\nabla_i, F_{jk}]\, dx_{ijk} \\ &= \sum_{i,j,k} [\nabla_i, [\nabla_j, \nabla_k]]\, dx_{ijk} \\ &= 2 \sum_{i<j<k} \big([\nabla_i, [\nabla_j, \nabla_k] + [\nabla_k, [\nabla_i, \nabla_j]] + [\nabla_k, [\nabla_i, \nabla_j]] \big)\, dx_{ijk} = 0. \end{aligned}$$

□

10.3 Gauge transformations

Definition 10.10 A *gauge transformation* u of a vector bundle $E \to M$ is a bundle automorphism

$$u\colon E \xrightarrow{\cong} E,$$

i.e., a fiberwise linear isomorphism covering id_M. Gauge transformations form a group $\mathcal{G}_E$. When E is given as a hermitian vector bundle, gauge transformations are taken to be fiberwise *unitary*.

There is a bundle of Lie groups $\mathsf{GL}(E) \subset \operatorname{End} E$, with fibers $\mathsf{GL}(E_x)$; and $\mathcal{G}_E$ is the group of its sections. However, $\mathsf{GL}(E)$ is not a *principal* bundle: it does not admit any natural action by $\mathsf{GL}(r, \mathbb{C})$. If one picks a basis $(e_1, \dots, e_r)$ for E_x, then an automorphism of E_x amounts to an invertible matrix $U \in \mathsf{GL}(r, \mathbb{C})$. Changing the basis conjugates U by the change-of-basis matrix. Hence, if $P_E \to M$ is the principal bundle of frames of E, then

$$\mathsf{GL}(E) \cong P_E \times_{\mathsf{GL}(r,\mathbb{C})} \mathsf{GL}(r, \mathbb{C}),$$

the bundle associated with the left action of $\mathsf{GL}(r, \mathbb{C})$ on itself by inner automorphisms, $g \cdot h = ghg^{-1}$.

Gauge transformations act on covariant derivatives, by $\nabla \mapsto u^*\nabla$:

$$(u^*\nabla)(s) = u\nabla(u^{-1}s).$$

Certainly $F_{u^*\nabla} = u^*F_\nabla = uF_\nabla u^{-1}$: this is an instance of the fact that curvature is compatible with pullbacks.

Lemma 10.11 *We have*

$$u^*\nabla = \nabla - (d_\nabla u)u^{-1}.$$

In a trivialized bundle,

$$u^*(d + A) = d - (du)u^{-1} + uAu^{-1}.$$

Proof We have

$$\nabla s = \nabla(uu^{-1}s) = (d_\nabla u)(u^{-1}s) + u\nabla(u^{-1}s).$$

Thus

$$(u^*\nabla)s = u\nabla(u^{-1}s) = \nabla s - (d_\nabla u)(u^{-1}s).$$

This gives the first formula, of which the second is an instance. □

10.4 Flat connections

Theorem 10.12 (flat connections are integrable) *If ∇ is a flat connection in $E \to M$ then, near any point of M, there is a local trivialization of E in which ∇ is the trivial connection d.*

This is an immediate consequence of

Proposition 10.13 *Let $H = (-1, 1)^n \subset \mathbb{R}^n$. Suppose ∇ is a (unitary) connection in the trivial bundle $\underline{\mathbb{R}}^r \to H$ such that $F_\nabla = 0$. Then there is a (unitary) gauge transformation u such that $u^*\nabla$ is trivial.*

Proof We have $\nabla = d + A = s + \sum A_i\, dx_i$ for (skew-hermitian) matrix-valued functions $A_i \colon H \to \operatorname{End}\mathbb{C}^r$. Assume inductively that $A_i = 0$ for $1 \leq i \leq p$ (the initial case, $p = 0$, is a vacuous assumption). The flatness assumption $F_{ij} = 0$ says that $\partial_i A_j - \partial_j A_i + [A_i, A_j] = 0$. Thus for $1 \leq i \leq p$,

$$\partial_i A_{p+1} = 0 :$$

A_{p+1} is independent of $(x_1, \ldots, x_p)$. We claim that there is a gauge transformation $u = u_{p+1}$ such that $u^*\nabla = d + \sum B_i\, dx_i$ where $B_i = 0$ for $1 \leq i \leq p+1$. From this claim the proposition follows.

Now,

$$B_i = -(\partial_i u)u^{-1} + uA_i u^{-1},$$

so what we want is a solution to the system of ODE for matrix-valued functions

$$\partial_i u = 0, \quad i = 1, \ldots, p;$$
$$\partial_{p+1} u + uA_{p+1} = 0.$$

The equation $\partial_{p+1} u + uA_{p+1} = 0$ is a linear ODE in the variable x_{p+1}, with coefficients varying smoothly in the parameters $(x_{p+2}, \ldots, x_n)$—and independent of $(x_1, \ldots, x_p)$. If we impose the initial condition that $u(x) = I$ when $x_{p+1} = 0$, the solution is unique, smooth as a function of x, and independent of $(x_1, \ldots, x_p)$, as required.

Finally, we address unitarity, assuming the A_i are skew-hermitian. We want $u(x)u(x)^\dagger = I$. But

$$\partial_{p+1} u(x)u(x)^\dagger = (\partial_{p+1} u)u^\dagger + u\partial_{p+1} u^\dagger = -uA_{p+1}u^\dagger - uA_{p+1}^\dagger u^\dagger = -u(A_{p+1} + A_{p+1}^\dagger)u^\dagger = 0,$$

and so from unitarity of $u(x)$ when $x_{p+1} = 0$ we deduce unitary for all x. □

10.5 Flat connections are local systems

A connection ∇ in a vector bundle E can be given by the following data.

First, we define E by specifying an open cover $\{U_\alpha\}$ of M. We take E to be trivialized over U_α, and we give transition functions $\tau_{\alpha\beta} \colon U_{\alpha\beta} \to \mathsf{GL}(r, \mathbb{C})$, forming a cocycle:

$$\tau_{\alpha\beta} = \tau_{\beta\alpha}^{-1}, \quad \tau_{\alpha\beta} \circ \tau_{\beta\gamma} \circ \tau_{\gamma\alpha} = I.$$

Over U_α, we have $\nabla = d + A_\alpha$, where

$$A_\alpha \in \Omega^1_{U_\alpha}(\operatorname{End}\mathbb{C}^r).$$

Thus the connection amounts to the collection of 1-forms $\{A_\alpha\}$. They must obey the transformation rule

$$A_\beta = \tau_{\alpha\beta} A_\alpha \tau_{\beta\alpha} - (d\tau_{\alpha\beta})\tau_{\beta\alpha}.$$

We saw last time that a *flat* connection is locally a trivial connection. Thus we can take $A_\alpha = 0$. The condition on the transition functions is then

$$d\tau_{\alpha\beta} = 0,$$

i.e., that $\tau_{\alpha\beta}$ is locally constant.

A vector bundle with a system of locally constant transition functions is called a *local system*. Thus a flat connection determines, and is determined by, a local system.

A vector bundle with flat connection can be is equivalent to a vector bundles with locally constant transition functions, also known as local systems.

11 $\mathsf{U}(1)$-connections

11.1 Connections and gauge transformations in line bundles

Let $L \to M$ be a hermitian line bundle. Let $\mathcal{A}_L$ be the space of unitary connections,

$$\mathcal{A}_L = \nabla + i\Omega^1_M.$$

Note that Ω^1_M has the structure of a topological vector space (it is a Fréchet space, its topology defined by a sequence of C^r-norms), and so $\mathcal{A}_L$ is an affine Fréchet space. Let $\mathcal{G}_L$ be the group of unitary gauge transformations. Since an automorphism of a 1-dimensional hermitian vector space is a unit scalar, one has

$$\mathcal{G}_L = C^\infty(M, \mathsf{U}(1).)$$

$\mathcal{G}_L$ also has a C^∞ topology, and acts continuously on $\mathcal{A}_L$. Introduce the orbit-space

$$\mathcal{B}_L = \mathcal{C}_L/\mathcal{G}_L.$$

It is by no means obvious that this quotient space is Hausdorff. It turns out that it is—we can identify it quite concretely with the product of a $b_1(M)$-dimensional torus and an affine Fréchet space. We establish this picture now.

11.1.1 Chern–Weil theory

Let $\nabla \in \mathcal{A}_L$. Then:

- *$F_\nabla \in \Omega^2(M; i\mathbb{R})$ is closed*, i.e., $dF_\nabla = 0$.
  Indeed, in a local trivialization we can write $\nabla = d + A$, with A an imaginary 1-form, and $F_\nabla = dA$ is locally exact.
- *The closed 2-form iF_∇ represents a class $c_L \in H^2_{DR}(M)$ independent of ∇.*
  Indeed, for $a \in \Omega^1_M(i\mathbb{R})$, we have $F_{\nabla+a} = F_\nabla + da$.
- *There is a universal constant $\lambda \in \mathbb{R}$ such that c_L is the image in $H^2_{DR}(M)$ of $\lambda c_1(L)$, where $c_1(L)$ is the Chern class.*
  The pullback formula $f^*F_\nabla = F_{f^*\nabla}$ shows that $c_{f^*L} = f^*c_L$, matching the fact that $f^*c_1(L) = c_1(f^*L)$. Every line bundle is the pullback of the tautological line bundle $\Lambda_N \to \mathbb{C}P^N$, for some N, via a map $f\colon M \to \mathbb{C}P^N$. So it suffices to show that $c_{\Lambda_N} = \lambda c_1(\Lambda_N)$. Moreover, $H^2_{DR}(\mathbb{C}P^N) = H^2_{DR}(\mathbb{C}P^1)$, and $\Lambda_N|_{\mathbb{C}P^1} = \Lambda_1$, so it suffices to show that $c_{\Lambda_1} = \lambda c_1(\Lambda_1)$. And $H^2_{DR}(\mathbb{C}P^1) = \mathbb{R}$, generated by $c_1(\Lambda)$, so there is some λ such that $c_\Lambda = \lambda c_1(\Lambda)$.

To find the constant λ, it suffices to check one particular line bundle over $\mathbb{C}P^1$. To do so, it is useful to take the point of view that ∇ is encoded as a 1-form on the unit circle bundle; one can then proceed as in Bott–Tu, *Differential forms in algebraic topology*, Ex. 6.44.1. The result is that

$$c_1(L) = \frac{1}{2\pi}[iF_\nabla].$$

11.1.2 Structure of $\mathcal{B}_L$

Gauge transformations u act on $\mathcal{A}_L$ by

$$u^*\nabla = \nabla - (du)u^{-1}.$$

Observations:

- The action is *semi-free*: let $\mathsf{U}(1) \subset \mathcal{G}_L$ be the group of constant gauge transformations. Then $\mathsf{U}(1)$ acts trivially, but the resulting action of $\mathcal{G}_L/\mathsf{U}(1)$ is free.
- One has

  $$\pi_0\mathcal{G}_L = H^1(M;\mathbb{Z})$$

  (homotopy classes of maps from M to the circle).
- The identity component $\mathcal{G}_L^\circ$ consists of maps $u = e^{i\xi}$ where $\xi \in C^\infty(M,\mathbb{R})$. We have $(e^{i\xi})^*\nabla = \nabla - i\,d\xi$. Hence the choice of ∇ defines an identification

  $$\mathcal{A}_L/\mathcal{G}_L^\circ \cong i(\Omega^1_M/d\Omega^0_M).$$
- Let $\mathcal{S}$ denote the 'Coulomb gauge slice',

  $$\mathcal{S} = \{\nabla + ia : d^*a = 0\},$$

  an affine Fréchet space. By the Hodge theorem, $\Omega^1_M/d\Omega^0_M = \ker d^*$. Thus projection

  $$\mathcal{S} \to \mathcal{A}_L/\mathcal{G}_L^\circ$$

  is a homeomorphism.
- The component group $\pi_0\mathcal{G}_L = H^1(X;\mathbb{Z})$ acts on $\mathcal{A}_L/\mathcal{G}_L^\circ$ (equally, on $\mathcal{S}_L$), and we next work out how this works. We have $u^*\nabla = \nabla - (du)u^{-1} = \nabla - d(\log u)$. The closed 1-form $d(\log u)$ represents the class of u in $H^1(X;\mathbb{Z})$. There is a unique cohomologous form, $d(\log u) + d\xi$, with $\nabla + d(\log u) + d\xi \in \mathcal{S}$; being both closed and co-closed, $d(\log u) + d\xi \in \mathcal{H}^1_g$. Thus we have

  $$\mathcal{B}_L \cong \mathcal{S}/\pi_0\mathcal{G}_L \cong \frac{H^1(M;\mathbb{R})}{H^1(M;\mathbb{Z})} \times \operatorname{im} d^*.$$

  The 'Picard torus' $P = H^1(M;\mathbb{R})/H^1(M;\mathbb{Z})$ is diffeomorphic to $(S^1)^{b_1(X)}$, while $\operatorname{im} d^*$ is a Fréchet space.
- The curvature is gauge-invariant (since $\mathsf{U}(1)$ is abelian), and so defines an affine-linear map

  $$iF\colon \mathcal{S} \to iF_\nabla + \operatorname{im} d \subset \Omega^2_M, \quad iF_{\nabla+ia} = iF_\nabla - da.$$

  Suppose we are interested in gauge-orbits connections with the *same* curvature as ∇. These form the subset

  $$P \times \{0\} \subset P \times \operatorname{im} d^*,$$

  since $\operatorname{im} d^* \cap \ker d = 0$. Thus *the gauge-orbits of* $\mathsf{U}(1)$ *connections of prescribed curvature (any representative of* $-2\pi i c_1(L)$*) form an affine copy of the Picard torus.* When $b_1(M) = 0$, the gauge-orbit is a single point.

11.2 $\mathsf{U}(1)$-instantons

Definition 11.1 Let X be a 4-manifold equipped with a conformal structure. A Yang–Mills *instanton*, or *anti-self-dual connection*, in the hermitian vector bundle $E \to X$, is a connection $\nabla \in \mathcal{A}_E$ such that

$$(F_\nabla)^+ = 0.$$

Here $(\cdot)^+$ is the self-dual projection $\frac{1}{2}(1+\star)$, mapping $\Omega^2(\mathfrak{u}(E))$ to $\Omega^+_g(\mathfrak{u}(E))$.

One has

$$(u^* F_\nabla)^+ = (u F_\nabla u^{-1})^+ = u F^+_\nabla u^{-1},$$

so $\mathcal{G}_E$ preserves the instantons.

The equation $(F_{\nabla+A})^+ = 0$ amounts to a first-order differential equation for A. Donaldson theory is the study of this equation; the focus is largely on the case of rank 2 bundles.

Our purpose here is to understand a simple case, that of instantons in line bundles $L \to X$, where X is compact.

Criterion for existence. The form $(i/2\pi)F_\nabla$ represents the integral cohomology class $c_1(L)$. If ∇ is an instanton then $(i/2\pi)F_\nabla$ is both closed and ASD, hence harmonic. Thus

$$c_1(L) \in \mathcal{H}^-_{[g]}(\mathbb{Z}) := \mathcal{H}^-_g \cap H^2(X;\mathbb{Z})'.$$

(Here $H^2(X;\mathbb{Z})' \subset H^2_{DR}(X)$ denotes the lattice of integer classes—a copy of $H^2(X;\mathbb{Z})/\mathsf{torsion}$.) Conversely, if $c_1(L) \in \mathcal{H}^-_{[g]}(\mathbb{Z})$ then, letting ω be the harmonic representative of $c_1(L)$—so $\omega^+ = 0$—one can find a connection ∇ with $F_\nabla = -2\pi i\omega$, and this is an instanton.

12 Instantons in $\mathsf{U}(1)$-bundles

12.1 Instantons in $\mathsf{U}(1)$-bundles

In lecture 11, we introduced the notion of an *instanton*, a unitary connection ∇ over an oriented Riemannian 4-manifold (X,g) with a conformal structure $[g]$, such that $F^+_\nabla = 0$. We began the study of instantons in line bundles $L \to X$. We take X to be compact.

So far we have given a criterion for existence of an instanton in L, namely, that $c_1(L) \in H^2_{DR}(X)$ lies in the intersection $\mathcal{H}^-_g(\mathbb{Z})$ of the integer lattice with the classes of closed, ASD 2-forms. Next we discuss uniqueness.

Suppose first that ∇ is an instanton in L. Then another connection $\nabla + ia$ is an instanton if and only if

$$a \in \ker(d^+ \colon \Omega^1_X \to \Omega^+_g).$$

Recall the self-duality complex $\mathcal{E}^*$,

$$0 \to \Omega^0_X \xrightarrow{d} \Omega^1_X \xrightarrow{d^+} \Omega^+_g \to 0.$$

We computed its cohomology, showing that

$$H^1(\mathcal{E}) = \frac{\ker d^+}{d\Omega^0_X} = \frac{\ker d}{d\Omega^0_X} = H^1_{DR}(X), \quad H^2(\mathcal{E}) = \frac{\Omega^+_g}{d\Omega^1_X} \cong \mathcal{H}^+_g.$$

Recall that $\mathcal{A}_L/\mathcal{G}^\circ_L$ is identified with $\Omega^1_X/d\Omega^0_X$, by mapping $[\nabla + ia]$ to $[a]$.

Let $\mathcal{I}_L = \nabla + \ker d^* \subset \mathcal{A}_L$ be the subspace of instantons. Then we see

$$\mathcal{I}_L/\mathcal{G}^\circ_L \cong H^1_{DR}(X).$$

The component group $\pi_0\mathcal{G}_L$ then acts, with quotient

$$\mathcal{I}_L/\mathcal{G}_L \cong P = \frac{H^1_{DR}(X)}{2\pi H^1_{DR}(X;\mathbb{Z})},$$

so the moduli space of instantons modulo gauge is the Picard torus.

One sees from this description that $\mathcal{I}_L/\mathcal{G}_L$ is a manifold, with tangent space $H^1_{DR}(X)$. It is worthwhile thinking through why, intrinsically, this is so. Let's work in the slice $\mathcal{S} = \nabla + \ker d^*$. We have $\mathcal{I}_L \cap \mathcal{S} = \nabla + \ker(d^* \oplus d^+)$. Let us look at how the self-duality complex, and the operator d^*, appear with respect to the Hodge decomposition:

$$\begin{tikzcd}
\mathcal{H}^0_g & \mathcal{H}^1_g & \mathcal{H}^{[g]}_+ \\
d^*\Omega^1_X \arrow[r, shift left, "d"] & d\Omega^0_X \arrow[l, shift left, "d^*"] & \\
 & d^*\Omega^2_X \arrow[r, "d^+"] & d^+\Omega^1_X
\end{tikzcd}$$

Note that $d^*\Omega^1_X$ is the space Ω^0_0 of functions f of mean zero $\int_X f \,\mathsf{vol} = 0$. The map $d^* \colon \Omega^1_X \to \Omega^0_0$ is an isomorphism, as is $d^+ \colon d^*\Omega^2_X \to d^+\Omega^1_X$. Thus we view $d^* \oplus d^+$ as an operator

$$d^* \oplus d^+ \colon \Omega^1 \to \Omega^0_0 \oplus \Omega^+_g,$$

and see that

$$\ker(d^* \oplus d^+) = \mathcal{H}^1_{[g]}, \quad \operatorname{coker}(d^* \oplus d^+) = \mathcal{H}^+_g.$$

Thus the instanton moduli space $\mathcal{I}_L.\mathcal{G}_L = (\mathcal{I}_L \cap \mathcal{S})/\pi_0\mathcal{G}_L$ is cut out as the zero-set of a smooth map which is not a submersion (unless $b^+(X) = 0$), but whose derivative has finite-dimensional cokernel of constant rank $b^+(X)$, and has tangent space $\mathcal{H}^1_{[g]}$.

In finite dimensions, clean level sets of smooth maps are submanifolds, by virtue of the inverse function theorem. In infinite dimensions, the inverse function theorem is available provided one works with Banach manifolds. For this reason (among others) it is customary to set up the problem we have been discussing in the framework of Sobolev spaces (and Hilbert manifolds). We will return to this point later.

12.1.1 Generic non-existence

$\mathcal{H}^-_{[g]}$ is the intersection of a fixed lattice in $H^2_{DR}(X)$ with a $b^-(X)$-dimensional subspace—a subspace which one might expect to be 'random', inasmuch as it varies with the conformal structure. Thus, provided that the subspace has positive codimension, one expects that the intersection will typically be zero.

A precise result on these lines is as follows:

Theorem 12.1 *For $k < b^+(X)$, and for any family of conformal structures $\{[g_t]\}_{t\in T}$ parameterized by a smooth compact k-manifold T, there are perturbations $[\hat{g}_t]$ to this family, arbitrarily close to $[g_t]$ in the C^r-norm induced by metrics on T and X, such that*

$$\mathcal{H}^-_{[\hat{g}_t]}(\mathbb{Z}) = 0 \quad \textit{for all } t \in T.$$

Recall that we regard the space $\mathcal{C}_X$ of conformal structures (of class C^r for some $r \geq 3$) as a Banach manifold by picking a reference conformal structure $[g_0]$ and identifying $\mathcal{C}_X$ it with an open subset of the bundle maps $m\colon \Lambda^- \to \Lambda^+$. The correspondence takes m to the unique $[g] \in \mathcal{C}_X$ for which $\Lambda^-[g] = \Gamma_m$.

The derivative of P is a linear map

$$D_{[g]}P\colon C^r(X, \mathsf{Hom}(\Lambda^-_{[g]}, \Lambda^+_{[g]})) \to \mathsf{Hom}(\mathcal{H}^-_{[g]}, \mathcal{H}^+_{[g]}),$$

and we found that

$$(D_{[g]}P)(m)(\alpha_-) = m(\alpha_-)_{\mathsf{harm}}.$$

It is this result that is key to the proof.

Any non-zero class $c \in H^2_{DR}(X)$ defines a closed subset

$$S_c = \{H \in \mathsf{Gr}^- : c \in H\} \subset \mathsf{Gr}^-.$$

If $H \in S_c$, H' is a subspace complementing to H, and $m\colon H \to H'$ a linear map, then the graph Γ_m lies in S_c if and only if $c \in \ker m$. Hence S_c is a submanifold with tangent space

$$T_H S_c = \{\mu \in \mathsf{Hom}(H, H') : \mu(c) = 0\}$$

and normal space

$$(N_{S_c})_H = T_H \mathsf{Gr}^- / T_H S_c \xrightarrow{\cong} H', \quad [\mu] \mapsto \mu(c),$$

so the codimension of S_c is $b^+(X)$.

Lemma 12.2 *P is transverse to the submanifold $S_c \subset \mathsf{Gr}^-$.*

Proof We must prove is that if $P[g] \in S_c$ then $\operatorname{im} D_{[g]}P$ spans the normal space to S_c in Gr^- at the point $\mathcal{H}^-[g]$. Unravelling the definitions, and applying our formula for DP, this means that *if $\alpha^- \in \mathcal{H}^-_{[g]}$ represents the class c then, for any $\alpha^+ \in \mathcal{H}^+_{[g]}$, one can find a bundle map $m\colon \Lambda^- \to \Lambda^+$ such that*

$$m(\alpha^-)_{\mathsf{harm}} = \alpha^+.$$

If this were *not* the case, there would be forms $\alpha^\pm \in \mathcal{H}^\pm_{[g]}$ such that α^+ which is L^2-orthogonal to $m(\alpha^-)_{\mathsf{harm}}$, for any m:

$$0 = \langle \alpha^+, m(\alpha^-)_{\mathsf{harm}} \rangle_{L^2} = \langle \alpha^+, m(\alpha^-) \rangle_{L^2} \quad \forall m.$$

Suppose then that there is some point $x \in X$ where $\alpha^+(x) \neq 0$ and $\alpha^-(x) \neq 0$. In a small geodesic ball B around x, we can choose m_0 so that $m_0(\alpha^-)|_B = \alpha^+|_B$. Choose a cutoff function χ, supported in B, non-negative and identically 1 on $\frac{1}{2}B$. Set $m = \chi m_0$ (a bundle map defined on X but supported in B). Then

$$\langle \alpha^+, m(\alpha^-) \rangle_{L^2} = \int_B \chi |\alpha^+|^2 \mathsf{vol}_g > 0,$$

contradiction. Certainly α^- and α^+ are non-zero, and so each of them must vanish on some open set.

The lemma now follows from the following unique continuation principle: namely, *a harmonic form vanishing on an open set vanishes everywhere*. This is an instance of a general unique continuation principle for elliptic equations, due to N. Aronszajn (*A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order*, 1957). □

Proof of the theorem. It follows from the lemma that, for any non-zero $c \in H^2_{DR}(X)$, $P^{-1}(S_c) \subset \mathcal{C}_X$ is a submanifold of codimension $b^+(X)$. This is by an application of the inverse function theorem, which is valid for smooth maps between Banach spaces. It is here that is becomes significant that we choose to work with C^r conformal structures.

Now take our manifold T of dimension $k < b^+(X)$ and C^r family of conformal structures g_t, $t \in T$. This family defines a smooth map $G\colon T \to \mathcal{C}_X$. Consider the space $\mathcal{G}$ of maps $T \to \mathcal{C}_X$ which lie within a fixed distance of G with respect to the C^r metric induced by Riemannian metrics on T and on X. For any non-zero class c, there is an open, dense set $U_c \subset \mathcal{G}$ of maps $\hat{G}$, transverse to $P^{-1}(S_c)$. Since $P^{-1}(S_c)$ has codimension $b^+(X) > \dim T$, this means that $\operatorname{im} \hat{G}$ misses S_c.

Let c vary over all classes in the integer lattice $H^2(X;\mathbb{Z})'$ with $c \cdot c < 0$ to obtain a countable intersection $\bigcap U_c$ of open dense subsets, representing maps $\hat{G}$ that miss all integer classes. By the Baire category theorem, applied in the complete metric space $\mathcal{G}$, this intersection is dense. □

13 Differential operators

13.1 First-order differential operators

Let E and F be real vector bundles over M. Let $\mathcal{D}_0(E,F)$ be the vector space of zeroth-order differential operators, that is, the $C^\infty(M)$-linear maps $\Gamma(M,E) \to \Gamma(M,F)$. There is a 'symbol' isomorphism

$$\sigma^0 \colon \mathcal{D}_0(E,F) \to \Gamma(M, \mathsf{Hom}(E,F)),$$

mapping the operator L to the well-defined bundle map $e \mapsto (Ls)_x$, where $s \in \Gamma(M,E)$ is any section such that $s(x) = e$.

Definition 13.1 A *first-order linear differential operator* from E to F is an $\mathbb{R}$-linear map $D \colon \Gamma(M,E) \to \Gamma(M,F)$ such that, for all functions $f \in C^\infty(M)$, the commutator $[D,f]$ is $C^\infty(M)$-linear.

The first-order linear operators form a vector space $\mathcal{D}_1(E,F) \supset \mathcal{D}_0(E,F)$.

Example 13.2 A covariant derivative $\nabla \colon \Gamma(M,E) \to \Gamma(T^*M \otimes E)$ is a first-order linear differential operator. For any f, one has $[\nabla, f]s = df \otimes s$.

The definition can be recast in terms of *jet bundles.*

Define J^1E to be the vector bundle of 1*-jets of sections of* E. A point of $(J^1E)_x$ is a pair $(x,[s])$, with $x \in M$, s the germ near x of a section of E, and $[s]$ is its 1-jet: $[s_1] = [s_2]$, when (i) $(s_1 - s_2)(x) = 0$ and (ii) $s_1 - s_2$ is tangent to 0 at x (that is: the derivative $D_x(s_1 - s_2) \colon T_xM \to T_{x,0}E$ maps to the zero-section $T_xM \subset T_{x,0}E$).

There is a short exact sequence

$$0 \to T^*M \otimes E \to J^1E \xrightarrow{ev} E \to 0.$$

Any section $s \in \Gamma(M,E)$ defines a section $j^1s \in \Gamma(M, J^1E)$. For functions f, one has the Leibniz rule

$$f(j^1s) - j^1(fs) = df(x) \otimes s(x) \in T_x^*M \otimes E_x \subset J_x^1E.$$

Definition 13.3 A *first-order jet-operator* is a map $D \colon \Gamma(M,E) \to \Gamma(M,F)$ of the form

$$(Ds)(x) = L(j_x^1 s), \quad L \in \Gamma(M, \mathsf{Hom}(J^1E,F)).$$

These form a vector space $\mathcal{D}_1(E,F)_{jet} \cong \Gamma(M, \mathsf{Hom}(J^1E,F))$.

In local coordinates $(x_1,\dots,x_n)$, and in local trivializations of E and F, first-order jet operators take the form

$$(Ds)_\alpha = \sum_{\beta,i} P^i_{\alpha\beta}(x)\frac{\partial}{\partial x_i}s_\beta + \sum_\beta Q_{\alpha\beta}(x)s_\beta. \tag{3}$$

The short exact sequence for J^1E induces a short exact sequence of vector spaces

$$0 \to \Gamma(M, \mathsf{Hom}(E,F)) \to \mathcal{D}_{jet}(E,F)_1 \xrightarrow{\mathsf{symb}} \Gamma(M, \mathsf{Hom}(T^*X \otimes E, F) \to 0.$$

The map symb, the 'principal symbol', is simply restriction to $T^*M \otimes E \subset J^1(E)$.

Theorem 13.4 *First-order jet operators are identical to first-order differential operators.*

Lemma 13.5 $\mathcal{D}_1(E,F)_{jet} \subset \mathcal{D}_1(E,F)$.

Proof Suppose that $D = L \circ j^1 \in \mathcal{D}_1(E,F)_{jet}$. Then D is $\mathbb{R}$-linear (since L and j^1 are); and, for functions f and g, one has

$$[D,f]s = L(df \otimes s)$$

and

$$\begin{aligned}[D,f](gs) &= L \circ j^1(fgs)) - f\,L(j^1(gs)) \\ &= L \circ \big(df \otimes gs + f \circ j^1(gs)\big) - f\,L(j^1(gs)) \\ &= sL \circ (df \otimes s) \\ &= g[D,f](s),\end{aligned}$$

i.e. $[[D,f],g] = 0$, as required. □

Our aim is now to prove the reverse inclusion, $\mathcal{D}_1(E,F) \subset \mathcal{D}_1(E,F)_{jet}$. Given $D \in \mathcal{D}(E,F)_1$, define

$$\sigma_D^1(f) = \sigma^0_{[f,D]} \in \Gamma(M, \mathsf{Hom}(E,F)).$$

Note that $[D,fg] = f[D,g] + [D,f]g$, from which it follows that

$$\sigma_D(fg) = f\,\sigma_D(g) + g\,\sigma_D(f).$$

Lemma 13.6 *Let $D \in \mathcal{D}(E,F)_1$, and suppose that $f(x) = 0$ and $df_x = 0$. Then $\sigma_D^1(f)_x = 0$.*

Proof Note that $f \in \mathfrak{m}_x^2$, where $\mathfrak{m}_x \subset C^\infty(M)$ is the maximal ideal of functions vanishing at x. So we can write $f = \sum_i g_i h_i$, where $g_i(x) = h_i(x) = 0$. But then $\sigma_D^1(f) = \sum g_i \sigma_D^1(h_i) + h_i \sigma_D^1(g)i)$, so $\sigma_D^1(f)_x = 0$. □

Also, $\sigma_D^1(c)_x = 0$ when c is a function that is constant near x. Thus $\sigma_D^1(f)$ really only depends on $df(x) \in T^*x_X$ (recall that $df(x)$ stands for the class of $f - f(x)$ in $\mathfrak{m}_x/\mathfrak{m}_x^2$). In view of this observation, we shall from now on write

$$\sigma_D^1(x;\xi) \in \mathsf{Hom}(E_x, F_x), \quad \xi \in T_x^*X,$$

to mean $\sigma_D^1(f)_x$ where $df(x) = \xi$.

The bundle map $\sigma_D^1 \colon T^*X \to \mathsf{Hom}(E,F)$ is called the *principal symbol* of the first-order differential operator D. (There is no clash here: when D is a jet operator, $\sigma_D^1 = \mathsf{symb}_D$.)

Note that $J^1E = J_M^1 \otimes E$, where $J_M^1 = J^1\underline{\mathbb{R}}$ is the bundle of 1-jets of functions; and that $J_M^1 = T^*M \oplus \underline{\mathbb{R}}$ (the second summand is given by the constant functions). Thus From the lemma, it follows that $\sigma_D(f)_x$ depends only on the 1-jet $j_x^1(f) \in (J_M^1)_x$.

Bearing in mind that $\mathsf{Hom}(J^1E,F) \cong \mathsf{Hom}(J_M^1, \mathsf{Hom}(E,F))$, we can think of σ_D^1 as defining a bundle map $J^1E \to F$, and hence a jet operator $\sigma_D^1 \circ j^1$. The following lemma is then a tautology:

Lemma 13.7 *$D - \sigma_D^1 \circ j^1$ is a zeroth order operator.*

But since zeroth order operators are jet operators, this shows that $\mathcal{D}_1(E,F) \subset \mathcal{D}_1(E,F)_{jet}$, completing the proof of the theorem.

13.2 Higher-order operators

One can recursively define an nth-order linear differential operator $D\colon \Gamma(M,E) \to \Gamma(M,F)$ to be an $\mathbb{R}$-linear map such that, for all functions f, $[D,f]$ is an $(n-1)$th order linear differential operator. These form a vector space $\mathcal{D}_n(E,F)$. Their union $\mathcal{D}(E,F)$ is a filtered vector space.

One can also define the bundle J^nE of n-jets of sections of E (in which sections of E are considered modulo the relation of having the same nth order Taylor expansion); and the map $j^n\colon \Gamma(M,E) \to \Gamma(M,J^nE)$ taking a section to its n-jet at each point. One defines $\mathcal{D}_n(E,F)_{jet}$ as the space of operators $s \mapsto L(j^n s)$ for $L \in \Gamma(M, \mathsf{Hom}(J^nE,F))$, and shows, much as before, that $\mathcal{D}_n(E,F) = \mathcal{D}_n(E,F)_{jet}$.

Truncation of jets defines a short exact sequence

$$0 \to \mathrm{Sym}^n(T^*X) \otimes E \to J^nE \to J^{n-1}E \to 0,$$

and restricting nth order jet operators to $\mathrm{Sym}^n\, T^*X \otimes E$ defines the *symbol isomorphism*

$$\sigma^n\colon \frac{\mathcal{D}_n(E,F)}{\mathcal{D}_{n-1}(E,F)} \to \Gamma(M, \mathsf{Hom}(\mathrm{Sym}^n(T^*X) \otimes E, F)).$$

The principal symbol is compatible with composition: if $D\colon \Gamma(E) \to \Gamma(E")$ has order m and, $D'\colon \Gamma(E') \to \Gamma(E'')$ has order m', then $D \circ D'$ is an $(m+m')$th order operator and

$$\sigma^{m+m'}_{D'\circ D} = \sigma^{m'}_{D'} \circ \sigma^m_D.$$

The composition on the right is the tensor product of two operations: multiplication in the graded ring $\mathrm{Sym}^\bullet(T^*X)$ of homogeneous polynomial functions on T^*_xX, and composition $\mathsf{Hom}(E',E'') \times \mathsf{Hom}(E,E') \to \mathsf{Hom}(E,E'')$. Thus, taking $E = F$, the principal symbol defines a graded ring isomorphism

$$\sigma^\bullet\colon \mathsf{gr}^\bullet\mathcal{D}(E,E) \to \Gamma\left(M, \mathrm{Sym}^\bullet(T^*X) \otimes \mathrm{End}\, E\right).$$

13.3 Examples of symbols

Example 13.8 For the exterior derivative $d\colon \Omega^k_X \to \Omega^{k+1}_X$, one has $[d,f]s = df \wedge s$. Therefore

$$\sigma^1_d(\xi) = \xi \wedge \cdot\colon \Lambda^k_X \to \Lambda^{k+1}_X.$$

Example 13.9 For a coupled exterior derivative $d_\nabla\colon \Omega^k_X(E) \to \Omega^{k+1}_X(E)$, one again has $[d_\nabla,f]\omega = df \wedge \omega$, so

$$\sigma^1_{d_\nabla}(\xi) = \xi \wedge \cdot\colon \Lambda^k_X \otimes E \to \Lambda^{k+1}_X \otimes E.$$

Note that $\sigma^1_{d_\nabla}(\xi) \circ \sigma^1_d(\xi) = 0$, consistent with the fact that $d^2_\nabla = F_\nabla \wedge \cdot$ is a zeroth-order operator, so has vanishing principal symbol.

Example 13.10 Formal adjoints. Suppose that M is compact and that E and F have euclidean metrics. Let $D^*\colon \Gamma(F) \to \Gamma(E)$ be a first-order linear differential operator that is formally adjoint to $D\colon \Gamma(E) \to \Gamma(F)$:

$$\langle t, Ds\rangle_{L^2(F)} = \langle D^*t, s\rangle_{L^2(E)}.$$

Then, for functions f, one has

$$\langle t, [f,D]s\rangle_{L^2(F)} = \langle [D^*,f], s\rangle_{L^2(E)},$$

so $[f,D]^* = [D^*,f]$ and

$$\sigma^1_{D^*}(\xi) = -(\sigma^1_D(\xi))^*.$$

Example 13.11 $d^*\colon \Omega^{k+1}_X \to \Omega^k_X$ has symbol given by the negative of contraction,

$$\sigma^1_{d^*}(\xi) = -\iota_\xi \wedge -\colon \Lambda^{k+1}_X \to \Lambda^k_X.$$

(check using bases). The formal adjoint to the coupled exterior derivative, $d^*_\nabla = \pm \star d_\nabla \star\colon \Omega^{k+1}(E) \to \Omega^k(E)$, has symbol

$$\sigma^1_{d^*}(\xi) = -\iota_\xi \wedge -\colon \Lambda^{k+1}_X(E) \to \Lambda^k_X(E).$$

Example 13.12 The Hodge Laplacian $\Delta = dd^* + d^*d$ has symbol

$$\sigma^2_\Delta(\xi,\xi)(\alpha) = -\iota_\xi(\xi \wedge \alpha) - \xi \wedge (\iota_\xi \alpha) = -|\xi|^2\alpha.$$

The covariant Laplacian acting on $\Omega^k(E)$,

$$\Delta_\nabla := d^*_\nabla d_\nabla + d_\nabla d^*_\nabla$$

likewise has symbol

$$\sigma^2_\Delta(\xi,\xi) = -|\xi|^2 \mathrm{id}.$$

13.4 Elliptic operators

Definition 13.13 (i) An *elliptic operator* is a linear differential operator $D \in \mathcal{D}_n(E,F)$ such that for all $x \in M$ and all non-zero $\xi = df(x) \in T^*_x M$, the symbol map

$$\sigma^n_D(\xi,\dots,\xi) = \frac{1}{n!}[\dots[[D,f],f],\dots,f]_x \in \mathsf{Hom}(E_x,F_x)$$

is an isomorphism.
(ii) A *generalized Laplacian* is an operator $\Delta \in \mathcal{D}_2(E,E)$ such that $\sigma^2_\Delta(\xi,\xi) = -|\xi|^2 \mathrm{id}_E$. Equivalently, one has $\frac{1}{2}[[\Delta,f],f] = -|df|^2$.
(iii) A *Dirac operator* is an operator $D \in \mathcal{D}_1(E,E)$ such that D^2 is a generalized Laplacian. Equivalently, $\sigma_D(\xi)^2 = -|\xi|^2 \mathrm{id}_E$.

Generalized Laplacians ∇ are evidently elliptic. While ∇ may not be formally self-adjoint, the fact that the symbol is self-adjoint implies that $\Delta - \Delta^* \in \mathcal{D}_1(E,E)$.
Likewise, Dirac D operators are elliptic, with $D - D^* \in \mathcal{D}_0(E,E)$.

Example 13.14 On a Riemannian 4-manifold X, the operator $d^* \oplus d^+\colon \Omega^1 \to \Omega^0 \oplus \Omega^+_g$ has symbol

$$\sigma^1(\xi)\colon \Lambda^1 \to \Lambda^0 \oplus \Lambda^+, \quad a \mapsto -\iota_\xi a + (\xi \wedge a)^+.$$

This map takes $\mathbb{R}\xi$ isomorphically to Λ^0 and $\xi^\perp$ isomorphically to Λ^+, hence is an isomorphism. So $d^* \oplus d^+$ is elliptic.

Example 13.15 The Laplacian on Ω^k_M and the covariant Laplacian on $\Omega^k_M(E)$ are generalized Laplacians.

Example 13.16 $d^* \oplus d\colon \Omega^*_X \to \Omega^*_X$ is a Dirac operator, since its square is the Laplacian Δ. Similarly, $d^*_\nabla \oplus d_\nabla\colon \Omega^*_X(E) \to \Omega^*_X(E)$ is a Dirac operator.

The symbol σ_D of a Dirac operator $D \in \mathcal{D}_1(E,E)$ squares to the symbol of the generalized Laplacian D^2, and therefore satisfies the *Clifford relation*

$$\sigma^1_D(\xi)^2 = -|\xi|^2.$$

The cotangent space T^*X_x, with its inner product g, defines a Clifford algebra

$$\mathsf{cl}(T^*X_x),$$

the associative algebra with unit $\mathbf{1}$ generated by T^*X_x and subject to the relation $\xi \cdot \xi = -|\xi|^2\mathbf{1}$. The symbol $\sigma = \sigma^1_D$ extends to a representation of the Clifford algebra

$$\sigma\colon \mathsf{cl}(T^*X_x) \to \operatorname{End} E_x.$$

14 Analysis of elliptic operators

14.1 Fredholm operators

Let B_1 and B_2 be separable Banach spaces, and $\mathcal{L}(B_1, B_2)$ the Banach space of bounded (equivalently, continuous) linear maps $T\colon B_1 \to B_2$ with the operator norm. If $T \in \mathcal{L}(B_1, B_2)$ is bijective, then it is invertible. Moreover, the invertible operators form an open set in $\mathcal{L}(B_1, B_2)$.

Definition 14.1 a *Fredholm operator* from B_1 to B_2 is a bounded linear map $T\colon B_1 \to B_2$ with finite-dimensional kernel and cokernel. The index of a Fredholm operator is

$$\operatorname{ind} T = \dim\ker T - \dim\operatorname{coker} T.$$

To say that T has finite-dimensional cokernel is to say that there is a finite-dimensional space $F \subset B_2$ with $F + \operatorname{im} T = B_2$. In this case, $\operatorname{im} T$ admits a finite-dimensional complement C.

Lemma 14.2 *If T is Fredholm then its image is closed.*

Proof Let C be a complement to $\operatorname{im} T$. Being finite-dimensional, C is closed. Consider the map

$$T'\colon (B_1/\ker T) \oplus C \to B_2, \quad T'([x], c) = Tx + c.$$

T' is bounded and bijective, hence an isomorphism of Banach spaces. Therefore it takes closed sets to closed sets. But $T'(B_1/\ker T \oplus 0) = T(B_1)$. □

Proposition 14.3 *$T \in \mathcal{L}(B_1, B_2)$ is Fredholm if and only if it is invertible modulo compact operators: that is, there exists $P \in \mathcal{L}(B_2, B_1)$ such that $TP - I$ and $PT - I$ are compact.*

Proof If T is Fredholm, we can write $B_1 = \ker T \oplus B_1'$ and $B_2 = \operatorname{im} T \oplus C$. The map $T' = T|_{B_1'}\colon B_1 \to \operatorname{im} T$ is a bounded bijection, hence an isomorphism. Define $P\colon \operatorname{im} T \oplus C \to B_1$ by $P(y, c) = T'^{-1}(y)$. Then $(TP - I)(y, c) = -c$, and $(PT - I)(k, x') = -k$, so both $TP - I$ and $PT - I$ have finite rank and are therefore compact.

Conversely, if $PT - I = K$ is compact, and (x_i) is a bounded sequence in $\ker T$, then $0 = PTx_i = (I + K)x_i$, and (passing to a subsequence) $Kx_i \to y$ (say); so $x_i \to -y$, and $Ty = 0$. So $\ker T$ is finite-dimensional. If $TP - I = K'$ is compact, then $\operatorname{im} P$ contains $\operatorname{im}(I + K')$, so it suffices to show that $I + K'\colon B_2 \to B_2$ has finite-dimensional cokernel. Now, K' is the norm-limit of finite rank operators; so we can write $I + K' = J + F$ where J is invertible and F has finite rank. Rewrite this again as $J + F = J^{-1}(I + J \circ F)$ to see that it suffices to treat the case where K' has finite rank. But $\operatorname{im}(I + K')$ contains the finite-codimensional space $\ker K'$, so $I + K'$ indeed has finite codimension. □

Proposition 14.4 *The following are equivalent for $T \in \mathcal{L}(B_1, B_2)$:*

(i) *T is Fredholm of index zero.*

(ii) *$T = J + F$ where J is invertible and F has finite rank.*

(iii) *$T = J + K$ where J is invertible and K is compact.*

Proof (i) $\Rightarrow$ (ii). Since the kernel and cokernel of T have the same dimension, we can find a map $f\colon \ker T \to B_2$ whose image represents $\operatorname{coker} T$. Since $\ker T$ is finite-dimensional, there exists a projection $\pi\colon B_1 \to \ker T$. Set $F = -f \circ \pi$. Then F has finite rank and $T - F$ is invertible.

(ii) $\Rightarrow$ (i): Take $T = J + F$ with J invertible, F finite rank. Then J^{-1} is an inverse modulo compact operators, so by the previous proposition, T is Fredholm. To show that T has index zero, we can reduce (as in the foregoing proof) to the case $T = I + F\colon B \to B$. Let A be a (finite-dimensional) complement to $\ker F$, and let $A' = A + \operatorname{im} F$. Let $B' \subset B$ be a complement to the finite-dimensional subspace A'. Then $I + F$ sends A' to itself and B' to itself—in the latter case, as the identity map. Thus we have $B = B' \oplus A'$ with A' finite-dimensional and $I + F = \mathrm{id}_B \oplus f$, for some $f \in \mathcal{L}(A', A')$. But the the kernel and cokernel of $I + F$ are the kernel and cokernel of f, hence have the same dimension. Thus the index is zero.

(ii) $\Rightarrow$ (iii): Finite rank implies compact.

(iii) $\Rightarrow$ (ii) There is some $\epsilon > 0$ such that $J + L$ remains invertible whenever $\|L\| < \epsilon$. But K is a norm-limit of finite rank operators, so one can write $K = F + L$ with F finite-rank and $\|L\| < \epsilon$. Then $T = F + (J + L)$. □

From the proposition we deduce:

Corollary 14.5 *The Fredholm operators $\mathcal{F}(B_1, B_2)$ form an open set in $\mathcal{B}(B_1, B_2)$, and the index*

$$\operatorname{ind}\colon \mathcal{F}(B_1, B_2) \to \mathbb{Z}$$

is a locally constant (equally, continuous) function.

14.2 Sobolev spaces and elliptic estimates

There are many references for Sobolev spaces and elliptic estimates; an appropriate one is L. Nicolaescu, *Lectures on the geometry of manifolds*.

14.2.1 Sobolev spaces

Let U be an open subset of $\mathbb{R}^n$. We consider the space $C_c^\infty(U; \mathbb{R}^r)$ of compactly supported $\mathbb{R}^r$-valued functions. Fix a real number $p > 1$ and an integer $k \geq 0$. We have the L^p norm

$$\|f\|_p = \left(\int_U |f|^p \mathrm{vol} \right)^{1/p},$$

and the Sobolev L^p_k norm

$$\|f\|_{p,k} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_p.$$

Here α is a multi-index $(i_1, \ldots, i_k)$ of size $k = |\alpha|$, where $1 \leq i_1 \cdots \leq i_k \leq n$, and $D^\alpha = \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}}$.

There is also the C^k norm,

$$\|f\|_{C^k} = \sum_{|\alpha| \leq k} \sup_U |D^\alpha f|.$$

Suppose now that M is a compact, oriented n-manifold, and $E \to M$ a real vector bundle of rank r. Choose a finite open covering $\{U_i\}$ for M, smooth embeddings $U_i \subset \mathbb{R}^n$, a partition of unity $\{\rho_i\}$

subordinate to the cover, and trivializations $\tau_i \colon E|_{U_i} \to \mathbb{R}^r$. We can then define a Sobolev norm on $\Gamma(M,E)$ by

$$\|s\|_{p,k} = \sum_i \|\rho_i(\tau_i \circ s)\|_{p,k}$$

defined using the coordinates on U_i. The Sobolev space $L^p_k(E)$ is defined to be the completion of $\Gamma(E)$ in the L^p_k norm. It is a Banach space; when $p = 2$, a Hilbert space.

While the Sobolev norm depends on various choices, different choices give rise to equivalent norms and so to identical Sobolev spaces. There is an alternative, and equivalent, approach based on choices of metrics in TX and E, and of a connection ∇ in E.

We can also define a C^k norm, on $C^k(M,E)$:

$$\|s\|_{C^k} = \sum_i \|\rho_i(\tau_i \circ s)\|_{C^k}.$$

It follows from the Arzela–Ascoli theorem, and the mean-value inequality in $\mathbb{R}^n$, that the inclusion $C^{k+1} \to C^k$ is *compact*: every bounded sequence (s_n) in C^{l+1} has subsequence that converges in C^l.

The basic facts about Sobolev spaces are these:

- **Sobolev inequality:** Define the scaling weight in n dimensions as
$$w(k,p) = k - \frac{n}{p}.$$
This is the weight with which, on $\mathbb{R}^n$, $\|D^\alpha f\|_p$ transforms under a dlilation $x \mapsto rx$ when $|\alpha| = k$. If
$$k > l \quad \text{and} \quad w(k,p) \geq w(l,q)$$
then there is a bounded inclusion of Sobolev spaces
$$L^p_k \to L^q_l.$$
- **Rellich lemma:** If $k > l$ and $w(k,p) < w(l,q)$, the inclusion $L^p_k \to L^q_l$ is compact. In particular, the inclusion
$$L^p_{k+1} \to L^p_k$$
is compact.
- **Morrey inequality:** Suppose $l \geq 0$ is an integer such that
$$l < w(p,k).$$
There is then a constant C such that $\|\cdot\|_{p,k} \leq C\|\cdot\|_{C^l}$, and therefore a bounded inclusion
$$L^p_k(E) \to C^l(E).$$
- **Smoothness:** One has
$$\bigcap_{k \geq k_0} L^p_k(E) = C^\infty(E).$$
(Indeed, $C^\infty(E)$ certainly lies in the intersection of Sobolev spaces, and is dense therein since it is dense in any one L^p_k; but a Cauchy sequence in $\bigcap_{k\geq k_0} L^p_k(E)$ has bounded derivatives of all orders, so its limit is C^∞.)

14.3 Elliptic estimates

Let $D\colon \Gamma(M,E)\to\Gamma(M,F)$ be a differential operator of order m over a closed, oriented, Riemannian manifold (M,g). The basic point is that D extends to a bounded linear map between Hilbert spaces:

$$D\colon L^2_{k+m}(E)\to L^2_k(F).$$

Theorem 14.6 (elliptic estimate) *If D is elliptic of order m, one has estimates on the L^2_k Sobolev norms for each $k\geq 0$:*

$$\|s\|_{2,k+m}\leq C_k\left(\|Ds\|_{2,k}+\|s\|_k\right).$$

Moreover

$$\|s\|_{2,k+m}\leq C_k\|Ds\|_{2,k}$$

for $s\in(\ker D)^\perp$ (L^2 orthogonal complement).

We will not prove the theorem, but we will note some important consequences (all of the first inequality, not the second):

Theorem 14.7 *If D is elliptic of order m, with formal adjoint D^*, then*

- *(i)* *Weyl's lemma holds: If $s\in\ker D\subset L^2_m(E)$, then s is C^∞.*
- *(ii)* *The unit ball in $\ker D\subset L^2_m(E)$ is compact, and hence $\ker D$ is finite-dimensional.*
- *(iii)* $\operatorname{im} D$ *is closed in $L^2(F)$.*
- *(iv)* $\operatorname{im} D=(\ker D^*)^\perp$ *(L^2 orthogonal complements).*
- *(v)* *One has $\operatorname{coker} D\cong\ker D^*$, finite-dimensional.*

Proof (i) The elliptic estimate tells implies that $s\in\ker D$ lies in $\bigcap_{k\geq m}L^2_k=C^\infty$.
(ii) If s_i is a bounded sequence in $\ker D$ then, by the elliptic estimate,

$$\|s_i\|_{2,m+1}\leq \mathsf{const};$$

so, by compactness of the inclusion, there is a subsequence that converges in L^2_m to a limit s_∞. Since D is a bounded operator $L^2_m\to L^2$, one has $\|Ds_\infty\|_{L^2}$=0, so $Ds_\infty=0$.
(iii) We may assume $m>0$. Take a sequence (t_i) in $\operatorname{im} D\subset L^2(F)$, converging to t_∞. Say $t_i=Ds_i$. Take a basis $(x_1,\dots,x_k)$ for $\ker D$. We may modify s_i (adding multiples of the x_j) to make it L^2-orthogonal to all the x_j; thus $s_i\perp\ker D$ in L^2.
There are now two possibilities: (s_i) is a bounded sequence in $L^2_m(E)$, or it is not. If it is bounded then, by compactness of the inclusion into L^2, it has a subsequence which converges in L^2. Passing to this subsequence, we have $\|s_i-s_j\|_{2,k+m}\leq C(\|t_i-t_j\|_{2,k}+\|s_i-s_j\|_2)$, so $\|s_i-s_j\|_{2,k+m}\to 0$ as $i,j\to\infty$. Then (s_i) is a Cauchy sequence, hence has a limit s_∞, and $Ds_\infty=t_\infty$, so $t_\infty\in\operatorname{im} D$.
On the other hand, if (s_i) is unbounded then we can pass to a subsequence with $\|s_i\|_{2,m}\to\infty$. Let $\hat{s}_i=s_i/\|s_i\|_{2,m}$. Then $D\hat{s}_i=t_i/\|s_i\|_{2,m}\to 0$. Since $(\hat{s}_i)$ is a bounded sequence in L^2_m, it has a convergent subsequence (by the case just considered). Its limit $\hat{s}$ is a unit vector in $(\ker D)^\perp$, yet $D\hat{s}=0$: contradiction.
(iv) Suppose $s\in L^2_m(F)$ is L^2-orthogonal to $\operatorname{im} D$. Then D^*s is L^2-orthogonal to $L^2(E)$, i.e. $D^*s=0$ in L^2. Thus $(\operatorname{im} D)^\perp=\ker D^*$. In general, for a subspace A of a Hilbert space, one has $A^{\perp\perp}=\overline{A}$. Since $\operatorname{im} D$ is closed, one has $\operatorname{im} D=(\ker D^*)^\perp$.
(v) D is again elliptic, so $\ker D^*$ is finite-dimensional; it complements $\operatorname{im} D$, hence is identified with $\operatorname{coker} D$.

□

Corollary 14.8 *An elliptic operator D of order m defines a Fredholm map $L^2_{k+m}(E) \to L^2_k(F)$ for any $k \geq 0$. Its index is independent of k. Moreover, its index depends only on the symbol of D.*

Proof We have seen that $\operatorname{im} D$ is closed in L^2; the higher k versions go similarly. We have also seen that $\ker D$ and $\operatorname{coker} D = \ker D^*$ are finite-dimensional. Since $\ker D$ and $\ker D^*$ comprise smooth sections, their dimensions are independent of k. If we add to D a differential operator K of order $< m$, then $K\colon L^2_{k+m} \to L^2_k$ factors as $L^2_{k+m} \xrightarrow{K} L^2_{k+1} \to L^2_k$. Since the inclusion $L^2_{k+1} \to L^2_k$ is compact, so is $K\colon L^2_{k+m} \to L^2_k$. Compact perturbations of a Fredholm operator do not affect its index. □

14.4 L^p bounds

Elliptic estimates also hold for L^p Sobolev spaces, for any $p > 1$. These estimates are by no means easy to prove; they depend on the Calderón–Zygmund theory of singular integral operators.

Theorem 14.9 (L^p elliptic estimate) *If D is elliptic of order m, one has estimates on the L^p_k Sobolev norms for each $p > 1$ and each $k \geq 0$:*

$$\|s\|_{p,k+m} \leq C_{p,k} \left(\|Ds\|_{p,k} + \|s\|_p \right).$$

Moreover, if s belongs to a complement to $\ker D$ in $L^p_m(E)$, one has

$$\|s\|_{p,k+m} \leq C_{p,k} \|Ds\|_{p,k}.$$

15 Clifford algebras, spinors and spin groups

This lecture is based on Deligne's *Notes on spinors* [4].

15.1 Clifford algebras

Let k be a commutative ring in which 2 is invertible.

Let (V, q) be a quadratic k-space: a k-module V with $q\colon V \to k$ a quadratic form. Thus the function $\langle \cdot, \cdot \rangle = V \times V \to k$ given by $\langle u, v \rangle = \frac{1}{2}(q(u+v) - q(u) - q(v))$ is k-bilinear, with $\langle v, v \rangle = q(v)$.

Typically, we will be interested in the case of positive-definite quadratic forms over $k = \mathbb{R}$, such as an inner product on a cotangent space $T^*_x M$. But the theory is algebraic, and other examples are sometimes of interest (with base ring $k = \mathbb{C}$, $\mathbb{R}[\epsilon]/\epsilon^2$, $C^\infty(M)$, etc.).

Definition 15.1 The *Clifford algebra* $\mathsf{cl}(V, q)$ is the associative k-algebra, with unit element $\mathbf{1}$, generated by V and subject to relations

$$v^2 = -q(v)\mathbf{1}, \quad v \in V. \tag{4}$$

The universal property underpinning this definition is as follows:

If A is an associative k-algebra with unit $\mathbf{1}_A$, and $f\colon V \to A$ a linear map such that $f(v)^2 = -q(v)\mathbf{1}_A$, then f extends to a unique homomorphism of unital k-algebras $\mathsf{cl}(V, q) \to A$.

A concrete construction—as for any presentation of a unital associative algebra—is to take the tensor algebra $\mathsf{T}V = \bigoplus_{n \geq 0} V^{\otimes n}$, and quotient by the 2-sided ideal I generated by the defining relations (in this case, by the elements $v \otimes v + q(v)\mathbf{1}$).

Warning. *Many texts use instead the relation $v^2 = +q(v)\mathbf{1}$.*

(4) is equivalent to the assertion that for all $u, v \in V$, we have

$$uv + vu = -2\langle u, v \rangle \mathbf{1}.$$

The following observations are immediate:

- The formation of Clifford algebras is compatible with extension of scalars $k \to K$, and is functorial in the k-module V. In particular, the linear action of the orthogonal group $\mathsf{O}(V, q)$ on V extends to an action $\mathsf{O}(V, q) \to \operatorname{Aut} \mathsf{cl}(V, q)$.
- The length l of a monomial $v_1 \cdots v_l \in \mathsf{cl}(V, q)$ is well-defined modulo 2, since (4) equates monomials of even length. This makes the Clifford algebra a $\mathbb{Z}/2$-graded algebra, also known as a *super-algebra*:
$$\mathsf{cl}(V, q) = \mathsf{cl}^0(V, q) \oplus \mathsf{cl}^1(V, q),$$
with $\mathsf{cl}^0(V, q)$ (resp. $\mathsf{cl}^1(V, q)$) spanned by monomials of even length (resp. odd length). The two summands are the ± 1-eigenspaces of the action of $-\mathrm{id}_V \in \mathsf{O}(V, q)$ on $\mathsf{cl}(V, q)$.
- Let $\mathsf{cl}(V, q)^{\mathsf{opp}}$ denote the opposite algebra (in the ungraded sense), which is the same k-module with the order-reversed product: $a \cdot^{\mathsf{opp}} b = b \cdot a$. The *principal anti-automorphism* is the unique homomorphism $\beta\colon\ \mathsf{cl}(V, q) \to \mathsf{cl}(V, q)^{\mathsf{opp}}$ extending id_V. It is an anti-involution.

- There is a notion of the opposite super-algebra $A^{\mathsf{opp}}_{\mathsf{super}}$ of the super-algebra A in which
$$a \cdot^{\mathsf{opp}}_{\mathsf{super}} b = (-1)^{|a||b|} b \cdot a.$$
We have $\mathsf{cl}(V,q)^{\mathsf{opp}}_{\mathsf{super}} = \mathsf{cl}(V,-q)$. In general, $\mathsf{cl}(V,q)$ and its opposite $\mathsf{cl}(V,-q)$ are not isomorphic algebras.
- Length defines an increasing filtration $\{F^\ell \,\mathsf{cl}(V,q)\}_{\ell\geq 0}$ of $\mathsf{cl}(V,q)$, with $F^\ell \cdot F^m \subset F^{\ell+m}$. In the associated graded algebra of the filtration $\mathrm{gr}\,\mathsf{cl}(V,q) = \bigoplus_\ell F^\ell/F^{\ell-1}$, one has $q(v) = 0$ for $v \in V$. This is the defining relation of the exterior algebra, and so defines an algebra epimorphism
$$i^\bullet \colon \Lambda^\bullet V \to \mathrm{gr}\,\mathsf{cl}(V,q), \quad v_1 \wedge \cdots \wedge v_k \mapsto [v_1]\cdots[v_k].$$
When V is a free module, it is clear that $i^\bullet$ is an isomorphism (actually it is true in general, as one can see by a localization argument). In particular, when V is has dimension d over a field k, we have
$$\dim_k \mathsf{cl}^0(V,q) = 2^{d-1} = \dim \mathsf{cl}^1(V,q).$$

15.1.1 Orthogonal sums

The theory of Clifford algebras and their representations is most elegant when we regard Clifford algebras as super-algebras (or even as super-algebras with an anti-automorphism of the underlying ungraded algebra). For example, define the super'tensor product of $\mathbb{Z}/2$-graded algebras as follows: $A\tilde{\otimes}B$ is $A \otimes_k B$ with the product
$$(a \otimes b)\cdot(a' \otimes b') = (-1)^{|b||a'|} aa' \otimes bb'.$$

Lemma 15.2 *When $(V,q) = (V_1,q_1) \oplus (V_2,q_2)$, there is a super-algebra isomorphism*
$$\theta\colon\ \mathsf{cl}(V_1,q_1)\,\tilde{\otimes}\,\mathsf{cl}(V_2,q_2) \to \mathsf{cl}(V,q),$$
respecting filtrations and lifting the isomorphism
$$\mathrm{gr}\,\theta = \wedge \colon \Lambda^\bullet V_1 \,\tilde{\otimes}\, \Lambda^\bullet V_2 \to \Lambda^\bullet(V_1 \oplus V_2).$$

Proof The inclusions $j_1\colon V_1 \to V$ and $j_2\colon V_2 \to V$ induce inclusions $\mathsf{cl}(j_i)\colon\ \mathsf{cl}(V_i,q_i) \to \mathsf{cl}(V,q)$. The map we want is $\theta(a\otimes b) = \mathsf{cl}(j_1)(a)\cdot\mathsf{cl}(j_2)(b)$. The map θ respects filtrations (where $F^k(A\,\tilde{\otimes}\,B) = \sum_i F^iA \otimes F^{k-i}B$). After passing to associated graded algebras, it induces the 'wedge' isomorphism given in the statement of the lemma. A filtered map inducing an isomorphism on the associated graded is an isomorphism. $\square$

15.2 Spinors

If $U = U^+ \oplus U^-$ is a $\mathbb{Z}/2$-graded k-module, the endomorphism algebra $\mathrm{End}_k\, U$ is a super-algebra $\mathsf{sEnd}\, U$, with $\mathsf{sEnd}^0\, U$ the parity-preserving endomorphisms $\mathrm{End}\, U^+ \times \mathrm{End}\, U^-$, and $\mathsf{sEnd}^1\, U$ the parity-reversing endomorphisms $\mathsf{Hom}(U^+,U^-) \to \mathsf{Hom}(U^-,U^+)$.

From now on in this section, k denotes a field (of characteristic $\neq 2$), and K an extension field of k.

Definition 15.3 Let (V,q) be even-dimensional and non-degenerate over k. A *spinor module*, defined over K, is a $\mathsf{cl}(V,q)$-super-module $S = S^+ \oplus S^-$ (i.e., a $\mathbb{Z}/2$-graded K-vector space with a representation $\rho\colon\ \mathsf{cl}(V,q) \to \mathsf{sEnd}_K\, S$) such that the K-linear extension
$$\rho_K\colon\ \mathsf{cl}(V_K,q_K) \to \mathsf{sEnd}_K\, S$$
is an isomorphism.

Definition 15.4 A *polarization* $P = (L, L')$ for an even-dimensional, non-degenerate quadratic form (U, Q) over K is a pair of K-subspaces of U such that $U = L \oplus L'$, and such that $Q|_L = 0$ and $Q|_{L'} = 0$.

One then has $L' \cong L^\vee$ (via Q), so $(U, Q) \cong (L \oplus L^\vee, \mathsf{ev})$.

- Over $\mathbb{C}$, or any algebraically closed field, a polarization for (U, Q) always exists.
- Over $\mathbb{R}$, a polarization exists iff q has signature zero.

Proposition 15.5 *Suppose that we are given* (V, q)*. A polarization* P *for* (V_K, q_K) *determines a spinor module over* K*.*

The construction is as follows: Define
$$S = S_P = \Lambda^\bullet L^\vee.$$
It is $\mathbb{Z}/2$-graded by the parity of $\bullet$: so $S = S^+ \oplus S^-$. For each $\lambda \in L$, define the *annihilation operator* $a(\lambda) = \iota_\lambda \in \mathsf{sEnd}^1 S$. For $\mu \in L^\vee$, define the *creation operator* $c(\mu) = \mu \wedge \cdot \in \mathsf{sEnd}^1 S$.These operators satisfy Heisenberg anticommutator relations
$$\{c(\lambda), c(\lambda')\} = 0, \quad \{a(\mu), a(\mu')\} = 0, \quad \{c(\lambda), a(\mu)\} = \mu(\lambda)\,\mathrm{id}.$$
The Clifford action on the spinors is now defined via a 'creation minus annihilation' formula:
$$\rho\colon \ \mathsf{cl}(L \oplus L^\vee, q) \to \operatorname{End} S, \quad \rho(\lambda, \mu) = c(\mu) - a(\lambda).$$

Proposition 15.6 ρ_K *is an isomorphism of* $\mathbb{Z}/2$*-graded* K*-algebras.*

Proof We work over K. Write $L = L_1 \oplus \cdots \oplus L_d$, a sum of lines. Then $L \oplus L^\vee$ decomposes as the orthogonal sum of $L_i \oplus L_i^\vee$, each summand having its evaluation pairing. Thus
$$\mathsf{cl}(L \oplus L^\vee) \cong \mathsf{cl}(L_1 \oplus L_1^\vee) \,\tilde{\otimes} \cdots \tilde{\otimes}\, \mathsf{cl}(L_d \oplus L_d^\vee).$$
We have
$$S = \Lambda^\bullet(L_1^\vee \oplus \cdots \oplus L_d^\vee) \cong \Lambda^\bullet L_1^\vee \otimes \cdots \otimes \Lambda^\bullet L_d^\vee,$$
and $\rho = \rho_L$ is the tensor product of the corresponding maps ρ_{L_i}. Thus it suffices to prove that ρ is an isomorphism when $\dim L = 1$. But in that case, $\mathsf{cl}^0(L \oplus L^\vee)$ and $\mathsf{sEnd}^0 S$ are both 2-dimensional, as are $\mathsf{cl}^1(L \oplus L^\vee)$ and $\mathsf{sEnd}^1 S$, and it is easy to check that ρ is an isomorphism. □

Corollary 15.7 *When* (V, q) *is polarized over the field* k*,*

(i) *any finite-dimensional, indecomposable (ungraded)* $\mathsf{cl}(V, q)$*-module is isomorphic to a sum of copies of* S*;*

(ii) *any finite-dimensional, indecomposable s* $\mathsf{cl}(V, q)$*-super-module is isomorphic to a sum of copies of* $S = S^+ \oplus S^-$ *and its parity-changed partner* $\Pi S = S^- \oplus S^+$*.*

Proof (i) A matrix-algebra $\operatorname{End} S$ over $k = \bar{k}$ has S as its only indecomposable module (see e.g. S. Lang, *Algebra*).
(ii) Let $T = T^0 \oplus T^1$ be a $\mathsf{cl}(V, q)$ super-module. If $(e_1, \ldots, e_{2m})$ is an orthonormal basis for V, the element
$$\omega = e_1 \ldots e_{2m} \in \mathsf{cl}^0(V)$$
anti-commutes with any $v \in V$, hence is central in $\mathsf{cl}^0(V)$. One has $\omega^2 = 1$. The ± 1 eigenspaces of ω in T are exactly the parity subspaces T^0 and T^1 (in some order). As an ungraded module we have $T \cong S^{\oplus r}$ for some r. The parity subspaces are then necessarily $T^0 = (S^+)^{\oplus r}$ and $T^1 = (S^-)^{\oplus r}$ (or vice versa), since they are determined by the action of ω. So $T = S^{\oplus r}$ as super-modules (up to parity shift). □

15.2.1 A quick note on spinors in odd dimension

A non-degenerate quadratic space (V, q) over an algebraically closed field $k = \bar{k}$ decomposes as an orthogonal sum $V = k \oplus V'$, where the quadratic form on k is $a \mapsto a^2$, and V' admits a polarization. Set

$$D_k = k[\varepsilon]/(\epsilon^2 + 1), \quad \varepsilon \in D^1 \text{ odd}.$$

Then $\mathsf{cl}(k) \cong D_k$. If S' is a spinor module for V' then

$$\mathsf{cl}(V, q) \cong D \,\tilde{\otimes}\, \mathsf{sEnd}_k S' \cong \mathsf{sEnd}_D(S' \otimes_k D).$$

With this in mind, we can define a spinor module for a non-degenerate quadratic space (V, q) over a field k, defined over the extension K, to be a representation ρ of $\mathsf{cl}(V, q)$ on a free, finite-rank D_K-supermodule S, such that that $\rho_K \colon \mathsf{cl}(V_K, q_K) \to \mathsf{sEnd}_{D_K} S$ is an isomorphism of K-superalgebras. We will not develop this notion here.

15.3 Projective actions

15.3.1 Projective action of the orthogonal group

Fix a spinor module S, over K, for the non-degenerate, $2m$-dimensional quadratic form (V, q) over k. The orthogonal group $\mathsf{O}(V)$ acts projectively on S, via a homomorphism

$$\Theta \colon \mathsf{O}(V) \to \mathsf{PGL}(S) = \mathrm{Aut}_K(S)/K^\times.$$

Construction of Θ: Any $g \in \mathsf{O}(V)$ extends to an automorphism $\mathsf{cl}(g)$ of the super-algebra $\mathsf{cl}(V)$, and so gives an irreducible representation $\rho \circ \mathsf{cl}(g) \colon \mathsf{cl}(V) \to \mathrm{Aut}\, S$. Now (S, ρ) is the unique irreducible super-module for $\mathsf{cl}(V)$, up to parity shift. Hence $(S, \rho \circ \mathsf{cl}(g))$ is isomorphic to (S, ρ), i.e., there is an intertwiner $\bar{g} \colon S \to S$, of either odd or even degree:

$$\rho \circ \mathsf{cl}(g) = \rho \circ \bar{g}.$$

$\bar{g}$ is unique up to scalars, and we set $\Theta(g) = [\bar{g}]$.

The center of $\mathsf{cl}^0(V)$ contains the volume element $\omega = e_1 \dots e_{2m}$. The automorphism $\mathsf{cl}(g)$ maps ω to $\det g \cdot \omega$. Since S^+ and S^- are the ± 1-eigenspaces of $\rho(\omega)$, the parity of $\bar{g}$ corresponds exactly to $\det g$. In other words: $\mathsf{SO}(V)$ is the parity-respecting subgroup.

Alternative construction of Θ: We have the action $\mathsf{cl} \colon \mathsf{O}(V) \to \mathrm{Aut}\, \mathsf{cl}(V)$ But $\mathsf{cl}(V)$ is a matrix algebra, and as such, all its automorphisms are inner, according to the Skolem–Noether theorem. *Thus $\mathsf{cl}(g)(a) = F(g) \cdot a \cdot F(g)^{-1}$, where $F(g) \in \mathsf{cl}(V)^\times$. Moreover $F(g)$, is well-defined modulo the center of $\mathsf{cl}(V)$, i.e., modulo $K^\times$. Thus we get a map $F \colon \mathsf{O}(V) \to \mathsf{cl}(V)^\times/K^\times$ such that $\mathsf{cl}(g) = \mathsf{Ad}\, F(g)$; and $\rho \circ F = \Theta$.*

Example 15.8 For the circle group $\mathsf{SO}(2)$, its projective action Θ on $S = \mathbb{C}^2$ amounts to the composite map

$$\mathsf{SO}(2) \xrightarrow{\text{square root}} \mathsf{SO}(2)/\{\pm 1\} \to \mathrm{Aut}\, \mathbb{C}P^1.$$

Note that Θ does not lift to a homomorphism $\mathsf{SO}(2) \to \mathrm{Aut}\, \mathbb{C}^2$.

15.3.2 Projective action of the orthogonal Lie algebra

In the same setting, the action $\mathsf{cl}\colon \mathsf{O}(V) \to \operatorname{Aut}\mathsf{cl}(V)$ induces, by differentiation, a map of Lie algebras

$$\delta = D\,\mathsf{cl}\colon \mathfrak{o}(V) \to \operatorname{Der}\mathsf{cl}(V).$$

On the left, we have the orthogonal Lie algebra (trace-free endomorphisms of V); on the right, the derivations of the algebra $\mathsf{cl}(V)$. To obtain this action, we extend scalars to $B = K[\epsilon]/\epsilon^2$, to obtain $\mathsf{cl}\colon \mathsf{O}_B(V[\epsilon]/\epsilon^2) \to \operatorname{Aut}_B\mathsf{cl}(V[\epsilon]/\epsilon^2)$. In $\mathsf{O}(V[\epsilon]/\epsilon^2)$ we have the elements of form $I + \xi\epsilon$, with $\xi \in \mathfrak{o}(V)$; and $\mathsf{cl}(I + \xi\epsilon) = I + \delta(\xi)\epsilon$.

Just as the action of $g \in \mathsf{O}(V)$ on the Clifford algebra is inner, $\mathsf{cl}(g) = \operatorname{Ad}F(g)$, so the derivations $\delta(\xi)$ are inner: $\delta(\xi) = [f(\xi), \cdot]$, where $f\colon \mathfrak{o}(V) \to \mathsf{cl}(V)/K$ is characterized by

$$\xi(v) = [f(\xi), v] \in V \subset \mathsf{cl}(V), \quad v \in V.$$

Moreover, we can lift f to land in $[\mathsf{cl}(V), \mathsf{cl}(V)] \subset \mathsf{cl}(V)$, since $[\mathsf{cl}(V), \mathsf{cl}(V)]$ projects isomorphically to $\mathsf{cl}(V)/K$ (because $A \xrightarrow{\cong} [A, A]/K$ when A is a full matrix algebra). Thus we have a map of Lie algebras

$$f\colon \mathfrak{o}(V) \to \mathsf{cl}(V),$$

landing in the commutator subalgebra and satisfying $\xi(v) = [f(\xi), v]$.

Define

$$\mathfrak{spin}(V) = f(\mathfrak{o}(V)).$$

Then $\mathfrak{spin}(V)$ is a Lie subalgebra of $\mathsf{cl}(V)$, and the map $f\colon \mathfrak{o}(V) \to \mathfrak{spin}(V)$ describes the action of $\mathfrak{o}(V)$ by derivations of $\mathsf{cl}(V)/K$.

We now make f explicit (our formula will show that $f(\mathfrak{o}(V)) \subset \mathsf{cl}^0(V)$).

Lemma 15.9 *There is a linear isomorphism*

$$\phi\colon \Lambda^2 V \to \mathfrak{o}(V), \quad x \wedge y \mapsto \phi(x \wedge y),$$

where $\phi(x \wedge y)$ is the trace-free endomorphism

$$\phi(x \wedge y)(v) = 2\left(\langle y, v\rangle x - \langle x, v\rangle y\right), \quad v \in V.$$

The inverse isomorphism is given, in terms of an orthonormal basis $(e_1, \ldots, e_n)$ for V, by

$$\phi^{-1}\colon \mathfrak{o}(V) \to \Lambda^2 V, \quad \xi \mapsto \frac{1}{4}\sum_{i,j} \xi_{ij} e_i \wedge e_j.$$

The proof is left to you. Now define $f\colon \mathfrak{o}(V) \to \mathsf{cl}^0(V)$ by

$$f(\xi) = \frac{1}{2}(xy - yx) = xy + \langle x, y\rangle \mathbf{1} \in \mathsf{cl}^0(V), \quad \xi = \phi(x \wedge y). \tag{5}$$

Lemma 15.10 *For $\xi \in \mathfrak{o}(V)$, one has*

$$\xi(v) = [f(\xi), v].$$

Moreover, f preserves Lie brackets.

Proof We need only check the relation $\xi(v) = [f(\xi), v]$:

$$[f(\xi), v] = [xy + \langle x, y\rangle \mathbf{1}, v] = [xy, v] = x(yv + vy) - (xv + vx)y = 2\langle y, v\rangle x - 2\langle x, v\rangle y = \xi(v).$$

□

The formula for f can be written still more explicitly as

$$f(\xi) = \frac{1}{4}\sum_{i,j} \xi_{ij} e_i \wedge e_j.$$

Note that it is valid even when (V, q) does not admit a polarization.

15.4 Spin groups

15.4.1 Clifford groups

Let (V,q) be a non-degenerate quadratic space over k.

The multiplicative group $\mathsf{cl}(V,q)^\times$ acts on $\mathsf{cl}(V,q)$ by inner automorphisms $u \mapsto (v \mapsto uvu^{-1})$. The *Clifford group* $G = G(V,q)$ is the normalizer of V, i.e., the subgroup of $\mathsf{cl}(V,q)^\times$ that preserves $V \subset \mathsf{cl}(V,q)$; let α_g be the action of $g \in G$ on V.

Note that $q(\alpha_g(v)) = -\alpha_g(v)^2 = -gv^2g^{-1} = q(v)$, i.e., $\alpha_v \in \mathsf{O}(V)$. Thus we have a homomorphism

$$\alpha\colon G \to \mathsf{O}(V).$$

If $u \in V$ and $q(u) \neq 0$ then $u \in G$: indeed, $u^{-1} = -\frac{1}{q(u)}u$, and

$$\alpha_u(v) = -\frac{1}{q(u)}uvu = -v + \frac{b(u,v)}{q(u)}u.$$

That is: $-\alpha_u$ *is the reflection in the hyperplane* $u^\perp$.

More generally, any product $w = u_1 \dots u_r$ of vectors with $q(u_i) \neq 0$ lies in G, with $\alpha_w \in \mathsf{O}(V)$.

Lemma 15.11 *The map $\alpha\colon G \to \mathsf{O}(V)$ is surjective.*

Proof The image of $G \to \mathsf{O}(V)$ contains the subgroup generated by reflections in hyperplanes $u^\perp$, where $q(u) \neq 0$. In the case of a positive-definite inner product space over $\mathbb{R}$, it is a familiar fact that $\mathsf{O}(V)$ is generated by reflections. It remains true, by a theorem of Cartan–Dieudonné, that for any non-degenerate quadratic form over a field of characteristic not 2, $\mathsf{O}(V)$ is generated by reflections. $\square$

The kernel of $G \to \mathsf{O}(V)$ is the intersection of G with the center of $\mathsf{cl}(V,q)$; thus it is the group of scalars $k^\times$, and we have a central extension

$$1 \to k^\times \to G \to \mathsf{O}(V) \to 1.$$

Lemma 15.12 *Every element of G is either a scalar, or a product $v_1 \dots v_r$ of elements of V, each having $q(v_i) \neq 0$.*

Proof Let $G' \subset G$ be the subgroup generated by $k^\times$ and V. Then the restriction to G' of $G \to \mathsf{O}(V)$ remains surjective, from which one sees that $G' = G$. The result follows. $\square$

Set $G^+ = G \cap \mathsf{cl}^+(V,q)$; it is the group generated by scalars and products v_1v_2. Since reflections have determinant -1, the image of G^+ in $\mathsf{O}(V)$ is exactly $\mathsf{SO}(V)$. Thus we get a central extension

$$1 \to k^\times \to G^+ \to \mathsf{SO}(V) \to 1.$$

15.4.2 Spin groups

Let $\beta\colon \mathsf{cl}(V,q) \to \mathsf{cl}(V,q)^{opp}$ be the principal anti-automorphism. If $g = cv_1 \dots v_r \in G$ (for $c \in k^\times$ and $r \geq 0$) we have $\beta(g)g = c^2 v_r \dots v_1 v_1 \dots v_r = c^2 \prod q(v_i) \in k^\times$, and this defines a homomorphism $\nu\colon G \to k^\times$, which is a version of the *spinor norm*.

The composite $k^\times \hookrightarrow G \xrightarrow{\nu} k^\times$ is the squaring map $c \mapsto c^2$.

We define the spin group

$$\mathsf{Spin}(V) = \ker \nu \cap G^+,$$

as the elements of G^+ of unit spinor norm. We have a central extension

$$1 \to \{\pm 1\} \to \mathsf{Spin}(V) \to \mathsf{SO}(V) \to 1.$$

As an algebraic group, $\mathsf{Spin}(V)$ has a Lie algebra $\mathfrak{spin}(V)$. If ϵ is a formal variable squaring to zero, $\mathfrak{spin}(V)$ is the kernel of the map $\mathsf{Spin}(V[\epsilon]) \to \mathsf{Spin}(V)$ given by setting ϵ to zero. Namely,

$$\mathfrak{spin}(V) = \{a \in \mathsf{cl}(V) : \ [a,V] \subset V, \ a + \tau(a) = 0\}.$$

The map $\mathfrak{spin}(V) \to \mathfrak{o}(V)$ induced by $\mathsf{Spin}(V) \to \mathsf{SO}(V)$ is an isomorphism.

The spin Lie algebra is exactly the Lie algebra $\mathfrak{spin}(V)$ which we constructed earlier: it consists of elements $a = \frac{1}{2}(xy - yx)$ where $x, y \in V$. Indeed, such elements a obey $[a,V] \subset V$ (as we saw earlier) and $a + \tau(a) = 0$, so lie in the spin Lie algebra as defined here; by dimension counting, they account for the whole of the spin Lie algebra.

15.4.3 Representations of spin

When (V,q) is even-dimensional over k and polarized over K, the K-linear spinor representation S of $\mathsf{cl}(V)$, restricted to $\mathsf{Spin}(V)$, lifts the projective representation of $\mathsf{O}(V)$:

$$\begin{tikzcd}
\mathsf{Spin}(V) \arrow[r, "\rho"] \arrow[d] & \mathrm{Aut}_K(S) \arrow[d] \\
\mathsf{SO}(V) \arrow[r, "\Theta"] & \mathrm{Aut}_K(S)/K^\times,
\end{tikzcd}$$

as becomes clear on unraveling the definitions. This diagram unveils a primary purpose of the Spin-groups: to realize the projective spinor representation of $\mathsf{SO}(V)$ as a linear representation of $\mathsf{Spin}(V)$. Notice that $-1 \in \mathsf{Spin}(V)$ maps under ρ to $-\mathrm{id}_S$.

Since $\mathsf{Spin}(V) \subset \mathsf{cl}^+(V)^\times$, the representation $\rho\colon \mathsf{Spin}(V) \to \mathrm{Aut}(S)$ decomposes as the direct sum of representations $\rho^\pm\colon \mathsf{Spin}(V) \to \mathrm{Aut}(S^\pm)$. Each of these is irreducible, because a subspace invariant under $\mathsf{Spin}(V)$ would be invariant under the whole of $\mathsf{cl}^0(V)$.

16 Spin groups and spin structures in low dimensions

16.1 The compact Lie groups $\mathsf{Spin}(n)$

16.1.1 The story so far

Assume $(V, |\cdot|^2)$ is a positive-definite inner product space over $\mathbb{R}$. Then $\mathsf{Spin}(V) = \mathsf{Spin}(V, |\cdot|^2)$ is the subgroup of the unit group $\mathsf{cl}(V)^\times$ in the Clifford algebra $\mathsf{cl}(V)$ formed from even products $e_1 \cdot \cdots \cdot e_{2r}$ of unit vectors $e_i \in V$. In this setting, $\mathsf{cl}(V)^\times$ can be regarded as a Lie group, and $\mathsf{Spin}(V)$ a closed Lie subgroup (they are the Lie groups associated with algebraic groups defined over $\mathbb{R}$).

We saw last time that the Lie algebra $\mathfrak{spin}(V)$ can also be realized inside the Lie algebra $\mathsf{cl}^0(V)$, with its commutator bracket; it is formed from the elements $[x, y]$ with $x, y \in V$. We wrote down an isomorphism $f\colon \mathfrak{o}(V) \to \mathfrak{spin}(V)$ such that $[f(\xi), \cdot]$ describes the infinitesimal action of ξ on $\mathsf{cl}(V, |\cdot|^2)$.

The group $\mathsf{Spin}\, V$ acts on V by inner automorphisms of the Clifford algebra, defining the homomorphism α in a short exact sequence

$$1 \to \pm 1 \to \mathsf{Spin}(V) \xrightarrow{\alpha} \mathsf{SO}(V) \to 1.$$

The existence of this short exact sequence makes clear that $\mathsf{Spin}(V)$ is *compact* (since $\mathsf{SO}(V)$ is). The derivative $D\alpha\colon \mathfrak{spin}(V) \to \mathfrak{o}(V)$, is exactly f^{-1}.

16.1.2 Exponentials

Since $\mathsf{SO}(V)$ is compact, its exponential map is surjective. That is, $\mathsf{SO}(V) = \exp \mathfrak{o}(V)$. Here $\exp \xi = \sum \frac{1}{n!}\xi^n$ can be computed inside $\operatorname{End} V$. Similarly, we have the subgroup

$$\exp(\mathfrak{spin}(V)) \subset \mathsf{Spin}\, V,$$

with the exponentials taken in $\mathsf{cl}^0(V)$.

Lemma 16.1 *There is a short exact sequence*

$$1 \to \pm 1 \to \exp(\mathfrak{spin}(V)) \xrightarrow{p} \mathsf{SO}(V) \to 1, \quad \dim V > 1,$$

whence

$$\mathsf{Spin}\, V = \exp(\mathfrak{spin}(V)).$$

Proof We have $\alpha(\exp f(\xi)) = \exp(D\alpha \circ f(\xi)) = \exp \xi$, for $\xi \in \mathfrak{o}(V)$. This proves exactness on the right (surjectivity). And if (e_1, e_2) are two orthonormal vectors in V then $\exp \frac{\pi}{2}[e_1, e_2] = -1$. □

Proposition 16.2 *When* $\dim V > 1$, *the Lie group homomorphism* $\mathsf{Spin}\, V \to \mathsf{SO}(V)$ *admits no continuous splitting.*

Proof Its realization by exponentials shows that $\mathsf{Spin}\, V$ is a connected Lie group, hence $\mathsf{Spin}(V) \to \mathsf{SO}(V)$ is a non-trivial covering map in the sense of topology. Thus it can have no continuous splitting. □

Consequently, $\mathsf{Spin}(n) := \mathsf{Spin}(\mathbb{R}^n, |\cdot|^2) \to \mathsf{SO}(n)$ can be characterized, for $n > 1$, as *the unique 2-to-1 homomorphism of Lie groups* $G \to \mathsf{SO}(n)$ *with* G *connected.*

16.2 Spinors

Recall also that, when $n = 2m$ is even, there is a spinor representation $\rho\colon \mathsf{cl}(V) \otimes \mathbb{C} \to \mathsf{sEnd}\, S$ on the spinors $S = S^+ \oplus S^-$. Since $\mathsf{Spin}(V)$ is contained in the even Clifford algebra $\mathsf{cl}^0(V)$, it preserves S^+ and S^-, and one has two 'half-spinor' representations

$$\rho^\pm\colon \mathsf{Spin}(2m) \to \mathrm{Aut}_{\mathbb{C}}\, S^\pm, \quad \dim\dim S^\pm = 2^{m-1}.$$

These are irreducible, because $\rho^\pm(\mathsf{Spin}(2m))$ spans $\rho^\pm(\mathsf{cl}^0(V)) = \mathrm{End}\, S^\pm$.

When $n = 2m+1$, there is a spinor representation $\rho\colon \mathsf{Spin}\, V \to \mathrm{Aut}_{\mathbb{C}}\, S$ on a complex vector space of dimension 2^m, which is in fact irreducible.

Warning: The spinors S are the unique irreducible representation of $\mathsf{cl}^0(V) \otimes \mathbb{C}$. However, there can be distinct irreducible complex representations of the real algebra $\mathsf{cl}^0(V)$. We will not treat this point systematically, but it will be implicit in treatments of orientations in what follows.

Lemma 16.3 *If (S, ρ) is a representation of the Clifford algebra $\mathsf{cl}(V, q)$ of a non-degenerate real quadratic form q, the following conditions on a hermitian form $(\cdot,\cdot)$ on S are equivalent:*

(1) $\rho(v) \in \mathfrak{u}(S)$ *for* $v \in V$, *i.e.,* $\rho(v)$ *is skew-hermitian:*

$$(\rho(v)s_1, s_2) + (s_1, \rho(v)s_2) = 0.$$

(2) $(\cdot,\cdot)$ *is spin-invariant, i.e.* $\rho(g)$ *is unitary for* $(\cdot,\cdot)$ *for all* $g \in \mathsf{Spin}(V)$.

(Neither q nor the hermitian form is not assumed positive-definite here.)

Proof If (i) holds then for $g \in \mathsf{cl}^0(V)$, one has $(\rho(g)s_1, s_2) = (s_1, \rho(\beta(g))s_2)$, where β is the principal anti-automorphism. So

$$(\rho(g)s_1, \rho(g)s_2)) = (s_1, \rho(g \cdot \beta(g))s_2).$$

For $g \in \mathsf{Spin}\, V$, one has $g \cdot \beta(g) = 1$, and so

$$(\rho(g)s_1, \rho(g)s_2)) = (s_1, s_2), \quad g \in \mathsf{Spin}(V),$$

i..e, $\rho(g)$ is unitary. Conversely, if (ii) holds then for $v \in V$ with $q(v) \neq 0$, one has

$$(\rho(v)s_1, s_2)) = -(\rho(v)s_1, q(v)^{-1}\rho(v)^2 s_2)) = q(v)^{-1}(s_1, \rho(v)^2 s_2) = -(s_1, \rho(v)s_2),$$

and by continuity the condition (i) holds even when $q(v) = 0$. □

Lemma 16.4 *When $q = |\cdot|^2$ is positive-definite, a spin-invariant, positive definite hermitian product exists on any representation (S, ρ).*

Proof Start with any positive-definite hermitian product $(\cdot,\cdot)_0$ on S, and then average over the compact group $\mathsf{Spin}\, V$ to obtain one that is spin-invariant. □

16.3 Low-dimensional cases

16.3.1 $\mathsf{Spin}(2)$

- The spin double covering of $\mathsf{SO}(2)$ is the angle-doubling map $\alpha\colon \mathsf{SO}(2) \to \mathsf{SO}(2)$.
- Let S^+ and S^- denote the representations of $\mathsf{SO}(2)$ on $\mathbb{C}$ in which rotation by θ acts as $e^{\pm i\theta}$ on $S^\pm$. Thus S^- is the dual (or conjugate) to S^+. Let $S = S^+ \oplus S^-$. We seek to realize S as the complex spinor representation of $\mathsf{cl}(\mathbb{R}^2)$ so that the standard hermitian product on S is spin-invariant.
- We need a Clifford map
$$\rho\colon \mathbb{R}^2 = \mathbb{C} \to \mathfrak{u}(S), \quad \rho(v) = \begin{bmatrix} 0 & \rho^-(v) \\ \rho^+(v) & 0 \end{bmatrix}, \quad \rho^-(v) = -\rho^+(v)^\dagger.$$
The Clifford relation $\rho^+(v)^\dagger \rho^+(v) = |v|^2$ is equivalent to the statement that
$$\rho^+\colon \mathbb{R}^2 = \mathbb{C} \to \mathsf{Hom}(S^+, S^-) = (S^-)^{\otimes 2}$$
is a $\mathbb{C}$-linear isometry.
- Define a *spin structure* on a 2-dimensional positive-definite inner product space V with an orientation (i.e. a hermitian line) to be a hermitian line L and a unitary isomorphism $\rho\colon V \to L^{\otimes 2}$. Given a spin structure, set $S^- = L$ and $S^+ = L^\vee$, and define ρ^+ and ρ^- as above. Then the resulting map $\mathsf{cl}(V) \otimes \mathbb{C} \to \mathsf{sEnd}(S^+ \oplus S^-)$ is an isomorphism, as one easily checks. Thus $S^+ \oplus S^-$ is a spinor representation.
- Given a spin structure in this sense, we can reconstruct the spin group as $\mathsf{Spin}\, V := \mathsf{U}(L)$. The map $\mathsf{Spin}\, V \to \mathsf{SO}(V)$ is $g \mapsto \rho \circ g^{\otimes 2} \circ \rho^{-1}$.

16.3.2 $\mathsf{Spin}(3)$

We shall use the quaternions $\mathbb{H}$, with the inclusion $\mathbb{C} \to \mathbb{H}$, $x + yi \mapsto x + yi$. There is the real inner product $\langle q_1, q_2 \rangle = \mathrm{Re}(q_1 \bar{q}_2)$ and norm-squared $|q|^2 := q\bar{q}$.
Consider a left $\mathbb{H}$-module E with a real inner product such that $(qe_1, e_2) = q(e_1, e_2) = (e_1, \bar{q}e_2)$ for $q \in \mathbb{H}$. The symmetry group of $(E, (\cdot,\cdot))$ is the compact symplectic group $\mathsf{Sp}(E)$.
We can regard E as a $\mathbb{C}$-vector space, with a hermitian product in the familiar sense. As such, it comes with a $\mathbb{C}$-antilinear isometry J with $J^2 = -1$ (namely, $Je = je$), and a complex symplectic form
$$\Omega \in \Lambda^2 E^*, \quad \Omega(e_1, e_2) = (e_1, Je_2).$$
Since J is determined by $(\cdot,\cdot)$ and Ω, the symmetries of $(E_{\mathbb{C}}, (\cdot,\cdot), \Omega)$—the intersection $\mathsf{U}(E) \cap \mathsf{Sp}(E, \Omega)$ of the unitary and complex symplectic groups—are exactly the symmetries $\mathsf{Sp}(E)$ of E as a hermitian quaternionic vector space. The basic instance of this principle is that $\mathsf{Sp}(1) := \mathsf{Sp}(\mathbb{H})$ coincides with $\mathsf{SU}(2)$ (in this case, Ω is a volume form).

- The spin covering of $\mathsf{SO}(3)$ is the map
$$\beta\colon \mathsf{Sp}(1) \to \mathsf{SO}(3),$$
given as follows: t An element of $\mathsf{Sp}(1)$ is a quaternion $q \in \mathbb{H}$ whose squared norm equals 1. There is an action $\mathsf{Sp}(1) \to \mathsf{SO}(\mathbb{H})$, $q \mapsto (x \mapsto qxq^{-1})$. This action preserves $1 \in \mathbb{H}$, hence also preserves the imaginary quaternions $\mathrm{Im}\,\mathbb{H} = 1^\perp = \mathbb{R}\{i, j, k\}$, and so defines $\beta\colon \mathsf{Sp}(1) \to \mathsf{SO}(\mathrm{Im}\,\mathbb{H})$. One has $\ker \beta = \{\pm 1\}$, and since $\dim \mathsf{Sp}(1) = 3 = \dim \mathsf{SO}(3)$, β is a local diffeomorphism, hence a 2-fold covering map. Note that $\mathsf{Sp}(1)$ is diffeomorphic to the 3-sphere S^3, hence connected, hence β is a non-trivial covering.

- One can also regard $\mathsf{Sp}(2)$ as $\mathsf{SU}(2)$, the symmetry group of $\mathbb{C}^2$ with its hermitian inner product and complex volume form $\Omega \in \Lambda^2\mathbb{C}^2$. We will see that the defining representation of $\mathsf{SU}(2)$ on $S = \mathbb{C}^2$ defines a spinor module S for $\mathsf{cl}(\mathbb{R}^3)$. We postulate that that the Clifford map
$$\rho\colon \mathbb{R}^3 \to \mathfrak{u}(S)$$
in fact maps to $\mathfrak{su}(S)$. The Clifford relations imply that ρ is an isometry when $\mathfrak{su}(2)$ has its inner product $\langle a, b\rangle = -\frac{1}{2}\operatorname{tr}(ab) = \frac{1}{2}\operatorname{tr}(a^\dagger b)$. Conversely, if ρ is an isometry then the Clifford relations holds. Indeed, $\rho(v)^2$ is a scalar matrix (like the square of any element of $\mathfrak{su}(2)$) of trace $-2|v|^2$, hence is $-|v|^2 I$.
- Thus if we have an isometry $\rho\colon \mathbb{R}^3 \to \mathfrak{su}(S)$ then we get a Clifford module (S, ρ), easily checked to be a spinor module.
- If (e_1, e_2, e_3) is an orthonormal basis then $\omega := e_1e_2e_3$ is central in $\mathsf{cl}(\mathbb{R}^3)$, with $\omega^2 = 1$. Hence $\rho(e_1)\rho(e_2)\rho(e_3)$ acts as $\pm I$ on S, and the sign changes according to whether or not the basis is oriented.
- Say V is a 3-dimensional oriented inner product space. Define a *spin structure* to be a 2-dimensional hermitian vector space S with a complex volume form $\Omega \in \Lambda^2 S$, and an isometry $\rho\colon V \to \mathfrak{su}(S)$ such that $\rho(e_1)\rho(e_2)\rho(e_3) = +I$ when (e_1, e_2, e_3) is an oriented orthonormal basis for V.
- $\mathsf{Spin}(V)$ is the symmetry group of the spin structure: the group of pairs $(g, \tilde{g})$ with $g \in \mathsf{SO}(V)$, $\tilde{g} \in \mathsf{SU}(S)$, so that the following diagram commutes:
$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "g"'] & \mathfrak{su}(S) \arrow[d, "\mathsf{Ad}(\tilde{g})"] \\ V \arrow[r, "\rho"] & \mathfrak{su}(S). \end{tikzcd}$$

16.4 Spin(4)

- There is a 2-1 covering
$$\gamma\colon \mathsf{Sp}(1) \times \mathsf{Sp}(1) \to \mathsf{SO}(4)$$
exhibiting an isomorphism $\mathsf{Spin}(4) \cong \mathsf{Sp}(1)\times\mathsf{Sp}(1)$, given as follows: left multiplication by unit quaternions preserves the norm on $\mathbb{H}$. Right multiplication by the conjugate of a unit quaternion also preserves the norm, and commutes with left multiplication. These actions together give the map $\alpha\colon \mathsf{Sp}(1) \times \mathsf{Sp}(1) \to \mathsf{SO}(\mathbb{H})$,
$$\gamma(q_1, q_2) = \{x \mapsto q_1 x q_2^{-1}\}.$$
There is a commutative diagram
$$\begin{tikzcd} \mathsf{Sp}(1) \arrow[r, "\beta"] \arrow[d, "\Delta"'] & \mathsf{SO}(\operatorname{Im}\mathbb{H}) \arrow[d] \\ \mathsf{Sp}(1) \times \mathsf{Sp}(1) \arrow[r, "\gamma"] & \mathsf{SO}(\mathbb{H}) \end{tikzcd}$$
with Δ the diagonal and the right vertical arrow the evident inclusion. If $(q_1, q_2) \in \ker\pi$ then (by taking $x = 1$) we see that $q_1 = q_2$, so $(q_1, q_2) \in \operatorname{im}\Delta$ and therefore $(q_1, q_2) = \pm(1, 1)$. Since $\dim \mathsf{Sp}(1) \times \mathsf{Sp}(1) = 6 = \dim \mathsf{SO}(4)$, we see that γ induces a Lie algebra isomorphism, and so is surjective.

- Let $S^+ = \mathbb{H}$ and $S^- = \mathbb{H}$. They are $\mathbb{H}$-vector spaces, and the underlying $\mathbb{C}$-vector spaces come with hermitian metrics. Let $\mathsf{Sp}(1) \times \mathsf{Sp}(1)$ act on S^+ via the first projection to $\mathsf{Sp}(1)$, and on S^- via the second projection. (These actions preserve the $\mathbb{H}$-structure and the hermitian metrics.) We will see that these are the half-spinor representations.
- The space $\mathsf{Hom}_{\mathbb{H}}(S^+, S^-)$ is a copy of $\mathbb{H}$, and carries an operator norm. The underlying real vector space is canonically oriented (since it has a complex structure). Let $\rho^+ : \mathbb{R}^4 \to \mathsf{Hom}_{\mathbb{H}}(S^+, S^-)$ be an $\mathbb{R}$-linear, orientation-preserving isometry. Define $\rho^-(v) = -\rho^+(v)^\dagger \colon \mathsf{Hom}_{\mathbb{H}}(S^-, S^+)$. Then one has $\rho^-(v)\rho^+(v) = |v|^2 \mathrm{id}_{S^+}$ and $\rho^+(v)\rho^-(v) = |v|^2 \mathrm{id}_{S^-}$, so ρ defines a Clifford module, in fact a spinor module.
- If V is a 4-dimensional oriented inner product space, define a *spin structure* on V to be a pair of rank 2 hermitian vector spaces S^+ and S^- with quaternionic structures, and an oriented isometry $\rho^+ \colon V \to \mathsf{Hom}_{\mathbb{H}}(S^+, S^-)$.
- $\mathsf{Spin}(V)$ is the symmetry group of the spin structure: the group of pairs $(g, \tilde{g})$ with $g \in \mathsf{SO}(V)$ and $\tilde{g} \in \mathsf{SU}(S^+) \times \mathsf{SU}(S^-) \subset \mathsf{SU}(S^+ \oplus S^-)$, making the following diagram commute:

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "g"'] & \mathrm{End}\, S \arrow[d, "\mathsf{Ad}(\tilde{g})"] \\ V \arrow[r, "\rho"] & \mathrm{End}\, S. \end{tikzcd}$$

17 Spin and Spin^c-structures: topology

17.1 Spin structures on vector bundles

Definition 17.1 Let $V \to M$ be rank n vector bundle. A *spin structure for* V, $\mathfrak{s} = (\mathsf{Spin}(V), \tau)$, is a principal $\mathsf{Spin}(n)$-bundle $\mathsf{Spin}(V) \to M$ together with an isomorphism

$$\tau\colon \mathsf{Spin}(V) \times_{\mathsf{Spin}(n)} \mathbb{R}^n \xrightarrow{\cong} V.$$

A *spin structure for* M is a spin structure for T^*M.

A spin structure for V induces a euclidean metric and an orientation in V. They are jointly characterized by the property that the principal $\mathsf{SO}(n)$-bundle $\mathsf{SO}(V)$ of orthonormal oriented frames is $\mathsf{Spin}(V) \times_{\mathsf{Spin}(n)} \mathsf{SO}(n)$. There is 2-fold covering $\mathsf{Spin}(V) \to \mathsf{SO}(V)$ over id_M.

Often, we think of the metric and orientation as being given in advance. Then one already has the principal bundle $\mathsf{SO}(V) \to M$. A spin structure amounts to a principal $\mathsf{Spin}(n)$-bundle $\mathsf{Spin}(V) \to M$ and an isomorphism $\mathsf{Spin}(V) \times_{\mathsf{Spin}(n)} \mathsf{SO}(n)$.

When n is even, there are associated half-spinor bundles $\mathbb{S}^{\pm} = \mathsf{Spin}(V) \times_{\mathsf{Spin}(n)} S^{\pm}$.[5] They are hermitian bundles and come with Clifford maps

$$\rho^+\colon V \to \mathsf{Hom}(\mathbb{S}^+, \mathbb{S}^-), \quad \rho^-\colon V \to \mathsf{Hom}(\mathbb{S}^-, \mathbb{S}^+),$$

with $\rho^-(v) = -\rho^+(v)^\dagger$, such that $\rho^-(v)\rho^+(v) = |v|^2\mathrm{id}$ and $\rho^+(v)\rho^-(v) = |v|^2\mathrm{id}$. When n is odd, one has a spinor bundle $\mathbb{S}$ but it does not split into half-spin bundles.

For low n, we defined a spin structure on a vector space in the previous lecture. Generalized to vector bundles, those definitions are equivalent to a spin structure as defined today:

- When $n = 2$, a spin structure in V amounts to square root of V as a hermitian line bundle.
- When $n = 3$, a spin structure in V amounts to a rank 2 hermitian vector bundle $\mathbb{S}$, a complex volume form $\Omega \in \Gamma(\Lambda^2\mathbb{S}^*)$ trivializing $\Lambda^2\mathbb{S}^*$, and an oriented isometry

$$\rho\colon V \to \mathfrak{su}(\mathbb{S}).$$

- When $n = 4$, a spin structure in V amounts to a pair $(\mathbb{S}^+, \mathbb{S}^-)$ of hermitian vector bundles, each with a quaternionic structure, and an oriented isometry

$$\rho^+\colon V \to \mathsf{Hom}_{\mathbb{H}}(\mathbb{S}_+, \mathbb{S}_-).$$

In each case, one must exhibit the principal spin bundle $\mathsf{Spin}\,V$. For $n = 4$, for instance, a point in $\mathsf{Spin}\,V_x$ consists of an oriented isometry $\theta\colon \mathbb{R}^4 \to V_x$ and unitary, quaternionic isomorphisms $\Theta_\pm\colon \mathbb{H} \to \mathbb{S}_\pm$ intertwining the map ρ^+ with the corresponding map in the model case.

[5] These are defined via a polarization of $V \otimes \mathbb{C}$. I ought to insert, in an earlier lecture, a brief explanation of how to obtain such polarizations in such a way as to ensure the spinors are canonically defined, as representations of the real Clifford algebra and not just of its complexification.

17.1.1 Uniqueness

Proposition 17.2 *If a spin structure in V exists, then the set of isomorphism classes of spin structures in V is a torsor for the group $H^1(M;\mathbb{Z}/2)$.*

Proof Suppose that we have a pair of spin structures $\mathfrak{s}$ and $\mathfrak{s}'$, in the oriented euclidean vector bundle V. An isomorphism $\mathfrak{s}\to\mathfrak{s}'$ is a map of $\mathsf{Spin}(n)$-spaces $\phi\colon \mathsf{Spin}(V)\to\mathsf{Spin}(V)'$ lying over the identity map of $\mathsf{SO}(V)$. For each $x\in M$, there are precisely 2 maps of $\mathsf{Spin}(n)$-spaces,

$$\phi_x\colon \mathsf{Spin}(V_x)\to\mathsf{Spin}(V_x)',$$

covering the identity of $\mathsf{SO}(V_x)$. As x varies, these isomorphisms form a 2-fold covering space $\mathsf{iso}(\mathfrak{s},\mathfrak{s}')\to M$. Via its monodromy $\pi_1(M)\to\mathbb{Z}/2$, this 2-fold covering space defines a class in $\delta(\mathfrak{s},\mathfrak{s}')\in H^1(X;\mathbb{Z}/2)$ which is the obstruction to finding a global section of $\mathsf{iso}(\mathfrak{s},\mathfrak{s}')\to M$, i.e. an isomorphism $\mathfrak{s}\to\mathfrak{s}'$.

For the converse, fix a good covering $\mathcal{U}=\{U_\alpha\}$ of M, and represent a given class in $H^1(M;\mathbb{Z}/2)$ by a $\{\pm1\}$-valued Čech cocycle $\delta=\{\delta_{\alpha\beta}\}$. Fix a trivialization of V over each open set U_α. A given spin structure $\mathfrak{s}$ has spin bundle $\mathsf{Spin}(V)$ with transition functions $\chi_{\alpha\beta}\colon U_{\alpha\beta}\to\mathsf{Spin}(V)$ forming a cocycle and inducing (via τ) the transition functions for V. The twisted transition functions $\tilde{\chi}_{\alpha\beta}=\delta_{\alpha\beta}\chi_{\alpha\beta}$ also form a cocycle, and still induce the transition functions for V, and so define a new spin structure. □

17.2 Existence and uniqueness in full

Let $V\to M$ be an oriented vector bundle. Fix a good covering $\mathcal{U}=\{U_\alpha\}$ of M; fix also a representative $\omega\in\check{C}^2(M,\mathcal{U};\mathbb{Z}/2)$ for the second Stiefel–Whitney class $w_2(V)$.

Theorem 17.3 *To give a spin structure in the oriented vector bundle $V\to M$ is equivalent to trivializing the second Stiefel–Whitney class $w_2(V)$. By this we mean the following:*

(i) A spin structure in V exists only if $w_2(V)=0$.

(ii) A 1-cochain $\eta\in\check{C}^1(M,\mathcal{U};\mathbb{Z}/2)$ with coboundary $\delta\eta=\omega$ determines a spin structure $\mathfrak{s}_\eta$.

(iii) If $\delta\eta=\omega=\delta\eta'$, then $\mathfrak{s}_\eta$ differs from $\mathfrak{s}_{\eta'}$ by the class $[\eta-\eta']\in H^1(M;\mathbb{Z}/2)$. Moreover, a 0-cocycle ζ with $\delta\zeta=\eta-\eta'$ determines an isomorphism $\mathfrak{s}_\eta\to\mathfrak{s}_{\eta'}$.

Example 17.4 One has a Bockstein exact sequence

$$H^1(X;\mathbb{Z}/2)\xrightarrow{\beta}H^2(X;\mathbb{Z})\xrightarrow{2}H^2(X;\mathbb{Z})\to H^2(X;\mathbb{Z}/2)$$

When $n=2$, the effect on the square root line bundle L of twisting the spin structure by $\delta\in H^1(X;\mathbb{Z}/2)$ is to add $\beta(\delta)$ to $c_1(L)$.

For clause (i), we note that $V|_{U_\alpha}$ admits a spin structure, and it is unique up to isomorphism (since $H^1(U_\alpha;\mathbb{Z}/2)=0$). Over an intersection $U_{\alpha\beta}$ we have two spin structures, and we can pick an isomorphism $\theta_{\alpha\beta}$ between them. We may assume $\theta_{\beta\alpha}=\theta_{\alpha\beta}^{-1}$. These isomorphisms are not necessarily consistent: they define a 2-cocycle $\omega_{\alpha\beta\gamma}=\theta_{\alpha\beta}\theta_{\beta\gamma}\theta_{\gamma\alpha}\in\check{C}^2(M,\mathcal{U};\mathbb{Z}/2)$. Its class $\omega(V)\in H^2(X;\mathbb{Z}/2)$ is independent of choices, because changing the isomorphisms $\theta_{\alpha\beta}$ by signs defining a 1-cochain $\eta_{\alpha\beta}$ changes ω to $\omega+\delta\eta$, and because the formation of $\omega(V)$ is compatible with refining $\mathcal{U}$.

The class $\omega(V)$ satisfies:

(1) $\omega(f^*V) = f^*\omega(V)$

(2) $\omega(U \oplus V) = \omega(U) + \omega(V)$, and in particular, $\omega(V \oplus \underline{\mathbb{R}}) = \omega(V)$.

Thus $\omega(V)$ is a characteristic class. Now, every rank r vector bundle is the pullback of the tautological vector bundle $\Lambda_r \to \mathsf{Gr}_r(\mathbb{R}^n)$, so to identify ω it will suffice to identify $\omega(\Lambda_r) \in H^2(\mathsf{Gr}_r(\mathbb{R}^n); \mathbb{Z}/2)$. Since ω is stable, it suffices to take n large.

But $H^2(\mathsf{Gr}_r(\mathbb{R}^n); \mathbb{Z}/2) = \mathbb{Z}/2$ for n large, the non-trivial element being the Stiefel–Whitney class w_2 (see e.g. A. Hatcher, *Vector bundles and K-theory*). In the case $r = 2$, we already know that w_2 obstructs spin structures, so $\omega = w_2$ when $r = 2$. The inclusion $\mathsf{Gr}_2(\mathbb{R}^n) \to \mathsf{Gr}_{2+k}(\mathbb{R}^{n+k})$ is an isomorphism on $H^2(\cdot, \mathbb{Z}/2)$, so

$$\omega = w_2.$$

and we deduce clause (i).

The details of (ii) and (iii) are left to the reader.

17.3 Spin^c-structures

Let $(V, |\cdot|^2)$ be a positive-definite real inner product space. We define the group $\mathsf{Spin}^c(V)$ to be the subgroup of $\mathsf{cl}(V \otimes \mathbb{C})^\times$ generated by $\mathsf{Spin}(V)$ and the unit-length scalars $\mathsf{U}(1)$. One has $\mathsf{Spin}\, V \cap \mathsf{U}(1) = \pm 1$, so

$$\mathsf{Spin}^c(V) \cong \frac{\mathsf{Spin}\, V \times \mathsf{U}(1)}{\pm(1,1)}.$$

There is a short exact sequence

$$1 \to \mathsf{U}(1) \to \mathsf{Spin}^c(V) \to \mathsf{SO}(V) \to 1$$

Set $\mathsf{Spin}^c(r) = \mathsf{Spin}^c(\mathbb{R}^r)$.

Definition 17.5 If $V \to M$ is a vector bundle of rank r, a Spin^c-structure on V is a principal $\mathsf{Spin}^c(r)$-bundle $\mathsf{Spin}^c(V) \to M$ and an isomorphism

$$\tau\colon\ \mathsf{Spin}^c(V) \times_{\mathsf{SO}(V)} \mathbb{R}^r \xrightarrow{\cong} V.$$

One still has a spinor representation $\rho\colon \mathsf{Spin}^c(r) \to \mathsf{U}(S)$, and if r is even, it is the direct sum of half-spinor representations S^+ and S^-. Hence a Spin^c-structure defines a spinor bundle $\mathbb{S} = \mathsf{Spin}^c(V) \times_{\mathsf{Spin}^c(r)} S$, which in the even-rank case is a direct sum $\mathbb{S}^+ \oplus \mathbb{S}^-$. And one has the Clifford map $\rho\colon V \to \mathfrak{u}(\mathbb{S})$, and when r is even, $\rho(v)$ exchanges $\mathbb{S}$ and $\mathbb{S}^-$.

Moreover, one can reconstruct the Spin^c-structure from the spinors: with $r = 2m$ even, say, suppose one is given hermitian vector bundles $\mathbb{S}^\pm \to M$, each of rank 2^{m-1}, and a map of vector bundles

$$\rho\colon V \to \mathsf{Hom}_{\mathbb{C}}(\mathbb{S}^+, \mathbb{S}^-)$$

such that

$$\rho(e)^\dagger \rho(f) + \rho(f)^\dagger)\rho(f) = 2\langle e, f\rangle \mathrm{id}_{\mathbb{S}^+}, \quad e, f \in V_x.$$

Let (S_{std}, ρ_{std}) be the spinors for $\mathbb{R}^r$. Then we can define $\mathsf{Spin}^c(V)_x$ as the space of pairs $(\theta, \tilde{\theta})$, where $\theta\colon \mathbb{R}^r \to V_x$ is an oriented isometry, and $\tilde{\theta} \in S_{std} \to \mathbb{S}_x$ a unitary map from the standard model for the spinors to $\mathbb{S}_x$, of even degree, such that

$$\rho(\theta(e)) = \tilde{\theta} \circ \rho_{std}(e) \circ \tilde{\theta}^{-1}.$$

One checks that $\mathsf{Spin}^c(V)_x$ is a torsor for the group $\mathsf{Spin}^c(r)$.

In low dimensions, one can say concretely what a Spin^c-structure on $V \to M$ amounts to. Assume V is already given with an orientation and euclidean metric.

- $n = 2$ (V a hermitian line bundle): a $\mathsf{Spin}^{\mathsf{c}}$-structure is a pair of hermitian line bundles $L^{\pm}$ and a $\mathbb{C}$-linear isometry $V \to \mathsf{Hom}_{\mathbb{C}}(L^+, L^-)$. (The spin case is where $L^+ = (L_-)^*$.)
- $n = 3$: a $\mathsf{Spin}^{\mathsf{c}}$-structure is a rank 2 hermitian vector bundle $\mathbb{S} \to M$ and an oriented isometry $\rho\colon V \to \mathfrak{su}(\mathbb{S})$ satisfying an orientation condition as in the spin case. (A spin structure comes also with a trivilization of $\Lambda^2\mathbb{S}$.)
- $n = 4$: a $\mathsf{Spin}^{\mathsf{c}}$-structure is a pair of rank 2 hermitian vector bundle $\mathbb{S}^{\pm} \to M$ and a map $\rho\colon V \to \mathsf{Hom}(\mathbb{S}^+, \mathbb{S}^-)$ such that $\rho(v)^\dagger \rho(v) = |v|^2 \mathrm{id}$.

17.4 Existence and uniqueness for $\mathsf{Spin}^{\mathsf{c}}$-structures

Proposition 17.6 *If $V \to M$ admits a $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$ then the $\mathsf{Spin}^{\mathsf{c}}$-structures form a torsor for $H^2(X;\mathbb{Z})$.*

Proof The pointwise $\mathrm{Spin}^{\mathrm{c}}$-isomorphisms from $\mathfrak{s}$ to $\mathfrak{s}'$ form a $\mathsf{U}(1)$-bundle $\mathsf{iso}(\mathfrak{s}, \mathfrak{s}') \to M$, and its Chern class $c = c_1(\mathfrak{s}, \mathfrak{s}') \in H^2(X;\mathbb{Z})$ is the obstruction to existence of a global isomorphism. Conversely, given $c \in H^2(X;\mathbb{Z})$, one can modify the transition functions by a Čech representative for c, just as in the spin case. □

Note that the effect of addition of $c \in H^2(X)$ on the spinors is to replace $(\mathbb{S}, \rho)$ by $(L_c \otimes \mathbb{S}, \mathrm{id} \otimes \rho)$, where L_c is the hermitian line bundle with $c_1(L_c) = c$.

There is a homomorphism $\lambda\colon\ \mathsf{Spin}^{\mathsf{c}}(r) \to \mathsf{U}(1)$, $[g, z] \mapsto z^2$; together with the covering $\alpha\colon\ \mathsf{Spin}^{\mathsf{c}}(r) \to \mathsf{SO}(r)$, this gives a 2-fold covering map

$$\mathsf{Spin}^{\mathsf{c}}(r) \xrightarrow{(\alpha,\lambda)} \mathsf{SO}(r) \times \mathsf{U}(1).$$

Via λ, any $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$ has an associated complex line bundle $L_{\mathfrak{s}}$. The homomorphism (α, λ) fits into a commutative diagram

$$\begin{tikzcd}
\mathsf{Spin}^{\mathsf{c}}(r) \arrow[r, "(\alpha,\lambda)"] \arrow[d] & \mathsf{SO}(r) \times \mathsf{U}(1) \arrow[d] \\
\mathsf{Spin}(r+2) \arrow[r, "\alpha"] & \mathsf{SO}(r+2).
\end{tikzcd}$$

The right vertical arrow is the inclusion of $\mathsf{SO}(r) \times \mathsf{U}(1) = \mathsf{SO}(r) \times \mathsf{SO}(2)$ into $\mathsf{SO}(r+2)$ given by the direct sum of matrices. The left vertical arrow is the inclusion $\mathsf{Spin}(r) \times \mathsf{U}(1) = \mathsf{Spin}(r) \times \mathsf{Spin}(2) \to \mathsf{Spin}(r+2)$. This diagram is a pullback square. From this we deduce the following:

Theorem 17.7 *The isomorphism classes $\mathsf{Spin}^{\mathsf{c}}$-structures $\mathfrak{s}$ on V, with $L_{\mathfrak{s}}$ a fixed oriented rank 2 real vector bundle L, are classified by spin structures on $V \oplus L$—hence by trivializations of $w_2(V \oplus L)$. Thus a $\mathsf{Spin}^{\mathsf{c}}$-structure exists if and only if there is some L with $w_2(L) = w_2(V)$, if and only if $w_2(V)$ admits a lift to $H^2(M;\mathbb{Z})$.*

17.4.1 The case of 4-manifolds

Theorem 17.8 (Hirzebruch–Hopf) *If X is a closed, oriented 4-manifold then $w_2(TX)$ admits a lift to integer coefficients.*

Hence X admits a Spin^c-structure. When $H_1(X;\mathbb{Z})$ has no 2-torsion, reduction $H^2(X;\mathbb{Z}) \to H^2(X;\mathbb{Z}/2)$ is surjective, so the point is to prove it without assumption on H_1.

The proof begins with a lemma valid for finite CW complexes Z. Write H^k and H_k for the kth integer (co)homology of Z, and $H^k(\mathbb{Z}/2)$ and $H_k(\mathbb{Z}/2)$ for the mod 2 (co)homology. Let $r\colon H^k \to H^k(\mathbb{Z}/2)$ and $r\colon H_k \to H_k(\mathbb{Z}/2)$ denote the maps induced by reduction $\mathbb{Z} \to \mathbb{Z}/2$.

Lemma 17.9 *There is a short exact sequence of abelian groups*

$$0 \to r(H^k_{tors}) \to H^k(\mathbb{Z}/2) \xrightarrow{\mathsf{ev}} \mathsf{Hom}(H_k, \mathbb{Z}/2) \to 0.$$

Proof Universal coefficients gives a pair of short exact sequences forming the rows in the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & \mathsf{Ext}^1(H_{k-1},\mathbb{Z}) \arrow[r] \arrow[d] & H^k \arrow[r, "\mathsf{ev}"] \arrow[d] & \mathsf{Hom}(H_k,\mathbb{Z}) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathsf{Ext}^1(H_{k-1},\mathbb{Z}/2) \arrow[r] & H^k(\mathbb{Z}/2) \arrow[r, "\mathsf{ev}"] & \mathsf{Hom}(H_k,\mathbb{Z}/2) \arrow[r] & 0.
\end{tikzcd}$$

The vertical maps arise make the diagram commute by a naturality property of universal coefficients sequences. The vertical map $\mathsf{Ext}^1(r)$ between the Ext-groups fits into a long exact sequence of Ext-groups, which continues

$$\mathsf{Ext}^1(H_{k-1},\mathbb{Z}) \xrightarrow{\mathsf{Ext}^1(r)} \mathsf{Ext}^1(H_{k-1},\mathbb{Z}/2) \to \mathsf{Ext}^2(H_{k-1},\mathbb{Z}) \to \dots,$$

but over the base ring $\mathbb{Z}$, $\mathsf{Ext}^2 = 0$, so $\mathsf{Ext}^1(r)$ surjects. In the upper row of the commutative diagram, $\mathsf{Ext}^1(H_{k-1},\mathbb{Z})$ is a torsion (in fact, finite) group, and its quotient $\mathsf{Hom}(H_k,\mathbb{Z})$ a torsion-free group. Thus the Ext group is the torsion subgroup of $H^k_{tors} \subset H^k$, and the Ext group in the lower row is $r(H^k_{tors})$. Thus we can rewrite the lower exact sequence as

$$0 \to r(H^k_{tors}) \to H^k(\mathbb{Z}/2) \xrightarrow{\mathsf{ev}} \mathsf{Hom}(H_k, \mathbb{Z}/2) \to 0,$$

as claimed. □

Proposition 17.10 *Let M be a compact n-manifold. Under the cup product pairing over $\mathbb{Z}/2$, on $H^k(M;\mathbb{Z}/2) \times H^{n-k}(M;\mathbb{Z}/2) \to \mathbb{Z}/2$, the $\mathbb{Z}/2$-subspaces*

$$r(H^k_{tors}) \subset H^k(M;\mathbb{Z}/2), \quad r(H^{n-k}) \subset H^{n-k}(M;\mathbb{Z}/2)$$

are mutual annihilators.

Proof Using integer-coefficient Poincaré duality, rewrite the exact sequence of the lemma as

$$0 \to r(H^k_{tors}) \to H^k(\mathbb{Z}/2) \xrightarrow{e} \mathsf{Hom}(H^{n-k}, \mathbb{Z}/2) \to 0,$$

where $e(x)(y) = \langle x \cup y, [M] \rangle$. Thus $\ker e$ is the annihilator of $r(H^{n-k})$. But $\ker e = r(H^k_{tors})$. The cup product pairing on mod 2 cohomology is perfect, by mod 2 Poincaréduality, and it follows that, reciprocally, $r(H^{n-k})$ is the annihilator of $r(H^k_{tors})$. □

We can now prove the result about closed, oriented 4-manifolds X. To show that $w_2(TX)$ admits an integer lift, it suffices to show that $w_2(TX) \cdot \bar{t} = 0$ whenever $\bar{t}$ is the mod 2 reduction of a torsion integer class t. But $w_2(TX) \cdot \bar{t} = \bar{t} \cdot \bar{t}$ by Wu's formula, and $\bar{t} \cdot \bar{t} = r(t \cdot t) = 0$.

Remark. Teichner and Vogt have observed that Spin^c-structures exist on arbitrary oriented 4-manifolds, not necessarily compact. For convenience, assume that X is homotopy-equivalent to a finite CW complex, so that our lemma applies. One then needs to show that w_2 annihilates the mod 2 reductions of torsion classes t in the compactly supported cohomology $H^2_c(X)$. Such classes are Poincaré dual to homology classes $h \in H_2(X;\mathbb{Z})$, and the relation $w_2(\bar{h}) = \bar{h}^2$ can then be checked by working in a tubular neighborhood of an oriented embedded surface representing h.

18 Dirac operators

18.1 The Levi-Civita connection

A Riemannian manifold M has a distinguished connection ∇ in TM, the Levi-Civita connection, uniquely characterized by two properties:

(i) ∇ is orthogonal: $d\langle u,v\rangle = \langle \nabla u, v\rangle + \langle u, \nabla v\rangle$; and
(ii) ∇ is torsion-free: $\nabla_u v - \nabla_v u - [u,v] = 0$.

The Riemann curvature tensor is the curvature of the Levi-Civita connection,

$$R = \nabla \circ \nabla \in \Omega^2_M(\mathfrak{so}(T^*M)), \quad R_{u,v} = \nabla_u\nabla_v - \nabla_v\nabla_u - \nabla_{[u,v]}.$$

Its components

$$R_{ijkl} = \langle R_{\partial_i,\partial_j}(\partial_k), \partial_l\rangle = \langle (\nabla_i\nabla_j - \nabla_j\nabla_i)\partial_k, \partial_l\rangle$$

transform under the symmetric group S_4 via the sign character $\varepsilon\colon S_4 \to \{\pm 1\}$. Moreover,

$$R_{i(jkl)} := R_{ijkl} + R_{jkil} + R_{kijl} = 0.$$

We find it convenient to work with the dual connection ∇^* in T^*M, $(\nabla^* e)(v) = e(\nabla v)$. The metric isomorphism $TM \to T^*M$ identifies ∇ with ∇^*, so they have identical curvature; henceforth we write ∇ to mean ∇^*.

18.2 Clifford connections

Definition 18.1 Let (M,g) be a Riemannian manifold, $\mathbb{S} \to M$ be a hermitian vector bundle, and $\rho\colon T^*M \to \mathfrak{u}(\mathbb{S})$ a Clifford map:

$$\rho(e)^2 = -|e|^2 \mathrm{id}_{\mathbb{S}}.$$

Let ∇ be the Levi–Civita connection in T^*M. A *Clifford connection* is a unitary connection $\tilde{\nabla}$ in $\mathbb{S}$ for which for which ρ is parallel:

$$[\tilde{\nabla}_v, \rho(e)] = \rho(\nabla_v e).$$

When ρ permutes the summands in a splitting $\mathbb{S} = \mathbb{S}^+ \oplus \mathbb{S}^-$, we require that $\tilde{\nabla}_v$ should preserve the summands.

Proposition 18.2 *When $\mathbb{S}$ is the spinor bundle of a spin structure $\mathfrak{s}$, there is a distinguished Clifford connection ∇^{spin} (it will be called the* spin connection*).*

Proof Work with an open covering $\{U_\alpha\}$ for M, and with local trivializations of the principal spin bundle $\mathsf{Spin}(V) \to M$. In the resulting local trivializations of V, ∇ is given over U_α by a 1-form $a_\alpha \in \Omega^1_{U_\alpha}(\mathfrak{so}(r))$ (namely, $\nabla = d + a$). These 1-forms satisfy a consistency condition with respect to the transition functions. A connection in $\mathbb{S}$ is given by 1-forms $A_\alpha \in \Omega^1_{U_\alpha}(\mathrm{End}\, S)$. The isomorphism $D\alpha\colon \mathfrak{spin}(r) \to \mathfrak{so}(r)$ gives rise to such a connection ∇^{spin}, with $A_\alpha = (D\alpha)^{-1} A_\alpha$. Here $\mathfrak{spin}(r)$ acts on the spinors S via the infinitesimal spin representation $D\rho\colon \mathfrak{spin}(r) \to \mathfrak{u}(S)$. By construction, ∇^{spin} is unitary and makes Clifford multiplication parallel; and it does not depend on the local trivializations, because under a change in these, it transforms in the same way as does ∇. $\square$

Remark. Had we developed the principal bundle perspective on connections, the construction of the spin connection would be perfectly transparent.

To give a formula for ∇^{spin}, we will need to invoke our formula for the Lie algebra isomorphism

$$f\colon \mathfrak{o}(V_x) \to \mathfrak{spin}(V_x)$$

inverse to the derivative of the covering map $\alpha\colon \mathsf{Spin}(V_x) \to \mathsf{SO}(V_x)$. In terms of an oriented orthonormal basis $(e_1, \dots, e_r)$ for V, we have

$$f\colon \mathfrak{so}(V_x) \to \mathfrak{spin}(V_x) \subset \mathsf{cl}^0(V_x), \quad f(\xi) = \frac{1}{4}\sum_{i,j}\langle \xi e_i, e_j\rangle \rho(e_i)\rho(e_j).$$

We give the formula for ∇^{spin}, using local coordinates $(x_1, \dots, x_n)$ and an oriented orthonormal local frame $(e_1, \dots, e_r)$ for V. Write

$$\nabla_{\partial_i} = \partial_i + A_i, \quad A_i(x) \in \mathfrak{so}(r).$$

Here A_i has entries $A_i^{\alpha\beta}$ Then

$$\nabla^{spin}_{\partial_i} = \partial_i + f(A_i) = \partial_i + \frac{1}{4}\sum_{\alpha,\beta} A_i^{\alpha\beta}\rho_\alpha\rho_\beta,$$

and

$$F(\nabla^{spin}) = f \circ R.$$

Proposition 18.3 *When $\mathbb{S}$ is the spinor bundle of a $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$, Clifford connections form an affine space modeled on $\Omega^1_M(i\mathbb{R})$.*

Proof Let $\tilde{\nabla}$ be a Clifford connection. Then any other takes the form $\tilde{\nabla} + A$, where $A \in \Omega^1_M(\mathfrak{u}(\mathbb{S}))$. For any tangent vector v, $A(v) \in \mathfrak{u}(\mathbb{S})$ is an infinitesimal automorphism of the $\mathbb{S}$ as a representation of the full Clifford algebra. Since $\mathbb{S}$ (or $\mathbb{S}^+$ and $\mathbb{S}^-$ is epresentation is irreducible, all its unitary automorphisms are in $\mathsf{U}(1)$, and its infinitesimal automorphisms in $i\mathbb{R}$. Moreover, $\tilde{\nabla} + a\mathrm{id}_{\mathbb{S}}$ *is* a Clifford connection when $a \in \Omega^1_M(i\mathbb{R})$.

For existence, notice that Clifford connections exist locally in M (we can use the spin connection); we can patch them together as usual, via partitions of unity. □

The Clifford connection determines, and is determined by, a unitary connection ∇° in the associated line bundle $L_{\mathfrak{s}}$ One has

$$(\tilde{\nabla} + a)^\circ = \tilde{\nabla}^\circ + 2a$$

(note the factor of 2!). For the case of a spin structure, $L_{\mathfrak{s}}$ is a trivialized bundle and $\tilde{\nabla}^{circ}$ the trivial connection. Write

$$F^\circ(\tilde{\nabla}) = F_{\tilde{\nabla}^\circ} \in \Omega^2_M(i\mathbb{R}),$$

and note that

$$F^\circ(\tilde{\nabla} + a) = F^\circ(\tilde{\nabla}) + 2\,da.$$

Locally in M, we can lift the $\mathsf{Spin}^{\mathsf{c}}$-structure to a spin structure, and compare $\tilde{\nabla}$ to the spin connection. We thereby see that

$$F_{\tilde{\nabla}} = F_{\nabla^{spin}} + \frac{1}{2}F^\circ(\tilde{\nabla}) \otimes \mathrm{id}_{\mathbb{S}} = f \circ R + \frac{1}{2}F^\circ(\tilde{\nabla}) \otimes \mathrm{id}_{\mathbb{S}}.$$

18.3 The Dirac operator

Let M be a manifold equipped with a $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$, and let $\tilde{\nabla}$ be a Clifford connection associated with the Levi-Civita connection ∇ in T^*M. The *Dirac operator* for $\tilde{\nabla}$ is the operator

$$D = \rho \circ \nabla \colon \Gamma(\mathbb{S}) \to \Gamma(\mathbb{S}),$$

namely, the composite

$$\Gamma(\mathbb{S}) \xrightarrow{\nabla} \Gamma(T^*M \otimes \mathbb{S}) \xrightarrow{\rho} \Gamma(\mathbb{S}).$$

In the case of a spin structure, there is a distinguished Clifford connection ∇^{spin}, and hence a distinguished Dirac operator D^{spin}.

In terms of a local frame $(e_1, \dots, e_n)$ for TM, with dual coframe $(e^1, \dots, e^n)$ for T^*M, one has $\nabla = \sum \nabla_{e_i} \otimes e^i$, and so

$$D\phi = \sum_i \rho(e^i) \circ \nabla_{e_i}.$$

In the even-dimensional case, it exchanges the half-spinor bundles, having components

$$D^{\pm} \colon \Gamma(\mathbb{S}^{\pm}) \to \Gamma(\mathbb{S}^{\mp}),$$

with D^- the formal adjoint to D^+.

If f is a function, one has

$$[D, f] = \rho \circ [\nabla, f] = \rho \circ df,$$

which shows that D has symbol

$$\sigma_D(\xi) = \rho(\xi) \in \mathrm{End}(\mathbb{S}.)$$

Thus D is indeed an example of a Dirac operator in the sense that its symbol satisfies the Clifford relation.

18.4 The formal adjoint to a covariant derivative

Let ∇ be an orthogonal covariant derivative in a euclidean vector bundle $V \to M$, and assume M is Riemannian.

Lemma 18.4 *For a vector field v, define its divergence $\mathsf{div}\, v$ to be the function $\mathsf{div}\, v = \star d(\iota_v \mathsf{vol}) = \star \mathcal{L}_v \mathsf{vol}$. Then, when M is compact,*

(1) *The operator*

$$\nabla_v^* = -\nabla_v - \mathsf{div}\, v$$

is the formal adjoint to ∇_v^*.

(2) *Define $\nabla^* \colon \Gamma(T^*M \otimes E) \to \Gamma(E)$ by*

$$\nabla^*(v^\sharp \otimes s) = \nabla_v^* s,$$

where v is a vector field and $v^\sharp$ the corresponding 1-form (so $g(u, v) = \iota_u(v^\sharp)$). Then ∇^* is the formal adjoint to ∇^*.

Proof (1) We will use the Cartan formula for the Lie derivative on forms, $\mathcal{L}_v = \iota_v \circ d + d \circ \iota_v$, the fact that $\mathcal{L}_v$ is a derivation, and Stokes's theorem. One has

$$\begin{aligned}
\int_M \big(\langle \nabla_v s_1, s_2\rangle + \langle s_1, \nabla_v s_2\rangle\big)\,\mathsf{vol} &= \int_M \iota_v d(s_1, s_2)\cdot \mathsf{vol}\\
&= \int_M \mathcal{L}_v(s_1, s_2)\cdot \mathsf{vol}\\
&= \int_M \mathcal{L}_v((s_1, s_2)\mathsf{vol}) - \int_M (s_1, s_2)\mathcal{L}_v \mathsf{vol}\\
&= \int_M d\circ \iota_v((s_1, s_2)\mathsf{vol}) - \int_M (s_1, s_2)\,\mathsf{div}(v)\mathsf{vol}\\
&= -\int_M (s_1, \mathsf{div}(v)s_2)\mathsf{vol}.
\end{aligned}$$

(2) We have

$$\begin{aligned}
\int_M \langle \nabla^*(v^\sharp \otimes s_1, s_2\rangle \mathsf{vol} &= \int_M \langle \nabla_v^* s_1, s_2\rangle \mathsf{vol}\\
&= \int_M \langle s_1, \nabla_v s_2\rangle \mathsf{vol}\\
&= \int_M \langle v^\sharp \otimes s_1, \nabla s_2\rangle \mathsf{vol}.
\end{aligned}$$

□

18.5 The Lichnérowicz formula

This formula is the result of what S.K. Donaldson has described as "one of the most fruitful calculations in differential geometry."

Theorem 18.5 (Lichnérowicz formula) *One has*

$$D^2 = \tilde\nabla^*\tilde\nabla + \frac{1}{4}\mathsf{scal}_g \cdot \mathsf{id}_{\mathbb{S}} + \frac{1}{2}\rho(F^\circ),$$

In this formula: (i) $\tilde\nabla^*\colon \Gamma(T^*M\otimes \mathbb{S}) \to \Gamma(\mathbb{S})$ is the formal adjoint

(ii) $\mathsf{scal}_g = \sum_{ij} R_{ijji}$ is the scalar curvature; (iii) 2-forms (such as $F_{\tilde\nabla^\circ}$) act on spinors via the map $\rho\colon \Lambda^2_M \to \mathrm{End}^0\,\mathbb{S}$, $\rho(e\wedge f) = \frac{1}{2}(\rho(e)\rho(f) - \rho(f)\rho(e))$.

This theorem is an example of a *Weitzenböck formula*: a formula that compares the square a certain Dirac operator in a bundle E with the covariant Laplacian $\nabla^*\nabla$ associated with a connection ∇ in E. The difference is necessarily a first-order operator, but a Weitzenböck formula identified a connection ∇ such that the difference is zeroth-order, and computes this difference.

Proof Work at the origin in local coordinates $(x_1, \dots, x_n)$, and let $e_i = dx_i(0)$. We have $\tilde\nabla = \sum_i \tilde\nabla_i \otimes dx_i$, where $\tilde\nabla_i = \nabla_{\partial_i}$.

It will be convenient to choose coordinates so that

$$\nabla_i(dx_j)(0) = 0.$$

Equivalently, in the local trivialization of T^*X induced by the coordinates, one has $\nabla = d + A$ where $A(0) = 0$. Geodesic coordinates (i.e. the coordinates induced by the exponential map of g) have this property.

In these coordinates, one has $\mathrm{div}(\partial_i) = 0$ at the origin, so $\tilde{\nabla}^* = -\sum_i \tilde{\nabla}_i \otimes \iota(\partial_i)$ at the origin. So $\tilde{\nabla}^*\tilde{\nabla} = -\sum \tilde{\nabla}_i \tilde{\nabla}_i$ at the origin.

Write $\{\cdot,\cdot\}$ for the anticommutator of operators, $\{A,B\} = AB + BA$.

We have $D = \sum_i \rho_i \nabla_i$, where $\rho_i = \rho(dx_i)$, and, at $x = 0$,

$$\begin{aligned} -\tilde{\nabla}^*\tilde{\nabla} + D^2 &= \sum_i \tilde{\nabla}_i\tilde{\nabla}_i + \sum_{i,j} \rho_i \tilde{\nabla}_i \circ \rho_j \tilde{\nabla}_j \\ &= \sum_i \tilde{\nabla}_i\tilde{\nabla}_i + \tfrac{1}{2}\sum_{i,j}\{\rho_i,\rho_j\}\tilde{\nabla}_i\tilde{\nabla}_j + \sum_{i,j}\rho_i[\tilde{\nabla}_i,\rho_j]\tilde{\nabla}_j + \tfrac{1}{2}\sum_{i,j}\rho_i\rho_j[\tilde{\nabla}_i,\tilde{\nabla}_j] \\ &= \sum_i \tilde{\nabla}_i\tilde{\nabla}_i - \sum_{i,j}\delta_{ij}\tilde{\nabla}_i\tilde{\nabla}_j + \sum_{i,j}\rho_i\rho(\nabla_i dx_j)\tilde{\nabla}_j + \sum_{i<j}\rho_i\rho_j(F_{\tilde{\nabla}})_{ij} \\ &= \sum_{i<j}\left(F_{\tilde{\nabla}}\right)_{ij}\rho_i\rho_j \end{aligned}$$

The the second equality is a matter of algebra; the third uses $\nabla_i(dx_j)(0) = 0$.

We have $F_{\tilde{\nabla}} = f(R) + \frac{1}{2}F^\circ(\tilde{\nabla}) \otimes \mathrm{id}_{\mathbb{S}}$. And $\sum_{i<j} F^\circ_{ij}\rho_i\rho_j = \rho(F^\circ)$. So it remains to show that

$$\sum_{i<j} f(R_{ij})\rho_i\rho_j = \frac{1}{4}\mathrm{scal}_g.$$

Well,

$$\sum_{i<j} f(R_{ij})\rho_i\rho_j = \frac{1}{4}\sum_{i<j;k,l} R_{lkij}\rho_i\rho_j\rho_k\rho_l = \frac{1}{8}\sum_{i,j,k,l} R_{lkij}\rho_i\rho_j\rho_k\rho_l.$$

The second equality uses the fact that $R_{lkij} = -R_{klij}$, and the Clifford relation $\{\rho_i,\rho_j\} = \delta_{ij}$.

We have mentioned that the curvature tensor has the symmetry $R_{l(kij)} := R_{lkij} + R_{lkji} + R_{lijk} = 0$. With this in mind, consider the sum

$$S_{ijkl} = \sum_{\sigma\in S_3} R_{l\sigma_k\sigma_i\sigma_j}\rho_{\sigma_i}\rho_{\sigma_j}\rho_{\sigma_k}.$$

We have

$$\sum_{i,j,k,l} R_{lkij}\rho_i\rho_j\rho_k\rho_l = \sum_l \sum_{i\le j\le k} S_{ijkl}\rho_l = \sum_l \sum_{i<j<k} S_{ijkl}\rho_l + \sum_l \sum_{i<j=k} S_{ijkl}.$$

Note here that $S_{ijkl} = 0$ when $i = j$, because $R_{lkii} = 0$. Now consider these two kinds of terms:

- $i < j < k$. Then $\rho_{\sigma_i}\rho_{\sigma_j}\rho_{\sigma_k} = \varepsilon(\sigma)\rho_i\rho_j\rho_k$, so
$$S_{ijkl} = \sum_{\sigma\in S_3} R_{l\sigma_k\sigma_i\sigma_j}\rho_{\sigma_i}\rho_{\sigma_j}\rho_{\sigma_k} = \left(\sum_{\sigma\in S_3}\varepsilon(\sigma)R_{l\sigma_k\sigma_i\sigma_j}\right)\rho_i\rho_j\rho_k = (R_{l(kij)} - R_{l(ijk)})\rho_i\rho_j\rho_k = 0.$$
- $j = k$. Then $\rho_{\sigma_i}\rho_{\sigma_j}\rho_{\sigma_k} = -\rho_{\sigma_i}$, and so
$$S_{ikkl} = -\sum_{\sigma} R_{l\sigma_k\sigma_i\sigma_k}\rho_{\sigma_i}.$$

Thus

$$\begin{aligned}\sum_{i,j,k,l} R_{lkij}\rho_i\rho_j\rho_k\rho_l &= \sum_l \sum_{i<k} S_{ikkl}\rho_l \\ &= -\sum_l \sum_{i<j=k} \sum_\sigma R_{l\sigma_k\sigma_i\sigma_k}\rho_{\sigma_i}\rho_l \\ &= -2\sum_l \sum_{i<k} R_{liki}\rho_k\rho_l - R_{lkki}\rho_i\rho_l \\ &= -2\sum_l \sum_{i<k} R_{liki}R_{liki}\rho_k\rho_l + R_{lkik}\rho_i\rho_l \\ &= -2\sum_{i,k,l} R_{liki}\rho_i\rho_l.\end{aligned}$$

Under the exchange $i \leftrightarrow l$, R_{liki} is symmetric, while $\rho_i\rho_l$ is antisymmetric if $i \neq l$. Thus in the sum we just obtained, $-2\sum_{i,k,l} R_{liki}\rho_i\rho_l$, it suffices to sum over $i = l$; we obtain.

$$\sum_{i,j,k,l} R_{lkij}\rho_i\rho_j\rho_k\rho_l = -2\sum_{i,k} R_{ikik} = -2\mathsf{scal}_g$$

which completes the proof. $\square$

Part III. The Seiberg–Witten equations

19 The Seiberg–Witten equations

19.1 Spin^c-structures in 4 dimensions

Let (X,g) be a closed 4-manifold, and let $\mathfrak{s}$ be a Spin^c-structure for (X,g). Thus one has a spinor bundle

$$\mathbb{S} = \mathbb{S}^+ \oplus \mathbb{S}^-,$$

where $\mathbb{S}^+$ and $\mathbb{S}^-$ are rank 2 hermitian vector bundles, and the Clifford map

$$\rho\colon T^*X \to \mathfrak{u}(\mathbb{S}),$$

where $\rho(e)$ exchanges the two summands of $\mathbb{S}$, satisfies $\rho(e)^2 = -|v|^2\,\mathrm{id}$, and the labeling of the summand $\mathbb{S}^\pm$ is compatible with the orientation in the following sense:

Let (e_1,e_2,e_3,e_4) be an oriented orthonormal basis for T^*_xX, and let $\omega = -e_1e_2e_3e_4 \in \mathsf{cl}^0(T^*_xX)$ (ω does not depend on the choice of such basis). Then ω anticommutes with T^*_xX, and so is central in $\mathsf{cl}^0(T^*_xX)$, and $\omega^2 = 1$; soω acts on $\mathbb{S}$ with eigenvalues ± 1, and its eigenspaces, being representations of $\mathsf{cl}^0(T^*_xX)$ exchnaged by T^*_xX, must be $\mathbb{S}^+$ and $\mathbb{S}^-$. The condition is that $\omega = \pm 1$ on $\mathbb{S}^\pm$.

These conditions are sufficient to ensure that ρ is modeled locally, in suitable bases, on left quaternionic multiplication. Thus such data determine a Spin^c-structure.

Lemma 19.1 *There is a canonical isomorphism* $\Lambda^2\mathbb{S}^+ \cong \Lambda^2\mathbb{S}^-$.

Proof Note that

$$\mathsf{Spin}^c(4) = \frac{\mathsf{SU}(2)\times\mathsf{SU}(2)\times\mathsf{U}(1)}{\pm(1,1,1)} \xrightarrow{\cong} G,$$

where

$$G = \{(A,B)\in \mathsf{U}(2)\times\mathsf{U}(2) : \det A = \det B\},$$

and the isomorphism with G is $[A,B,z]\mapsto (zA,zB)$. The spinor bundles are $\mathbb{S}^\pm = \mathsf{Spin}^c(T^*X)\otimes_{G,p_\pm}\mathbb{C}^2$, the bundles associated to the two projections $p_\pm\colon G\to \mathsf{U}(2)$. Recall the homomorphism $\lambda\colon\ \mathsf{Spin}^c(4)\to \mathsf{U}(1)$, $[A,B,z]\mapsto z^2$. Viewed as a homomorphism on G, one has $\lambda = \det\circ p_+ = \det\circ p_-$. Thus the hermitian line bundle

$$L_{\mathfrak{s}} = \mathsf{Spin}^c(T^*X)\otimes_{G,\lambda}\mathbb{C}$$

is identified with both $\Lambda^2\mathbb{S}^+$ and with $\Lambda^2\mathbb{S}^-$. $\square$

We define

$$\det(\mathfrak{s}) = \Lambda^2\mathbb{S}^+.$$

19.2 Spin^c-structures and self-duality

We will need to understand how the Clifford map ρ interacts with self-duality. Note that Clifford multiplication can be defined on 2-forms:

$$\rho\colon \Lambda^2T^*X \to \mathsf{sEnd}^0\mathbb{S} = \operatorname{End}\mathbb{S}^+\times\operatorname{End}\mathbb{S}^-, \quad \rho(e\wedge f) = \tfrac12[\rho(e),\rho(f)].$$

This map should be interpreted as the composite

$$\Lambda^2T^*X \xrightarrow{\cong} \mathfrak{so}(T^*_xX) \xrightarrow{f} \mathfrak{spin}(T^*_xX) \subset \mathsf{Cl}^0(T^*_xX) \xrightarrow{\rho} \mathsf{sEnd}^0\,\mathbb{S},$$

where the first map is the isomorphism $e \wedge f \mapsto \{x \mapsto \langle x,f\rangle e - \langle x,e\rangle f\}$. So the action of 2-forms on spinors is just the action of the spin Lie algebra, in disguise. In particular,

$$\rho(\Lambda^2 T^*X) \subset \mathfrak{su}(\mathbb{S}).$$

Left multiplication by $\omega = -e_1e_2e_3e_4 \in \mathsf{cl}^0(T_x^*X)$ in $\mathsf{cl}^0(T_x^*X)$ preserves $\mathfrak{spin}(T_x^*X)$, and corresponds to the Hodge star on $\Lambda^2 T_x^*X$ (because one checks that $\omega \cdot [e_1, e_2] = [e_3, e_4]$). Thus the ± 1 eigenspaces of $\star$ (namely, $\Lambda_x^\pm$) map under ρ to $\mathbb{S}^\pm$. Thus

$$\rho(\Lambda^+) = \begin{bmatrix} \mathfrak{su}(\mathbb{S}^+) & 0 \\ 0 & 0 \end{bmatrix}, \quad \rho(\Lambda^-) = \begin{bmatrix} 0 & 0 \\ 0 & \mathfrak{su}(\mathbb{S}^-) \end{bmatrix},$$

and the maps $\rho\colon \Lambda^\pm \to \mathfrak{su}(\mathbb{S}^\pm)$ are isomorphisms.

19.3 The configuration space

Let $\mathcal{A}_{cl}(\mathbb{S}^+)$ denote the space of Clifford connections in $\mathbb{S}^+$, and $\mathcal{A}(\det \mathfrak{s})$ the unitary connections in $\Lambda^2\mathbb{S}$. Both are $\Omega^1_X(i\mathbb{R})$-torsors, and one has a map

$$\mathcal{A}_{cl}(\mathbb{S}^+) \to \mathcal{A}(\det \mathfrak{s}), \quad \nabla \mapsto \nabla^\circ,$$

whose effect on connection matrices (in local trivializations) is

$$(\nabla + a\,\mathrm{id}_{\mathbb{S}^+})^\circ = \nabla^\circ + 2a.$$

The bundle of groups $\operatorname{Aut} \mathfrak{s}$ has fibers $\operatorname{Aut} \mathfrak{s}_x$ consisting of all unitary automorphisms u of $\mathbb{S} = \mathbb{S}^+ \oplus \mathbb{S}^-$ such that $\rho(v) \circ u = u \circ \rho(v)$ for all $v \in T_x^*X$. Such a u is necessarily a unit scalar; in particular, it preserves $\mathbb{S}^+$). One has a gauge group $\mathcal{G} = \mathcal{G}(\mathfrak{s}) = \Gamma(\operatorname{Aut} \mathfrak{s})$, which is simply the group $C^\infty(X, \mathsf{U}(1))$. There is a homomorphism

$$\mathcal{G}(\mathfrak{s}) \to \mathcal{G}(\det \mathfrak{s}), \quad u \mapsto \det(u|_{\mathbb{S}^+});$$

viewed as a homomorphism from $C^\infty(X, \mathsf{U}(1))$ to itself, it is simply given by $u \mapsto u^2$.

The Seiberg–Witten configuration space—the domain of the map defining the equations—is

$$\mathcal{A}_{cl}(\mathbb{S}^+) \times \Gamma(\mathbb{S}^+).$$

The gauge group $\mathcal{G}$ acts by $u \cdot (\nabla, \phi) = (u^*\nabla, u\phi) = (u \circ \nabla \circ u^{-1}, u\phi)$.

19.4 The Seiberg–Witten equations

19.4.1 The Dirac equation

We find it convenient now to adopt the standard gauge theory convention of writing A to mean a connection—thus A defines a procedure which defines, in each local trivialization τ of one's bundle, V over an open set U, a 1-form $A_\tau \in \Omega^1_U(\operatorname{End} V)$. We write ∇_A for A viewed as a covariant derivative.

If $A \in \mathcal{A}_{cl}(\mathbb{S}^+)$, we write D_A for the Dirac operator $\rho \circ \nabla_A$, and $D_A^+ \colon \Gamma(\mathbb{S}^+) \to \Gamma(\mathbb{S}^-)$ for its restriction to positive spinors.

The first of the two Seiberg–Witten equations is the *Dirac equation*

$$D_A^+\phi = 0, \quad A \in \mathcal{A}_{cl}(\mathbb{S}^+), \quad \phi \in \Gamma(\mathbb{S}^+). \tag{6}$$

The Dirac operator D_A^+ is the map $\Gamma(\mathbb{S}^+) \to \Gamma(\mathbb{S}^-)$—half of the full Dirac operator D_A.

For a fixed A,the Dirac equation is linear elliptic equation in ϕ, with symbol $\sigma_\xi = \rho(\xi)\colon \mathbb{S}^+ \to \mathbb{S}^-$. (It is not affine-linear as a function of (A, ϕ), however.)

The Dirac equation provides no constraint on A. However, it is invariant under the gauge group $\mathcal{G}$, since

$$D_{u^*A} = \rho \circ \nabla_{u^*A} = \rho \circ (u\nabla_A u^{-1}) = u \circ \rho \circ \nabla_A \circ u^{-1},$$

so

$$D_{u^*A}(u \cdot \phi) = u(D_A\phi).$$

19.4.2 The curvature equation

The second Seiberg–Witten equation—the *curvature equation*—now constraints the gauge-orbit of A:

$$\frac{1}{2}\rho(F(A^\circ)^+) - (\phi\phi^*)_0 = 0 \quad \text{in } i \cdot \mathfrak{su}(\mathbb{S}^+). \tag{7}$$

Explanation of the curvature equation: $F(A^\circ)^+$, the curvature of the connection A° in $\det \mathfrak{s}$, lies in $i\Omega^+$. We view $\rho(F^\circ(A)^+$ as lying in $i\mathfrak{su}(\mathbb{S}^+)$: it is a trace-free, hermitian endomorphism of $\mathbb{S}^+$. Next, the spinor $\phi \in \Gamma(\mathbb{S}^+)$ define a rank-1 endomorphism $\phi\phi^*$ of $\mathbb{S}^+$ by

$$(\phi\phi^*)(\chi) = (\chi, \phi)\phi.$$

Note that the hermitian product is conjugate-linear on right, so if $\psi = u\phi$ for a function u then $\psi\psi^* = |u|^2(\phi\phi^*)$. We denote by $(\cdot)_0$ the trace-free component of an endomorphism of $\mathbb{S}^+$:

$$(\theta)_0 = \theta - \frac{\operatorname{tr}\theta}{2}\mathrm{id}.$$

We have

$$(\phi\phi^*) = \phi\phi^* - \tfrac{1}{2}|\phi|^2\mathrm{id}.$$

The left-hand-sides of the Dirac and curvature equations together define a map

$$\mathfrak{F}\colon \mathcal{A}_{cl}(\mathbb{S}) \times \Gamma(\mathbb{S}^+) \to i\mathfrak{su}(\mathbb{S}^+) \times \Gamma(\mathbb{S}_-),$$

and so Seiberg–Witten equations together read

$$\mathfrak{F}(A, \phi) = 0.$$

For $u \in \mathcal{G}$ one has

$$\mathfrak{F}(u^*A, u\phi) = u \cdot \mathfrak{F}(A, \phi),$$

where u acts on the target by $u \cdot (\theta, \chi) = (\theta, u\chi)$.

We will also frequently need a perturbation of the curvature equation, defined by a self-dual 2-form $\eta \in \Omega^+_X$. The perturbed curvature equation is

$$\frac{1}{2}\rho(F(A^\circ)^+ - 4i\eta) - (\phi\phi^*)_0 = 0. \tag{8}$$

The solution space is $\mathfrak{F}_\eta = 0$, where $\mathfrak{F}_\eta$ is the evident adjustment to $\mathfrak{F}$. It is gauge-invariant in the same sense as the unpertubed equation.

One has

$$\mathfrak{F}_\eta(A + a \cdot \mathrm{id}, \phi + \chi) - \mathfrak{F}_\eta(A, \phi) = \begin{bmatrix} \rho(d^+a) - (\phi\chi^* + \chi\phi^* + \chi\chi^*)_0, \\ D_A\chi + \frac{1}{2}\rho(a)(\phi + \chi) \end{bmatrix}.$$

The linearization of $\mathfrak{F}_\eta$ is therefore

$$(D_{(A,\phi)}\mathfrak{F}_\eta)\begin{bmatrix} a \\ \chi \end{bmatrix} = \begin{bmatrix} \rho\circ d^+ & 0 \\ 0 & D_A \end{bmatrix}\begin{bmatrix} a \\ \chi \end{bmatrix} + \begin{bmatrix} (\phi\chi^* + \chi\phi^*)_0 \\ \frac{1}{2}\rho(a)\phi \end{bmatrix}$$

We often impose, in addition to the SW equations, the Coulomb gauge equation

$$d^*a = 0, \quad a = A^\circ - A_0^\circ \in i\Omega^1_X. \tag{9}$$

Here $A_0 \in \mathcal{A}_{cl}(\mathbb{S})$ is a reference Clifford connection. The linearization of the SW map together with the Coulomb condition is

$$D_{(A,\phi)}(d^*, \mathfrak{F}_\eta)\begin{bmatrix} a \\ \chi \end{bmatrix} = \begin{bmatrix} d^*a \\ \rho\circ d^+ a \\ D_A\chi \end{bmatrix} + \begin{bmatrix} 0 \\ (\phi\chi^* + \chi\phi^*)_0 \\ \frac{1}{2}\rho(a)\phi \end{bmatrix}.$$

The symbol of this operator is the same as that for the first-order term, dropping the zeroth-order term on the right. It is the direct sum of the symbol of $d^* \oplus \rho\circ d^+$ and that of D_A.

19.5 The index

We have not yet discussed the analytical consequences of ellipticity. We will do so in a subsequent lecture, but the bare facts are as follows: Let $\delta\colon \Gamma(E) \to \Gamma(F)$ be an $\mathbb{R}$-linear elliptic operator over a closed manifold M. Let δ^* be the formal adjoint operator, defined with respect to euclidean metrics in E and F. In practice one works in Sobolev spaces, but for the present we can state results using the C^k topology, for any large k.

- $\ker\delta^* \cong \operatorname{coker}\delta$.
- δ is Fredholm: $\operatorname{im}\delta$ is closed and $\ker\delta$ and $\ker\delta^*$ are finite-dimensional. Moreover, these kernels are comprised of C^∞ sections.
- The index $\operatorname{ind}\delta = \dim\ker\delta - \dim\ker\delta^*$ depends only on the symbol σ_δ.

The Atiyah–Singer index theorem gives a formula for $\operatorname{ind}\delta$. In the case of Dirac operators, it reads as follows. Assume M even-dimensional, and suppose that $E = E^+ \oplus E^-$ is a super Clifford module for T^*M (over $\mathbb{C}$). Let $W = \operatorname{End}_{\mathsf{cl}(T^*M)} E$. Let D be $\mathbb{C}$-linear Dirac operator with components $D^\pm\colon \Gamma(E^\pm) \to \Gamma(E^\mp)$. Then

$$\operatorname{ind}_{\mathbb{C}} D^+ = \int_M \hat{A}(TM)\cdot \mathsf{ch}\,(W)$$

where ch is the *Chern character*,

$$\mathsf{ch} = \mathsf{rank}\cdot 1 + c_1 + \frac{1}{2}(c_1^2 - 2c_2) + \cdots \in H^{even}(M),$$

and $\hat{A}(TM) \in H^{4*}(M)$ is the A-hat series, a certain function of the Pontryagin classes $p_k = p_k(TM)$:

$$\hat{A} = 1 - \frac{1}{24}p_1 + \dots.$$

In the case of the spin Dirac operator of a spin structure on a 4-manifold X, W is a trivial complex line bundle, and so

$$\operatorname{ind}_{\mathbb{C}} D^{spin} = \int_X \hat{A}(X) = -\frac{1}{24}\int_M p_1(TX) = -\frac{1}{8}\tau(X).$$

If we now twist the spin structure by a line bundle L to get a $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$ with $\det\mathfrak{s} = L^2$, we have $W = L$ and

$$\operatorname{ind}_{\mathbb{C}} D_A = \int_X \hat{A}(X) = \int_X \left(1 + c_1(L)) + \frac{1}{2}c_1(L)^2\right)\left(1 - \frac{1}{24}p_1(TX)\right) = \frac{1}{8}\left(c_1(\mathfrak{s})^2[X] - \tau(X)\right).$$

The latter formula is valid for Dirac operators in $\mathsf{Spin}^{\mathsf{c}}$ 4-manifolds, regardless of whether a spin structure exists.

We have already studied the elliptic operator

$$\delta = d^* \oplus d^+ \colon \Omega^1_X \to \Omega^0_X \oplus \Omega^+_X$$

using Hodge theory, so we do not need to appeal to the Atiyah–Singer formula. The kernel and cokernel of δ are the odd and even cohomology of the self-duality complex ($\Omega^0 \xrightarrow{d} \Omega^1 \xrightarrow{d^+} \Omega^+$), and therefore its (real) index is

$$\mathrm{ind}_{\mathbb{R}}\, \delta = -\chi(\mathcal{E}^*) = b_1(X) - 1 - b^+(X).$$

Up to zeroth-order terms which do not affect the index, the linearized SW operator with Coulomb gauge fixing is the direct sum of D_A and δ (more precisely, not δ but $d^* \oplus \rho \circ d^+$; the isomorphism $\rho\colon \Lambda^+ \to \mathfrak{su}(\mathbb{S}^+)$ is irrelevant to the index). Thus its index is

$$\mathrm{ind} = \mathrm{ind}_{\mathbb{R}}\, D_A + \mathrm{ind}_{\mathbb{R}}\, \delta = \frac{1}{4}\left(c_1(\mathfrak{s})^2[X] - \tau(X)\right) + b_1(X) - 1 - b^+(X).$$

We arrive at the following formula:

Theorem 19.2 *The index of the linearized SW equation with Coulomb gauge fixing is the number*

$$d(\mathfrak{s}) = \frac{1}{4}\left(c_1(\mathfrak{s})^2[X] - 2\chi(X) - 3\tau(X)\right).$$

19.5.1 A reinterpretation of the Seiberg–Witten index

Lemma 19.3 *Let $S \to M$ be a rank 2 complex vector bundle. Then $c_1(\mathrm{End}\, S) = 0$ and $c_2(\mathrm{End}\, S) = -c_1(S)^2 + 4c_2(S)$.*

Proof Use the splitting principle to reduce to the case where S is a sum of line bundles: $S = \lambda_1 \oplus \lambda_2$. Then

$$\mathrm{End}\, S \cong \mathrm{End}\, \lambda_1 \oplus \mathrm{End}\, \Lambda_2 \oplus \lambda_1^* \otimes \lambda_2 \oplus \lambda_2^* \otimes \lambda_1,$$

and $\mathrm{End}\, \Lambda_i = \mathbb{C}$. Let $\ell_i = c_1(\lambda_i)$. Then $c(S) = (1+\ell_1)(1+\ell_2)$, while $c(\lambda_1^* \otimes \lambda_2) = 1 + \ell_2 - \ell_1$. Thus $c(\mathrm{End}\, S) = c(\lambda_1^*\lambda_2)\cdot c(\lambda_2^*\lambda_1) = (1+\ell_2-\ell_1)(1+\ell_1-\ell_2) = 1-(\ell_1+\ell_2)^2+4\ell_1\ell_2 = 1-c_1(S)^2+4c_2(S)$. □

In the case of the spinor bundle $\mathbb{S}^+ \to X$ of a $\mathsf{Spin}^{\mathsf{c}}$-structure on a 4-manifold, we have $\mathfrak{sl}(\mathbb{S}^+) = \mathfrak{su}(\mathbb{S}^+) \otimes \mathbb{C} \cong \Lambda^+ \otimes \mathbb{C}$. Thus

$$p_1(\Lambda^+) = -c_2(\Lambda^+ \otimes \mathbb{C}) = -c_2(\mathfrak{sl}(\mathbb{S}^+)) = -c_2(\mathrm{End}\, \mathbb{C}) = c_1(\mathbb{S})^2 - 4c_2(\mathbb{S}).$$

Lemma 19.4

$$p_1(\Lambda^+) = p_1(TX) + 2e(TX).$$

Sketch proof. This formula is from Hirzebruch–Hopf, *Felder von Flächenelementen auf vierdimensionaler Mannigfaltigkeiten*.

One can identify real characteristic classes for oriented rank n real vector bundles with the space $\mathbb{C}[\mathfrak{so}(\mathfrak{n})]^{\mathsf{SO}(n)}$ of $\mathsf{SO}(n)$-invariant polynomial on the Lie algebra $\mathfrak{so}(\mathfrak{n})$. The identification is via the Chern–Weil homomorphism, $f \mapsto c_f$. Here f is an invariant polynomial on $\mathfrak{so}(\mathfrak{n})$, and the characteristic class c_f evaluates on the bundle V as $c_f(V) = [f(F)]$: $F = F_A$ is the normalized curvature of a connection A in the bundle V, and $f(F)$ the resulting real-valued closed 2-form.

The Chern–Weil homomorphism is natural in the following sense: if $\theta\colon\ \mathsf{SO}(m) \to \mathsf{SO}(n)$ is a homomorphism, one has $c_{\theta^* f}(V) = c_f(P_V \times_\theta \mathbb{R}^n)$. Here P_V is the principal frame bundle of the $\mathsf{SO}(m)$-bundle V, and $P_V \times_\theta \mathbb{R}^n$ the associated bundle.

In the case of an $\mathsf{SO}(n)$-bundle, one has $p_1 = [\operatorname{tr} F^2]$ and $e = \det F$.

Now, Λ^+ arises from the principal $\mathsf{SO}(4)$-bundle P_{TX} via the homomorphism $\lambda^+\colon\ \mathsf{SO}(4) \to \mathsf{SO}(3)$. The invariant polynomial $p_1(F) = \operatorname{tr} F^2$ on $\mathfrak{so}(4)$ pulls back under λ^+ to an invariant polynomial $q(F) = p_1((\lambda^+)^* F)$ on $\mathfrak{so}(3)$. On the Lie algebra of the maximal torus $T = \mathsf{SO}(2) \times \mathsf{SO}(2) \subset \mathsf{SO}(4)$, one can write $p_1(t_1, t_2) = t_1^2 + t_2^2$, with t_i a coordinate on the ith copy of $\mathfrak{so}(2)$. But λ^+ maps T to $T' = \mathsf{SO}(2) \times \{1\} \subset \mathsf{SO}(3)$, and if t' is the coordinate on the Lie algebra of T', one checks that $(\lambda^+)^* t' = t_1 + t_2$. Thus $(\lambda^+)^* p_1(t') = (t_1 + t_2)^2 = (t_1^2 + t_2^2) + 2t_2 t_2$, i.e., $q(F) = \operatorname{tr} F^2 + 2 \det F$. Thus $p_1(\Lambda^+) = p_1(TX) + 2e(TX)$. □

Thus
$$c_1(\mathbb{S}^+)^2 - 4c_2(\mathbb{S}^+) = p_1(TX) + 2e(TX),$$
from which we see
$$c_2(\mathbb{S}^+)[X] = \frac{1}{4}\left(c_1(\mathbb{S}^+)^2[X] - 2\chi(X) - 3\sigma(X)\right) = d(\mathfrak{s}).$$
The second Chern number of $\mathbb{S}^+$ is also its Euler number—the signed count of zeros of a generic positive spinor.

Remark. The fact that the Seiberg–Witten index $d(\mathfrak{s})$ is the Euler number of $\mathbb{S}^+$ cries out for a direct explanation. There *is* such an explanation: see M. Maridakis, *Spinor pairs and the concentration principle for Dirac operators*, arXiv 1510.07004. It draws on ideas of Witten and of Taubes, in particular of Taubes's approach to the Riemann–Roch formula, for which see C. Wendl, *Lectures on Holomorphic Curves in Symplectic and Contact Geometry*.

20 The Seiberg–Witten equations: bounds

The space of solutions to the Seiberg–Witten equations is *compact*: for any sequence of solutions (A_i, ϕ_i) there is a sequence of gauge transformations u_i such that $(u_i^* A_i, u_i \phi_i)$ converges to a smooth limiting solution. Today we begin working towards a proof of that remarkable result.

20.0.2 An inequality for Laplacians

Let (M, g) be an oriented Riemannian manifold. Let ∇ be an orthogonal covariant derivative in a real, euclidean vector bundle $E \to M$. In Lecture 17, we saw that ∇ has a formal adjoint operator Define $\nabla^* \colon \Gamma(T^*M \otimes E) \to \Gamma(E)$.

The first two parts of the following lemma are restatements from Lecture 17, wherein they were proved:

Lemma 20.1

$$\tfrac{1}{2} d^* d(|s|^2) = \langle \nabla^* \nabla s, s \rangle - |\nabla s|^2.$$

Proof For functions f of compact support, we have

$$\begin{aligned}
\int_M f \langle \nabla^* \nabla s, s \rangle \mathsf{vol} &= \int_M \langle \nabla^* \nabla s, fs \rangle \mathsf{vol} \\
&= \int_M \langle \nabla s, \nabla(fs) \rangle \mathsf{vol} \\
&= \int_M \langle \nabla s, f \nabla s + df \otimes s \rangle \mathsf{vol} \\
&= \int_M f |\nabla s|^2 \mathsf{vol} + \int_M \langle \nabla s, df \otimes s \rangle \mathsf{vol}.
\end{aligned}$$

(Angle-brackets are used to denote the inner products both on E and on $T^*X \otimes E$.) It suffices, then, to show that $\int_M \langle \nabla s, df \otimes s \rangle \mathsf{vol} = \int_M \frac{1}{2} d^* d(|s^2|) f \, \mathsf{vol}$.

Since ∇ is an orthogonal connection, $d(|s|^2) = 2\langle \nabla s, s \rangle$, and

$$g\left(d(|s|^2), df\right) = 2\langle \nabla s, df \otimes s \rangle.$$

Thus

$$\int_M \langle \nabla s, df \otimes s \rangle \mathsf{vol} = \int_M \tfrac{1}{2} g\left(d|s|^2, df\right) \, \mathsf{vol} = \int_M \tfrac{1}{2} d^* d(|s^2|)) \cdot f \, \mathsf{vol}.$$

□

We deduce the following very useful inequality:

Proposition 20.2 *We have*

$$\tfrac{1}{2} d^* d(|s|^2) \leq \langle \nabla^* \nabla s, s \rangle.$$

Lemma 20.3 *If p is a local maximum of the function f then $(d^* df)(p) \geq 0$.*

Proof Recall that $d^* = \pm \star d \star$. One can choose local coordinates $(x_1, \ldots, x_n)$ in which the metric is $g = g_{std} + \sum_{r_{ij}} x_i x_j$ for smooth functions r_{ij}, g_{std} being the usual metric on $\mathbb{R}^n$. The Hodge star differs from that of g_{std} by a terms involving $x_i x_j$. Consequently, the action of d^* on 1-forms takes the following shape:

$$d^*(a_i\, dx_i) = -\partial_i a + \sum_j h_{ij} x_j a_i,$$

for smooth functions h_{ij}. So, for functions f, one has $d^* df = -\sum \partial_i^2 f + \sum h_{ij} x_j (\partial_i f)$. Thus

$$(d^* df)(0) = -\sum (\partial_i^2 f)(0),$$

so the lemma follows from the corresponding statement for the Laplacian $-\sum \partial_i^2$. □

20.1 *A priori* bounds for solutions to the SW equations

The following lemma is an easy calculation:

Lemma 20.4 *For $\phi \in \Gamma(\mathbb{S}^+)$, one has $((\phi\phi^*)_0 \chi, \chi) = |\chi|^2 (\chi, \phi) - \frac{1}{2}|\chi|^2|\phi|^2$. In particular,*

$$\big((\phi\phi^*)_0 \phi, \phi\big) = \frac{1}{2}|\phi|^4.$$

Lemma 20.5 *For $\eta \in \Omega^2_X$, and $\phi \in \Gamma(\mathbb{S})$, one has $(\rho(\eta)\phi, \phi) \le |\eta||\phi|^2$.*

Proof It suffices to take $\eta = e \wedge f$ for orthogonal unit vectors e and f. One then has $(\rho(\eta)\phi, \phi) = \frac{1}{2}([\rho(e), \rho(f)]\phi, \phi) = (\rho(e)\phi, \rho(f)\phi) \le |\rho(e)\phi| \cdot |\rho(f)\phi| \le |\phi|^2$. □

We return now to the Seiberg–Witten equations

$$D_A^+ \phi = 0, \quad \tfrac{1}{2}\rho(F(A^\circ)^+ - 4i\eta) - (\phi\phi^*)_0 = 0.$$

The left hand side defines the expression $\mathfrak{F}_\eta(A, \phi)$.

Lemma 20.6 (basic pointwise estimate) *If $\mathfrak{F}_\eta(A, \phi) = 0$ then*

$$d^* d|\phi|^2 + \tfrac{1}{2}(\mathsf{scal}_g - 8|\eta|)|\phi|^2 + |\phi|^4 \le 0, \tag{10}$$

Proof We use the inequality $\frac{1}{2} d^* d(|s|^2) \le \langle \nabla^* \nabla s, s\rangle$, followed by the Lichérowicz formula, to obtain

$$\begin{aligned}\tfrac{1}{2} d^* d|\phi|^2 &\le (\nabla_A^* \nabla_A \phi, \phi) \\ &= \left(D_A^- D_A^+ \phi - \frac{1}{4}\mathsf{scal}_g \phi - \frac{1}{2}\rho(F(A^\circ)^+)\phi, \phi\right).\end{aligned}$$

Now impose the two SW equations to get

$$\begin{aligned}\tfrac{1}{2} d^* d|\phi|^2 &\le -\frac{1}{4}\mathsf{scal}_g|\phi|^2 - \big(2i\rho(\eta)\phi + (\phi\phi^*)_0\phi, \phi\big) \\ &\le -\frac{1}{4}\mathsf{scal}_g|\phi|^2 - \frac{1}{2}|\phi|^4 + 2|\eta||\phi|^2,\end{aligned}$$

where in the second step we use the two lemmas above. □

Define

$$s = \max(8|\eta| - \mathsf{scal}_g, 0):$$

a continuous, non-negative function on X.

Theorem 20.7 (pointwise bound on ϕ) *If $\mathfrak{F}_\eta(A,\phi)=0$ then*
$$\max|\phi|^2 \le \tfrac{1}{2}\max s.$$

Proof We have
$$d^*d|\phi|^2 + |\phi|^4 \le \frac{s}{2}|\phi|^2.$$
At a point x where $|\phi(x)|^2$ is maximized, one has $(d^*d|\phi|^2)(x) \ge 0$ (as we showed above), so
$$|\phi(x)|^4 \le \frac{s(x)}{2}|\phi(x)|^2.$$
If $|\phi(x)|=0$ then the claimed result is trivially true; if not, we have
$$|\phi|^2 \le |\phi(x)|^2 \le \frac{1}{2}s(x) \le \frac{1}{2}\max s.$$
$\square$

It is worth noting that the calculations that led to this bound were sensitive to the precise form of the equations. If one flipped the sign of the quadratic term $(\phi\phi^*)_0$, we would not be able to obtain a bound on $|\phi|$ (though of course ellipticity would be unaffected).

We can immediately make the following conclusion:

Corollary 20.8 *If $\mathsf{scal}_g \ge 0$ then the only solutions to the unperturbed equation $\mathfrak{F}(A,\phi)=0$ are those with $\phi=0$.*

For in this case, $s=0$.

Note that the solutions to $\mathfrak{F}_\eta=0$ with $\phi=0$ are the Clifford connections A such that $F(A^\circ)^+-4i\eta=0$. Taking $\eta=0$, we get the equation for abelian instantons, $F(A^\circ)^+=0$, which we analyzed some time ago.

Having obtained pointwise control on $|\phi|$, the next step is to control $|F(A^\circ)^+|$:

Proposition 20.9 *If $\mathfrak{F}_\eta(A,\phi)=0$ then*
$$|F(A^\circ)^+ - 4i\eta| \le \frac{1}{4}\max s.$$

Proof We have $\rho(F(A^\circ)^+-4i\eta)=(\phi\phi^*)_0$, which implies that $|F(A^\circ)^+-4i\eta| \le |\rho(F(A^\circ)^+-4i\eta)|_{op} \le |(\phi\phi^*)_0|_{op} \le \frac{1}{2}|\phi|^2 \le \frac{1}{4}\max s$. $\square$

20.2 Finiteness

Proposition 20.10 *Among those Spin^c-structures $\mathfrak{s}$ with $d(\mathfrak{s}) \ge d_0$, only finitely many isomorphism classes contain solutions to $\mathfrak{F}=0$. (Here the metric g is fixed, and $\eta=0$.)*

Proof $c_1(\mathfrak{s})$ is represented by $(i/2\pi)F(A^\circ)$. Writing $F=iF(A^\circ)$, we have
$$c_1(\mathfrak{s})^2[X] = \frac{1}{4\pi^2}\int_X F\wedge F = \frac{1}{4\pi^2}\int_X (F^+\wedge F^+ + F^-\wedge F^-) = \frac{1}{4\pi^2}\int_X (|F^+|^2 - |F^-|^2)\mathsf{vol}_g.$$

Thus

$$\frac{1}{4\pi^2}\int_X |F|^2 \mathsf{vol}_g = \frac{1}{4\pi^2}\int_X (|F^+|^2 + |F^-|^2)\mathsf{vol}_g$$
$$= -c_1(\mathfrak{s})^2[X] + \frac{1}{2\pi^2}\int_X |F^+|^2 \mathsf{vol}_g.$$

Set $S = \max s$. Integrating the pointwise bound on F^+, we have $\int_X |F^+|^2 \mathsf{vol}_g \leq \frac{S^2}{4}\mathsf{vol}(X)$; so

$$\frac{1}{4\pi^2}\int_X |F|^2 \mathsf{vol}_g \leq -c_1(\mathfrak{s})^2[X] + \frac{S^2\cdot \mathsf{vol}(X)}{8\pi^2}.$$

Now, $d(\mathfrak{s}) = \frac{1}{4}(c_1(\mathfrak{s})^2[X] - 2\chi - 3\tau) \geq d_0$, so $c_1(\mathfrak{s})^2[X] = 4d(\mathfrak{s}) + 2\chi + 3\tau \geq 4d_0 + 2\chi + 3\tau$. Thus

$$\frac{1}{4\pi^2}\int_X |F|^2 \mathsf{vol}_g \leq -4d_0 - 2\chi - 3\tau + \frac{S^2\cdot \mathsf{vol}(X)}{8\pi^2},$$

a bound dependent only on (X,g) and d_0.

The finite-dimensional vector space $H^2_{DR}(X)$ has a norm: $\|c\|$ is the infimum of the L^2-norm of closed 2-forms representing c (or equally the L^2-norm of c_{harm}). When a solution to $\mathfrak{F} = 0$ exists, we have

$$\|c_1(\mathfrak{s})\|^2 \leq \|\frac{1}{2\pi}F\|^2_{L^2} \leq -4d_0 - 2\chi - 3\tau + \frac{S^2\cdot \mathsf{vol}(X)}{8\pi^2},$$

So $c_1(\mathfrak{s})$ lies in the intersection of the integer lattice with a ball. Thus there are only finitely many options for $c_1(\mathfrak{s})$. But Spin^c-structures with $c_1(\mathfrak{s})$ fixed are in bijection with the 2-torsion in $H^2(X;\mathbb{Z})$, a finite group. Thus only finitely many $\mathfrak{s}$ with $d \geq d_0$ admit solutions. □

Remark. When $\eta \neq 0$, there can be infinitely many Spin^c-structures with solutions to $\mathfrak{F}_\eta = 0$. We will find that this is true for $\mathbb{C}P^2$, when $\int \eta^2$ is sufficiently large.

21 The compactness theorem

21.1 Statement of the theorem

We work as usual over a closed, oriented Riemannian 4-manifold (X, g), with a chosen Spin^c-structure $\mathfrak{s}$, and a chosen self-dual 2-form η.

Theorem 21.1 *Let (A_j, ϕ_j) be any sequence of solutions to the Seiberg–Witten equation $\mathfrak{F}_\eta = 0$. Assume (i) that A_j is in Coulomb gauge relative to A_0, i.e., $A_j = A_0 + a_j \cdot \mathrm{id}$ with $d^* a_j = 0$; and (ii) that the harmonic projections $(a_j)_{harm}$ form a bounded sequence in $\mathcal{H}^1_g(X)$ (with respect to its L^2-norm). Then a subsequence converges in C^∞ to a smooth limiting solution (A, ϕ).*

Note that both assumptions (i) and (ii) can be achieved by gauge transformations. One satisfies (ii) by using topologically non-trivial gauge transformations to make sure that $(a_j)_{harm}$ lies in a fundamental domain for the action of $H^1(X;\mathbb{Z})$ on $\mathcal{H}^1_g(X)$. One then acts by gauge transformations in the identity component of $\mathcal{G}$ to satisfy (ii), noticing that these have no effect on $(a_j)_{harm}$. Thus:

Corollary 21.2 *Solutions to $\mathfrak{F}_\eta = 0$ project to a compact of the space of configrations (A, ϕ) modulo gauge transformations.*

(We shall look more closely at the space of configurations mod gauge in the subsequent lecture. Its C^∞ topology arises from a metric, so the use of sequential compactness is legitimate.)

My exposition leans heavily on the one in Kronheimer–Mrowka's book *Monopoles and 3-manifolds*, which covers the generalization of the compactness theorems to manifold with boundary.

21.2 Sobolev multiplication

We will work over our closed, oriented Riemannian 4-manifold (X, g). We shall be using the Sobolev spaces $L^p_k(X)$. Among the equivalent choices for Sobolev norm, the one we find convenient here is one defined by a metric g and its Levi-Civita connection ∇:

$$\|f\|^p_{p,k} = \int_X (|f|^p + |\nabla f|^p + \cdots + |\nabla^{(k)} f|^p)\mathrm{vol}_g.$$

Here, $\nabla f = df$, and for higher j, $\nabla^{(j)} f \in \Gamma((T^*X)^{\otimes j})$ is the j-fold iterated covariant derivative. The pointwise norm $|\nabla^j f|$ is defined via the tensor product metric $g^{\otimes j}$.

We note that the scaling weight for L^p_k in 4 dimensions is $w(p,k) = k - 4/p$; thus $w(2,k) = k-2$, and $w(0,p) = -4/p$, and $w(p,1) = 1 - 4/p$. Thus we have embeddings

$$L^2_k \to C^0, \quad k \geq 3;$$

and

$$L^2_1 \to L^4, \quad L^2_2 \to L^4_1 \to L^p; \quad L^3_1 \to L^{12},$$

for example.

Lemma 21.3 *Ove X^4, multiplication of smooth functions extends to a bounded map*

$$L^2_1(X) \otimes L^2_1(X) \to L^4(X).$$

Proof There is a bounded inclusion $L^2_1 \to L^4$; the product fg of L^4 functions is L^2. □

Lemma 21.4 *Over X^4, multiplication of smooth functions extends to a bounded map*

$$L^2_k(X) \otimes L^2_l(X) \to L^2_l(X)$$

provided that $k \geq 3$ and $k \geq l$. In particular, $L^2_k(X)$ is an algebra for $k \geq 3$.

Proof The proof is based on the previous lemma, and the bounded embedding $L^2_k \to C^0$ valid for $k \geq 3$. We want to prove

$$\|ab\|_{2,l} \leq \mathsf{const.}\|a\|_{2,k}\|b\|_{2,l}.$$

Now, $\nabla^{(m)}(ab) = \sum_i \binom{m}{i} \nabla^{(i)}a \otimes \nabla^{(m-i)}b$, so

$$\|ab\|^2_{2,l} \leq \sum_{i+j\leq l} \binom{i+j}{i} \|\nabla^{(i)}a \otimes \nabla^{(j)}b\|_2.$$

We have

$$\|(\nabla^{(k)}a) \otimes b\|_2 \leq \|b\|_{C^0} \cdot \|a\|_{2,k} \leq \|b\|_{C^0}\|a\|_{2,l} \leq \mathsf{const.}\|a\|_{2,k}\|b\|_{2,l}.$$

One similarly handles terms $\|a \otimes \nabla^{(l)}b\|_2$. For each of the remaining terms $\|\nabla^{(i)}a \otimes \nabla^{(j)}b\|_2$, one has $i \leq k-1$ and $j| \leq l-1$. The L^2_1 norms of $\nabla^{(i)}f$ and $\nabla^{(j)}b$ are bounded respectively by constants times $\|a\|_{2,k}$ and $\|b\|_{2,l}$, and so by (i), $\|\nabla^{(i)}a \otimes \nabla^{(j)}b\|_2 \leq \mathsf{const.}\|a\|_{2,k} \cdot \|b\|_{2,l}$. □

It follows that, for euclidean vector bundles E and F, one has a bounded multiplication

$$L^2_k(E) \otimes L^2_l(F) \to L^2_l(E \otimes F), \quad k \geq 3, \quad k \geq l.$$

There are also bounded multiplication maps for the lower-regularity Sobolev spaces in 4 dimensions, but these bring in Sobolev spaces with $p > 2$. Specifically, we shall use the following instances:

Lemma 21.5 *There are bounded multiplications*

$$L^2_1 \otimes L^3_1 \to L^3, \tag{11}$$

$$L^2_2 \otimes L^3_1 \to L^2_1, \tag{12}$$

$$L^2_3 \otimes L^2_2 \to L^2_2. \tag{13}$$

Proof (1) One has a tensor product of Sobolev embeddings $L^2_1 \otimes L^3_1 \to L^4 \otimes L^{12}$. Hölder's inequality, $\|uv\|_1 \leq \|u\|_{4/3} \cdot \|v\|_4$, applied to $u = |a|^3$ and $v = |b|^3$, now gives $\|ab\|^3_3 \leq \|a\|^{3/4}_4 \cdot \|b\|^{1/4}_{12}$, so we have an embedding $L^4 \otimes L^{12} \to L^3$.

I omit the proofs of (2) and (3). □

21.3 The Seiberg–Witten equations in Sobolev spaces

From the fact that L^2_k is an algebra for $k \geq 3$, with continuous multiplication, we see that there is a topological group $\mathcal{G}_{2,k}$ of Spin^c gauge transformations of class L^2_k. By the same token, it acts continuously on L^2_k configurations (A, ϕ). (Here A differs from a reference A^∞ connection A_0 by an L^2_k 1-form)—again, for any $k \geq 3$.

We can also formulate the Seiberg–Witten map for $\mathfrak{s}$ in Sobolev spaces. It is convenient to build in Coulomb gauge fixing. Thus we fix $A_0 \in \mathcal{A}_{cl}(\mathbb{S}^+)$ and consider the Seiberg–Witten map $\mathfrak{F}_\eta$ with the Coulomb gauge condition built in:

$$\mathfrak{F}'_\eta \colon L^2_k(iT^*X) \times L^2_k(\mathbb{S}^+) \to L^2_{k-1}(X)_0 \times L^2_{k-1}(i\mathfrak{su}(\mathbb{S}^+)) \times L^2_k(\mathbb{S}^-).$$

Here the subscript zero in $L^2_{k-1}(X)_0$ means functions of mean zero; and

$$\mathfrak{F}'_\eta(a,\phi) = (d^*a, \mathfrak{F}_\eta(A,\phi)).$$

The point is that the quadratic terms $\rho(a)\phi$ and $(\phi\phi^*)_0$ in $\mathfrak{F}_\eta(A,\phi))$ lie in L^2_k because of the Sobolev mutliplication property, while the differential operators take L^2_k to L^2_{k-1}.

We can write $\mathfrak{F}'_\eta$ as follows:

$$\mathfrak{F}'_\eta(a,\phi) = \mathfrak{D}(a,\phi) + \mathfrak{q}(a,\phi) + \mathfrak{c},$$

where

$$\mathfrak{D}(a,\phi) = \begin{bmatrix} d^*a \\ \rho\circ d^+a \\ D^+_{A_0}\phi \end{bmatrix}, \quad \mathfrak{q}(a,\phi) = \begin{bmatrix} 0 \\ -(\phi\phi^*)_0 \\ \frac{1}{2}\rho(a)\phi \end{bmatrix}, \quad \mathfrak{c} = \mathfrak{c}(\eta, A_0) = \begin{bmatrix} 0 \\ \frac{1}{2}\rho(F(A_0^\circ)^+ + 4i\eta) \\ 0 \end{bmatrix}.$$

Notice that $\mathfrak{D}$ is a first-order linear elliptic operator, while $\mathfrak{q}$ is quadratic, and $\mathfrak{c}$ a constant.

21.4 Elliptic estimates and the proof of compactness

21.4.1 The positive feedback loop

The feedback loop is usually called elliptic bootstrapping, and it works as follows:

Proposition 21.6 *Fix some $k \geq 3$, and consider a set of solutions $\gamma = (a,\phi)$ to $\mathfrak{F}'_\eta\gamma = 0$ for which there is a uniform bound $\|\gamma\|_{2,k} \leq c_k$. There is then also a bound*

$$\|\gamma\|_{2,k+1} \leq c_{k+1}.$$

Proof The elliptic estimate for $\mathfrak{D}$ gives a bound

$$\|\gamma\|_{2,k+1} \leq c'_{k+1}\left(\|\mathfrak{D}\gamma\|_{2,k} + \|\gamma\|_{2,k}\right).$$

Thus using the equations and the postulated L^2_k bound, we get

$$\|\gamma\|_{2,k+1} \leq c'_{k+1}\left(\|\mathfrak{q}(\gamma) + \mathfrak{c}\|_{2,k} + c_k\right).$$

From the form of $\mathfrak{q}_\eta(a,\phi)$ and the Sobolev multiplication in L^2_k, we get

$$\|\mathfrak{q}_\eta(\gamma)\|_{2,k} \leq c''_k(1 + \|\gamma\|^2_{2,k}) \leq 2c''_k(1 + c_k^2).$$

The result follows. □

From the compact embeddings $C^l \to L^2_k$ (where $k \geq l+4$) we deduce:

Corollary 21.7 *If $\{\gamma_j = (a_j,\phi_j)\}$ is a sequence of solutions to $\mathfrak{F}'_\eta\gamma = 0$, converging in L^2_3 to a limit (a,ϕ) of class L^2_3, then (a,ϕ) is C^∞ and for each l there is a subsequence converging to (a,ϕ) in C^l.*

It is in fact unnecessary to pass to a subsequence; we will return to this point later.

21.5 L^2_1 bounds

Recall our *a priori* bounds on solutions: there is a constant $\kappa = \kappa(X, g, \eta)$ such that for any $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$, and any (A, ϕ) such that $\mathfrak{F}_\eta(A, \phi) = 0$, one has

$$\|\phi\|^2_{C^0} \leq \kappa, \quad \|F(A^\circ)^+\|_{C^0} \leq \kappa.$$

We now select a reference Clifford connection $A_0 \in \mathcal{A}_{cl}(\mathbb{S}^+)$, and write $A = A_0 + a\,\mathrm{id}_{\mathbb{S}}$, where $a \in i\Omega^1_X$, and assume $\mathfrak{F}'_\eta(a, \phi) = 0$, i.e., we have a SW solution with $d^*a = 0$. There is an L^2 orthogonal decomposition

$$a = a_{harm} + a',$$

with a_{harm} harmonic and $a' \in \mathrm{im}\, d^*$.

Lemma 21.8 *There is a uniform bound*

$$\|a'\|_{2,1} \leq K.$$

on solutions (a, ϕ) to $\mathfrak{F}'_\eta(a, \phi) = 0$. Here K depends on (X, g, η) and on A_0.

Note: There are similar bounds on $\|a'\|_{p,1}$ for any $p > 1$.

Proof We have

$$F(A^\circ)^+ = F(A_0^\circ)^+ + 2d^+a,$$

so there is a pointwise bound

$$|d^+a| \leq \tfrac{1}{2}\kappa + c(A_0),$$

where $c(A_0) = \max |F(A_0^\circ)^+|$.

The operator $d^* \oplus d^+$ is elliptic, with kernel $\mathcal{H}^1_g$ the harmonic 1-forms. On the complement $(\mathcal{H}^1)^\perp = \mathrm{im}\, d \oplus \mathrm{im}\, d^* \subset \Omega^1_X$, one has an elliptic estimate

$$\|b\|_{2,1} \leq C\|(d^* \oplus d^+)b\|_2.$$

(This is the one place that we shall make use of the sharp form of the elliptic estimate, valid on the orthogonal complement to the kernel of the elliptic operator.) Thus

$$\|a'\|_{2,1} \leq C\|(d^* \oplus d^+)a\|_2 \leq C\left(\tfrac{1}{2}\kappa + c(A_0)\right) \mathsf{vol}(X)^{1/2},$$

which is a bound of the kind we are seeking. □

Lemma 21.9 *Consider a set of solutions $\gamma = (a, \phi)$ to $\mathfrak{F}'_\eta(a, \phi) = 0$ such that a_{harm} is bounded in $H^1_{DR}(X)$. There are then bounds*

$$\|\gamma\|_{2,1} \leq K'.$$

Proof The assumption $a_{harm} \leq C$, and the previous lemma, together show that $\|a\|_{2,1} \leq K + C$. We then use the Dirac equation $D^+_{A_0}\phi + \frac{1}{2}\rho(a)\phi = 0$, and the elliptic estimate for $D^+_{A_0}$, to get $\|\phi\|_{2,1} \leq K''(\|\rho(a)\phi\|_2 + \|\phi\|_2) \leq K''(\kappa\|\rho(a)\|_2 + \kappa\mathsf{vol}(X)^{1/2})$, so the result following from the bound on $\|a\|_2$. □

21.5.1 From L^2_1 to L^2_3

The question now is how to pass from L^2_1 bounds to L^2_3 convergence. The strategy we shall use has two stages: first, we show that L^2_1 convergence implies L^2_3 convergence, and then we show how to obtain L^2_1 convergence from L^2_1 bounds.

The proof that L^2_1 convergence implies L^2_3 convergence is based on the boundedness of the multiplication maps (1–3) asserted earlier:

Lemma 21.10 *Consider a sequence of solutions $\gamma_j = (a_j, \phi_j)$ to $\mathfrak{F}'_\eta = 0$, converging in L^2_1 to a limit $\gamma = (a, \phi)$ of class L^2_1. Then γ_j converges in L^2_3, and γ is of class L^2_3.*

Proof The L^3_1 elliptic estimate for $\mathfrak{D}$ and the equations give

$$\|\gamma_i - \gamma_j\|_{3,1} \leq C(\|\mathfrak{q}(\gamma_i) - \mathfrak{q}(\gamma_j)\|_3 + \|\gamma_i - \gamma_j\|_3).$$

This is the one place that we will invoke an elliptic estimate for $p \neq 2$.

The sequence γ_j is Cauchy in L^2_1. So, fixing $\epsilon > 0$, we can find i_0 such that $\|\gamma_i - \gamma_{i_0}\|_{L^2_1} \leq \epsilon$ for $i \geq i_0$.

Now, $\mathfrak{q}$ is the quadratic form associated with a symmetric bilinear form $\mathfrak{b}$ (defined pointwise over X). We have $\mathfrak{q}(\gamma_i) - \mathfrak{q}(\gamma_j) = \mathfrak{b}(\gamma_i - \gamma_j, \gamma_i + \gamma_j) = \mathfrak{b}(\gamma_i - \gamma_j, \gamma_i + \gamma_j - 2\gamma_{i_0}) + 2\mathfrak{b}(\gamma_i - \gamma_j, \gamma_{i_0})$, so

$$\|\gamma_i - \gamma_j\|_{3,1} \leq C'' \left(\|\mathfrak{b}(\gamma_i - \gamma_j, \gamma_i + \gamma_j + 2\gamma_{i_0})\|_3 + 2\|\gamma_{i_0}\|_{C^0} \|\gamma_i - \gamma_j\|_3 + \|\gamma_i - \gamma_j\|_3 \right).$$

Using the bounded multiplication $L^2_1 \times L^3_1 \to L^3$, we turn this into bounds

$$\begin{aligned} \|\gamma_i - \gamma_j\|_{3,1} &\leq C'' \left(\|\gamma_i - \gamma_j\|_{3,1} \|\gamma_i + \gamma_j + 2\gamma_{i_0}\|_{2,1} + 2\|\gamma_{i_0}\|_{C^0} \|\gamma_i - \gamma_j\|_3 + \|\gamma_i - \gamma_j\|_3 \right) \\ &\leq C'' \left(2\epsilon \|\gamma_i - \gamma_j\|_{3,1} + 2\|\gamma_{i_0}\|_{C^0} \|\gamma_i - \gamma_j\|_3 + \|\gamma_i - \gamma_j\|_3 \right), \end{aligned}$$

so rearranging,

$$(1 - 2C''\epsilon)\|\gamma_i - \gamma_j\|_{3,1} \leq C''(2\|\gamma_{i_0}\|_{C^0} + 1)\|\gamma_i - \gamma_j\|_3.$$

Taking $\epsilon = 1/(4C'')$, and fixing a corresponding i_0, we can treat $\|\gamma_{i_0}\|_{C^0}$ as a constant; so

$$\|\gamma_i - \gamma_j\|_{3,1} \leq C''' \|\gamma_i - \gamma_j\|_3 \leq C'''' \|\gamma_i - \gamma_j\|_{2,1},$$

via the embedding $L^3 \to L^2_1$. Thus (γ_j) is Cauchy in L^3_1.

We now repeat virtually the same argument to see that (γ_j) is Cauchy in L^2_2. This time, we use the L^2_2 elliptic estimate, the bounded multiplication $L^2_2 \times L^3_1 \to L^2_1$, and we bound $\|\mathfrak{b}(\gamma_i - \gamma_j, \gamma_{i_0})\|_{2,1}$ by a constant times $\|\gamma_{i_0}\|_{C^1} \cdot \|\gamma_i - \gamma_j\|_{2,1}$.

Repeat once again to see that (γ_j) is Cauchy in L^2_3. Use the L^2_3 elliptic estimate, boudned multiplication $L^2_3 \times L^2_2 \to L^2_2$,, and bound $\|\mathfrak{b}(\gamma_i - \gamma_j, \gamma_{i_0})\|_{2,2}$ by a constant times $\|\gamma_{i_0}\|_{C^2} \cdot \|\gamma_i - \gamma_j\|_{2,2}$. The result now follows. □

A very similar argument proves the following variant on the positive feedback loop:

Lemma 21.11 *If a sequence of solutions (a_j, ϕ_j) to $\mathfrak{F}'_\eta = 0$ converges in L^2_3 then it converges in L^2_k for all k.*

21.5.2 Weak and strong L^2_1 convergence

We recall a general principle of Hilbert space theory. Let H be a separable Hilbert space, and (x_n) a sequence in H. We say that x_n converges weakly to x if $\langle x_n - x, y\rangle \to 0$ for all $y \in H$. We shall write $x_n \rightsquigarrow x$. Weak limits are unique, and of course ordinary ('strong') limits are also weak limits.

Lemma 21.12 *If (x_n) is a bounded sequence in H then it has a weakly convergent subsequence $x'_n \rightsquigarrow x$. One has $\|x\| \leq \liminf \|x'_n\|$. If $\|x_n\| \to \|x\|$ then $x'_n \to x$ (i.e. the convergence is strong).*

Proof The functional $y \mapsto \liminf\langle x_n, y\rangle$ is bounded—its norm is at most $\liminf \|x_n\|$—and so, by Riesz's representation lemma, is represented as $\langle x, \cdot\rangle$ for a unique $x \in H$. Pass to a subsequence x'_n such that $\lim \|x'_n\| = \liminf \|x_n\|$. Then $x'_n \rightsquigarrow x$. Uniqueness follows from non-degeneracy of the inner product. The fact that $\|x\| = \sup_{\|y\|=1}\langle x, y\rangle$ shows that the norm of a weak limit is bounded above by $\liminf \|x_n\|$. If $\|x_n\| \to \|x\|$ then $\|x_n - x\|^2 = |x_n|^2 + |x|^2 - 2\langle x_n, x\rangle \to 0$, so $x_n \to x$. □

Lemma 21.13 *If $L\colon H \to H'$ is a bounded linear map between Hilbert spaces, and if $x_n \rightsquigarrow x$ in H, then $Lx_n \rightsquigarrow Lx$ in H'.*

Proof $\langle Lx_n, y\rangle_{H'} = \langle x_n, L^*y\rangle_H \to \langle x, L^*y\rangle_H = \langle Lx, y\rangle_{H'}$. □

Lemma 21.14 *If a sequence of solutions $\gamma_j = (a_j, \phi_j)$ to $\mathfrak{F}'_\eta = 0$ converges weakly in L^2_1, it converges strongly in L^2_1, and its limit is also a solution.*

Proof Let (a, ϕ) be the weak L^2_1 limit, and let $A = A_0 + a$.

Step 1. $F(A_j^\circ)^+ \rightsquigarrow F(A^\circ)^+$ in L^2, while $d^*a = 0$.

Indeed, applying the bounded linear map $d^+\colon L^2_1(T^*X) \to L^2(\Lambda^+)$ to a_j, and using the previous lemma, we see that $d^+a_j \rightsquigarrow d^+a$ in L^2. Likewise $d^*a_j = 0 \rightsquigarrow d^*a$.

Step 2. We can find a subsequence $\gamma'_j = (a'_j, \phi'_j)$ converging strongly in L^2 to a limit $\gamma' = (a', \phi')$.

This follows from compactness of the inclusion $L^2_1 \to L^2$,

Step 3. Refining our subsequence, we may assume that $F((A_0 + a'_j)^\circ)$ has a weak L^2 limit.

Indeed, when we proved finiteness of Spin^c-structures admitting SW solutions, we noted the identity

$$\|F(A^\circ)\|_2^2 = -4\pi^2 c_1(\mathfrak{s})^2)[X] + 2\|F(A^\circ)^+\|_2^2.$$

Now $F((A_0 + a_j)^\circ)^+$ is bounded in L^2, so $F((A_0 + a_j)^\circ)$ is again bounded in L^2.

Step 4. We can further assume that $(\phi'_j\phi'^*_j)_0$ and $\nabla_{A'_j}\phi'_j$ have weak L^2 limits.

Indeed, both sequences are bounded in L^2. (In the latter case, this can be checked using the Lichnérowicz formula.)

Step 5. $(\phi'_j\phi'^*_j)_0 \rightsquigarrow (\phi\phi^*)_0$ and $\nabla_{A'_j}\phi'_j \rightsquigarrow \nabla_A\phi$ in L^2.

Indeed, $\nabla_{A_j}\phi_j = \nabla_{A_0}\phi_j + \rho(a_j)\phi_j$, and $\nabla_{A_0}\phi_j$ converges weakly to $\nabla_{A_0}\phi$ (since ϕ_j converges weakly to ϕ in L^2_1). Thus the weak limit of $\rho(a_j)\phi_j$ exists. But $a_j \to a$ and $\phi_j \to \phi$ strongly in L^2, so $\rho(a_j)\phi_j \to \rho(a)\phi$ strongly in L^1, and therefore $\rho(a_j)\phi_j$ converges to $\rho(a)\phi$ weakly in L^2. This shows that $\nabla_{A_j}\phi_j$ converges to $\nabla_A\phi$ weakly in L^2, as claimed. Similarly, the strong L^2 convergence $\phi_j \to \phi$ implies strong L^1 convergence $(\phi_j\phi_j^*)_0 \to (\phi\phi^*)_0$, and therefore also weak L^2 convergence.

Step 6. $\mathfrak{F}'_\eta(a,\phi) = 0$.

Follows from steps 1 and 5.

Step 7. $F((A_0 + a'_j)^\circ)^+$, $(\phi'_j\phi'^*_j)_0$ and $\nabla_{A'_j}\phi'_j$ all converge strongly in L^2.

For one can now check that the L^2-norms of the (known) weak limits are the limits of the L^2-norms.

Step 8. (a'_j, ϕ'_j) converges in L^2_1.

Follows from strong L^2 convergence and step 7.

Step 8. (a_j, ϕ_j) converges in L^2_1 to the solution (a, ϕ).

Indeed, it converges weakly to (a,ϕ) and a subsequence converges strongly. □

Assembling the various stages of this argument, we arrive at a proof of the following sharper statement of the compactness theorem:

Theorem 21.15 *Let $\gamma_j = (a_j, \phi_j)$ be a sequence of solutions to $\mathfrak{F}'_\eta = 0$, such that $(a_j)_{harm}$ is bounded in $H^1_{DR}(X)$. Then*

(i) If (a_j, ϕ_j) converges weakly in L^2_1, it converges in C^∞, and its limit is again a solution.

(ii) (a_j, ϕ_j) always has a subsequence which converges weakly in L^2_1.

22 Transversality

22.1 Reducible solutions

22.1.1 Reducible configurations

We consider ‘configurations’ $(A,\phi)\in\mathcal{A}_{cl}(\mathbb{S}^+)\times\Gamma(\mathbb{S}^+)$. Let $\mathcal{C}(\mathfrak{s})_k$ denote the space of such configurations, of Sobolev class L^2_k relative to a smooth reference connection A_0, where $k\geq 4$. It carries an action of $\mathcal{G}_{k-1}$, the L^2_{k-1} gauge group (the action is continuous because of the Sobolev multiplication available for L^2_{k-1}).

We must distinguish now between two types of configurations:

- The locus $\mathcal{C}^{irr}(\mathfrak{s})_k$ of *irreducible* configurations (A,ϕ): those where ϕ is not identically zero. Notice that the gauge group $\mathcal{G}_{k-1}$ acts freely on these (since it acts freely on the spinor component).
- The locus $\mathcal{C}^{red}(\mathfrak{s})_k$ of *reducible* configurations $(A,0)$. The stabilizer of the gauge action is $\mathsf{U}(1)$, the group of constant gauge transformations.

The *based* gauge group (relative to a point $x\in X$),

$$\mathcal{G}^x_{k-1}:=\{u\in\mathcal{G}_{k-1}:u(x)=1\}$$

acts freely on $\mathcal{C}(\mathfrak{s})_k$. (Notice that $L^2_{k-1}\subset C^0$ when $k-1\geq 3$, so $u(x)$ makes sense.)

Dropping the Sobolev subscripts now, observe that $\mathcal{G}=\mathsf{U}(1)\times\mathcal{G}^x$. The circle group $\mathsf{U}(1)$ acts *semi-freely* on $\mathcal{C}(\mathfrak{s})/\mathcal{G}^x$: it acts freely on $\mathcal{C}^{irr}(\mathfrak{s})$ and trivially on $\mathcal{C}^{red}(\mathfrak{s})$.

22.1.2 Reducible solutions and abelian instantons

We consider reducible solutions to the SW equations $\mathfrak{F}'_\eta=0$. These amount to connections $A=A_0+a$ such that $d^*a=0$ and

$$F(A^\circ)^+-2i\eta=0.$$

From our study of the self-duality complex, we note that we can write the self-dual 2-form η as

$$\eta=\eta_{harm}+\eta',$$

where η_{harm} is a self-dual harmonic form, and $\eta'\in\operatorname{im}d^+$. And indeed, our study of abelian instantons (with minor adjustments to account for η) shows:

Proposition 22.1 *There is a reducible solution to $\mathfrak{F}_\eta=0$ if and only if*

$$c_1(\mathfrak{s})+(1/\pi)\eta_{harm}\in\mathcal{H}^-.$$

When non-empty, the space of reducible solutions modulo $\mathcal{G}_x$ is an affine space for $H^1(X;\mathbb{R})/H^1(X;\mathbb{Z})$.

Notice that existence of reducible depends on g only through its conformal structure $[g]$. And we already have a theorem on non-existence of abelian instantons for generic conformal structures. From it (again, making minor adaptations to accommodate η) we deduce:

Theorem 22.2 *Suppose $b^+(X)>0$. Fix $([g],\eta)$. Then, for any $r\geq 2$, $[g]$ can be approximated in C^r by C^r-conformal structures $[g_i]$ such that for no Spin^c-structure $\mathfrak{s}$ does there exist a reducible solution to $\mathfrak{F}'_\eta=0$.*

(With a little extra thought one can take the $[g_i]$ and the approximation to be C^∞, but I will not explain this point.)

Let $\mathcal{V}$ be the space of pairs $([g],\eta)$ where $[g] \in \mathsf{conf}_X$ is a conformal structure and $\eta \in \Omega^+_{[g]}$. Take both of class C^r. Then $\mathcal{V}$ is a family of Banach spaces $\Omega^+_{[g]}$, parametrized by the contractible Banach manifold conf_X.

Define

$$\mathcal{W}(\mathfrak{s}) = \{([g],\eta) \in \mathcal{V} : c_1(\mathfrak{s}) + (1/\pi)\eta_{harm,g} \in \mathcal{H}^-_g \subset H^2_{DR}(X)\}.$$

This is a continuous family of Banach subspaces inside $\mathcal{V}$; the fiber $\mathcal{W}(\mathfrak{s})_{[g]}$ over $[g]$ of the projection $\mathcal{W}(\mathfrak{s}) \to \mathsf{conf}_X$ is an affine-linear subspace of $\Omega^+_{[g]}$ of codimension $b^+(X)$.

In the case $b^+(X) = 1$, the codimension-1 subspace $\mathcal{W}(\mathfrak{s})$ is called the *wall*. Its complement has two path-components, called *chambers*. (These chambers are distinguishable globally, not just fiberwise, since the base conf_X is contractible.) When $b^+(X) > 1$, the complement to $\mathcal{W}$ is path-connected.

Proposition 22.3 *Suppose $b^+(X) > 0$. Fix $([g_0],\eta_0)$ and $([g_1],\eta_1)$, not on $\mathcal{W}(\mathfrak{s})$.*

(i) If $b^+(X) > 1$ then any interpolating path $([g_t],\eta_t)$ can be approximated by one which avoids $\mathcal{W}(\mathfrak{s})$.

(ii) If $b^+(X) = 1$ then any interpolating path $([g_t],\eta_t)$ can be approximated by one transverse to the wall $\mathcal{W}(\mathfrak{s})$.

Again, this is a minor variant on our earlier transversality theorem for abelian instantons (except that it is based on the standard transversality theorem relative to a subspace); the proof essentially carries over.

22.2 Transversality for irreducible solutions

We shall quote without proof the following result from elliptic theory:

Theorem 22.4 (unique continuation) *Let L be a linear elliptic operator over a connected manifold X, and suppose that $Lu = 0$. If $u|_U = 0$ for an open set $U \subset X$ then $u = 0$.*

The case of the classical Laplacian is a familiar instance, and we used another instance to prove transversality for abelian instantons. See Donaldson–Kronheimer's book, for example, for a proof.

To obtain transversality for irreducible solutions to the SW equations, we will allow η to vary. We will be a little more economical, however, in that we will fix η_{harm} and let $\eta' \in \operatorname{im} d^+$ vary.

For this purpose, fix $\omega \in \mathcal{H}^-_g$, and define the parametric Seiberg–Witten map (with Coulomb gauge fixing relative to A_0),

$$\mathfrak{F}^{par}_\omega \colon d^+(L^2_{k+1}(T^*X)) \times L^2_k(iT^*X) \times L^2_k(\mathbb{S}^+) \to L^2_{k-1}(X)_0 \times L^2_{k-1}(i\mathfrak{su}(\mathbb{S}^+)) \times L^2_{k-1}(\mathbb{S}^-)$$

to be

$$\mathfrak{F}^{par}_\omega(\eta', a, \phi) \mapsto \mathfrak{F}'_{\omega+\eta'}(a,\phi).$$

Theorem 22.5 *If $\mathfrak{F}^{par}_\omega(\eta, A, \phi) = 0$, where $\phi \not\equiv 0$, then the derivative $D_{(\eta,A,\phi)}\mathfrak{F}^{par}_\omega$ is surjective.*

Proof The derivative $\mathsf{D} = D_{(\eta,A,\phi)}\mathfrak{F}^{par}_{\omega}$ is given by the following formula:

$$\mathsf{D}\begin{bmatrix}\delta\\ b\\ \chi\end{bmatrix} = \begin{bmatrix} d^*b\\ \rho(d^+b+2i\delta)-(\chi\phi^*+\phi\chi^*)_0\\ D_A^+\chi+\frac{1}{2}\rho(b)(\phi)\end{bmatrix}$$

This is a Fredholm map between Hilbert spaces.

Our claim is that L^2-orthogonal complement to $\operatorname{im}\mathsf{D}$ is zero. It follows that $\operatorname{im}\mathsf{D}$ is dense in L^2; hence dense in L^2_{k-1}; and so—the image being closed in L^2_{k-1}—D is surjective.

Take (f,α,ψ) L^2-orthogonal to $\operatorname{im}\mathsf{D}$. We must prove that it vanishes. We have

$$\mathsf{D}\begin{bmatrix}\delta\\ 0\\ 0\end{bmatrix} = \begin{bmatrix} 0\\ 2i\rho(\delta)\\ 0\end{bmatrix},$$

so α is L^2-orthogonal to $i\rho(\delta)$ for arbitrary $\delta\in\operatorname{im}d^+$. Now, $\rho\colon i\Lambda^+\to i\mathfrak{su}(\mathbb{S}^+)$ is a bundle isomorphism; the finite-codimensional subspace $\operatorname{im}d^+\subset i\Omega^+$ therefore maps to a dense subspace of $L^2_k(i\mathfrak{su}(\mathbb{S}^+))$. Hence $\alpha=0$.

Next observe (by taking $\delta=b=0$) that ψ is L^2-orthogonal to the image of D_A^+, and therefore lies in $\ker D_A^+$. By unique continuation, to show $\psi=0$ it will suffice to prove it on an open set.

We have

$$\mathsf{D}\begin{bmatrix}0\\ b\\ 0\end{bmatrix} = \begin{bmatrix} d^*b\\ \rho(d^+b)\\ \frac{1}{2}\rho(b)(\phi)\end{bmatrix}.$$

Taking $b=d^*\beta$ for a 2-form β, so that $d^*b=0$, we see that ψ is L^2_{k-1}-orthogonal to $\rho(d^*\beta)\phi$. By assumption, $\phi(x)\neq 0$ for some $x\in X$. Fix such an x, and a small ball B centered on x, on which $\phi(x)\neq 0$. On B, we can manufacture a 1-form b so that $\rho(b)\phi=\psi$. Let $\sigma\in C^\infty(X)$ be a cutoff function for B, equal to 1 near x and zero outside B. Then, if ψ is not identically zero on B, we have $\langle\rho(\sigma\cdot b)\phi,\psi\rangle_{L^2}>0$. We can approximate $\sigma\cdot b$ by co-exact form $d^*\beta$, so that $\langle\rho(d^*\beta)\phi,\psi\rangle_{L^2}>0$. But this contradicts our assumption that ψ is L^2 orthogonal to the image of D. Thus $\psi\equiv 0$ on B.

Now,

$$\mathsf{D}\begin{bmatrix}0\\ 0\\ \chi\end{bmatrix} = \begin{bmatrix} 0\\ (\phi\chi^*+\chi^*\phi)_0\\ D_A^+\chi\end{bmatrix},$$

and so $\langle D_A^+\chi,\psi\rangle_{L^2}=0$ for all χ. Hence $\langle\chi,D_A^-\psi\rangle_{L^2}=0$.

we can manufacture β_1 and β_2 such that $(\rho(d^*\beta_1)\phi,\rho(d^*\beta_2)\phi$ form an orthonormal basis for $\mathbb{S}^+_x$, and so span $\mathbb{S}^+$ on a neighborhood U_x of x. Thus $\psi=0$ on U_x, and hence on X.

Finally, we have $\int_X fd^*a\,\mathrm{vol}=0$ for all a, so $df=0$ and therefore f (having mean zero) must vanish too. □

23 Transversality, continued

23.1 Generic transversality for irreducibles

23.1.1 Previously

We want to show that, for a fixed self-dual harmonic form ω, and for generic perturbations $\eta' = d^+\beta$, the irreducible part of the SW moduli space with 2-form $\omega + \eta'$ is cut out transversely. That is, we want to show that, for generic η', for any pair $(A = A_0 + a, \phi)$, with $\phi \neq 0$, such that

$$\mathfrak{F}'_{\omega+\eta'}(a,\phi) = 0,$$

the linearized operator

$$D_{(a,\phi)}\mathfrak{F}'_{\omega+\eta'}$$

surjects. (Recall that $\mathfrak{F}'_{\omega+\eta'} = 0$ incorporates the Coulomb gauge condition $d^*a = 0$, and we take the codomain of the d^*-term to be mean-zero functions.)

Notice that the equation $\mathfrak{F}'_{\omega+\eta'}(a,\phi) = 0$ could be rewritten as an inhomogeneous equation

$$\mathfrak{F}'_{\omega}(a,\phi) = \begin{bmatrix} 0 \\ \eta' \\ 0 \end{bmatrix},$$

so really we want $(0,\eta',0)$ to be a regular value of $\mathfrak{F}'_\omega$ (on irreducible configurations).

What we have shown so far is that the parametric SW map,

$$\mathfrak{F}^{par}_\omega \colon (\eta', a, \phi) \mapsto \mathfrak{F}'_{\omega+\eta'},$$

restricted to irreducibles, has 0 as a regular value.

23.1.2 The schematic argument

The argument we will use is a rather general one, so we run through it in a model of our situation.

We have a smooth map $F = (f_1, f_2) \colon U \to V \times P$, where U, V and P are Banach spaces. We think of P as a parameter space. For each $p \in P$, we have a space $M_p = F^{-1}(0,p)$, and we want to show that it is cut out transversely (meaning $(0,p)$ is a regular value of p) for generic p. We consider the map $\widehat{F} \colon U \times P \to V \times P$, $\widehat{F}(x,p) = (f_1(x), f_2(x) - p)$, and we suppose that $\widehat{M} := \widehat{F}^{-1}(0,0)$ is cut out transversely; that is, if $f_1(x) = 0$ and $f_2(x) = p$ then $D_{(x,p)}\widehat{F}$ is surjective.

The argument is as follows. There is a smooth projection map $\Pi \colon \widehat{M} \to P$, and $M_p = \Pi^{-1}(p)$.

Lemma 23.1 *If $F(x) = (0,p)$ then $\ker D_xF = \ker D_{x,p}\Pi$ and $\operatorname{coker} D_xF \cong \operatorname{coker} D_{x,p}\Pi$. Hence $(0,p)$ is a regular value for F if and only if p is a regular value for Π.*

Proof The tangent space $T_{x,p}\widehat{M}$ is the space of pairs $(\dot{x},\dot{p})$ with $\dot{x} \in \ker D_xf_1$ and $\dot{p} = D_xf_2(\dot{x})$. And $D\Pi(\dot{x},\dot{p}) = \dot{p}$, so $\ker D_{x,p}\Pi = \ker D_xf$.

The assumption that $\widehat{M}$ is cut out transversely means that the equations $(D_xf_1)(\dot{x}) = v$, $(D_xf_2)(\dot{x}) - \dot{p} = q$ are solvable for $(\dot{x},\dot{p})$, given $(x,p) \in \widehat{M}$ and a tangent vector (v,q). This amounts to the statement that D_xf_1 is surjective.

In general, if $L_1\colon U \to V$ and $L_2\colon U \to P$ are linear maps, with L_1 surjective, the inclusion $P \to V \times P$ induces an isomorphism

$$\frac{P}{L_2(\ker L_1)} \to \operatorname{coker}(L_1 \times L_2).$$

In the case at hand, we see that

$$\operatorname{coker} D\Pi = \frac{P}{Df_2(\ker Df_1)} \cong \operatorname{coker} DF.$$

□

If the Banach spaces are finite-dimensional, the next step is to invoke Sard's theorem:

Theorem 23.2 (Sard, v.1) *The critical values of a C^∞ map $f\colon M \to N$ between finite-dimensional manifolds are of measure zero in N; equivalently, the regular values are of full measure.*

This is the most traditional statement of Sard's theorem. The notion of 'measure zero' makes sense in a finite-dimensional normed space, and is invariant under diffeomorphisms and countable unions, and so makes sense on N. However, it does not make sense in an infinite-dimensional Banach space, since volume is not meaningful. However, there is a variant which is better suited to our purposes, and which Smale adapted to *Fredholm maps* in infinite dimensions.

Definition 23.3 A subspace U of a topological space T is called *Baire* if it is the intersection of a countable collection of open dense subsets.

The Baire category theorem says when the topology of T arises from a complete metric space, Baire subspaces are dense.

Theorem 23.4 (Sard, v.12) *The regular values of a C^∞ map $f\colon M \to N$ between finite-dimensional manifolds form a Baire subspace.*

Banach spaces obviously admit complete metrics, but so too do Fréchet spaces. Thus, if X is a closed manifold and $E \to X$ a vector bundle, the Baire category theorem applies to all the following spaces T:

$$L^p_k(E); \quad C^r(E); \quad C^\infty(E).$$

Definition 23.5 Let Y and Z be C^∞ Banach manifolds. Let $\Phi\colon Y \to Z$ be a smooth map. We call Φ a *Fredholm map* if its derivative $D_x\Phi\colon T_xY \to T_{f(x)}Z$ is a Fredholm map for all x.

Theorem 23.6 (Smale) *The regular values of a Fredholm map Φ form a Baire subset of Z.*

As in the finite-dimensional case, there are sharper versions allowing for finite differentiability. We will give a proof presently.

Returning to our map F, if we suppose that F is a Fredholm map then Π is Fredholm too, by the lemma. Thus, by Smale–Sard, the regular values of Π form a Baire subset. And these regular values are exactly the parameters p for which $F^{-1}(0,p)$ is cut out transversely.

The inverse function theorem *is* available:

Theorem 23.7 *Let $f\colon M \to N$ be a smooth map of Banach spaces, and suppose that $D_x f$ is surjective. Then there is a neighborhood U of x such that $f^{-1}(M) \cap U$ is a smooth submanifold of U. In particular, if y is a regular value of f then $f^{-1}(y)$ is a smooth submanifold of M.*

So when p is regular for Π, $F^{-1}(0,p)$ is a manifold, with tangent spaces $\ker D_x F$ (and $\operatorname{coker} D_x F = 0$). Its dimension is the Fredholm index of $D_x F$.

We can now state a generic transversality theorem:

Theorem 23.8 *Fix a metric g, a g-harmonic form ω, a $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$, and an integer $k \geq 3$. Let P be the space of 2-forms $\eta' = d^+\beta \in \operatorname{im} d^+$ where β is of class L^2_{k+1}. The subset $P_{reg} \subset P$ of forms η' for which all irreducible solutions to $\mathfrak{F}'_{\omega+\eta'} = 0$ are cut out transversely is a Baire subspace. When $\eta' \in P_{reg}$, the space of irreducible solutions*

$$M_{\omega+\eta'} = (\mathfrak{F}'_{\eta})^{-1}(0)$$

is a smooth manifold of finite dimension $d(\mathfrak{s}) + 1$, with a free $\mathsf{U}(1)$-action.

The proof of the theorem exactly follows the template we have discussed, with $F = \mathfrak{F}^{par}_{\omega}$ and $P = \operatorname{im} d^+$. We use the parametric Seiberg–Witten map to define a parametric solution space M^{par}_{ω}, with a Fredholm projection $\Pi\colon M^{par}_{\omega} \to \operatorname{im} d^+$. Regular values of Π form a Baire subspace of $\operatorname{im} d^+$, by Sard–Smale, and η' is regular for Π just when the solution space $\mathfrak{F}'_{\omega+\eta'})^{-1}(0)$ is cut out transversely.

The dimension of the smooth manifolds of solutions is given by the Fredholm index of the linearized operator. Notice that for present purposes, the codomain of the d^* component of the SW map is $L^2_{k-1}(X)_0$, the functions of mean zero. Earlier, we computed the index to be $d(\mathfrak{s})$, but in that situation we did not impose the mean-zero condition. This discrepancy explains why the relevant index here is $d(\mathfrak{s}) + 1$.

The action of $\mathsf{U}(1)$ is that by constant gauge transformations; of course, the orbit-space has dimension $d(\mathfrak{s})$.

Corollary 23.9 *When $b^+ > 0$, for generic metrics g and generic g-self-dual perturbations η with fixed g-harmonic class $\eta_{harm} = w \in H^2(X)$, the SW moduli spaces $\mathfrak{F}'_{\omega_g+\eta'} = 0$ are cut out transversely as manifolds of dimension $d(\mathfrak{s}) + 1$, carrying a free action of S^1.*

23.2 Fredholm maps and the Sard–Smale theorem

Definition 23.10 Let Y and Z be smooth Banach manifolds (with second countable topology). A *Fredholm map* from Y to Z is a smooth map $\Phi\colon Y \to Z$ whose derivative $D_y\Phi\colon T_yY \to T_zZ$ is Fredholm for every $y \in Y$.

In local charts, Φ amounts to a map $B_1 \to B_2$ between Banach spaces; the derivatives $D_y\Phi\colon B_1 \to B_2$ vary continuously in the operator-norm topology, and so their indices $\operatorname{ind} D_y\Phi = \dim\ker D_y\Phi - \dim\operatorname{coker} D_y\Phi$ are locally constant in y. Thus we have a locally constant function $\operatorname{ind}\colon Y \to \mathbb{Z}$.

Now let Y and Z be Banach manifolds, and $\Phi\colon Y \to Z$ be a Fredholm map. Fix $y \in Y$. Consider charts near y given by embeddings $i\colon (T,0) \to Y$, $i(0) = y$. Here $T \subset T_yY$ is a neighborhood of 0, and we require $D_0 i = \mathrm{id}_{T_yY}$. Similarly consider charts near $\Phi(y)$ given by embeddings $i'\colon (T',0) \to Z$,

$i(0) = \Phi(y)$, where $T' \subset T_{\Phi(y)}Z$ is a neighborhood of 0, and we require $D_0 i' = \mathrm{id}_{T_{\Phi(y)}Z}$. Let us call such embeddings 'tangential charts'.

In tangential charts, Φ becomes a map $\tilde{\Phi} = i' \circ \Phi \circ i^{-1} \colon T \to T'$ with $D_0\Psi = D_0\Phi$.

Let $K = \ker D_y\Phi \subset T_yY$ (finite-dimensional), and let U be a complement to K.

Let $V = \operatorname{im} D_y\Phi \subset T_{\Phi(y)}Z$ and let C be a (finite-dimensional) complement to V.

Then, in tangential charts, $\tilde{\Phi}$ takes the form of a map

$$\widehat{\Phi} \colon (U \oplus K, 0) \to (V \oplus C, 0).$$

More precisely, the germ of $\tilde{\Phi}$ near 0 is the germ of $\widehat{\Phi}$ near 0; henceforth we conflate maps with their germs near 0.

Lemma 23.11 *Once U and C and a tangential chart near $\Phi(y)$ are chosen, there exists a tangential chart near y in which $\widehat{\Phi}$ takes the form*

$$\widehat{\Phi} \colon U \oplus K \to V \oplus C, \quad \widehat{\Phi}(x,k) = (Lx, \phi(x,c)),$$

where $L = D_y\Phi(x) \colon U \to V$ (a linear isomorphism).

Proof We have

$$D_y\Phi = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix}.$$

Work with the chosen tangential chart near $\Phi(y)$ and an arbitrary initial tangential chart near y. Thus we view Φ as a map $U \oplus K \to V \oplus C$. Define

$$\Psi \colon U \oplus K \to V \oplus K, \quad \Psi(x,k) = (\mathrm{pr}_V \circ \Phi(x,k), k).$$

Then

$$D_0\Psi = \begin{bmatrix} L & 0 \\ 0 & I_K \end{bmatrix},$$

so by the inverse function theorem, Ψ has smooth inverse Ψ^{-1} near 0. The composite

$$\Phi \circ \Psi^{-1} \circ D_0\Psi \colon U \oplus K \to V \oplus C$$

takes the form $(x,k) \mapsto (Lx, \phi(x,k))$; so, adjusting the chosen tangential chart near y via $\Psi^{-1} \circ D_0\Psi$, one obtains the claimed result. $\square$

Lemma 23.12 *A Fredholm map Φ is locally closed: every point $y \in Y$ has a neighborhood N such that $\Phi|_N$ maps closed sets in N to closed sets in Z.*

Proof We may work in the charts provided by the last lemma, so that Φ takes the form $U \oplus K \to V \oplus C$, $\widehat{\Phi}(x,k) = (Lx, \phi(x,c))$ with L a linear isomorphism. Let N be of the form $U \times B_K$, where B_K is a neighborhood of 0 in K. We claim that $\Phi|_N$ is a closed mapping. Let $A \subset N$ be a closed set, and take a sequence in $\Phi(A)$, say $(y_i, c_i) = (Lx_i, \phi(x_i, k_i))$, converging to a limit (y, c). We can write $y = Lx$, so $L(x - x_i) \to 0$, so $x_i \to x$. Since B_K is bounded, we can pass to a subsequence for which $k_i \to k \in K$; since A is closed, $(x,k) \in A$. And $c = \phi(x,k)$, so $\Phi(A)$ is closed too. $\square$

We can now prove the Sard–Smale theorem. Let $\Phi\colon Y \to Z$ be a Fredholm map. We assume Y second countable, so there is a countable cover of Y by open sets U_α on which $\Phi|_{U_\alpha}\colon U_\alpha \to Z$ is closed, and on which Φ takes the special form $\Phi(x,k) = (Lx, \phi(x,c))$, as above, in suitable charts. It will suffice to show that the regular values of $\Phi|_{U_\alpha}$ are open and dense. The critical points of $\Phi|_{U_\alpha}$ are closed in U_α, so the critical values are closed in N, and the regular values open. The derivative of Φ, given in this local form, is

$$D_{(x,k)}\Phi = \begin{bmatrix} L & 0 \\ D_{(x,k)}\phi|_U & D_{(x,k)}\phi|_K \end{bmatrix},$$

and this is surjective if and only if $D_{(x,k)}\phi|_K$ is surjective. Thus the regular values of $\Phi|_{U_\alpha}$ are precisely those of $k \mapsto \phi(0,k)$, and these are dense by Sard.

24 The diagonalization theorem

24.1 Statement and a preliminary reduction

Theorem 24.1 *Suppose that X is a closed, oriented 4-manifold with positive-definite intersection form Q_X. Then $Q_X \cong \langle 1\rangle^{b_2(X)}$, i.e., Q_X is diagonalizable over $\mathbb{Z}$.*

In the case of fundamental groups without non-trivial representations into SU(2), the theorem was proved by Donaldson (*An application of Yang–Mills theory to 4-dimensional topology*, 1982). His proof, using instanton moduli spaces, remains one of the most beautiful things in modern geometry. Variants of the argument allowed somewhat more general fundamental groups, but still somewhat restricted. The proof presented here, using Seiberg–Witten moduli spaces, is also elegant, and requires no assumption on π_1. It is due to Kronheimer–Mrowka (unpublished), but expositions appear in the Seiberg–Witten theory texts of L. Nicolaescu and D. Salamon, for instance.

Preliminaries:

(1) It is equivalent to assert that negative-definite intersection forms are diagonalizable over $\mathbb{Z}$.

(2) *Claim:* Given a closed, oriented 4-manifold X with $b_1(X) > 0$, there is a 4-manifold Y with $Q_Y \cong Q_X$ and $b_1(Y) < b_1(X)$. Hence it suffices to prove the theorem when $b_1(X) = 0$.

We prove the claim using surgery. Suppose $b_1(X) > 0$, and consider a class $h \in H_1(X)$ which is non-torsion and primitive. By transversality theory, we may represent h by a smoothly embedded loop γ. Surgery along γ excises a narrow tubular neighborhood $N_\gamma = S^1 \times D^3$ to form a manifold X° with boundary $S^1 \times S^2$; and then glues in $D^2 \times S^2$ along its boundary $S^1 \times S^2$ to obtain a new manifold Y in which γ bounds a disc.

Lemma 24.2 *One has $H_1(Y) = H_1(X)/\langle h\rangle$.*

Proof The inclusion-induced maps $H_1(X^\circ) \to H_1(X)$ and $H_1(X^\circ) \to H_1(Y)$ are surjective, by transversality theory or the Mayer–Vietoris sequence. The latter sequence shows that their respective kernels are the images in $H_1(X^\circ)$ of the classes in $H_1(S^1 \times S^2)$ that become trivial in $H_1(S^1 \times D^3)$ or in $H_1(D^2 \times S^2)$. Thus $H_1(X^\circ) = H_1(X)$ while $H_1(X^\circ)/[\gamma] \cong H_1(Y)$. □

Lemma 24.3 *$Q_Y \cong Q_X$.*

Proof Take a basis for $H_2(X)'$, and represents it by a collection of oriented surfaces $(\Sigma_1, \ldots, \Sigma_{b_2})$ in X. We can arrange, by making small perturbations, that all Σ_i are disjoint from γ, and we can then choose N_γ disjoint from the Σ_i too. The intersection numbers $\Sigma_i \cdot \Sigma_j$ can be computed in X°, so we have an isometric copy of $H_2(X)'$ inside $H_2(Y)'$. We have $\chi(Y) = \chi(X^\circ) + 2 = \chi(X) + 2$, so $b_2(Y) = b_2(X)$, and therefore $H_2(X)'$ has finite index in $H_2(Y)'$. But both are unimodular, so the index is 1. □

The claim follows from the two lemmas.

24.2 Seiberg–Witten moduli spaces

Assume now that $b_1 = 0$. We take Q_X is *negative*-definite; then the operator $d^+ \oplus d^* \colon \Omega^1 \to \Omega^+ \oplus \Omega^0$ has cokernel $H^0(X) = \mathbb{R}$ (trivial cokernel if we take mean-zero functions in the codomain); and also trivial kernel. This explains the preference for negative- over positive-definite lattices.

Fix a Spin^c-structure $\mathfrak{s}$ with Chern class $c = c_1(\mathfrak{s})$. Its Dirac operators $D_A^+ \colon \Gamma(\mathbb{S}^+) \to \Gamma(\mathbb{S}^+)$ have index

$$\operatorname{ind}_{\mathbb{R}} D_A^+ = \frac{1}{4}(c^2[X] - \tau_X) = \frac{1}{4}(c^2[X] + b_2).$$

Thus

$$d(\mathfrak{s}) = \frac{1}{4}(c_1(\mathfrak{s})^2[X] + b_2) - 1.$$

Note that $d(\mathfrak{s})$ is odd, because c is a characteristic vector, hence $c^2 \equiv \tau$ modulo 8.

Consider the space of irreducible SW solutions

$$M_\eta^{irr}(\mathfrak{s}) = \tilde{M}_\eta^{irr}(\mathfrak{s})/\mathsf{U}(1).$$

Here $\eta \in \operatorname{im} d^+$. We take the harmonic 2-form ω to be 0.

For generic η, $\tilde{M}_\eta^{irr}(\mathfrak{s})$ is cut out transversely as a manifold of dimension $d(\mathfrak{s}) + 1$. Pick such an η.

If $d(\mathfrak{s}) < 0$, there are then no irreducible solutions to the SW equations. So it is natural to suppose $d(\mathfrak{s}) \geq 0$. That is, we make the

Hypothesis 24.4 *The* Spin^c*-structure* $\mathfrak{s}$ *has* $c_1(\mathfrak{s})^2[X] + b_2(X) > 0$.

Write $d(\mathfrak{s}) = 2k - 1$. Since $b^+ = 0$ and $b_1 = 0$, there is a unique gauge-orbit of reducible solutions to the η-SW equations, say $[A_0, 0]$. (Thus $F(A_0^\circ)^+ = 2i\eta$.) Thus the full moduli space $\tilde{M}_\eta(\mathfrak{s})$ is a compact space containing a distinguished point R (the reducible solution) whose complement is a manifold of dimension $2k$. It comes with $\mathsf{U}(1)$ action fixing R and free in its complement.

The important point is to understand the structure of $\tilde{M}_\eta(\mathfrak{s})$ near R.

Let $\mathsf{D} = D_{(A_0,0)}\mathfrak{F}'_\eta$. Then

$$\mathsf{D}\begin{bmatrix} b \\ \chi \end{bmatrix} = \begin{bmatrix} d^*b \\ \rho(d^+b) \\ D_{A_0}^+\chi \end{bmatrix}.$$

Since we are dealing with $\tilde{M}_\eta$, we take the first factor in the codomain to be $L^2_{k-1}(X)_0$, the mean-zero functions.

Thus $\ker \mathsf{D} = \mathcal{H}^1 \times \ker D_{A_0}^+ = \ker D_{A_0}^+$, and $\operatorname{coker} \mathsf{D} = \mathcal{H}^+ \times \operatorname{coker} D_{A_0}^+ = \ker D_{A_0}^-$.

Hypothesis 24.5 *The reducible solution* $R = (A_0, 0)$ *is* regular, *meaning* $\operatorname{coker} \mathsf{D} = 0$ *(i.e.,* $\operatorname{coker} D_{A_0}^+ = 0$*).*

Things will be straightforward when R is regular. Later we will explain how to arrange regularity.

While this hypothesis may seem a bland extension of the generic regularity of irreducibles that we have already established, it actually has a remarkable consequence. For, in the regular case, $\tilde{M}_\eta$ is cut out transversely near R, as an odd-dimensional manifold. That means that irreducible solutions *exist* (recall that we have no yet constructed any such solutions).

The circle acts on the tangent space at the fixed point, $T_R\tilde{M}_\eta = \ker D_{A_0}^+$, by scalar multiplication.

Lemma 24.6 *Suppose that Q is a $2k$-manifold with an action of $\mathsf{U}(1)$, and that q is a fixed point of the action. Assume that the resulting action on T_qQ has a single weight N, i.e. $T_qQ \cong \mathbb{C}^{\otimes N} \otimes \mathbb{R}^k$ as a representation. Then there is a neighborhood of q and an equivariant chart $(\mathbb{C}^{\otimes N} \otimes \mathbb{R}^k, 0) \to (Q, q)$.*

Proof Locally near q, choose an invariant Riemannian metric (possible by averaging), and use its exponential map to define the chart. □

Thus R has a $\mathsf{U}(1)$-equivariant neighborhood modeled on a neighborhood of 0 in $\mathbb{C}^k$, with the action of $\mathsf{U}(1)$ by scalar multiplication.

Now remove a small equivariant ball around R, so as to obtain a compact $2k$-manifold $\tilde{N}$ bounding a sphere S^{2k-1}, with a free action of $\mathsf{U}(1)$ given by scalar multiplication on the boundary. The quotient N has $\partial N = \mathbb{C}P^{k-1}$.

When $k-1$ is even, this immediately generates a contradiction: $\mathbb{C}P^{k-1}$ is not a boundary. (An obstruction to bounding arises from the mod 2 Euler characteristic, which is also the top Stiefel–Whitney class evaluated on the fundamental class.)

Regardless of the parity of k, we obtain a slightly less obvious contradiction. The map $\tilde{N} \to N$ is a principal $\mathsf{U}(1)$-bundle, so its associated complex line bundle E has a Stiefel–Whitney class $w_2(E) \in H^2(N; \mathbb{Z}/2)$ (which is the mod 2 reduction of $c_1(E)$). Now $E|_{\partial N}$ is the tautological bundle over $\mathbb{C}P^{k-1}$, so $w_2(E)$ is the generator of $H^2(\mathbb{C}P^{k-1}; \mathbb{Z})$. Thus $w_2(E)^{k-1}[\partial N] = 1$. But this contradicts the triviality of $[\partial N]$ is trivial in $H_{2k-1}(N; \mathbb{Z}/2)$.

So—modulo the issue of regularity of R—we see that there can be no characteristic vector c for Q_X such that $-c^2 < b_2(X)$. Equivalently, passing to the positive-definite lattice $\Lambda = -Q_X$, we see that Λ admits no 'short' characteristic vectors: every characteristic vector c satisfies $|c|^2 \geq b_2 = \mathsf{rank}\,\Lambda$.

Which lattices Λ does this argument preclude?

Theorem 24.7 (N. Elkies) *Suppose that Λ is a positive-definite unimodular lattice, of rank N, whose shortest characteristic vectors c have $c \cdot c \geq N$. Then Λ is isomorphic to the standard $\mathbb{Z}^N$ lattice $N\langle 1 \rangle$.*

Note that the shortest characteristic vectors for $\mathbb{Z}^N$ are $(\pm 1, \dots, \pm 1)$, of length-squared N. The proof of Elkies' theorem falls outside the main topic of this course, but for those who are curious I have written an account of it below.

24.3 Generic regularity of the reducible solution

Lemma 24.8 *Fix $k \geq 3$. Introduce the parametric Dirac map*

$$\mathcal{D}^{par} \colon L^2_k(iT^*X) \times L^2_k(\mathbb{S}^+) \to L^2_{k-1}(\mathbb{S}^-), \quad (a, \phi) \mapsto D^+_{A_0+a}\phi = D^+_{A_0}\phi + \rho(a)\phi.$$

Restrict the domain to

$$L^2_k(iT^*X) \times (L^2_k(\mathbb{S}^+ \setminus \{0\}).$$

Then $\mathcal{D}^{par}$ has 0 as a regular value.

Proof One has

$$(D_{a,\phi}\mathcal{D}^{par})(b,\chi) = D^{+}_{A_0}\chi + \rho(a)\chi + \rho(b)\phi.$$

We need to show $D_{a,\phi}\mathcal{D}^{par}$ is surjective. Take ψ to be L^2-orthogonal to its image; it will suffice to prove $\psi \equiv 0$. Taking $b = 0$, we see that ψ is orthogonal to $D^{+}_{A_0+a}\chi$ for all χ; hence $D^{-}_{A_0+a}\psi = 0$. By unique continuation, therefore, it will suffice to prove that ψ vanishes on an open set. By assumption, $\phi(x) \neq 0$ for some x. Now take $\chi = 0$ to see that ψ is orthogonal to $\rho(b)\phi$ for all b. This forces $\chi = 0$ on a neighborhood of x where ϕ is nowhere-vanishing; and we are done. □

We now follow the same model argument as we used to establish generic regularity for irreducible solutions to the SW equations. Let H^{par} be the zero-set of $\mathcal{D}^{par}$ (where we assume $\phi \not\equiv 0$). Thus we have a projection map $\Pi\colon H^{par} \to L^2_k(iT^*X)$, with fibers $\Pi^{-1}(a) = (\ker D^{+}_{A_0+a}) \setminus \{0\}$. The map $\mathcal{D}^{par}$ is Fredholm, and therefore so is Π. The index of Π is $\operatorname{ind} D^{+}_{A}$. By the Sard–Smale theorem, the regular values of Π are a Baire subset of the 1-forms. These regular values are precisely the 1-forms for which $\operatorname{coker} D^{+}_{A_0+a} = 0$. Thus we conclude:

Proposition 24.9 *For generic a,* $\operatorname{coker} D^{+}_{A_0+a} = 0$.

We now proceed as follows. Writing $A = A_0 + a$, define η by

$$\eta(a) = \frac{1}{2i}F(A^{\circ})^{+}.$$

Since $\mathcal{H}^{+} = 0$, $\Omega^{+} = \operatorname{im} d^{+}$, and so $\eta \in \operatorname{im} d^{+}$. We impose two simultaneous conditions on a: (i) $\operatorname{coker} D^{+}_{A} = 0$; and (ii) η is regular for irreducible SW solutions modulo $\mathcal{G}_x$. Condition (i) is defined by a Baire subspace of the 1-forms; condition (ii) by a Baire subset of $\operatorname{im} d^{+}$; both can be achieved simultaneously.

Note that we did not impose $d^*a = 0$, so the most convenient gauge choice for solutions $(A_0 + a', \phi)$ to the SW equations is to require $d^*a' = d^*a$. This makes no significant changes to the analysis.

Remark. In this proof we have seen something remarkable: starting from a reducible solution—a solutions to a linear PDE—we have obtained a compact manifold's worth of solutions to a non-linear PDE. This then leads to a contradiction, and all our hard-won irreducible solutions disappear in a puff of logic. Nonetheless, there is a closely-related situation in which the same mechanism serves to provide *actual* solutions to the irreducible SW equations, and that is *wall-crossing formula*, valid when $b^{+} = 1$ (it takes a simple form when also $b_1 = 0$).

24.4 Proof of Elkies' theorem

The proof uses θ-functions, their modularity properties under a certain subgroup of $PSL_2(\mathbb{Z})$, and their failure to be modular under the whole of $PSL_2(\mathbb{Z})$. J.-P. Serre's *A course in arithmetic* covers the relevant background. With the machinery of θ-functions in place, the argument is very simple.

24.4.1 θ-functions of lattices

View Λ as an integral lattice in $\mathbb{R}^N$, and write $u \cdot v$ for the dot product of elements in Λ. The θ-series for Λ is the q-series

$$\theta_\Lambda = \sum_{v\in\Lambda} q^{|v|^2} = \sum_{n\geq} r_\Lambda(n) q^n.$$

Here $r_\Lambda(n)$ is the number of vectors v with $v \cdot v = n$. One has

$$\theta_{\Lambda_1\oplus\Lambda_2} = \theta_{\Lambda_1} \cdot \theta_{\Lambda_2}.$$

24.4.2 Modularity

Writing $q = e^{\pi i\tau}$, the series defining θ_Λ converges uniformly when τ varies in a compact subset of the upper half-plane $\mathbb{H}$, and so defines a holomorphic function $\theta_\Lambda(\tau)$ on $\mathbb{H}$ (i.e. for $0 < |q| < 1$) which is holomorphic also at $\tau = i\infty$ (i.e. at $q = 0$).

By construction, $\theta_\Lambda(\tau + 2) = \theta_\Lambda(\tau)$.

Lemma 24.10 *When Λ is unimodular,*

$$\theta_\Lambda(\tau) = (\tau/i)^{N/2}\theta_\Lambda(-1/\tau).$$

Here $(\tau/i)^{1/2}$ is defined to be positive for $\tau \in i\mathbb{R}_+$.

Proof The Poisson summation formula implies [cf. Serre, *op. cit.*] that if $\Lambda \subset \mathbb{R}^N$ is an integral lattice, and $\Lambda^\vee \subset \mathbb{R}^N$ its dual—comprising the vectors $u \in \mathbb{R}^N$ whose inner products $u\cdot v$ with elements $v \in \Lambda$ are integers—and if $f \in C^\infty(\mathbb{R}^N)$ is a rapidly decreasing function with Fourier transform $\hat{f} \in C^\infty(\mathbb{R}^N)$, then

$$\mathsf{det}\,\Lambda \cdot \sum_{v\in\Lambda} f(v) = \sum_{u\in\Lambda^\vee} \hat{f}(u).$$

Here $\mathsf{det}\,\Lambda$, the determinant of the symmetric matrix representing Λ, is also the volume of the quotient $\mathbb{R}^N/\Lambda$. If one takes $f(x) = e^{-\pi|x|^2}$ then $\hat{f} = f$. Take $t > 0$, and apply the formula to the rescaled lattice $t^{1/2}\Lambda$, whose dual is $t^{-1/2}\Lambda$, to obtain

$$\mathsf{vol}(\mathbb{R}^N/\Lambda) \cdot \theta_\Lambda(it) = t^{N/2}\theta_{\Lambda^\vee}(1/(it)).$$

In the case of a unimodular lattice, $\mathsf{vol}(\mathbb{R}^N/\Lambda) = 1$ and $\Lambda^\vee = \Lambda$. In this case

$$\theta_\Lambda(it) = t^{N/2}\theta_\Lambda(1/(it)).$$

Putting $\tau = it$, we get the stated identity—valid on $\mathbb{H}$ and not just on $i\mathbb{R}_+$ since the two sides are holomorphic. $\square$

When Λ is *even* unimodular, its θ-series has only even powers of q, and so $\theta_\Lambda(\tau + 1) = \theta_\Lambda(\tau)$. Being in addition holomorphic at $i\infty$, θ_Λ is a modular form for the full modular group $PSL_2(\mathbb{Z})$.

In general, the lemma implies that θ_Λ transforms as a modular form of half-integer weight $N/2$ for the group $\Gamma^+ \subset PSL_2(\mathbb{Z})$ generated by S and T^2 : here $T(z) = z + 1$, and $S(z) = -1/z$. The modular curve $\mathbb{H}/\Gamma$ is a sphere with two punctures. Indeed, Γ^+ has a fundamental domain whose closure D^+ in $\mathbb{H}$ is the ideal hyperbolic triangle with vertices 1, -1 and $i\infty$:

$$D^+ = \{z = x + iy \in \mathbb{H} : |x| \leq 1 \leq |z|^2\}.$$

Translation T^2 identifies the two vertical edges, while S maps the arc from 1 to i to the arc from -1 to i. Thus $\mathbb{H}/\Gamma$ has punctures ('cusps') at $i\infty$ and 1. We can compactify $\mathbb{H}/\Gamma$ canonically, to a genus 0 Riemann surface X_{Γ^+}, adding in the two cusps. Since it is defined by non-negative powers of q, θ_Λ is holomoprhic at $i\infty$. We do not assert that θ_Λ a modular form for Γ^+, since the definition includes holomorphy at the second cusp, which we have not checked.

What will be crucial for us is that $\theta_\Lambda(\tau+1)$ is related in an interesting way to the characteristic vectors. Let c_0 be a characteristic vector for Λ (i.e., $c\cdot v \equiv v\cdot v \mod 2$). Define a generating function for characteristic vectors,

$$\chi_\Lambda(\tau) = \sum_{c\in c_0+2\Lambda} q^{|c|^2/4} = \sum_c e^{\pi i \tau (c\cdot c)/4}.$$

Again, $\chi_{\Lambda_1\oplus\Lambda_2} = \chi_{\Lambda_1}\cdot\chi_{\Lambda_2}$

Lemma 24.11

$$(\tau/i)^{N/2}\theta_\Lambda(\tau+1) = \chi_\Lambda(-1/\tau).$$

Proof Observe that

$$\theta_\Lambda(\tau+1) = \sum_v (-q)^{v\cdot v} = \sum_v (-1)^{c_0\cdot v} q^{v\cdot v} = \sum_{v\in\Lambda} e^{\pi i(|v|^2\tau + c\cdot v)}.,$$

where c_0 is a characteristic vector. With this in mind, apply Poisson summation to the rescaled lattice $t^{1/2}\Lambda$, and to the function $f(x) = e^{-\pi|x+t^{-1/2}\frac{c}{2}|^2}$. Again, $\hat{f}=f$, because of the translation-invariance of the Fourier transform. And again, use the rescaled lattice $t^{1/2}\Lambda$. We get

$$t^{N/2}\sum_{v\in\Lambda} e^{-\pi(t|v|^2 + c\cdot v + t^{-1}|c|^2/4)} = \sum_{u\in\Lambda} e^{-\pi t^{-1}|u+\frac{c}{2}|^2},$$

whence the result. □

24.4.3 The lattice $\mathbb{Z}^N$

For the standard lattice $\langle 1\rangle$ (i.e. $\mathbb{Z}\subset\mathbb{R}$) one has

$$\theta_{\mathbb{Z}} = 1 + 2\sum_{n>0} q^{n^2}.$$

Jacobi's triple product identity says that

$$\theta_{\mathbb{Z}} = \prod_{n\geq 1}(1-q)^{2n}(1+q^{2m-1})(1+q^{2m-1}),$$

It implies that $\theta_{\mathbb{Z}}$ is non-zero for $|q|<1$. Thus $\theta_{\mathbb{Z}}(\tau)$ is non-vanishing on $\mathbb{H}$. Thus $1/\theta_{\mathbb{Z}}$ is holomorphic on $\mathbb{H}$, as is $1/\theta_{\mathbb{Z}^N} = 1/\theta_{\mathbb{Z}}^N$.

24.4.4 The ratio of θ-functions

Consider the ratio

$$R(\tau) = \theta_\Lambda/\theta_{\mathbb{Z}^N} = \theta_\Lambda/\theta_{\mathbb{Z}}^N.$$

It is a holomorphic function on $\mathbb{H}$, invariant under Γ^+, and so defines a holomorphic function on $\mathbb{H}/\Gamma^+$. Since $\theta_{\mathbb{Z}^N} = 1+O(q)$ as $q\to 0$, R is holomorphic at the cusp $i\infty$. We examine its behavior at the other cusp 1.

The formula for $\theta_\Lambda(\tau+1)$ implies that

$$R(\tau) = \frac{\chi_\Lambda}{\chi_{\mathbb{Z}}^N}\left(\frac{1}{1-\tau}\right).$$

Thus the behavior of R at $\tau = 1$ is governed by the asymptotics of $\chi_\Lambda/\chi_{\mathbb{Z}}^N$ at $\tau = i\infty$.

The leading term in χ_Λ (viewed as a q-series) is $r_\Lambda(\ell)q^{\ell/4}$, where ℓ is the length-squared of the *shortest* characteristic vectors. The characteristic vectors of $\mathbb{Z}$ are the odd integers, so $\chi_{\mathbb{Z}} = 2q^{1/4}(1+q^2+q^6+....)$, and $\chi_{\mathbb{Z}}^N = 2^N q^{N/4}(1+Nq^2+\dots)$. Thus

$$\frac{\chi_\Lambda}{\chi_{\mathbb{Z}}^N} = \frac{r_\Lambda(\ell)}{2^N} q^{(\ell-N)/4}(1+O(q))$$

as $q \to 0$.

We can now prove Elkies' theorem. The assumption is that $\ell \geq N$. We see that $\chi_\Lambda/\chi_{\mathbb{Z}}^N$ is holomorphic at $q = 0$, so R is holomorphic at the cusp 1. But then R is holomorphic on the compact Riemann surface X_{Γ^+}, and hence constant. Since $R(i\infty) = 1$, we have $R \equiv 1$. Thus $\theta_\Lambda = \theta_{\mathbb{Z}^N}$.

So Λ has $2N$ vectors of length 1, i.e. N pairs $(\pm e_1, \dots, \pm e_N)$ with $|e_j| = 1$. If $i \neq j$, e_i is linearly independent from e_j, and so $|e_i \cdot e_j| < 1$. But $e_i \cdot e_j$ is an integer, so $e_i \cdot e_j = 0$. Thus the e_j span a lattice $\mathbb{Z}^N \subset \Lambda \subset \mathbb{R}^N$. Any $v \in \Lambda$ can be written as $v = \sum (v \cdot e_j)e_j$, and since $v \cdot e_j \in \mathbb{Z}$, we have $v \in \Lambda$.

25 Seiberg–Witten invariants

25.1 Preliminaries

25.1.1 Homology orientations

A *homology orientation* for the closed, oriented 4-manifold X is the equivalence class σ of a triple (H^+, H^-, o) consisting of a splitting $H^2_{DR}(X) = H^+ \oplus H^-$ into positive- and negative-definite subspaces, and an orientation of $H^1_{DR}(X) \oplus H^0_{DR}(X)^* \oplus (H^+)^*$, ie., the ray $\mathbb{R}_+ o$ spanned by an isomorphism

$$o\colon \mathbb{R} \to \det(H^1_{DR}(X) \oplus H^0_{DR}(X)^* \oplus (H^+)^*)$$

Given another triple (K^+, K^-, o'), projection defines an isomorphism $H^+ \to K^+$, given by projection, and therefore an isomorphism $\det(H^1_{DR}(X) \oplus H^0_{DR}(X)^* \oplus (H^+)^*) \to \det(H^1_{DR}(X) \oplus H^0_{DR}(X)^* \oplus (K^+)^*)$; thus it makes sense to compare $\mathbb{R}_+ o$ to $\mathbb{R}_+ o'$. If they are equal, we deem (H^+, H^-, o) and (K^+, K^-, o') equivalent.

We note:

(i) The set o_X of homology orientations has exactly two elements.

(ii) An orientation-preserving diffeomorphism $\phi\colon X \to X'$ induces a $o_\Phi\colon o_{X'} \to o_X$. One has $o_{\mathrm{id}_X} = \mathrm{id}$, and $o_{\Phi_2 \circ \Phi_1} = o_{\Phi_1} \circ o_{\Phi_2}$.

25.2 Conjugation of Spin^c-structures

Let $\mathfrak{s} = (\mathbb{S}, \rho)$ be a Spin^c-structure on an oriented Riemannian 4-manifold X. Its conjugate $\bar{\mathfrak{s}}$ is $(\overline{\mathbb{S}}, \rho)$. That is: one changes the action of $\mathbb{C}$ on the spinor bundles by composing with conjugation $\mathbb{C} \to \mathbb{C}$: i now acts as $-i$ did. Note that for $e \in T^*X$, $\rho(e)$ is still $\mathbb{C}$-linear and skew-adjoint with $\rho(e)^2 = -|e|^2$; the splitting $\mathbb{S} = \mathbb{S}^+ \oplus \mathbb{S}^-$ is unaffected, as is the orientation condition. A similar process applies in any dimension.

One has $c_1(\bar{\mathfrak{s}}) = -c_1(\mathfrak{s})$.

In Lie group-theoretic terms, conjugation of Spin^c-structures, in dimension n, arises from the automorphism of $\mathsf{Spin}^c(n) = \frac{\mathsf{U}(1)\times\mathsf{Spin}(n)}{\pm(1,1)}$ given by conjugation on the $\mathsf{U}(1)$-factor.

25.3 Formulation of the invariants

Assume that $b^+(X) \geq 0$. The invariants then take the form of a map

$$\mathsf{sw}_{X,\sigma}\colon \mathsf{Spin}^c(X) \to \mathbb{Z},$$

from the isomorphism classes of Spin^c-structures to the integers, and depending only on the oriented manifold X and a homology orientation $\sigma \in o_X$.

They have the following basic properties:

(I) Signs: If $-\sigma$ is the other homology orientation,

$$\mathsf{sw}_{X,-\sigma} = -\,\mathsf{sw}_{X,\sigma}\,.$$

(II) Diffeomorphism-invariance: If $\Phi\colon X' \to X$ is an oriented diffeomorphism, and $\Phi^*\colon \mathsf{Spin}^c(X) \to \mathsf{Spin}^c(X')$ the resulting bijection, then

$$\mathsf{sw}_{X', o_\Phi \sigma}(\Phi^* \mathfrak{s}) = \mathsf{sw}_{X,\sigma}(\mathfrak{s}).$$

(III) sw_X has finite support.

(IV) Conjugation-invariance: $\mathsf{sw}_{X,\sigma}(\bar{\mathfrak{s}}) = (-1)^{1-b_1+b^+}\,\mathsf{sw}_{X,\sigma}(\mathfrak{s})$.

(V) Dimension: If $\mathsf{sw}_{X,\sigma}(\mathfrak{s}) \neq 0$ then $d(\mathfrak{s})$ is non-negative and even.

In the case where $d(\mathfrak{s}) = 0$—which in practice is by far the most important case[6]—$\mathsf{sw}_{X,\sigma}(\mathfrak{s})$ is a signed count of the finite set of gauge-equivalence classes of solutions to the SW equations.

In the case where $b^+(X) = 1$, there are still SW invariants but they depend on an additional datum, namely, a *chamber*. Recall that there is a space $\mathcal{V}$ of pairs $([g], \eta)$, where g is a conformal structure and η a g-self-dual 2-form. The space $\mathsf{conf}(X)$ of conformal structures is contractible (it can be identified with a convex set in a Fréchet space), and $\mathcal{V} \to \mathsf{conf}_X$ is a vector bundle. The *wall* $\mathcal{W}(\mathfrak{s}) \subset \mathcal{V}$ is the codimension 1 affine subbundle defined by

$$c_1(\mathfrak{s}) \in -2i\eta_{harm,\,g} + \mathcal{H}^-_{[g]} \subset H^2_{DR}(X).$$

Let $\mathsf{Spin}^c(X)_{ch}$ be the set of pairs $(\mathfrak{s}, c)$, where $\mathfrak{s} \in \mathsf{Spin}^c(X)$ and c is a chamber for $\mathfrak{s}$.

The SW invariant now takes form of a map

$$\mathsf{Spin}^c(X)_{ch} \to \mathbb{Z}.$$

It still satisfies properties (I), (II) and (V). It also satisfies (IV) if one understands conjugation to map the chamber of $([g], \eta)$ to that of $([g], -\eta)$. The finite support property is no longer true.

(VI) Wall-crossing formula. Suppose $b^+(X) = 1$ and $b_1(X) = 0$, and that $d(\mathfrak{s})$ is non-negative and even. If c_+ and c_- are the two chambers then

$$|\,\mathsf{sw}_{X,\sigma}(\mathfrak{s}, c_+) - \mathsf{sw}_{X,\sigma}(\mathfrak{s}, c_+)| = 1.$$

A more precise result specifies the sign of the difference, in terms of σ. A more complicated wall-crossing formula is available when $b_1 > 0$.

25.4 Configuration spaces

Let $\mathcal{C} = \mathcal{C}(\mathfrak{s})$ be the space of configurations. It has various quotients:

- $\mathcal{B}_x = \mathcal{C}(\mathfrak{s})/\mathcal{G}^x$ the quotient by the based gauge group.
- $\mathcal{B} = \mathcal{B}_x/\mathsf{U}(1)$.
- $\widehat{\mathcal{B}}_x = \mathcal{C}(\mathfrak{s})/\mathcal{G}^x_\circ$ (quotient by the identity component of the based gauge group). Note that the Coulomb slice projects diffeomorphically to this quotient.
- $\widehat{\mathcal{B}} = \widehat{\mathcal{B}}_x/\mathsf{U}(1)$.

[6] A manifold with $b^+ > 1$ and $b_1 = 0$ is said to have simple type if $d(\mathfrak{s}) = 0$ on $\operatorname{supp} \mathsf{sw}_X$. The simple type conjecture, which is open, claims that if X is simply connected with $b^+ > 1$ it is of simple type.

We add superscripts ^{irr} when we want to consider only irreducible configurations.

The projection $\mathcal{B}^{irr}_x(\mathfrak{s}) \to \mathcal{B}^{irr}(\mathfrak{s})$ is a principal $\mathsf{U}(1)$ bundle. Its associated line bundle has a first Chern class $c \in H^2(\mathcal{B}^{irr}(\mathfrak{s}))$.

When $b^+(X) > 0$, generic pairs (g, η) have the property that there are no reducible solutions to the SW equations $\mathfrak{F}_\eta = 0$ for $\mathfrak{s}$, and the irreducible solutions are cut out transversely in Coulomb gauge. Their gauge-orbits then form a compact manifold

$$M_\eta(\mathfrak{s}) \subset \mathcal{B}^{irr}(\mathfrak{s}),$$

of dimension $d(\mathfrak{s})$.

We saw that $\mathcal{G}^x_\circ$ orbits of $\mathsf{U}(1)$-connections are parametrized by

$$H^1_{DR}(X) \times \operatorname{im} d^*,$$

and $\mathcal{G}^x$-orbits by

$$\frac{H^1(X;\mathbb{R})}{H^1(X;\mathbb{Z})} \times \operatorname{im} d^*.$$

Hence

$$\begin{aligned} \widehat{\mathcal{B}}_x &\cong H^1_{DR}(X) \times \operatorname{im} d^* \times \Gamma(\mathbb{S}^+); \\ \widehat{\mathcal{B}} &\cong H^1_{DR}(X) \times \operatorname{im} d^* \times \Gamma(\mathbb{S}^+)/\mathsf{U}(1); \\ \mathcal{B}_x &\cong \frac{H^1(X;\mathbb{R})}{H^1(X;\mathbb{Z})} \times \operatorname{im} d^* \times \Gamma(\mathbb{S}^+); \\ \mathcal{B} &\cong \frac{H^1(X;\mathbb{R})}{H^1(X;\mathbb{Z})} \times \operatorname{im} d^* \times \Gamma(\mathbb{S}^+)/\mathsf{U}(1). \end{aligned}$$

In the corresponding spaces of irreducible configurations, one removes 0 from $\Gamma(\mathbb{S}^+)$. In particular,

$$\mathcal{B}^{irr} \cong \frac{H^1(X;\mathbb{R})}{H^1(X;\mathbb{Z})} \times \operatorname{im} d^* \times \mathbb{P}\Gamma(\mathbb{S}^+) \times (0, \infty),$$

and thus $\mathcal{B}^{irr}$ deformation-retracts to a copy of

$$\frac{H^1(X;\mathbb{R})}{H^1(X;\mathbb{Z})} \times \mathbb{P}\Gamma(\mathbb{S}^+).$$

In this discussion, one can work with Sobolev L^2_k configurations.

Lemma 25.1 *Let H be a separable complex Hilbert space with orthonormal basis $(e_1, e_2, \dots)$. Consider the projective space $\mathbb{P}H$ and its subspace $\mathbb{C}P^\infty \bigcup_n \mathbb{P}\mathbb{C}\{e_1, \dots, e_n\}$. The inclusion of $\mathbb{C}P^\infty \to \mathbb{P}H$ is a weak homotopy equivalence.*

Proof We quote the facts that the sphere $S(H)$ and the union of finite-dimensional spheres S^∞ are contractible. The inclusion $\mathbb{C}P^\infty \to \mathbb{P}H$ lifts to a $\mathsf{U}(1)$-equivariant inclusion $S^\infty \to S(H)$. There are homotopy long exact sequences for the two $\mathsf{U}(1)$-bundles $S^\infty \to \mathbb{C}P^\infty$ and $S(H) \to \mathbb{P}H$, and the inclusions induce a map from one exact sequence to the other. This shows that $\mathbb{C}P^\infty \to \mathbb{P}H$ induces isomorphisms on π_i for all i. □

Thus the cohomology ring of $\mathbb{P}H$ is a polynomial ring. Applying this to $H = L^2_k(\mathbb{S}^+)$, one sees that for L^2_k configurations one has

$$H^*(\mathcal{B}^{irr}) \cong \Lambda\, H^1(X)^* \otimes \mathbb{Z}[c].$$

25.5 The construction

We have not yet considered orientations for M_η. The picture that will emerge is as follows:

Proposition 25.2 *There is a real line bundle* $\underline{\det \mathrm{ind}} \to \mathcal{B}$ *and a canonical isomorphism* $\underline{\det \mathrm{ind}}|_{M_\eta} \cong \det TM_\eta$ *for any* (g,η).

The space $\widehat{\mathcal{B}}$ is simply connected, so $\underline{\det \mathrm{ind}}$ is necessarily trivial over $\widehat{\mathcal{B}}$. It is in fact trivial over $\widehat{\mathcal{B}}$ too. A homology orientation σ determines a trivialization for $\underline{\det \mathrm{ind}}$, and therefore an orientation for TM_η.

Thus M_η carries a fundamental homology class

$$[M_\eta] \in H_{d(\mathfrak{s})}(\mathcal{B}^{irr}),$$

whose sign depends on σ. We define

$$\mathrm{sw}_{X,\sigma}(\mathfrak{s}) = \langle c^{d(\mathfrak{s})/2}, [M_\eta] \rangle$$

when $d(\mathfrak{s})$ is even and non-negative; and

$$\mathrm{sw}_{X,\sigma}(\mathfrak{s}) = 0$$

otherwise.

When $b_1 = 0$, this number is equivalent information to the class $[M_\eta]$.

When $b_1 > 0$, the homology class contains more information, and this can be packaged as an invariant valued in $\Lambda^* H^1(X)$; we shall not use this construction.

The crucial issue is well-definedness—invariance from the choice of (g,η). A path $(g_t,\eta_t)_{t\in[0,1]}$ determines a parametric SW moduli space N, formed from of pairs $(t,[\gamma])$, where $t \in [0,1]$ and γ a solution to $\mathfrak{F}_{\eta_t} = 0$. When $b^+ > 1$, this moduli space is, for generic paths, a smooth cobordism from M_{η_0} to M_{η_1}. The ambient space $[0,1] \otimes \mathcal{B}^{irr}$ carries the line bundle $\mathrm{pr}_1^* T[0,1] \otimes \mathrm{pr}_2^* \underline{\det \mathrm{ind}}$, which is oriented by $\partial_t \otimes o_\sigma$, o being the orientation of $\underline{\det \mathrm{ind}}$ corresponding to the homology orientation σ. With this convention, one sees that $[M_{\eta_0}] = [M_{\eta_1}]$ and so

$$\langle c^{d(\mathfrak{s})/2}, [M_{\eta_0}] \rangle = \langle c^{d(\mathfrak{s})/2}, [M_{\eta_1}] \rangle.$$

When $b^+ = 1$, the same argument works so long as the path stays in one chamber.

25.5.1 Properties

Consider the properties (I–V). Two of them (the sign rule (I) and the dimension constraint (V)) are immediate consequence of the definition. Diffeomorphism-invariance (II) is also immediate, once sw is well-defined independent of choices—for we can pull back the metric, 2-form and the Spin^c-structure using the chosen oriented diffeomorphism, so the equations pull back in the same way, and so do their spaces of solutions.

As for the finite support property (III), we saw that for a fixed metric, and for $\eta = 0$, there are only finitely many Spin^c-structures for which irreducible solutions exist. One can easily adapt the argument (I ought to have done so at the time) to show that the same is true for any *fixed* η (but *not* if one allows η to depend on the Spin^c-structure).

As for conjugation-invariance (I), the point is that the equations are unchanged by conjugation. Thus there is a homeomorphism $\mathcal{B}^{irr}(\mathfrak{s}) \to \mathcal{B}^{irr}(\mathfrak{s})$, respecting the class c, and mapping the moduli spaces $M_\eta(\mathfrak{s})$ to $M_\eta(\overline{\mathfrak{s}})$ by a diffeomorphism. The only question, then, is how this diffeomorphism acts on orientations. We postpone consideration of that point until we have discussed the determinant index bundle.

The wall-crossing formula lies a little deeper, and we will not prove it in this lecture.

25.6 Orientations

25.6.1 Virtual vector bundles

Fix a Hilbert space H (over $\mathbb{R}$ or $\mathbb{C}$), and let $\mathcal{U}$ denote the partially-ordered set of finite-dimensional subspaces $U \subset H$ under inclusion. A subset $\mathcal{U}' \subset \mathcal{U}$ is *cofinal* if every $U \in \mathcal{U}$ is contained in a member of $\mathcal{U}'$. In what follows, $\mathcal{U}$ will be the 'universe' from which our vector spaces are drawn.

Define a (real or complex) *stable vector bundle* over a space Z to be a pair $(V \to Z, U)$, where $V \to Z$ is a vector bundle and $U \in \mathcal{U}$. We write this bundle formally as $V - \underline{U}$. An isomorphism f of stable bundles, $V_0 - \underline{U}_0 \to V_1 - \underline{U}_1$, is a system of vector bundle isomorphisms $f_W \colon V_0 \oplus \underline{U}_1 \oplus \underline{W} \to V_1 \oplus \underline{U}_0 \oplus \underline{W}$, defined for a cofinal set of $W \in \mathcal{U}$, and compatible with orthogonal direct sums in H in that $f_{W\oplus W'} = f_W \oplus \mathrm{id}_{W'}$. Two isomorphisms are considered equal if they are equal for a cofinal set of W.

Now define a *virtual vector bundle* $V \to Z$ as follows. There is a collection of non-empty open subsets of $S \subset Z$ which we shall call 'small'. Smallness is preserved by passage to subsets, and every $z \in Z$ has a small neighborhood. For each small open set S, V assigns a stable vector bundle $V_S - \underline{U_S}$. For any two small open sets S_1 and S_2 one has a stable isomorphism $\theta_{S_1,S_2} \colon V_{S_1} - \underline{U_{S_1}} \to V_{S_1} - \underline{U_{S_1}}$ over $S_1 \cap S_2$, and these satisfy a cocycle condition on triple overlaps.

A formal difference of vector bundles $V_1 - V_2$ defines a virtual vector bundle, and when Z one can always so represent a virtual vector bundle.

25.6.2 The virtual index bundle

Let H_0 and H_1 be Hilbert spaces, and $\mathsf{Fred}(H_0,H_1)$ the space of Fredholm operators $L \colon H_0 \to H_1$. (Again, there are $\mathbb{R}$- and $\mathbb{C}$-linear versions.) It is an open subspace of $\mathcal{B}(H_0,H_1)$, the Banach space of bounded linear maps. It carries a continuous (i.e., locally constant) function $\mathrm{ind} \colon \mathsf{Fred}(H_0,H_1) \to \mathbb{Z}$. However, more is true: there is a well-defined *virtual vector bundle*

$$\underline{\mathrm{ind}} \to \mathsf{Fred}(H_0,H_1)$$

such that, for any $L \in \mathsf{Fred}(H_0,H_1)$, there is a stable isomorphism

$$\underline{\mathrm{ind}}_L = \ker L - \operatorname{coker} L.$$

To construct it, work with the universe $\mathcal{U}$ of finite-dimensional subsets of H_2. Consider a point $L_0 \in \mathsf{Fred}(H_0,H_1)$. Write $\operatorname{coker} L_0$ to mean $(\operatorname{im} L_0)^\perp$. The restriction $L_0^\perp := L_0 \colon (\ker L_0)^\perp \to \operatorname{im} L_0$ is an isomorphism. Let O_{L_0} be the open neighborhood of L_0 consisting of operators L such that the projected operator $L^\perp = \pi_{\operatorname{im} L_0} \circ L|_{(\ker L_0)^\perp} \colon (\ker L_0)^\perp \to \operatorname{im} L_0$ is an isomorphism.

We will define $\underline{\mathrm{ind}}|_{O(L_0)}$ as a stable vector bundle on O_{L_0}, namely

$$\underline{\mathrm{ind}}|_{O(L_0)}|_{O(L_0)} = K_{L_0} - \underline{C}_{L_0}.$$

Here

$$K_{L_0} = \{(L,x) \in O_{L_0} \times H_0 : Lx \in C\} = \{(L,x) : x \in \ker \pi_{C^\perp} \circ L)\},$$

and

$$C_{L_0} = (\operatorname{im} L_0)^\perp.$$

Note that $K_{L_0} \to O_{S_0}$ is a vector bundle, since $\ker \pi_{C^\perp} \circ L$ is the graph of a linear map $\ker L_0 \to (\ker L_0)^\perp$ depending continuously on L. Because of this, we find that on overlaps $O_{L_0} \cap O_{L_1}$, there are canonical identifications $K_{L_0} - C_{L_0} = K_{L_1} - C_{L_1}$, and these behave coherently on overlaps.

25.6.3 The determinant index bundle

Any stable vector bundle $V-\underline{U}$ over Z defines a determinant line bundle $\det V-\underline{U} = \det V \otimes \det \underline{U}^* \to Z$. Stable isomorphisms of stable vector bundles defines isomorphisms of determinant line bundles. Thus a virtual vector bundle V also defines a determinant line bundle $\det V$.

In particular, we have a well-defined determinant line bundle

$$\underline{\det \mathrm{ind}} \to \mathsf{Fred}(H_0, H_1).$$

In the case where H_0 and H_1 are complex, we have a continuous inclusion

$$i\colon \mathsf{Fred}_{\mathbb{C}}(H_0, H_1) \to \mathsf{Fred}_{\mathbb{R}}(H_0, H_1)$$

of the $\mathbb{C}$-linear into the $\mathbb{R}$-linear Fredholms. The kernel and cokernel of a $\mathbb{C}$-linear Fredholm L are canonically oriented $\mathbb{R}$-vector spaces. For this reason, there is a canonical homotopy class of trivializations

$$i^*\underline{\det \mathrm{ind}}_{\mathbb{R}} \cong \underline{\mathbb{R}}.$$

25.6.4 A family of Fredholm operators

On our $\mathsf{Spin}^{\mathsf{c}}$ 4-manifold, let

$$H_0 = L^2_k(iT^*X \oplus \mathbb{S}^+); \quad H_1 = L^2_{k-1}(\underline{\mathbb{R}} \oplus \Lambda^+ \oplus \mathbb{S}^-).$$

Define the map

$$\mathfrak{D}\colon \mathcal{C}(\mathfrak{s}) \to \mathsf{Fred}(H_0, H_1)$$

by

$$\mathfrak{D}_{a,\phi}(b,\chi) = \left(d^*b, \rho(d^+b) + (\phi\chi^* + \chi\phi^*)_0, D_{A_0}\chi + \rho(a)\chi + \rho(b)\phi\right),$$

and notice that, at a solution $(A_0 + a, \phi)$ to the SW equations $\mathcal{F}'_\eta = 0$ (with Coulomb gauge fixing), $\mathfrak{D}_{a,\phi}$ is the derivative of $\mathcal{F}'$.

One has a real line bundle $\tilde{\mathcal{L}} = \mathfrak{D}^*(\underline{\det \mathrm{ind}})$. Suppose that the space $\tilde{M}_\eta$ of solutions to $\mathcal{F}'_\eta = 0$ is cut out transversely; so $\operatorname{coker} \mathfrak{D}_{a,\phi} = 0$ for $(a,\phi) \in \tilde{M}_\eta$. Then $\mathfrak{D}^*(\underline{\det \mathrm{ind}})|_{M_\eta} = \det TM_\eta$.

The group $\mathsf{U}(1)$ of constant gauge transformations acts on H_0 (by $u \cdot (a,\phi) = (a, u\phi)$) and on H_1 (by $u\cdot(f,\alpha,\chi) = (f,\alpha,u\chi)$. One has $\mathfrak{D}_{u\cdot(a,\phi)} = u\mathfrak{D}_{(a,\phi)}$. Therefore $u \in \mathsf{U}(1)$ maps $\ker \mathfrak{D}_{(a,\phi)} \to \ker \mathfrak{D}_{u(a,\phi)}$ and $\operatorname{coker} \mathfrak{D}_{(a,\phi)} \to \operatorname{coker} \mathfrak{D}_{u(a,\phi)}$. The action of $\mathsf{U}(1)$ on $\mathcal{C}(\mathfrak{s})$ lifts to an action in the determinant line bundle $\tilde{\mathcal{L}}$. Thus $\tilde{\mathcal{L}}$ descends to a well-defined line bundle $\mathcal{L} \to \mathcal{C}(\mathfrak{s})/\mathsf{U}(1)$. When M^{irr}_η is cut out transversely, one has $\mathcal{L}|_{M^{irr}_\eta} = \det TM^{irr}_\eta$.

To orient M^{irr}_η, then, is to trivialize $\mathcal{L}$ over M^{irr}_η.

There is a homotopy $\{\mathfrak{D}_t\}_{t\in[0,1]}$ of $\mathsf{U}(1)$-equivariant maps $\mathcal{C}(\mathfrak{s}) \to \mathsf{Fred}(H_0, H_1)$, starting from $\mathfrak{D}_0 = \mathfrak{D}$, as follows:

$$\mathfrak{D}_{t;a,\phi}(b,\chi) = \left(d^*b, \rho(d^+b) + (1-t)(\phi\chi^* + \chi\phi^*)_0, D_{A_0}\chi + \rho(a)\chi + (1-t)\rho(b)\phi\right).$$

Thus

$$\mathfrak{D}_{1;a,\phi}(b,\chi) = \left(d^*b, \rho(d^+b), D_{A_0+a}\chi\right).$$

Pullbacks of vector bundles by homotopic maps are isomorphic; an equivariant trivialization of $\mathfrak{D}_1^* \det \mathrm{ind}$ determines an equivariant trivialization of $\mathfrak{D}_0^* \det \mathrm{ind}$ (up to homotopy).

Notice that $\mathfrak{D}_1$ is the sum of two maps: the constant map $(a,\phi) \mapsto d^* \oplus \rho \circ d^+$, and the map $D\colon (a,\phi) \mapsto D_{A+a_0}$. Thus $\mathfrak{D}_1^*(\det \operatorname{ind}) = \det\ker(d^* \oplus \rho d^+) \otimes \det\operatorname{coker}(d^* \oplus d^+)^* \otimes \det\operatorname{ind} D$. Now D maps to the $\mathbb{C}$-linear Fredholm operators, and as such, $\det\operatorname{ind} D$ is trivial. Thus to trivialize $\mathfrak{D}_1^*(\det \operatorname{ind})$ is to trivialize the line $\det\ker(d^* \oplus \rho d^+) \otimes \det\operatorname{coker}(d^* \oplus d^+)^*$ Now this line is

$$\det\ker(d^* \oplus \rho d^+) \otimes \det\operatorname{coker}(d^* \oplus d^+)^* = \det \mathcal{H}^1 \otimes \det(\mathbb{R} \oplus \mathcal{H}^+)^* = \det(\mathcal{H}^1 \oplus \mathbb{R}^* \oplus (\mathcal{H}^+)^*).$$

A trivialization of it is exactly a homology orientation.

This construction explains the orientability of M_η.

There is one small omission in this account: I have not explained how to see that the action of $\pi_0\mathcal{G} = H^1(X;\mathbb{Z})$ respects the orientation.

The sign arises in the conjugation-invariance property (II) arises because conjugation acts $\mathbb{C}$-antilinearly in $\ker D$ and $\operatorname{coker} D$, and therefore acts in $\det\operatorname{ind} D$ with sign $(-1)^{\operatorname{ind} D_A^+}$. Conjugation acts trivially in $\det\operatorname{ind}(d^* \oplus d^+)$. Since $d(\mathfrak{s})$ is even, the parities of $\operatorname{ind} D_A^+$ and of $\operatorname{ind}(d^* \oplus d^+)$ agree, so the overall sign for the action on $\mathfrak{D}^* \det\operatorname{ind}$ is $(-1)^{\operatorname{ind}(d^* \oplus d^+)} = (-1)^{1-b_1+b^+}$.

26 Taubes's constraints on symplectic 4-manifolds

26.1 The canonical $\mathsf{Spin}^{\mathsf{c}}$-structure

Let $(V, \langle\cdot,\rangle)$ be a $2n$-dimensional positive-definite inner product space, and $J \in \mathsf{SO}(V)$ an orthogonal complex structure: $J^2 = -\mathrm{id}_V$. Then $V^* \otimes \mathbb{C} = \mathsf{Hom}_{\mathbb{R}}(V, \mathbb{C})$ comes with a polarization:

$$V \otimes \mathbb{C} = V^{1,0} \oplus V^{0,1}.$$

Here $V^{1,0}$ is the $+i$-eigenspace of J^*, that is, the space of $\mathbb{C}$-linear maps $V \to \mathbb{C}$; and $V^{0,1}$ the $-i$-eigenspace, the antilinear maps. Both $V^{1,0}$ and $V^{0,1}$ are isotropic.

In Lecture 15, we described how a polarization gives rise to a spinor representation. In this case, it is on

$$S = \Lambda^{\bullet} V^{0,1}.$$

Tracing through the 'creation minus annihilation formula' established there, one finds that the Clifford map

$$\rho\colon V^* \to \mathrm{End}_{\mathbb{C}} S$$

is given by

$$\rho(e) = \sqrt{2}\left(e^{0,1} \wedge \cdot - \iota(\overline{e^{1,0}})\right).$$

Here the contraction operation uses the metric: for $f \in V^* \otimes \mathbb{C}$, $\iota(f)$ is the derivation of degree -1 such that $\iota(f)e = \langle e, f\rangle$ for $e \in V^*$.

The splitting of the spinors, $S = S^+ \oplus S^-$, is the even/odd splitting of the $(0, \bullet)$-forms.

Globally, if (M^{2n}, g) is a Riemannian manifold, and $J \in \Gamma(\mathsf{SO}(TM))$ an almost complex structure ($J^2 = -\mathrm{id}$), then J determines an orientation on M and a Clifford module,

$$\mathbb{S} = \Lambda^{\bullet,0}(T^*M),$$

with Clifford map

$$\rho\colon T^*M \to \mathrm{End}\,\mathbb{S}, \quad \rho(e) = \sqrt{2}\left(\iota(e^{0,1}) \wedge \cdot - \iota(\overline{e^{1,0}})\right).$$

This module is actually a $\mathsf{Spin}^{\mathsf{c}}$-structure; in Lie group terms, it arises because the natural inclusion $\mathsf{U}(n) \to \mathsf{SO}(2n)$ factors through $\mathsf{Spin}^{\mathsf{c}}(2n)$.

Homotopic almost complex structures determine isomorphic $\mathsf{Spin}^{\mathsf{c}}$-structures.

Lemma 26.1 *In the case of a closed, oriented 4-manifold X, if $\mathfrak{s}$ is the $\mathsf{Spin}^{\mathsf{c}}$-structure arising from an almost complex structure, compatible with the given orientation, then $d(\mathfrak{s}) = 0$.*

Proof One has $c_1(\mathfrak{s}) = c_1(\Lambda_J^{0,2}) = c_1((\Lambda^{2,0})^*) = c_1(\Lambda^2_{\mathbb{C}}(TX, J)) = c_1(TX, J)$. Since TX is complex,

$$p_1(TX)[X] = (c_1(TX)^2 - 2c_2(TX))[X] = c_1(\mathfrak{s})^2[X] - 2\chi(X),$$

so by Hirzebruch's signature theorem, $c_1(\mathfrak{s})^2[X] - 2\chi(X) = 3\tau(X)$, hence $d(\mathfrak{s}) = 0$. □

Remark. In fact, an obstruction-theoretic calculation (found in Hirzebruch–Hopf, *Felder von Flächenementen in vier-dimensionalen Mannigfaltigkeiten*, section 4.6, but attributed there to Wu Wen-Tsun) shows that a sufficient condition for a 4-manifold to be almost complex is that there is a characteristic vector c with $c^2 = p_1(TX) + 2e(TX)$, so if $d(\mathfrak{s}) = 0$ then $\mathfrak{s}$ arises from an almost complex structure.

If (M, ω) is now a *symplectic* manifold—so ω is a non-degenerate 2-form with $d\omega = 0$—then M admits compatible almost complex structures: those for which $\omega(u, Ju) > 0$ for $u \geq 0$ and for which $\omega(u, Jv) = -\omega(Ju, v)$, so that if we put $g(u, v) = \omega(u, Jv)$ then g is a Riemannian metric. Compatible almost complex structures form a contractible space (see e.g. D. McDuff and D. Salamon, *Introduction to symplectic topology*). Thus a symplectic manifold admits a canonical $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}_{can}$, well-defined up to isomorphism (and pinned down precisely by a choice of compatible J).

26.2 Statement of the constraints

In general, on a symplectic manifold (M, ω), the *canonical line bundle* is the complex line bundle $K_X = \mathsf{det}_{\mathbb{C}}(T^*X, J)$, where J is a compatible almost complex structure (different J's result in isomorphic line bundles).

When (X^4, ω) is symplectic, the canonical Spin^c-structure $\mathfrak{s}_{can}$ has positive spinors $\mathbb{S}^+_{can} = \Lambda^{0,0}_J \oplus \Lambda^{0,2}_J$.

Theorem 26.2 (Taubes) *Let (X, ω) be a closed symplectic 4-manifold with $b^+ > 1$ and $K = c_1(\Lambda^{2,0}_J)$. Then there is a canonical solution (A_{can}, ϕ_{can}) to the Dirac equation for $\mathfrak{s}_{can}$. Moreover, for $\tau > 0$, $(A_{can}, \tau^{1/2}\phi_{can})$ is a solution to the SW equations with 2-form*

$$\eta(\tau) = iF(A^\circ_{can})^+ + \tfrac{1}{2}\tau\omega.$$

Check constants

We will call $(A_{can}, \tau^{1/2}\phi_{can})$ the 'Taubes monopole' with 'Taubes parameter' τ.

It is convenient to trivialize the $H^2(X)$-torsor $\mathsf{Spin}^c(X)$ using $\mathfrak{s}_{can}$. Thus $\mathsf{Spin}^c(X) \cong H^2(X)$ and $\mathfrak{s}_{can} \mapsto 0$. Note that conjugation of Spin^c-structures corresponds to $c \mapsto K - e$, where $K = c_1(\Lambda^{2,0}_J)$.

We view sw_X as a function of $H^2(X)$.

Theorem 26.3 (Taubes) *Let (X, ω) be a closed symplectic 4-manifold with $b^+ > 1$. Then there is a canonical homology orientation σ for which the following hold:*

(1) *For τ sufficien tly large, the Taubes monopole is the unique solution to $\mathcal{F}_{\eta(\tau)} = 0$ (modulo gauge), and for generic J it is regular. One has*

$$\mathsf{sw}_{X,\sigma}(0) = 1,$$

and therefore $\mathsf{sw}_{X,\sigma}(K) = (-1)^{1+b_1+b^+}$.

(2) *If $\mathsf{sw}_X(c) \neq 0$ then one has*

$$0 \leq c \cdot [\omega] \leq K \cdot [\omega],$$

and if one of the inequalities is an equality then c is 0 or K.

If one takes a closed, oriented manifold M^{2n}, and a class $w \in H^2(M; \mathbb{R})$, and asks whether there exists a symplectic form ω, compatible with the orientation, with $[\omega] = w$, there are two basic constraints: one must have $w^n[M] > 0$, and M must admit almost complex structures compatible with the orientation. Amazingly, in dimension $2n \geq 6$, these are the *only* known constraints. In dimension 4, the almost complex structure J and the class w must obey the constraints dictated by (1) and (2). When X is simply connected, these are the only known further constraints.[7] However, there are examples of manifolds obeying Taubes's constraints where the existence of symplectic forms is an open problem (for instance knot surgery on a K3 surface along a knot with monic Alexander polynomial).

The following corollary is remarkable:

Corollary 26.4 *Let (X, ω) be a closed symplectic 4-manifold with $b^+ > 1$. Then $K \cdot [\omega] \geq 0$.*

[7] When X is not simply connected, Taubes's constraints apply also on the finite coverings of X.

This corollary is the starting point for a successful classification of 4-manifolds admitting symplectic forms with $K_X \cdot [\omega] < 0$—they have $b^+ = 1$, by the theorem. They turn out to be diffeomorphic to certain complex surfaces, namely, blow-ups of $\mathbb{C}P^2$, $\mathbb{C}P^1 \times \mathbb{C}P^1$, or ruled surfaces.

Taubes's constraints are only the simpler part of the picture established by Taubes, in which SW invariants are in fact counts of J-holomorphic curves in X. For instance, a deeper reason why $c_1(K_X) \cdot [\omega] \geq 0$: it is that there exists a C^∞-section of K_X whose (transverse) zero-set S—an oriented surface in X—is symplectic: $\omega|_{TS} > 0$.

Taubes's constraints are sometimes sufficient to compute SW invariants. For instance:

Corollary 26.5 *Let (X, ω) be a closed symplectic 4-manifold with $b^+ > 1$ and K_X torsion in $H^2(X; \mathbb{Z})$. Then $\mathsf{sw}_X(\mathfrak{s}_{can}) = \pm 1$, and $\mathfrak{s}_{can}$ is the only Spin^c-structure with non-vanishing Seiberg–Witten invariant.*

This applies to K3 surfaces, for instance.

26.3 Geometry of almost complex manifolds

An almost complex structure J makes TM into a complex vector bundle. It also defines the splitting of the complexified cotangent bundle,

$$(T^*M) \otimes \mathbb{C} = \mathsf{Hom}_{\mathbb{R}}(TM, \mathbb{C}) = T^{1,0}M \oplus T^{0,1}M$$

into the complex-linear maps and the complex anti-linear maps $TM \to \mathbb{C}$. The projection

$$\pi^{1,0} = \tfrac{1}{2}(1 - iJ) \colon T^*M \to T^{1,0}M$$

is an $\mathbb{C}$-linear isomorphism.

The real exterior powers of the cotangent bundle, after complexification, split up into (p,q)-forms:

$$\Lambda^k_{\mathbb{R}}(T^*M) \otimes \mathbb{C} = \Lambda^k_{\mathbb{C}}\left(T^*M \otimes \mathbb{C}\right) = \Lambda^k_{\mathbb{C}}\left(T^{1,0}M \oplus T^{0,1}M\right) = \bigoplus_{p+q=k} \Lambda^{p,q}$$

where

$$\Lambda^{p,q} = \Lambda^p_{\mathbb{C}} T^{1,0} \otimes \Lambda^q_{\mathbb{C}} T^{0,1}.$$

We write $\Omega^{p,q} = \Gamma(M, \Lambda^{p,q})$.

The exterior derivative d has components

$$\partial_J = \tfrac{1}{2}(1 - iJ) \circ d \colon \Omega^{p,q} \to \Omega^{p+1,q}, \quad \bar{\partial}_J = \tfrac{1}{2}(1 + iJ) \circ d \colon \Omega^{p,q+1} \to \Omega^{p,q+1}.$$

The *Nijenhuis tensor*

$$N_J \colon \Lambda^2 TX \to TX$$

is defined on vector fields by

$$N_J(u, v) = [Ju, Jv] - [u, v] - J[Ju, v] - J[u, Jv],$$

but is $C^\infty(X)$-bilinear, hence represented by a tensor. Extend N_J to a $\mathbb{C}$-linear map $\Lambda^2_{\mathbb{C}} TX_{\mathbb{C}} \to TX_{\mathbb{C}}$. It then takes $\Lambda^2 T_{0,1}$ to $T_{1,0}$; that is, if $Ju = -iu$ and $Jv = -iv$, then $JN_J[u, v] = iN_J[u, v]$, as one easily checks.

Dualizing N_J, we get a map $N_J^* \colon \Lambda^1 \to \Lambda^2$ sending $\Lambda^{1,0}$ to $\Lambda^{0,2}$.

Lemma 26.6 $\bar\partial_J^2 f = -\frac{1}{4}N_J^* \circ \partial_J f$.

Proof Recall that for a 1-form α and vector fields u and v, one has $(d\alpha)(u,v) = u \cdot \alpha(v) - v \cdot \alpha(u) - \alpha([u,v])$. The fact that $d^2 = 0$ says that $u \cdot df(v) - v \cdot df(u) = df([u,v])$.

Since $\bar\partial_J^2 f$ is a $(0,2)$-form, it suffices to check it on complex vector fields u and v to type $(0,1)$, i.e. $-i$-eigenvectors for J. Under this assumption,

$$\begin{aligned}(\bar\partial^2 f)(u,v) &= (d\,\bar\partial f)(u,v)\\ &= u((\bar\partial f)(v)) - v((\bar\partial f)(u)) - \bar\partial f([u,v])\\ &= u(df(v)) - v(df(u)) - \bar\partial f([u,v])\\ &= \partial f([u,v])\\ &= \partial f([u,v]^{1,0}).\end{aligned}$$

On the other hand, $N_J(u,v) = -\frac{1}{2}(1+Ji)[u,v] = -[u,v]^{1,0}$, whence the lemma. □

An almost complex structure J on M is *integrable* if M^{2n} is J is induced by local holomorphic charts. That is, any point of M lies in the image of a diffeomorphism $\Phi\colon U \to \tilde U \subset M$ from an open set $U \subset \mathbb{C}^n$ to an open subset of M, such that $i \circ D\Phi = D\Phi \circ J$. So M becomes a complex manifold.

If J is integrable, with holomorphic coordinates $(z_1,\dots,z_n)$, one has

$$\bar\partial_J\left(f\,dz_{i_1}\wedge\cdots\wedge dz_{i_p}\wedge d\bar z_{j_1}\dots d\bar z_{j_q}\right) = \sum_k \frac{\partial f}{\partial \bar z_k}\wedge dz_{i_1}\wedge\cdots\wedge dz_{i_p}\wedge d\bar z_{j_1}\dots d\bar z_{j_q},$$

from which it follows that

$$\bar\partial_J \circ \bar\partial_J = 0.$$

From the lemma, we see that if J is integrable, $N_J = 0$. The *Newlander–Nirenberg theorem* asserts that, conversely, $N_J = 0$ implies integrability.

26.4 Almost Kähler manifolds

Now consider a manifold M^{2n} with a *compatible triple* $(g, J\omega)$, consisting of a Riemannian metric g, an almost complex structure J, and a non-degenerate 2- form ω, related by

$$g(u,v) = \omega(u, Jv),$$

At a point $x \in X^2$, one can find an orthonormal basis $(e_1,\dots,e_n,f_1,\dots,f_n)$ for T_x^*X such that

$$\omega(x) = \sum e_k \wedge f_k, \quad Je_k = f_k,$$

A compatible triple defines an *almost Kähler structure* if $d\omega = 0$ (i.e., ω is symplectic). Note that $\omega(Ju,Jv) = \omega(u,v)$, so $\omega \in \Omega_J^{1,1}$. An almost Kähler structure is emphKähler if J is integrable.

Lemma 26.7 *Let $(E, (\cdot,\cdot)_E)$ be a hermitian vector bundle over the almost Kahler mänifold M. For a unitary connection A in E, one has*

$$(\bar\partial_A)^2 = -\frac{1}{4}N_J^* \circ \partial_A + F_A^{0,2}.$$

Proof The proof is similar to the one for $\bar\partial^2 f = -\frac{1}{4}N_J^*(\partial f)$. It suffices to evaluate both sides on complex vector fields u and v to type $(0,1)$. In that case,

$$\begin{aligned}
(\bar\partial^2\,\alpha)(u,v) &= (d_A\,\bar\partial_A\,\alpha)(u,v)\\
&= d_{A,u}((\bar\partial_A\,\alpha)(v)) - d_{B,v}((\bar\partial_A\,\alpha)(u)) - (\bar\partial_A\,\alpha)([u,v])\\
&= d_{A,u}((\nabla_A\alpha)(v)) - d_{B,v}((\nabla_A\alpha)(u)) - (\bar\partial_A\,\alpha)([u,v])\\
&= F_A(u,v)\alpha + (\partial_A\alpha)([u,v])\\
&= F_A^{0,2}(u,v)\alpha + (\partial_A\alpha)([u,v]^{1,0})\\
&= F_A^{0,2}(u,v)\alpha - \frac{1}{4}N_J^*\circ(\partial_A\alpha).
\end{aligned}$$

□

The following identities are cases of the 'Kähler identities', but they remain true in the almost Kähler case.

Let $L = \omega\wedge\cdot \in \mathrm{End}\,\Lambda_X^*$, and let L_ω^* be the g-adjoint operator.

Lemma 26.8 *Let (g,J,ω) be an almost Kähler structure on M^{2n}. Let ∇_A be a unitary connection in a hermitian vector bundle $(E,(\cdot,\cdot)_E)$ over the almost Kähler manifold M. Then one has identities*

(i) $\bar\partial_A^* = iL_\omega^*\circ\partial_A$ *acting on* $\Omega_X^{0,1}(E)$;

(ii) $\partial_A^* = -iL_\omega^*\circ\bar\partial_A$ *acting on* $\Omega^{1,0}(E)$;

Textbooks on complex manifolds often prove these identities by a computation valid on $\mathbb{C}^n$ with its standard Kähler structure, and then deduce them for Kähler manifolds by the principle that Kähler metrics are standard up to first order in suitable holomorphic coordinates. That route is not available here, but another argument (indicated briefly in Donaldson–Kronheimer's book in the Kähler setting) is:

Proof Identity (ii) is merely the conjugate of (i); we prove (i). We can conjugate $\alpha\in\Omega^{0,1}(E)$ to give $\bar\alpha\in\Omega^{1,0}(E)$ (note that this operation does not send E to $\bar E$!). We then have pointwise equations

$$n(\alpha\wedge\bar\alpha)_E\wedge\omega^{n-1} = -i|\alpha|^2\omega^n = -in!|\alpha|^2\mathrm{vol}_g,\quad \alpha\in\Omega_X^{0,1}(E).$$

Hence

$$\langle\alpha_1,\alpha_2\rangle_{L^2} = \langle\!\langle\alpha_1,\alpha_2\rangle\!\rangle := -\frac{i}{(n-1)!}\int_M(\alpha_1\wedge\overline{\alpha}_2)_E\wedge\omega^{n-1}.$$

The relevance of this observation is that the formal adjoint $\bar\partial_A^*$, which is defined using the L^2 hermitian product on E-valued $(0,1)$-forms, can equally be defined using $\langle\!\langle\cdot,\cdot\rangle\!\rangle$. We apply integration by parts to the derivative of the $(n-1,n)$-form $(\alpha,s)_E\wedge\omega^{n-1}$:

$$\begin{aligned}
0 &= \int_M d((\alpha,s)_E\wedge\omega^{n-1})\\
&= \int_M d((\alpha,s)_E\wedge\omega^{n-1}\\
&= \int_M (d_A\alpha,s)_E\wedge\omega^{n-1} + \int_M(\alpha\wedge\nabla_A s)_E\wedge\omega^{n-1}\\
&= \int_M (\partial_A\alpha,s)_E\wedge\omega^{n-1} + \int_M(\alpha\wedge\partial_A s)_E\wedge\omega^{n-1}.
\end{aligned}$$

Now, for $\gamma \in \Omega^{1,1}_X(E)$, one has $(\gamma, s)_E \wedge \omega^{n-1} = (n-1)! g(\gamma, s\omega)_E \, \mathsf{vol}_g$, where $g(\cdot,\cdot)_E$ combines the metric on Ω^2 with the hermitian product on E. We deduce that

$$\langle \bar\partial_A^* \alpha, s\rangle_{L^2} = \langle\!\langle \alpha, \bar\partial_A s\rangle\!\rangle = \langle\!\langle \alpha, \overline{\partial_A s}\rangle\!\rangle = i\int_M g(\partial_A\alpha, s\omega)_E \mathsf{vol} = i\langle L^*_\omega \partial_A \alpha, s\rangle_{L^2},$$

which shows that $\bar\partial_A^* = iL^*_\omega \partial_A$. □

Proposition 26.9 *In the same situation, the following Weitzenböck formula holds on* $\Gamma(E)$*:*

$$\tfrac{1}{2}\nabla_A^*\nabla_A = \bar\partial_A^* \bar\partial_A + iL^*_\omega(F_A).$$

Proof We see from the Kähler identities (i) and (ii) that

$$\nabla_A^* = (\partial_A + \bar\partial_A)^* = iL^*_\omega(-\bar\partial_A + \partial_A),$$

so

$$\nabla_A^*\nabla_A = iL^*_\omega(-\bar\partial_A + \partial_A)\circ(\partial_A + \bar\partial_A) = iL^*_\omega \circ [\partial_A, \bar\partial_A],$$

the latter equality being because L^*_ω kills the $(2,0)$ and $(0,2)$-forms. On the other hand,

$$2\,\bar\partial_A^*\,\bar\partial_A = 2iL^*_\omega \circ \partial_A\,\bar\partial_A,$$

and $F_A^{1,1} = \bar\partial_A\,\partial_A + \partial_A\,\bar\partial_A$, so

$$\nabla_A^*\nabla_A - 2\,\bar\partial_A^*\,\bar\partial_A = -2iL^*_\omega \circ (\bar\partial_A\,\partial_A + \partial_A\,\bar\partial_A) = -2iL^*_\omega(F_A).$$

□

26.5 Symplectic 4-manifolds

Now consider a 4-manifold X with a *compatible triple* $(g, J\omega)$, consisting of a Riemannian metric g, an almost complex structure J, and a symplectic form ω, related by

$$g(u, v) = \omega(u, Jv),$$

At a point $x \in X^2$, one can find an orthonormal basis (e_1, e_2, e_3, e_4) for $T^*_x X$ such that

$$\omega(x) = e_1 \wedge e_2 + e_3 \wedge e_4, \quad Je_1 = e_2, \quad Je_3 = e_4.$$

Thus we see that

$$|\omega|^2_g \equiv 2,$$

and that ω is g-self-dual, hence harmonic.

The splitting of the complex 2-forms is related to self-duality:

Lemma 26.10 *One has*

$$\Lambda^+_g \otimes \mathbb{C} = \Lambda^{2,0} \oplus \mathbb{C}\cdot\omega \oplus \Lambda^{0,2},$$
$$\Lambda^-_g \otimes \mathbb{C} = \Lambda^{1,1}_0,$$

where $\Lambda^{1,1}_0 = \omega^\perp \subset \Lambda^{1,1}$.

Proof With the e_j as before, let $\eta = e_1 + ie_2$ and $\eta' = e_3 + ie_4$. Then it is easy to check that $\star(\eta \wedge \eta') = \eta \wedge \eta'$. Conjugating, $\star(\overline{\eta} \wedge \overline{\eta}') = \overline{\eta} \wedge \overline{\eta}'$. We already know $\star\omega = \omega$. This gives Λ^+_g; takes its orthogonal complement to obtain $\Lambda^- \otimes \mathbb{C}$. □

Note that $L^*_\omega \omega = 2$, and $L^*(\Lambda^{1,1}_0) = 0$. Thus, if $\eta \in \Omega^+_X$ is a real self-dual form, one has

$$\eta = \eta^{2,0} + \tfrac{1}{2}(L^*\eta)\omega + \overline{\eta^{2,0}}.$$

On the other hand, the canonical $\mathsf{Spin}^{\mathsf{c}}$-structure has

$$\begin{aligned}\mathbb{S}^+ &= \Lambda_J^{\bullet,0} = \Lambda^{0,0} \oplus \Lambda^{0,2} = \underline{\mathbb{C}} \oplus \Lambda^{0,2},\\ \mathbb{S}^- &= \Lambda^{0,1}.\end{aligned}$$

One checks using the standard form of $\omega(x)$ that the element $\rho(\omega) \in \mathfrak{su}(\mathbb{C} \oplus \Lambda^{0,2})$ is diagonal:

$$\rho(\omega) = \begin{bmatrix} -2i & 0 \\ 0 & 2i \end{bmatrix}.$$

One has

$$\rho(\beta) = 2\begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}, \quad \rho(\beta) = 2\begin{bmatrix} 0 & \overline{\beta} \\ 0 & 0 \end{bmatrix}, \quad \beta \in \Lambda^{0,2}.$$

(Here $\overline{\beta} \in \Lambda^{2,0} = (\Lambda^{2,0})^*$.)

27 Taubes's constraints, continued

27.1 The canonical solution to the Dirac equation

There is a distinguished positive spinor

$$\phi_{can} = 1_X \in \Gamma(\Lambda_J^{0,0}),$$

nowhere vanishing on X. There is also a distinguished Clifford connection $A \in \mathcal{A}_{cl}(\mathbb{S}^+)$. Indeed, if $A \in \mathcal{A}_{cl}(\mathbb{S}^+)$ is a Clifford connection, with covariant derivative, then $\nabla_A \phi_{can} \in \Omega^1_X(\mathbb{S}^+)$. Any other Clifford connection takes the form $A + a \cdot \mathrm{id}_{\mathbb{S}^+}$, with $a \in i\Omega^1_X$; and

$$\nabla_{A+a}\phi_{can} = \nabla_A \phi_{can} + \rho(a)\phi_{can} = \nabla_A \phi_{can} + a \otimes 1_X.$$

Thus one has an affine-linear isomorphism

$$\mathcal{A}_{cl}(\mathbb{S}^+) \to i\Omega^1_X$$

assigning to A the component of $\nabla_A \phi_{can}$ in $\Omega^1_X(\Lambda^{0,0}) = \Omega^1_X(\mathbb{C})$.

The distinguished Clifford connection A_{can} is the one such that the component of $\nabla_A \phi_{can}$ in $\Omega^1_X(\Lambda^{0,0})$ is 0.

Remark. Note that a Clifford connection in $\mathbb{S}^+$ determines, and is determined by, a Clifford connection in $\mathbb{S}^-$ (since either one of them is equivalent to a unitary connection in $\Lambda^2\mathbb{S}^+ = \Lambda^2\mathbb{S}^-$). Thus we get from A_{can} a connection in $\Lambda^{0,1}$, or equally, in $T^*X \otimes \mathbb{C}$. This is *not* in general the Levi-Civita connection, for its torsion turns out to be N_J. It is the Levi-Civita connection precisely when J is integrable.

Now let $D^+ = D^+_{A_{can}}$ be the Dirac operator for the distinguished Clifford connection.

Theorem 27.1 *(i) One has $D^+\phi_{can} = 0$.*
(ii) One has $D^+ = \sqrt{2}\left(\bar\partial_J \oplus \bar\partial_J^*\right)$.

To prove the theorem, we will need a general fact about Clifford connections ∇. By definition, they make ρ parallel with respect to the Levi-Civita connection:

$$\nabla_v(\rho(e)\phi) - \rho(e)\nabla_v\phi = \rho(\nabla_v^{LC} e)\phi, \quad e \in \Omega^1_X.$$

We can make k-forms act on spinors, by the linear extension of the rule

$$\rho(e_1 \wedge \cdots \wedge e_k)\phi = \frac{1}{k!}\sum_{\sigma \in S_k} \mathsf{sign}(\sigma)\rho(e_{\sigma_1})\dots\rho(e_{\sigma_k})\phi.$$

(We have previously only considered the cases of 1-forms and 2-forms.) If α is a 1-form, one has

$$\rho(\alpha \wedge \beta) = \rho(\alpha)\rho(\beta) - (-1)^{\deg\beta}\rho(\beta)\rho(\alpha).$$

One can use this to prove the following lemma (I leave this to you):

Lemma 27.2 *The identity*

$$\nabla_v(\rho(\gamma)\phi) - \rho(\gamma)\nabla_v\phi = \rho(\nabla_v^{LC}\gamma)\phi.$$

remains true for all forms γ.

Lemma 27.3

$$\tilde{\rho}(\nabla(\rho(\gamma)\phi) - \rho(\delta\gamma)\phi = \tilde{\rho}(\rho(\gamma)\nabla\phi).$$

for k-forms γ and spinors ϕ. Here $\delta = d + d^*$, and $\tilde{\rho}$ is the composite

$$T^*X \otimes \mathbb{S} \xrightarrow{\rho\otimes\mathrm{id}} \operatorname{End}\mathbb{S} \otimes \mathbb{S} \xrightarrow{\mathrm{ev}} \mathbb{S}.$$

(Note that $\tilde{\rho}(\nabla(\rho(\gamma\phi) = D(\rho(\gamma)\phi)$, where D is the Dirac operator.)

Proof Fix ϕ and consider the left-hand-side as a defining a first-order operator L_ϕ in γ. Check using the known symbols of D and δ that L_ϕ is $C^\infty(X)$-linear in γ. It will suffice, then, to verify the identity at $x = 0$ when $\gamma = dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ in local coordinates x_i. We can choose these coordinates so that $(d^* dx_i)(0) = 0$. In that case, $\delta dx_I = 0$. In view of the previous lemma, the identity holds for dx_I if and only if

$$\tilde{\rho}(\rho(\nabla^{LC} dx_I)\phi) = 0.$$

Since ∇^{LC} is a graded derivation, it suffices to prove it for the 1-forms dx_i. Write

$$\nabla^{LC}(dx_i) = \sum_{j,k} \Gamma_i^{jk} dx_j \otimes dx_k,$$

and recall the symmetry $\Gamma_i^{jk} = \Gamma_i^{kj}$ arising from torsion-freeness, and $\Gamma_i^{jk} = -\Gamma_j^{ik}$ arising from orthogonality. Thus $\Gamma_i^{jk} = 0$ when $k = j$. We have

$$\tilde{\rho}(\rho(\nabla^{LC} dx_i) = \sum_{j,k} \Gamma_i^{jk} \rho(dx_j)\rho(dx_k) = -\sum_j \Gamma_i^{jj} \rho(dx_j)\rho(dx_j) = 0.$$

□

Proof of the theorem (i) We want to show $D^+\phi_{can} = 0$. Let $\Omega = \frac{1}{2i}\rho(\omega) \in \mathfrak{su}(\mathbb{S}^+)$. Then $\Omega\phi_{can} = -\phi_{can}$; and $\Omega(\nabla_v \phi_{can}) = \Omega(\nabla_v \phi_{can})$. Now $\delta\omega = 0$. So, by the lemma, we have

$$D^+(\Omega\phi_{can}) = \tilde{\rho} \circ \nabla(\Omega \cdot \phi_{can}) = \tilde{\rho}(\Omega \cdot \nabla\phi_{can}) = \tilde{\rho}(\nabla\phi_{can}) = D^+\phi_{can}.$$

But $D^+(\Omega\phi_{can}) = D^+(-\phi_{can}) = -D^+\phi_{can}$. So $D^+\phi_{can} = 0$.

(ii) The differential operators D^+ and $\sqrt{2}(\bar{\partial}_J + \bar{\partial}_J^*)$ have the same symbol, and both annihilate ϕ_{can}. Their difference A is algebraic, and annihilates ϕ_{can}. Hence $D^+\alpha = \sqrt{2}\,\bar{\partial}_J\,\alpha$ for $\alpha \in \Gamma(\Lambda^{0,0})$. For $\beta \in \Gamma(\Lambda^{0,2})$, one has $\beta = \frac{1}{2}\rho(\beta)\phi_{can}$, and $\rho(\beta)\nabla\phi_{can} = 0$ (because $(2,0)$-forms annihilate $(2,0)$-forms). Hence

$$D^+\beta = \tfrac{1}{2}\tilde{\rho}(\nabla(\rho(\beta)\phi_{can})) = \tfrac{1}{2}\rho(\delta\beta)\phi_{can} + \tfrac{1}{2}\tilde{\rho}(\rho(\beta)\nabla\phi_{can}) = \tfrac{1}{2}\rho(\delta\beta)\phi_{can}.$$

Now,

$$\rho(d^*\beta)\phi_{can} = \sqrt{2}(d^*\beta)^{0,1} = \sqrt{2}\,\bar{\partial}^*\,\beta,$$

and $\delta\beta = (1 + \star)d^*\beta = 2(d^*\beta)^+$, so $\rho(\delta\beta)\phi_{can} = 2\rho(d^*\beta)\phi_{can} = 2\sqrt{2}\,\bar{\partial}^*\,\beta$, which shows that

$$D^+\beta = \sqrt{2}\,\bar{\partial}^*\,\beta.$$

□

27.2 The SW equations

Work with $\mathfrak{s} = L \otimes \mathfrak{s}_{can}$. A Clifford connection A in $\mathbb{S}^+ = (\Lambda^{0,0} \oplus \Lambda^{2,0}) \otimes L$ then takes the form $\nabla_{can} \otimes \mathrm{id}_L + \mathrm{id} \otimes \nabla_B$, with ∇_{can} the distinguished connection in $\Lambda^{0,0} \oplus \Lambda^{2,0}$ and ∇_B a unitary connection in L.

Thinking of $\mathbb{S}^+$ as $L \oplus (\Lambda^{2,0} \otimes L)$, the Dirac operator for this connection is

$$D_B^+ = \sqrt{2}\left(\bar{\partial}_B \oplus \bar{\partial}_B^*\right),$$

with $\bar{\partial}_B = \nabla_B^{0,1}$ and $\bar{\partial}_B^*$ its formal adjoint. This assertion follows from the 'untwisted' case with $L = \underline{\mathbb{C}}$ and B trivial, which we have already established.

Now consider the curvature $F(A^\circ)$. One has $F(A^\circ) = F(\nabla_{can}^\circ) + 2F(B^\circ)$.

The SW curvature equation involves the self-dual part F^+ of $F = F(A^\circ)$. Since this is an imaginary, self-dual 2-form, one has

$$F^+ = F^{2,0} + (L^* F)\,\omega + \overline{F^{2,0}}.$$

Thus F^+ determines, and is determined by, $F^{0,2}$ and the function L^*F.

We have

$$\phi\phi^* = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \begin{bmatrix} \overline{\alpha} & \overline{\beta} \end{bmatrix} = \begin{bmatrix} |\alpha|^2 & \alpha\overline{\beta} \\ \overline{\alpha}\beta & |\beta|^2 \end{bmatrix}$$

and so

$$(\phi\phi^*)_0 = \begin{bmatrix} \frac{1}{2}(|\alpha|^2 - |\beta|^2) & \alpha\overline{\beta} \\ \overline{\alpha}\beta & \frac{1}{2}(|\beta|^2 - |\alpha|^2). \end{bmatrix}.$$

Write $\eta = F(A_{can}^\circ)^+/(4i) + \eta'$. Then the curvature equation $\frac{1}{2}\rho(F^+ - 4i\eta) = (\phi\phi^*)_0$ says that

$$\rho(F_B^+ - 2i\eta') = (\phi\phi^*)_0.$$

Taubes's choice is to take

$$\eta_0 = -\frac{1}{4}\tau\omega.$$

Then the SW equations can be written as

$$\bar{\partial}_B\,\alpha = -\,\bar{\partial}_B^*\,\beta, \tag{14}$$

$$F_B^{0,2} = \tfrac{1}{2}\overline{\alpha}\beta, \tag{15}$$

$$L_\omega^*(iF_B) = \frac{1}{4}\left(|\beta|^2 - |\alpha|^2 + \tau\right). \tag{16}$$

So the Taubes monopole is the solution with L trivial, B trivial, $\alpha = \tau^{1/2}$, $\beta = 0$.

27.3 Proof of Taubes's constraints

The main point is to prove the following

Proposition 27.4 *There is a constant $C = C(X, g, J)$ such that if $e = c_1(L)$, $e \cdot [\omega] \leq 0$, and (B, α, β) is a solution to the SW equations with Taubes parameter $\tau > C$ for the Spin^c-structure $L \otimes \mathfrak{s}_{can}$, then $e = 0$ and (B, α, β) is gauge-equivalent to the Taubes monopole.*

Proof For a solution (B,α,β) to the equations, with Taubes parameter τ, we compute

$$\begin{aligned}
\tfrac{1}{2}\|\nabla_B\alpha\|^2_{L^2} &= \tfrac{1}{2}\langle\nabla_B^*\nabla_B\alpha,\alpha\rangle_{L^2}\\
&= \langle\bar\partial_B^*\,\bar\partial_B\,\alpha,\alpha\rangle_{L^2} + \int_X L_\omega^*(iF_B)|\alpha|^2\,\mathsf{vol}_g\\
&= -\langle\bar\partial_B^*\,\bar\partial_B^*\,\beta,\alpha\rangle_{L^2} + \frac{1}{4}\int_X(|\beta|^2-|\alpha|^2+\tau)|\alpha|^2\,\mathsf{vol}_g\\
&= -\langle\beta,\bar\partial_B\,\bar\partial_B\,\alpha\rangle_{L^2} + \frac{1}{4}\int_X(|\beta|^2-|\alpha|^2+\tau)|\alpha|^2\,\mathsf{vol}_g\\
&= -\langle\beta,F_B^{0,2}\alpha\rangle_{L^2} + \frac{1}{4}\int_X(\beta,N_J^*\circ\partial_B\alpha) + \frac{1}{4}\int_X(|\beta|^2-|\alpha|^2+\tau)|\alpha|^2\,\mathsf{vol}_g\\
&= -\frac{1}{2}\int_X|\alpha|^2|\beta|^2\mathsf{vol}_g + \frac{1}{4}\int_X(\beta,N_J^*\circ\partial_B\alpha) + \frac{1}{4}\int_X(|\beta|^2-|\alpha|^2+\tau)|\alpha|^2\,\mathsf{vol}_g\\
&= \frac{1}{4}\int_X(\beta,N_J^*\circ\partial_B\alpha) - \frac{1}{4}\int_X\left((\tau-|\alpha|^2)^2+\tau(\tau-|\alpha|^2)+2|\alpha|^2|\beta|^2\right)\mathsf{vol}_g.
\end{aligned}$$

Our assumption is that $e\cdot[\omega]\le 0$. Thus

$$c_1(L)\wedge[\omega] = \frac{1}{2\pi}\int_X iF_B\wedge\omega = \frac{1}{4\pi}\int_X L_\omega^*(iF_B)\,\omega\wedge\omega\le 0,$$

and so

$$\int_X(|\beta|^2-|\alpha|^2+\tau)\,\mathsf{vol}_g\le 0,$$

with equality iff $e=0$. Thus when $e=0$,

$$\tfrac{1}{2}\|\nabla_B\alpha\|^2_{L^2} + \frac{1}{4}\int_X\left((\tau-|\alpha|^2)^2+\tau|\beta|^2+2|\alpha|^2|\beta|^2\right)\mathsf{vol}_g = \frac{1}{4}\int_X(\beta,N_J^*\circ\partial_B\alpha)\mathsf{vol}_g,$$

and the left-hand side is a sum of non-negative terms. More generally, when $e\cdot[\omega]\le 0$, this equality becomes an inequality $\le$. The right-hand side is bounded above by $C\|\beta\|_{L^2}\|\nabla_B\alpha\|_{L^2}$, for a constant $C=C(X,g,J)$. Thus

$$\|\nabla_B\alpha\|^2_{L^2} + \frac{\tau}{2}\|\beta\|^2_{L^2} + \frac{1}{2}\|(|\alpha|^2-\tau)\|^2_{L^2} \le 2C\|\beta\|_{L^2}\cdot\|\nabla_B\alpha\|_{L^2}\le C^2\|\beta\|^2_{L^2}+\|\nabla_B\alpha\|^2_{L^2},$$

i.e.

$$\left(\frac{\tau}{2}-C^2\right)\|\beta\|^2_{L^2} + \frac{1}{2}\|(|\alpha|^2-\tau)\|^2_{L^2}\le 0.$$

Taking $\tau>2C^2$, we find $\beta=0$ and $|\alpha|^2=\tau$. Since L admits a nowhere-zero section, it is a trivial line bundle, so $e=0$. Moreover, going back to our multi-line calculation above, we see that $\|\nabla_B\alpha\|^2_{L^2}=0$, so $\nabla_B\alpha=0$. Thus B admits a covariant-constant section, and so is a trivial connection. Thus (B,α,β) is gauge-equivalent to the Taubes monopole. □

This proposition shows that if $\mathsf{sw}_X(e)\neq 0$ then $e\cdot[\omega]\ge 0$, and if equality holds, $e=0$. Conjugation-invariance $\mathsf{sw}_X(K-e)=\pm\,\mathsf{sw}_X(e)$ then tells us that if $\mathsf{sw}_X(e)\neq 0$, $e\cdot[\omega]\le K\cdot[\omega]$, and that if equality holds, $e=K$.

Two points remain in the proof of Taubes's constraints. One is to prove that the Taubes monopole is regular for $\tau\gg 0$, so that $\mathsf{sw}_X(0)=\pm 1$. The other is to exhibit a canonical homology orientation such that $\mathsf{sw}_X(0)=+1$, and which is symplectomorphism-invariant. These notes presently omit both points.

28 The symplectic Thom conjecture

28.1 The minimal genus problem

Let X be a closed, orientable 4-manifold, and $\sigma \in H_2(X;\mathbb{Z})$. We know that σ can be represented as the fundamental class $[\Sigma]$ of an oriented, embedded surface $\Sigma \subset X$.

Lemma 28.1 *One can choose Σ to be path connected, and one can make its genus arbitrarily large.*

For one can lower the number of connected components of Σ by choosing an embedded path connecting points x_0 and x_1 on different components of Σ, removing small discs from Σ near x_0 and x_1, and inserting a thin tube around the path; this operation is homologically trivial. In $\mathbb{R}^4$ (or for that matter, $\mathbb{R}^3$) one can find embedded, null-homologous surfaces of arbitrary genus; inserting these into a chart, then connecting them to Σ as before, raises the genus.

Problem: What is the minimal genus of a connected surface representing σ?

A variant is to allow disconnected surfaces $\Sigma = \coprod \Sigma_i$, and to minimize the 'complexity'

$$\chi_-(\Sigma) = \sum_{g(\Sigma_i)>0} (2g(\Sigma_i) - 2)$$

over representatives Σ of σ.

This problem was attacked by Kronheimer–Mrowka, initially using instanton gauge theory, and later (by them and others) using Seiberg–Witten theory.

Definition 28.2 Assume X is oriented with $b^+(X) > 1$. A class $c \in H^2(X;\mathbb{Z})$ is called a *basic class* if it arises as $c_1(\mathfrak{s})$ for a Spin^c-structure $\mathfrak{s}$ such that $\mathsf{sw}_X(\mathfrak{s}) \neq 0$.

Theorem 28.3 *(adjunction inequality) Assume X is oriented with $b^+(X) > 1$, and suppose σ is represented by an embedded, oriented surface Σ with non-negative normal bundle (that is, each component Σ_i has non-negative self-intersection). Then, for every basic class c, one has*

$$\chi_-(\Sigma) \geq c \cdot \sigma + \sigma \cdot \sigma.$$

A generalization, allowing surfaces of negative self-intersection, is proved in P. Ozsváth and Z. Szabó's paper *The symplectic Thom conjecture*.

We shall use Taubes's results to deduce the following as a corollary:

Theorem 28.4 (Symplectic Thom conjecture) *Suppose that $b^+(X) > 1$ and that σ is represented by an embedded, oriented surface Σ with non-negative normal bundle and without spherical components. If there exists a symplectic form on X, compatible with the orientation, for which $\omega|_{T\Sigma} > 0$, then Σ minimizes χ_- among representatives of σ.*

This applies in particular to smooth complex curves Σ in a Kähler surface with $h^{20} > 1$.

Ozsváth–Szabó were able to relax the non-negativity assumption on Σ. There is a variant statement for the case $b^+ = 1$, which includes the case of $\mathbb{C}P^2$, but this case requires some care.

28.2 Proving the adjunction inequality

The proof of the adjunction inequality starts with the case where the normal bundle is trivial. The case of positive self-intersection is then deduced by a blow-up procedure, relying on a blow-up formula for the SW invariants. We have not proved such a formula, but our application to symplectic topology will circumvent this part of the story.

Proposition 28.5 *Assume $b^+(X) > 1$, and suppose σ is represented by an embedded, oriented surface Σ with trivial normal bundle. Then, for every basic class c, one has*

$$\chi_-(\Sigma) \geq \langle c, \sigma \rangle.$$

Proof The starting point is the Gauss–Bonnet formula, which tells us[8] that for a metric h on Σ, one has

$$\int_\Sigma \mathsf{scal}(h)\,\mathsf{vol}_h = 4\pi\chi(\Sigma).$$

The fact that c is a basic class tells us that, for any metric and any self-dual 2-from η (we will take $\eta = 0$), the SW equations admit a solution (for if there were not solution for (g, η) then the moduli space, being empty, would be cut out transversely). The relation between the SW equations and scalar curvature will allow us to deploy the Gauss–Bonnet formula.

Fix an oriented diffeomorphism of a neighborhood of Σ with neighborhood $\Sigma \times \bar{D}(2)$ (with $\Sigma \subset X$ appearing as $\Sigma \times \{0\}$). Here $\bar{D}(2)$ is the closed disc in $\mathbb{R}^2$ of radius 2. The region $\Sigma \times (\bar{D}(2) \setminus D(1))$ can then be identified, by a change of coordinates on the second factor, with

$$\Sigma \times S^1 \times [0, 1].$$

Let $y \in \mathbb{R}/\mathbb{Z} = S^1$ and $z \in [0, 1]$ be the standard coordinates. Let g_1 be a metric on X such that, on the cylinder $\Sigma \times S^1 \times [0, 1]$, g_1 takes the form

$$g_1 = h \oplus dy^2 \oplus dz^2,$$

with h a metric on Σ of constant scalar curvature and $\mathsf{vol}_h(\Sigma) = 1$. For $t \geq 1$, let X_t be the manifold obtained by 'stretching the neck', namely, replacing $\Sigma \times S^1 \times [0, 1]$ by $\Sigma \times S^1 \times [0, t]$ in the obvious fashion, and let g_t be metric $h \oplus ds^2 \oplus dt^2$ on X_t. There is a family of diffeomorphisms $X_1 \to X_t$ beginning at the identity for $t = 1$, and so for all t. the SW equations on (X_t, g_t) admits solutions for some $\mathsf{Spin}^{\mathsf{c}}$-structure $\mathfrak{s}$ with $c_1(\mathfrak{s}) = c$.

Notice that $\mathsf{scal}(g_t) = 4\pi(2g(\Sigma_i) - 2)$ on the cylindrical region around the component Σ_i. Let $s_-(t) = \max(0, -\mathsf{scal}(g_t))$. Then its L^2-norm over the cylindrical region $C_t = \Sigma \times S^1 \times [0, 1]$ is

$$\| - s_-(t)\|_{C_t, g_t} = -4\pi\chi_-(\Sigma)t^{1/2}.$$

If (A_t, ϕ_t) is a solution to the g_t-SW equations for $\mathfrak{s}$, we have pointwise bounds

$$|F_t^+|^2_{g_t} \leq \frac{1}{8} s_-(t)^2,$$

where $s_t = \max(0, -\mathsf{scal}(g_t))$ and $F_t = F(A_t^\circ)$; and resulting L^2 bounds over the cylinder,

$$\|F_t^+\|^2_{C_t, g_t} \leq \frac{1}{8}\|s_-(t)\|^2_{C_t, g_t} = 2\pi^2\chi_-^2 t.$$

Thus

$$\|F_t^+\|^2_{X_t, g_t} \leq 2\pi^2\chi_-^2 t + S,$$

[8] The most standard phrasing invokes the Gauss curvature $k = 2\,\mathsf{scal}$.

where S can be taken to be the norm-squared of s_- over the complement of C_t, a quantity that is independent of t.

Recall next that

$$\|F_t\|^2_{X_t,g_t} = \|F_t^+\|^2_{X_t,g_t} + \|F_t^-\|^2_{X_t,g_t} = 2\|F_t^+\|^2_{g_t} + 4\pi^2 c^2[X],$$

so

$$\|F_t\|^2_{X_t,g_t} \leq 4\pi^2\chi_-^2 t + C,$$

where C is independent of t, and

$$\|F_t\|_{X_t,g_t} \leq 2\pi\chi_- t^{1/2} + C^{1/2}.$$

On the other hand, if ω is any closed 2-form, one has

$$\|\omega\|_{g_t} \geq t^{1/2}\int_\Sigma \omega.$$

Indeed, $\int_{\Sigma\times\{(y,z)\}}\omega$ is independent of $(y,z)\in S^1\times[0,t]$, so

$$\begin{aligned}
\mathsf{vol}(S^1\times[0,t])\int_\Sigma\omega &= \int_{S^1\times[0,t]}\left(\int_\Sigma\omega\right)dy\,dz\\
&\leq \int_{S^1\times[0,t]}\left(\int_\Sigma|\omega|\mathsf{vol}_h\right)dy\,dz\\
&= \int_{\Sigma\times S^1\times[0,t]}|\omega|\mathsf{vol}_{g_t}\\
&\leq \mathsf{vol}(S^1\times[0,t])^{1/2}\left(\int_{\Sigma\times S^1\times[0,t]}|\omega|^2\mathsf{vol}_{g_t}\right)1/2\\
&\leq \mathsf{vol}(S^1\times[0,t])^{1/2}\|\omega\|_{g_t}.
\end{aligned}$$

Thus

$$\|F_t\|_{X_t,g_t} \geq 2s\pi t^{1/2}\langle c,\sigma\rangle,$$

and so

$$\langle c,\sigma\rangle \leq \chi_-(\Sigma) + \frac{C^{1/2}}{2\pi t^{1/2}}.$$

Taking $t\to\infty$, we get $\langle c,\sigma\rangle\leq\chi_-(\Sigma)$. □

28.2.1 Blowing up

One can blow up the open set $U\subset\mathbb{C}^2$ at a point $z\in U$ to obtain the complex surface $\tilde{U}$ and the blow-down map $\pi\colon\tilde{U}\to U$. The map π maps $\pi^{-1}(Y\setminus\{z\})$ biholomorphically to $U\setminus\{z\}$, but the fiber over z is $E=\mathbb{P}(T_zB)$, a 2-sphere. Its normal bundle is the tautological line bundle, hence its self-intersection is -1.

A smooth complex curve $C\subset B$ passing through z has a proper transform $\tilde{C}\subset\tilde{B}$, again a smooth complex curve, and $\pi\colon\tilde{C}\to C$ is biholomorphic.

In complex geometry, the importance of blowing up is partly its naturality—the blow-up is independent of coordinates, even though one way to construct goes via holomorphic coordinates. One can apply the blow-up operation near a point $x\in X$ in an oriented 4-manifold, using a complex chart near x, with the understanding the precise manifold constructed will depend on the choice of coordinates. It still comes, however, with a blow-down map $\pi\colon \tilde{X} \to X$.

The result $\tilde{X}$ is always diffeomorphic to the connected sum $X \# \overline{\mathbb{C}P^2}$. (It is the union, along S^3, of X minus a 4-ball and a tubular neighborhood ν of E; the complement of a ball in $\overline{\mathbb{C}P^2}$ is a 2-disc bundle over a 2-sphere C with self-intersection -1, and so is identified with ν).

Since $\tilde{X}$ is a connected sum, $H_2(\tilde{X})$ is the orthogonal direct sum $H_2(X) \oplus H_2(\overline{\mathbb{C}P^2}) = H_2(X) \oplus \mathbb{Z} \cdot e$ (here $e = [E]$). Moreover, the map $\pi_*\colon H_2(\tilde{X}) \to H_2(X)$ acts as the identity on the summand $H^2(X)$, while $\pi_* e = 0$.

Given an embedded surface Σ, one can choose the ball near x so that either Σ misses the chart altogether, or else Σ pass through x, and in that case, appears as a complex curve near x. Either way, it has a proper transform $\tilde{\Sigma}$.

Lemma 28.6 *If Σ passes through x, one has $[\tilde{\Sigma}] = \sigma - e$.*

Proof The fact that π restricts to a diffeomorphism $\tilde{\Sigma} \to \Sigma$ gives that $[\tilde{\Sigma}] = \Sigma + me$ for some m. And $\tilde{\Sigma}$ intersects E transversely at a point (namely, the complex line $T_x\Sigma$), and (like all transverse intersections in complex surfaces) the sign is $+1$; so $m = -1$. □

We have

$$(\sigma - e)^2 = \sigma \cdot \sigma - 1.$$

Thus by successively blowing up points of Σ, one can make a positive normal bundle trivial. The number of blow-ups required is $m = \sigma \cdot \sigma$.

In the adjunction inequality

$$\chi_-(\Sigma) \geq \langle c, \sigma \rangle + \sigma \cdot \sigma,$$

the two sides remain unchanged under blowing up, replacing c by $\tilde{c} = c + e$, Σ by its proper transform $\tilde{\Sigma}$, and therefore σ by $\tilde{\sigma} = [\tilde{\Sigma}] = \sigma - e$.

Thus it suffices to prove the inequality (for $\tilde{c} = c + e_1 + \cdots + e_m$) in the m-fold blow-up, in which $\tilde{\Sigma}$ has trivial normal bundle—the case we have just addressed. However, one needs to know tha $\tilde{c}$ is a basic class.

One has

$$\tilde{c}^2 = (c + e)^2 = c^2 - 1.$$

On the other hand $(2\chi + 3\tau)(\tilde{X}) = (2\chi + 3\tau)(X) - 1$. Thus if $\tilde{\mathfrak{s}}$ and $\mathfrak{s}$ are $\mathsf{Spin}^{\mathsf{c}}$-structures on $\tilde{X}$ and X respectively, which agree on the complement of a ball around x, and with $c_1(\tilde{\mathfrak{s}}) = c + s$, $c_1(\mathfrak{s}) = c$, one has

$$d(\tilde{\mathfrak{s}}) = d(\mathfrak{s}).$$

Theorem 28.7

$$\mathsf{sw}_{\tilde{X}}(\tilde{\mathfrak{s}}) = \pm\, \mathsf{sw}_X(\mathfrak{s}).$$

Thus $\tilde{c}$ is a basic class when c is, and we are done.

We will not prove this formula. The essential point is that $\overline{\mathbb{C}P^2}$, being negative-definite and admitting a metric of negative scalar curvature, has just one, reducible solution to the SW equations for that metric. One proves that one can 'glue' that to a solution to the equations for X, using a long neck $S^3 \times [-T, T]$, and that for T sufficiently large, this procedure accounts for all solutions. However, as far as the Thom conjecture is concerned, one can circumvent this blow-up formula.

28.3 The symplectic Thom conjecture

Theorem 28.8 *When Σ is a symplectic surface in the symplectic manifold (X,ω), one has the* adjunction formula

$$\chi(\Sigma) = \langle K_X, \sigma\rangle + \sigma \cdot \sigma.$$

Proof One can find an ω-compatible almost complex structure J such that $J(T\Sigma) = T\Sigma$ (first choose J on Σ, then extend; cf. the treatment of almost complex structures in McDuff–Salamon, *Introduction to symplectic topology*. One then has a short exact sequence of complex vector bundles

$$0 \to T\Sigma \to TX|_\Sigma \to N_\Sigma \to 0,$$

which implies that for each component Σ_i (with $\sigma_i = [|\Sigma_i]$), one has $\langle c_1(TX), \sigma_i\rangle = \langle c_1(T\Sigma_i), \sigma_i\rangle + \langle c_1(N_{\Sigma_i}), \sigma_i\rangle$, which is to say

$$2g(\Sigma_i) - 2 = \langle K_X, \sigma_i\rangle + \sigma_i \cdot \sigma_i.$$

□

On the other hand, if $b^+(X) > 1$, the fact that $K_{\tilde{X}}$ is a basic class (by Taubes) implies an adjunction inequality

$$\chi_-(S) \geq \langle K_X, \sigma\rangle + \sigma \cdot \sigma$$

in X, valid for any representative S of σ, symplectic or not. Thus, if the symplectic surface Σ has no spherical components, we see that it saturates the adjunction inequality, and so minimizes complexity.

To avoid invoking the blow-up formula for SW invariants, we note that blowing up is a symplectic operation (see McDuff–Salamon). Briefly, the point is that one can take a Darboux neighborhood of x—a closed 4-ball $D^4(r)$ of some (possibly small) radius r, and 'symplectic width' $\rho = \pi r^2$; and replace it by a suitable symplectic disc bundle over S^2, where S^2 has its $\mathsf{SO}(3)$-invariant volume-form normalized to have volume ρ. The result is a symplectic form $\tilde{\omega}$ over $\tilde{X}$, in which E is symplectic of volume ρ.

Lemma 28.9 $K_{\tilde{X}} = K_X + e$.

Proof Certainly $K_{\tilde{X}} - \pi^* K_X$ vanishes on $\tilde{X} \setminus E$, and so $K_{\tilde{X}} = \pi^* K_X + me$ for some $m \in \mathbb{Z}$. By the adjunction formula, $\langle K_{\tilde{X}}, e\rangle + e \cdot e = -2$, i.e. $\langle K_{\tilde{X}}, e\rangle = -1$, and this shows that $m = 1$. □

Thus we can apply Taubes in a suitable blow-up of X, instead of X itself, and obtain the same conclusion.

29 Wish-list

These are topics I should have liked to cover, given a little more time:

- Proof of the wall-crossing formula
- The Bauer–Furuta invariants, and Furuta's $5/4$-theorem.
- Spin 4-manifolds, $\mathrm{pin}(2)$-symmetry, and symplectic manifolds with $c_1 = 0$.
- Vanishing for connected sums.
- Gluing along 3-tori, and the knot surgery formula.
- The invariants of Kähler surfaces.
- $SW = Gr$ on symplectic 4-manifolds.

The bibliography is still under construction...

References

[1] D. Barden, *Simply connected 5-manifolds*, Annals of Math.

[2] C. Bohr, B. Hanke, D. Kotschick, *Cycles, submanifolds, and structures on normal bundles*.

[3] R. Bott, L. Tu,

[4] P. Deligne, *Notes on spinors*

[5] S. Donaldson, *An application of Yang–Mills theory to 4-dimensional topology*

[6] S. Donaldson, P. Kronheimer, *The geometry of four-manifolds*

[7] Earle, Eells

[8] M. Freedman, *Simply connected 4-manifolds*

[9] A. Hatcher, *Vector bundles and K-theory*

[10] J. Jost, *Riemannian geometry and geometric analysis*

[11] M. Kervaire, J. Milnor, *Groups of homotopy spheres*

[12] P. Kronheimer, T. Mrowka, *Monopoles and 3-manifolds*

[13] W. Lueck,

[14] J. Lurie,

[15] P. May, *A concise course in algebraic topology*

[16] J. Milnor, J. Stasheff, *Characteristic classes*

[17] J. Morgan, *The Seiberg–Witten equations and applications to smooth 4-manifolds*

[18] J.-P. Serre, *A course in arithmetic*

[19] R. Thom, *Quelques propriétés des variétés différentiables*

[20] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*.

[21] S. Weinberger, *Computers, rigidity and moduli*

[22] E. Witten, *Monopoles and four-manifolds*