Introductory notes on Donaldson and Seiberg-Witten theories and differential topology in four dimensions

Ben Mares

March 21, 2017

Abstract

These notes are based on my lecture courses given at SISSA for first-year graduate students during Winter 2013 and Winter 2015. The aim is to give a gentle introduction to gauge theory applied to four-dimensional topology. The emphasis is on the interplay between functional analysis and topology.

Contents

Chapter 1

Introduction

1.1 About these notes

These notes cover the following three subjects related to mathematical gauge theory applied to four-dimensional topology.

Background and context Before students can fully comprehend the proofs, they must first learn some formidable technical machinery. Additionally, appreciation of the theorems requires context. These notes aim to provide context while introducing the most crucial parts of the technical machinery.

Donaldson theory This is the study of SU(2) and SO(3) anti-self-dual instantons. Conceptually it is fairly simple, but the technicalities are extremely difficult. The functional analysis required to tackle these difficulties will be introduced but not fully developed.

Seiberg-Witten theory This is the theory of a U(1) gauge field coupled to a spinor. Conceptually it is more complicated, but since the technicalities are vastly simpler, a more thorough treatment will be given.

Historically, topological field theory (in the form of supersymmetric path integrals) was instrumental to the discovery of Seiberg-Witten theory, but sadly this aspect is beyond the scope of these notes. Nor will we discuss the very fruitful algebraic geometric approach to Donaldson theory (semistable sheaves).

Since Donaldson theory has since been mostly replaced by Seiberg-Witten theory, one might question why it should be presented. Not only is it historically interesting, but since Seiberg-Witten theory has been the primary focus of research for the past two decades, Donaldson theory now seems to be coming back into fashion.

In order to make these notes accessible to those with less rigorous mathematical backgrounds, some elementary topics are mentioned. For topics which are already well-covered by the existing literature, these notes summarize the most important aspects, often from a particular perspective. Be warned

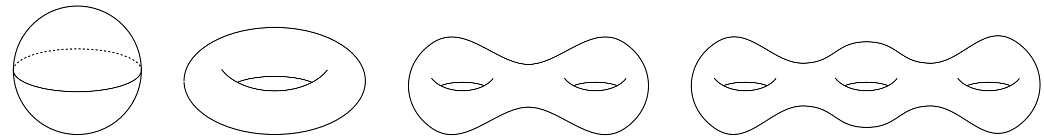


Figure 1.1: Every closed oriented two-dimensional manifold is determined up to diffeomorphism by its genus $g(\Sigma)$, which counts the number of handles.

that in order to emphasize certain intuitions, some definitions may be left incomplete, given only by example. However, such definitions are readily found in the references, and often in Wikipedia.

Notation and conventions are outlined in Section A.

1.2 A brief survey

Kevin Iga [Iga02] nicely summarizes the development of gauge theory as a tool in four-dimensional differential topology:

> In 1983, Donaldson shocked the topology world by using instantons from physics to prove new theorems about four-dimensional manifolds, and he developed new topological invariants. In 1988, Witten showed how these invariants could be obtained by correlation functions for a twisted $\mathcal{N} = 2$ SUSY gauge theory. In 1994, Seiberg and Witten discovered dualities for such theories, and in particular, developed a new way of looking at four-dimensional manifolds that turns out to be easier, and is conjectured to be equivalent to, Donaldson theory.

In conjunction with the results of Freedman and Taubes, Donaldson theory provided:

- Large classes of topological four-manifolds which admit no smooth structure
- Examples of "exotic" topological four-manifolds which admit multiple smooth structures, including $\mathbb{R}^4$
- Invariants which are sometimes capable of distinguishing smooth structures

Due to technical obstacles, the proofs from Donaldson theory were quite cumbersome. Upon the discovery of Seiberg-Witten theory, four-dimensional differential topology was revitalized with simpler proofs, simpler invariants, and new theorems. When the four-manifold is symplectic, Taubes proved that the Seiberg-Witten invariant corresponds to a certain count of pseudo-holomorphic curves (with respect to a generic almost-complex structure). In this case, the definition is very similar to that of the Gromov-Witten invariant.

One particular application of gauge theory I plan to focus on was one of the first major triumphs of Seiberg-Witten theory. Recall the classical result that the topology of any closed[1] connected

[1] A manifold is said to be *closed* when it is compact and has empty boundary.

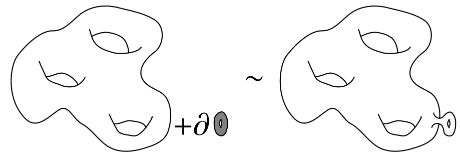


Figure 1.2: It is always possible to increase the genus of an embedded curve within a homology class adding the boundary of a small solid torus. Here the genus increases from three to four.

oriented two-dimensional manifold Σ is determined by its genus, as in Figure 1.1. The *minimal genus problem* asks: for a given smooth four-manifold X and a connected oriented two-dimensional submanifold Σ, when is it possible to reduce the genus of Σ without changing the topological class of Σ in the homology $H_2(X;\mathbb{Z})$?[2] Some simpler homology classes might be representable by smoothly embedded spheres or tori, while more complicated classes may require surfaces of higher genus.

In the case when our four-manifold X is a complex manifold and $\Sigma \hookrightarrow X$ is holomorphically embedded, the genus $g(\Sigma)$ is exactly determined by the *adjunction formula*

$$2\,g(\Sigma) - 2 = K([\Sigma]) + [\Sigma] \cdot [\Sigma]\,.$$

Here $g(\Sigma)$ is the genus, K is the canonical class,[3] and $[\Sigma] \cdot [\Sigma]$ is the self-intersection number.[4] For example on $\mathbb{CP}^2$, consider the surface Σ defined as the zero set of a generic homogeneous polynomial of degree $d > 0$. Then $[\Sigma] = d\,[H]$, where the hyperplane class $[H]$ generates $H_2(X;\mathbb{Z})$. Since two lines intersect in a point, $[H] \cdot [H] = 1$. The canonical class is $K(d\,[H]) = -3d$. Solving for $g(\Sigma)$ yields the classical formula $g(\Sigma) = \frac{1}{2}(d-1)(d-2)$.

One interesting consequence of the adjunction formula is that it's impossible to holomorphically increase the genus as in Figure 1.2.

Since smooth (but non-holomorphic) curves Σ can have genus greater than predicted by the adjunction formula, it's natural to wonder whether or not there are smooth curves of lower genus. For specific four-manifolds X, some specific results were known using geometric techniques. More partial results were achieved using difficult techniques from Donaldson theory.

The *Thom conjecture* asserts that inside $\mathbb{CP}^2$, any smooth surface Σ with $[\Sigma] = d\,[H]$ for $d > 0$ satisfies the adjunction inequality $g(\Sigma) \geq \frac{1}{2}(d-1)(d-2)$. Kronheimer and Mrowka solved the Thom conjecture by discovering a simple proof based on Seiberg-Witten theory [KM94]. Their idea yields a beautiful extension of the adjunction formula to an *adjunction inequality* for any four-manifold with some nonvanishing Seiberg-Witten invariant.

An even simpler proof of the adjunction inequality appears in [KM07, §40], which proves the main assertion by simple estimates using the Seiberg-Witten action (1.5). Indeed, by assuming some simple properties of Seiberg-Witten invariants and some identities, a proof is presented already in Section 1.4.

[2]Roughly speaking, an element of $H_2(X;\mathbb{Z})$ is like a closed, oriented, two-dimensional submanifold up to bordism within X. See Chapter B for more details.

[3]For our purposes, the canonical class is just some particular linear function $H^2(X;\mathbb{Z}) \to \mathbb{Z}$.

[4]The self-intersection number is described in Section B.7.

1.3 Overview of Donaldson theory

In four dimensions, there are infinite families of smooth manifolds which are homeomorphic, but not diffeomorphic.

$$X_i \overset{\text{diffeo}}{\not\simeq} X_j \quad X_i \overset{\text{homeo}}{\simeq} X_j$$

While the topology is the same, the notion of *calculus* is different. Thus to distinguish such manifolds, it makes sense to study solutions to differential equations. Gauge theory provides specific differential equations which are sufficiently natural and simple for this purpose.

The most basic differential equation from gauge theory is the *Yang-Mills equation* (1.1). It is a second-order PDE for a connection, which is denoted by A. In physics, a connection is called a *gauge potential*. The curvature of the connection is denoted by F_A, which physicists call the *gauge field*. Suppose X is any smooth manifold equipped with a metric g. The Yang-Mills equation is

$$d_A \star F_A = 0, \tag{1.1}$$

where $\star$ is the Hodge star operator associated to the metric g, and d_A is the exterior covariant derivative. All these concepts will be explained thoroughly in Section 1.8. All notations are summarized in Section A.

The manifold X corresponds physically to the space or spacetime. Geometrically, connections live in something called a principal bundle, which is a type of fiber bundle $P \to X$ whose fibers are a fixed compact Lie group G.

When the manifold X is Minkowski space, and when $G = \mathrm{U}(1)$, the group of unit complex numbers, the connection A can be identified with a one-form on X, which physically corresponds to the electromagnetic potential. The curvature $F_A = dA$ is the electromagnetic field. The Yang-Mills equation is equivalent to Maxwell's equations (with no charge or current). It describes the propogation of electromagnetic waves.

Our primary focus will be the next simplest case, where the gauge group is $\mathrm{SU}(2)$, so that A is a gauge potential corresponding physically to the weak force. Locally, A is represented by a one-form on X with values in the Lie algebra $\mathfrak{su}(2)$. The curvature of A (gauge field) is defined locally by [5]

$$F_A = dA + \tfrac{1}{2}\left[A \wedge A\right].$$

For our purposes, X will be a closed oriented Riemannian manifold. In contrast with the wave-type equation which arises on Minkowski space, the Yang-Mills equation with a positive-definite metric corresponds to a Laplace-type equation which describe statics.

In either case, the Yang-Mills equation arises as the Euler-Lagrange equation for stationary points of the Yang-Mills action functional

$$S_{\mathrm{YM}}(A) := \int_X |F_A|^2 .$$

[5] On first glance, the expression $[A \wedge A]$ looks like it must vanish. The Lie bracket is antisymmetric, and the wedge product of differential forms of odd-degree is also antisymmetric. However, the tensor product of two antisymmetric bilinear maps is a *symmetric* bilinear map, and there can be nonzero cross-terms. For instance, if $A = \alpha_1 \otimes \xi_1 + \alpha_2 \otimes \xi_2$, then $[A \wedge A] = 2\alpha_1 \wedge \alpha_2 \otimes [\xi_1, \xi_2]$.

In physics it is common to decompose the electromagnetic field into orthonormal electric and magnetic parts $|F_A|^2 = |E|^2 + |B|^2$, which depend on a choice of orthonormal basis for the spacetime. In contrast, when the metric is positive-definite, there is an orthonormal splitting $F_A = F_A^+ + F_A^-$ which is basis-independent. This leads to the a*nti-self-dual Yang-Mills equation* or *instanton equation*

$$F_A^+ = 0, \tag{1.2}$$

which is first-order in A.

The decomposition satisfies

$$\int_X \left(\left|F_A^-\right|^2 - \left|F_A^+\right|^2\right) = C_{\text{top}}, \tag{1.3}$$

where C_{top} is a constant depending only on the topology of the principal bundle P. Consequently,

$$S_{\text{YM}}(A) = \int_X \left(\left|F_A^+\right|^2 + \left|F_A^-\right|^2\right) = C_{\text{top}} + \int_X 2\left|F_A^+\right|^2. \tag{1.4}$$

From (1.4) we conclude that $S_{\text{YM}}(A) \geq C_{\text{top}}$ since the integral on the right hand side is nonnegative. Furthermore, $\mathrm{S}_{\text{YM}} = C_{\text{top}}$ is equivalent to $\int_X \left|F_A^+\right|^2 = 0$, which implies that F_A^+ is identically zero. Thus all solutions to (1.2) are absolute minimizers of S_{YM}, hence stationary points, and thus also satisfy the second-order Yang-Mills equation (1.1).

For any fixed bundle P, the space of connections $\mathscr{A}_P$ is an affine linear space, which is rather boring. However, the gauge group

$$\mathscr{G}_P := \operatorname{Aut}(P)$$

acts on $\mathscr{A}_P$, and the quotient space of physical states

$$\mathscr{B}_P := \mathscr{A}_P/\mathscr{G}_P$$

has interesting topology.

Informally speaking, the equation $F_A^+ = 0$ cuts out a finite-dimensional oriented submanifold

$$\mathscr{M}_{\text{ASD}} \subset \mathscr{B}_P.$$

Thus it defines a homology class

$$[\mathscr{M}_{\text{ASD}}] \subset H_d(\mathscr{B}_P),$$

where

$$d = \dim \mathscr{M}_{\text{ASD}}.$$

In reality, the situation is much more complicated due to technical issues, however this is the guiding idea. We can produce numbers by evaluating this on cohomology classes in $H^d(\mathscr{B}_P)$. This gives us the Donaldson invariants, which often detect smooth structures.

Note The ASD equation (1.2) depends on the choice of Riemannian metric on X, which determines the decomposition $F_A = F_A^+ + F_A^-$. Thus the moduli space $\mathscr{M}_{\text{ASD}}$ also depends on the metric. However, the Donaldson invariants are constructed in such a way that they are (usually) independent of this choice. Thus they depend only on the underlying smooth structure on X.

Warning These invariants are often called *topological* invariants. However, they are *not* well-defined in the sense of point-set topology! Instead, they are *differential-topological* invariants.

1.4 Overview of Seiberg-Witten theory

Seiberg-Witten invariants follow the same scheme but with different equations:

$$\begin{aligned} F_A^+ &= q(\phi), \\ \partial\!\!\!/_A \phi &= 0, \end{aligned}$$

where ϕ is a spinor, q is a quadratic map (unique up to a constant), and A is a U(1) connection (or electromagnetic potential). While this equation is conceptually more complicated due to the coupling with the spinor, the technicalities are vastly simpler. Seiberg-Witten invariants are conjectured to encode the same information as Donaldson invariants, and this conjecture has been rigorously established in most cases.

Since Donaldson theory was mostly replaced by Seiberg-Witten theory, one might question why I should present it. Not only is it historically interesting, but the past few years have seen a resurgence in Donaldson theory.

By assuming a few identities, right now I can prove some powerful theorems using Seiberg-Witten theory.

Define the Seiberg-Witten action

$$S_{\mathrm{SW}}(A, \phi) := \int_X \left(|\partial\!\!\!/_A \phi|^2 + \left| F_A^+ - q(\phi) \right|^2 \right). \tag{1.5}$$

Then $S_{\mathrm{SW}}(A, \phi) \geq 0$, with equality exactly when the Seiberg-Witten equations are satisfied. Expanding,

$$S_{\mathrm{SW}}(A, \phi) = \int_X \left(|\partial\!\!\!/_A \phi|^2 + \left| F_A^+ \right|^2 + |q(\phi)|^2 - 2 \left\langle F_A^+, q(\phi) \right\rangle \right).$$

There's an identity from differential geometry called the Lichnerowicz–Weitzenböck formula, which states that

$$0 = \int_X \left(|\nabla_A \phi|^2 - |\partial\!\!\!/_A \phi|^2 + \tfrac{1}{4} \mathrm{Sc}\, |\phi|^2 + 2 \left\langle F_A^+, q(\phi) \right\rangle \right),$$

where the function Sc is scalar curvature. Taking this as given, and adding it to $L_{\mathrm{SW}}(A, \phi)$, we get

$$S_{\mathrm{SW}}(A, \phi) = \int_X \left(|\nabla_A \phi|^2 + \tfrac{1}{4} \mathrm{Sc}\, |\phi|^2 + \left| F_A^+ \right|^2 + |q(\phi)|^2 \right).$$

Another useful identity is $|q(\phi)|^2 = \frac{1}{8} |\phi|^4$.

$$S_{\mathrm{SW}}(A, \phi) = \int_X \left(|\nabla_A \phi|^2 + \left| F_A^+ \right|^2 + \tfrac{1}{4} \mathrm{Sc}\, |\phi|^2 + |\phi|^4 \right).$$

Note that if $\mathrm{Sc} \geq 0$ everywhere, then the only way this integral can be zero is if $\phi \equiv 0$. In most generic cases, it is impossible to have solutions with $\phi \equiv 0$. Thus, if $\mathrm{Sc} \geq 0$, then all the Seiberg-Witten invariants must vanish.

Recall from (1.3) that $\int_X \left| F_A^+ \right|^2 = -\frac{1}{2} C_{\mathrm{top}} + \frac{1}{2} \int_X |F_A|^2$. Thus

$$S_{\mathrm{SW}}(A, \phi) = -\tfrac{1}{2} C_{\mathrm{top}} + \int_X \left(|\nabla_A \phi|^2 + \tfrac{1}{2} |F_A|^2 + \tfrac{1}{4} \mathrm{Sc}\, |\phi|^2 + \tfrac{1}{8} |\phi|^4 \right).$$

Completing the square,

$$S_{\mathrm{SW}}(A,\phi) = \int_X \left(|\nabla_A\phi|^2 + \tfrac{1}{2}|F_A|^2 + \tfrac{1}{8}\left(|\phi|^2 - (-\mathrm{Sc})\right)^2 \right) - C,$$

where

$$C := \left(\tfrac{1}{2} C_{\mathrm{top}} + \tfrac{1}{8}\int_X \mathrm{Sc}^2 \right).$$

This is an extremely powerful form of the action. Note that C depends only on the topology of the bundle and the geometry of X. The rest of the action is the sum of positive terms. It follows that for any solution,

$$\int_X |\nabla_A\phi|^2\,,\ \tfrac{1}{2}\int_X |F_A|^2\,,\ \tfrac{1}{8}\int_X \left(|\phi|^2 - (-\mathrm{Sc})\right)^2 \le C.$$

In a certain sense, $\nabla_A\phi$ and F_A cannot be too large, and $|\phi|$ cannot be too far from $\sqrt{-\mathrm{Sc}}$. With some simple Sobolev theory, these bounds imply compactness of the moduli space.

Seiberg-Witten invariants depend on a choice of something called a Spin^c-structure. Often[6] a Spin^c-structure $\mathfrak{s}$ is determined by its Chern class $c_1(\mathfrak{s}) \in H^2(X;\mathbb{Z})$. Assuming a mild topological condition on X (that $b^+(X) > 1$), then there is a map $\mathrm{SW} : \mathrm{Spin}^c(X) \to \mathbb{Z}$ which gives a signed count of solutions to the Seiberg-Witten equations.

Definition. For a smooth oriented closed 4-manifold X with $b^+(X) > 1$, a *Seiberg-Witten basic class* is a cohomology class $\kappa \in H^2(X;\mathbb{Z})$ such that there is a Spin^c-structure $\mathfrak{s}$ with $\mathrm{SW}(\mathfrak{s}) \neq 0$ and $c_1(\mathfrak{s}) = \kappa$.

If κ is a Seiberg-Witten basic class, then for any metric on X, there must exist solutions to the Seiberg-Witten equations associated to κ.

Theorem 1 (Adjunction inequality). *Suppose X is a smooth oriented closed 4-manifold with $b^+(X) > 1$ and a Seiberg-Witten basic class κ. If Σ is an embedded surface of genus $g \geq 1$ with $[\Sigma]\cdot[\Sigma] \geq 0$, then*

$$2g - 2 \geq |\kappa \cdot [\Sigma]| + [\Sigma]\cdot[\Sigma]\,.$$

The idea of the proof is quite nice. If $[\Sigma]\cdot[\Sigma] > 0$, then we can reduce to the case of $[\Sigma]\cdot[\Sigma] = 0$ Assuming for now that $[\Sigma]\cdot[\Sigma] = 0$, we wish to prove

$$2g - 2 \geq |\langle \kappa, [\Sigma]\rangle|\,.$$

Since X and Σ are oriented, the normal bundle to Σ is oriented. Rank two oriented vector bundles are classified up to isomorphism by their Euler class. Since the Euler class of the normal bundle is $[\Sigma]\cdot[\Sigma] = 0$, the normal bundle of Σ must be trivial. Thus Σ contains a tubular neighborhood of the form $\Sigma \times D^2 \hookrightarrow X$. Inside of D^2 we can find an annulus $S^1 \times [0,1]$. Inside this annulus is a smaller disk $\tilde{D}^2$. We can decompose X into three pieces: the annulus, and the two complementary pieces on either side:

$$X = X' \cup \left(\Sigma \times S^1 \times [0,1]\right) \cup \left(\Sigma \times \tilde{D}^2\right),$$

[6] when $H^2(X;\mathbb{Z})$ has no 2-torsion

where $X' := X \backslash (\Sigma \times D^2)$. Under the assumption that the Seiberg-Witten invariant is nonzero, we are guaranteed at least one solution to the Seiberg-Witten equations for *any* choice of Riemannian metric. Along the neck $N := \Sigma \times S^1 \times [0,1]$ we choose a product metric where S^1 has length 1, $[0,1]$ has length given by a parameter L, and Σ has area 1 and constant sectional curvature $2\pi(2-2g)$. Such a choice of metric on Σ is possible by uniformization and Gauss-Bonet. On the complement N^c of N, we fix some arbitrary metric. We will see that by sending $L \to \infty$, the behavior of the solution along the neck becomes dominant.

$$\begin{aligned} S_{\mathrm{SW}}(A,\phi) &\geq -\tfrac{1}{2}C_{\mathrm{top}} + \tfrac{1}{2}\int_X \left(|F_A|^2 - (\tfrac{1}{2}\mathrm{Sc})^2\right) \\ &\geq -\tfrac{1}{2}C_{\mathrm{top}} - \tfrac{1}{8}\int_{N^c} s^2 + \tfrac{1}{2}\int_N \left(|F_A|^2 - (\tfrac{1}{2}\mathrm{Sc})^2\right). \end{aligned}$$

The scalar curvature counts the sectional curvature twice, so $(\frac{1}{2}\mathrm{Sc})^2 = (2\pi(2-2g))^2$ along N.

Note that F_A is a cohomology representative of $-2\pi i c_1(\kappa)$. It follows that if $i_{\theta,t}$ denotes the inclusion $\Sigma \hookrightarrow \Sigma \times \{\theta\} \times \{t\}$, then $\int_\Sigma i^*_{\theta,t} F_A = -2\pi i\,(\kappa \cdot \Sigma)$. Since $i^*_{\theta,t}$ is a restriction which projects out components not parallel to Σ, it follows that $(2\pi\,(\kappa \cdot [\Sigma]))^2 \leq \left(\int_{\Sigma\times\{\theta\}\times\{t\}} |F_A|\right)^2 \leq \int_{\Sigma\times\{\theta\}\times\{t\}} |F_A|^2$. Thus

$$S_{\mathrm{SW}}(A,\phi) \geq \left(-\tfrac{1}{2}C_{\mathrm{top}} - \tfrac{1}{8}\int_{N^c} s^2\right) + \tfrac{1}{2}(2\pi)^2\left((\kappa \cdot [\Sigma])^2 - (2g-2)^2\right)L.$$

If (A,ϕ) is a solution, then $S_{\mathrm{SW}}(A,\phi) = 0$, so

$$0 \geq -C + \tfrac{1}{2}(2\pi)^2\left((\kappa \cdot [\Sigma])^2 - (2g-2)^2\right)L, \quad \forall L.$$

Since L can be made arbitrarily large, $(\kappa \cdot \Sigma)^2 - (2g-2)^2 \leq 0$. Since $g \geq 1$, we have $2g - 2 \geq 0$, and thus taking square roots,

$$|\kappa \cdot \Sigma| \leq 2 - 2g.$$

Now for the blowup argument. Suppose $[\Sigma]\cdot[\Sigma] \geq 0$. Let $\tilde{X} = X \# \overline{\mathbb{CP}^2}$, $\tilde{\Sigma} = \Sigma \# E$, and $\tilde{\kappa} = \kappa \pm [E]$, where $E \cong S^2 \cong \mathbb{CP}^1 \subset \mathbb{CP}^2$ is the exceptional curve. There is a "blowup formula" which implies that $\tilde{\kappa}$ is a basic class for $\tilde{X}$. We have $[E]\cdot[E] = -1$ and $[E]\cdot[\Sigma] = 0$, and $\langle \kappa, [E]\rangle = 0$. Thus

$$\left[\tilde{\Sigma}\right]\cdot\left[\tilde{\Sigma}\right] = ([\Sigma] + [E])\cdot([\Sigma]+[E]) = [\Sigma]\cdot[\Sigma] - 1,$$

and

$$\left|\left\langle \tilde{\kappa}, \left[\tilde{\Sigma}\right]\right\rangle\right| = |\langle \kappa \pm [E]\,, [\Sigma] + [E]\rangle| = |\langle \kappa, [\Sigma]\rangle \mp 1| = |\langle \kappa, [\Sigma]\rangle| + 1$$

upon choosing the appropriate sign. Thus

$$|\langle \kappa, [\Sigma]\rangle| + [\Sigma]\cdot[\Sigma] = \left|\left\langle \tilde{\kappa}, \left[\tilde{\Sigma}\right]\right\rangle\right| + \left[\tilde{\Sigma}\right]\cdot\left[\tilde{\Sigma}\right].$$

The genus of $\tilde{\Sigma}$ is the same as that of Σ. Thus by blowing up, we have replaced the adjunction inequality with an equivalent inequality where $\left[\tilde{\Sigma}\right]\cdot\left[\tilde{\Sigma}\right]$ is reduced by one. Iterating, we reduce to $\left[\tilde{\Sigma}\right]\cdot\left[\tilde{\Sigma}\right] = 0$. See [Law97] for more details, for instance the case $g = 0$.

1.5 Recommended references

For an easy-to-read panoramic view of the field, the best reference is certainly [Sco05]. It covers background, Donaldson theory, Seiberg-Witten theory, has several images, many nice geometric proofs, and plenty of useful references. However, the leisurely style comes at the cost of omitting analysis.

One of my favorite references for the basic theory of manifolds and differential forms from a geometric and physical perspective is [Fra12]. A much more sophisticated book which develops homological algebra and cohomology from the perspective of differential forms is [BT82]. Another introductory text which develops sheaf theory and Hodge theory is [War83].

For Donaldson theory, there are not many good references at the introductory level. Two of the best introductory textbooks are [FU91] and [Law85]. There are also the lecture notes [Mor98], together with more advanced topics in the same volume. Another important reference is [DK90] which presents a mixture of introductory and advanced topics. Advanced references include [FM94],

For Seiberg-Witten theory, there are many easy introductory references... [Mor96] [Sal] [Tau98] [Mar99] [Moo01] For spinors, [LM89]

1.6 Euclidean vs Lorentzian

1.7 Transversality

1.8 Principal bundles

1.9 Manifolds

Definition 2. A *topological manifold of dimension n* is a set equipped with an n-dimensional atlas, which is Hausdorff and second-countable.

An *$n-$dimensional atlas* on a set X is a cover $\{U_\alpha\}$ of X, and charts $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ such that

- each $V_\alpha \subset \mathbb{R}^n$ is open,
- each ϕ_α is a bijection, and
- each transition function $\phi_{\alpha\beta} := \phi_\beta \circ \phi_\alpha^{-1} : V_\alpha \to V_\beta$ is a homeomorphism.

Remark 3. Abstractly, manifolds begin life as a set, and inherit all their properties from their atlas. For example, subset of a manifold is *open* if it is open in each chart.

Remark. Given two different atlases on the same set, if their union is still an atlas, then the atlases are called *compatible*, and the resulting manifolds are considered equivalent.

Definition 4. An manifold is *smooth* if the transition functions are required instead to be diffeomorphisms.

Remark. Functions on smooth manifolds are *smooth* if they are smooth in each chart.

Definition 5. A smooth manifold is *orientable* if all transition functions $\phi_{\alpha\beta}$ can be chosen to be *orientation-preserving,* i.e. they satisfy

$$\det\left(\frac{\partial}{\partial x^j}\phi^i_{\alpha\beta}\right) > 0.$$

Remark 6. It's complicated, but one can extend this definition to topological manifolds.

Definition 7. Given an oriented manifold X, we define the *orientation-reversed manifold* $\bar{X}$ to be the same smooth manifold, but with the opposite orientation.

Complex conjugation on $\mathbb{CP}^n$ reverses orientation only for n odd.

1.10 Examples of exotic manifolds

Topological/smooth manifolds, together with continuous/smooth maps, form a *category*. This means that every manifold has an identity map, and maps can be composed. In any category, there is a notion of *isomorphism*, which is a map with a two-sided inverse.

Definition 8. A continuous map of topological manifolds $f : X_1 \to X_2$ is a *homeomorphism* if it is an isomorphism of topological manifolds, i.e. there exists a continuous $f^{-1} : X_2 \to X_1$ such that $f^{-1} \circ f = \mathrm{Id}_{X_1}$ and $f \circ f^{-1} = \mathrm{Id}_{X_2}$.

Definition 9. A smooth map of smooth manifolds $f : X_1 \to X_2$ is a *diffeomorphism* if it is an isomorphism of smooth manifolds.

It's easy to place multiple smooth structures on the same topological manifold. For example, consider two smooth atlases on the same copy of $\mathbb{R}$, giving two smooth manifolds which we denote by X_1 and X_2. On X_1 we use the atlas with the single chart $\phi = \mathrm{Id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$. On X_2 we use the single chart $\psi : \mathbb{R} \to \mathbb{R}$ by $\psi(x) = x^3$. Individually, these are each clearly smooth atlases, since the only transition function is the identity. These two atlases are compatible topologically, since $\psi \circ \phi^{-1} = x \mapsto x^3$ and $\phi \circ \psi^{-1} = x \mapsto x^{1/3}$ are homeomorphisms. Thus X_1 and X_2 are the same topological manifold. However, they are not smoothly the compatible, since $x^{1/3}$ is not smooth.

We should not get too excited, since X_1 and X_2 are diffeomorphic. In particular, the map $X_1 \to X_2$ given by $x \mapsto x^{1/3}$ is a diffeomorphism. (Remember, smoothness of a map is defined in terms of coordinate charts!)

What we really want to understand is the difference between *diffeomorphism classes* of smooth manifolds, and *homeomorphism classes* of topological manifolds. Visualizing examples is not easy, due to the following result:

Theorem 10 (Moise's Theorem (with others)). *Let X be a topological manifold of dimension $d \leq 3$. Then X admits a smooth structure, unique up to diffeomorphism.*

The first examples of exotic smooth structures were discovered by Milnor on the 7-sphere S^7. There are 28 distinct smooth structures on S^7. They can be realized explicitly as the manifolds obtained by the equations

$$\begin{aligned} a^2 + b^2 + c^2 + d^3 + e^{6k-1} &= 0, \\ |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 &= \varepsilon, \\ (a, b, c, d, e) &\in \mathbb{C}^5, \end{aligned}$$

for $\varepsilon > 0$ small, and $k = 1, \ldots, 28$. Perhaps it is best to view exotic structures as distinct manifolds which are "accidentally" homeomorphic.

Chapter 2

Classification of manifolds

Now let's examine classification of manifolds. In dimensions three and below, homeomorphims classes and diffeomorphism classes agree, so we sloppily refer to "isomorphism" of manifolds to avoid the distinction.

We might as well start in dimension zero, where a manifold is by definition a (countable) collection of discrete points. Any manifold is a disjoint union of its connected components, so it makes sense to study only connected manifolds. The only connected 0-manifold is a point.

Connected 1-manifolds are diffeomorphic to either $\mathbb{R}$ or S^1. This is a good point to mention the notion of a *manifold with boundary*.

Definition 11. A *manifold with boundary* is a manifold which is locally isomorphic to "relatively open" subsets of the closed half-plane $\left\{\vec{x} \in \mathbb{R}^n | x^1 \geq 0\right\}$. Points of the manifold corresponding in some (any) chart to points with $x^1 = 0$ are called *boundary points*.

Remark. A profound statement which somewhat underlies the foundation of homology theory is this: given a manifold with boundary, its boundary is a manifold without boundary. Symbolically, $\partial^2 = 0$, where ∂ is the operator which takes a manifold and gives its boundary. The relation $d^2 = 0$ is the dual statement under Stokes' theorem.

There are a total of four connected 1-manifolds with boundary.

	compact	noncompact
empty boundary	S^1	$(0,1)$
nonempty boundary	$[0,1]$	$[0,1)$

(Note that $(0,1] \cong [0,\infty)$ by the diffeomorphism $x \mapsto x^{-1} - x$.)

In dimension 2 (surfaces) can be quite nasty in general. (Consider for example $\mathbb{R}^2$ – cantor set.) The situation becomes much nicer if we restrict to compact manifolds. Compact surfaces with boundary must have boundary which is compact with empty boundary, i.e. finitely many copies of S^1. For simplicity, we consider only surfaces without boundary.

For example, we have S^2, the torus $T^2 = T$, as well as our first examples of non-orientable surfaces $S^2/\text{antipodal} = \mathbb{RP}^2 = P$, and the Klein bottle K.

Using the operation of connected sum, we can form a composite surface from two new ones. This induces an abelian monoid (=group without inverse axiom) structure on isomorphism classes of surfaces. We obtain the relations

$$\begin{aligned} S^2 \# X &= X, \quad \forall X, \\ P \# P &= K, \\ P \# P \# P = K \# P &= T \# P. \end{aligned}$$

Remark. S^2 is the identity of the monoid.

Remark. These generators and relations are complete, i.e. the resulting monoid is isomorphic to the monoid of isomorphism classes of connected compact surfaces.

Remark. The monoid is generated by T and P (the second relation eliminates K). Given a word in T and P, if P appears, then by the last relation we can trade T for P^2.

Thus the isomorphism classes correspond to the orientable surfaces

$$\Sigma_g := T^{\# g}, \quad g \geq 0, \quad \left(\Sigma_0 := S^2\right)$$

plus the non-orientable surfaces

$$P^{\# k}, \quad k > 0.$$

From here, we would want to show two things:

- every compact connected 2-manifold is isomorphic to one of these examples, and
- these examples are distinct.

There are various ways to prove the first statement, but they all tend to be fairly combinatorial, so they are of little interest to us. Furthermore, the corresponding statement in four dimensions is hopeless, since there is no conjectured enumeration of four-manifolds.

The second statement is far more interesting for our purposes.

Definition 12. For $p = 0, \ldots, n$, the p-th *Betti number* of a manifold X is $b^p(X) := \dim H^p(X)$.

The Betti numbers satisfy many nice properties:

- $b^0(X) = \#$components. Thus if X is connected, then $b^0(X) = 1$.
- (Poincaré duality) If X is an oriented compact n-manifold, then $b^p = b^{n-p}$.
- If X is a connected compact n-manifold, then

$$b^n(X) = \begin{cases} 1 & \text{if } X \text{is orientable,} \\ 0 & \text{if } X \text{is non-orientable.} \end{cases}$$

When all the Betti numbers are finite, it's possible to define the Euler characteristic $\chi = b^0 - b^1 + b^2 - \cdots$. This has some especially nice properties which make it very easy to compute. From any long exact sequence

$$\cdots \to H^{p-1}(Z) \to H^p(X) \to H^p(Y) \to H^p(Z) \to H^{p+1}(X) \to \cdots,$$

it follows that $\chi(Y) = \chi(X) + \chi(Z)$. For instance, from the Mayer-Vietoris sequence, if A and B are open, then

$$\chi(A) + \chi(B) = \chi(A \cap B) + \chi(A \cup B).$$

If there is a finite-dimensional chain complex $\mathscr{C}^\bullet(X)$ which computes the cohomology of X, then $\chi(\mathscr{C}^\bullet(X)) = \chi(X)$. For example, based on a triangulation, simplicial cohomology gives $\chi(\Sigma) = V - E + F$ for any triangulated surface with V vertices, E edges, and F faces.

Also, for an unbranched $n : 1$ covering $X' \to X$ where X has a finite triangulation, $\chi(X') = n\chi(X)$.

One easily computes that

$$\chi(P^{\#k}) = 2 - k, \quad \chi(\Sigma_g) = 2 - 2g.$$

From the properties of Betti numbers and this Euler characteristic computation, it follows that for the connected compact surfaces,

$$\begin{aligned} b^0(P^{\#k}) &= 1, & b^0(\Sigma_g) &= 1, \\ b^2(P^{\#k}) &= 0, & b^2(\Sigma_g) &= 1, \\ b^1(P^{\#k}) &= k - 1, & b^1(\Sigma_g) &= 2g. \end{aligned}$$

In particular, $\left\{b^1, b^2\right\}$ form a complete set of invariants, equivalent to $\{\text{orientability}(X), \chi(X)\}$.

Assuming the classification, we observe that the pair consisting of $\{\text{orientability}(X), \chi(X)\}$ is a complete invariant, meaning that two manifolds are isomorphic iff they have the same such invariants.

Every nonorientable manifold X has an orientable double-cover. Consider the real line bundle $\Lambda^n T^*X$, and remove the zero section $\Lambda^n T^*X - 0$. An orientation corresponds to a section. To convert this to a double-cover, we want to collapse the two rays of each fiber to points. The fiber over x can be identified with $(\Lambda^n T_x^* X - \{0\})/\mathbb{R}_+$, where $\mathbb{R}_+$ denotes the multiplicative group of positive numbers. If X were oriented, then this cover would have a section, and thus the double-cover would be two disjoint copies of X. Instead, in the nonorientable case, the double-cover is connected. In two dimensions, the double-cover is $\Sigma_{k-1} \to P^{\#k}$. The Euler characteristic is multiplicative under covering spaces, so we verify $\chi(\Sigma_{k-1}) = 2\chi(P^{\#k})$.

If X is non-orientable with orientable double-cover $X' \to X$, then the involution $\sigma : X' \to X'$ which swaps the fibers is fixed-point-free and orientation-reversing. Furthermore, any such involution on an orientable manifold determines a non-orientable manifold X'/σ.

Since a non-orientable manifold is equivalent to an orientable manifold with an involution, we now focus on only on the orientable case.

In our case of oriented surfaces, note that b^1 is always even. We can explain this via a slight refinement of Poincaré duality.

Recall that $H^\bullet_{\mathrm{dR}}(X)$ is a skew-commutative graded ring, meaning that for $a \in H^p_{\mathrm{dR}}(X)$, $b \in H^q_{\mathrm{dR}}(X)$,

$$a \smile b = (-1)^{pq}\, b \smile a \in H^{p+q}_{\mathrm{dR}}(X).$$

We get the following refinement of Poincaré duality:

Theorem 13. *If X is a compact, oriented, connected n-manifold, then for each p, the cup product is a nondegenerate bilinear map*

$$H^p_{\mathrm{dR}}(X) \times H^{n-p}_{\mathrm{dR}}(X) \to H^n_{\mathrm{dR}}(X) \cong \mathbb{R}.$$

Specifically, if we use the natural identification of $H^n_{\mathrm{dR}}(X)$ with $\mathbb{R}$, then the cup product induces a map

$$\begin{aligned} H^p_{\mathrm{dR}}(X) &\to \left(H^{n-p}_{\mathrm{dR}}(X)\right)^* \\ a &\mapsto (b \mapsto a \smile b \in \mathbb{R}). \end{aligned}$$

Definition 14. A bilinear map is *nondegenerate* when this map is an isomorphism.

Our previous notion of Poincaré duality follows from

$$b^p = \dim H^p_{\mathrm{dR}}(X) = \dim \left(H^{n-p}_{\mathrm{dR}}(X)\right)^* = b^{n-p}.$$

This enhanced version of Poincaré duality also detects certain intrinsic constraints on the cup product structure of $H^\bullet(X)$. For example, if $n/2$ is an odd integer, i.e. $n = 2, 6, 10, \ldots$, then the cup product

$$H^{n/2}_{\mathrm{dR}}(X) \times H^{n/2}_{\mathrm{dR}}(X) \to \mathbb{R}$$

is antisymmetric! Choosing any basis of $H^{n/2}_{\mathrm{dR}}(X)$, we obtain a nondegenerate antisymmetric matrix. Thus the eigenvalues are nonzero, purely imaginary, and come in conjugate pairs, so $b^{n/2}_{\mathrm{dR}}(X)$ is even. Using the antisymmetric version of Graham-Schmidt, it is possible to choose a "symplectic basis" such that the matrix takes the form

$$\begin{pmatrix} & -1 & & & \\ 1 & & & & \\ & & & -1 & \\ & & 1 & & \\ & & & & \ddots \end{pmatrix}.$$

The operation of connected sum amounts to a direct sum (block diagonal composition) on the level of intersection forms. The intersection form of T is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so the intersection form of $\Sigma_g = \underbrace{T\# \cdots \# T}_{g}$ is

$$\underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{g}.$$

In this way, the structure of Σ_g is reflected in its cohomology.

If $n/2$ is an even integer, i.e. $n = 4, 8, 12, \ldots$, then

$$H^{n/2}_{\mathrm{dR}}(X) \times H^{n/2}_{\mathrm{dR}}(X) \to \mathbb{R}$$

is symmetric, so the eigenvalues are real and nonzero. We can choose a basis so that the matrix is

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix},$$

and we get two new invariants $b^+(X) = \#\text{positive eigenvalues}$ and $b^-(X) = \#\text{negative eigenvalues}$ which satisfy $b^+ + b^- = b^2$. The combination $\sigma(X) = b^+(X) - b^-(X)$ is called the *signature*.

Remark. Technically, it's wrong to talk about the eigenvalues of a bilinear form, since a bilinear form is not an endomorphism. The transformation law is different. (Under a change of basis, the matrix of a bilinear form transforms as $\mathbf{Q} \mapsto \mathbf{G}^T\mathbf{Q}\mathbf{G}$, while an endomorphism transforms as $\mathbf{L} \mapsto \mathbf{G}^{-1}\mathbf{L}\mathbf{G}$.) The actual eigenvalues depend on the choice of basis, but the number of positive eigenvalues of any matrix representing a bilinear form gives the maximal dimension of any positive-definite subspace.

A more sophisticated way of stating the chirality proof is that the intersection form is an invariant of oriented diffeomorphism. Orientation reversal flips the sign of the intersection form. Since the intersection forms (+1) of $\mathbb{CP}^2$ and (−1) of $\overline{\mathbb{CP}^2}$ are inequivalent as bilinear forms, $\mathbb{CP}^2$ and $\overline{\mathbb{CP}^2}$ cannot be oriented-diffeomorphic.

More generally, if X is smooth, compact, oriented, and dimension n with $n/2$ is even, then the intersection form Q is symmetric, and

$$b^2(X) \text{ odd} \implies \sigma(X) \neq 0 \iff Q \nsim -Q \implies X \text{ chiral}.$$

(Remarkably, the $\iff$ in the middle also holds true over the integers.)

Gauge theory is sensitive to orientation. The moduli spaces for $\mathbb{CP}^2$ and $\overline{\mathbb{CP}^2}$ look completely different. Thus we will consider them distinct oriented manifolds.

If we use integer cohomology, then we get even more structure. The cup product induces an isomorphism of $\mathbb{Z}$-modules

$$\frac{H^p(X;\mathbb{Z})}{\text{torsion}} \to \left(\frac{H^{n-p}(X;\mathbb{Z})}{\text{torsion}}\right)^*.$$

If n is even, then choosing a basis,

$$\frac{H^{n/2}(X;\mathbb{Z})}{\text{torsion}} \simeq \mathbb{Z}^{b^{n/2}}.$$

The dual basis gives

$$\left(\frac{H^{n/2}(X;\mathbb{Z})}{\text{torsion}}\right)^* \simeq \left(\mathbb{Z}^{b^{n/2}}\right)^* \cong \mathbb{Z}^{b^{n/2}}.$$

Thus the matrix $\mathbf{Q}$ corresponding cup product is a square matrix of length $b^{n/2}$ with integer entries which is invertible over the integers. An integer bilinear form which is invertible over the integers is called *unimodular*.

In terms of the basis, the cup product $v \smile w$ corresponding to vectors $\mathbf{v}$ and $\mathbf{w}$ is $\mathbf{v}^T\mathbf{Q}\mathbf{w}$. If $\mathbf{v} \mapsto \mathbf{G}\mathbf{v}$, then $\mathbf{Q} \mapsto \left(\mathbf{G}^{-1}\right)^T \mathbf{Q}\mathbf{G}^{-1}$ to preserve

$$(\mathbf{G}\mathbf{v})^T \left(\left(\mathbf{G}^{-1}\right)^T \mathbf{Q}\mathbf{G}^{-1}\right)(\mathbf{G}\mathbf{w}) = \mathbf{v}^T\mathbf{Q}\mathbf{w}.$$

In contrast, an endomorphism $\mathbf{L}$ transforms as $\mathbf{L} \mapsto \mathbf{G}\mathbf{L}\mathbf{G}^{-1}$.

Remark. An integer matrix is invertible over the integers iff the determinant is ± 1. This can be seen explicitly via the formula $A^{-1} = \operatorname{adj}(A)/\det A$, where $\operatorname{adj}(A)$ is the transpose of the cofactor matrix (no division). It is natural to denote such matrices by $\mathrm{GL}(n;\mathbb{Z})$, but not in this particular context. Here it is more natural to interpret the "invertibility" as the condition that the duality map $H^{n/2}(X;\mathbb{Z}) \to \left(H^{n/2}(X;\mathbb{Z})\right)^*$ is an isomorphism of $\mathbb{Z}$-modules.

If $n/2$ is even, then Q is symmetric, and if $n/2$ is odd then Q is antisymmetric.

Unimodular antisymmetric bilinear forms are boring. They are just

$$\begin{pmatrix} & -1 & & & \\ 1 & & & & \\ & & & -1 & \\ & & 1 & & \\ & & & & \ddots \end{pmatrix}.$$

Classification of unimodular symmetric bilinear forms is a rich number-theoretic subject.

Theorem 15 (Freedman's classification of topological four-manifolds)**.** *Every equivalence class of integral unimodular (*$\det = \pm 1$*) symmetric bilinear form corresponds to either 1 or 2 homeomorphism classes of* simply connected *compact topological 4-manifolds. (It corresponds to 1 such homeomorphism class iff the diagonal entries of the matrix are all even.) For any given form, at most one homeomorphism class can admit smooth structures.*

The simplest matrices are $Q_{\mathbb{CP}^2} = (+1)$, $Q_{\overline{\mathbb{CP}^2}} = (-1)$, and $Q_{S^2\times S^2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Much more interesting is the E_8 Cartan matrix

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}.$$

This is symmetric, unimodular, and positive-definite, yet it is not equivalent to a diagonal matrix! This is easy to see since it's even.

2.1 Functorial view of classification

In dimension 2, we saw that cohomological classification coincides with topological classification (which also corresponds to smooth classification).

As we move into dimension 3, although smooth and topological classifications still coincide, we will encounter other levels of classification which do not. Namely,

$$\text{smooth oriented} \subset \text{smooth} \subset \text{topological} \subset \text{homotopy} \subset \text{cohomology}.$$

For example, the Poincare homology 3-sphere $\mathrm{SO}(3)/I$, where I is the icosahedral group of order 60 has the same integral cohomology as S^3, but is not homotopy equivalent to it.

To be specific, there are functors between categories, where all manifolds are assumed to be closed and connected

$$\left\{\begin{matrix}\text{oriented smooth manifolds}\\ \text{orientation-preserving diffeomorphisms}\end{matrix}\right\} \Longrightarrow \left\{\begin{matrix}\text{smooth manifolds,}\\ \text{smooth maps}\end{matrix}\right\}$$
$$\left\{\begin{matrix}\text{smooth manifolds,}\\ \text{smooth maps}\end{matrix}\right\} \Longrightarrow \left\{\begin{matrix}\text{topological manifolds,}\\ \text{continuous maps}\end{matrix}\right\}$$
$$\left\{\begin{matrix}\text{topological manifolds,}\\ \text{continuous maps}\end{matrix}\right\} \Longrightarrow \left\{\begin{matrix}\text{topological manifolds,}\\ \text{homotopy classes of continuous maps}\end{matrix}\right\}$$
$$\left\{\begin{matrix}\text{topological manifolds,}\\ \text{homotopy classes of continuous maps}\end{matrix}\right\} \Longrightarrow \left\{\begin{matrix}\text{graded-commutative rings,}\\ \text{degree-preserving homomorphisms}\end{matrix}\right\}$$

Functors are morphisms of categories. They send objects of one category to objects of another, and similarly for morphisms. They preserve identity morphisms and composition. The main consequence is that functors preserve isomorphisms. Suppose X_1 and X_2 are isomorphic in some category $\mathscr{C}$, i.e. there are $f : X_1 \to X_2$ and $g : X_2 \to X_1$ such that $g \circ f = \mathrm{Id}_{X_1}$ and $f \circ g = \mathrm{Id}_{X_2}$. Suppose $F : \mathscr{C} \to \mathscr{D}$ is a functor, so that $Y_1 = F(X_1)$ and $Y_2 = F(X_2)$. Then there are morphisms $F(f) : Y_1 \to Y_2$ and $F(g) : Y_2 \to Y_1$ so that $F(g) \circ F(f) = F(g \circ f) = F(\mathrm{Id}_{X_1}) = \mathrm{Id}_{Y_1}$, and similarly $F(f) \circ F(g) = \mathrm{Id}_{Y_2}$. Thus isomorphic objects will remain isomorphic under a functor. But non-isomorphic objects might become isomorphic in the image of a functor.

Functors take isomorphisms to isomorphisms, and isomorphism classes to isomorphism classes. The induced maps on isomorphism classes need not be injective or surjective. For example, upon forgetting orientation, $\left[\mathbb{CP}^2\right]$ and $\left[\overline{\mathbb{CP}^2}\right]$ both map to the same $\left[\mathbb{CP}^2\right]$. In two dimensions, $\left[\Sigma_g\right]$ maps to $\left[\Sigma_g\right]$, but nothing maps to $\left[T^{\#k}\right]$.

2.2 Notions of chirality

Chirality is the study of lack of orientation reversal. An excellent overview of the subject is presented in the thesis [Mül08]. Many details and generalizations can be found therein.

We can extend our diagram from last time:

$$\begin{array}{ccccccc}
\text{orientation-preserving diffeomorphic} & \Rightarrow & \text{orientation-preserving homeomorphic} & \Rightarrow & \text{positively homotopy equivalent} & & \\
\Downarrow & & \Downarrow & & \Downarrow & & \\
\text{diffeomorphic} & \Rightarrow & \text{homeomorphic} & \Rightarrow & \text{homotopy equivalent} & \Rightarrow & \text{isomorphic cohomology}
\end{array}$$

and all these equivalences are induced by functors.

It can happen that a manifold is smoothly chiral, but topologically achiral (admitting an orientation-reversing homeomorphism but not diffeomorphism). Several exotic spheres provide examples. Oriented exotic n-spheres form an abelian monoid (group without inverses) under connected sum. When $n > 4$, there is an inverse is given by orientation reversal, making it an abelian group Θ_n.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Θ_n	0	0	0	?	0	0	$\mathbb{Z}_{28}$	$\mathbb{Z}_2$	$\mathbb{Z}_2^3$	$\mathbb{Z}_6$	$\mathbb{Z}_{992}$	0	$\mathbb{Z}_3$	$\mathbb{Z}_2$	$\mathbb{Z}_{8128} \oplus \mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2^4$

It is unknown whether exotic S^4 exist.

Fix some $n > 4$ so that Θ_n is a group. For each $a \in \Theta_n$ fix some manifold X_a representing a so that up to oriented diffeomorphism, $X_a \overset{\text{diff}}{\simeq} X_b \iff a = b \in \Theta_n$, and $\overline{X_a} \overset{\text{diff}}{\simeq} X_{-a}$. I claim that X_a is smoothly achiral iff $2a = 0$.

$$X_a \overset{\text{diff}}{\simeq} \overline{X_a} \iff X_a \overset{\text{diff}}{\simeq} X_{-a} \iff a = -a \iff 2a = 0.$$

For example, for the exotic 7-spheres group $\mathbb{Z}_{28}$, only X_0 and X_{14} are achiral, and the remaining 26 are chiral.

On the other hand, up to homeomorphism, each X_a is homeomorphic to S^n, and $S^n \overset{\text{homeo}}{\cong} \overline{S^n}$, so all exotic spheres are homeomorphically achiral.

2.3 Fundamental group

Consider a manifold X with a specified point $x_0 \in X$. Define $\pi_1(X, x_0)$ to be the set of homotopy classes of parameterized loops starting and ending at x_0, where any homotopy is also required to fix the endpoints at x_0. It's routine to check that $\pi_1(X, x_0)$ is a group, where composition corresponds to concatenation of loops, and reversal of a path gives its inverse. When X is connected, $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$ are isomorphic with an isomorphism induced from choice of a path connecting x_0 to x_1. However, the isomorphism depends on the homotopy class of such a path. The group $\pi_1(X, x_0)$ is called the *fundamental group* of X, and it is a homotopy invariant of X. (A homotopy equivalence of manifolds induces an isomorphism of fundamental groups.)

If X is connected and $\pi_1(X, x_0) = 0$, then X is said to be *simply connected*. For example, S^n is simply connected for $n \geq 2$.

In principle, this would be a good time to discuss covering spaces, but in the interest of time, we will skip them for now. Most manifolds we consider will be simply connected anyway. Instead, I leave you with a theorem which should suffice for most of our purposes.

Theorem 16. *If G is a group acting freely and properly discontinuously on a simply connected manifold X, then X/G is a manifold with $\pi_1(X/G, [x_0]) = G$.*

2.4 Poincaré homology sphere

The symmetries of the icosahedron form a subgroup $I \subset SO(3)$ of order 60 called the icosahedral group. The Poincaré homology sphere is the quotient space $P = SO(3)/I$. Topologically, the group $SO(3)$ is $\mathbb{RP}^3 = S^3/\text{antipodal map}$. (We will explain this in a moment.) There is a group $\tilde{I}$ of order 120 called the binary icosahedral group such that $P = SO(3)/I = SU(2)/\tilde{I}$. Since $SU(2) \simeq S^3$ is simply connected, it follows from the above theorem that $\pi_1(P) = \tilde{I}$. The group $\tilde{I}$ is *perfect*, meaning that it is generated by its commutators. Consequently, the abelianization of $\tilde{I}$ (the group obtained by imposing commutativity) is trivial. Once we understand more about homology and cohomology, we will see how this implies that the cohomology ring of P is isomorphic to that of S^3. That's why P is called a (co)homology sphere.

2.5 Lens spaces

Let p and q be coprime integers. The lens space $L(p; q)$ is the quotient space of the unit sphere $S^3 \subset \mathbb{C}^2$ under the $\mathbb{Z}_p$ action generated by

$$(z_1, z_2) \sim (e^{2\pi i/p} z_1, e^{2\pi i q/p} z_2).$$

Thus $\pi_1(L(p; q), x_0) = \mathbb{Z}_p$. There is a natural orientation induced from S^3, and $\overline{L(p; q)} \cong L(p; -q)$ by conjugating the second coordinate.

Lens spaces provide many useful examples for understanding different levels of classification. For example, two lens spaces $L(p; q_1)$ and $L(p; q_2)$ are orientation-preserving homeomorphic (equivalently diffeomorphic) iff $q_1 q_2^{\pm 1} \equiv 1 \pmod{p}$. More generally, they are homotopy equivalent iff $q_1 q_2^{\pm 1} \equiv n^2 \pmod{p}$ for some integer n.

For example, $L(5; 4) \cong \overline{L(5; 1)}$, so $L(5; 4)$ and $L(5; 1)$ are orientation-reversing homeomorphic. However, they are not orientation-preserving homeomorphic since $4 \cdot 1^{\pm 1} \equiv 4 \not\equiv 1 \pmod{5}$. Thus they are "topologically chiral." In contrast, they are orientation-preserving homotopy-equivalent since $4 = 2^2$, so they are "homotopically achiral."

As another example, since $3 \cdot 1^{\pm 1} \equiv 3 \not\equiv 1 \pmod{7}$, it follows that $L(7; 3)$ and $L(7; 1)$ are not orientation-preserving homeomorphic. Nor are they orientation-reversing homeomorphic since $-3 \cdot 1^{\pm 1} \equiv -3 \not\equiv 1 \pmod{7}$. They are orientation-reversing homotopy equivalent since $L(7; 3) \cong \overline{L(7; 4)}$, and $L(7; 1)$ is orientation-preserving homotopy-equivalent to $L(7; 4)$ since $4 \cdot 1^{\pm 1} \equiv 2^2 \pmod{7}$. However, $L(7; 3)$

and $L(7;1)$ are not orientation-preserving homotopy-equivalent since 3 is not a square modulo 7.

$$\begin{tikzcd}
\text{o.p. homeomorphic} \arrow[r, "L(5;1)", "L(5;4)"'] \arrow[d, "L(5;1)"', "L(5;4)"] & \text{o.p. homotopy-equivalent} \arrow[d, "L(7;1)"', "L(7;3)"] \\
\text{homeomorphic} \arrow[r, "L(7;1)", "L(7;3)"'] & \text{homotopy-equivalent}
\end{tikzcd}$$

2.6 Notes on classification

Classification of 3-manifolds is based on Thurston's geometrization conjecture. Roughly, this states that every three-manifold can be decomposed in terms of certain "geometric" pieces. Perelman showed that these pieces can be obtained via the Ricci flow

$$\partial_t g_{ij} = -2R_{ij},$$

where g_{ij} is a Riemannian metric, and R_{ij} is the Ricci curvature tensor. This is an evolution equation which behaves like a heat equation, tending to uniformize the curvature. Singularities develop, for instance, as the various geometric pieces pinch off, and one of the major technical obstacles is understanding how to deal with these singularities so that the flow can continue. Perelman's results essentially reduce the classification problem to understanding the geometric pieces and their possible quotients. Consequently, the theory of 3-manifolds involves much group theory related to the possible fundamental groups which arise.

In higher dimensions, the group theory becomes literally impossible. Any finitely presented group can appear as $\pi_1(X)$ for compact X when $\dim X \geq 4$. The classification of finitely presented groups is undecidable. Philisophically, the idea is that given any fixed axiom system, it's possible to manufacture a group presentation which effectively encodes a statement such as, "triviality of this group is equivalent to a proof with your axioms that this group is nontrivial." Assuming consistency of your axioms, such a group must be nontrivial. However, your axioms cannot provide a proof. Thus it does not lead to inconsistency to make an axiom which declares that the constructed group is either trivial or nontrivial.

Since the general classification problem is doomed from the start, typically one focuses on classifying simply connected manifolds in these dimensions. When $n \geq 5$, classification of simply-connected smooth manifolds is generally considered well-understood due to surgery theory, which essentially reduces classification to an algebraic problem thanks to the h-cobordism theorem. However, in dimension 4, things go wrong due to failure of the "Whitney trick." Given two submanifolds P and Q of complementary dimension, they can be perturbed to intersect transversely to meet in finitely many points. If everything is oriented, then these intersection points have signs. One wants to be able to cancel intersection points which have opposite signs. The strategy is to form a loop by taking a path inside each of P and Q between the intersection points. We wish to to fill this in with a smoothly embedded "Whitney disk," which then allows us to slide apart the surfaces. In dimension 4, such disks will generally have self-intersections. Roughly speaking, Freedman's classification of simply connected topological 4-manifolds uses an infinite sequence of modifications, called *Casson handles*, to eliminate self-intersections, but not smoothly.

2.7 Classification of unimodular symmetric bilinear forms

We can summarize the results with the following table:

	odd	even
indefinite	$m\,(+1) \oplus n(-1)$	$\pm m\,E_8 \oplus n\,H$
definite	too difficult, but we only encounter $m\,(\pm 1)$	

A unimodular bilinear form Q is *positive/negative definite* if $Q(x,x)$ is always positive/negative for nonzero x. If Q is neither positive definite nor negative definite, then Q is called *indefinite*. If $Q(x,x)$ is always even, then Q is called *even*. Otherwise, Q is called *odd*. For example,

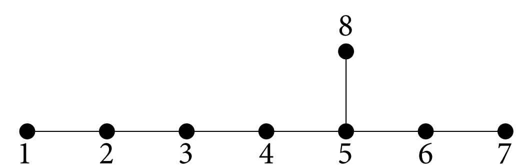
$$H := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is even and indefinite, since

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1x_2.$$

Note that Q is even iff its diagonal elements are always even. (This is because off-diagonal entries automatically acquire a factor of two.)

An even positive-definite form arises via the Cartan matrix for the Lie algebra E_8:



$$E_8 := 2I - A_{E_8} = \begin{pmatrix} 2 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & & -1 \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & \\ & & & & -1 & & & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & & & & & & \\ 1 & 2 & 1 & & & & & \\ & 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & & \\ & & & 1 & 2 & 1 & & 1 \\ & & & & 1 & 2 & 1 & \\ & & & & & 1 & 2 & \\ & & & & 1 & & & 2 \end{pmatrix} = 2I + A_{E_8}.$$

Classification of unimodular definite forms is not understood, and the numbers grow rapidly with rank. Thankfully we are saved from this hopeless situation by

Theorem 17 (Donaldson). *If X is a simply-connected 4-manifold with Q_X definite, then*

$$Q_X \sim \pm \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = m\,(\pm 1)\,.$$

Indefinite forms are much easier to classify. If Q is indefinite and odd, then Q is diagonal:

$$Q \sim m\,(+1) \oplus n(-1),$$

with $m, n > 0$. If Q is indefinite and even, then

$$Q \sim \pm m\,E_8 \oplus n\,H,$$

where $n > 0$, $m \geq 0$, and $-E_8$ is to be understood as E_8 with the opposite sign.

Remark. Indefinite forms are completely classified by rank, signature, and type (even/odd). For example, $E_8 \oplus -E_8 \sim 8H$ since it is even of signature zero. Also, $E_8 \oplus (-1) = 8\,(+1) \oplus (-1)$ which is odd of signature 7.

Now that we understand the possible cohomology of closed oriented 4-manifolds, we can try and ascend our classification hierarchy to understand smooth 4-manifolds. Recall that we restrict to *simply-connected* closed 4-manifolds because the general classification problem would encompass the impossible classification of all finitely presented groups, which arise as fundamental groups. Now we attempt to use cohomology to ascend the classification hierarchy. The first step proceeds without difficulty. Hatcher gives a complete proof of:

Proposition 18 (Algebraic Topology, 4C.3). *For a simply-connected closed topological* 4*-manifold, cohomology determines homotopy type.*

Thanks to the incredible work of Freedman, we can ascend to homeomorphism classification:

Theorem 19 (Freedman). *For any unimodular symmetric bilinear form Q, there is a closed simply-connected topological* 4*-manifold with Q as its intersection form. Furthermore,*

- if Q is even, the manifold is unique up to homeomorphism,
- if Q is odd, there are two homeomorphism classes, at least one of which is not smoothable.

Note that for any intersection form Q, there is at most one homeomorphism class containing a smooth manifold. Consequently, two simply-connected smooth 4-manifolds X_1 and X_2 are homeomorphic iff $Q_{X_1} \sim Q_{X_2}$!

At this point, the natural question is: given an intersection form, how many smooth 4-manifolds does it correspond to?

For all definite forms except for the diagonal, the answer is zero by Donaldson's theorem.

Next we list what we have: $Q_{\mathbb{CP}^2} = (+1)$, $Q_{\overline{\mathbb{CP}^2}} = (-1)$. $Q_{S^2 \times S^2} = H$. These realize all intersection forms without a $\pm E_8$ factor.

Note that if Q_X is odd, then X is homeomorphic to $m\mathbb{CP}^2 \# n\overline{\mathbb{CP}^2}$.

To see examples of manifolds with an E_8 factor, we go to complex geometry. The family Calabi-Yau manifolds of complex dimension 2 is called the K3 surfaces. Since they are all diffeomorphic, differential topologists call them *the* K3 surface. One description is the quartic hypersurface in $\mathbb{CP}^3$ defined by

$$z_0^4 + z_1^4 + z_2^4 + z_3^4 = 0, \quad (z_0 : z_1 : z_2 : z_3) \in \mathbb{CP}^3.$$

It turns out that

$$Q_{\mathrm{K3}} = -2E_8 \oplus 3H.$$

It's not possible to find a smooth 4-manifold with a single copy of E_8. For now, we state without proof that for X a simply connected closed 4-manifold,

$$Q_X \text{ even} \iff \text{tangent bundle of } X \text{admits a spin structure} =: X \text{ is spin}$$

There are several interesting theorems on spin manifolds.

Theorem 20 (Rokhlin). *If X is a smooth spin 4-manifold, then the signature satisfies $\sigma(X) \equiv 0 \pmod{16}$.*

Since a closed simply-connected spin 4-manifold has intersection form

$$Q \sim \pm m\, E_8 \oplus n\, H,$$

we compute

$$\sigma(Q) = \pm 8m.$$

Thus Rokhlin's theorem implies m is even.

Next we ask whether it is possible to reduce the number of H in K3. Furuta used the Seiberg-Witten equations to prove

Theorem 21 (Furuta (2001)). *If X is a closed oriented spin 4-manifold with $b_2(X) \neq 0$, then*

$$b_2(X) \geq \frac{10}{8} \left|\sigma(X)\right| + 2.$$

Substituting $b_2(X) = 8m + 2n$ and $|\sigma(X)| = 8m$, the above inequality is equivalent to $n \geq m + 1$. Thus for K3, $n \geq 3$, so we have the minimal number of H.

Closely related

Conjecture 22 ($\frac{11}{8}$). *If X is a closed oriented spin 4-manifold, then*

$$b_2(X) \geq \frac{11}{8} \left|\sigma(X)\right|.$$

This is equivalent to $n \geq \frac{3}{2}m$. By Freedman's classification, this is equivalent to the conjecture that any simply-connected closed oriented spin 4-manifold be homeomorphic to

$$\tfrac{m}{2}\, \mathrm{K3} \# \left(n - \tfrac{3}{2}m\right) (S^2 \times S^2),$$

where of course the number of copies of each type is a nonnegative integer.

Assuming the $\frac{11}{8}$ conjecture, all smooth closed simply connected 4-manifolds are homeomorphic to connected sums of $\mathbb{CP}^2$, $\overline{\mathbb{CP}^2}$, K3, $\overline{\mathrm{K3}}$, and $S^2 \times S^2 = \overline{S^2 \times S^2}$. Furthermore, based on the classification theorem, we can read off all the relations

$$\begin{aligned}
\mathrm{K3}\#\overline{\mathrm{K3}} &= 22\, S^2 \times S^2, \\
\mathrm{K3}\#\mathbb{CP}^2 &= 4\mathbb{CP}^2\#19\overline{\mathbb{CP}^2}, \\
\mathrm{K3}\#\overline{\mathbb{CP}^2} &= 3\mathbb{CP}^2\#20\overline{\mathbb{CP}^2}, \\
\mathbb{CP}^2\#S^2 \times S^2 &= 2\mathbb{CP}^2\#\overline{\mathbb{CP}^2},
\end{aligned}$$

plus the corresponding identities obtained from the above by reversing the orientations. Thus the smooth classification problem is focused on classifying exotic structures on these connected sums.

Chapter 3

Bundle theory

3.1 Čech cohomology

It is extremely useful to be able to switch perspectives on cohomology. De Rham cohomology relates to calculus of differential forms. Singular cohomology relates to submanifolds. Čech cohomology will relate to fiber bundles. The equivalence of these theories provides deep connections between these subjects.

Suppose we have an open cover $\{U_\alpha\}$ of our manifold X. We define the chain complex $\check{C}^p(\{U_\alpha\}\,;A)$ as follows. Denote multiple intersections by

$$U_{\alpha\beta} := U_\alpha \cap U_\beta, \quad U_{\alpha\beta\gamma} := U_\alpha \cap U_\beta \cap U_\gamma, \quad \text{etc.}$$

A Čech p-cochain ϕ associates to each $p+1$-fold intersection $U_{\alpha_0\alpha_1\cdots\alpha_p}$ a locally constant function $\phi_{\alpha_0\alpha_1\cdots\alpha_p} : U_{\alpha_0\alpha_1\cdots\alpha_p} \to A$.

$$\check{C}^p(\{U_\alpha\}\,;A) := \left\{\phi = \left\{\phi_{\alpha_0\alpha_1\cdots\alpha_p} : U_{\alpha_0\alpha_1\cdots\alpha_p} \to A \text{ locally constant}\right\}\right\}.$$

The coboundary map is

$$\begin{aligned} d : \check{C}^p(\{U_\alpha\}\,;A) &\to \check{C}^{p+1}(\{U_\alpha\}\,;A) \\ (d\phi)_{\alpha_0\alpha_1\cdots\alpha_{p+1}} &:= \sum_{k=0}^{p+1}(-1)^k\,\phi_{\alpha_0\cdots\widehat{\alpha_k}\cdots\alpha_{p+1}}, \end{aligned}$$

where $\widehat{\alpha_k}$ denotes omission of α_k. It's easy to verify that $d^2 = 0$. If $\phi \in \check{C}^0(\{U_\alpha\}\,;A)$, then ϕ defines a collection of locally constant functions $\phi_\alpha : U_\alpha \to A$. If $d\phi = 0$, then for each $U_\alpha \cap U_\beta$, we have $0 = \phi_\alpha - \phi_\beta$, so the ϕ_α agree on the overlaps and determine a locally constant function $\phi : X \to A$. As usual, we define

$$\check{H}^p(\{U_\alpha\}\,;A) := \frac{\ker d}{\operatorname{image} d}.$$

But we want cohomology to depend on X rather than a given cover. A different open cover $\left\{V_\beta\right\}_{\beta\in J}$ is called a *refinement* of $\{U_\alpha\}_{\alpha\in I}$ if each V_β is contained in some U_α. Fixing a choice $V_\beta \subset U_{\tau(\beta)}$ of some function $\tau : J \to I$ induces a restriction map

$$\check{C}^p(\{U_\alpha\};A) \to \check{C}^p(\left\{V_\beta\right\};A).$$

The induced map on cohomology $\check{H}^p(\{U_\alpha\};A) \to \check{H}^p(\left\{V_\beta\right\};A)$ does not depend on the choice of τ. Note that any two open covers $\{U_\alpha\}_{\alpha\in I}$ and $\left\{V_\beta\right\}_{\beta\in J}$ have a common refinement $\left\{U_\alpha \cap V_\beta\right\}_{(\alpha,\beta)\in I\times J}$. We define

$$\check{H}^p(X;A) := \underset{\{U_\alpha\}\text{ open cover}}{\text{dir-lim}} \check{H}^p(\{U_\alpha\};A).$$

This means that any element of $\check{H}^p(X;A)$ is represented as a Čech cocycle with respect to some specific cover $\{U_\alpha\}$, and two elements in $\check{H}^p(X;A)$ are equal iff they become equal under a common refinement. Thankfully, we don't have to worry about this direct limit in practice.

A cover $\{U_\alpha\}$ is called a *good cover* if each U_α is contractible, as well as each finite intersection $U_{\alpha_0\cdots\alpha_p}$. If $\{U_\alpha\}$ is a good cover, then $\check{H}^p(X;A) = \check{H}^p(\{U_\alpha\};A)$.

Good covers always exist on a smooth manifold. We can pick a Riemannian metric, and then use metric balls which are sufficiently small to be *geodesically convex*, meaning that any two points are joined by a unique geodesic. Any geodesically convex subset is contractible, and intersections of geodesically convex subsets are geodesically convex. (The naive strategy would be to use convex coordinate charts, however convexity is not preserved under coordinate change. But geodesic convexity with respect to a fixed Riemannian metric is.)

Like our previous cohomology theories, $\check{H}^p(X;A)$ is canonically isomorphic to sheaf cohomology, and thus can be identified with de Rham and singular cohomology.

3.2 Bundle theory

A *smooth Euclidean vector bundle of rank k over a manifold X* is a projection map $\pi : E \to X$ such that each fiber $\pi^{-1}(X)$ is a Euclidean vector space, and they fit together via smooth local trivializations

$$\begin{tikzcd} E\,|_{U_\alpha} \arrow[dr, "\pi"'] & & U_\alpha \times \mathbb{R}^k \arrow[ll, "\cong", "\phi_\alpha"'] \arrow[dl, "\pi_1"] \\ & U_\alpha & \end{tikzcd},$$

where the $\{U_\alpha\}$ cover X, and the ϕ_α parameterize fiberwise linear isometries $\phi_\alpha(x) \in \mathrm{Iso}(\mathbb{R}^k, E|_x)$. Note that $\phi = \{\phi_\alpha\}$ defines a map

$$\begin{tikzcd} E \arrow[dr, "\pi"'] & & \coprod_\alpha U_\alpha \times \mathbb{R}^k \arrow[ll, "\phi"'] \arrow[dl, "\pi_1"] \\ & X & \end{tikzcd}$$

which covers E, and two points $(x,v)_\alpha \in U_\alpha \times \mathbb{R}^k$ and $(x,w)_\beta \in U_\beta \times \mathbb{R}^k$ map to the same point in E iff

$$\phi_\alpha(x)v = \phi_\beta(x)w \iff v = \phi_\alpha^{-1}(x)\phi_\beta(x)w = \phi_{\alpha\beta}(x)w,$$

where

$$\phi_{\alpha\beta}(x) := \phi_\alpha^{-1}(x)\phi_\beta(x) \in \mathrm{Isom}(\mathbb{R}^k \leftarrow \mathbb{R}^k) =: \mathrm{O}(k).$$

As $\phi_{\alpha\beta}$ ranges over all possible x, it defines a function $\phi_{\alpha\beta} \in C^\infty(U_{\alpha\beta}; \mathrm{O}(k))$ called a *transition function*. After identifying corresponding points $\left[x, \phi_{\alpha\beta}(x)v\right]_\alpha \sim [x, v]_\beta$, the map ϕ induces an isometry of Euclidean vector bundles

$$\begin{tikzcd}
E \arrow[dr, "\pi"'] & & \coprod_\alpha U_\alpha \times \mathbb{R}^k / \sim . \arrow[ll, "\cong", "\phi"'] \arrow[dl, "\pi_1"] \\
& X &
\end{tikzcd}$$

Thus every smooth Euclidean vector bundle is isomorphic (=isometric) to a vector bundle determined by transition functions $\phi_{\alpha\beta} \in C^\infty(U_{\alpha\beta}; \mathrm{O}(k))$. According to the definition $\phi_{\alpha\beta}(x) := \phi_\alpha^{-1}(x)\phi_\beta(x)$, the transition functions are readily verified to satisfy the cocycle condition on the triple-overlaps $U_{\gamma\alpha\beta}$:

$$\phi_{\alpha\beta}\phi_{\gamma\beta}^{-1}\phi_{\gamma\alpha} = \mathrm{Id}. \tag{3.1}$$

Simple consequences include $\phi_{\alpha\alpha} = \mathrm{Id}$, and $\phi_{\beta\alpha} = \phi_{\alpha\beta}^{-1}$.

Conversely, any collection of transition functions $\phi_{\alpha\beta} \in C^\infty(U_{\alpha\beta}; \mathrm{O}(k))$ for any open cover $\{U_\alpha\}$ satisfying (3.1) defines a vector bundle $\coprod_\alpha U_\alpha \times \mathbb{R}^k / \sim$.

3.3 Frame bundles

Often various operations with vector spaces involve expressions which make use of a basis. Usually it is possible to find a more insightful alternative construction which does not involve a basis. In particular situations, it may be unwieldly or impossible to avoid such constructions. But making a choice of basis is ugly, and there is a clever way around it.

Rather than choose a particular basis, the more satisfying and mathematically invariant approach is to parameterize constructions over all such choices of bases. This leads to the notion of a *torsor* and its parameterized version, the *principal bundle*. To motivate the definitions, we examine the case of a vector bundle.

For simplicity, we first study the set of bases of a single vector space. Let V be a Euclidean vector space of rank k. Observe that an orthonormal basis for V is equivalent to a linear isometry $V \leftarrow \mathbb{R}^k$. Specifically, the image of the standard basis $\{e_i\}_{i=1}^k$ of $\mathbb{R}^k$ determines an orthonormal basis of V. Thus we define the *orthonormal frames* on V to be the isometries

$$\mathrm{Fr}^{\mathrm{O}}(V) := \mathrm{Isom}(V \leftarrow \mathbb{R}^k).$$

This perspective makes clear that there is a natural right action of $\mathrm{O}(k) := \mathrm{Isom}(\mathbb{R}^k \leftarrow \mathbb{R}^k)$ on $\mathrm{Fr}^{\mathrm{O}}(V)$ by composition. Specifically, if $\phi \in \mathrm{Fr}^{\mathrm{O}}(V)$ and $g \in \mathrm{O}(k)$, then

$$\phi g \in \mathrm{Isom}(V \leftarrow \mathbb{R}^k \leftarrow \mathbb{R}^k) = \mathrm{Fr}^{\mathrm{O}}(V).$$

For any $\phi_0, \phi_1 \in \mathrm{Fr}^{\mathrm{O}}(V)$, there is a unique $g \in \mathrm{O}(k)$ such that $\phi_1 = \phi_0 g$, namely

$$\phi_1 = \phi_0 \underbrace{(\phi_0^{-1}\phi_1)}_{\in \mathrm{O}(k)}.$$

Thus if we fix a basepoint ϕ_0, we see that $\mathrm{Fr}^{\mathrm{O}}(V)$ is in bijection with $\mathrm{O}(k)$ via $\phi \mapsto \phi_0^{-1}\phi$. But apart from such an identification, it makes no sense to compose elements of $\mathrm{Fr}^{\mathrm{O}}(V)$, so it is not a group. Instead, we say that $\mathrm{Fr}^{\mathrm{O}}(V)$ is a *right* $\mathrm{O}(k)$ *torsor*, where a torsor is a space with a free and transitive right group action. More concretely, a torsor is like a copy of a group that has "lost its identity," but still knows how to act on itself by left or right multiplication.

Simultaneously, $\mathrm{Fr}^{\mathrm{O}}(V)$ is a left $\mathrm{Aut}(V) = \mathrm{O}(V) := \mathrm{Isom}(V \leftarrow V)$ torsor via composition on the other side:

$$\mathrm{O}(V) \circlearrowright \mathrm{Fr}^{\mathrm{O}}(V) \circlearrowleft \mathrm{O}(k),$$

and these actions clearly commute.

Although $\mathbb{R}^k$ and V are not naturally isomorphic, the two corresponding trivial bundles over $\mathrm{Fr}^{\mathrm{O}}(V)$ are. There is a natural isomorphism

$$\mathrm{Fr}^{\mathrm{O}}(V) \times \mathbb{R}^k \longrightarrow \mathrm{Fr}^{\mathrm{O}}(V) \times V$$

given by

$$(\phi, x) \mapsto (\phi, \phi(x)).$$

This natural isomorphism achieves the desired goal of identifying $\mathbb{R}^k$ and V (over the space of frames) without a choice of basis.

3.4 Associated bundles

The isomorphism of the previous section allows a powerful construction for transfering structures on $\mathbb{R}^k$ to structures on the abstract vector space V as follows. There is a diagonal left action of $\mathrm{O}(k)$ on $\mathrm{Fr}^{\mathrm{O}}(V) \times \mathbb{R}^k$ given by

$$g(\phi, x) = (\phi g^{-1}, gx).$$

The map

$$\begin{aligned} \mathrm{Fr}^{\mathrm{O}}(V) \times \mathbb{R}^k &\longrightarrow V \\ (\phi, x) &\mapsto \phi(x) \end{aligned}$$

is invariant under the action of $\mathrm{O}(k)$:

$$g(\phi, x) = (\phi g^{-1}, gx) \mapsto \phi(g^{-1}gx) = \phi(x).$$

It is straightforward to verify that we get an isomorphism

$$\frac{\mathrm{Fr}^{\mathrm{O}}(V) \times \mathbb{R}^k}{\mathrm{O}(k)} \xrightarrow{\cong} V.$$

Indeed, fixing any $\phi_0 \in \mathrm{Fr}^{\mathrm{O}}(V)$, every point $[\phi, x]$ in the quotient is equivalent to a (unique) point of the form $[\phi_0, x']$ for $x' \in \mathbb{R}^n$. To verify, note that equivalence on the quotient can be rephrased as $[\phi g, x] \sim [\phi, gx]$. Then

$$[\phi, x] = \left[\phi_0\left(\phi_0^{-1}\phi\right), x\right] = \left[\phi_0, \left(\phi_0^{-1}\phi\right)x\right].$$

These representatives $[\phi_0, x]$ are mapped isomorphically to V as $[\phi_0, x] \mapsto \phi_0(x)$.

There is an important generalization of this construction called an *associated space*. Suppose F is a space with an action of $\mathrm{O}(k)$, i.e. we have a homomorphism $\rho : \mathrm{O}(k) \to \mathrm{Aut}(F)$. We combine the right action on $\mathrm{Fr}^{\mathrm{O}}(V)$ with the left action of F to define

$$\mathrm{Fr}^{\mathrm{O}}(V) \times_\rho F := \frac{\mathrm{Fr}^{\mathrm{O}}(V) \times F}{\mathrm{O}(k)},$$

where equivalence is given by $[\phi g, f] \sim [\phi, \rho(g)f]$. This allows us to associate structures from the model space $\mathbb{R}^k$ to an abstract copy V, so long as the structure is $\mathrm{O}(k)$-invariant. For example, when ρ_{st} is the standard representation on $\mathbb{R}^k$, from the previous computation we get

$$\mathrm{Fr}^{\mathrm{O}}(V) \times_{\rho_{\mathrm{st}}} \mathbb{R}^k = V.$$

If ρ_{Λ^p} is the representation $\mathrm{O}(k) \to \mathrm{O}(\Lambda^p\mathbb{R}^k)$ on the p-th exterior power, then

$$\mathrm{Fr}^{\mathrm{O}}(V) \times_{\rho_{\Lambda^p}} \Lambda^p\mathbb{R}^k = \Lambda^p V.$$

If $\mathrm{Ad} : \mathrm{O}(k) \to \mathrm{Aut}(\mathrm{O}(k))$ is the adjoint action $g \mapsto (h \mapsto ghg^{-1})$, then

$$\mathrm{Fr}^{\mathrm{O}}(V) \times_{\mathrm{Ad}} \mathrm{O}(k) \cong \mathrm{O}(V)$$

via

$$[\phi, h] \mapsto \phi h \phi^{-1} \in \mathrm{Isom}(V \leftarrow \mathbb{R}^k \leftarrow \mathbb{R}^k \leftarrow V).$$

The adjoint action is appropriate since $[\phi g, h]$ and $\left[\phi, ghg^{-1}\right]$ correspond to the same element of $\mathrm{O}(V)$.

We can repeat these constructions fiberwise for a smooth Euclidean vector bundle of rank k, $\pi : E \to X$. The orthonormal frame bundle $\mathrm{Fr}^{\mathrm{O}}(E)$ is the fiber bundle over X whose fiber at a point x is $\mathrm{Fr}^{\mathrm{O}}(E|_x)$. It carries commuting actions

$$(\mathscr{G}_E := \Gamma(\mathrm{Aut}(E))) \circlearrowright \mathrm{Fr}^{\mathrm{O}}(E) \circlearrowleft C^\infty(X; \mathrm{O}(k)),$$

where $\Gamma(X; \mathrm{Aut}(E))$ denotes smooth sections of the bundle whose fiber over any point x are the isometries of $E|_x$. These are also known as *gauge transformations*.

Given $\rho : \mathrm{O}(k) \to \mathrm{Aut}(F)$, we can form the associated bundle

$$\mathrm{Fr}^{\mathrm{O}}(E) \times_\rho F := \frac{\mathrm{Fr}^{\mathrm{O}}(E) \times F}{\mathrm{O}(k)}$$

with the same equivalence relation fiberwise.

Given a local trivialization $\{\phi_\alpha\}$, the any associated bundle may be reconstructed via transition functions:

$$\begin{tikzcd} \mathrm{Fr}^{\mathrm{O}}(E) \times_\rho F \arrow[dr, "\pi"'] & & \coprod_\alpha U_\alpha \times F/\sim , \arrow[ll, "\cong"', "\{\phi_\alpha\}"] \arrow[dl, "\pi_1"] \\ & X & \end{tikzcd}$$

where $\phi_\alpha\left([x,f]_\alpha\right) = [\phi_\alpha(x), f]$. We compute that the necessary equivalence relation on $\coprod_\alpha U_\alpha \times F/\sim$ must be $[x,f]_\beta \sim \left[x, \rho(\phi_{\alpha\beta})f\right]_\alpha$ by equating equivalent points in the image:

$$\begin{aligned} & [x,f']_\alpha \sim [x,f]_\beta \\ \Longleftrightarrow\ & \phi_\alpha\left([x,f']_\alpha\right) \sim \phi_\beta\left([x,f]_\beta\right) \\ \Longleftrightarrow\ & [\phi_\alpha(x), f'] \sim \left[\phi_\beta(x), f\right] \\ \Longleftrightarrow\ & [\phi_\alpha(x), f'] \sim \left[\phi_\alpha(x)\phi_\alpha^{-1}(x)\phi_\beta(x), f\right] \\ \Longleftrightarrow\ & [\phi_\alpha(x), f'] \sim \left[\phi_\alpha(x), \rho\left(\phi_{\alpha\beta}\right) f\right] \\ \Longleftrightarrow\ & f' = \rho\left(\phi_{\alpha\beta}\right) f. \end{aligned}$$

Thus the associated bundle uses the same transition functions, but they are represented on a different fiber.

One important example is that the bundle $\mathrm{Aut}(E) := \mathrm{Fr}^{\mathrm{O}}(E) \times_{\mathrm{Ad}} \mathrm{O}(k)$, whose whose fiber over a point x is $\mathrm{O}(E|_x)$, and whose global sections are $\mathscr{G}_E := \Gamma(\mathrm{Aut}(E))$.

A *principal G-bundle* is a fiber bundle associated with the action $\rho_L : G \to \mathrm{Aut}(G)$ given by left multiplication $g \mapsto (h \mapsto gh)$. For example,

$$\mathrm{Fr}^{\mathrm{O}}(E) \times_{\rho_L} \mathrm{O}(k) \cong \mathrm{Fr}^{\mathrm{O}}(E)$$

by the map

$$[\phi(x), g] \mapsto \phi(x)g.$$

Thus $\mathrm{Fr}^{\mathrm{O}}(E)$ is a principal $\mathrm{O}(k)$ bundle.

More generally, given any fiber bundle $\pi : H \to X$ with fiber F, there is a principal $\mathrm{Aut}(F)$ bundle P such that the fiber $P|_x$ is $\mathrm{Iso}(H|_x \leftarrow F)$. There is clearly a right action on P by $\mathrm{Aut}(F)$ (which generalizes to an action of $C^\infty(X; \mathrm{Aut}(F))$). If ρ is the standard representation $\rho : \mathrm{Aut}(F) \to \mathrm{Aut}(F)$, then $P \times_\rho F = H$. In the case when $F = \mathbb{R}^k$ with the standard Euclidean structure, then H is a Euclidean vector bundle, and P is the orthonormal frame bundle.

The moral is that any fiber bundle H with fiber F is equivalent to a pair (P, ρ) where P is a principal G-bundle, and $\rho : G \to \mathrm{Aut}(F)$.

3.5 Čech cohomology revisited

Recall that Čech cohomology is described by

$$\check{C}^p(\{U_\alpha\}; A) := \left\{\phi = \left\{\phi_{\alpha_0\alpha_1\cdots\alpha_p} : U_{\alpha_0\alpha_1\cdots\alpha_p} \to A \text{ locally constant}\right\}\right\},$$

with differential

$$d : \check{C}^p(\{U_\alpha\}\,; A) \to \check{C}^{p+1}(\{U_\alpha\}\,; A)$$
$$(d\phi)_{\alpha_0\alpha_1\cdots\alpha_{p+1}} := \sum_{k=0}^{p+1}(-1)^k\,\phi_{\alpha_0\cdots\widehat{\alpha_k}\cdots\alpha_{p+1}}.$$

There is no reason to restrict to locally constant functions valued in an abelian group. In the context of principal bundles, we consider more general sheaves (i.e. classes of functions or sections) and try to make sense of Čech cohomology.

Recall that a local trivialization for a Euclidean vector bundle is a smooth map which for each $x \in U_\alpha$ gives an isometry $\phi_\alpha(x) : \mathbb{R}^k \to E|_x$. Each $\phi_\alpha(x)$ is an orthonormal frame, so it is equivalent to say that $\phi_\alpha \in \Gamma(U_\alpha; P)$ is a smooth section of the orthonormal frame bundle.

For a general principal bundle P, a local trivialization is equivalent to a local section. A system of local trivializations covering P is equivalent to a collection of local sections $\{\phi_\alpha \in \Gamma(U_\alpha; P)\}$. In the Čech framework,

$$\phi = \{\phi_\alpha\} \in \check{C}^0(X; P).$$

Transition functions are $\phi_{\alpha\beta} := \phi_\alpha^{-1}\phi_\beta \in C^\infty(U_{\alpha\beta}; G)$. This is the nonabelian version of an alternating sum with omitted indices, so we should interpret

$$d\,\{\phi_\alpha\} := \left\{\phi_{\alpha\beta} := \phi_\alpha^{-1}\phi_\beta\right\} \in \check{C}^1(X; C^\infty(G)).$$

The relation $d^2 = 0$ still holds if we interpret

$$d\left\{\phi_{\alpha\beta}\right\} := \left\{\phi_{\alpha\beta\gamma} := \phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta}\right\}.$$

The condition

$$\phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta} = 1$$

is the *cocycle condition* for transition functions, and incorporates all the constraints for general transition functions. Usually the constraints are written with three formulas, but this form encodes them all into one. Setting $\alpha = \beta = \gamma$, we get

$$1 = \phi_{\alpha\alpha}\phi_{\alpha\alpha}^{-1}\phi_{\alpha\alpha} = \phi_{\alpha\alpha}.$$

Setting only $\gamma = \alpha$, we get

$$\phi_{\beta\alpha}\phi_{\alpha\alpha}^{-1}\phi_{\alpha\beta} = 1 \implies \phi_{\beta\alpha} = \phi_{\alpha\beta}^{-1}.$$

Finally,

$$\phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta} = 1 \iff \phi_{\alpha\beta}\phi_{\beta\gamma}\phi_{\gamma\alpha} = 1.$$

If $\phi \in \check{C}^p(X; \mathcal{S})$, this means that on some open cover $\{U_\alpha\}$, there is a collection of functions $\left\{\phi_{\alpha_0\cdots\alpha_p} \in \mathcal{S}(U_{\alpha_0\cdots\alpha_p})\right\}$, where the sheaf $\mathcal{S}$ determines some class of functions on each overlap $U_{\alpha_0\cdots\alpha_p} := U_{\alpha_0} \cap \cdots \cap U_{\alpha_p}$. Whenever $\mathcal{S}$ is a sheaf of abelian groups, then

$$d\phi = \left\{\phi_{\alpha_0\cdots\alpha_{p+1}} := \sum_{k=0}^{p+1}(-1)^k\phi_{\alpha_0\cdots\widehat{\alpha_k}\cdots\alpha_{p+1}}\right\}.$$

In other cases, we must find a suitable alternative interpretation in the same spirit.

Given a system of local trivializations $\{\phi_\alpha\}$ of a fiber bundle, each ϕ_α is equivalent to a local section of the principal bundle P over U_α. Thus a local trivialization is equivalent to an element of $\check{C}^0(X;P)$.

The transition functions are $\left\{\phi_{\alpha\beta} := \phi_\alpha^{-1}\phi_\beta\right\}$, which should be thought of as a Čech differential

$$\left\{\phi_{\alpha\beta}\right\} = d\left\{\phi_\alpha\right\} \in \check{C}^1(X;G_{C^\infty}),$$

with smooth G-valued functions (as opposed to locally-constant G-valued functions).

Not all elements of $\check{C}^1(X;G_{C^\infty})$ arise in this way from some principal bundle. The relation $d^2 = 0$ still holds if we interpret

$$d\left\{\phi_{\alpha\beta}\right\} := \left\{\phi_{\alpha\beta\gamma} := \phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta}\right\}.$$

The condition

$$\phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta} = \mathrm{Id}$$

is the *cocycle condition* for transition functions. Usually the constraints are written with three formulas, but this form encompases them all. Setting $\alpha = \beta = \gamma$, we get

$$\mathrm{Id} = \phi_{\alpha\alpha}\phi_{\alpha\alpha}^{-1}\phi_{\alpha\alpha} = \phi_{\alpha\alpha}.$$

Setting only $\gamma = \alpha$, we get

$$\phi_{\beta\alpha}\phi_{\alpha\alpha}^{-1}\phi_{\alpha\beta} = \mathrm{Id} \implies \phi_{\beta\alpha} = \phi_{\alpha\beta}^{-1}.$$

Finally,

$$\phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta} = \mathrm{Id} \implies \phi_{\alpha\beta}\phi_{\beta\gamma}\phi_{\gamma\alpha} = \mathrm{Id}.$$

It's convenient to introduce the notation $\ker d := \check{Z}^p \subset \check{C}^p$ where $\check{Z}^p$ denotes the *cocycles* which are the cochains with trivial differential. In summary, we have a map

$$\begin{gathered} d : \check{C}^0(X;P) \to \check{Z}^1(X;C^\infty(G)). \\ \text{local trivializations} \mapsto \substack{\text{transition functions} \\ \text{(satisfying cocycle condition)}} \end{gathered}$$

We now proceed to systematically develop bundle theory by asking natural homological questions.

Any element $\left\{\phi_{\alpha\beta}\right\} \in \check{Z}^1(X;C^\infty(G))$ is represented as the image of d from some principal bundle P'. In particular, the cocycle condition is precisely the consistency condition required to carry out the gluing construction $P' := \coprod_\alpha U_\alpha \times G/\sim$ which realizes the transition functions $\left\{\phi_{\alpha\beta}\right\}$.

Next we should ask how the choice of local trivialization affects the transition functions. As a guiding principle, we will discover an analogue of the exact sequence in ordinary cohomology:

$$0 \to Z^{i-1} \hookrightarrow C^{i-1} \xrightarrow{d} Z^i \twoheadrightarrow H^i \to 0.$$

We will interpret an analogue of this with $i = 1$ and coefficients in G_{C^∞}.

We call an element of $\check{C}^0(X;G_{C^\infty})$ a *change of trivialization.* To explain, we obtain a right action of $\check{C}^0(X;C^\infty(G))$ on $\check{C}^0(X;P)$ by

$$\{\phi_\alpha\} \cdot \{g_\alpha\} := \{\phi_\alpha g_\alpha\}.$$

Furthermore, the action is transitive (one orbit) since, after possibly refining the open cover, any $\{\phi'_\alpha\}$ is obtained by $\{\phi'_\alpha\} = \{\phi_\alpha\} \cdot \left\{\left(\phi_\alpha^{-1}\phi'_\alpha\right)\right\}$.

There is also a right action of $\check{C}^0(X; G_{C^\infty})$ on $\check{C}^1(X; G_{C^\infty})$ given by $\left\{\phi_{\alpha\beta}\right\} \cdot \{g_\alpha\} := \left\{g_\alpha^{-1}\phi_{\alpha\beta}g_\beta\right\}$. Furthermore, under these actions the Čech differential is equivariant:

$$d\left(\{\phi_\alpha\} \cdot \{g_\alpha\}\right) = d\left(\{\phi_\alpha g_\alpha\}\right) = \left\{g_\alpha^{-1}\phi_\alpha^{-1}\phi_\beta g_\beta\right\} = d\left(\{\phi_\alpha\}\right) \cdot \{g_\alpha\}.$$

In summary,

$$\begin{tikzcd} \check{C}^0(X;P) \arrow[dr, "d"] & \\ \check{C}^0(X;G_{C^\infty}) \arrow[u, squiggly] \arrow[r, squiggly] & \check{Z}^1(X;G_{C^\infty}) \end{tikzcd}$$

where a squiggly arrow indicates a group action instead of an actual map. Note that since $\check{C}^0(X; P)$ consists of a single orbit, by equivariance, the image of d is the corresponding orbit in $\check{Z}^1(X; G_{C^\infty})$.

Suppose two principal bundles P_1 and P_2 share a common element in $\check{Z}^1(X; G_{C^\infty})$. Then P_1 and P_2 correspond to the same $\check{C}^0(X; G_{C^\infty})$ orbit. Furthermore, they are both isomorphic to the gluing construction P'. Thus P_1 and P_2 must be isomorphic. This establishes a correspondence

$$\begin{matrix}\text{iso classes of}\\ \text{principal } G\text{-bundles}\end{matrix} = \frac{\check{Z}^1(X; G_{C^\infty})}{\text{action of } \check{C}^0(X; G_{C^\infty})} =: \check{H}^1(X; G_{C^\infty}).$$

$$\begin{tikzcd} \check{C}^0(X;P) \arrow[dr, "d"] & & & \\ \check{C}^0(X;G_{C^\infty}) \arrow[u, squiggly] \arrow[r, squiggly] & \check{Z}^1(X;G_{C^\infty}) \arrow[r] & \check{H}^1(X;G_{C^\infty}) \arrow[r] & 0 \end{tikzcd}$$

The next question is when does a change of trivialization act trivially on transition functions. For a meaningful answer to this question, we should fix local trivializations $\{\phi_\alpha\} \in \check{C}^0(X; P)$. Then it could happen that our change of trivialization $\{g_\alpha\} \in \check{C}^0(X; G_{C^\infty})$ satisfies

$$\begin{aligned} g_\alpha^{-1}\phi_{\alpha\beta}g_\beta &= \phi_{\alpha\beta} \\ \iff g_\alpha &= \phi_{\alpha\beta}g_\beta\phi_{\alpha\beta}^{-1}, \end{aligned}$$

i.e. the g_α transform via the adjoint representation. To cast this in the correct language, we should interpret $\check{Z}^0$ for any associated bundle

$$\{f_\alpha \in C^\infty(U_\alpha, F)\} \in \check{Z}^0(X; P \times_\rho F) = \check{H}^0(X; P \times_\rho F)$$

not in terms of the kernel of some differential, but rather as a collection of local sections which agree on the overlaps, thus defining a global section. In corresponding trivializations, $f_\alpha = \rho(\phi_{\alpha\beta})f_\beta$. Thus in our case, $\{g_\alpha\} \in \check{Z}^0(X; P \times_{\mathrm{Ad}} G) = \check{Z}^0(X; \mathrm{Aut}(P)) = \mathcal{G}_P$. So the changes of trivialization which preserve transition functions are the gauge transformations.

$$\begin{tikzcd}
 & & \overbrace{\check{C}^0(X;P)}^{\text{local trivializations}} \arrow[dr, "d"] & & & \\
0 \arrow[r] & \underbrace{\check{Z}^0(X;\mathrm{Aut}(P))}_{\text{bundle automorphisms}} \arrow[r] & \underbrace{\check{C}^0(X;G_{C^\infty})}_{\text{changes of trivialization}} \arrow[u, squiggly] \arrow[r, squiggly] & \underbrace{\check{Z}^1(X;G_{C^\infty})}_{\text{transition functions}} \arrow[r] & \underbrace{\check{H}^1(X;G_{C^\infty})}_{\text{isomorphism classes of smooth principal } G\text{-bundles}} \arrow[r] & 0
\end{tikzcd}$$

We should mention the caveat that our identification of $\check{Z}^0(X;\mathrm{Aut}(P))$ with a subset of $\check{C}^0(X;G_{C^\infty})$ depends on the choice of local trivializations $\{\phi_\alpha\}$.

3.6 Sequences from coefficients

In ordinary cohomology, a short exact sequence of abelian groups $0 \to A \to B \to C \to 0$ gives a short exact sequence of chain complexes $0 \to C^\bullet(X;A) \to C^\bullet(X;B) \to C^\bullet(X;C) \to 0$ which gives a long exact sequence of cohomology

$$\cdots \to H^i(X;A) \to H^i(X;B) \to H^i(X;C) \to H^{i+1}(X;A) \to H^{i+1}(X;B) \to \cdots$$

We can imitate this with principal bundles. For example, consider

$$0 \to \mathbb{Z} \to \mathbb{R} \xrightarrow{e^{2\pi i x}} \mathrm{U}(1) \to 0.$$

This gives

$$\cdots \to \check{H}^1(X;\mathbb{R}_{C^\infty}) \to \check{H}^1(X;\mathrm{U}(1)_{C^\infty}) \to \check{H}^2(X;\mathbb{Z}) \to \check{H}^2(X;\mathbb{R}_{C^\infty}) \to \cdots$$

Beware that $\check{H}^p(X;\mathbb{R}_{C^\infty})$ is *not* equivalent to $H^p(X;\mathbb{R})$. Resolving the space of locally constant $\mathbb{R}$-valued functions is much more interesting than resolving the space of *smooth* $\mathbb{R}$-valued functions. Since $C^\infty(\mathbb{R})$ already has a partition of unity, it is resolved by

$$C^\infty(X;\mathbb{R}) \to 0 \to 0 \to \cdots .$$

Thus $\check{H}^0(X;\mathbb{R}_{C^\infty}) = C^\infty(X;\mathbb{R})$, and $\check{H}^p(X;\mathbb{R}_{C^\infty}) = 0$ for other p. On the other hand, since $\mathbb{Z}$ has the discrete topology, smooth $\mathbb{Z}$-valued functions are locally constant. Thus

$$0 \to \check{H}^1(X;\mathrm{U}(1)_{C^\infty}) \xrightarrow{c_1} H^2(X;\mathbb{Z}) \to 0$$

so the group of isomorphism classes of smooth principal U(1) bundles is equivalent to the group $H^2(X;\mathbb{Z})$. Via the standard representation, principal U(1) bundles correspond to complex line bundles. Group composition on $\check{H}^1(X;\mathrm{U}(1)_{C^\infty})$ corresponds to multiplying together the U(1)-valued transition functions, which is equivalent to tensor product. The isomorphism labeled c_1 is called the first Chern class. The homological algebra makes c_1 explicit. The recipe to compute it is as follows.

- Fix a "good cover" $\{U_\alpha\}$ of X so that all intersections are contractible.
- Given an isomorphism class $[P]$, choose a representative P. Pick local trivializations to obtain transition functions $\left\{\phi_{\alpha\beta}\right\} \in \check{Z}^1(X; \mathrm{U}(1)_{C^\infty})$ for P.
- Since each U_α is contractible, we can choose branches for $\left\{\eta_{\alpha\beta} = (2\pi i)^{-1}\log\phi_{\alpha\beta}\right\} \in \check{C}^1(X; \mathbb{R}_{C^\infty})$.
- Consider $d\left\{\eta_{\alpha\beta}\right\} = \left\{\eta_{\alpha\beta\gamma} = \eta_{\beta\gamma} - \eta_{\alpha\gamma} + \eta_{\alpha\beta}\right\}$. By the cocycle condition for $\left\{\phi_{\alpha\beta}\right\}$, $e^{2\pi i\eta_{\alpha\beta\gamma}} = 1$. Thus $\left\{\eta_{\alpha\beta\gamma}\right\} \in \check{Z}^2(X; \mathbb{Z})$.
- The cohomology class $c_1([P]) \in \check{H}^2(X; \mathbb{Z})$ is represented by $\left\{\eta_{\alpha\beta\gamma}\right\}$.

It's tedious but routine to verify that the result is independent of choices.

3.7 Extension of structure group

Recall that a central extension $\tilde{G}$ of G is a short exact sequence of the form

$$0 \to A \to \tilde{G} \to G \to 0,$$

where A is an (abelian) subgroup of $\tilde{G}$. We want to know when it's possible to lift transition functions from a structure group to an extension. The prototypical example is

$$0 \to \mathbb{Z}_2 \to \mathrm{Spin}(k) \to \mathrm{SO}(k) \to 0.$$

The group $\mathrm{Spin}(k)$ for $k > 1$ is characterized as the unique nontrivial $\mathbb{Z}_2$ extension of $\mathrm{SO}(k)$: Thus $\mathrm{Spin}(k)$ is the total space of a principal $\mathbb{Z}_2$-bundle over $\mathrm{SO}(k)$. These are classified topologically by

$$H^1(\mathrm{SO}(k); \mathbb{Z}_2) = \mathrm{Hom}(H_1(\mathrm{SO}(k); \mathbb{Z}); \mathbb{Z}_2) \oplus 0 = \mathrm{Hom}(\pi_1(\mathrm{SO}(k))^{\mathrm{ab}}; \mathbb{Z}_2).$$

We know that $\pi_1(\mathrm{SO}(2)) = \pi_1(S^1) = \mathbb{Z}$, and $\pi_1(\mathrm{SO}(3)) = \pi_1(S^3/\mathbb{Z}_2) = \mathbb{Z}_2$. It's easy to show that $\pi_1(\mathrm{SO}(k+1)) = \pi_1(\mathrm{SO}(k))$ for $k \geq 3$. Thus $H^1(\mathrm{SO}(k); \mathbb{Z}_2) = \mathbb{Z}_2$ has a unique nontrivial element corresponding topologically to $\mathrm{Spin}(k)$.

The part of the long exact sequence of Čech cohomology which makes sense is given by

$$H^1(X; \mathbb{Z}_2) \rightsquigarrow \check{H}^1(X; \mathrm{Spin}(k)_{C^\infty}) \to \check{H}^1(X; \mathrm{SO}(k)_{C^\infty}) \xrightarrow{w_2} H^2(X; \mathbb{Z}_2).$$

An isomorphism class $[P] \in \check{H}^1(X; \mathrm{SO}(k)_{C^\infty})$ comes from an element of $\check{H}^1(X; \mathrm{Spin}(k)_{C^\infty})$ if and only if $w_2([P]) = 0 \in H^2(X; \mathbb{Z}_2)$. There can be several isomorphism classes of principal $\mathrm{Spin}(k)$ bundles lifting the same class of $\mathrm{SO}(k)$ bundles. The action of $H^1(X; \mathbb{Z}_2)$ is transitive, but not always free. However this action becomes free if we refine our notion of lift. This refinement is an essential subtlety for the definition of a spin structure.

As before, suppose $0 \to A \to \tilde{G} \to G \to 0$ is a central extension, and P is some fixed principal G-bundle. A lift of P to the structure group $\tilde{G}$ is a principal $\tilde{G}$-bundle $\tilde{P}$ equipped with an isomorphism of P with the G-bundle associated to the quotient of $\tilde{P}$ by A. Two lifts are equivalent if they are related by an isomorphism of $\tilde{P}$ which induces *the identity* on P. (A general isomorphism of $\tilde{P}$ induces an isomorphism on P which is not necessarily the identity!) Be warned that it is possible for inequivalent lifts to be isomorphic as principal $\tilde{G}$-bundles.

To understand the equivalence classes of lifts of such a bundle P, suppose that $\left\{\phi_{\alpha\beta}\right\} \in \check{Z}^1(X;G)$ is a Čech cocycle representing the the transition functions for P relative to some local trivialization. If the cochain $\left\{\tilde{\phi}_{\alpha\beta}\right\} \in \check{Z}^1(X;\tilde{G})$ is an arbitrary choices for lifts to $\tilde{G}$, then the combination $w_2 := \left[\left\{\tilde{\phi}_{\alpha\beta}\tilde{\phi}_{\beta\gamma}\tilde{\phi}_{\gamma\alpha}\right\}\right] \in \check{H}^2(X;A)$ is called the generalized second Stiefel-Whitney class, and depends only on the isomorphism class of P, i.e. $\left[\left\{\phi_{\alpha\beta}\right\}\right] \in \check{H}^1(X;G)$. The cochain $\left\{\tilde{\phi}_{\alpha\beta}\right\}$ can be chosen to be a cocycle if and only if $w_2([P]) = 0$. Such a cochain then corresponds to a lift $\tilde{P}$ of P. Any other lift is of the form $\left\{a_{\alpha\beta}\tilde{\phi}_{\alpha\beta}\right\}$ for $\left\{a_{\alpha\beta}\right\} \in \check{Z}^1(X;A)$, and two such lifts are isomorphic if and only if they represent the same element of $\check{H}^1(X;A)$. In this manner, the space of lifts of P up to equivalence is an $\check{H}^1(X;A)$-torsor when $w_2([P]) = 0$, and empty otherwise.

Recall from last time, we derived the "exact sequence"

$$\begin{tikzcd}
 & & \overbrace{\check{C}^0(X;P)}^{\substack{\text{local}\\ \text{trivializations}\\ \{\phi_\alpha\}}} \arrow[drr, "d=\{\phi_\alpha\}\mapsto\{\phi_{\alpha\beta}:=\phi_\alpha^{-1}\phi_\beta\}"] & & & \\
0 \arrow[r] & \underbrace{\check{Z}^0(X;\mathrm{Aut}(P))}_{\substack{\text{bundle}\\ \text{automorphisms}\\ \{g_\alpha\}\mid g_\alpha=\phi_{\alpha\beta}g_\beta\phi_{\alpha\beta}^{-1}}} \arrow[r] & \underbrace{\check{C}^0(X;G_{C^\infty})}_{\substack{\text{changes of}\\ \text{trivialization}\\ \{g_\alpha\}}} \arrow[u, rightsquigarrow, "\{\phi_\alpha\}\cdot\{g_\alpha\}:=\{\phi_\alpha g_\alpha\}"] \arrow[rr, rightsquigarrow, "\{\phi_{\alpha\beta}\}\cdot\{g_\alpha\}:=\{g_\alpha^{-1}\phi_{\alpha\beta}g_\beta\}"] & & \underbrace{\check{Z}^1(X;G_{C^\infty})}_{\substack{\text{transition}\\ \text{functions}\\ \{\phi_{\alpha\beta}\}\mid\phi_{\beta\gamma}\phi_{\alpha\gamma}^{-1}\phi_{\alpha\beta}=1}} \arrow[r] & \underbrace{\check{H}^1(X;G_{C^\infty})}_{\substack{\text{isomorphism}\\ \text{classes of smooth}\\ \text{principal } G\text{-bundles}}} \arrow[r] & 0
\end{tikzcd}$$

Note that bundle automorphisms transform via the adjoint representation. Also, local trivializations correspond to local sections of P. In particular, a global section of P corresponds to a global trivialization of P, so P has global sections iff it is isomorphic to the trivial principal bundle $P \cong X \times G$.

Another important question is when does Čech cohomology with smooth coefficients $\check{H}^k(X;G_{C^\infty})$ coincide with ordinary cohomology $H^k(X;G)$ with locally constant coefficients. In order for $H^k(X;G)$ to make sense, G should be abelian. In this case, in order to naturally identify $\check{H}^k(X;G_{C^\infty})$ with $H^k(X;G)$, we want to identify the sheaf G_{C^∞} with the locally constant sheaf G. But G_{C^∞} corresponds with the locally constant sheaf whenever G has the discrete topology. For example, smooth $\mathbb{Z}$-valued functions are necessarily locally constant. In contrast, $\check{H}^k(X;\mathbb{R}_{C^\infty}) = 0$ for $k > 0$ thanks to partitions of unity, while $H^k(X;\mathbb{R})$ is usually nontrivial.

3.8 Reduction of structure group

Suppose P is a principal G-bundle, and $H \subset G$ is a subgroup (not necessarily normal). Then we have a "short exact sequence"

$$0 \to H \to G \to G/H \to 0.$$

(When H is not normal, we should think of "exactness" as the property that each coset of G has a free and transitive action of H.)

To keep in mind a concrete example, consider

$$0 \to \mathrm{O}(k) \to \mathrm{GL}(k) \to \mathrm{Met}(k) \to 0,$$

where

$$\mathrm{Met}(k) := \{\text{symmetric positive-definite matrices}\} = \mathrm{GL}(k)/\mathrm{O}(k).$$

How is $\mathrm{GL}(k)/\mathrm{O}(k)$ identified with positive-definite matrices? Consider the map $\mathrm{GL}(k) \to \mathrm{Met}(k)$ is given by $g \mapsto (g^T)^{-1}\,\mathrm{Id}_{k\times k}\,g^{-1} = \rho_{\mathrm{Met}}(g)\,\mathrm{Id}_{k\times k}$. (What's the reason for this transformation law? Hint: we want the inner product of two vectors to be independent of frame.) Clearly if h is orthogonal, then gh has the same image as g, so we get a well-defined map from left cosets $\mathrm{GL}(k)/\mathrm{O}(k) \to \mathrm{Met}(k)$.

Exercise. Verify that the map $M \mapsto M^{-1/2}$ is the two-sided inverse $\mathrm{Met}(k) \to \mathrm{GL}(k)/\mathrm{O}(k)$ by assuming the polar decomposition $g = ph$ for any $g \in \mathrm{GL}(k)$, p positive-definite, and h orthogonal.

The associated bundle $P \times_{\rho_L} \mathrm{GL}(k)/\mathrm{O}(k) = P \times_{\rho_{\mathrm{Met}}} \mathrm{Met}(k)$ then corresponds to the bundle of Euclidean metrics on the corresponding fibers. A global section

$$s \in \check{H}^0(X; P \times_{\rho_{\mathrm{Met}}} \mathrm{Met}(k)) = \Gamma(X; P \times_{\rho_{\mathrm{Met}}} \mathrm{Met}(k))$$

corresponds to a metric on the associated standard vector bundle $E = P \times_{\rho_{\mathrm{st}}} \mathbb{R}^k$. A metric then determines an $\mathrm{O}(k)$ structure on P. It picks out the subset of local trivializations which are orthonormal, and upon restriction to these, the transition functions take values in $\mathrm{O}(k)$.

More abstractly, given the principal G-bundle P, we seek to modify it so that the transition functions take values in H. Specifically, suppose $\{\phi_\alpha\}$ are local trivializations such that $\left\{\phi_{\alpha\beta}\right\} \in \check{Z}^1(X; G_{C^\infty})$. We seek a change of trivialization $\{g_\alpha\} \in \check{C}^0(X; G_{C^\infty})$ such that $\left\{\phi_{\alpha\beta}\right\} \cdot \{g_\alpha\}$ belongs to $\check{Z}^1(X; H_{C^\infty})$. Thus $g_\alpha^{-1}\phi_{\alpha\beta}g_\beta = h_{\alpha\beta}$ for some $h_{\alpha\beta}$ with values in H. Equivalently, $g_\alpha = \phi_{\alpha\beta}g_\beta h_{\alpha\beta}^{-1} \in \phi_{\alpha\beta}g_\beta H$, so

$$\{g_\alpha H\} \in \check{H}^0(X; P \times_{\rho_L} G/H).$$

Conversely, given any section of $P \times_{\rho_L} G/H$, it's easy to see that if we can locally locally we can lift to $\{g_\alpha\} \in \check{C}^0(X; G_{C^\infty})$, then the corresponding

$$h_{\alpha\beta} := g_\alpha^{-1}\phi_{\alpha\beta}g_\beta \in C^\infty(U_{\alpha\beta}; H)$$

so that

$$\left\{h_{\alpha\beta}\right\} \in \check{Z}^1(X; H_{C^\infty})$$

determines a principal H-bundle. Homologically, we have the following “exact sequence”:

$$\begin{tikzcd}
\mathscr{G}_P = \underbrace{\check{H}^0(X; P \times_{\mathrm{Ad}} G)}_{\{g_\alpha\} | g_\alpha = \phi_{\alpha\beta} g_\beta \phi_{\alpha\beta}^{-1}} \arrow[r, rightsquigarrow, "\{\tilde{g}_\alpha\} \cdot \{g_\alpha H\} := \{\tilde{g}_\alpha g_\alpha H\}"] & \underbrace{\check{H}^0(X; P \times_{\rho_L} G/H)}_{\substack{\text{reductions} \\ \{g_\alpha H\} | g_\alpha H = \phi_{\alpha\beta} g_\beta H}} \arrow[r, "\{g_\alpha H\} \mapsto \{g_\alpha^{-1} \phi_{\alpha\beta} g_\beta\}"] & \check{H}^1(X; H_{C^\infty}) \arrow[dl] \\
& \check{H}^1(X; G_{C^\infty}) \arrow[r] & \check{H}^1(X; (G/H)_{C^\infty})
\end{tikzcd}$$

Note that $\check{H}^1(X; (G/H)_{C^\infty})$ only really makes sense when H is normal, since otherwise there is no clear interpretation of the cocycle condition $\phi_{\beta\gamma} H \left(\phi_{\alpha\gamma} H\right)^{-1} \phi_{\alpha\beta} H = 1$. In this case, $P \times_{\rho_L} G/H$ is the associated principal G/H bundle. Homological algebra dictates that reductions should only exist when the corresponding principal bundle $P \times_{\rho_L} G/H$ is trivial. Indeed, reductions correspond to global sections of this principal bundle, so they exist iff it is trivial.

Regardless of whether or not H is normal, what really counts is the space of reductions $\Gamma(X; P \times_{\rho_L} G/H)$. These reductions, up to the action by gauge transformations, parameterize the isomorphism classes of smooth principal H-bundles over P.

A more sophisticated application of this formalism is the following:

Theorem 23. *Over a complex manifold X, a reduction from a smooth vector bundle $E \to X$ to a holomorphic vector bundle is equivalent to a $\overline{\partial}_\alpha$-operator on E which satisfies $\overline{\partial}_\alpha^2 = 0$.*

To understand why, consider the sequence of sheaves given by

$$0 \to \mathscr{O}(\mathrm{GL}(k; \mathbb{C})) \to C^\infty(\mathrm{GL}(k; \mathbb{C})) \to \mathrm{Hol}(k) \to 0,$$

where $\mathscr{O}(\mathrm{GL}(k; \mathbb{C}))$ denotes the sheaf of holomorphic functions valued in $\mathrm{GL}(k; \mathbb{C})$, and $\mathrm{Hol}(k)$ denotes the space of operators

$$\overline{\partial}_\alpha : \Omega^0(U; \mathbb{C}^k) \to \Omega^{0,1}(U; \mathbb{C}^k)$$

subject to the additional constraints

- $\overline{\partial}_\alpha(f\, s) = \left(\overline{\partial} f\right) s + f\, \overline{\partial}_\alpha s$ for all $f \in \Omega^0(U; \mathbb{C})$ and $s \in \Omega^0(U; \mathbb{C}^k)$,
- $0 = \overline{\partial}_\alpha^2 s \in \Omega^{0,2}(U; \mathbb{C})$ for all $s \in \Omega^0(U; \mathbb{C}^k)$.

The map $C^\infty(\mathrm{GL}(k)) \to \mathrm{Hol}(k)$ is given by

$$g \mapsto g \circ \overline{\partial} \circ g^{-1} = g \circ \left(\overline{\partial} g^{-1} + g^{-1} \overline{\partial}\right) = \overline{\partial} - (\overline{\partial} g) g^{-1},$$

where $\overline{\partial}$ is the coordinatewise standard $\overline{\partial}$ operator.

Assuming the exactness of the above short exact sequence of sheaves, we expect an exact sequence in Čech cohomology of the form

$$\check{H}^0(X; P \times_\rho \mathrm{Hol}(k)) \to \check{H}^1(X; \mathrm{GL}(k; \mathbb{C})_{\mathscr{O}}) \to \check{H}^1(X; \mathrm{GL}(k; \mathbb{C})_{C^\infty}).$$

Here we have principal bundles $\check{H}^1(X; \mathrm{GL}(k)_{\mathcal{O}})$ which correspond to holomorphic vector bundles, i.e. vector bundles with holomorphic transition functions. Then $\check{H}^0(X; P \times_\rho \mathrm{Hol}(k))$ corresponds to a $\overline{\partial}_\alpha$ operator on a smooth vector bundle P, whose kernel selects the "holomorphic sections."

The bulk of the work for this picture amounts to showing exactness of:

$$0 \to \mathcal{O}(\mathrm{GL}(k;\mathbb{C})) \to C^\infty(\mathrm{GL}(k;\mathbb{C})) \to \mathrm{Hol}(k) \to 0.$$

Exactness at the left is obvious, since holomorphic functions are a subspace of smooth functions. Exactness at the center is simply the statement that

$$g\overline{\partial}g^{-1} = \tilde{g}\overline{\partial}\tilde{g}^{-1} \iff \overline{\partial}(g^{-1}\tilde{g}) = 0.$$

The hardest part is surjectivity. The condition $\overline{\partial}_\alpha^2 s = 0$ is an integrability condition which ensures that we can find g such that $\overline{\partial}_\alpha = g\overline{\partial}g^{-1}$. For details of the integrability theorem, see Donaldson and Kronheimer, 2.1.53.

3.9 Classification of principal bundles on a 4-manifold

A connected Lie group G is called *simple* if it is non-abelian, and the Lie algebra $\mathfrak{g}$ of G has no non-trivial ideals besides $0, \mathfrak{g}$. For example, $\mathrm{U}(1)$ is not simple since it is abelian. More generally, $\mathrm{U}(k)$ is not simple since its Lie algebra contains $\mathfrak{u}(1)$ as an ideal. However, $\mathrm{SU}(k)$ is simple. The special orthogonal groups $\mathrm{SO}(k)$ are simple for $k = 3$ and $k \geq 5$. Compact simple simply-connected Lie groups $\tilde{G}$ are in bijection with admissible Dynkin diagrams. Classification of Dynkin diagrams yields the list of possible groups

$$\tilde{G} \in \{\mathrm{SU}(k), \mathrm{Spin}(k), \mathrm{Sp}(2k), E_6, E_7, E_8, F_4, G_2\}.$$

The center $Z(\tilde{G})$ is the subgroup of elements which commute with everything else, and is determined by

$$Z(\tilde{G}) = \Lambda_{\mathrm{weight}}/\Lambda_{\mathrm{root}},$$

which is a finite abelian group.

More generally, any compact simple non-simply-connected Lie group G is the quotient of $\tilde{G}$ by a subgroup of its center. This subgroup lifts to an intermediate lattice given by the kernel of the exponential map:

$$\Lambda_{\mathrm{root}} \subset \ker(\exp) \subset \Lambda_{\mathrm{weight}}.$$

Then $Z(G) = \Lambda_{\mathrm{weight}}/\ker(\exp)$, and $\pi_1(G) = \ker(\exp)/\Lambda_{\mathrm{root}}$ are both finite abelian groups.

For example, consider $\tilde{G} = \mathrm{SU}(k)$. A maximal torus $T \subset \tilde{G}$ is the diagonal matrices of determinant 1, which is a copy of $\mathrm{U}(1)^{k-1} \subset \mathrm{U}(1)^k$. The center $Z(\tilde{G})$ is isomorphic to $\mathbb{Z}_k$, consisting of multiples of the identity matrix $e^{2\pi i/k}I$. The Lie algebra $\mathfrak{t}$ of T is $\left\{(\theta_1, \ldots, \theta_{n+1}) \in \mathbb{R}^{n+1} | \sum \theta_i = 0\right\}$. The root lattice is the kernel of the exponential map, which is the subset intersecting $2\pi\mathbb{Z}^{n+1}$. The weight lattice is where the exponential map hits the center, i.e. the intersection of $\mathfrak{t}$ with

$$2\pi\mathbb{Z}^{n+1} + 2\pi/k\,(1, \ldots, 1)\,\mathbb{Z}.$$

The group $\mathrm{PU}(k) := \mathrm{SU}(k)/Z(\mathrm{SU}(k))$ has fundamental group $\pi_1(\mathrm{PU}(k)) = Z(\mathrm{SU}(k)) = \mathbb{Z}_k$. Note that $\mathrm{Spin}(3) = \mathrm{SU}(2)$, and $\mathrm{PU}(2) = \mathrm{SU}(2)/\mathbb{Z}_2 = \mathrm{SO}(3)$.

...

Chapter 4

Connections

4.1 Connections on principal bundles

Definition 24. A *connection* on a principal G-bundle $P \to X$ is an element of $\{A_\alpha\} \in \check{C}^0(X; T^*X \otimes \mathfrak{g})$ which satisfies

$$A_\alpha = \mathrm{Ad}_{\phi_{\alpha\beta}} A_\beta - (d\phi_{\alpha\beta})\phi_{\alpha\beta}^{-1}. \tag{4.1}$$

If $\rho : G \to \mathrm{Aut}(V)$ is a linear representation, $E = P \times_\rho V$, and and $s \in \Gamma(E)$, then the induced connection $\nabla^A s \in \Gamma(T^*X \otimes E)$ is well-defined, given by

$$(\nabla^A s)_\alpha := \nabla(s_\alpha) + \rho(A_\alpha)s_\alpha,$$

where $\rho : \mathfrak{g} \to \mathrm{Lie}(\mathrm{Aut}(V))$ also denotes the induced Lie algebra map.

Example 25. If $\rho_{\mathrm{st}} : \mathrm{O}(n) \to \mathrm{Iso}(\mathbb{R}^n)$, then in an orthonormal frame, A_α is $\mathfrak{o}(n)$-valued, and hence an antisymmetric matrix of one-forms. In this case, it is automatically compatible with the inner product since in $\Gamma(U_\alpha; T^*X)$,

$$\left\langle \nabla^A s, t \right\rangle + \left\langle s, \nabla^A t \right\rangle = \langle \nabla s_\alpha, t_\alpha \rangle + \langle A_\alpha s_\alpha, t_\alpha \rangle + \langle s_\alpha, \nabla t_\alpha \rangle + \langle s_\alpha, A_\alpha t_\alpha \rangle = \nabla \langle s_\alpha, t_\alpha \rangle = \nabla \langle s, t \rangle .$$

Conversely, since ρ_{st} is an isomorphism, any ordinary compatible connection on the standard associated vector bundle determines, in any local orthonormal frame ϕ_α, an antisymmetric connection 1-form A_α which satisfies the transformation law.

We want to show that connections exist on any principal bundle. This is the done in the same way that ordinary connections are proven to exist on vector bundles. The key observation is that if ∇^A and ∇^B are two connections, and $f \in C^\infty(X)$, then $f\nabla^A + (1-f)\nabla^B$ still satisfies the product rule, since

$$f\nabla^A(gs) + (1-f)\nabla^B(gs) = f\,dg \otimes s + fg\nabla^A s + (1-f)\,dg \otimes s + (1-f)g\nabla^B s = dg \otimes s + g(f\nabla^A + (1-f)\nabla^B)s.$$

Now we graft together the trivial connections over each U_γ. Choose a partition of unity $\sum_\gamma f_\gamma = 1$ such that f_γ is supported in U_γ. Denote by ∇^{A^γ} the connection over U_γ given by $(A^\gamma)_\gamma = 0$, so that $(\nabla^{A^\gamma})_\gamma = \nabla + \rho(0)$. Then use the connection $\sum_\gamma f_\gamma \nabla^{A^\gamma}$, which obeys the product rule since $\sum_\gamma f_\gamma = 1$. Now let's determine the formula for A_α. First we compute $(A^\gamma)_\alpha = \mathrm{Ad}_{\phi_{\alpha\gamma}} 0 - d\phi_{\alpha\gamma}\phi_{\alpha\gamma}^{-1}$. Thus $A_\alpha = -\sum_\gamma f_\gamma d\phi_{\alpha\gamma}\phi_{\alpha\gamma}^{-1}$.

Exercise: Prove that this A_α satisfies (4.1).

Now we know that connections always exist. Now it is easy to classify them.

For any fixed connection A_0, the space of all connections is given by $\mathscr{A}_P = A_0 + \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$, where $\mathfrak{g}_{\mathrm{Ad}} := P \times_{\mathrm{Ad}} \mathfrak{g}$.

Proof: $(A - A_0)_\alpha = \mathrm{Ad}_{\phi_{\alpha\beta}}(A - A_0)_\alpha \iff A - A_0 \in \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$.

Now we want to try to understand how to find distinguished connections on a principal bundle.

Suppose A is a $\mathfrak{gl}(n)$ connection on $\mathrm{Fr}(TX)$. Then the torsion tensor

$$T^A(X,Y) := \nabla_X^A Y - \nabla_Y^A X - [X,Y].$$

Fundamental theorem of Riemannian geometry: If X is a Riemannian manifold, then there is a unique $\mathfrak{o}(n)$ connection on $\mathrm{Fr}^O(TX)$ called the Levi-Civita connection, denoted by LC, such that $T^{\mathrm{LC}} = 0$.

Example: Suppose G is a compact Lie group. Then, by averaging, there exists a metric which is invariant under both left and right multiplication by any $g \in G$, denoted respectively by L_g and R_g. We identify any $\xi \in \mathfrak{g}$ with the left-invariant vector field whose value at g is $(L_g)_*\xi \in T_gG$. In this way, the Lie bracket of vector fields corresponds to the Lie algebra. (Using right-invariant fields would introduce a minus sign into the Lie bracket.) The Levi-Civita connection on G is given by $\nabla_\xi \eta = \frac{1}{2}[\xi, \eta]$. Torsion vanishes by antisymmetry of the Lie bracket. Compatibility follows from bi-invariance, which is equivalent to $\langle \chi, [\xi,\eta] \rangle = \langle [\chi, \xi], \eta \rangle$.

The convention is to implicitly use the Levi-Civita connection on all tensors associated to the tangent bundle.

For example, consider $\alpha \otimes s \in \Gamma(T^*X \otimes E)$, and we want to compute $\nabla_A(\alpha \otimes s)$ using some connection A on the principal bundle for E. We use the tensor representation

$$\rho_{\mathrm{st}} \otimes \rho : O(n) \times G \to \mathrm{Aut}(T^*X \otimes E).$$

The infinitesimal version is

$$(\rho_{\mathrm{st}} \otimes \mathrm{Id} + \mathrm{Id} \otimes \rho) : \mathfrak{o}(n) \oplus \mathfrak{g} \to \mathrm{Lie}(\mathrm{Aut}(T^*X \otimes E)).$$

Thus the covariant derivative is given locally by

$$\nabla_\alpha = \nabla + A_\alpha^{\mathrm{LC}} \otimes \mathrm{Id} + \mathrm{Id} \otimes A_\alpha.$$

This allows us to define, for instance, the second covariant derivative

$$(\nabla^A)^2 s \in \Gamma(T^*X \otimes T^*X \otimes E).$$

Most importantly, we have curvature

$$R^A(X,Y)s := (\nabla^A)^2_{X \otimes Y - Y \otimes X} s = (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})s.$$

It is convenient to introduce the operator $d_A : \Omega^p(X;E) \to \Omega^{p+1}(X;E)$ defined by Λd_A, where

$$\Lambda : T^*X \otimes \Lambda^p T^*X \to \Lambda^{p+1}T^*X$$

by the wedge product. Then

$$R^A(X,Y)s = (d_A d_A s)(X,Y),$$

and

$$(d_A s)_\alpha = d + \rho(A_\alpha)\wedge .$$

Defining

$$[(\xi_1 \otimes \omega_1) \wedge (\xi_2 \otimes \omega_2)] := [\xi_1, \xi_2] \otimes \omega_1 \wedge \omega_2 = (-1)^{1+\deg\omega_1 \deg\omega_2}\,[(\xi_2 \otimes \omega_2) \wedge (\xi_1 \otimes \omega_1)]$$

we compute

$$\begin{aligned} d_A d_A s &= (d + \rho(A_\alpha)\wedge)(d + \rho(A_\alpha)\wedge) \\ &= d^2 + [d, \rho(A_\alpha)\wedge] + \tfrac{1}{2}\,[\rho(A_\alpha)\wedge, \rho(A_\alpha)\wedge] \\ &= 0 + \rho(dA_\alpha)\wedge + \rho\left(\tfrac{1}{2}\,[A_\alpha \wedge A_\alpha]\right)\wedge \\ &= \rho((F_A)_\alpha)\wedge, \end{aligned}$$

where

$$(F_A)_\alpha := dA_\alpha + \tfrac{1}{2}\,[A_\alpha \wedge A_\alpha] \in \Omega^2(U_\alpha; \mathfrak{g}).$$

Some computation shows that F_A does not depend on the Riemannian metric, and it transforms according to the adjoint representation, so $F_A \in \Omega^2(X; \mathfrak{g}_{\mathrm{Ad}})$. It follows that $F_A(X,Y) \in \Gamma(\mathfrak{g}_{\mathrm{Ad}})$, so $\rho(F_A(X,Y))$ gives an endomorphism on each fiber, and

$$(\nabla^A)^2_{X\otimes Y - Y \otimes X} s = R^A(X,Y)s = (d_A^2 s)(X,Y) = \rho(F_A(X,Y))(s).$$

For example, on a Lie group G,

$$R^{\mathrm{LC}}(\xi,\eta)\chi = \tfrac{1}{4}\,[\xi,[\eta,\chi]] - \tfrac{1}{4}\,[\eta,[\xi,\chi]] - \tfrac{1}{2}\,[[\xi,\eta],\chi] = -\tfrac{1}{4}\,[[\xi,\eta],\chi] = \mathrm{ad}\left(-\tfrac{1}{4}\,[\xi,\eta]\right)\chi.$$

4.2 Flat bundles

Definition 26. A *flat* principal G-bundle on X is an element of $\check{Z}^1(X; G_{\mathrm{const}})$ whose transition functions are constant.

Theorem 27. *If X is connected, then isomorphism classes of flat principal G-bundles over X correspond to conjugacy classes of homomorphsims $\pi_1(X)^{\mathrm{op}} \to G$.*

Remark 28. The superscript "op" on $\pi_1(X)$ denotes the *opposite group*, which is the original group with the order of multiplication reversed. The standard convention for composition of loops in $\pi_1(X)$ is that $[\gamma_1] \cdot [\gamma_2]$ is represented by a path which traces firstly γ_1 and secondly γ_2, while in $\pi_1(X)^{\mathrm{op}}$ the order is reversed, which is more like the ordering in function composition where the rightmost function is applied first. Any group π is naturally isomorphic to π^{op} via the map $x \mapsto x^{-1}$.

Remark 29. Normally it is sloppy to omit the basepoint from the notation $\pi_1(X, x_0)$ because, while there are isomorphisms from $\pi_1(X, x_0)$ to $\pi_1(X, x_1)$ induced by homotopy classes of paths from x_0 to x_1, the isomorphism depends up to conjugation on the chosen path. But since we are interested only in homomorphisms up to conjugation, the choice of basepoint doesn't matter (assuming that X is connected).

Proof. First we construct a map $\check{Z}^1(X; G_{\text{const}}) \to \text{Hom}(\pi_1(X, x_0)^{\text{op}}, G)/\text{conj}$. Suppose $\pi : P \to X$ is a flat principal G-bundle. For any continuous path $\gamma : [0, 1] \to X$, define $\text{Hol}(\gamma) \in \text{Iso}(P|_{\gamma(0)}, P|_{\gamma(1)})$ as follows. By compactness, cover the image of γ by finitely many $U_{\alpha_1}, \ldots, U_{\alpha_k}$, so that $\gamma : [t_{i-1}, t_i] \to U_{\alpha_i}$ for some $0 = t_0 < t_1 < \cdots < t_k = 1$. For $i = 1, \ldots, k$, define an element $\tau_i \in \text{Iso}(P|_{\gamma(t_{i-1})}, P|_{\gamma(t_i)})$ according to the rule $\tau_i(p_{i-1}) = p_i$. Thus $(\tau_i)_\alpha^\beta (p_{i-1})_\beta = (p_i)_\alpha$. In particular, $(\tau_i)_{\alpha_i}^{\alpha_{i-1}} \cdot (p_{i-1})_{\alpha_{i-1}} = (p_i)_{\alpha_i} = \phi_{\alpha_i\alpha_{i-1}} \cdot (p_i)_{\alpha_{i-1}}$, so $(\tau_i)_{\alpha_i}^{\alpha_{i-1}} = \phi_{\alpha_i\alpha_{i-1}}$. Finally, define $\text{Hol}(\gamma) := \tau_k \circ \tau_{k-1} \circ \cdots \circ \tau_1$, so that

$$\text{Hol}(\gamma)_{\alpha_k}^{\alpha_1} = \phi_{\alpha_k\alpha_{k-1}}\phi_{\alpha_{k-1}\alpha_{k-2}} \cdots \phi_{\alpha_2\alpha_1} \in G.$$

Given a change of trivialization $\left\{\phi_{\beta_i} = \phi_{\alpha_i} g_i^{-1}\right\}_{i=1}^k$ so that $\phi_{\beta_i\beta_{i-1}} = \phi_{\beta_i}^{-1}\phi_{\beta_{i-1}} = g_i \phi_{\alpha_i\alpha_{i-1}} g_{i-1}^{-1}$, it follows that

$$\text{Hol}(\gamma)_{\beta_k}^{\beta_0} = g_k\phi_{\alpha_k\alpha_{k-1}} g_{k-1}^{-1} g_{k-1}\phi_{\alpha_{k-1}\alpha_{k-2}} g_{k-2}^{-1} g_{k-2} \cdots g_2^{-1} g_2 \phi_{\alpha_2\alpha_1} g_1^{-1} = g_k \text{Hol}(\gamma)_{\alpha_k}^{\alpha_0} g_0^{-1}.$$

We are interested in the particular case where γ is a loop based at x_0 so that

$$\text{Hol}(\gamma) \in \text{Iso}(P|_{x_0}, P|_{x_0}) = \text{Ad}(P|_{x_0}),$$

given by

$$\text{Hol}(\gamma)_{\alpha_1}^{\alpha_1} = \phi_{\alpha_1\alpha_k}\phi_{\alpha_k\alpha_{k-1}}\phi_{\alpha_{k-1}\alpha_{k-2}} \cdots \phi_{\alpha_2\alpha_1},$$

which clearly transforms under the adjoint representation. Composition of loops corresponds to composition of the corresponding sequence of $\phi_{\alpha_i\alpha_{i-1}}$ in the right-to-left order of $\pi_1(X, x_0)^{\text{op}}$. Evidently, small perturbations of γ and the $\{t_i\}$ do not change the result. Nor does refinement of the cover. Finally, any homotopy of γ can be covered by finitely many U_α such that the initial path and final path are related by a sequence of refinements. Thus $\text{Hol}_{\alpha_1}^{\alpha_1}$ descends to a map $\pi_1(X, x_0)^{\text{op}} \to G$. It is straightforward to verify that this map is a homomorphism, and that change of trivialization by $\{g_{\alpha_1}\}$ acts as conjugation by $g_{\alpha_1}^{-1}$. Thus the desired map is well-defined.

To perform the inverse of the above construction, for some homomorphism, consider an open cover $\{U_\alpha\}$ such that

1. each U_α is simply-connected,
2. each $U_{\alpha\beta}$ is either empty or connected,
3. for each α, there are choices of both a basepoint $x_\alpha \in U_\alpha$, and a path γ_α from x_0 to x_α, and
4. if $x_0 \in U_\alpha$, then $x_\alpha = x_0$ and γ_α is chosen to be the constant path.

For each $U_{\alpha\beta} \neq \emptyset$, consider the element $\left[\gamma_{\alpha\beta}\right] \in \pi_1(X, x_0)^{\text{op}}$ represented by the following construction. Starting from x_0, follow γ_α to x_α. Then choose a path inside $U_\alpha \cup U_\beta$ from x_α to x_β. Then follow the reverse of γ_β back to x_0. The resulting homotopy class is independent of the choice of path from x_α to x_β, as a consequence of $\pi_1(U_\alpha \cup U_\beta) = \{1\}$ by the Seifert-van Kampen theorem. (The difference of any two chosen paths is a loop, which is nullhomotopic.) Furthermore, if $x_0 \in U_\alpha \cap U_\beta$, then $\gamma_{\alpha\beta}$ remains inside $U_\alpha \cup U_\beta$, and so $\left[\gamma_{\alpha\beta}\right] = 1$.

Given any homomorphism $h : \pi_1(X, x_0)^{\mathrm{op}} \to G$, we construct a corresponding flat bundle as follows. Define a bundle by the transition functions $\phi_{\alpha\beta} := h\left(\left[\gamma_{\alpha\beta}\right]\right)$ when $U_{\alpha\beta} \neq \emptyset$. To verify that this determines a bundle, we must check the cocycle condition that

$$\phi_{\alpha_1\alpha_2}\phi^{-1}_{\alpha_3\alpha_2}\phi_{\alpha_3\alpha_1} = e$$

when $U_{\alpha_1\alpha_2\alpha_3} \neq \emptyset$. It suffices to show that if $U_{\alpha_1\alpha_2\alpha_3} \neq \emptyset$ then $[\gamma_{\alpha_1\alpha_2}]\,[\gamma_{\alpha_2\alpha_3}]\,[\gamma_{\alpha_3\alpha_1}] = e$. By cancelling the paths γ_{α_2} and γ_{α_3} with their reverses, and changing basepoint to x_{α_1} along γ_{α_1}, this class is represented by a loop from x_{α_1} to x_{α_3} to x_{α_2} and back to x_{α_1} which is contained in $U_{\alpha_1} \cup U_{\alpha_2} \cup U_{\alpha_3}$, which we seek to show is nullhomotopic. This will follow from the fact that $U_{\alpha_1} \cup U_{\alpha_2} \cup U_{\alpha_3}$ is simply-connected. To see why, note that $(U_{\alpha_1} \cup U_{\alpha_2}) \cap U_{\alpha_3}$ is connected since

$$(U_{\alpha_1} \cup U_{\alpha_2}) \cap U_{\alpha_3} = U_{\alpha_1\alpha_3} \cup U_{\alpha_2\alpha_3}$$

is the union of connected sets whose intersection $U_{\alpha_1\alpha_2\alpha_3}$ is nonempty. Thus by another application of the Seifert-van Kampen theorem, $\pi_1\left((U_{\alpha_1} \cup U_{\alpha_2}) \cup U_{\alpha_3}\right) = \{1\}$. Therefore $\phi_{\alpha\beta}$ defines a cocycle.

To verify that this is an inverse to the previous map, consider any loop γ based at x_0. Cover γ as before by $U_{\alpha_1}, \ldots, U_{\alpha_k}$ with $t_1, \ldots, t_k$ as before. We express $[\gamma]$ in terms of the $\left[\gamma_{\alpha\beta}\right]$ as follows. Without changing the $\{U_{\alpha_i}\}$ or the $\{t_i\}$, it is possible to homotope γ to pass through $x_{\alpha_1}, \ldots, x_{\alpha_k}$. Noting that $\left[\gamma_{\alpha_k\alpha_1}\right] = e$ since $x_0 \in U_{\alpha_1\alpha_k}$, we have $\gamma = \left[\gamma_{\alpha_1\alpha_k}\right]\left[\gamma_{\alpha_k\alpha_{k-1}}\right]\cdots[\gamma_{\alpha_3\alpha_2}]\,[\gamma_{\alpha_2\alpha_1}]$. Thus

$$\begin{aligned}
\mathrm{Hol}([\gamma])^{\alpha_1}_{\alpha_1} &= h\left(\left[\gamma_{\alpha_1\alpha_k}\right]\right) h\left(\left[\gamma_{\alpha_k\alpha_{k-1}}\right]\right)\cdots h\left([\gamma_{\alpha_3\alpha_2}]\right)\cdots h\left([\gamma_{\alpha_2\alpha_1}]\right)\\
&= h\left(\left[\gamma_{\alpha_1\alpha_k}\right]\left[\gamma_{\alpha_k\alpha_{k-1}}\right]\cdots[\gamma_{\alpha_3\alpha_2}]\,[\gamma_{\alpha_2\alpha_1}]\right)\\
&= h([\gamma]),
\end{aligned}$$

so the holonomy reproduces the given homomorphism, inverting the construction as desired. □

4.3 Flat connections

Note that any flat principal G-bundle has a canonical connection given by $\{A_\alpha = 0\}$.

Exercise 30. Verify that $\{A_\alpha = 0\}$ defines a connection on a flat bundle. Show the converse is also true: if $\{A_\alpha = 0\}$ is a connection on a principal G-bundle, then that bundle must be flat.

Definition 31. A connection A is said to be *flat* if $F_A = 0$.

This definition is justified by the following theorem.

Theorem 32. *If $P \to X$ is a smooth principal G-bundle, then a flat connection A determines a reduction of P to a flat principal G-bundle.*

Since two reductions are equivalent iff they are related by a gauge transformation $g \in \mathscr{G}_P = \Gamma(X; \mathrm{Ad}\, P)$, the previous two theorems establish a bijection between

$$\{A \mid A \in \mathscr{A}_P,\ F_A = 0\} / \mathscr{G}_P \longleftrightarrow \mathrm{Hom}(\pi_1(X), G)/\mathrm{conj}.$$

From the theory of reductions, in order to prove Theorem 32, it suffices to prove

Theorem 33. *For any open $V \subset \mathbb{R}^n$, and for any $x \in V$, there exists $U \subset V$ such that $x \in U$ and*

$$0 \to G \to C^\infty(U; G) \to Flat_G(U) \to 0$$

is exact, where $Flat_G(U) = \left\{A \in \Omega^1(U; \mathfrak{g}) \mid dA + \frac{1}{2}[A \wedge A] = 0\right\}$, and the map $C^\infty(U; G) \to Flat_G(U)$ is defined by $g \mapsto -(dg)g^{-1}$.

The only nontrivial part of Theorem 33 is surjectivity of $g \mapsto -(dg)g^{-1}$. A change of trivialization by g then transforms such a connection to the trivial connection $A = 0$.

The case $n = 1$ is already mildly interesting. Note that since two-forms vanish automatically in one dimension, all connections are flat, so $\mathrm{Flat}_G([0,1]) = \Omega^1([0,1]\,; \mathfrak{g})$. In particular, if $A = -\chi(t)\, dt$, then we seek a map $g : [0,1] \to G$ such that

$$\frac{dg}{dt} g^{-1} = \chi(t).$$

Such a g exists by

Theorem 34. *For any Lie group G and smooth map $\chi : [0,1] \to \mathfrak{g}$, there exists a unique function $\mathrm{OE}[\chi] : [0,1] \to G$ such that $\mathrm{OE}[\chi](0) = e$, and*

$$\frac{dg}{dt} = (R_{g(t)})_* \chi(t),$$

where R_g denotes right-multiplication by g, and $(R_g)_* : \mathfrak{g} = T_eG \to T_gG$ is the pushforward map. Furthermore, $\mathrm{OE}[\chi]$ is smooth, and depends smoothly on χ. The general solution is $\mathrm{OE}[\chi]g_0$ for arbitrary $g_0 \in G$.

Thus

$$\frac{d}{dt}\mathrm{OE}[\chi] = \chi(t)\mathrm{OE}[\chi].$$

This construction will allow us to define a holonomy map for a general connection A, not necessarily flat.

First note that for $f : Y \to X$, there is a nicely behaved pullback principal bundle $f^*(P) \to Y$, and corresponding pullback connections $f^*(A)$ such that $f^*(F_A) = F_{f^*(A)}$.

Definition 35. For smooth $\gamma : [0,1] \to X$, define $\mathrm{Hol}(A, \gamma) \in \mathrm{Iso}(P|_{\gamma(0)}, P|_{\gamma(1)})$ by the following procedure. Consider the connection $\gamma^*(A)$ on $\gamma^*(P) \to [0,1]$, which is flat. Pick an arbitrary trivialization ϕ_α of $\gamma^*(P)$ in which $\gamma^*(A)$ is represented by $-\chi(t)\, dt \in \Omega^1([0,1]\,; \mathfrak{g}_{\mathrm{Ad}})$. Then change the trivialization to $\phi_\beta = \phi_\alpha \mathrm{OE}[\chi(t)]$, so that $\gamma^*(A)$ is represented by 0 in the trivialization ϕ_β. The fibers

$\phi_\beta(0)$ and $\phi_\beta(1)$ are identified with $P|_{\gamma(0)}$ and $P|_{\gamma(1)}$ respectively. Define $\mathrm{Hol}(A,\gamma) := \phi_\beta(1)\phi_\beta(0)^{-1} \in \mathrm{Iso}(P|_{\gamma(0)}, P|_{\gamma(1)})$.

This is well-defined since any other trivialization in which $\gamma^*(A) = 0$ is given by $\phi_\beta g_0$ for an arbitrary constant $g_0 \in G$, and $\phi_\beta(1)g_0\left(\phi_\beta(0)g_0\right)^{-1} = \phi_\beta(1)\phi_\beta(0)^{-1}$. It's easy to show that this extends the definition of $\mathrm{Hol}([\gamma])$ for flat bundles via a subdivision argument.

We would like to show that when F_A is flat, then $\mathrm{Hol}(A,\gamma)$ is independent of homotopy of γ. Assuming Theorem 34, the result follows from the case where X is a square.

Lemma 36. *Consider $A = A_1(t,s)\,ds + A_2(t,s)\,dt \in \Omega^1([0,1]\times[0,1];\mathfrak{g})$. There exists a change of trivialization g in which the connection is $A' = A_1'(t,s)\,ds + 0\,dt$ such that*

- $g(0,s) = e$,
- $A_1'(0,s) = A_1(0,s)$,
- $F_{A'} = \frac{\partial}{\partial t}A_1'(t,s)\,dt\wedge ds$.

In particular, if $F_A = 0$, then A' is independent of t and of the form $A_1'(s)\,ds$.

Proof. Consider $g(t,s) = \mathrm{OE}[t_0 \mapsto -A_2(t_0,s)](t)$. It's evident that $g(0,s) = e$. A change of trivialization by g is equivalent to a transition by g^{-1}. The ds component of A' is

$$A_1' = g^{-1}A_1 g - \frac{\partial g^{-1}}{\partial s}g = g^{-1}\frac{\partial g}{\partial s} + g^{-1}A_1 g.$$

Restricted to $s = 0$ we have $g = e$ and $\frac{\partial g}{\partial s} = 0$, and thus $A_1'(t,0) = A_1(t,0)$.

Similarly, the dt component of A' is, applying the differential equation for the ordered exponential,

$$g^{-1}\frac{\partial g}{\partial t} + g^{-1}A_2 g = -g^{-1}A_2 g + g^{-1}A_2 g = 0.$$

It follows that $A' = A_1'(t,s)\,ds$, and consequently

$$F_{A'} = \frac{\partial}{\partial t}A_1'(t,s)\,dt\wedge ds.$$

Since $F_{A'} = g^{-1}F_A g$, it follows that if $F_A = 0$, then $F_{A'} = 0$, and so $A_1(t,s)$ is independent of t. □

Corollary 37. *Suppose $F_A = 0$. Then* $\mathrm{Hol}(A,\gamma)$ *is independent of enpoint-fixing homotopies of γ.*

Proof. Suppose $\gamma_s : [0,1] \to X$ is a smooth homotopy for $s \in [0,1]$ which fixes the endpoints. Then $\gamma^*(A)$ is a connection on $[0,1]\times[0,1]$ which, in a general trivialization ϕ_α over $[0,1]\times[0,1]$ has the form $A_1(t,s)\,ds + A_2(t,s)\,dt$. The fibers of $\gamma^*(P)|_{t=0}$ are identified with the single fiber $P|_{\gamma_s(0)}$, and similarly the fibers of $\gamma^*(P)|_{t=1}$ are identified with the fiber $P|_{\gamma_s(1)}$. It's easy to arrange that ϕ_α be constant when restricted to both $t = 0$ and $t = 1$. Since γ is also constant on this set, this implies that both $A_1|_{t=0}$ and $A_1|_{t=1}$ must vanish. Applying Lemma 36, we can change the trivialization by some

$g = g(t,s)$ so that in $\phi_\beta = \phi_\alpha \cdot g$ with the hypotheses of Lemma 36 satisfied. It follows that $\phi_{\beta\alpha} = g^{-1}$. Let A' denote A transformed to ϕ_β. Then $A' = A_1'(s)\,ds$. Since the dt component of A' vanishes, A' pulls back to $[0,1] \times \{s\}$ as zero, and hence

$$\mathrm{Hol}(\gamma_s) = \phi_\beta(1,s)\phi_\beta(0,s)^{-1}.$$

It remains to show that this is independent of s. Since $g|_{t=0} = e$, it follows that $\phi_\beta|_{t=0} = \phi_\alpha|_{t=0} \cdot e$ which is constant. Since $A_1|_{t=0} = 0$, it follows that $A_1'(s) = A_1'(s,0) = A_1(s,0) = 0$, and thus $A' = 0$. Finally, it follows that $A = -(dg)g^{-1}$. Using the fact that $0 = A_1|_{t=1} = -\frac{\partial g}{\partial s} g^{-1}|_{t=1}$, it follows that $g|_{t=1}$ is constant. Thus $\phi_\beta|_{t=1} = \phi_\alpha g|_{t=1}$ is constant since both $\phi_\alpha|_{t=1}$ and $g|_{t=1}$ are constant. □

Corollary 38. *If $P \to X$ is a principal G-bundle with flat connection A, and if X is simply-connected, then there exists a trivialization $\phi \in \Gamma(X;P)$ such that the connection form $\phi^*(A)$ is $0 \in \Omega^1(X;\mathfrak{g})$.*

Proof. Choose $x_0 \in X$ and $\phi_0 \in P|_{x_0}$. Define

$$\phi(x) := \mathrm{Hol}(A,\gamma)(\phi_0),$$

where γ is any choice of path from x_0 to x. Since A is flat, the result does not depend on the homotopy class of γ. Since X is simply-connected, there is a single homotopy class, and this map is well-defined.

Consider $\phi^*(A)$. Pulling back further to any path $\gamma : [0,1] \to X$, this gives the zero connection form on $[0,1]$. Now A must itself be zero, since otherwise there would be some vector on which A is nonzero, and hence some path representing it which would pull back to something nonzero. □

It remains only to prove Theorem 34.

Lemma 39. *Fix an arbitrary inner product on $\mathfrak{g}$. There exists $\epsilon > 0$ such that if $\chi : [0,1] \to \mathfrak{g}$ is continuous and satisfies $|\chi(t)| < \epsilon$, then there exists a unique continuous $g : [0,1] \to G$ such that $g(0) = e$ and $\frac{dg}{dt} g^{-1} = \chi$, which depends smoothly on χ.*

Remark 40. The notation "$\frac{dg}{dt}$" means $g_*(\partial_t) \in T_{g(t)}G$. Then "$\frac{dg}{dt}g^{-1}$" means $(R_{g^{-1}})_* \left(\frac{dg}{dt}\right) \in T_eG = \mathfrak{g}$.

Proof of Theorem 34 assuming Lemma 39. The idea is to rescale Theorem 34 so that the hypotheses of Lemma 39 are satisfied. For $m, M \in \mathbb{Z}$ with $0 \le m < M$, let $\chi_{m,M} : [0,1] \to \mathfrak{g}$ denote the function $\chi_{m,M}(t) = \chi((t+m)/M)/M$. Assuming that $\mathrm{OE}[\chi]$ exists for all χ, it must satisfy

$$\mathrm{OE}[\chi]((t+m)/M) = (\mathrm{OE}[\chi_{m,M}](t))\,(\mathrm{OE}[\chi_{m-1,M}](1)) \cdots (\mathrm{OE}[\chi_{1,M}](1))\,(\mathrm{OE}[\chi_{0,M}](1)),$$

since both sides satisfy the same differential equation for all m, M, and $t \in [0,1]$, and respect the continuity of $\mathrm{OE}\,[\chi]$. By choosing $M > \max_t |\chi(t)|/\epsilon$, it follows from Lemma 39 that the right hand side exists, is unique, and depends smoothly on χ. Therefore $\mathrm{OE}\,[\chi]$ exists, is unique, and depends smoothly on χ. □

Proof of Lemma 39. The strategy is to apply the Banach fixed-point theorem in some coordinate chart. Fix some number $1 < r < 1 + \frac{1}{10}$. Let $B_r(e)$ denote an open ball around $e \in G$ which is identified with an open ball of radius r in $\mathbb{R}^d$, where $d = \dim G$. Let $V : \mathfrak{g} \times B_r(e) \to \mathbb{R}^d$

denote the components of the right-invariant vector fields. Thus if $\chi \in \mathfrak{g}$ and the right-invariant vector field χ^R has components $(\chi^R)^1, \ldots, (\chi^R)^d$ in $B_r(e)$, then for $x \in B_r(e)$ we have $V(\chi, x) = ((\chi^R)^1(x), \ldots, (\chi^R)^d(x))$. The coordinate version of the desired differential equation is thus

$$\frac{dg}{dt}(t) = V(\chi(t), g(t)), \quad g(0) = 0.$$

Integrating, any solution to this differential equation gives a solution to the integral equation

$$g(t) = \int_0^t V(\chi(t_0), g(t_0))\, dt_0,$$

and is thus a fixed-point of the operator

$$\Gamma(g)(t) := \int_0^t V(\chi(t_0), g(t_0))\, dt_0.$$

Existence and uniqueness of a solution will then follow by showing that Γ is a contraction mapping on the appropriate function space.

For $\lambda > 0$, let $B_\lambda(0) \subset \mathfrak{g}$ denote the open ball $|\chi| < \lambda$. By compactness of $\overline{B_1(0)} \times \overline{B_1(e)}$, it follows that the restriction of V has

$$|V| + |\nabla V| < L$$

for some $L \in \mathbb{R}$. In particular,

$$\begin{aligned} |V(\chi(t), y) - V(\chi(t), x)| &= \left| \int_0^1 \nabla_{y-x} V(\chi(t), (1-s)x + sy) \right| \leq L\, |y - x|, \\ |V(\chi(t), x)| &\leq L\, |x|. \end{aligned}$$

Note that since $V(\lambda\chi, x) = \lambda V(\chi, x)$, it follows that on the restriction of V to $\mathfrak{g} \times \overline{B_1(e)}$, we have

$$\begin{aligned} |V(\chi(t), y) - V(\chi(t), x)| &\leq L\, |\chi(t)|\, |y - x|, \\ |V(\chi(t), x)| &\leq L\, |\chi(t)|\, |x|. \end{aligned}$$

Now take $\epsilon < 1/3L$, so that $|\chi(t)| < 1/3L$, and

$$\begin{aligned} |V(\chi(t), y) - V(\chi(t), x)| &< \tfrac{1}{3}\, |y - x|, \\ |V(\chi(t), x)| &< \tfrac{1}{3}\, |x|. \end{aligned}$$

We use this estimate to construct a fixed point on the following space. Consider the vector space $\mathscr{C}([0,1]\,; \mathbb{R}^d)$ of continuous maps $g : [0,1] \to \mathbb{R}^d$. This has a norm given by $\|g\| := \max_t |g(t)|$. This norm is always finite by continuity and compactness of $[0,1]$. Furthermore, all Cauchy sequences converge, meaning that if $\{g_i\}_{i=1}^\infty$ satisfies $\left\|g_i - g_j\right\| \to 0$ as $\min(i,j) \to \infty$, then $\lim_{i\to\infty} g_i$ exists and is continuous. This gives $\mathscr{C}([0,1]\,; \mathbb{R}^d)$ the structure of a Banach space.

Let $B_{\mathscr{C}}$ denote the unit ball in $\mathscr{C}([0,1]\,; \mathbb{R}^d)$, which is the space of continuous functions from $[0,1]$ to the unit ball in $\mathbb{R}^d$. We will show that the operator Γ defines a contraction mapping $\Gamma : B_{\mathscr{C}} \to B_{\mathscr{C}}$.

In order for $\Gamma(g)$ to be well-defined, it's important that $g \in B_{\mathscr{C}}$ so that $V(\chi(t), g(t))$ is well-defined and satisfies the bounds. Its image is in $B_{\mathscr{C}}$ since

$$|\Gamma(g)(t_0)| \leq \int_0^1 |V(\chi(t), g(t))| \, dt < \tfrac{1}{3}, \text{ so } \|\Gamma(g)\| < \tfrac{1}{3}.$$

Furthermore, Γ is a contraction mapping, meaning that it reduces distance by at least some fixed factor:

$$|\Gamma(g_2)(t_0) - \Gamma(g_1)(t_0)| \leq \int_0^1 |V(\chi(t), g_2(t)) - V(\chi(t), g_1(t))| \, dt < \tfrac{1}{3} \|g_2 - g_1\|,$$

so

$$\|\Gamma(g_2) - \Gamma(g_1)\| < \tfrac{1}{3} \|g_2 - g_1\|.$$

Now we employ the Banach fixed point theorem. Consider now the sequence $\{0, \Gamma(0), \Gamma(\Gamma(0)), \ldots\}$. This is Cauchy since

$$\left\|\Gamma^i(0) - \Gamma^j(0)\right\| \leq (\tfrac{1}{3})^{\min i,j} \left\|\Gamma^{|i-j|}(0) - 0\right\| \leq 2(\tfrac{1}{3})^{\min i,j} \to 0.$$

Thus there exists some continuous limit g_∞. This limit g_∞ is easily verified to be a unique fixed point $\Gamma(g_\infty) = g_\infty$. Thus g_∞ satisfies the integral equation

$$g_\infty(t) = \int_0^t V(\chi(t_0), g_\infty(t_0)) \, dt_0.$$

Differentiating, one finds that

$$\begin{aligned}
\frac{dg_\infty}{dt}(t) &= V(\chi(t), g_\infty(t)), \\
\frac{d^2 g_\infty}{dt^2}(t) &= \frac{\partial V}{\partial \chi}\frac{d\chi}{dt} + \frac{\partial V}{\partial g}\frac{dg_\infty}{dt} \\
&= \frac{\partial V}{\partial \chi}\frac{d\chi}{dt} + \frac{\partial V}{\partial g} V(\chi(t), g_\infty(t)).
\end{aligned}$$

Continuing to differentiate, we observe that the higher-order derivatives of g_∞ can be expressed in terms of the higher-order derivatives of V and χ. Thus if V and χ are smooth, then so is g_∞, and g_∞ depends smoothly on χ. □

4.4 Matrix groups

All compact groups G can be realized as a matrix group. In particular, for any G, there exists some integer N and a faithful (injective) representation $\rho : G \to \mathrm{GL}(N; \mathbb{R})$. There is the associated vector bundle $E := P \times_\rho \mathbb{R}^N$ which comes equipped with a natural embedding

$$P \subset \mathrm{Fr}(E),$$

given by $\phi \mapsto ([\phi, e_1], \ldots, [\phi, e_N])$.

From ρ there is an associated representation $\rho_{\mathfrak{gl}} : G \to \mathrm{GL}(N \times N; \mathbb{R})$ which acts on an $N \times N$ matrix M according to

$$\begin{aligned}\rho_{\mathfrak{gl}}(g) \cdot M &= \rho(g) M \rho(g)^{-1} \\ &= g M g^{-1},\end{aligned}$$

where in the last line we have left ρ implicit.

The associated vector bundle is $\mathfrak{gl}(E) = P \times_{\rho_{\mathfrak{gl}}} \mathbb{R}^{N \times N}$, whose fibers are endomorphisms of the fibers of E. There is a fiber subbundle $\mathrm{GL}(E) \subset \mathfrak{gl}(E)$ consisting of the invertible endomorphisms. Finally there is $G_{\mathrm{Ad}} \subset \mathrm{GL}(E)$, which describes the automorphisms which preserve P in each fiber, which is fiberwise a copy of G. The fiberwise Lie algebra is thus naturally a subbundle of $\mathfrak{gl}(E)$. Thus

$$\begin{aligned}G_{\mathrm{Ad}} \subset \mathrm{GL}(E) &\subset \mathfrak{gl}(E), \\ \mathfrak{g}_{\mathrm{Ad}} &\subset \mathfrak{gl}(E).\end{aligned}$$

The infinitesimal version of $\rho_{\mathfrak{gl}} : \mathfrak{g} \to \mathrm{GL}(N \times N; \mathbb{R})$ is given by

$$\begin{aligned}\rho_{\mathfrak{gl}}(\chi) \cdot M &= \rho(\chi) M - M \rho(\chi) \\ &= \chi M - M \chi,\end{aligned}$$

where again the last line leaves ρ implicit.

4.5 Gauge transformations and stabilizers

Suppose $P \to X$ is a principal G-bundle, where G is compact. Automorphisms of P are given by elements of $\mathscr{G}_P := \Gamma(X; G_{\mathrm{Ad}})$. There is a natural induced action

$$\nabla_{g \cdot A} := g \nabla_A g^{-1}.$$

We wish to understand the space $\mathscr{B}_P := \mathscr{A}_P / \mathscr{G}_P$, the space of orbits. The first thing to understand is a single orbit $\mathscr{G}_P \cdot A$. Next we understand the neighborhood of a single orbit. Finally, we will survey the global topology of $\mathscr{B}_P$.

The key to understanding a single orbit is

Definition 41. For any connection $A \in \mathscr{A}_P$, its *stabilizer* $\mathrm{Stab}(A) \subset \mathscr{G}_P$ is $\{g \in \mathscr{G}_P \mid g \cdot A = A\}$.

The structure of a single orbit $\mathscr{G}_P \cdot A$ is determined by $\mathscr{G}_P \cdot A = \mathscr{G}_P / \mathrm{Stab}_A$. Thus we wish to understand solutions to $g \cdot A = A$. For this, we need to find a formula for $g \cdot A$.

Viewing g as a G_{Ad}-valued section of $\Gamma(X; \mathfrak{gl}(E))$, and leaving ρ implicit, we compute in U_α that

$$\begin{aligned}\left(g \nabla_A g^{-1}\right)_\alpha &= g_\alpha (\nabla + A_\alpha) g_\alpha^{-1} \\ &= \nabla + g_\alpha A_\alpha g_\alpha^{-1} - (\nabla g_\alpha) g_\alpha^{-1}.\end{aligned}$$

Thus

$$(g \cdot A)_\alpha = g_\alpha A_\alpha g_\alpha^{-1} - (\nabla g_\alpha) g_\alpha^{-1}, \tag{4.2}$$

which is the same formula as a change of trivialization of A by g_α^{-1}.

It's possible to get a more invariant formula without passing to a local trivialization:

$$g \nabla_A g^{-1} = \nabla_A - (\nabla_A g) g^{-1}.$$

Thus

$$g \cdot A = A - (\nabla_A g) g^{-1},$$

and so $g \cdot A = A$ iff

$$0 = g \cdot A - A = -(\nabla_A g) g^{-1} \in \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}).$$

To reconcile this with (4.2), note that

$$(\nabla_A g)_\alpha = \nabla g_\alpha + \rho_{\mathfrak{gl}}(A_\alpha) g_\alpha = \nabla g_\alpha + A_\alpha g_\alpha - g_\alpha A_\alpha,$$

and thus

$$(-(\nabla_A g) g^{-1})_\alpha = -A_\alpha + g_\alpha A_\alpha g_\alpha^{-1} - (\nabla g_\alpha) g_\alpha^{-1}.$$

Remark 42. Identifying $\mathscr{G}_P$ with the corresponding G_{Ad}-valued sections of $\mathfrak{gl}(E)$,

$$\mathrm{Stab}(A) = \{g \in \mathscr{G}_P \mid \nabla_A g = 0\}.$$

If g is a solution to $\nabla_A g = 0$, then it is connected to holonomy. If γ is any path from x_0 to x_1, then in the holonomy trivialization $\phi : [0,1] \to \gamma^*(P)$, we have $\phi^*(A) = 0$, and

$$g(\gamma(1)) = \mathrm{Hol}(A, \gamma) g(\gamma(0)) \mathrm{Hol}(A, \gamma)^{-1}. \tag{4.3}$$

Definition 43. For a given basepoint $x_0 \in X$, the *holonomy subgroup* $\mathrm{Hol}(A)_{x_0} \subset G_{\mathrm{Ad}}|_{x_0}$ is the image of $\mathrm{Hol}(A, \gamma)$ over all loops γ based at x_0.

Definition 44. Given a subset of a group $S \subset G$, the *centralizer* of S is denoted

$$Z_G(S) := \left\{ g \in G \mid g = sgs^{-1} \; \forall s \in S \right\}.$$

Note that $Z_G(G)$ is the center $Z(G)$.

Theorem 45. *The stabilizer of any $A \in \mathscr{A}_P$ is the centralizer of the holonomy subgroup. More precisely, if X is connected, then for any $x_0 \in X$, then there is an isomorphism*

$$\mathrm{Stab}(A) \to Z_{G_{\mathrm{Ad}}|_{x_0}}(\mathrm{Hol}(A)_{x_0})$$

given by $g \mapsto g(x_0) \in G_{\mathrm{Ad}}|_{x_0}$.

Proof. Suppose $g \in \mathrm{Stab}(A)$. First we must show that the restriction $g(x_0) \in Z_{G_{\mathrm{Ad}}|_{x_0}}(\mathrm{Hol}(A)_{x_0})$. This follows directly from applying (4.3) to any loop based at x_0. Thus $g(x_0)$ is fixed by $\rho_{\mathrm{Ad}}(\mathrm{Hol}(A, \gamma))$, so $g(x_0) \in Z_{G_{\mathrm{Ad}}|_{x_0}}(\mathrm{Hol}(A)_{x_0})$. Conversely, if $g(x_0) \in Z_{G_{\mathrm{Ad}}|_{x_0}}(\mathrm{Hol}(A)_{x_0})$, then

$$g(x) := \rho_{\mathrm{Ad}} \left(\mathrm{Hol}(A, \gamma) \right) g(x_0)$$

for any choice of γ from x_0 to x is a well-defined unique solution to $\nabla_A g = 0$. ◻

Thus we see that $\mathrm{Stab}(A) \subset \mathscr{G}_P$ is isomorphic to a finite-dimensional subgroup of $G_{\mathrm{Ad}}|_{x_0}$. Moreover, this subgroup is a centralizer. This is a very strong constraint.

Theorem 46. *If $H \subset G$ is a centralizer subgroup, and Z denotes the center of G, then $Z \subset H$.*

Theorem 47. *The only subgroups of* $\mathrm{SU}(2)$ *which arise as centralizers are isomorphic to* $\mathbb{Z}_2$, $\mathrm{U}(1)$, *or* $\mathrm{SU}(2)$.

Proof. Centralizer subgroups are the intersection of the centralizers of single elements. The centralizer of $\pm\mathrm{Id}_{2\times 2}$ is $\mathrm{SU}(2)$. for any $\theta_0 \in (0,\pi)$,

$$Z_{\mathrm{SU}(2)}\begin{pmatrix} e^{i\theta_0} & \\ & e^{-i\theta_0}\end{pmatrix} = \left\{\begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta}\end{pmatrix} \mid \theta \in \mathbb{R}\right\} = \mathrm{U}(1).$$

Thus the centralizers in $\mathrm{SU}(2)$ are intersections of $\mathrm{U}(1)$ subgroups. But the intersection of two or more distinct $\mathrm{U}(1)$ subgroups is $\mathbb{Z}_2$. □

Definition 48. If $\mathrm{Stab}(A) = Z(G)$, then A is said to be *irreducible*. Otherwise, A is said to be *reducible*.

Example 49. If $G = \mathrm{SU}(2)$ and A_0 is a trivial connection $\nabla_{A_0} = \nabla + 0$ on the trivial bundle, then $\mathrm{Hol}(\gamma)_{x_0} = \{e\}$, and $\mathrm{Stab}(A) = \mathrm{SU}(2)$ consists of all constant gauge transformations.

Lemma 50. *If $S \subset T$ then $Z_G(T) \subset Z_G(S)$. Also, $S \subset Z_G(T) \iff T \subset Z_G(S)$.*

Corollary 51. *$S \subset Z_G(Z_G(S))$, and $Z_G(S) = Z_G(Z_G(Z_G(S)))$.*

Thus Z_G is an involution on centralizers.

For example, when $G = \mathrm{SU}(2)$, the centralizers are

$$\mathbb{Z}_2 \subset \mathrm{U}(1) \subset \mathrm{SU}(2),$$

and Z_G exchanges $\mathbb{Z}_2$ and $\mathrm{SU}(2)$, but fixes $\mathrm{U}(1)$. Note that there is actually an $\mathbb{RP}^2$ worth of conjugate $\mathrm{U}(1)$ subgroups, each of which are fixed by Z_G.

Corollary 52. *If A is reducible then* $\mathrm{Hol}(A)|_{x_0} \neq G$.

Proof. Suppose $\mathrm{Hol}(A)|_{x_0} = G$. Then $\mathrm{Stab}(A) = Z_G(\mathrm{Hol}(A)|_{x_0}) = Z_G(G)$ which is the center $Z(G)$. This implies that A is irreducible. □

Remark 53. It can be that A is irreducible but $\mathrm{Hol}(A)|_{x_0} \neq G$. For instance, if $G = \mathrm{SU}(2)$ and $\mathrm{Hol}(A)|_{x_0}$ is any non-abelian proper subgroup, then $\mathrm{Hol}(A)|_{x_0} \not\subset \mathrm{U}(1)$, so $\mathrm{Stab}(A) \not\supset Z_G(\mathrm{U}(1)) = \mathrm{U}(1)$, so $\mathrm{Stab}(A) = \mathbb{Z}_2$.

If $\mathrm{Hol}(A)|_{x_0} \neq G$, then holonomy defines a natural space of reductions to $H = \mathrm{Hol}(A)|_{x_0}$ parameterized by $G_{\mathrm{Ad}}|_{x_0}/H$, and A descends to this reduction.

Remark 54. Suppose G is simple and A is reducible with $\mathrm{Stab}(A) = G$, Then $\mathrm{Hol}(A)|_{x_0} \subset Z_G(\mathrm{Stab}(A)) = Z(G)$, so A comes from a connection on a $Z(G)$-bundle. Since the center is discrete, $\mathrm{Lie}(Z(G)) = 0$, and thus A must be flat.

Remark 55. Suppose A is a connection in a principal SU(2) bundle, and $\mathrm{Stab}(A) \neq \mathbb{Z}_2$. Then by the classification of centralizers, $\mathrm{Stab}(A) = \mathrm{U}(1)$ or $\mathrm{Stab}(A) = \mathrm{SU}(2)$. In either case, $\mathrm{Stab}(A) \supset \mathrm{U}(1)$, so

$$\mathrm{Hol}(A)|_{x_0} \subset Z_G(Z_G(\mathrm{Hol}(A)|_{x_0})) = Z_G(\mathrm{Stab}(A)) \subset Z_G(\mathrm{U}(1)) = \mathrm{U}(1).$$

If $E = P \times_{\rho_{\mathrm{st}}} \mathbb{C}^2$, choose a frame $\{e_1, e_2\}$ for $E|_{x_0}$ in which holonomy is of the form

$$\mathrm{Hol}(A)|_{x_0} \subset \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \mid \theta \in \mathbb{R} \right\}.$$

Consider the parallel transport of e_1. Thanks to the holonomy, parallel transport along any loop takes e_1 back to a multiple of itself. Thus parallel transport spans a subbundle L of E. Similarly, parallel transport of e_2 spans a complementary subbundle L'. Thus $E = L \oplus L'$. Being an SU(2) bundle, the determinant bundle $\Lambda^2 E$ must be canonically trivial, so $\Lambda^2 E \cong \mathbb{C}$. (Equivalently, the fibers of E are equipped with a skew-symmetric bilinear form.) If $E = L \oplus L'$, then $\Lambda^2 E \cong L \otimes L'$, so $L' \cong L^{-1}$.

Suppose abstractly that some group $\mathscr{G}$ acts smoothly on some space $\mathscr{A}$, and we wish to understand the orbit space $\mathscr{B} := \mathscr{A}/\mathscr{G}$. Of course the context will be $\mathscr{G} = \mathscr{G}_P$ for some principal bundle P, and $\mathscr{A} = \mathscr{A}_P$, but for now, let's think of these as finite-dimensional spaces. Given some $A_0 \in \mathscr{A}$ with some stabilizer, we wish to understand a neighborhood of $[A_0] \in \mathscr{B}$. The idea is to construct in some manner a small "slice" $\mathcal{S}_{A_0} \subset \mathscr{A}$ which transversely intersects the orbit $\mathscr{O}_{A_0} := \mathscr{G} \cdot A_0$ through the point A_0. In particular, we want

$$T_{A_0}\mathscr{A} = T_{A_0}\mathscr{O}_{A_0} \oplus T_{A_0}\mathcal{S}_{A_0}.$$

Assume that the action of $\mathscr{G}$ on $\mathscr{A}$ restricts to an action of $\mathrm{Stab}(A_0)$ on $\mathcal{S}_{A_0}$. The desired description of a neighborhood of $[A_0]$ in $\mathscr{B}$ is given by the quotient $\mathcal{S}_{A_0}/\mathrm{Stab}(A_0)$. This follows directly from the key lemma that

$$\begin{aligned} \mathfrak{m} : (\mathscr{G} \times \mathcal{S}_{A_0})/\mathrm{Stab}(A_0) &\to \mathscr{A}, \\ \mathfrak{m}(g, A) &:= g \cdot A \end{aligned}$$

is a $\mathscr{G}$-equivariant diffeomorphism onto its image.

To see why this should be a local diffeomorphism, consider the linearization of $\mathfrak{m}$ at (e, A_0). If this linearization is an isomorphism, then it follows by the inverse function theorem that $\mathfrak{m}$ is a local diffeomorphism in a neighborhood of (e, A_0). It's important that the slice $\mathcal{S}_{A_0}$ be chosen to be small enough so that the inverse function theorem will apply to a much larger region. The linearization is

$$\begin{aligned} d_{(e,A_0)}\mathfrak{m} : \left(T_e\mathscr{G}_P \oplus T_{A_0}\mathcal{S}_{A_0}\right)/T_e\mathrm{Stab}(A_0) &\to T_{A_0}\mathscr{A}_P \\ d_{(e,A_0)}\mathfrak{m}(\chi, a) &= \chi \cdot A_0 + a. \end{aligned}$$

Recall now that the restriction of the action to $\mathscr{G} \times \{A_0\}$ gives a diffeomorphism from $\mathscr{G}/\mathrm{Stab}(A_0)$ to $\mathscr{O}_{A_0}$. This induces an isomorphism

$$T_e\mathscr{G}/T_e\mathrm{Stab}(A_0) \to T_{A_0}\mathscr{O}_{A_0}.$$

Thus

$$\mathrm{Im}(d_{(e,A_0)}\mathfrak{m}) = T_{A_0}\mathscr{O}_{A_0} + T_{A_0}\mathcal{S}_{A_0}.$$

Under our assumption about the slice,

$$d_{(e,A_0)}\mathfrak{m} : T_{A_0}\mathcal{O}_{A_0} \oplus T_{A_0}\mathcal{S}_{A_0} \to T_{A_0}\mathcal{A}$$

is an isomorphism. Indeed, there is some neighborhood $U \subset \mathcal{G}$ of e such that the neighborhood of (e, A_0) can be chosen to be a product neighborhood $U' \times \mathcal{S}_{A_0}$, where $U' = U \cdot \mathrm{Stab}(A_0)$. By $\mathcal{G}$-equivariance, it follows that $\mathfrak{m}$ is a local diffeomorphism from $((U' \cdot g) \times \mathcal{S}_{A_0}) / \mathrm{Stab}(A_0)$ to a neighborhood of $g \cdot A_0$ for all $g \in \mathcal{G}$.

We wish to show that $\mathfrak{m}$ is not just a local diffeomorphism, but an actual diffeomorphism onto its image. Thus we need to show that it's injective. This requires choosing $\mathcal{S}_{A_0}$ to be small enough such that $\mathcal{S}_{A_0}$ and $g \cdot \mathcal{S}_{A_0}$ intersect only when $g \in \mathrm{Stab}(A_0)$.

If it's not possible to choose $\mathcal{S}_{A_0}$ small enough, then there exist sequences $\{g_i \in \mathcal{G} \backslash U'\}$ and $\{A_i \in \mathcal{S}_{A_0}\}$ such that $g_i \cdot A_i \in \mathcal{S}_{A_0}$, and both $A_i \to A_0$, and $g_i \cdot A_i \to A_0$. Assuming that it's possible to pass to some convergent subsequence, then $g_i \to g$ with $g \cdot A_0 = A_0$, thus $g \in \mathrm{Stab}(A_0)$. But $\mathrm{Stab}(A_0)$ is in the interior of U', which is a contradiction.

Now we consider the case $\mathcal{A} = \mathcal{A}_P$ and $\mathcal{G} = \mathcal{G}_P$. Given $A_0 \in \mathcal{A}_P$, we wish to understand the tangent space to the orbit $\mathcal{O}_{A_0}$. This is determined by the map $\mathcal{G}_P \to \mathcal{A}_P$ given by $g \mapsto g \cdot A_0 = A_0 - d_A g$. Since $\mathcal{G}_P = \Omega^0(X; G_{\mathrm{Ad}})$, the Lie algebra is $\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})$, and the linearization $T_e\mathcal{G}_P \to T_{A_0}\mathcal{A}_P$, or equivalently $\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}}) \to \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$ is given by

$$\chi \mapsto -d_{A_0}\chi.$$

Thus

$$T_{A_0}\mathcal{O}_{A_0} = \mathrm{Im}(d_{A_0} : \Omega^0(X; \mathfrak{g}_{\mathrm{Ad}}) \to \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})).$$

We seek a complementary subspace

$$\begin{aligned} T_{A_0}\mathcal{A} &= T_{A_0}\mathcal{O}_{A_0} \oplus T_{A_0}\mathcal{S}_{A_0}, \\ \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) &= \mathrm{Im}(d_{A_0}) \oplus C. \end{aligned}$$

The easiest way to construct such a subspace is to choose a Riemannian metric on X and an invariant metric on $\mathfrak{g}$ to define an inner product on $\Omega^p(X; \mathfrak{g}_{\mathrm{Ad}})$ by $\langle\langle \alpha, \beta \rangle\rangle := \int_X \langle \alpha \cdot \beta \rangle_{\mathfrak{g}}\, d\mathrm{vol}$, and define

$$C := \mathrm{Im}(d_{A_0})^\perp.$$

There is a good way to characterize C as follows. If $\alpha \in C$, then for all $\beta \in \Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})$,

$$0 = \langle\langle \alpha, d_{A_0}\beta \rangle\rangle = \left\langle\left\langle d_{A_0}^*\alpha, \beta \right\rangle\right\rangle,$$

where $d_{A_0}^* : \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) \to \Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})$ is the formal adjoint. Unlike d_{A_0}, the operator $d_{A_0}^*$ depends on the choice of Riemannian metric. It is determined by integration by parts. The only way for $0 = \left\langle\left\langle d_{A_0}^*\alpha, \beta \right\rangle\right\rangle$ for all β is if $d_{A_0}^*\alpha = 0$. Thus $\mathrm{Im}(d_{A_0})^\perp = \ker d_{A_0}^*$. Thus

$$\mathrm{Im}(d_{A_0}) \oplus \ker(d_{A_0}^*) \subset \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}).$$

It's not at all obvious whether or not this is an equality. There is no simple procedure for decomposing a general element of $\Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$ as such a sum.

More generally, given a differential operator $D : \Gamma(E) \to \Gamma(F)$, and its adjoint $D^* : \Gamma(F) \to \Gamma(E)$, we wish to know when it is possible to write

$$\Gamma(F) = \mathrm{Im}(D) \oplus \ker D^*.$$

This is certainly not always possible. Consider the zeroth order differential operator $D : \Omega^0(S^1) \to \Omega^0(S^1)$ given by multiplication by $(Df)(\theta) := f(\theta)\sin\theta$. It's clear that $D = D^*$, and $\ker D^* = \{0\}$ is the space of all smooth functions on S^1 which vanish except when $\theta = 0$ or $\theta = \pi$. However, $\mathrm{Im}(D)$ is the subspace of functions which vanish at $\theta = 0$ and $\theta = \pi$. In particular, the constant function $f(\theta) = 1$ is missing from $\mathrm{Im}(D) \oplus \ker D^*$.

There is a class of operators called elliptic. D is elliptic iff D^* is elliptic. If D is elliptic, then $\ker D$ is finite-dimensional. In this case, there is a finite Gram-Schmidt process which writes $\Gamma(F) = (\ker D^*)^\perp \oplus \ker D^*$.

This situation is reminiscent of the Hodge decomposition

$$\begin{aligned}
\Omega^p(X) &= \mathrm{Im}(d : \Omega^{p-1}(X) \to \Omega^p(X)) \oplus \ker(d^* \Omega^p(X) \to \Omega^{p-1}(X)) \\
&= \mathrm{Im}(d^* : \Omega^{p+1}(X) \to \Omega^p(X)) \oplus \ker(d : \Omega^p(X) \to \Omega^{p+1}(X)) \\
&= \mathrm{Im}(d : \Omega^{p-1}(X) \to \Omega^p(X)) \oplus \mathrm{Im}(d^* : \Omega^{p+1}(X) \to \Omega^p(X)) \oplus \\
&\qquad \oplus \ker(d + d^* : \Omega^p(X) \to \Omega^{p+1}(X) \oplus \Omega^{p-1}(X)).
\end{aligned}$$

Consider a principal U(1) bundle P. Since U(1) is abelian, the adjoint action is trivial, and $G_{\mathrm{Ad}} = X \times \mathrm{U}(1)$, and $\mathfrak{g}_{\mathrm{Ad}} = X \times \sqrt{-1}\,\mathbb{R}$. Given any $A_0 \in \mathscr{A}_P$, we have $\mathscr{A}_P = A_0 + \sqrt{-1}\,\Omega^0(X)$.

Chapter 5

Hodge decomposition

5.1 Hodge star

Let X be a closed oriented Riemannian n-manifold X. Consider the de Rham cohomology $H^k(X;\mathbb{R})$ defined by

$$\frac{\ker d \subset \Omega^k(X)}{\text{image } d}.$$

We wish to find a natural subspace $\mathscr{H}^k \subset \Omega^k(X)$ such that $\mathscr{H}^k \cong H^k(X;\mathbb{R})$ via the map $\omega \mapsto [\omega]$. In other words, we want to trade our quotient space $H^k(X;\mathbb{R})$ for a subspace $\mathscr{H}^k$.

If V is a finite-dimensional Euclidean vector space, and if $W \subset V$ is a subspace, then we can naturally represent the quotient V/W by $W^\perp$. Specifically, each coset in V/W intersects a unique vector in $W^\perp$, so we get an isomorphism $W^\perp \to V/W$ by $v \mapsto [v]$. Of course $W^\perp$ is not the only subspace with this property.

Definition. A subspace $S \subset V$ is called a *slice* for the quotient V/W if the quotient map restricts to S as an isomorphism.

The idea of the Hodge decomposition is simply to imitate this construction in the infinite dimensional setting of de Rham theory.

The first ingredient we need is an inner product on $\Omega^p(X)$. For this, consider $\mathbb{R}^n$ equipped with the standard $\mathrm{SO}(n)$ structure, i.e. the standard Euclidean metric and orientation, so that $\left\{e^1,\ldots,e^n\right\}$ is an orthonormal basis. We define a Euclidean metric on $\Lambda^p\mathbb{R}^n$ by declaring $e^{i_1}\wedge\cdots\wedge e^{i_p}$ to be an orthonormal basis. More invariantly, one can define

$$\left\langle v^1\wedge\cdots\wedge v^p, w^1\wedge\cdots\wedge w^p\right\rangle := \det\left\langle v^i, w^j\right\rangle,$$

and the right hand side is clearly invariant under $\mathrm{O}(n)$. (The action is $v_1\wedge\cdots\wedge v_p \mapsto gv_1\wedge\cdots\wedge gv_p$.)

We can define a map on the exterior powers of $\mathbb{R}^n$ by $\star : \Lambda^p\mathbb{R}^n \to \Lambda^{n-p}\mathbb{R}^n$ characterized by the relation

$$\alpha\wedge\star\beta = \langle\alpha,\beta\rangle\; e^1\wedge\cdots\wedge e^n.$$

This characterization is clearly invariant under $\mathrm{SO}(n)$, and one computes that

$$\star\left(e^{i_1}\wedge\cdots\wedge e^{i_k}\right)=\pm e^{\tilde{i}_1}\wedge\cdots\wedge e^{\tilde{i}_{n-p}},$$

where $\tilde{i}$ denotes the indices complementary to i, and $\pm$ is determined by

$$e^{i_1}\wedge\cdots\wedge e^{i_k}\wedge e^{\tilde{i}_1}\wedge\cdots\wedge e^{\tilde{i}_{n-p}}=\pm e^1\wedge\cdots\wedge e^n.$$

One verifies that

$$\star^2=(-1)^{p(n-p)}:\Lambda^p\mathbb{R}^n\to\Lambda^p\mathbb{R}^n,$$

Furthermore, $\star$ encodes the orientation and metric via the identities

$$\begin{aligned}\star 1&=e^1\wedge\cdots\wedge e^n,\\ \langle\alpha,\beta\rangle&=\star(\alpha\wedge\star\beta)=\star(\beta\wedge\star\alpha).\end{aligned}$$

The Hodge star map is equivariant under $\mathrm{SO}(n)$, i.e. $\star(g\alpha)=g(\star\alpha)$. For any vector space V equipped with a reduction to $\mathrm{SO}(n)$, i.e. V is equipped with an orientation and a Euclidean metric, the Hodge star determines a map $\star:\Lambda^pV\to\Lambda^{n-p}V$, where

$$\Lambda^pV:=\mathrm{Fr}^{\mathrm{SO}}(V)\times_\rho\Lambda^p\mathbb{R}^n.$$

Of course this also makes sense for any principal $\mathrm{SO}(n)$ bundle. Suppose X is a smooth n-manifold equipped with a reduction of the cotangent bundle T^*X to a $\mathrm{SO}(n)$ structure, i.e. X is oriented Riemannian. (A Riemannian metric determines an isomorphism $TX\to T^*X$, so reductions of T^*X or TX are equivalent.) In particular, $\star$ induces a bundle map $\Lambda^pT^*X\to\Lambda^{n-p}T^*X$. Differential forms are sections $\Omega^p(X)=\Gamma(\Lambda^pT^*X)$, so we get a map $\star:\Omega^p(X)\to\Omega^{n-p}(X)$ which acts fiberwise.

Finally, we define a Euclidean inner product on $\Omega^p_c(X)$ (p-forms with compact supports) by

$$\alpha\cdot\beta:=\int_X\alpha\wedge\star\beta.$$

Define

$$\begin{aligned}d^*&:\Omega^p(X)\to\Omega^{p-1}(X),\\ d^*\alpha&:=(-1)^{n(p+1)+1}\star d\star\alpha.\end{aligned}$$

This satisfies $(d^*)^2=0$ as a consequence of $d^2=0$ and $\star^2=\pm1$.

Theorem. *The operator d^* is the formal metric adjoint of d, i.e. up to a boundary term,*

$$\langle d\alpha,\beta\rangle=\langle\alpha,d^*\beta\rangle.$$

Proof. Suppose $\alpha\in\Omega^{p-1}(X)$ and $\beta\in\Omega^p(X)$. Then

$$d(\alpha\wedge\star\beta)=d\alpha\wedge\star\beta+(-1)^{p-1}\alpha\wedge d\star\beta=d\alpha\wedge\star\beta-\alpha\wedge\star d^*\beta.$$

Integrating, we obtain

$$\int_{\partial X}\alpha\wedge\star\beta=\langle d\alpha,\beta\rangle-\langle\alpha,d^*\beta\rangle.$$

Whenever the boundary term vanishes (e.g. if X is closed), we have $\langle d\alpha,\beta\rangle=\langle\alpha,d^*\beta\rangle$. □

Now let's return to representing cohomology classes on a closed manifold. To find a slice for $\ker d/\text{image } d$, we want to consider

$$(\text{image } d)^\perp \subset \ker d \subset \Omega^p(X).$$

We have that

$$\alpha \in (\text{image } d)^\perp \iff \forall \beta,\ 0 = \langle \alpha, d\beta \rangle = \langle d^*\alpha, \beta \rangle \iff \alpha \in \ker d^*.$$

Since we want to look at the kernel of d^* inside the kernel of d, we are led to study

$$\mathscr{H}^p(X) := \ker d \cap \ker d^* \subset \Omega^p(X).$$

There are a few alternative characterizations. Note that $\mathscr{H}^p(X) = \ker(d \oplus d^*) = \ker(d + d^*)$, since $d + d^* : \Omega^p \to \Omega^{p+1} \oplus \Omega^{p-1}$. For the other characterization, define the Hodge Laplacian

$$\Delta := (d + d^*)^2 = \cancel{d^2} + dd^* + d^*d + \cancel{(d^*)^2}.$$

Over $\mathbb{R}^n$ with the standard metric, Δ on $\Omega^0(X)$ is given by

$$\Delta = -\sum_{i=1}^{n} \left(\frac{\partial}{\partial x^i} \right)^2.$$

(The minus sign is the geometer's convention, which makes Δ act positively on $e^{i\xi\cdot x} \mapsto \xi^2 e^{i\xi\cdot x}$.)

Note that Δ is formally self-adjoint, since up to boundary terms,

$$\langle \alpha, \Delta\beta \rangle = \langle \alpha, (d + d^*)(d + d^*)\beta \rangle = \langle (d^* + d)\alpha, (d + d^*)\beta \rangle = \langle \Delta\alpha, \beta \rangle.$$

Clearly

$$\alpha \in \mathscr{H}^p(X) \implies (d + d^*)\alpha = 0 \implies \Delta\alpha = 0.$$

But conversely,

$$\Delta\alpha = 0 \implies \langle \alpha, \Delta\alpha \rangle = 0 \implies \langle (d + d^*)\alpha, (d + d^*)\alpha \rangle = 0 \implies (d + d^*)\alpha = 0 \implies \alpha \in \mathscr{H}^p(X).$$

For this reason, $\mathscr{H}^p(X)$ is called the *space of harmonic p-forms*.

We wish to form the decomposition

$$\begin{aligned} \Omega^p(X) &= (\text{image } \Delta) \oplus (\text{image } \Delta)^\perp \\ &= (\text{image } \Delta) \oplus (\ker \Delta^*) \\ &= (\text{image } \Delta) \oplus \mathscr{H}^p(X). \end{aligned} \tag{5.1}$$

We can further decompose image $\Delta \subset (\text{image } d) + (\text{image } d^*)$. It is a simple exercise to verify that the images of d and d^* are orthogonal, so we obtain the orthogonal decomposition

$$\Omega^p(X) = d\Omega^{p-1}(X) \oplus d^*\Omega^{p+1}(X) \oplus \mathscr{H}^p(X)$$

known as the Hodge decomposition.

Unfortunately the previous argument was not rigorous, and requires substantial effort. What's wrong with it?

The problem is the first line of (5.1). Consider the operator $m_x : C^\infty([-1,1]) \to C^\infty([-1,1])$ given by $(m_x f)(x) := xf(x)$. The image $\mathrm{im}(m_x)$ is the set of smooth functions which vanish at $x = 0$. The orthogonal complement $\mathrm{im}(m_x)^\perp$ is the set of functions g such that $\int_{-1}^{1} x\, f(x)\, g(x)\, dx = 0$ for all $f \in C^\infty([-1,1])$. This implies that $x\, g(x) = 0$ for all x. By continuity, this means that $g(x) = 0$ for all x. Thus $\mathrm{im}(m_x)^\perp = \{0\}$, and

$$\mathrm{im}(m_x) \oplus \mathrm{im}(m_x)^\perp \subsetneq C^\infty([-1,1]).$$

We must rule out the possibility that $(\text{image } \Delta) \oplus (\text{image } \Delta)^\perp \subsetneq \Omega^p(X)$.

5.2 Hodge decomposition for elliptic operators

Let $E \to X$ and $F \to X$ be vector bundles over a closed Riemannian manifold X. Consider a linear differential operator $D : \Gamma(E) \to \Gamma(F)$. Whenever E and F are equipped with inner products, there is another linear differential operator $D^* : \Gamma(F) \to \Gamma(E)$ called the formal adjoint. It satisfies $\langle\langle Ds_1, s_2\rangle\rangle = \langle\langle s_1, D^* s_2\rangle\rangle$, where $s_1, s_3 \in \Gamma(E), s_2 \in \Gamma(F)$, and $\langle\langle s_1, s_3\rangle\rangle := \int \langle s_1, s_2\rangle \; d\mathrm{vol}$. Also, $D^{**} = D$, and $\ker D^* = (\mathrm{im} D)^\perp$. We wish to show that

$$\begin{aligned}\Gamma(F) &= \mathrm{im}(D) \oplus \ker(D^*),\\ \Gamma(E) &= \mathrm{im}(D^*) \oplus \ker(D^*).\end{aligned}$$

This will not hold for all D, but it holds for a class of operators D called elliptic. Before giving the definition of an elliptic operator, we note that if D is elliptic, then $\ker D$ is finite-dimensional, and D^* is also elliptic. Thus the above two decompositions are equivalent under $D \mapsto D^*$, and the kernels have finite dimension.

The problem we wish to solve is as follows. Suppose D is an elliptic operator. Given any $s \in \Gamma(F)$, we wish to write $s = s_1 + Ds_0$ with $D^* s_1 = 0$. The procedure is as follows. Since $\ker(D^*)$ is finite-dimensional, a finite Gram-Schmidt process writes $s = s_1 + s_2$, with $D^* s_1 = 0$ and $s_2 \perp \ker(D^*)$. Then it suffices to show that for any $s_2 \perp \ker(D^*)$, there exists a solution to the equation $Ds_0 = s_2$.

The decomposition

$$\Omega^p(X) = d\Omega^{p-1}(X) \oplus d^*\Omega^{p+1}(X) \oplus \mathscr{H}^p(X)$$

is then a corollary of the simple facts that

- Δ is elliptic,
- $\Delta = \Delta^*$,
- $\mathscr{H}^p(X) := \ker \Delta$,
- $\mathrm{im}(\Delta) = d\Omega^{p-1}(X) \oplus d^*\Omega^{p+1}(X)$.

Definition 56. A *smooth differential operator D of degree d* is a map $D : \Gamma(E) \to \Gamma(F)$ which in local coordinates has the form

$$D = \sum_{|I| \leq d} a^I(x) \partial_I,$$

where the $a^I(x)$ are smooth functions with values in $\mathrm{Hom}(E|_x, F|_x)$, and $I = (I_1, \dots, I_k)$ is a multiindex, so that $\partial_I = \partial_{I_1} \cdots \partial_{I_k}$ and $|I| = k$. Specifically, given local frames $\{e_a\}_{a=1}^k$ and $\{f_b\}_{b=1}^\ell$ of E and F respectively, we can locally write $s \in \Gamma(X; E)$ as $s = \sum_{a=1}^k s^a(x) e_a$ for smooth scalar functions $s^a(x)$. Then

$$Ds = \sum_{|\alpha| \leq d} \sum_{i,j} a^I(x)_a^b \left(\partial_I s^a(x) \right) f_b.$$

Example 57. A connection A defines a smooth differential operator $\nabla_A : \Gamma(E) \to \Gamma(T^*X \otimes E)$ of degree one.

Differential operators have complicated transformation laws, generalizing the transformation law of a connection. However, the transformation law for the highest order part is simple. For local coordinates $x_1, \dots, x_n$ on X, there are corresponding local coordinates on T^*X given by $(x_1, \dots, x_n, p_1, \dots, p_n)$, where p_i are the coordinates dual to the dx^i.

Definition 58. The *principal symbol* of a smooth differential operator $D : \Gamma(E) \to \Gamma(F)$ of degree d is the map $\sigma(D, x, p) : T_x^*X \to \mathrm{Hom}(E|_x, F|_x)$ given by

$$\sigma(D, x, p) := \sum_{|I| = d} a^I(x) p_I.$$

The principal symbol is well-defined, since under change of coordinates and change of trivialization, the corrections are of order $d - 1$ and lower. It is similar to the Fourier transform since it replaces differentiation by multiplication.

Example 59. Consider the operator $d : \Omega^k(X) \to \Omega^{k+1}(X)$. The symbol is

$$\sigma(d, x, p) = p \wedge \bullet : \Lambda^k T_x^* X \to \Lambda^{k+1} T_x^* X.$$

To understand why,

$$d\omega = d(\omega_I) \wedge dx^I = dx^i \wedge (\partial_i \omega_I)\, dx^I,$$

so $\sigma(d, x, p) = dx^i p_i \wedge \bullet = p \wedge \bullet$.

Example 60. The principal symbol of ∇_A is the map $\sigma(\nabla_A, x, p) : E|_x \to T_x^* X \otimes E|_x$ given by $s \mapsto p \otimes s$.

Definition 61. Suppose a manifold X is equipped with a Riemannian metric, $E \to X$ and $F \to X$ are vector bundles equipped with inner products, and $D : \Gamma(E) \to \Gamma(F)$ is a differential operator of degree d. The *formal adjoint D^** is the differential operator given by the formula

$$D^* = \sum_{|I| \leq d} (-1)^{|I|} \partial_I \left(a^I(x)^* \sqrt{\det g} \right),$$

where $a^I(x)^* \in \mathrm{Hom}(F|_x, E|_x)$ is the adjoint of $a^I(x)$, and $\sqrt{\det g}$ is the volume form.

The formal adjoint satisfies

$$\begin{aligned}
\langle\langle Ds_1, s_2\rangle\rangle &= \int \left(a^I(x)(\partial_I s_1)\cdot s_2\right)\sqrt{\det g}\,dx\\
&= \int \partial_I s_1 \cdot \left(a^I(x)^* s_2\sqrt{\det g}\right)\,dx\\
&= (-1)^{|I|}\int s_1\cdot\partial_I\left(a^I(x)^* s_2\sqrt{\det g}\right)\,dx\\
&= \langle\langle s_1, D^* s_2\rangle\rangle\,,
\end{aligned}$$

where the second-to-last line is integration by parts, assuming that s_1 and s_2 are supported in the coordinate chart. Since everything is linear and we can use a partition of unity to write sections in terms of linear combinations supported in coordinate charts, we can ignore boundary terms.

The symbol of the formal adjoint is $\sigma(D^*, x, p) = (-1)^d\sigma(D, x, p)^*$ since the multiplication operator $a^I(x)^*$ and differentiation operator ∂_α commute modulo operators of order $d-1$.

Example 62. Consider $d^* : \Omega^k(X)\to\Omega^{k-1}(X)$. Its symbol is

$$\sigma(d, x, p) = -i_p : \Lambda^k T_x^* X \to \Lambda^{k-1}T_x^* X,$$

where i_p is the contraction map is defined as the alternating sum

$$i_p(p_1\wedge\cdots\wedge p_k) := \langle p, p_1\rangle p_2\wedge\cdots\wedge p_k - \langle p, p_2\rangle p_1\wedge p_3\wedge\cdots\wedge p_k + \cdots .$$

It satisfies $i_p\circ i_p = 0$ and $i_p(p\wedge\omega) = |p|^2\omega - p\wedge i_p\omega$.

Example 63.

$$\sigma(\nabla_A^*, x, p)(\alpha\otimes s) = -(p\cdot\alpha)s.$$

Example 64. Consider the operator $d + d^* : \Omega^{\text{even}}(X)\to\Omega^{\text{odd}}(X)$. Its symbol is

$$\sigma(d+d^*, x, p) = c(p) := -i_p(\bullet) + p\wedge\bullet : \Lambda^{\text{even}}T_x^*X\to\Lambda^{\text{odd}}T_x^*X.$$

It follows that

$$c(p)^2\omega = i_p^2\omega + p\wedge p\wedge\omega - i_p(p\wedge\omega) - p\wedge i_p(\omega) = 0 + 0 - |p|^2\,\omega.$$

This is known as the Clifford algebra relation

$$c(p)^2 = -|p|^2\,\cdot$$

Since $(d+d^*)^2 = \Delta$ and symbols compose as expected under composition of operators, it follows that

$$\sigma(\Delta, x, p) = -|p|^2\,\mathrm{Id}_{\Lambda^\bullet}.$$

Definition 65. An operator D is *elliptic* if for all $x\in X$ and for all $p\neq 0$ the symbol $\sigma(D, x, p)$ is in $\mathrm{Iso}(E|_x, F|_x)$.

Example 66. The Hodge Laplacian Δ is elliptic, since $-|p|^2\,\mathrm{Id}_{\Lambda^\bullet}$ has inverse $-|p|^{-2}\,\mathrm{Id}_{\Lambda^\bullet}$ whenever $p\neq 0$. Similarly, $d+d^*$ is elliptic since $c(p)$ has inverse $-c(p/|p|^2)$.

Furthermore, D is elliptic iff D^* is elliptic, since $\sigma(D^*, x, p)^{-1} = (-1)^d\left(\sigma(D, x, p)^{-1}\right)^*$.

Example 67. The coarse Laplacian is the map $\nabla_A^*\nabla_A : \Gamma(E)\to\Gamma(E)$. Its symbol is $\sigma(\nabla_A^*\nabla_A, x, p)(s) = \sigma(\nabla_A^*, x, p)(p\otimes s) = -|p|^2\,s$. Thus the symbols of $\nabla_{\mathrm{LC}}^*\nabla_{\mathrm{LC}}$ and Δ coincide. The difference turns out to be a zeroth-order operator in terms of the curvature.

5.3 Sobolev spaces

To prove the main theorems of elliptic theory, we need to introduce the L^2 Sobolev spaces L^2_s for $s \in \mathbb{R}$. Roughly speaking, L^2_s consists of the space of functions (or sections) whose derivatives up to order s are in L^2. However, these are not "functions" in the traditional sense, but rather distributions.

Suppose $E \to X$ is a vector bundle over a closed Riemannian manifold X. Then $L^2_s(E)$ for $s \in \mathbb{R}$ is the completion of $\Gamma(E)$ with respect to the topology induced by a particular Hilbert space norm. Equivalently, one can define $L^2_s(X)$ to be the completion of $C^\infty(X)$, and then define $L^2_s(E)$ to be sections of E with coefficients in $L^2_s(X)$, i.e. $L^2_s(E) := L^2_s(X) \otimes_{C^\infty(X)} \Gamma(E)$. Thus it suffices to concentrate on $C^\infty(X)$ and its completion $L^2_s(X)$.

Thanks to the Fourier transform, when $X = T^n$ is the n-torus, the definitions of $L^2_s(T^n)$ are especially simple. Once one understands the Sobolev spaces $L^2_s(T^n)$, it is not difficult to define $L^2_s(X)$ for any closed manifold X. Indeed, when X is closed, the spaces $L^2_s(X)$ are characterized locally. It suffices to use a finite partition of unity subordinate to some cover of X by balls, and then consider those balls as open subsets of T^n.

Before we begin, let's recall some basic properties of Banach spaces.

(i) All Hilbert spaces are Banach spaces

(ii) Two norms $\|\bullet\|$ and $\|\bullet\|'$ on a Banach space B are *equivalent* if there exists $C > 1$ such that $C^{-1}\|f\| \leq \|f\|' \leq C\|f\|$ for all $f \in B$. In this case we write $\|\bullet\| \sim \|\bullet\|'$.

(iii) A subset $S \subset B$ is *bounded* if there exists C such that $\|f\| \leq C$ for all $f \in S$.

(iv) A subset $S \subset B$ is *closed* if every Cauchy sequence in S converges to a point in S.

(v) A subspace of a Banach space is itself a Banach space iff it is closed.

(vi) A linear map $L : B_1 \to B_2$ is *bounded* if there exists C such that $\|Lf\| \leq C\|f\|$ for all $f \in B_1$.

(vii) A linear map L is bounded iff it is continuous.

(viii) A linear map L is *compact* if L takes bounded sets to sets with compact closure.

(ix) The kernel of any continuous linear map is closed.

(x) If $\|\bullet\| \sim \|\bullet\|'$ then the corresponding topologies and notions of boundedness are equivalent.

(xi) If $L : B_1 \to B_2$ is injective and has closed range, then the inverse map $L^{-1} : \operatorname{ran}(L) \to B_1$ is continuous.

(xii) A Banach space B is finite-dimensional iff its unit ball is compact.

The Sobolev spaces L^2_s have the following properties.

(A) L^2_s is a Hilbert space.

(B) C^∞ is dense in L^2_s.

(C) $L^2_0 = L^2$.

(D) If $s < t$ then $L^2_s \supset L^2_t$, and this inclusion is compact.

(E) If $D : \Gamma(E) \to \Gamma(F)$ is a smooth differential operator of order d, then $D : L^2_{s+d}(E) \to L^2_s(F)$. Consequently, $D : \mathscr{D}(E) \to \mathscr{D}(F)$.

(F) For $\alpha \in (0,1)$ and $r \in \mathbb{Z}_{\geq 0}$, there is a continuous embedding $L^2_{r+\alpha+n/2} \subset C^{r,\alpha}$ into the Hölder space. In particular, since $C^{r,\alpha} \subset C^r$, it follows that $L^2_s \subset C^r$ whenever $s > r + n/2$.

(G) $\cap_s L^2_s = C^\infty$, and $\cup_s L^2_s = \mathscr{D}$, the space of distributions on X.

(H) The pairing $f \cdot g := \int_X f\bar{g}\, d\mathrm{vol}$ for $f, g \in C^\infty$ extends to a continuous duality pairing $f \cdot g$ for all $f \in L^2_s$ and $g \in L^2_{-s}$ satisfying $|f \cdot g| \leq \|f\|_s \, \|g\|_{-s}$. For every continuous $\alpha \in (L^2_s)^*$ there exists a unique $g \in L^2_{-s}$ such that $\alpha(f) = \int_X f\bar{g}\, d\mathrm{vol}$ for all $f \in L^2_s$.

(I) If $D : \Gamma(E) \to \Gamma(F)$ is elliptic, then the map

$$\begin{aligned} \Upsilon : L^2_{s+d} &\to L^2_s \oplus L^2_{s+d-1} \\ f &\mapsto (Df, f) \end{aligned} \tag{5.2}$$

has closed image.

Although the L^2_s are Hilbert spaces, we will never make use of the inner product, only the corresponding norm, which we denote $\|\bullet\|_s$. Instead, the inner product symbol will be used only for the duality pairing of (H).

The last point (I) is extremely powerful. It implies by (xi) that the inverse map $(Df, f) \mapsto f$ to Υ is continuous, which implies by (vii) and (vi) the famous elliptic estimate

$$\|f\|_{s+d} \leq C_{D,s} \left(\|Df\|_s + \|f\|_{s+d-1} \right), \tag{5.3}$$

for some constant $C_{D,s}$ independent of f (but depending on D and s).

Theorem 68. *The elliptic estimate* (5.3) *is equivalent to (I).*

Proof. We have just seen how (5.3) implies (I). For the converse, note that (5.3) implies that the inverse to Υ of (5.2) is continuous. Furthermore, Υ is continuous by (E) and (D). Thus (5.3) is a homeomorphism from L^2_{s+d} to its image. In particular, the image of Υ must be closed. $\square$

The elliptic estimate implies all the interesting properties of elliptic operators. We shall take this approach in Section 5.5 after defining the Sobolev spaces L^2_s.

5.4 Fourier theory on T^n

In what follows, $C^\infty(T^n)$ will be short for the smooth complex-valued functions $C^\infty(T^n;\mathbb{C})$.

Definition 69. The n-torus is defined by

$$T^n := \mathbb{R}^n/(2\pi\mathbb{Z})^n.$$

Definition 70. For $f, g \in C^\infty(T^n)$, the inner product $f \cdot g$ is defined by

$$f \cdot g := \int_{T^n} f\bar{g}\, d\text{vol},$$

with $d\text{vol} := (2\pi)^{-n} dx^1 \cdots dx^n$ so that $\int_{T^n} 1\, d\text{vol} = 1$.

Remark 71. With respect to this inner product, the functions $\left\{e^{k\cdot x}\right\}_{k\in\sqrt{-1}\,\mathbb{Z}^n}$ are orthonormal.

Definition 72. Define L to be the lattice $L := \sqrt{-1}\,\mathbb{Z}^n$.

Definition 73. For $f \in C^\infty(T^n;\mathbb{C})$, define $\Delta f := -\sum_{i=1}^n \partial_i^2 f$.

Remark 74. Integration by parts shows that $(\partial_i f)\cdot g = -f \cdot \partial_i g$. It follows that $(\Delta f)\cdot g = f\cdot(\Delta g)$. More generally, if $p(z)$ is any polynomial, then $p(\Delta)f\cdot g = f\cdot p(\Delta)g$. Finally,

$$p(\Delta)e^{k\cdot x} = p(|k|^2)e^{k\cdot x}.$$

The minus sign in the definition of Δ ensures that it is positive-semidefinite.

Definition 75. The Schwarz space $\mathcal{S}(L)$, often abbreviated as $\mathcal{S}$, is defined to be the space of complex-valued functions on L which decay faster than any polynomial. Specifically,

$$\mathcal{S}(L) := \left\{c : L \to \mathbb{C} \mid \forall a \in \mathbb{Z}_{\geq 0} \exists K_a \text{ such that } |c(k)| \leq K_a(1+|k|^2)^{-a}\right\}.$$

Definition 76. The inner product $c_1 \cdot c_2$ on $\mathcal{S}(L)$ is defined by

$$c_1 \cdot c_2 := \sum_{k\in L} c_1(k)\overline{c_2(k)}.$$

Remark 77. The sum in $c_1 \cdot c_2$ is absolutely convergent since $\sum(1+|k|^2)^{-a}$ converges for $a > \frac{1}{2}n$, and one can choose bounds on c_1 and c_2 so that $\left|c_1(k)\overline{c_2(k)}\right| \leq K(1+|k|^2)^{-a}$ for any $a > \frac{1}{2}n$.

Definition 78. The *inverse Fourier transform* $\mathcal{F}^{-1} : \mathcal{S}(L) \to C^\infty(T^n;\mathbb{C})$ is given by

$$\mathcal{F}^{-1}(c)(x) := \sum_{k\in L} c(k)e^{k\cdot x}.$$

Lemma 79. *The inverse Fourier transform is a well-defined isometry onto its image.*

Proof. First we must check that the image of $\mathscr{F}^{-1}$ consists of C^∞ functions. Thanks to the rapid decay, all sums will be absolutely convergent, and thus it is justified to swap orders of derivatives, summations, and integrals. In particular, $\sum_k c(k)e^{k\cdot x} \in C^\infty$ because

$$\partial_I \sum_k c(k)e^{k\cdot x} = \sum_k k^I c(k)e^{k\cdot x}$$

is absolutely convergent for any multiindex I. For the isometry claim,

$$\begin{aligned}
\mathscr{F}^{-1}(c_1) \cdot \mathscr{F}^{-1}(c_2) &= \int_{T^n} \left(\sum_{k_1} c_1(k_1)e^{k_1\cdot x}\right)\left(\sum_{k_2} \overline{c_2(k_2)}e^{-k_2\cdot x}\right) d\mathrm{vol} \\
&= \sum_{k_1,k_2} c_1(k_1)\overline{c_2(k_2)} \int_{T^n} e^{(k_1-k_2)\cdot x}\, d\mathrm{vol} \\
&= c_1 \cdot c_2.
\end{aligned}$$

In particular, this shows that $\mathscr{F}^{-1}$ is injective. □

Definition. The *Fourier transform* $\mathscr{F} : C^\infty(T^n;\mathbb{C}) \to \mathcal{S}(\sqrt{-1}\,\mathbb{Z}^n)$ is $\mathscr{F}(f) = c_f$ with $c_f(k) := f \cdot e^{k\cdot x}$.

Theorem 80. *The Fourier transform $\mathscr{F}$ is a well-defined two-sided inverse for $\mathscr{F}^{-1}$.*

Proof. To show that $\mathscr{F}$ takes $C^\infty(T^n;\mathbb{C})$ to $\mathcal{S}(L)$, we must show that c_f decays quickly. Note that for any $a \in \mathbb{Z}_{\geq 0}$,

$$e^{k\cdot x} = (1+|k|^2)^{-a}(1+\Delta)^a e^{k\cdot x},$$

so that

$$\left|c_f(k)\right| = \left(1+|k|^2\right)^{-a}\left|((1+\Delta)^a f)\cdot e^{k\cdot x}\right| \leq \left(1+|k|^2\right)^{-a}\left\|(1+\Delta)^a f\right\|_{C^0}.$$

Therefore, to show that $c_f \in \mathcal{S}(\sqrt{-1}\,\mathbb{Z}^n)$, it suffices to take $K_a = \|(1+\Delta)^a f\|_{C^0}$.

To show that $\mathscr{F}$ is a left-inverse to $\mathscr{F}^{-1}$, compute

$$(\mathscr{F}\mathscr{F}^{-1}c)(k_0) = \left(\sum_k c(k)e^{k\cdot x}\right)\cdot e^{k_0\cdot x} = \sum_k c(k)\delta^k_{k_0} = c(k_0).$$

The interesting direction is to show that

$$\mathscr{F}^{-1}\mathscr{F}f = f.$$

In particular, plugging in $x = 0$, we wish to show that

$$\sum_k c_f = f(0).$$

This is clear in the case that f is constant, in which case $c_f = f(0)\delta_0$, where δ_0 is the Kronecker delta function. It is also clear if f is of the form $f(x) = \left(e^{k_1\cdot x} - e^{k_2\cdot x}\right)g(x)$ for $g \in C^\infty$, since then

$c_f(k) = c_g(k - k_1) - c_g(k - k_2)$, which cancels after summing over k. For general f, using Taylor's theorem with remainder, it is not difficult to write

$$f(x) = f(0) + \sum_{i=1}^{n} (e^{\sqrt{-1}x^i} - 1) g_i(x)$$

for smooth functions $g_i \in C^\infty$, from which the theorem follows from the previous cases. □

Definition 81. For $s \in \mathbb{R}$, define $(1 + \Delta)^{s/2} : C^\infty(T^n) \to C^\infty(T^n)$ by

$$(1 + \Delta)^{s/2} f := \mathscr{F}^{-1} \left((1 + |k|)^{s/2} c_f(k) \right).$$

Definition 82. Let $L^2(T^n)$ denote the L^2 completion of $C^\infty(T^n)$, and let $\ell^2(L)$ denote the square-summable sequences on L.

Remark 83. Since $\mathscr{F}$ is an isometry from $C^\infty(T^n)$ to $\mathscr{S}(L)$, it extends to an isometry on the L^2 completions

$$\mathscr{F} : L^2(T^n; \mathbb{C}) \to \ell^2(\sqrt{-1}\,\mathbb{Z}^n).$$

Definition 84. The *Sobolev space* $L^2_s(T^n; \mathbb{C})$ is the L^2 completion of $C^\infty(T^n; \mathbb{C})$ with respect to the norm $\|f\|_s := \left\|(1 + \Delta)^{s/2} f\right\|_0$, where $\|f\|_0$ denotes the standard L^2 norm.

Remark 85. One can identify $f \in L^2_s(T^n; \mathbb{C})$ with the sequence c_f such that $(1 + |k|^2)^{s/2} c_f(k) \in \ell^2$. If $s \geq 0$, then c_f can be identified with the Fourier transform of some actual element $f \in L^2$.

Definition 86. For $s \in \mathbb{R}$, let $\ell^2_s(L)$ denote $(1 + |k|^2)^{-s/2} \ell^2(L)$, specifically

$$\ell^2_s(L) := \left\{ c : L \to \mathbb{C} \,\middle|\, \sum_k (1 + |k|^2)^s \, |c(k)|^2 < \infty \right\}.$$

Remark 87. If $s < t$, then $\ell^2_s \supset \ell^2_t \supset \mathscr{S}$. Correspondingly, $L^2_s \supset L^2_t \supset C^\infty$.

Lemma 88. *Each $\{\partial_i\}_{i=1}^n$ induces a map $L^2_{s+1} \to L^2_s$ for each $s \in \mathbb{R}$.*

Proof.

$$\|\partial_i f\|^2_{s+1} = \left\|(1 + \Delta)^{(s+1)/2} (\partial_i f)\right\|^2_0 = \sum_k \frac{k_i^2}{(1 + |k|^2)} (1 + |k|^2)^s \left|c_f(k)\right|^2 \leq \|f\|^2_s.$$

□

Definition 89. For $c_1, c_2 \in \mathscr{S}(L)$, the *convolution product* $c_1 * c_2$ is

$$(c_1 * c_2)(k) := \sum_{k_1 + k_2 = k} c_1(k_1) c_2(k_2).$$

Lemma 90. *If $f, g \in C^\infty(T^n; \mathbb{C})$, then $\mathscr{F}(fg) = \mathscr{F}(f) * \mathscr{F}(g)$.*

Lemma 91. *If $g \in C^\infty$, then multiplication by g induces a bounded map $L^2_s \to L^2_s$ for each $s \in \mathbb{R}$.*

Proof.

$$\|fg\|_s = \left\|(1+\Delta)^{s/2}(fg)\right\|_0 = \left\|(1+|k|)^{s/2}(c_f * c_g)(k)\right\|_0$$
$$\ldots$$

□

Theorem 92. *If $D : C^\infty(T^n;\mathbb{C}) \to C^\infty(T^n;\mathbb{C})$ is a differential operator of degree d, then D induces a bounded linear map $L^2_{s+d} \to L^2_s$.*

Theorem 93. *If $s > n/2$ and $f \in L^2_s$, then $f \in C^0$ with $\|f\|_{C^0} \leq K_s \|f\|_{L^2_s}$.*

Proof. Compute

$$\|f\|_{C^0} = \left\|\sum_k c_f(k)e^{k\cdot x}\right\|_{C^0} \leq \sum_k \left|c_f(k)\right| \leq \sum_k (1+|k|^2)^{s/2}\left|(1+|k|^2)^{-s/2}c_f(k)\right| \leq \sqrt{\sum_k (1+|k|^2)^{-s}}\,\|f\|_{L^2_s} = K_s\|f\|_{L^2_s}.$$

□

Corollary 94. *For every $\varepsilon > 0$, $L^2_k \supset C^k \supset L^2_{k+n/2+\varepsilon}$. In particular,*

$$\bigcup_{s\to\infty} L^2_s(T^n) = C^\infty(T^n).$$

Theorem 95. *For each s, the inner product $C^\infty(T^n)\times C^\infty(T^n) \to \mathbb{C}$ naturally extends to a perfect duality pairing $L^2_s \times L^2_{-s} \to \mathbb{C}$, given by $f\cdot g = \sum_k c_f(k)\overline{c_g(k)}$. It satisfies the estimate $|f\cdot g| \leq \|f\|_s\|f\|_{-s}$. For every bounded linear functional $\phi : L^2_s \to \mathbb{C}$, there exists a unique $g \in L^2_{-s}$ such that $\phi(f) = f\cdot g$.*

Proof. The bound follows from

$$f\cdot g = (1+\Delta)^{s/2}f\cdot(1+\Delta)^{-s/2}g.$$

Now both $(1+\Delta)^{s/2}f$ and $(1+\Delta)^{-s/2}g$ are in L^2 with $\left\|(1+\Delta)^{s/2}f\right\|_0 = \|f\|_s$. The existence of g given ϕ follows from applying the Riesz representation theorem. □

Definition 96. Let $\mathcal{S}'(\sqrt{-1}\,\mathbb{Z}^n)$ denote the space of sequences with polynomial growth. Specifically,

$$\mathcal{S}'(\sqrt{-1}\,\mathbb{Z}^n) := \left\{c : \sqrt{-1}\,\mathbb{Z}^n \to \mathbb{C} \mid \exists a, K \text{ such that } |c(k)| \leq K(1+|k|^2)^a\right\}.$$

Remark 97.

$$\bigcap_{s\to-\infty} \ell^2_s(T^n) = \mathcal{S}'(\sqrt{-1}\,\mathbb{Z}^n).$$

Remark 98. The space $\mathcal{S}'(\sqrt{-1}\,\mathbb{Z}^n)$ is the dual space of

Example 99. Consider the constant function $1 \in \mathcal{S}'(\sqrt{-1}\,\mathbb{Z}^n)$. The pairing $c_f \cdot 1$ gives

$$c_f \cdot 1 = \sum_k c_f(k) = f(0).$$

In the function domain, we should think of this as $f \cdot \delta = f(0)$, where δ denotes the Dirac delta function on T^n.

Definition 100. The space of *distributions* $\mathcal{S}'(T^n)$ is the set of linear functions $\phi : C^\infty(T^n; \mathbb{C}) \to \mathbb{C}$ such that for every ϕ there exist constants a_ϕ and K_ϕ so that

$$|\phi(f)| \leq K_\phi \left\|(1+\Delta)^{a_\phi} f\right\|_{C^0} \;\; \forall f \in C^\infty(T^n; \mathbb{C}).$$

5.5 Elliptic theory

Theorem 101. *If D is elliptic, then any distributional solution to $Df = 0$ is smooth.*

Proof. Suppose $f \in \mathscr{D}$ and $Df = 0$. Then by (G), $f \in L^2_t$ for some $t \in \mathbb{R}$. Thus $\|f\|_t < \infty$. The elliptic estimate implies

$$\|f\|_{s+d} \leq C_{D,s} \|f\|_{s+d-1} \tag{5.4}$$

for any $s \in \mathbb{R}$. Taking $s = t - d + 1$, it follows that $\|f\|_{t+1} < \infty$, so $f \in L^2_{t+1}$. By induction, $f \in L^2_s$ for all $s \in \mathbb{R}$. From (G) it follows that $f \in C^\infty$. □

In general, one would expect, as s decreases so that L^2_{s+d} grows, that the subspace $\ker(D : L^2_{s+d} \to L^2_s) \subset L^2_{s+d}$ is also likely to grow. However, Theorem 101 shows that this is not the case when D is elliptic, and the kernel is independent of s when viewed as a subspace of distributions $\mathscr{D}$. Indeed, $\ker D$ is a fixed subspace of C^∞.

Theorem 102. *If D is elliptic, then* $\ker D$ *is finite dimensional.*

We present two proofs. The first uses sequences and estimates, while the second uses more abstract Banach space methods.

Proof. Suppose for a contradiction that $\ker D$ is infinite-dimensional. Since $\ker D \subset C^\infty$, the inner products $f \cdot g$ make sense for any $f, g \in \ker D$. In particular, one can construct an infinite orthonormal sequence $\{f_i\} \subset \ker D$. Thus the L^2_0 norms are $\|f_i\|_0 = 1$. It follows from (5.3) with $s = 1 - d$ that $\{f_i\}$ is bounded in L^2_1. By (D), $\{f_i\}$ must admit a convergent subsequence in L^2_0. However $\{f_i\}$ is not even Cauchy in L^2_0 since $\left\|f_i - f_j\right\|_0 = 2$ for any $i \neq j$ by orthonormality. □

Alternative proof. Consider the subspace

$$K := \ker(D : L^2_{s+d} \to L^2_s) \subset L^2_{s+d}.$$

This subspace is closed by (ix). Thus by (v), K is a Banach space. Recall that the map Υ from (5.2) is a homeomorphism onto its image from the proof of Theorem 68. Therefore the restriction $\Upsilon|_K$ is also a homeomorphism onto its image, given by $f \mapsto (0, f)$, $L^2_{s+d} \to \{0\} \oplus L^2_s$. By (D), $\Upsilon|_K$ is compact. In particular, we have shown that K admits a compact linear homeomorphism. Thus the unit ball of K must be compact by (viii). It follows from (xii) that K is finite-dimensional. □

Theorem 103. *If $D : \Gamma(E) \to \Gamma(F)$ is elliptic, then there is a decomposition $\mathscr{D}(E) = \ker D \oplus (\ker D)^{\perp}$.*

Proof. Suppose $\{k_1, \dots, k_r\}$ is an orthonormal basis for $\ker D$. Then for any $f \in \mathscr{D}(E)$, $f \in L^2_s$ for some $s \in \mathbb{R}$. Since $\{k_i\} \subset C^{\infty} \subset L^2_{-s}$, it follows that $f \cdot k_i$ is well-defined. Then using the Gram-Schmidt procedure,

$$f = (f \cdot k_1)k_1 + \cdots + (f \cdot k_r)k_r + (f - (f \cdot k_1)k_1 + \cdots + (f \cdot k_r)k_r) \in \ker D \oplus (\ker D)^{\perp}.$$

□

Theorem 104 (Poincaré inequality). *If D is elliptic of order d, then there exist constants $C'_{D,s}$ such that for all $f \in (\ker D)^{\perp}$,*

$$\|f\|_{s+d} \leq C'_{D,s} \|Df\|_s .$$

Proof. Suppose for contradiction that the inequality fails. Then there exists a sequence $\{f_i\}$ such that $f_i \in (\ker D)^{\perp}$

$$\frac{\|f_i\|_{s+d}}{\|Df_i\|_s} \to \infty.$$

Rescaling the f_i so that $\|f_i\|_{s+d} = 1$, this is equivalent to $\|Df_i\|_s \to 0$. Since $\{f_i\}$ is bounded in L^2_{s+d}, it follows that there is a subsequence which converges to some $f_\infty \in L^2_{s+d-1}$. For simplicity, denote this subsequence also by $\{f_i\}$, so that $f_i \to f_\infty$ in L^2_{s+d-1}.

First we wish to show that $f_\infty \in (\ker D)^{\perp}$. For any $k \in \ker D$, consider $f_i \cdot k$. This is zero since $f_i \in (\ker D)^{\perp}$. Next we take the limit as $i \to \infty$. Since $k \in C^{\infty}$, $k \in (L^2_{s+d-1})^* = L^2_{-s-d+1}$, we can use the continuity of the duality pairing $L^2_{s+d-1} \times L^2_{-s-d+1} \to \mathbb{C}$ to compute

$$0 = \lim_{i \to \infty} (f_i \cdot k) = f_\infty \cdot k.$$

Thus $f_\infty \in (\ker D)^{\perp}$.

Next we will show that $Df_\infty = 0$ by showing that $Df_\infty \cdot \phi = 0$ for any $\phi \in C^{\infty}$. Consider

$$Df_\infty \cdot \phi = f_\infty \cdot D^*\phi = \lim_{i \to \infty} f_i \cdot D^*\phi,$$

where once again we use continuity of the same duality pairing. Thus

$$Df_\infty \cdot \phi = \lim_{i \to \infty} (f_i \cdot D^*\phi) = \lim_{i \to \infty} (Df_i \cdot \phi) \leq \lim_{i \to \infty} \|Df_i\|_s \|\phi\|_{-s} = 0.$$

Thus $Df_\infty = 0$. Since $f_\infty \in (\ker D) \cap (\ker D)^{\perp}$, it follows that $f_\infty = 0$. However, this contradicts the elliptic estimate

$$1 = \|f_i\|_{s+d} \leq C_{D,s} \left(\|Df_i\|_s + \|f_i\|_{s+d-1} \right)$$

since both $\|Df_i\|_s \to 0$ and $\|f_i\|_{s+d-1} \to \|f_\infty\|_{s+d-1} = 0$.

□

The Poincaré inequality allows us to strengthen the elliptic estimate. Let $\pi_k : L^2_{s+d}(E) \to L^2_{s+d}(E)$ be the orthogonal projection to $\ker D$. For any orthonormal basis $\{k_1, \ldots, k_r\}$ of $\ker D$,

$$\pi_k(f) := \sum_{i=1}^{r} (f \cdot k_i) k_i.$$

For any $f \in L^2_{s+d}(E)$, we can write $f = \pi_k(f) + f_\perp$, where $f_\perp := f - \pi_k(f)$ is orthogonal to $\ker D$. In this case,

$$\|f\|_{s+d} \le \|f_\perp\|_{s+d} + \|\pi_k(f)\|_{s+d} \le C'_{D,s} \|Df_\perp\|_s + \|\pi_k(f)\|_{s+d} = C'_{D,s} \|Df\|_s + \|\pi_k(f)\|_{s+d}.$$

Now note that the term $\|\pi_k(f)\|_{s+d}$ consists of the restriction of the $\|\bullet\|_{s+d}$ norm to the finite-dimensional subspace $\ker D$. All norms on a finite-dimensional vector space are equivalent. Thus, after possibly changing the constant $C'_{D,s}$, we obtain

$$\|f\|_{s+d} \le C'_{D,s} \left(\|Df\|_s + \|\pi_k(f)\| \right)$$

for any fixed norm $\|\bullet\|$ on $\ker D$. This also implies

$$\|f\|_{s+d} \le C_{D,s,s'} \left(\|Df\|_s + \|f\|_{s'} \right)$$

for any s', not just for $s' = s + d - 1$.

Suppose $D : \Gamma(E) \to \Gamma(F)$ is an elliptic differential operator. Then $D^* : \Gamma(F) \to \Gamma(E)$ is also elliptic, and there are induced maps on the completions $D : L^2_{s+d}(E) \to L^2_s(F)$ and $D^* : L^2_{-s}(F) \to L^2_{-s-d}(E)$. Furthermore, $\ker D^*$ is finite-dimensional, spanned by finitely many smooth sections in $\Gamma(F)$. We wish to prove the Hodge decomposition

$$\Gamma(F) = \ker D^* \oplus \operatorname{im} D.$$

We will prove more generally that

$$L^2_s(F) = \ker D^* \oplus \operatorname{im} D.$$

Since $\ker D^*$ is finite-dimensional, we can use a finite Gram-Schmidt process to write

$$L^2_s(F) = \ker D^* \oplus (\ker D^*)^\perp.$$

It remains to show that $(\ker D^*)^\perp = \operatorname{im} D$. In particular, for any $\alpha \in L^2_s(F)$ which satisfies $\alpha \perp \ker D^*$, we want to construct a solution $\omega \in L^2_{s+d}(E)$ to the PDE

$$D\omega = \alpha \in L^2_s(F). \tag{5.5}$$

By duality, this is equivalent to

$$\psi \cdot D\omega = \psi \cdot \alpha \quad \forall \psi \in L^2_{-s}(F),$$

which is the same as

$$D^*\psi \cdot \omega = \psi \cdot \alpha \quad \forall \psi \in L^2_{-s}(F).$$

Now $\Gamma(F) \subset L^2_{-s}(F)$ is dense since $L^2_{-s}(F)$ is a completion of $\Gamma(F)$. Thus our equation is equivalent to

$$D^*\psi \cdot \omega = \psi \cdot \alpha \quad \forall \psi \in \Gamma(F).$$

This is called the "weak form" of $D\omega = \alpha$.

Philisophically, the "weak form" corresponds to viewing ω as a *distribution*, i.e. a continuous functional on the space of smooth sections $\Gamma(E) \to \mathbb{C}$. Any section α of a vector bundle E with coefficients in L^2_s determines a unique functional ℓ_α by

$$\ell_\alpha : \Gamma(E) \to \mathbb{C},$$
$$\ell_\alpha(\phi) := \phi \cdot \alpha,$$

which satisfies the inequality

$$\ell_\alpha(\phi) \leq \|\alpha\|_s \|\phi\|_{-s}.$$

Conversely, suppose $\ell : \Gamma(E) \to \mathbb{C}$ satisfies $|\ell(\phi)| \leq C \|\phi\|_{-s}$ for some $C, s \in \mathbb{R}$. This implies that ℓ is bounded with respect to the L^2_{-s} topology on $\Gamma(E)$. By the Hahn-Banach theorem, ℓ extends to a bounded map $\ell : L^2_{-s}(E) \to \mathbb{C}$. Since $\Gamma(E) \subset L^2_{-s}(E)$ is dense, this extension is unique. Thus $\ell \in (L^2_{-s}(E))^* = L^2_s(E)$, so $\ell = \ell_\alpha$ for some $\alpha \in L^2_s(E)$. This motivates the following definition.

Definition 105. The space of *distributional sections* $\mathscr{D}'(E)$ is the space of continuous linear functionals $\Gamma(E) \to \mathbb{C}$, where $\ell : \Gamma(E) \to \mathbb{C}$ is *continuous* if there exists $C, s \in \mathbb{R}$ such that $|\ell(\phi)| \leq C \|\phi\|_{-s}$.

From this definition, it is clear that $\mathscr{D}'(E) = \cup_{s \to -\infty} L^2_s(E)$.

For any differential operator $D : \Gamma(E) \to \Gamma(F)$ with smooth coefficients, we want to make sense of D in terms of distributions. Assuming that boundary terms vanish (i.e. X is closed), then

$$\ell_{D\alpha}(\psi) = \psi \cdot D\alpha = D^*\psi \cdot \alpha = \ell_\alpha(D^*\psi).$$

Thus distributions over closed manifolds satisfy

$$\ell_{D\alpha}(\psi) = \ell_\alpha(D^*\psi), \ \forall \psi \in \Gamma(F).$$

Recall that we wish to construct $\omega \in L^2_{s+d}(E)$ which solves $D\omega = \alpha$ given any $\alpha \in L^2_s(F)$ which satisfies $\alpha \perp \ker D^*$. This is equivalent to the weak form

$$\ell_\omega(D^*\psi) = \psi \cdot \alpha \quad \forall \psi \in \Gamma(F).$$

To solve this equation, we define a distribution $\ell : \Gamma(E) \to \mathbb{C}$ which corresponds formally to $\ell =$ "$\ell_{D^{-1}\alpha}$".

To make the correct definition, we should understand this formal correspondence. It should satisfy

$$\text{“}\ell_{D^{-1}\alpha}(D^*\psi) = D^*\psi \cdot D^{-1}\alpha = \psi \cdot DD^{-1}\alpha = \psi \cdot \alpha.\text{”}$$

Thus we define for $\alpha \perp \ker D^*$,

$$\ell : (\text{image } D^* \subset \Gamma(E)) \to \mathbb{C},$$
$$\ell(D^*\psi) := \psi \cdot \alpha.$$

First we must check that this is well-defined, independent of our choice of ψ. If $D^*\psi_1 = D^*\psi_2$, then $D^*(\psi_1 - \psi_2) = 0$ so $\psi_1 - \psi_2 \in (\ker D^*)^\perp$. Since $\alpha \in (\ker D^*)^\perp$, it follows that $\psi_1 \cdot \alpha = \psi_2 \cdot \alpha$, and ℓ is indeed well-defined. Next we wish to extend from $\ell : \operatorname{im} D^* \to \mathbb{C}$ to $\tilde{\ell} : \Gamma(E) \to \mathbb{C}$. For this we will use the Hahn-Banach theorem. In order for it to apply, we must show that ℓ is bounded in some Sobolev topology.

Lemma 106. *The functional $\ell : (\operatorname{im} D^*) \to \mathbb{C}$, is bounded in the L^2_{-s-d} topology, i.e.*

$$|\ell(\phi)| \leq C_{D,\alpha,s} \|\phi\|_{L^2_{-s-d}} \qquad \forall \phi \in \Gamma(E)$$

for some $C_{D,\alpha,s}$ which is independent of ϕ.

Assuming this lemma, then by the Hahn-Banach theorem, ℓ extends to a bounded linear functional $\tilde{\ell} : L^2_{-s-d}(E) \to \mathbb{C}$. By the Riesz representation theorem, there is some $\omega \in L^2_{s+d}(E)$ such that $\tilde{\ell}(\phi) = \phi \cdot \omega$. This ω then satisfies $D\omega = \alpha \in L^2_s(F)$ since for all $\phi \in \Gamma(F)$,

$$\phi \cdot D\omega = D^*\phi \cdot \omega = \tilde{\ell}(D^*\phi) = \ell(D^*\phi) = \phi \cdot \alpha.$$

This lemma follows quickly from the Poincaré inequality. Assuming that $\phi = D^*\psi$, it is simple to rearrange that $\psi \perp \ker D^*$. Then gives

$$|\ell(\phi)| = |\ell(D^*\psi)| = |\psi \cdot \alpha| \leq \|\psi\|_{-s} \|\alpha\|_s \leq C'_{D^*,-s} \|D^*\psi\|_{-s-d} \|\alpha\|_s \leq \left(C'_{D^*,-s} \|\alpha\|_s\right) \|\phi\|_{-s-d},$$

proving the lemma.

Now recall our original motivation. We wanted to produce a slice $\mathcal{S}_{A_0}$ through some connection A_0 which is transverse to the action of gauge transformations.

$$\begin{aligned} T_{A_0}\mathcal{A} &= T_{A_0}\mathcal{O}_{A_0} \oplus T_{A_0}\mathcal{S}_{A_0}, \\ \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) &= \operatorname{Im}(d_{A_0}) \oplus C. \end{aligned}$$

We want to choose $C = \operatorname{Im}(d_{A_0})^\perp$. To find the appropriate decomposition of $\Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$, we need an elliptic operator. We take $D = d_{A_0} + d^*_{A_0}$ from sections of the bundle $\Lambda^\bullet T^*X \otimes \mathfrak{g}_{\mathrm{Ad}}$ to sections of the same bundle. The symbol is the Clifford map $c(p) := -i_p(\bullet) + p \wedge \bullet$, acting as the identity on the $\mathfrak{g}_{\mathrm{Ad}}$ factor, which is invertible away from $p \neq 0$ since $c(p)^2 = -|p|^2 \cdot$. It follows that $\ker D$ is finite-dimensional, and $\Omega^\bullet(X; \mathfrak{g}_{\mathrm{Ad}}) = \ker D \oplus \operatorname{im} D$. On any Sobolev completion, D acts as a homeomorphism on $\operatorname{im} D$. In particular, D takes any closed space to a closed space. For example, the image of D on the closed subspace $\Omega^p(X; \mathfrak{g}_{\mathrm{Ad}})$ is some closed subspace of

$$\Omega^{p-1}(X; \mathfrak{g}_{\mathrm{Ad}}) \oplus \Omega^{p+1}(X; \mathfrak{g}_{\mathrm{Ad}}).$$

Restricting to each of the two factors, we find that the image of $d^*_{A_0}$ is a closed subspace of $\Omega^{p-1}(X; \mathfrak{g}_{\mathrm{Ad}})$ and that the image of d_{A_0} is a closed subspace of $\Omega^{p+1}(X; \mathfrak{g}_{\mathrm{Ad}})$. It follows that

$$\operatorname{im} d_{A_0} = \overline{\operatorname{im} d_{A_0}} = (\operatorname{im} d_{A_0})^{\perp\perp} = (\ker d^*_{A_0})^\perp,$$

and similarly with d_{A_0} and $d^*_{A_0}$ exchanged. Thus

$$\Omega^p(X; \mathfrak{g}_{\mathrm{Ad}}) = \operatorname{im} d_{A_0} \oplus \ker d^*_{A_0} = \operatorname{im} d^*_{A_0} \oplus \ker d_{A_0}.$$

Note that when A_0 is not flat, we do not generally have that $\operatorname{im} d_{A_0} \perp \operatorname{im} d_{A_0}^*$ since

$$d_{A_0}\alpha \cdot d_{A_0}^*\beta = d_{A_0}^2\alpha \cdot \beta = [F_{A_0} \wedge \alpha] \cdot \beta.$$

Thus we don't have the more general decomposition

$$\Omega^p(X; \mathfrak{g}_{\mathrm{Ad}}) \neq \operatorname{im} d_{A_0} \oplus \operatorname{im} d_{A_0}^* \oplus (\ker d_{A_0} \cap \ker d_{A_0}^*).$$

5.6 L_k^p Sobolev spaces

For $1 \leq p < \infty$, the space $L^p(E)$ is the Banach space completion of $\Gamma(E)$ with respect to the norm

$$\|s\|_{L^p} := \left(\int_X |s|^p\right)^{1/p}.$$

Each $s \in L^p(E)$ is determined by a measurable section. Two measurable sections which are equal almost everywhere determine the same element of $L^p(E)$. There is also $L^\infty(E)$ which is defined on measurable sections of finite norm

$$\|s\|_{L^\infty} = \lim_{p\to\infty} \|s\|_{L^p} = \operatorname{ess\,sup} |s|.$$

Note that $\Gamma(E)$ is not dense in $L^\infty(E)$. The completion of $\Gamma(E)$ with respect to the L^∞ norm is simply $C^0(E)$.

The L^p norms satisfy the estimate

$$|f \cdot g| \leq \|f\|_{L^p} \|g\|_{L^q}, \quad 1 \leq p, q \leq \infty,\ p^{-1} + q^{-1} = 1.$$

Indeed, for $1 \leq p < \infty$, the dual space $(L^p)^* = L^q$, where $1 < q \leq \infty$ satisfies $p^{-1} + q^{-1} = 1$, so that any $\ell \in (L^p)^*$ is given by $\ell = \ell_g$ for some $g \in L^q$ so that $\ell_g(f) = f \cdot g$. For any p, the corresponding value of q is called the dual exponent. Note that the dual space $(L^\infty)^*$ is larger than L^1.

By iteration, if $\sum_{i=1}^k p_i^{-1} = 1$, then $\left|\int f_1 \cdots f_k\right| \leq \|f_1\|_{p_1} \cdots \|f_k\|_{p_k}$.

For $k \in \mathbb{Z}_{\geq 0}$, define the L_k^p norms by

$$\|f\|_{L_k^p}^p := \sum_{i=0}^k \left\|\nabla^i f\right\|_{L^p}^p.$$

When $p = 2$, $X = T^n$, and $s \in \mathbb{Z}_{\geq 0}$,

$$\|f\|_{L_s^2}^2 = \sum_{i=0}^s \left\|\nabla^i f\right\|_{L^2}^2 = \sum_{\|I\|\leq s} \|\partial_I f\|_{L^2}^2 = \sum_{\|I\|\leq s} \sum_{k\in L} \left|k^I c(k)\right|^2 = \sum_{k\in L} \left(\sum_{\|I\|\leq s} \left|k^I\right|^2\right) |c(k)|^2.$$

It is readily verified that this is equivalent to

$$\|f\|_s^2 = \sum_{k\in L} (1 + |k|^2)^s \, |c(k)|^2.$$

In this sense, the L^p_k spaces generalize the L^2_s spaces.

If E is a vector bundle, and A_0 is any smooth connection on E, then define

$$\|s\|^p_{L^p_{k,A_0}} := \sum_{i=0}^{k} \left\|\nabla^i_{A_0} s\right\|^p_{L^p}.$$

It can be verified that for any two connections A_0 and A_1, the norms $\|s\|_{L^p_{k,A_0}}$ and $\|s\|_{L^p_{k,A_1}}$ are equivalent.

5.7 Slice theorem

Now let's return to the proof that

$$\begin{aligned} \mathfrak{m} : (\mathscr{G} \times \mathscr{S}_{A_0})/\mathrm{Stab}(A_0) &\to \mathscr{A}, \\ \mathfrak{m}(g, A) &:= g \cdot A \end{aligned}$$

is a local diffeomorphism onto its image. We will assume that $n = \dim X$ satisfies $n \geq 3$. This is to ensure that the solution r to $n^{-1} + 2^{-1} + r^{-1} = 1$ satisfies $1 \leq r < \infty$, and the required Sobolev embedding exists.

Recall that we need

$$\begin{aligned} T_{A_0}\mathscr{A} &= T_{A_0}\mathscr{O}_{A_0} \oplus T_{A_0}\mathscr{S}_{A_0}, \\ \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) &= \operatorname{im} d^*_{A_0} \oplus \ker d^*_{A_0}. \end{aligned}$$

Thus we take

$$\mathscr{S}_{A_0}(\varepsilon) = \left\{ A_0 + a \mid a \in \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}),\ d^*_{A_0} a = 0,\ \|a\|_{L^n} < \varepsilon \right\}.$$

Our goal will be to show that when ε is sufficiently small, that under the appropriate Sobolev completions, $\mathfrak{m}$ indeed defines a local diffeomorphism onto its image.

First we verify that the linearization $D_{(A,1)}\mathfrak{m}$ is an isomorphism from $\left(T_1\mathscr{G} \times T_{A_0}\mathscr{S}_{A_0}(\varepsilon)\right)/T_1\mathrm{Stab}(A_0) \to T_{A_0}\mathscr{A}$. Recall that $T_1\mathscr{G} = \Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})$, $T_{A_0}\mathscr{S}_{A_0}(\varepsilon) = \ker d^*_{A_0} \subset \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$, and $T_1\mathrm{Stab}(A_0) = \ker d_{A_0} \subset \Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})$, and $T_{A_0}\mathscr{A} = \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})$. Recall that

$$\mathfrak{m}(g, A_0 + a) = g \cdot (A_0 + a) = A_0 + gag^{-1} - (d_A g)g^{-1}.$$

Thus the map $D_{(A,1)}\mathfrak{m}(\xi, a) = a - d_A\xi$. (Consider the linear term in t when replacing $g \mapsto 1 + t\xi$, $g^{-1} \mapsto 1 - t\xi$, $a \mapsto ta$.)

Then $D_{(A,1)}\mathfrak{m}$ is clearly surjective by the Hodge decomposition $\Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) = \operatorname{im} d_{A_0} \oplus \ker d^*_{A_0}$. If (ξ, a) is in the kernel, then by the Hodge decomposition, $a = 0$ and $d_A\xi = 0$, so $\xi \in T_1\mathrm{Stab}(A_0)$. This shows that $D_{(A,1)}\mathfrak{m}$ is injective. To show that $\mathfrak{m}$ is a diffeomorphism onto its image, we must show that it is injective. Thus we must show that if

$$g \cdot (A_0 + a) = A_0 + b, \text{ where } A_0 + a, A_0 + b \in \mathscr{S}_{A_0}(\varepsilon),$$

then $g \in \mathrm{Stab}(A_0)$. This is equivalent to

$$\begin{aligned} b &= gag^{-1} - (d_{A_0}g)g^{-1}, \\ d_{A_0}g &= ga - bg. \end{aligned}$$

It will now be useful to consider $\mathfrak{g}_{\mathrm{Ad}}$ and G_{Ad} as fixed subbundles of $\mathfrak{gl}(E)$ for some vector bundle E associated to a faithful representation. Thus g and a can be viewed as sections of $\mathfrak{gl}(E)$.

Now we use the identity

$$d_A^*(ga) = - \star d_A(g \star a) = - \star ((d_A g) \wedge \star a) - g \star d_A \star a = g d_A^* a - (d_A g) \cdot a.$$

In the last term, $d_A g \cdot a$ denotes the combination of inner product of one-forms and matrix multiplication in $\mathfrak{gl}(E)$.

Applying $d_{A_0}^*$, and using the fact that $d_{A_0}^* a = d_{A_0}^* b = 0$, we obtain

$$d_{A_0}^* d_{A_0} g = b \cdot d_A g - (d_A g) \cdot a.$$

Now use the Hodge decomposition of $\Omega^0(X; \mathfrak{gl}(E))$ to write $g = g_k + g_\perp$, where $d_{A_0} g_k = 0$. The condition $g \in \mathrm{Stab}(A_0)$ is equivalent to $d_{A_0} g = 0$, or $g_\perp = 0$. Towards proving that $g_\perp = 0$, take the $\mathfrak{gl}(E)$-inner product of both sides with $g_\perp$:

$$\begin{aligned} \|d_{A_0} g_\perp\|_{L^2}^2 &= \mathrm{Tr}(b \cdot d_A g_\perp \, g_\perp^* - (d_{A_0} g_\perp) \cdot a \, g_\perp^*) \\ &\leq \|b\|_{L^n} \|d_{A_0} g_\perp\|_{L^2} \|g_\perp\|_{L^r} + \|d_{A_0} g_\perp\|_{L^2} \|a\|_{L^n} \|g_\perp\|_{L^r}, \quad n^{-1} + 2^{-1} + r^{-1} = 1, \\ &= (\|a\|_{L^n} + \|b\|_{L^n}) \|g_\perp\|_{L^r} \|d_{A_0} g_\perp\|_{L^2} \\ &\leq 2\varepsilon C \sqrt{\|g_\perp\|_{L^2_{1,A_0}}^2} \, \|d_{A_0} g_\perp\|_{L^2}, \end{aligned}$$

where we used the fact that there is a continuous embedding $L^2_{1,A_0}(\mathfrak{gl}(E)) \hookrightarrow L^r(\mathfrak{gl}(E))$ with some constant C, by the Sobolev embedding theorem. Now

$$\|g_\perp\|_{L^2_{1,A_0}}^2 = \|g_\perp\|_{L^2}^2 + \|d_{A_0} g_\perp\|_{L^2}^2 = g_\perp \cdot (1 + d_{A_0}^* d_{A_0}) g_\perp,$$

and

$$(1 + d_{A_0}^* d_{A_0}) g_\perp = d_{A_0}^* d_{A_0} \left((d_{A_0}^* d_{A_0})^{-1} + 1 \right) g_\perp \leq \left(\lambda_1(A_0)^{-1} + 1 \right) d_{A_0}^* d_{A_0} g_\perp,$$

where $\lambda_1(A_0)$ denotes the smallest nonzero eigenvalue of $d_{A_0}^* d_{A_0}$. Thus

$$\|g_\perp\|_{L^2_{1,A_0}}^2 \leq \left(\lambda_1(A_0)^{-1} + 1 \right) \|d_{A_0} g_\perp\|_{L^2}^2.$$

It follows that

$$\|d_{A_0} g_\perp\|_{L^2}^2 \leq 2C\varepsilon \sqrt{\lambda_1(A_0)^{-1} + 1} \, \|d_{A_0} g_\perp\|_{L^2}^2.$$

If ε is chosen small enough so that $2C\varepsilon\sqrt{\lambda_1(A_0)^{-1} + 1} < 1$, then $d_{A_0} g_\perp = 0$. Thus $g_\perp \in \ker d_{A_0}$, so $g_\perp = 0$ as desired, and $\mathfrak{m}$ is injective with $\mathcal{S}_{A_0}(\varepsilon)$.

One may wonder whether the bound C for the map $L^2_{1,A_0}(E) \hookrightarrow L^r(E)$ for vector bundles necessarily depends on A_0. The answer is no. The constant C from functions $L^2_1 \hookrightarrow L^r$ suffices. In particular, for $s \in L^2_{1,A_0}(E)$,

$$\|s\|_{L^r} = \||s|\|_{L^r} \le C \||s|\|_{L^2_1} = C\sqrt{\|s\|^2_{L^2} + \|\nabla |s|\|^2_{L^2}} \le C \|s\|_{L^2_{1,A_0}} .$$

The last inequality follows directly from Kato's inequality

$$|\nabla |s|| \le |\nabla_{A_0} s| .$$

Wherever $s \neq 0$, we have

$$\nabla |s| = \nabla \sqrt{s \cdot s} = \frac{\nabla(s \cdot s)}{2\sqrt{s \cdot s}} = \frac{s}{|s|} \cdot \nabla_{A_0} s,$$

so

$$|\nabla |s|| \le \left|\frac{s}{|s|}\right| |\nabla_{A_0} s| = |\nabla_{A_0} s| .$$

To deal with the case where s has zeroes, one argues by some sort of approximation argument. For example, one can replace $|s|$ by $\sqrt{\varepsilon + s \cdot s}$ and send $\varepsilon \to 0$. Or one can approximate s by a sequence of smooth sections transverse to the zero section. That way, the zero set of s has measure zero and does not contribute to any L^p norm estimate.

The remaining ingredient to make this proof rigorous is to show that there are sensible Sobolev completions of $\mathscr{A}_P$ and $\mathscr{G}_P$ so that the completion of $\mathscr{G}_P$ acts on the completion of $\mathscr{A}_P$. For this, we need Sobolev multiplication.

5.8 Sobolev multiplication

Recall that there are embeddings

$$L^p_k \hookrightarrow \begin{cases} C^0 & \text{if } k/n - 1/p > 0, \\ L^{\frac{1}{1/p - k/n}} & \text{if } k/n - 1/p < 0, \\ L^r \ \forall r & \text{if } k/n - 1/p = 0. \end{cases}$$

In a nutshell, the idea of Sobolev multiplication is as follows. We want to understand when multiplication

$$\begin{aligned} L^p_k \times L^q_\ell &\to L^r_m \\ (f, g) &\mapsto fg \end{aligned}$$

is continuous. Derivatives don't magically appear, so we take $0 \le m \le \min(k, \ell)$. Thus we seek a bound on the L^r norm of $\nabla^m(fg)$. By the product rule, this expands as a sum

$$\nabla^m(fg) = (\nabla^m f)g + m \cdot (\nabla^{m-1} f) \otimes \nabla g + \cdots + f \nabla^m g.$$

We wish to show that each term of this sum is in L^r. It turns out that it suffices to check both end terms $(\nabla^m f)g$ and $f\nabla^m g$ are L^r. For the first term, $\nabla^m f \in L^p_{k-m}$ and $g \in L^q_\ell$. We apply the Sobolev embedding to each of $\nabla^m f$ and g to obtain functions which are in either C^0 or an L^p space. Then it is straightforward to check whether the product is bounded in L^r. The same must be checked for $f \in L^p_k$ and $\nabla^m g \in L^q_{\ell-m}$. If everything is within the appropriate ranges, then $L^p_k \times L^q_\ell \to L^r_m$ is continuous.

Sobolev multiplication below the borderline

Whenever $k/n - 1/p \leq 0$, the Sobolev space L^p_k contains discontinuous functions, and thus does not embed into C^0 or even L^∞.

Theorem 107. *Consider two Sobolev spaces L^p_k and L^q_ℓ such that* strict *inequality holds: $k/n - 1/p < 0$ and $\ell/n - 1/q < 0$. Then multiplication of functions extends to a continuous map of Sobolev spaces*

$$L^p_k \times L^q_\ell \to L^r_m$$

whenever $m \in \mathbb{Z}$ with $0 \leq m \leq \min(k, \ell)$, and r is such that

$$0 < m/n + (1/p - k/n) + (1/q - \ell/n) \leq 1/r \leq 1.$$

In other words, there exists a constant $C_{Xpqrk\ell m}$ (depending only on X and the Sobolev indices) such that

$$\|fg\|_{L^r_m} \leq C_{Xpqrk\ell m} \|f\|_{L^p_k} \|g\|_{L^q_\ell}.$$

Remark. Originally we defined the Sobolev space L^r_m for $m \in \mathbb{R}$ and $r \in [1, \infty]$ via the spectral decomposition of the Laplacian:

$$L^r_m := \left\{ \text{distributions } f \mid (1+\Delta)^{m/2} f \in L^r \right\}.$$

But here we will assume without proof that for $m \in \mathbb{Z}_{\geq 0}$, the L^r_m norm is equivalent to

$$\|f\|_{L^r_m} \sim \sum_{i=0}^{m} \left\| \nabla^i f \right\|_{L^r}.$$

Proof. To prove our desired estimate, it suffices (by the iterated product rule and triangle inequality) to find an estimate

$$\left\| (\nabla^a f)(\nabla^b g) \right\|_{L^r} \leq C_{Xpqrk\ell ab} \|f\|_{L^p_k} \|g\|_{L^q_\ell}$$

for each pair of integers $a, b \geq 0$ with $a + b \leq m$.

We have continuous maps

$$L^p_k \xrightarrow{\nabla^a} L^p_{k-a} \hookrightarrow L^{\frac{1}{1/p-(k-a)/n}}$$

and

$$L^q_\ell \xrightarrow{\nabla^b} L^q_{\ell-b} \hookrightarrow L^{\frac{1}{1/q-(\ell-b)/n}}.$$

By Hölder's inequality

$$\|fg\|_{L^{1/(u+v)}} \leq \|f\|_{L^{1/u}} \|g\|_{L^{1/v}},$$

so multiplication is continuous on

$$L^{1/u} \times L^{1/v} \to L^{1/(u+v)}.$$

(This assumes that $a + b \leq 1$ so that $L^{1/(a+b)}$ is still a Banach space.) Thus mutiplication is continuous on

$$L^{\frac{1}{1/p-(k-a)/n}} \times L^{\frac{1}{1/q-(\ell-b)/n}} \to L^{\frac{1}{-(m-a-b)/n+m/n+(1/p-k/n)+(1/q-\ell/n)}} \hookrightarrow L^r,$$

and composition of these maps gives a continuous map

$$L^p_k \times L^q_\ell \to L^r.$$

Our desired constant $C_{X^{pqr}_{k\ell ab}}$ is by definition the operator norm of this map. $\square$

Sobolev multiplication above the borderline

Theorem 108. *Suppose that $k/n - 1/p > 0$ and $L^p_k \hookrightarrow L^q_\ell$ (i.e. $k \geq \ell$ and $k/n - 1/p \geq \ell/n - 1/q$). If $\ell \in \mathbb{Z}_{\geq 0}$, then multiplication of functions extends to a continuous map of Sobolev spaces*

$$L^p_k \times L^q_\ell \to L^q_\ell.$$

Proof. As in the previous section, we want to check

$$\left\|(\nabla^a f)(\nabla^b g)\right\|_{L^q} \leq C_{Xpqk\ell ab} \|f\|_{L^p_k} \|g\|_{L^q_\ell}$$

for all nonnegative integers a, b such that $a + b \leq \ell$. We have to deal with the cases

$$L^p_k \xrightarrow{\nabla^a} L^p_{k-a} \hookrightarrow \begin{cases} C^0 & \text{if } (k-a)/n - 1/p > 0, \\ L^{\frac{1}{1/p-(k-a)/n}} & \text{if } (k-a)/n - 1/p < 0, \\ L^r \ \forall r & \text{if } (k-a)/n - 1/p = 0, \end{cases}$$

and

$$L^q_\ell \xrightarrow{\nabla^b} L^q_{\ell-b} \hookrightarrow \begin{cases} C^0 & \text{if } (\ell-b)/n - 1/q > 0, \\ L^{\frac{1}{1/q-(\ell-b)/n}} & \text{if } (\ell-b)/n - 1/q < 0, \\ L^r \ \forall r & \text{if } (\ell-b)/n - 1/q = 0. \end{cases}$$

□

- If both $(k-a)/n - 1/p < 0$ and $(\ell-b)/n - 1/q < 0$, then as in the previous proof,

  $$L^{\frac{1}{1/p-(k-a)/n}} \times L^{\frac{1}{1/q-(\ell-b)/n}} \hookrightarrow L^q$$

  as desired since

  $$1/p - (k-a)/n + 1/q - (\ell-b)/n \overset{a+b\leq\ell}{\leq} 1/p - k/n + 1/q \overset{k/n-1/p>0}{<} 1/q.$$

- Suppose instead that $(\ell-b)/n - 1/q < 0$ and $(k-a)/n - 1/p \geq 0$. Then

  $$L^q_{\ell-b} \hookrightarrow L^{\frac{1}{1/q-(\ell-b)/n}},$$

  and

  $$L^p_{k-a} \hookrightarrow C^0 \text{ or } L^r \ \forall r.$$

  - In the case $b = \ell$, we must have $a = 0$, so $(k-a)/n - 1/p = k/n - 1/p > 0$ and thus $L^p_{k-a} \hookrightarrow C^0$. Therefore,

    $$L^p_{k-a} \times L^q_{\ell-b} \hookrightarrow C^0 \times L^{\frac{1}{1/q-(\ell-b)/n}} = C^0 \times L^q \to L^q.$$

– If $b \neq \ell$, then $L^p_{k-a} \hookrightarrow L^r$ for any r, so

$$L^q_{\ell-b} \times L^p_{k-a} \to L^{\frac{1}{1/q-(\ell-b)/n+1/r}} = L^q \text{ for } r = n/(\ell-b).$$

- Suppose $(k-a)/n - 1/p < 0$ and $(\ell-b)/n - 1/q \geq 0$. Then

$$L^p_{k-a} \hookrightarrow L^{\frac{1}{1/p-(k-a)/n}} \hookrightarrow L^q$$

  because

$$1/p - (k-a)/n \leq 1/p - (k-\ell)/n \leq 1/q.$$

  For equality to hold, both $a = \ell$ and $k/n - 1/p = \ell/n - 1/q$, thus $b = 0$ and $(\ell-b)/n - 1/q = k/n - 1/p > 0$. Therefore, $L^q_{\ell-b} \hookrightarrow C^0$.

- Finally, suppose $(k-a)/n - 1/p \geq 0$ and $(\ell-b)/n - 1/q \geq 0$. Then choosing $r = 2q$, we obtain $L^p_{k-a} \hookrightarrow L^{2q}$ and $L^q_{\ell-b} \hookrightarrow L^{2q}$, so $L^p_{k-a} \times L^q_{\ell-b} \to L^q$.

Sobolev multiplication on the borderline

Theorem 109. *Suppose that $k/n - 1/p = 0$ and $L^p_k \hookrightarrow L^q_\ell$ (i.e. $k \geq \ell$ and $\ell/n - 1/q \leq 0$). If $\ell \in \mathbb{Z}_{\geq 0}$, then multiplication of functions extends to a continuous map of Sobolev spaces*

$$(L^p_k \cap L^\infty) \times (L^q_\ell \cap L^\infty) \to (L^q_\ell \cap L^\infty).$$

Furthermore, if the other function is below the borderline $\ell/n - 1/q < 0$, then we have the stronger result

$$(L^p_k \cap L^\infty) \times L^q_\ell \to L^q_\ell.$$

Proof. We will proceed by assuming $\ell/n - 1/q < 0$, proving continuity of the second multiplication, and realizing $\ell/n - 1/q = 0$ as an exceptional case.

Continuity of $(L^p_k \cap L^\infty) \times L^q_\ell \to L^q_\ell$ is equivalent to estimates of the form

$$\left\|(\nabla^a f)(\nabla^b g)\right\|_{L^q} \leq C_{Xpqabk\ell}\left(\|f\|_{L^p_k} + \|f\|_{L^\infty}\right)\|g\|_{L^q_\ell}$$

for all nonnegative integers a, b such that $a + b \leq \ell$. For the case $a = 0$ we use the L^∞, and for the case $a > 0$ we use the Sobolev embedding theorem to obtain

$$(L^p_k \cap L^\infty) \xrightarrow{\nabla^a} \begin{cases} L^\infty & \text{if } a = 0 \\ L^{\frac{1}{a/n}} & \text{if } a > 0 \end{cases} = L^{n/a}, \text{ where } n/0 := \infty.$$

In particular, for all a we have an estimate of the form

$$\|\nabla^a f\|_{L^{n/a}} \leq C\left(\|f\|_{L^p_k} + \|f\|_{L^\infty}\right).$$

Note that the $\|f\|_{L^\infty}$ term on the right is essential to cover the case $a = 0$ since $L^p_k \not\hookrightarrow L^\infty$.

For the other factor,

$$L^q_\ell \xrightarrow{\nabla^b} L^q_{\ell-b} \hookrightarrow \begin{cases} L^r \ \forall r < \infty & \text{if } b = 0 \text{ and } \ell/n - 1/q = 0, \\ L^{\frac{1}{1/q-(\ell-b)/n}} & \text{else.} \end{cases} \tag{5.6}$$

In the latter case when either $b \neq 0$ or $\ell/n - 1/q \neq 0$,

$$(L^p_k \cap L^\infty) \times L^q_\ell \hookrightarrow L^{\frac{1}{1/q-(\ell-a-b)/n}} \hookrightarrow L^q,$$

as desired. This proves continuity of

$$(L^p_k \cap L^\infty) \times L^q_\ell \to L^q_\ell \text{ when } \ell/n - 1/q < 0.$$

In the exceptional case where g is also borderline, so that $\ell/n - 1/q \neq 0$, we should take $g \in L^q_\ell \cap L^\infty$ so that we can effectively set $r = \infty$ in (5.6) and obtain a bound for $\nabla^b g$ in $L^{\frac{1}{1/q-(\ell-b)/n}} = L^{n/b}$ even when $b = 0$. This proves continuity of

$$(L^p_k \cap L^\infty) \times (L^q_\ell \cap L^\infty) \to L^q_\ell.$$

For the original claim

$$(L^p_k \cap L^\infty) \times (L^q_\ell \cap L^\infty) \to (L^q_\ell \cap L^\infty),$$

it remains to prove

$$(L^p_k \cap L^\infty) \times (L^q_\ell \cap L^\infty) \to L^\infty,$$

but this is obvious since

$$L^\infty \times L^\infty \to L^\infty.$$

□

Remark. These estimates can be generalized to non-integral Sobolev spaces by using interpolation theory.

Exercise 110. Suppose that $1 < p < \infty$ and $k \in \mathbb{Z}$ with $k \geq 1$ are such that $(k+1)/n - 1/p \geq 0$. (In particular, if $n = 4$ and $k = 1$, then $p \geq 2$.) Show that if $g \in \mathscr{G}^p_{k+1}$ and $A \in \mathscr{A}^p_k$, then

- $g \cdot A \in \mathscr{A}^p_k$,
- $F_A \in L^p_{k-1}(X; \Lambda^2 T^* X)$

5.9 Slice theorem

We wish to prove the existence of Coulomb gauge. For any smooth connection A_0, we wish to show that any "nearby" connection $A_0 + a$ is gauge-equivalent to $A_0 + b$, which is also nearby, and satisfies the Coulomb gauge condition $d^*_{A_0} b = 0$. Furthermore, both the gauge transformation and b are uniquely determined by a, up to a gauge transformation which stabilizes A_0.

More precisely, the idea is as follows. Suppose that X is a closed Riemannian n-manifold with $n = 4$, equipped with a principal G-bundle $P \to X$, where G is compact. Fix some smooth connection $A_0 \in \mathscr{A}$. We wish to study the space of $\mathscr{G}$-orbits in an L^2_1-neighborhood of A_0. Towards this goal, we seek an orthogonal complement to the tangent space of the $\mathscr{G}$-orbit at A_0:

Definition 111. The *slice* at a connection A_0 of radius δ is defined to be

$$\mathcal{S}_{A_0,\delta} := \left\{ a \in \Omega^1(X; \mathfrak{g}_{\mathrm{Ad}}) \mid d_{A_0}^* a = 0,\ \|a\|_{L^2_{1,A_0}} \leq \delta \right\}.$$

If δ is sufficiently small, in a manner which depends on A_0, then the claim is that the left quotient of $\mathcal{S}_{A_0,\delta}$ by $\mathrm{Stab}(A_0)$ serves as a chart for $\mathcal{B} = \mathcal{A}/\mathcal{G}$ around $[A_0]$ in the following sense. Consider the map

$$\mathfrak{m}_{A_0} : \frac{\mathcal{G} \times \mathcal{S}_{A_0,\delta}}{\mathrm{Stab}(A_0)} \to \mathcal{A}$$

given by

$$\mathfrak{m}_{A_0}(g, a) := g \cdot (A_0 + a) = A_0 + a - (\nabla_{A_0+a} g) g^{-1} = A_0 + gag^{-1} - (\nabla_{A_0} g) g^{-1},$$

and where $g_0 \in \mathrm{Stab}(A_0)$ acts on $\mathcal{S}_{A_0,\delta}$ by the adjoint action $a \mapsto g_0 a g_0^{-1}$, and on $\mathcal{G}$ by $g \mapsto g g_0^{-1}$.

Theorem 112. *For every smooth connection $A_0 \in \mathcal{A}$, there exists a δ depending on A_0 such that in the specified Sobolev completions, $\mathfrak{m}_{A_0,\delta}$ is a homeomorphism onto its image:*

$$\mathfrak{m}_{A_0} : \frac{\mathcal{G}_{L^2_2} \times \mathcal{S}_{A_0,\delta,L^2_1}}{\mathrm{Stab}(A_0)} \to \mathcal{A}_{L^2_1}.$$

The first thing to check is that the definition of $\mathfrak{m}_{A_0}$ makes sense: namely that $\mathrm{Stab}(A_0)$ actually maps $\mathcal{S}_{A_0,\delta}$ to itself, and that $\mathfrak{m}_{A_0}$ is invariant under the action of $\mathrm{Stab}(A_0)$. This is simple to verify:

Lemma 113. *If $g_0 \in \mathrm{Stab}(A_0)$, and $a \in \mathcal{S}_{A_0,\delta}$, then $g_0 a g_0^{-1} \in \mathcal{S}_{A_0,\delta}$ and $g g_0^{-1} \cdot (A_0 + g_0 a g_0^{-1}) = g \cdot (A_0 + a)$.*

Proof. Suppose $g \in \mathrm{Stab}(A_0)$. This is equivalent to $\nabla_{A_0} g = 0$. In order to verify that $g \cdot (A_0 + a) \in \mathcal{S}$, we must check the two defining conditions from Definition 111 for $b \in \mathcal{S}$, where $g \cdot (A_0 + a) = A_0 + b$. Since $g \in \mathrm{Stab}(A_0)$ is equivalent to $\nabla_{A_0} g = 0$, it follows that $b = gag^{-1}$. To check that $d_{A_0}^* b = 0$, use $d_{A_0}^* = -\iota_i \nabla_{A_0,i}$, and the product rule for ∇_{A_0} to conclude that

$$d_{A_0}^*(gag^{-1}) = g(d_{A_0}^* a) g^{-1} = 0.$$

Finally, note that $\|b\|_{L^2_1}^2 = \left\| gag^{-1} \right\|_{L^2}^2 + \left\| g(\nabla_{A_0} a) g^{-1} \right\|_{L^2}^2 = \|a\|_{L^2_1}^2 \leq \delta_1$. These two conditions verify that $b \in \mathcal{S}$. ☐

For the proof of Theorem 112, we would like to argue that $\mathfrak{m}_{A_0,\delta_1}$ is a diffeomorphism onto its image. We would use the inverse function theorem to prove that $\mathfrak{m}_{A_0,\delta_1}$ is a local diffeomorphism for δ_1 sufficiently small. Then we would prove that $\mathfrak{m}_{A_0,\delta_1}$ is also injective for δ_1 sufficiently small. We would be done since any injective local diffeomorphism is an actual diffeomorphism onto its image. Unfortunately we cannot make this argument directly, since $\mathcal{G}_{L^2_2}$ is not a smooth Lie group in the L^2_2 topology, as we explain in the next section.

5.9.1 Smoothness problems with the borderline Lie groups

Suppose $P \to X$ is a principal $\mathrm{O}(k)$ bundle with standard vector bundle $E \to X$ and endomorphism bundle $\mathfrak{gl}(E) \to X$. Then $\mathscr{G}_P$ can be identified with sections of $O(E) \subset \mathfrak{gl}(E)$. Specifically, a section $s \in \mathscr{G}_P$ is the same as a section $s \in \Gamma(X; \mathfrak{gl}(E))$ which satisfies the equation $ss^T = \mathrm{Id}_E \in \Gamma(X; \mathrm{Sym}(E))$, where $\mathrm{Sym}(E)$ denotes the subbundle of $\mathfrak{gl}(E)$ consisting of the endomorphisms of E which are symmetric (with respect to the metric on E).

We would like to be able to take the completion of $\mathscr{G}_P$ with respect to a Banach space topology, and show that the completion is a Banach Lie group. It suffices to show that Id_E is a regular value of the map $F : \Gamma(X; \mathfrak{gl}(E)) \to \Gamma(X; \mathrm{Sym}(E))$ given by $F(s) := ss^T$, and that this map is smooth. The derivatives are $d_sF(t) = ts^T + st^T$, $d_s^2F(t,u) = tu^T + ut^T$, and $d_s^kF = 0$ for $k \geq 3$. Smoothness is guaranteed so long as the Banach space is an algebra under multiplication. Thus we should take an L^p_k topology which is above the borderline, or the $L^p_k \cap L^\infty$ topology on the borderline.

To show that Id_E is a regular value, we must show that d_sF is surjective for any s in the appropriate completion of $\Gamma(X; \mathfrak{gl}(E))$ such that $ss^T = \mathrm{Id}_E$. Surjectivity follows from observing that $d_sF(\frac{1}{2}vs) = v$ for any $v \in \Gamma(X; \mathrm{Sym}(E))$. This shows that in the borderline case, $\mathscr{G}_{P,L^p_k}$ is a smooth Banach Lie group in the $L^p_k \cap L^\infty$ topology.

In this borderline case, the Lie algebra consists of $L^p_k \cap L^\infty$ sections of $\mathfrak{o}(E)$. This is problematic for the Hodge decomposition

$$\Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_{k-1}} = \ker d^*_{A_0} \oplus d_{A_0}\left(\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_k}\right).$$

Specifically, we want to interpret the second factor as the tangent space of the $\mathscr{G}_{P,L^p_k}$ orbit, and $\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_k}$ as the Lie algebra of $\mathscr{G}_{P,L^p_k}$. Then $\ker d^*_{A_0}$ is a complementary subspace, so it can be used as the tangent space to the slice. However, $L^p_k \cap L^\infty$ is missing functions such as $\ln\ln(\exp(1)/|x|)$, which are in L^p_k but not L^∞. Thus

$$\Omega^1(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_{k-1}} \neq \ker d^*_{A_0} \oplus d_{A_0}\left(\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_k \cap L^\infty}\right).$$

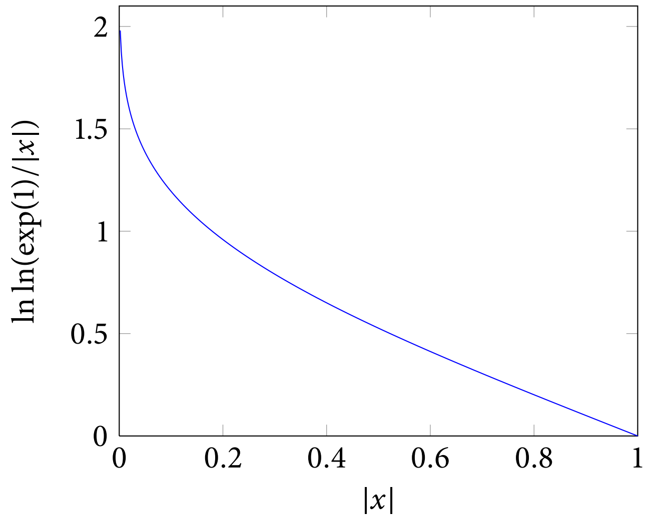
Specifically, while

$$g(x) = \exp\left(\ln\ln(\exp(1)/|x|)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right),$$

suitably extended from local coordinates, is a perfectly fine element of the borderline $\mathscr{G}_{P,L^p_k \cap L^\infty}$, the corresponding direction tangent to the orbit given by

$$(d_{A_0}g)g^{-1} = d_{A_0}\left(\ln\ln(\exp(1)/|x|)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)$$

is missing from $d_{A_0}\left(\Omega^0(X; \mathfrak{g}_{\mathrm{Ad}})_{L^p_k \cap L^\infty}\right)$. The moral is that while elliptic operators behave nicely between Sobolev spaces, they fail to behave nicely on other spaces such as C^k or $L^p_k \cap L^\infty$.



5.9.2 Calculus on small slices

While we cannot directly argue via the inverse function theorem for the L^2_1 completion of the slice $\mathscr{S}_{A_0,\delta}$ due to the borderline issues of the previous section, we can use differential topological arguments above the borderline. We can then transfer these arguments to the borderline via a combination of the continuity method and density arguments.

Consider the map

$$\mathfrak{m}_{A_0,\delta} : \frac{\mathscr{G}_{L^2_{5/2}} \times \mathscr{S}_{A_0,\delta,L^2_{3/2}}}{\mathrm{Stab}(A_0)} \to \mathscr{A}_{L^2_{3/2}}. \tag{5.7}$$

Theorem 114. *For δ sufficiently small, the map $\mathfrak{m}_{A_0,\delta}$ is a diffeomorphism onto its image. Furthermore, the image contains the $L^2_{3/2}$ points of some L^2_1 ball around A_0.*

The strategy of the proof will be to first show that $d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}$ is an isomorphism. Next we will show that $d_{(\mathrm{Id},a)}\mathfrak{m}_{A_0,\delta}$ is an isomorphism for all $a \in \mathscr{S}_{A_0,\delta,L^2_{3/2}}$ when δ is sufficiently small. By $\mathscr{G}$-equivariance, this establishes that $\mathfrak{m}_{A_0,\delta}$ is a local diffeomorphism. Then we will show that $\mathfrak{m}_{A_0,\delta}$ is injective. Since an injective local diffeomorphism is a genuine diffeomorphism, it remains only to show that the image contains an L^2_1 ball. For this we use a continuity argument.

Definition 115. Let $B_{A_0,\varepsilon,L^2_{3/2}}$ denote the $L^2_{3/2}$ points of an L^2_1 ball of radius ε around A_0, equipped with the $L^2_{3/2}$ topology.

Since $B_{A_0,\varepsilon,L^2_{3/2}}$ is contractible, it consists of a single connected component. Thus in order to show that the image of $\mathfrak{m}_{A_0,\delta}$ contains $B_{A_0,\varepsilon,L^2_{3/2}}$, it suffices to show that the image of $\mathfrak{m}_{A_0,\delta}$ is open, closed, and nonempty in $B_{A_0,\varepsilon,L^2_{3/2}}$ with respect to the $L^2_{3/2}$ topology.

Lemma 116. *The map $d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}$, under the Sobolev completion specified in* (5.7)*, is an isomorphism.*

Proof. This essentially follows from the Hodge theorem. We compute

$$d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}(\xi,\alpha) = -d_{A_0}\xi + \alpha.$$

The Hodge decomposition gives

$$\Omega^1(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{3/2}} = d_{A_0}\Omega^0(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{5/2}} \oplus \ker\left(\Omega^1(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{3/2}} \overset{d^*_{A_0}}{\to} \Omega^0(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{1/2}}\right).$$

Now $-d_{A_0}\xi$ is a general element of the first factor, and α is a general element of the second factor, hence $d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}$ is surjective. The kernel of $d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}$ consists of $\xi \in \mathrm{Lie}(\mathrm{Stab}(A_0))$. Thus $d_{(\mathrm{Id},0)}\mathfrak{m}_{A_0,\delta}$ is an isomorphism. □

Lemma 117. *If δ sufficiently small, then $d_{(\mathrm{Id},a)}\mathfrak{m}_{A_0,\delta}$, under the Sobolev completion specified in* (5.7), *is an isomorphism for all $a \in \mathcal{S}_{A_0,\delta,L^2_{3/2}}$.*

Proof. We compute

$$d_{(\mathrm{Id},a)}\mathfrak{m}_{A_0,\delta}(\xi,\alpha) = -d_{A_0+a}\xi + \alpha = -d_{A_0}\xi - [a,\xi] + \alpha.$$

We would like to say that $[a,\xi]$ is an $L^2_{3/2}$-small perturbation when $\|a\|_{L^2_1} \le \delta$. This would imply that $d_{(\mathrm{Id},a)}\mathfrak{m}_{A_0,\delta}$ remains an isomorphism under perturbation. However this is not true since we cannot bound an $L^2_{3/2}$ norm in terms of the lower regularity of an L^2_1 bound on a factor. Thus we need a trick. The trick is that by the Hodge decomposition, it suffices to bound just the part which lies in the $\mathrm{im}(d_{A_0})$ component of the Hodge decomposition, since the $\ker(d^*_{A_0})$ component is automatically surjective, regardless of how big the perturbation is. The $\mathrm{im}\,(d_{A_0})$ component is orthogonal to $\ker d^*_{A_0}$, so

$$\left\|[a,\xi]_{\mathrm{im}(d_{A_0})}\right\|_{L^2_{3/2}} \le C\left\|d^*_{A_0}\,[a,\xi]\right\|_{L^2_{1/2}} = C\left\|\left[d^*_{A_0}a,\xi\right] - [a\cdot d_{A_0}\xi]\right\|_{L^2_{1/2}} = C\,\|[a\cdot d_{A_0}\xi]\|_{L^2_{1/2}} \le C\,\|a\|_{L^2_1}\,\|\xi\|_{L^2_{3/2}} \le \delta C$$

Thus for δ sufficiently small, the map $\Omega^0(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{5/2}} \to d_{A_0}\Omega^0(X;\mathfrak{g}_{\mathrm{Ad}})_{L^2_{5/2}}$ given by $\xi \mapsto -d_{A_0}\xi - [a,\xi]_{\mathrm{im}(d_{A_0})}$ is an isomorphism. □

Corollary 118. *If δ is sufficiently small, then the image of $\mathfrak{m}_{A_0,\delta}$ is open in $B_{A_0,\varepsilon,L^2_{3/2}}$ with respect to the $L^2_{3/2}$ topology.*

Lemma 119. *If δ is sufficiently small, then $\mathfrak{m}_{A_0,\delta}$, under the Sobolev completion specified in* (5.7), *is injective.*

Proof. We must show that if $a,b \in \mathcal{S}_{A_0,\delta,L^2_{3/2}}$ with $g\cdot(A_0+a) = A_0+b$, then $g \in \mathrm{Stab}_{A_0}$, i.e. $d_{A_0}g = 0$. We want to prove this using Sobolev estimates. The strategy is as follows. We have a Hodge decomposition of

$$g \in \Omega^0(X;\mathrm{End}\,E) \xrightarrow{d_{A_0}} \Omega^1(X;\mathrm{End}\,E)$$
$$g = g_0 + g_1 \in \ker(d_{A_0}) \oplus \ker(d_{A_0})^{\perp}.$$

$$d_{A_0} g = 0 \iff g_1 = 0 \iff d_A g_1 = 0 \iff \|d_{A_0} g_1\|_{L^2}^2 \le C \|d_{A_0} g_1\|_{L^2}^2 \text{ for } C < 1.$$

Recall that

$$g \cdot (A_0 + a) = A_0 + b \iff d_{A_0} g = ga - bg.$$

In four dimensions, $d_A^* = - \star d_A \star$, so recalling that $d_{A_0}^* a = 0 = d_{A_0}^* b$, we have

$$\begin{aligned} d_{A_0}^* d_{A_0} g &= - \star d_{A_0} \left(g \star a - (\star b) g \right) \\ &= - \star \left((d_{A_0} g) \wedge \star a + (\star b) d_{A_0} g \right) . \end{aligned}$$

Using the fact that $d_{A_0} g = d_{A_0} g_1$, take the inner product with g_1 to get

$$\begin{aligned} \|d_{A_0} g_1\|_{L^2}^2 &\le \|g_1\|_{L^4} \left(\|a\|_{L^4} + \|b\|_{L^4} \right) \|d_{A_0} g_1\|_{L^2} \\ &\le C_{L^2_{1,A_0}(\mathfrak{gl}(E)) \hookrightarrow L^4} \|g_1\|_{L^2_{1,A_0}} \left(\|a\|_{L^4} + \|b\|_{L^4} \right) \|d_{A_0} g_1\|_{L^2} . \end{aligned}$$

We want to replace $\|g_1\|_{L^2_{1,A_0}}$ with $\|d_{A_0} g_1\|_{L^2}$. Recall that

$$\|g_1\|_{L^2_{1,A_0}}^2 = \|g_1\|_{L^2}^2 + \|d_A g_1\|_{L^2}^2 .$$

Since $g_1 \perp \ker d_{A_0}$, we have a Poincare inequality

$$\|d_{A_0} g_1\|_{L^2}^2 = \left\langle g_1, d_{A_0}^* d_{A_0} g_1 \right\rangle \ge \lambda_1(A_0) \|g_1\|_{L^2}^2 ,$$

where $\lambda_1(A_0)$ is the first nonzero eigenvalue of the Laplacian $d_A^* d_A$ on $\Omega^0(X; \mathfrak{gl}(E))$. Thus

$$\|g_1\|_{L^2_{1,A_0}}^2 \le \left(1 + \lambda_1(A_0)^{-1} \right) \|d_{A_0} g_1\|_{L^2}^2 .$$

Consequently,

$$\|d_{A_0} g_1\|_{L^2}^2 \le C \sqrt{(1 + \lambda_1(A_0)^{-1})} \left(\|a\|_{L^4} + \|b\|_{L^4} \right) \|d_{A_0} g_1\|_{L^2}^2 .$$

Thus, as long as

$$\|a\|_{L^4}, \|b\|_{L^4} < \frac{1}{2 C_{L^2_{1,A_0}(\mathrm{End}\, E) \hookrightarrow L^4} \sqrt{(1 + \lambda_1(A_0)^{-1})}},$$

we have injectivity. Recall that $L^2_1 \hookrightarrow L^4$. Thus if $\|a\|_{L^2_1}, \|b\|_{L^2_1} < \delta$ for δ sufficiently small, then the above inequality holds, so $\mathfrak{m}_{A_0,\delta}$ is injective on the slice. □

5.9.2.1 Gauge equivalence is preserved under weak L^2_1 limits

Lemma 120. *Suppose $a_i \overset{L^2_1}{\rightharpoonup} a$, $b_i \overset{L^2_1}{\rightharpoonup} b$, and there exist $g_i \in \mathscr{G}_{L^2_2}$ such that $g_i \cdot (A_0 + a_i) = A_0 + b_i$. Then there exists a subsequence such that $g_i \overset{L^2_2}{\rightharpoonup} g \in \mathscr{G}_{L^2_2}$ and $g \cdot (A_0 + a) = b$.*

Proof. From $a_i \overset{L^2_1}{\rightharpoonup} a$, $b_i \overset{L^2_1}{\rightharpoonup} b$ it follows that a_i and b_i are bounded in L^2_1 and hence also in L^4. Now from

$$d_{A_0} g_i = g_i a_i - b_i g_i$$

it follows that $\|d_{A_0} g_i\|_{L^4} \leq \|g_i\|_{L^\infty} (\|a_i\|_{L^4} + \|b_i\|_{L^4})$. Thus g_i is bounded in $L^4_{1,A_0} \cap L^\infty$. Applying this bound to the same inequality,

$$\|d_{A_0} g_i\|_{L^2_{1,A_0}} \leq C \|g_i\|_{L^4_{1,A_0} \cap L^\infty} \left(\|a_i\|_{L^2_{1,A_0}} + \|b_i\|_{L^2_{1,A_0}} \right).$$

Thus g_i is L^2_{2,A_0}-bounded. Therefore, passing to a subsequence, $g_i \overset{L^2_2}{\rightharpoonup} g$. We must verify that $g \in \mathscr{G}_{L^2_2}$. Note that $L^2_2 \hookrightarrow L^p$ compactly for all $1 \leq p < \infty$. Choosing any fixed p and passing to a further subsequence, $g_i \to g$ pointwise almost everywhere. Thus $g(x) \in G_{\mathrm{Ad}} |_x \subset \mathfrak{gl}(E)$ for almost all x, and so $g \in \mathscr{G}_{L^2_2}$.

It remains to show that $g \cdot (A_0 + a) = b$. This is equivalent to $d_{A_0} g = ga - bg$. Note that by compactness of the embeddings $L^2_1 \hookrightarrow L^2$ and $L^2_2 \hookrightarrow L^2$, on passing to a subsequence, we have strong L^2 convergence $g_i \to g$, $a_i \to a$, and $b_i \to b$. It follows that that $g_i a_i - b_i g_i \to ga - bg$ in L^1 since

$$\begin{aligned}
\|g_i a_i - b_i g_i - (ga - bg)\|_{L^1} &\leq \|g_i a_i - g a_i\|_{L^1} + \|g a_i - ga\|_{L^1} + \|b_i g_i - b_i g\|_{L^1} + \|b_i g - bg\|_{L^1} \\
&\leq \|g_i - g\|_{L^2} \|a_i\|_{L^2} + \|g\|_{L^2} \|a_i - a\|_{L^2} + \|b_i\|_{L^2} \|g_i - g\|_{L^2} + \|b_i - b\|_{L^2} \|g\|_{L^2}.
\end{aligned}$$

Also, by compactness of the embedding $L^2_1 \hookrightarrow L^1$ it follows that after passing to a subsequence, $d_{A_0} g_i \overset{L^1}{\to} d_{A_0} g$. Thus by uniqueness of limits in L^1, it follows by taking the L^1 limit of both sides of $d_{A_0} g_i = g_i a_i - b_i g_i$ that $d_{A_0} g = ga - bg$. □

Theorem 121. *If δ is sufficiently small, then the image of $\mathfrak{m}_{A_0,\delta}$ is closed in $B_{A_0,\varepsilon,L^2_{3/2}}$ with respect to the $L^2_{3/2}$ topology.*

Proof. Suppose $A_i = A_0 + a_i$ is a sequence in the image of $\mathfrak{m}_{A_0,\delta}$, and $a_i \overset{L^2_{3/2}}{\to} a$. By virtue of being in the image of $\mathfrak{m}_{A_0,\delta}$, it follows that there exist gauge transformations $g_i \in \mathscr{G}_{L^2_{5/2}}$ such that $g_i(A_0 + a_i) = A_0 + b_i$ with $b_i \in \mathscr{S}_{A_0,\delta,L^2_{3/2}}$, i.e. $d^*_{A_0} b_i = 0$ and $\|b_i\|_{L^2_1} \leq \delta$. It follows that for some subsequence, $b_i \overset{L^2_1}{\rightharpoonup} b$. From Lemma 120, after passing to a further subsequence, there exists $g \in \mathscr{G}_{L^2_2}$ such that $g \cdot (A_0 + a) = A_0 + b$. To show that $A_0 + a$ is in the image of $\mathfrak{m}_{A_0,\delta}$, we must show that $g \in \mathscr{G}_{L^2_{5/2}}$. We compute

$$\begin{aligned}
d^*_{A_0} d_{A_0} g &= - \star d_{A_0} \left(g \star a - (\star b) g\right) \\
&= -(d_{A_0} g) \cdot a + g d^*_{A_0} a + b \cdot d_{A_0} g.
\end{aligned}$$

The first two terms on the right are bounded in $L^2_{1/2}$, however the last term is only bounded in $L^{8/5}_{1/2}$. This allows us to conclude that $g \in L^{8/5}_{5/2}$, which is still borderline. Continuing to bootstrap in this manner does not allow one to show that g is above the borderline. We need another trick.

Consider the operator $L_b : L^2_{5/2} \to L^2_{1/2}$ given by $L_b(g) := d^*_{A_0} d_{A_0} g - b \cdot d_{A_0} g$. This operator is elliptic, possibly with kernel. The equation is

$$L_b g = -(d_{A_0} g) \cdot a + g d^*_{A_0} a,$$

and the right hand side is bounded in $L^2_{1/2}$. This allows us to conclude that $g \in L^2_{5/2}$. However, note that we can't conclude any bound on $\|g\|_{L^2_{5/2}}$ unless for some reason we know that g is orthogonal to the kernel of L_b.

□

5.10 Theta functions

Let V denote a real vector space of dimension n, equipped with a symmetric bilinear form $Q : V \otimes V \to \mathbb{R}$.

Note that Q induces a map, which we denote also by $Q : V \to V^*$ given by $v \mapsto Q(v, -)$.

Definition 122. A symmetric bilinear form Q is said to be *non-degenerate* if $Q : V \to V^*$ is invertible. In this case, we denote by Q^{-1} both the inverse $Q^{-1} : V^* \to V$ and the corresponding symmetric bilinear form $Q^{-1} : V^* \otimes V^* \to \mathbb{R}$ given by $Q^{-1}(\alpha, \beta) := Q(Q^{-1}(\alpha), Q^{-1}(\beta))$.

Remark 123. The bilinear form Q^{-1} satisfies $Q^{-1}(\alpha, \beta) = \alpha(Q^{-1}(\beta)) = \beta(Q^{-1}(\alpha))$.

Definition 124. A *lattice* (L, Q) *of rank* n is a free abelian group L of rank n equipped with a symmetric bilinear form $Q : L \otimes L \to \mathbb{R}$.

Remark 125. The set of isomorphisms $\mathbb{Z}^n \to L$ is a $\mathrm{GL}(n; \mathbb{Z})$-torsor, where $\mathrm{GL}(n; \mathbb{Z})$ denotes the set of $n \times n$ integer matrices with determinant ± 1.

Remark 126. One often considers L as a discrete subgroup of maximal rank inside some real vector space V with a compatible inner product. The canonical choice of V is $V = L \otimes_{\mathbb{Z}} \mathbb{R}$, and there is a natural $\mathbb{R}$-bilinear extension of Q from L to V. When Q is positive-definite, (L, Q) embeds into Euclidean $\mathbb{R}^n$, uniquely up to an orthogonal transformation.

Definition 127. If (L, Q) is a lattice with Q non-degenerate, then the *dual lattice* is the non-degenerate lattice (L^*, Q^{-1}), where $L^* := \mathrm{Hom}(L, \mathbb{Z})$.

Remark 128. The double-dual of (L, Q) is canonically isomorphic to (L, Q) induced by the evaluation map $\mathrm{ev} : L \to L^{**}$ given by $v \mapsto (\alpha \mapsto \alpha(v))$, and $Q^{-1^{-1}} = Q$.

Remark 129. An invariant volume measure on V corresponds to an element of $|\Lambda^n V^*|$, which transforms according to the representation $\mathrm{GL}(n) \to \mathrm{GL}(1)$ given by $g \mapsto |\det g|$. There is a duality map $|\Lambda^n V^*| \otimes |\Lambda^n V| \to \mathbb{R}$ which associates to any invariant volume measure μ on V the invariant volume measure μ^{-1} on V^*.

Definition 130. The *standard volume measure* of a non-degenerate lattice (L, Q) is the positive measure denoted by $\sqrt{\det Q}$, such that the hypercube of an orthonormal basis in V has unit volume.

Theorem 131 (Poisson summation formula)**.** *As distributions on the vector space V, dual to the space of Schwarz functions $\mathcal{S}(V)$, written as a function of the variable $v \in V$, we have the identity*

$$\sum_{x \in L} \delta(v - x) = \frac{\mu}{\mu(V/L)} \sum_{\alpha \in L^*} e^{2\pi i \alpha(v)},$$

where μ denotes an arbitrary invariant positive measure on V, and $\mu/\mu(V/L)$ is the unique invariant measure on V which is normalized so that the torus V/L has unit measure.

Remark 132. This formula amounts to the statement that the Fourier transformation of the "Dirac comb" supported on L is the Dirac comb of L^*. Since the Dirac comb is periodic, it should be given by a Fourier series, and the coefficients of this Fourier series, when suitably normalized, are all 1.

Proof. Suppose $\phi \in \mathcal{S}(V)$ is a test function. Then we must show that the distributions evaluate to the same value on ϕ. We introduce the "periodicized" function

$$\phi_P(v) := \sum_{x \in L} \phi(x + v) \in C^\infty(V/L).$$

This sum is absolutely convergent since ϕ decays rapidly. Note that

$$\phi_P(0) = \int \sum_{x \in L} \delta(v - x)\phi(v),$$

which is the left-hand side of the identity. It suffices to show that $\phi_P(0) = \int \sum_{\alpha \in L^*} e^{2\pi i \alpha(v)} \phi(v)\, \mu/\mu(V/L)$. For this, we expand $\phi_P(v)$ in a Fourier series $\phi_P(v) = \sum_{\alpha \in L^*} c_\alpha e^{-2\pi i \alpha(v)}$. To compute the c_α, take F to be a fundamental domain for V/L. Then

$$\begin{aligned}
c_\alpha &= \int_F e^{2\pi i \alpha(v)} \phi_P(v)\, \mu/\mu(V/L) \\
&= \sum_{x \in L} \int_F e^{2\pi i \alpha(v)} \phi(v + x)\, \mu/\mu(V/L) \\
&= \sum_{x \in L} \int_{F - x} e^{2\pi i \alpha(v)} \phi(v)\, \mu/\mu(V/L) \\
&= \int_V e^{2\pi i \alpha(v)} \phi(v)\, \mu/\mu(V/L).
\end{aligned}$$

Plugging in $v = 0$ into the Fourier transform gives

$$\phi_P(0) = \sum_{\alpha \in L^*} c_\alpha = \int_V \sum_{\alpha \in L^*} e^{2\pi i \alpha(v)} \phi(v)\, \mu/\mu(V/L),$$

which agrees with the right hand side. □

Remark 133. It will also be useful to have the shifted version of Poisson summation, corresponding to the substitution $v \mapsto v - x_0$, which yields

$$\sum_{x \in L + x_0} \delta(v - x) = \frac{\mu}{\mu(V/L)} \sum_{\alpha \in L^*} e^{-2\pi i \alpha(x_0)} e^{2\pi i \alpha(v)}.$$

We also have the dual version

$$\sum_{\alpha\in L^*+\alpha_0} \delta(k-\alpha) = \mu^{-1}\mu(V/L)\sum_{x\in L} e^{-2\pi i\alpha_0(x)}e^{2\pi i k(x)}.$$

Definition 134. A lattice (L, Q) is *positive-definite* if $Q(x, x) > 0$ for all nonzero $x \in L$.

Definition 135. A lattice (L, Q) is *integral* if $Q : L \otimes L \to \mathbb{Z}$.

Lemma 136. *(L, Q) is integral iff $Q(L) \subset L^*$.*

Definition 137. A lattice (L, Q) is *unimodular* if $Q(L) = L^*$.

Remark 138. If (L, Q) is unimodular, then it is integral and non-degenerate.

Remark 139. Suppose $\{e_1, \ldots, e_n\}$ is a basis for L. Let $\underline{Q}$ denote the matrix for Q in this basis. Then Q is non-degenerate iff $\det \underline{Q} \neq 0$. The matrix of Q^{-1} in the dual basis is given by $\underline{Q}^{-1}$. Furthermore, (L, Q) is integral iff $\underline{Q}$ has integer entries. When (L, Q) is integral, the index of $Q(L) \subset L^*$ is the absolute value of $\det \underline{Q}$. Finally, Q is unimodular iff $\underline{Q}$ both has integer entries and $\det \underline{Q} = \pm 1$.

Example 140. The triangular lattice A_2 has rank 2, and in some basis, $\underline{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. We can take $V = \mathbb{R}^2$ with $Q = dx^2 + dy^2$, and $e_1 = (\sqrt{2}, 0)$, $e_2 = (-\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}})$. The dual basis is $e^1 = (\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{6}})$, $e^2 = (0, \sqrt{\frac{2}{3}})$. The components of $Q(e_i)$ in the dual basis are the same as the components of e_i. We have $Q(e_1) = 2e^1 - e^2$, $Q(e_2) = -e^1 + 2e^2$, $\det \underline{Q} = 3$, and $Q^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Taking $\mu = \sqrt{\det Q} = |dx \wedge dy|$, we have $\mu(\mathbb{R}^2/A_2) = \mu(e_1, e_2) = \sqrt{3}$, $\mu^{-1} = \sqrt{\det Q^{-1}} = \left|\partial_x \wedge \partial_y\right|$, and $\mu^{-1}((\mathbb{R}^2)^*/A_2^*) = \mu^{-1}(e^1, e^2) = \sqrt{\frac{1}{3}} = 1/\mu(\mathbb{R}^2/A_2)$.

Definition 141. The *theta function* of a positive-definite lattice (L, Q) is the function

$$\theta_L(\tau) := \sum_{x\in L} e^{(2\pi i\tau)\frac{1}{2}Q(x,x)}.$$

Remark 142. θ_L converges on the upper-half-plane. Under the change of variables $\tau = it$, it becomes the real-valued function

$$\theta_L(t) := \sum_{x\in L} e^{-\pi t Q(x,x)},$$

defined for $\Re(t) > 0$. Note the abuse of notation, since the function differs depending on the variable used. If (L, Q) is integral, then under the change of variables $q = e^{\pi i\tau}$, it becomes

$$\theta_L(q) := \sum_{x\in L} q^{Q(x,x)} = \sum_{n\in\mathbb{Z}_{\geq 0}} \#\{x \in L \mid Q(x, x) = n\}\, q^n.$$

Note: $\theta_{\mathbb{Z}}(q) = 1 + 2q + 2q^4 + 2q^9 + 2q^{16} + \cdots$. Furthermore, $\theta_{L_1\oplus L_2} = \theta_{L_1}\theta_{L_2}$. In particular, $\theta_{\mathbb{Z}^n}(q) = \theta_{\mathbb{Z}}(q)^n$.

Lemma 143. *If (L, Q) is an integral and positive-definite lattice, then $\theta_L(\tau+2) = \theta_L(\tau)$.*

Proof. This is a simple computation: $\theta_L(\tau+2) = \sum_{x\in L} e^{(2\pi i\tau)\frac{1}{2}Q(x,x)+2\pi i Q(x,x)} = \theta_L(\tau)$. □

Lemma 144. *If (L, Q) is a positive-definite lattice (integral or not), then $\theta_{L^*}(-1/\tau) = \sqrt{\det Q}(V/L)\,(\tau/i)^{n/2}\theta_L(\tau)$, where $n = \operatorname{rank}(L)$.*

Proof. For simplicity consider $t = \tau/i$, so that $\Re(t) > 0$, and consider the double-integral

$$\int_{k\in V^*}\sum_{\alpha\in L^*}\delta(k-\alpha)\int_{v\in V}\sqrt{\det(tQ)}e^{-2\pi i k(v)-2\pi t\cdot\frac{1}{2}Q(v,v)}.$$

$$\int_{v\in V}\delta_L(v)\int_{k\in V^*}\sqrt{\det(tQ)^{-1}}e^{-2\pi i k(v)-2\pi t^{-1}\cdot\frac{1}{2}Q^{-1}(k,k)}.$$

Note that

$$1 = \int_V e^{-2\pi\cdot\frac{1}{2}Q(v,v)}\sqrt{\det Q}.$$

Completing the square, $t\cdot\frac{1}{2}Q(v,v)+ik(v) = \frac{1}{2}tQ(v+it^{-1}Q^{-1}(k), v+it^{-1}Q^{-1}(k))+\frac{1}{2}t^{-1}Q(Q^{-1}(k), Q(k))$. Thus

$$\int_{v\in V} e^{-2\pi i k(v)-2\pi t\cdot\frac{1}{2}Q(v,v)}\sqrt{\det(tQ)} = e^{-2\pi t^{-1}\cdot\frac{1}{2}Q(k,k)}\int_{v'=v+it^{-1}Q^{-1}(k)\in V} e^{-2\pi t\frac{1}{2}Q(v',v')}\sqrt{\det(tQ)} = e^{-2\pi t^{-1}\cdot\frac{1}{2}Q^{-1}(k,k)}.$$

Thus the double-integral is equal to

$$\int_{k\in V^*}\sum_{\alpha\in L^*}\delta(k-\alpha)e^{-2\pi t^{-1}\cdot\frac{1}{2}Q^{-1}(k,k)} = \sum_{\alpha\in L^*}e^{-2\pi t^{-1}\cdot\frac{1}{2}Q^{-1}(\alpha,\alpha)} = \theta_{L^*}(t^{-1}).$$

On the other hand, reversing the order of integration, the double integral becomes

$$\int_{v\in V}\sqrt{\det(tQ)}\sum_{\alpha\in L^*}e^{-2\pi i\alpha(v)}e^{-2\pi t\cdot\frac{1}{2}Q(v,v)}.$$

Poisson summation gives

$$= \sqrt{\det(tQ)}(V/L)\int_{v\in V}\sum_{x\in L}\delta(v-x)e^{-2\pi t\cdot\frac{1}{2}Q(v,v)} = \sqrt{\det Q}(V/L)\,t^{n/2}\theta_L(t).$$

The transformation $t\mapsto 1/t$ is equivalent to $\tau\mapsto -1/\tau$. Changing variables, we get the desired result. □

Remark 145. If (L, Q) is unimodular, then $\sqrt{\det Q}(V/L) = 1$ and $\theta_{L^*} = \theta_L$. Thus θ_L satisfies $\theta_L(\tau) = \theta_L(\tau+2) = (\tau/i)^{-n/2}\theta_L(-1/\tau)$, so θ_L is a modular form of weight $n/2$, but only under a subgroup of the full modular group. If Q is even, then $\theta_L(\tau) = \theta_L(\tau+1)$ and θ_L does transform under the full modular group.

For $k \in \mathbb{Z}_{\geq 3}$, define the Eisenstein series

$$E_k(\tau) := (2\zeta(k))^{-1} \sum_{(m,n)\in\mathbb{Z}^2\setminus(0,0)} (m + n\tau)^{-k}.$$

If k is odd, then $E_k(\tau) = 0$. The ring of modular forms under the full modular group is freely generated by E_4 and E_6. Note that

$$E_4(q) = 1 + 240q^2 + 2160q^4 + 6720q^6 + 17520q^8 + 30240q^{10} + O(q^{12}),$$

$$E_6(q) = 1 - 504q^2 - 16632q^4 - 122976q^6 + O(q^8).$$

Remark: Another common convention is to take $q = e^{2\pi i\tau}$, which has the effect of $q \mapsto \sqrt{q}$ in the above formulas.

Note that $\theta_{E_8}(q) = E_4(q)$, since E_8 is even, and $E_4(q)$ is the only modular form of weight 4, up to a scaling factor, under the full modular group.

Appendix A

Notation and conventions

This section is intended for reference only. Notations are explained, however the definitions and explanations occur later in this chapter. Most of this section should make no sense to a beginner, so the reader should not be intimidated.

The number systems $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ respectively denote the integers, rationals, reals, and complex numbers. The integers modulo k are denoted by the quotient $\mathbb{Z}/k\mathbb{Z}$, and often abbreviated as $\mathbb{Z}_k$. An equivalence class is denoted by square brackets, so $[2] = [7]$ in $\mathbb{Z}_5$.

Unless otherwise specified, X will be a smooth compact oriented Riemannian manifold without boundary of dimension n. Usually $n = 4$. The Riemannian metric is denoted by g. The vector space of real-valued functions on X whose partial derivatives exist to all orders is denoted $C^\infty(X)$.

Integration $\int_X f$ of a function f on a Riemannian manifold X is assumed to be with respect to the Riemannian volume measure which has the local coordinate expression $d\mathrm{vol}_g := \sqrt{\det g_{ij}}dx^1 \cdots dx^n$.

If $E \to X$ is a complex vector bundle (which is automatically assumed to be finite-dimensional), then E^* denotes the dual bundle and $\bar{E}$ denotes the conjugate bundle. The space of smooth sections of a vector bundle $E \to X$ is denoted by $\Gamma(E)$. For $s \in \Gamma(E)$ and $\alpha \in \Gamma(E^*)$, the natural metric-independent pairing of α and s denoted by $\alpha \cdot s$ or $s \cdot \alpha$ which gives a function in $C^\infty(X)$. There is a natural antilinear map $E \to \bar{E}$ which is written $s \mapsto \bar{s}$. If E is equipped with a Hermitian metric h, then the inner product of s with itself is denoted by $h(s,s) = s \cdot s = |s|^2 \in C^\infty(X)$. A Hermitian metric determines a complex linear isomorphism between $\bar{E}$ and E^* given by $\bar{s} \mapsto h(-,s)$, where $h(-,s)$ is shorthand for the complex linear functional $t \mapsto h(t,s)$.

Given $s_1, s_2 \in \Gamma(E)$ and $\alpha_1, \alpha_2 \in \Gamma(E^*)$, there is a natural evaluation map $(\alpha_1 \otimes \alpha_2)(s_1, s_2) := (\alpha_1 \cdot s_1)(\alpha_2 \cdot s_2)$ also sometimes denoted by $(s_1 \otimes s_2) \cdot (\alpha_1 \otimes \alpha_2)$.

Two objects A and B are *isomorphic*, denoted $A \simeq B$, if there exist $f : A \to B$ and $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$. If there is a specific isomorphism in mind, then A and B can be identified with each other, and this is denoted by $A \cong B$. For example, if E is a real vector bundle without metric, then $E \simeq E^* \simeq E^{**}$, while $E \cong E^{**}$ because the evaluation map $\mathrm{ev} : E \to E^{**}$ given by $\mathrm{ev}(s) := \alpha \mapsto \alpha \cdot s$ for $\alpha \in E^*$ is a distinguished isomorphism.

The exterior algebra of a vector space or bundle is denoted by $\Lambda^\bullet E$. Differential forms are defined to be $\Omega^\bullet(X) := \Gamma(\Lambda^\bullet T^*X)$, where T^*X denotes the cotangent bundle. Differential forms with values in

E are denoted by $\Omega^\bullet(X;E) := \Gamma((\Lambda^\bullet T^*X)\otimes E)$. Often $\mathbb{R}$ or $\mathbb{C}$ also denote the corresponding trivial vector bundle over X. There are natural identifications of real-valued forms $\Omega^\bullet(X;\mathbb{R}) \cong \Omega^\bullet(X)$, smooth functions $C^\infty(X) \cong \Omega^0(X) \cong \Gamma(\mathbb{R})$, and one-forms $\Omega^1(X) \cong \Gamma(T^*X)$.

The symbol $\bullet$ in a subscript or superscript is shorthand for some index p, quantified over all sensible values of p. For instance, $d : \Omega^\bullet(X) \to \Omega^{\bullet+1}(X)$ means that $d : \Omega^p(X) \to \Omega^{p+1}(X)$ for $p = 0, 1, \ldots, n-1$. It is also used to denote a cochain complex $\mathscr{C}^\bullet$ implicitly equipped with a differential, usually denoted by d. Similarly, $\mathscr{C}_\bullet$ can denote a chain complex with differential ∂.

If $V \in \Gamma(TX)$ is a vector field then $\iota_V : \Omega^\bullet(X;E) \to \Omega^{\bullet-1}(X;E)$ denotes the contraction operator. If $\alpha \in \Omega^1(X)$, then $\epsilon_\alpha : \Omega^\bullet(X;E) \to \Omega^{\bullet+1}(X;E)$ denotes the wedge operator $\epsilon_\alpha(\beta) := \alpha \wedge \beta$. They obey the algebraic relation $\epsilon_\alpha \iota_V + \iota_V \epsilon_\alpha = (\alpha \cdot V)\cdot$, where $(\alpha \cdot V)\cdot$ denotes the multiplication operator corresponding to the function $\alpha \cdot V$. Evaluation of vectors on $\omega \in \Omega^p(X;E)$ is written as either $\omega(V_1, \ldots, V_p)$ or $(V_1 \otimes \cdots \otimes V_p) \cdot \omega$, and is normalized to be $\iota_{V_p} \cdots \iota_{V_1} \omega$. Consequently, for one-forms α_1 and α_2,

$$\begin{aligned}(V_1 \otimes V_2) \cdot (\alpha_1 \wedge \alpha_2) &= (\alpha_1 \wedge \alpha_2)(V_1, V_2)\\ &= (\alpha_1 \cdot V_1)(\alpha_2 \cdot V_2) - (\alpha_2 \cdot V_1)(\alpha_1 \cdot V_2)\\ &= (V_1 \otimes V_2 - V_2 \otimes V_1) \cdot (\alpha_1 \otimes \alpha_2).\end{aligned}$$

The operator Λ replaces tensor products of one-forms by wedge products, so

$$\Lambda(\alpha_1 \otimes \cdots \otimes \alpha_p \otimes \omega) = \alpha_1 \wedge \cdots \wedge \alpha_p \wedge \omega,$$

where $\omega \in \Omega^q(X;E)$, and thus $\Lambda : \Omega^q(X;(T^*X)^{\otimes p} \otimes E) \to \Omega^{p+q}(X;E)$, where $(T^*X)^{\otimes p}$ denotes the p-fold tensor product. In particular, the conventions dictate that

$$\Lambda(\alpha_1 \otimes \alpha_2)(V_1, V_2) = (V_1 \otimes V_2 - V_2 \otimes V_1) \cdot (\alpha_1 \otimes \alpha_2). \tag{A.1}$$

When X is oriented and Riemannian, the Hodge star operator is denoted by $\star : \Omega^\bullet(X;E) \to \Omega^{n-\bullet}(X;E)$. If α and β are $\mathbb{R}$-valued differential forms of the same degree, then their inner product $\alpha \cdot \beta$ is determined by $(\alpha \cdot \beta)\, d\mathrm{vol}_X = \alpha \wedge \star\beta$. In the case $n = 4$ and $\omega \in \Omega^2(X;E)$, there are orthogonal projection operators $\omega^+ := \frac{1}{2}(\omega + \star\omega)$ and $\omega^- := \frac{1}{2}(\omega - \star\omega)$ to self-dual forms $\Omega^+(X;E)$ and anti-self-dual forms $\Omega^-(X;E)$.

The covariant derivative with respect to the Levi-Civita connection is denoted by either ∇, or ∇_{LC} for emphasis. For any bundle E associated to the tangent bundle, the induced connection uses the same symbols. If $s \in \Gamma(E)$, then $\nabla s \in \Gamma(T^*X \otimes E)$.

Usually P denotes a smooth principal[1] bundle with gauge group G. It is assumed that G is a compact Lie group with Lie algebra $\mathfrak{g}$, and that $\mathfrak{g}$ is equipped with a bi-invariant Euclidean metric denoted $\langle \xi, \chi \rangle$. The adjoint bundles of P are denoted by G_{Ad} and $\mathfrak{g}_{\mathrm{Ad}}$. The group of smooth automorphisms of P is denoted by $\mathscr{G} := \Gamma(G_{\mathrm{Ad}})$. The space of connections in P is denoted by $\mathscr{A}_P$ or simply $\mathscr{A}$. Usually $A \in \mathscr{A}$. The induced covariant derivative on any associated bundle is denoted by ∇_A. Often $A_0 \in \mathscr{A}$ denotes a chosen basepoint, sometimes known as a "fiducial connection." The difference $A - A_0 \in \Omega^1(X;\mathfrak{g}_{\mathrm{Ad}})$ is then a Lie-algebra-valued one-form. The curvature is denoted by $F_A \in \Omega^2(X;\mathfrak{g}_{\mathrm{Ad}})$.

[1]According to an old joke regarding the common misspelling, a princip*le* bundle has a moral fiber. This joke is due to Allen Knutson and popularized in J. P. Serre's classic lecture "How to write mathematics badly."

For Lie-algebra-valued differential forms α and β, infix notation specifies an operation on forms, while outfix notation specifies an operation on the Lie algebra. For example, $\langle \alpha \wedge \beta \rangle$ is a real-valued differential form obtained by wedging the form part while taking the inner product on the Lie algebra. Another example is $[\alpha \cdot \beta]$, which is a Lie-algebra-valued function obtained from taking the inner product of differential forms and applying the Lie bracket.

If $E_1 \to X$ and $E_2 \to X$ are vector bundles associated to principal bundles $P_1 \to X$ and $P_2 \to X$ with connections $A \in \mathscr{A}_{P_1}$ and $B \in \mathscr{A}_{P_2}$, then the induced covariant derivative on $E_1 \otimes E_2$ is

$$\nabla_{A\otimes B} := \nabla_A \otimes \mathrm{Id}_B + \mathrm{Id}_A \otimes \nabla_B.$$

If one of the connections is the Levi-Civita connection, then $\nabla_{A\otimes \mathrm{LC}}$ is abbreviated to ∇_A.

The exterior covariant derivative on a vector bundle E associated to a principal bundle with connection A is defined as

$$d_A := \Lambda \circ \nabla_A : \Omega^\bullet(X; E) \to \Omega^{\bullet+1}(X; E).$$

In the case $E = \mathbb{R}$ and A is the trivial connection, $d_A = d$ is the ordinary exterior derivative. Curvature satisfies

$$F_{A+a} = F_A + d_A a + \tfrac{1}{2} [a \wedge a] .$$

Suppose P is a principal bundle with structure group G and connection A, and G has a vector space representation $\rho : G \to \mathrm{End}(V)$. Then on the associated vector bundle $E := P \times_\rho V$, the second exterior derivative d_A^2 acts as $\rho(F_A)\wedge$.

Suppose V is a vector field and ∇ is a covariant derivative on a vector bundle E, where the connection A is left implicit. For a section $s \in \Gamma(E)$, it is common to let $\nabla_V s$ denote the contraction $V \cdot \nabla s$ of V with the T^*X factor of ∇s. The second covariant derivative is $\nabla^2 s \in \Gamma(T^*X \otimes T^*X \otimes E)$. Iterated covariant derivatives satisfy

$$\nabla_V \nabla^k_{W_1\otimes\cdots\otimes W_k} = \nabla^{k+1}_{V\otimes W_1\otimes\cdots\otimes W_k} + \nabla^k_{\nabla_V (W_1\otimes\cdots\otimes W_k)}.$$

For example,

$$\nabla^2_{V\otimes W} = \nabla_V \nabla_W - \nabla_{\nabla_V W}.$$

The Riemannian curvature, denoted $\mathrm{Rm} \in \Omega^2(X; \mathfrak{so}(TX))$, is the curvature F_{LC} of the orthonormal frame bundle of X. For a real vector bundle $E \to X$ with Euclidean metric, $\mathfrak{so}(E)$ denotes the bundle of skew-symmetric endomorphisms of E.

To compute the standard action of Rm on TX, contract vector fields V and W with the two-form part to obtain $\mathrm{Rm}(V, W) \in \Gamma(\mathfrak{so}(TX))$. Evaluating the endomorphism on another vector field S produces the vector field $\mathrm{Rm}(V, W)S \in \Gamma(TX)$, given by

$$\mathrm{Rm}(V, W)S = (d^2_{\mathrm{LC}} S)(V, W) = (\Lambda \nabla^2 S)(V, W) \overset{(A.1)}{=} \nabla^2_{V\otimes W - W\otimes V} S.$$

The Ricci tensor is

$$\mathrm{Ric}(W, S) := \mathrm{Tr}(V \mapsto \mathrm{Rm}(V, W)S) \in \Gamma(\mathrm{Sym}^2 T^*X),$$

where Tr denotes the trace, and $\mathrm{Sym}^p(E^*)$ denotes homogeneous symmetric polynomials of degree p on the fibers of E. The scalar curvature Sc is the metric contraction of the Ricci curvature.

Suppose V is a real vector space (again assumed to be finite-dimensional). Then $\mathfrak{gl}(V)$ denotes the Lie algebra of endomorphisms of V. The dual representation $L \mapsto -L^*$ is an isomorphism of Lie algebras $\mathfrak{gl}(V) \cong \mathfrak{gl}(V^*)$. If V has a non-degenerate symmetric bilinear form g, then $V \cong V^*$ by the "musical isomorphisms" $v \mapsto v^\flat := g(v, -)$ and inverse $\alpha \mapsto \alpha^\sharp$. The dual representation is compatible with the musical isomorphisms precisely when $(L\alpha^\sharp)^\flat = -L^*\alpha$, which is the definition of $L \in \mathfrak{so}(V)$. Consequently, as $\mathrm{Rm} \in \Omega^2(X; \mathfrak{so}(TX))$,

$$\nabla^2_{V\otimes W - W\otimes V}\alpha = \left(\mathrm{Rm}(V,W)\alpha^\sharp\right)^\flat = -\mathrm{Rm}(V,W)^*\alpha.$$

Now suppose $\{e_1, \ldots, e_n\}$ is a basis for V, and let $\left\{e^1, \ldots, e^n\right\}$ denote the dual basis. For $L \in \mathfrak{gl}(V)$, the components are $L^i_j := e^i \cdot Le_j$. If $v = v^j e_j$ (where repeated indices are implicitly summed), then $Lv = \left(L^i_j v^j\right) e_i$. If $\alpha = \alpha_i e^i \in V^*$, then the dual map is $L^*\alpha = \left(L^i_j \alpha_i\right) e^j$. If $g(e_i, e_j) = g_{ij}$, then $(v^\flat)_i = g_{ij}v^j$ and $(\alpha^\sharp)^j = \alpha_i g^{ij}$, where g^{ij} is the inverse matrix so that $g_{ij}g^{jk} = \delta^k_i$, where δ^k_i is the Kronecker delta. The condition $L \in \mathfrak{so}(V)$ is equivalent to $-L^i_j = g_{i\ell}L^\ell_k g^{ki}$. If $\{e_i\}$ is orthonormal and g is positive-definite so that $g_{ij} = \delta_{ij}$, then the condition becomes simply $-L^i_j = L^j_i$.

If $\{e_i\}$ is a local frame for TX with local coframe $\left\{e^i\right\}$, then the components and contractions of Riemannian curvature are defined by

$$\begin{aligned}
R_{ij}{}^k{}_\ell &:= e^k \cdot \mathrm{Rm}(e_i, e_j)e_\ell, \\
R_{j\ell} &:= R_{ij}{}^i{}_\ell = \mathrm{Ric}(e_j, e_\ell), \\
\mathrm{Sc} &:= R_{jk}g^{jk}.
\end{aligned}$$

Appendix B

Cohomology

B.1 How to invent de Rham cohomology

Cohomology of manifolds is essentially the deep study of locally-constant functions. To begin, $H^0(X;R)$ is defined to be the space of locally-constant R-valued functions on a manifold X. We take R to be a commutative ring, usually $\mathbb{R}$, $\mathbb{Z}$, or $\mathbb{Z}_2$. To obtain a locally-constant function on X, we may freely assign any value to each connected component of X. Thus $H^0(X;R)$ is the Cartesian product of several copies of R. Denoting the set of connected components of X by $\pi_0(X)$, we summarize this by

$$H^0(X;R) \cong \prod_{\pi_0(X)} R = R^{\pi_0(X)}.$$

We note several important properties of $H^0(X;R)$ which will extend to higher cohomology.

- $H^0(X;R)$ is a ring.
- If X is compact, then $\pi_0(X)$ is finite, and hence $H^0(X;R)$ is a finite-dimensional vector space (or module when R is not a field) over R.
- A ring homomorphism $h : R \to R'$ induces a ring homomorphism, also denoted by h,

  $$h : H^0(X;R) \to H^0(X;R'),$$

  given by

  $$h(f) := h \circ f = (x \mapsto h(f(x))) \in H^0(X;R').$$

- A smooth map $\phi : X \to Y$ of manifolds induces a ring homomorphism in the opposite direction

  $$\phi^* : H^0(Y;R) \to H^0(X;R).$$

  In particular, if $f \in H^0(Y;R)$, then

  $$\phi^*(f) := f \circ \phi = (x \mapsto f(\phi(x))) \in H^0(X;R).$$

It would seem that this is the end of the story for locally-constant functions. However, by examining how locally-constant functions restrict to open subsets, we will discover a rigid structure which leads naturally to the definition of $H^p(X;R)$ for $p > 0$.

Suppose A and B are open subsets of X such that $A \cup B = X$. We have natural inclusion maps which are smooth

$$\begin{tikzcd} & A \arrow[dr, hook, "i_A"] & \\ A \cap B \arrow[ur, hook, "j_A"] \arrow[dr, hook, "j_B"'] & & X \\ & B \arrow[ur, hook, "i_B"'] & \end{tikzcd}$$

The induced maps on $H^0(X;R)$ are restriction maps, which point in the opposite direction.

$$\begin{tikzcd} & H^0(A;R) \arrow[dl, "j_A^*"'] & \\ H^0(A \cap B;R) & & H^0(X;R) \arrow[ul, "i_A^*"'] \arrow[dl, "i_B^*"] \\ & H^0(B;R) \arrow[ul, "j_B^*"] & \end{tikzcd}$$

We combine the restriction maps i_A^* and i_B^* into a linear map

$$i_A^* \times i_B^* : H^0(X;R) \to H^0(A;R) \times H^0(B;R).$$

For concreteness, let's consider $X = S^1$, with A and B two arcs which cover X, and $R = \mathbb{R}$. Then the diagram becomes

$$\begin{tikzcd} & \mathbb{R} \arrow[dl, "j_A^*"'] & \\ \mathbb{R}^2 & & \mathbb{R} \arrow[ul, "i_A^*"'] \arrow[dl, "i_B^*"] \\ & \mathbb{R} \arrow[ul, "j_B^*"] & \end{tikzcd}$$

and $i_A^* \times i_B^* : \mathbb{R} \to \mathbb{R}^2$ is given by $\lambda \mapsto \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}$, because the locally constant function $f(x) = \lambda$ restricts as λ to both A and B.

Given two locally-constant functions f_A and f_B on A and B, one can ask whether the pair extends to some function on X. This is equivalent to asking whether the pair (f_A, f_B) is in the image of $i_A^* \times i_B^*$. The image of $i_A^* \times i_B^*$ determines a subspace of $H^0(A;R) \times H^0(B;R)$ which, in our example, corresponds to the span of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

With this subspace in mind, we note that there are two primary ways to specify a vector subspace. A subspace is determined either parametrically, as the image of some linear map such as $i_A^* \times i_B^*$, or instead via some constraint equations, which amount to the kernel of some linear map. A major theme of homological algebra is that whenever a subspace appears as either an image or kernel, one should seek a description of the alternate form. Thus we seek to characterize the image of $i_A^* \times i_B^*$ as the kernel of some other map.

The answer is easily expressed in words. Two locally-constant functions f_A and f_B extend to X iff they agree on the overlap $A \cap B$. Equivalently, we require that the difference of the restriction maps vanishes:

$$j_A^*(f_A) - j_B^*(f_B) = 0.$$

In our example, if $f_A(a) = \lambda_A$ and $f_B(b) = \lambda_B$, then the pair (f_A, f_B) is represented by the vector $\begin{pmatrix} \lambda_A \\ \lambda_B \end{pmatrix}$, and $j_A^* - j_B^*$ is represented by the vector $\begin{pmatrix} \lambda_A - \lambda_B \\ \lambda_A - \lambda_B \end{pmatrix}$. In summary, we have

$$H^0(X;R) \xrightarrow{i_A^* \times i_B^*} H^0(A;R) \times H^0(B;R) \xrightarrow{j_A^* - j_B^*} H^0(A \cap B;R),$$

which for our example in matrix notation is

$$\mathbb{R} \xrightarrow{\begin{pmatrix} 1 \\ 1 \end{pmatrix}} \mathbb{R}^2 \xrightarrow{\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}} \mathbb{R}^2.$$

When we are in this general situation

$$M_1 \xrightarrow{d_1} M_2 \xrightarrow{d_2} M_3$$

with $\operatorname{im} d_1 = \ker d_2$, we say that this diagram is *exact* at M_2. Half of this equality is very simple to check: $\operatorname{im} d_1 \subset \ker d_2$, which amounts to the fact that $d_2 \circ d_1 = 0$. In our example, we easily verify that

$$\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Exactness is a convenient language to convey several properties. For instance,

$$0 \xrightarrow{0} M_2 \xrightarrow{d_2} M_3$$

means that d_2 is injective, and

$$M_1 \xrightarrow{d_1} M_2 \xrightarrow{0} 0$$

means that d_1 is surjective.

An *exact sequence* is a chain of maps which are exact. For instance, exactness of

$$0 \longrightarrow M_1 \xrightarrow{d_1} M_2 \longrightarrow 0$$

means that d_1 is simultaneously injective and surjective, hence an isomorphism. Exactness of

$$0 \longrightarrow M_1 \xrightarrow{d_1} M_2 \xrightarrow{d_2} M_3 \longrightarrow 0$$

means that d_2 is surjective, d_1 is injective, and that $M_3 \cong M_2/M_1$. Such an exact sequence with three nonzero terms sandwiched by zeroes arises frequently, and is called a *short exact sequence*.

We wish to extend everything to an exact sequence. Note that $i_A^* \times i_B^*$ is injective, since A and B cover X, and the only way for a function to restrict as zero to both A and B is if that function is zero everywhere. Hence we can add a zero on the left:

$$0 \longrightarrow H^0(X;R) \xrightarrow{i_A^* \times i_B^*} H^0(A;R) \times H^0(B;R) \xrightarrow{j_A^* - j_B^*} H^0(A \cap B;R). \tag{B.1}$$

In general, the image of $j_A^* - j_B^*$ is a proper subspace of $H^0(A \cap B; R)$. In our example, it is the span of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. To continue the exact sequence, we must find some natural space, and a map to this space, such that the kernel is the image of $j_A^* - j_B^*$. This space will be $H^1(X;R)$.

The failure of $j_A^* - j_B^*$ to be surjective is due to the fact that locally constant functions are very rigid. Continuous functions are more flexible. Denote by $C(X)$ the ring of continuous functions on X. Then

Theorem 146.

$$0 \longrightarrow C(X) \xrightarrow{i_A^* \times i_B^*} C(A) \times C(B) \xrightarrow{j_A^* - j_B^*} C(A \cap B) \longrightarrow 0$$

is a short exact sequence.

Proof. We have already demonstrated everything except that $j_A^* - j_B^*$. To prove this, we construct a right-inverse.

To construct a right-inverse, it suffices to construct functions $\phi_A, \phi_B \in C(X)$ such that $\phi_A + \phi_B = 1$, ϕ_A vanishes outside of A, and ϕ_B vanishes outside of B. Given $f \in C(A \cap B)$, consider the map $C(A \cap B) \to C(A) \times C(B)$ given by $f \mapsto \begin{pmatrix} \phi_B \cdot f \\ -\phi_A \cdot f \end{pmatrix}$. Note that f is only defined on $A \cap B$. But $\phi_B f$ makes sense as a continuous function on all of A since $\phi_B \equiv 0$ on $A - (A \cap B)$, and we can extend by zero. Similarly, $-\phi_A f$ makes sense on B. Finally we verify that $(j_A^* - j_B^*) \circ \begin{pmatrix} \phi_B \cdot f \\ -\phi_A \cdot f \end{pmatrix} = f$, so that our map is indeed a right-inverse. We obtain for $x \in A \cap B$,

$$(j_A^*(\phi_B \cdot f))(x) + (j_B^*(\phi_A \cdot f))(x) = \phi_B(j_A(x)) \cdot f(j_A(x)) + \phi_A(j_B(x)) \cdot f(j_B(x)) = \phi_B(x) f(x) + \phi_A(x) f(x) = 1 \cdot f(x),$$

as desired. The proof is finished up to the following lemma. □

Lemma 147. *Let X be a smooth manifold, and suppose that $\{U_i\}_{i \in I}$ is an open cover of X. Then there exists a collection of functions $\{\phi_i\}_{i \in I}$ with the following properties.*

- *Each $\phi_i \geq 0$, and the closure of the set where $\phi_i > 0$ is contained in U_i.*
- *For any $x \in X$ only finitely many $\phi_i(x)$ are nonzero.*

- *For any $x \in X$, the finite sum $\sum_{i\in I} \phi_i(x) = 1$.*
- *Each $\phi_i \in C^\infty(X)$, meaning that ϕ_i and all its partial derivatives are continuous.*

Such a collection of functions is called a *partition of unity subordinate to the cover $\{U_i\}_{i\in I}$*. The construction is standard, and thus not included here. It relies on the existence of smooth cutoff functions such as

$$\chi(x) := \begin{cases} e^{-1/x} & x > 0, \\ 0 & x \le 0. \end{cases}$$

By the same argument, it follows that

$$0 \longrightarrow C^\infty(X) \xrightarrow{i_A^* \times i_B^*} C^\infty(A) \times C^\infty(B) \xrightarrow{j_A^* - j_B^*} C^\infty(A \cap B) \longrightarrow 0$$

is exact.

To summarize the situation up to now, smooth functions have partitions of unity while locally-constant functions do not. We need to introduce $H^p(X; R)$ with $p > 0$ in order to continue the exact sequence (B.1).

Let $B \subset \mathbb{R}^n$ be the unit n-ball. Let's try to extend

$$0 \longrightarrow H^0(B; \mathbb{R}) \longrightarrow C^\infty(B; \mathbb{R}),$$

where the map on the right is the inclusion of constant functions as smooth functions. A smooth function is locally constant when

$$\nabla f = \frac{\partial f}{\partial x^1} dx^1 + \cdots + \frac{\partial f}{\partial x^n} dx^n = 0.$$

We define

$$\Omega^0(X) := C^\infty(X; \mathbb{R}),$$

and

$$\Omega^1(B) := \left\{ \alpha_1 dx^1 + \cdots + \alpha_n dx^n \mid \text{each } \alpha_i \in \Omega^0(B) \right\}.$$

Thus

$$0 \longrightarrow H^0(B; \mathbb{R}) \longrightarrow \Omega^0(B) \xrightarrow{\nabla} \Omega^1(B)$$

is exact. But the operator ∇ is not surjective. When $n = 2$ or 3, the image of ∇ is characterized as the set of $\alpha \in \Omega^1(B)$ such that curl $\alpha = 0$. In general, if $\alpha_i = \frac{\partial f}{\partial x^i}$, then

$$\frac{\partial \alpha_i}{\partial x^j} = \frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial \alpha_j}{\partial x^i}$$

by symmetry of mixed partial derivatives, which holds for all smooth functions. Thus we must require symmetry of the matrix $\left(\frac{\partial \alpha_i}{\partial x^j}\right)_{i,j}$. Equivalently, we must require that the antisymmetric part of $\left(\frac{\partial\alpha_i}{\partial x^j}\right)_{i,j}$ must vanish. Symbolically, we define

$$\begin{aligned}\nabla(\alpha_1 dx^1 + \cdots + \alpha_n dx^n) := & \frac{\partial\alpha_1}{\partial x^1}dx^1 \otimes dx^1 + \cdots + \frac{\partial\alpha_1}{\partial x^n}dx^n \otimes dx^1 + \\ & + \cdots + \\ & + \frac{\partial\alpha_n}{\partial x^1}dx^1 \otimes dx^n + \cdots + \frac{\partial\alpha_n}{\partial x^n}dx^n \otimes dx^n.\end{aligned}$$

Next we define Λ to be the operator which replaces the tensor product with the antisymmetrized tensor product, known as the wedge product $\wedge$. Thus

$$d\alpha := \Lambda\nabla(\alpha) = \sum_{i<j}\left(\frac{\partial\alpha_j}{\partial x^i} - \frac{\partial\alpha_i}{\partial x^j}\right) dx^i \wedge dx^j.$$

We are led naturally to the definitions

$$\Omega^p(B) := \left\{ \sum_{i_1<i_2<\cdots<i_p} \alpha_{i_1 i_2\cdots i_p} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \right\},$$

$$d : \Omega^p(B) \to \Omega^{p+1}(B).$$

Since $\Omega^p(B) = 0$ for $p > n$, we are led to the sequence

$$0 \longrightarrow H^0(B;\mathbb{R}) \longrightarrow \Omega^0(B) \xrightarrow{d} \Omega^1(B) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n(B) \longrightarrow 0.$$

By symmetry of mixed partial derivatives, it follows that $d \circ d = 0$.

Lemma 148 (Poincaré lemma)**.** *This sequence is exact.*

This lemma is standard, so we omit it. Since $d^2 = 0$, we know that $\operatorname{im} d \subset \ker d$. To show that $\operatorname{im} d \supset \ker d$, one constructs for any $\alpha \in \Omega^p(B)$ such that $d\alpha = 0$ a $\eta \in \Omega^{p-1}(B)$ which satisfies $d\eta = \alpha$.

We wish to replace B by a more general manifold X. For this, we must define the transformation law for dx^i under change of coordinates. If $\phi : Y \to X$ is a smooth map, so that $x = \phi(y)$, then

$$\phi^* : \Omega^1(Y) \leftarrow \Omega^1(X)$$

according to

$$\phi^*(dx^i) = \frac{\partial x^i}{\partial y^j}dy^j = \frac{\partial\phi^i}{\partial y^j}dy^j.$$

This transformaton is known as pull-back, and reverses the direction of arrows. It extends tensorialy, and, for example on $\Omega^n(X)$,

$$\phi^*(dx^1 \wedge \cdots \wedge dx^n) = \frac{\partial x^1}{\partial y^{j_1}}dy^{j_1} \wedge \cdots \wedge \frac{\partial x^n}{\partial y^{j_n}}dy^{j_n} = \det\left(\frac{\partial x^i}{\partial y^j}\right) dy^1 \wedge \cdots \wedge dy^n.$$

We can view $\Omega^1(X)$ as sections of the cotangent bundle T^*X. The cotangent space at a point x is denoted T_x^*X, and is defined as the vector space of smooth functions $\{f \in C^\infty(X) \mid f(x) = 0\}$ which vanish at x, under the quotient relation that $[f] \sim [g]$ iff f and g agree to first-order, i.e. in any (equivalently every) coordinate chart, $(df)|_x = (dg)|_x$. This is coordinate-independent, and defines a vector space of dimension $n = \dim X$.

The tangent space...

It is no longer true that

$$0 \longrightarrow H^0(X;\mathbb{R}) \longrightarrow \Omega^0(X) \xrightarrow{d} \Omega^1(X) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n(X) \longrightarrow 0$$

is exact. However, since d is defined locally, the equation $d^2 = 0$ continues to hold, so $\operatorname{im} d \subset \ker d$. A chain of maps

$$\cdots \longrightarrow M^{p-1} \xrightarrow{d} M^p \xrightarrow{d} M^{p+1} \xrightarrow{d} \cdots$$

for which $d^2 = 0$ is called a *cochain complex*. The defect from a cochain complex being exact is measured by the cohomology

$$H^p(M^\bullet, d) := \frac{\ker(M^p \xrightarrow{d} M^{p+1})}{\operatorname{im}(M^{p-1} \xrightarrow{d} M^p)}.$$

We define the de Rham cohomology

$$H^p_{\mathrm{dR}}(X) := H^p(\Omega^\bullet(X), d),$$

which is the cohomology of the de Rham cochain complex

$$0 \longrightarrow \Omega^0(X) \xrightarrow{d} \Omega^1(X) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n(X) \longrightarrow 0.$$

The first important property of the de Rham complex is that $H^0_{\mathrm{dR}}(X) = H^0(X;\mathbb{R})$. The second important property is that since one can multiply elements of $\Omega^p(X)$ by smooth functions, techniques involving partitions of unity apply. To summarize these two properties, one says that the de Rham complex is a "fine resolution of the constant sheaf."

There is a notion called "sheaf cohomology" and de Rham cohomology is a means of computing it. In this sense, $H^p(X;\mathbb{R}) = H^p_{\mathrm{dR}}(X)$.

Suppose once again that $X = A \cup B$ for open subsets A and B. By the same argument as before,

$$0 \longrightarrow \Omega^p(X) \xrightarrow{i_A^* \times i_B^*} \Omega^p(A) \times \Omega^p(B) \xrightarrow{j_A^* - j_B^*} \Omega^p(A \cap B) \longrightarrow 0$$

is exact. Moreover, these maps fit together into a huge commutative diagram

$$
\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & \Omega^{p+1}(X) \arrow[u, "d"] \arrow[r, "i_A^*\times i_B^*"] & \Omega^{p+1}(A)\times\Omega^{p+1}(B) \arrow[u, "d"] \arrow[r, "j_A^*-j_B^*"] & \Omega^{p+1}(A\cap B) \arrow[u, "d"] \arrow[r] & 0 \\
0 \arrow[r] & \Omega^{p}(X) \arrow[u, "d"] \arrow[r, "i_A^*\times i_B^*"] & \Omega^{p}(A)\times\Omega^{p}(B) \arrow[u, "d"] \arrow[r, "j_A^*-j_B^*"] & \Omega^{p}(A\cap B) \arrow[u, "d"] \arrow[r] & 0 \\
0 \arrow[r] & \Omega^{p-1}(X) \arrow[u, "d"] \arrow[r, "i_A^*\times i_B^*"] & \Omega^{p-1}(A)\times\Omega^{p-1}(B) \arrow[u, "d"] \arrow[r, "j_A^*-j_B^*"] & \Omega^{p-1}(A\cap B) \arrow[u, "d"] \arrow[r] & 0 \\
 & \vdots \arrow[u, "d"] & \vdots \arrow[u, "d"] & \vdots \arrow[u, "d"] &
\end{tikzcd}
$$

This is abbreviated by

$$0 \longrightarrow \Omega^\bullet(X) \xrightarrow{i_A^*\times i_B^*} \Omega^\bullet(A)\times\Omega^\bullet(B) \xrightarrow{j_A^*-j_B^*} \Omega^\bullet(A\cap B) \longrightarrow 0\ ,$$

which is called a short exact sequence of cochain complexes. The big result from commutative algebra is that a short exact sequence of cochain complexes induces a long exact sequence on cohomology. Specifically, if

$$0 \longrightarrow C^\bullet \xrightarrow{f} D^\bullet \xrightarrow{g} E^\bullet \longrightarrow 0$$

is a short exact sequence of cochain complexes, then there is an exact sequence of the form

$$\cdots \longrightarrow H^p(C^\bullet, d) \xrightarrow{f} H^p(D^\bullet, d) \xrightarrow{g} H^p(E^\bullet, d) \xrightarrow{f^{-1}\circ d\circ g^{-1}} H^{p+1}(C^\bullet, d) \xrightarrow{f} \cdots$$

Thus we obtain our desired extension

$$
\begin{tikzcd}
0 \arrow[r] & H^0(X;\mathbb{R}) \arrow[r, "i_A^*\times i_B^*"] & H^0(A;\mathbb{R})\times H^0(B;\mathbb{R}) \arrow[r, "j_A^*-j_B^*"] & H^0(A\cap B;\mathbb{R}) \arrow[dll, rounded corners] \\
 & H^1_{\mathrm{dR}}(X) \arrow[r] & H^1_{\mathrm{dR}}(A)\times H^1_{\mathrm{dR}}(B) \arrow[r] & H^1_{\mathrm{dR}}(A\cap B) \arrow[dll, rounded corners] \\
 & H^2_{\mathrm{dR}}(X) \arrow[r] & H^2_{\mathrm{dR}}(A)\times H^2_{\mathrm{dR}}(B) \arrow[r] & H^2_{\mathrm{dR}}(A\cap B) \arrow[dll, rounded corners] \\
 & \vdots & \vdots & \vdots
\end{tikzcd}
$$

B.2 How to compute with de Rham cohomology

Functoriality. Identity maps to identity. Composition maps to composition.

B.2.1 Homotopy invariance

Homotopy

Definition. A (smooth) homotopy between two maps $f_i : X \to Y$, $i \in \{0,1\}$ is a (smooth) continuous map $f : X \times [0,1] \to Y$ such that $f_i = f|_{X\times\{i\}}$.

Both homotopy and smooth homotopy are equivalence relations. Between two smooth manifolds, one could choose between either smooth homotopy or regular continuous homotopy. Conveniently, these notions are essentially equivalent.

Theorem (Smooth approximation)**.** *Any continuous map $f : X \to Y$ between smooth manifolds is homotopic to a smooth map. Moreover, if f_0 and f_1 are homotopic, then they are smoothly homotopic.*

Thus homotopy classes of maps between smooth manifolds are equivalent, regardless of whether or not the maps are required to be smooth.

In the homotopy category of continuous maps modulo homotopy, let's examine the notion of isomorphism, which is called *homotopy equivalence*. The morphism represented by $f : X \to Y$ is an isomorphism if there exists a map $g : Y \to X$ such that $g \circ f \simeq \mathrm{Id}_X$ and $f \circ g \simeq \mathrm{Id}_Y$. For example, the inclusion of $f : S^1 \hookrightarrow \mathbb{R}^2 - \{0\}$ is a homotopy equivalence since $g : \mathbb{R}^2 - \{0\} \to S^1$ defined by $x \mapsto x/\,|x|$ satisfies $g \circ f = \mathrm{Id}_{S^1}$, and $f \circ g : \mathbb{R}^2 - \{0\} \to \mathbb{R}^2 - \{0\}$ by $x \mapsto x/\,|x|$ is homotopic to the identity by

$$x \mapsto \frac{x}{t + (1-t)\,|x|}.$$

As in this example, a noncompact manifold can be homotopy equivalent to a manifold of a lower dimension. Any contractible manifold is homotopy equivalent to a point.

In order to compute cohomology, it is very important to be able to deform things. The key idea is that homotopic maps produce identical maps on cohomology.

Definition 149. Two smooth maps f_0 and $f_1 : X \to Y$ are *smoothly homotopic* if there exists a smooth function $f : [0,1]_t \times X \to Y$ such that $f_0 = f_{t=0}$ and $f_1 = f_{t=1}$. In this case, we write $f_0 \simeq f_1$.

Theorem 150. *If $f_0 \simeq f_1$, then $f_0^* = f_1^* : H^\bullet_{\mathrm{dR}}(Y) \to H^\bullet_{\mathrm{dR}}(X)$.*

The symbol $\mathscr{C}^\bullet$ is an abbreviation for "$\mathscr{C}^p$ for all p."

As an important consequence, suppose $f : X \to Y$ and $g : Y \to X$ satisfy $f \circ g \simeq \mathrm{Id}_Y$ and $g \circ f \simeq \mathrm{Id}_X$. Then $g^* \circ f^* = \mathrm{Id}_{\Omega^\bullet(Y)}$ and $f^* \circ g^* = \mathrm{Id}_{\Omega^\bullet(X)}$. Thus f^* is invertible, and thus an isomorphism. Such a map f is called a *homotopy equivalence*. Thus a homotopy equivalence induces an isomorphism on cohomology.

We call two spaces X and Y *homotopy equivalent* if there exists a homotopy equivalence $f : X \to Y$. We write $X \simeq Y$.

$B^n \simeq \mathrm{pt}$.

$\mathbb{R}^n - \{0\} \simeq S^{n-1}$.

Use Mayer-Vietoris to prove cohomology of spheres. Cohomology of $\mathbb{CP}^2$.

Suppose we have an arbitrary linear map $L : \Omega^p(Y) \to \Omega^q(X)$. Consider the relevant portion of the de Rham complexes

$$\begin{tikzcd} \Omega^{p-1}(Y) \arrow[r, "d"] & \Omega^p(Y) \arrow[r, "d"] \arrow[d, "L"'] & \Omega^{p+1}(Y) \\ \Omega^{q-1}(X) \arrow[r, "d"] & \Omega^q(X) \arrow[r, "d"] & \Omega^{q+1}(X) \end{tikzcd}$$

In order for K to induce a map on cohomology, it must map $\ker\left(\Omega^p(Y) \to \Omega^{p+1}(Y)\right)$ to $\ker\left(\Omega^q(X) \to \Omega^{q+1}(X)\right)$. Then we can try to write $L\,[\omega] := [L\omega]$. To check that this is well-defined, independent of the choice of representative, we must verify the following equivalent conditions

- $[L\omega] = \left[L\left(\omega + d\eta\right)\right]$ for all $\eta \in \Omega^{q-1}(X)$,
- $\left[Ld\eta\right] = 0$ for all $\eta \in \Omega^{q-1}(X)$,
- For all $\eta \in \omega^{q-1}(X)$ there exists some $\nu \in \Omega^{p-1}(X)$ such that $Ld\eta = d\nu$,
- L maps im $\left(\Omega^{p-1}(Y) \to \Omega^p(Y)\right)$ to im $\left(\Omega^{q-1}(X) \to \Omega^q(X)\right)$.

There are two simple conditions under which these conditions are trivially satisfied.

Lemma 151. *For arbitrary linear maps K and K' as in*

$$\begin{tikzcd} \Omega^{p-1}(Y) \arrow[r, "d"] & \Omega^p(Y) \arrow[r, "d"] \arrow[dl, "K"'] & \Omega^{p+1}(Y) \arrow[dl, "K'"'] \\ \Omega^{q-1}(X) \arrow[r, "d"] & \Omega^q(X) \arrow[r, "d"] & \Omega^{q+1}(X) \end{tikzcd}$$

the compositions $d \circ K$ and $K' \circ d$ both induce the zero map $H^p(Y) \to H^q(X)$. In particular, $d \circ K + K' \circ d$ induce the zero map.

Proof. First, for any $\omega \in \Omega^p(Y)$ by definition $[d(K\omega)] = 0$. Second, if $\omega \in \ker d$, then $[K'd\omega] = [K'0] = 0$. □

Associated to each homotopy $f : [0,1]_t \times X \to Y$ there is a collection of operators $K : \Omega^\bullet(Y) \to \Omega^{\bullet-1}(X)$ such that the difference $f_1^* - f_0^* : \Omega^\bullet(Y) \to \Omega^\bullet(X)$ is given by

$$f_1^* - f_0^* = dK + Kd. \tag{B.2}$$

Thus the difference must vanish on cohomology.

This equation (B.2) has the following dual geometric interpretation. Consider a cylinder as a geometric shape. Then f_1^* and f_0^* represent the top and bottom caps of the cylinder. Considering K as an extrusion operator acting on the disc, Kd represents the extrusion of the circle, which is the side of the cylinder, while dK represents the boundary of the extruded disc, which is the boundary of the cylinder itself. By an orientation convention, the side of the cylinder cancels, leaving just the top and bottom remaining.

The operator K is assembled as follows. For each t, denote by i_t^* the restriction of $\Omega^\bullet([0,1]_t \times X)$ to $\Omega^\bullet(X)$. There is a composition

$$\Omega^\bullet(Y) \xrightarrow{f^*} \Omega^\bullet([0,1]_t \times X) \xrightarrow{\iota_{\partial_t}} \Omega^{\bullet-1}([0,1]_t \times X) \xrightarrow{i_t^*} \Omega^{\bullet-1}(X).$$

Define $K := \int_0^1 i_t^* \circ \iota_{\partial_t} \circ f^* \, dt$.

Lemma 152 (Cartan homotopy formula). *If V is a vector field on X, then $\mathscr{L}_V \omega = d\iota_V \omega + \iota_V d\omega$.*

We compute

$$\begin{aligned} f_1^* - f_0^* &= (f \circ i_1)^* - (f \circ i_0)^* \\ &= \int_0^1 \frac{d}{dt} (f \circ i_t)^* \, dt \\ &= \int_0^1 \frac{d}{dt} i_t^* \circ f^* \, dt \\ &= \int_0^1 i_t^* \circ \mathscr{L}_{\partial_t} \circ f^* \, dt \\ &= \int_0^1 \left(i_t^* \circ d \circ \iota_{\partial_t} \circ f^* + i_t^* \circ \iota_{\partial_t} \circ d \circ f^* \right) \, dt \\ &= \int_0^1 \left(d \circ i_t^* \circ \iota_{\partial_t} \circ f^* + i_t^* \circ \iota_{\partial_t} \circ f^* \circ d \right) \, dt \\ &= dK + Kd. \end{aligned}$$

Integration

Let X be an n-manifold. Given two overlapping coordinate charts $\phi_i : U_i \to \mathbb{R}^n$, $i \in \{1,2\}$, the transition function $\phi_2 \circ \phi_1^{-1}$ is invertible, and hence the Jacobian determinant is nowhere zero. It is orientation-preserving if the Jacobian is positive, and orientation-reversing if the Jacobian is negative. X is said to be *oriented* if it comes equipped with an atlas where all the transition functions are orientation-preserving. Thus the coordinate charts define a consisitent orientation.

The Jacobian determinant arises in differential forms as follows. Given a change of coordinates $f : Y \to X$,

$$f^*(dx^1 \wedge \cdots \wedge dx^n) = \frac{\partial x^1}{\partial y^{j_1}} dy^{j_1} \wedge \cdots \wedge \frac{\partial x^n}{\partial y^{j_n}} dy^{j_n} = \det\left(\frac{\partial x^i}{\partial y^j}\right) dy^1 \wedge \cdots \wedge dy^n.$$

The determinant is characterized by multilinearity and antisymmetry.

Any nowhere-vanishing $\omega \in \Omega^n(X)$ determines an orientation, where a coordinate chart is positively-oriented if $dx^1 \wedge \cdots \wedge dx^n$ is a positive function multiple of ω. Conversely, any orientation is determined by such a nowhere-vanishing $\omega \in \Omega^n(X)$.

If X is oriented, then $\omega \in \Omega^n(X)$ transforms as an integrand: if $\omega = \omega(x)\, dx^1 \wedge \cdots \wedge dx^n$, then

$$f^*(\omega) = \omega(x(y)) \left|\det\left(\frac{\partial x^i}{\partial y^j}\right)\right| dy^1 \wedge \cdots \wedge dy^n.$$

If X is compact, possibly with boundary, then the integrand is bounded. Then integration makes sense as a linear map $\int_X : \Omega^n(X) \to \mathbb{R}$. Furthermore, Stokes' theorem is the statement that for the inclusion $i : \partial X \to X$, this diagram commutes:

$$\begin{tikzcd} \Omega^{n-1}(X) \arrow[r, "i^*"] \arrow[d, "d"] & \Omega^{n-1}(\partial X) \arrow[d, "\int_{\partial X}"] \\ \Omega^n(X) \arrow[r, "\int_X"] & \mathbb{R} \end{tikzcd}, \quad \text{i.e. } \int_{\partial X} i^*(\omega) = \int_X d\omega.$$

Theorem. *Suppose X is a compact smooth oriented manifold with boundary ∂X. There does not exist a smooth retraction map $r : X \to \partial X$, i.e. a map such that*

$$\begin{tikzcd} \partial X \arrow[r, "i"'] \arrow[rr, bend left, "\mathrm{Id}_{\partial X}"] & X \arrow[r, "r"'] & \partial X \end{tikzcd}.$$

Proof. Suppose such an r exists. Apply the functor Ω^{n-1}. Then

$$\begin{tikzcd} \Omega^{n-1}(\partial X) \arrow[r, "r^*"'] \arrow[rr, bend left, "\mathrm{Id}^*_{\partial X}"] & \Omega^{n-1}(X) \arrow[r, "i^*"'] & \Omega^{n-1}(\partial X) \end{tikzcd}.$$

The cochain diagram and Stokes' theorem give linear maps

$$\begin{tikzcd} \Omega^{n-1}(\partial X) \arrow[r, "r^*"'] \arrow[rr, bend left, "\mathrm{Id}^*_{\partial X}"] \arrow[d, "d"] & \Omega^{n-1}(X) \arrow[r, "i^*"'] \arrow[d, "d"] & \Omega^{n-1}(\partial X) \arrow[d, "\int_{\partial X}"] \\ \Omega^n(\partial X) \arrow[r, "r^*"] & \Omega^n(X) \arrow[r, "\int_X"] & \mathbb{R} \end{tikzcd}.$$

The orientation on X induces an orientation on ∂X. Let $\omega \in \Omega^{n-1}(\partial X)$ be a form inducing this orientation. Thus ω is nowhere-vanishing, and positive in each coordinate chart for ∂X. Consider ω in the upper-left of this diagram, and its eventual image in $\mathbb{R}$ in the lower-right. We can follow the arrows in any direction. Going across the top, we get $\mathrm{Id}^*_{\partial X}\omega = \omega$ in the upper-right. By the positivity of ω in each coordinate chart, $\int_{\partial X} \omega > 0$. Alternatively, following d on the left, by the Pauli exclusion principle $\Omega^n(\partial X) = \{0\}$ since $n > \dim \partial X$. Since the maps are linear, we must get zero going along the bottom row. This is the contradiction. ☐

What can be done with cohomology classes?

Suppose $[\omega] \in H^k(X; \mathbb{R})$. Suppose Y is a compact oriented k-manifold without boundary, and suppose $f : Y \to X$. I claim we can "evaluate" $f(Y)$ on $[\omega]$ by the following formula:

$$\langle f, [\omega] \rangle := \int_Y f^*(\omega).$$

To show that this is well-defined, we must show that a different representative, $[\omega + d\eta] = [\omega]$, leads to the same result. Indeed, the difference

$$\langle f, [d\eta]\rangle = \int_Y f^*(d\eta) = \int_Y df^*(\eta) = \int_{\partial Y = \emptyset} f^*(\eta) = 0.$$

Moreover, I claim that the result depends only on the oriented bordism class of f. Specifically, if W is a $k+1$-dimensional compact oriented manifold with $\partial W = \bar{Y} \amalg Y'$, and $\tilde{f} : W \to X$ with $\tilde{f}|_Y = f$, then $\langle f, [\omega]\rangle = \left\langle \tilde{f}\ |_{Y'}, [\omega]\right\rangle$. This also follows from Stokes' theorem:

$$\left\langle \tilde{f}\ |_{Y'}\right\rangle - \langle f, [\omega]\rangle = \int_{\partial W} \tilde{f}^*(\omega) = \int_W d\tilde{f}^*(\omega) = \int_W \tilde{f}^*(d\omega) = 0.$$

This shows we have a well-defined pairing

$$\Omega_k^{\mathrm{SO}}(X) \times H^k(X; \mathbb{R}) \to \mathbb{R}.$$

This pairing is functorial in that if $\phi : X \to X'$, then

$$\langle [f], \phi^* [\omega]\rangle = \int_Y f^*(\phi^*(\omega)) = \langle [\phi \circ f], [\omega]\rangle =: \langle \phi_* [f], [\omega]\rangle.$$

Orientation of $\mathbb{CP}^2$

One important property of cohomology is that the wedge product on forms induces a product on cohomology called the cup product.

$$[\alpha] \smile [\beta] := [\alpha \wedge \beta].$$

This is well-defined and functorial.

Recall the example of $X = \mathbb{CP}^2$. It has a Kähler form $\omega \in \Omega^2(\mathbb{CP}^2)$. It satisfies $d\omega = 0$, and $\omega \wedge \omega \in \Omega^4(\mathbb{CP}^2)$ is a volume form, which is positive for the usual orientation of $\mathbb{CP}^2$. Indeed,

$$\left\langle \left[\mathbb{CP}^2\right], [\omega] \smile [\omega]\right\rangle = \int_{\mathbb{CP}^2} \omega \wedge \omega > 0.$$

To repeat the previous argument, suppose $\phi : \overline{\mathbb{CP}^2} \to \mathbb{CP}^2$ is an orientation-preserving diffeomorphism. Then $\phi^*([\omega]) = \lambda [\omega]$ for some $\lambda \in \mathbb{R} - \{0\}$. It follows that

$$\left\langle \left[\mathbb{CP}^2\right], [\omega] \smile [\omega]\right\rangle = \left\langle \phi_* \left[\overline{\mathbb{CP}^2}\right], [\omega] \smile [\omega]\right\rangle = \left\langle \left[\overline{\mathbb{CP}^2}\right], \phi^* ([\omega] \smile [\omega])\right\rangle = -\lambda^2 \left\langle \left[\mathbb{CP}^2\right], [\omega] \smile [\omega]\right\rangle.$$

Triangulation

A more involved study of cohomology shows that, for similar reasons, $\mathbb{CP}^2$ is not the boundary of any 5-manifold. For this reason, we have that

$$\left[\mathbb{CP}^2\right] \neq 0 \in \Omega_4^{\mathrm{SO}}(\mathrm{pt}).$$

Before knowing this fact, one may suspect that the pairing

$$\Omega_k^{\mathrm{SO}}(X) \otimes H^k(X;\mathbb{R}) \to \mathbb{R}$$

is nondegenerate, i.e. that $\Omega_k^{\mathrm{SO}}(X) \otimes \mathbb{R}$ is dual to $H^k(X;\mathbb{R})$. However, since $H^k(\mathrm{pt};\mathbb{R}) = 0$ for $k > 0$, this is clearly not the case.

A better candidate for the dual to cohomology is called homology. It amounts to bordism plus triangulation. Any smooth manifold admits a triangulation. If X is a convex subset of $\mathbb{R}^n$, then $H_k(X) = 0$ for $k > 0$. Given a triangulation (for example, consider $f : \mathbb{CP}^2 \to X$) and a choice of vertex $v \in X$, we can form the cone Kf. As a linear combination of simplices (hypertriangles), it satisfies the relation $f = \partial Kf - K\partial f$. In particular, if $\partial f = 0$, then $f = \partial Kf$, so every closed surface is a boundary in homology.

There is a map $\Omega_k^{\mathrm{SO}}(X) \to H_k(X)$, and the pairing factors as

$$\Omega_k^{\mathrm{SO}}(X) \otimes H^k(X;\mathbb{R}) \to H_k(X) \otimes H^k(X;\mathbb{R}) \to \mathbb{R}.$$

This map is nondegenerate.

Singular cohomology

We take an alternative approach to locally constant functions which leads to triangulations from the outset.

Define $\mathscr{C}^0(X)$ to be the vector space of arbitrary functions

$$\mathscr{C}^0(X) := \{f : X \to \mathbb{R}\}.$$

A function $f \in \mathscr{C}^0(X)$ is locally constant iff every continuous path $\gamma : [0,1] \to U$ satisfies $f(\gamma(1)) - f(\gamma(0)) = 0$.

Define $\mathscr{C}^1(X)$ to be the vector space of arbitrary functions

$$\mathscr{C}^1(X) := \{g : \{\text{cts paths in } X\} \to \mathbb{R}\}$$

and define the linear map

$$\begin{aligned} d : \mathscr{C}^0(X) &\to \mathscr{C}^1(X), \\ df &:= [\gamma \mapsto f(\gamma(1)) - f(\gamma(0))]. \end{aligned} \tag{B.3}$$

Then $f \in \mathscr{C}^0(X)$ is constant iff $df \equiv 0$. We proceed in the same way as before. Suppose $X = B$ is some ball. We resolve $\operatorname{im}(d) \subset \mathscr{C}^1(B)$.

Let Δ be a model triangle with oriented edges

$$\begin{tikzcd} & {} \arrow[dr, "\partial_3"] & \\ {} \arrow[ur, "\partial_1"] \arrow[rr, "\partial_2"'] & & {} \end{tikzcd}.$$

Define $\mathscr{C}^2(U)$ to be the vector space of arbitrary functions

$$\mathscr{C}^2(B) := \{h : \{\text{continuous maps } \sigma : \Delta \to B\} \to \mathbb{R}\}.$$

Define

$$\begin{aligned} d : \mathscr{C}^1(B) &\to \mathscr{C}^2(B), \\ dg &:= \left[\sigma \mapsto g(\sigma|_{\partial_1}) - g(\sigma|_{\partial_2}) + g(\sigma|_{\partial_3})\right]. \end{aligned} \tag{B.4}$$

Suppose Δ has vertices

$$\begin{tikzcd} & b \arrow[dr] & \\ a \arrow[ur] \arrow[rr] & & c, \end{tikzcd}$$

and $g = df$. Then

$$\begin{aligned} d^2 f = dg &= [\sigma \mapsto g(\sigma|_{\partial_1}) - g(\sigma|_{\partial_2}) + g(\sigma|_{\partial_3})] \\ &= [\sigma \mapsto df(\sigma|_{\partial_1}) - df(\sigma|_{\partial_2}) + df(\sigma|_{\partial_3})] \\ &= \left[\sigma \mapsto \Big(f(\sigma(b)) - f(\sigma(a))\Big) - \Big(f(\sigma(c)) - f(\sigma(a))\Big) + \Big(f(\sigma(c)) - f(\sigma(b))\Big)\right] \\ &= [\sigma \mapsto 0] \\ &= 0. \end{aligned} \tag{B.5}$$

If U is a ball, then $dg = 0 \iff g = df$. Using simplices, we can continue the resolution, and obtain the complex of *singular cochains*

$$0 \to \mathscr{C}^0(U) \to \mathscr{C}^1(U) \to \mathscr{C}^2(U) \to \cdots$$

resolving $\mathbb{R} \subset \mathscr{C}^0(U)$.

For a manifold M which is not necessarily a ball, we form the complex

$$0 \to \mathscr{C}^0(M) \to \mathscr{C}^1(M) \to \mathscr{C}^2(M) \to \cdots$$

with singular cohomology

$$H^p_{\text{sing}}(M; \mathbb{R}) \cong H^p(M; \mathbb{R}).$$

Singular cohomology makes sense for arbitrary topological spaces, not just manifolds. Singular cohomology also makes sense with values any ring in place of $\mathbb{R}$.

There are other variants of cohomology which are naturally isomorphic to de Rham cohomology and singular cohomology. For example, Čech cohomology arises from characterizing locally constant functions f as those for which, given an open cover of X, there exists a refinement of that cover such that $f\ |_U$ is constant for each U.

There are variants of cohomology which are distinct from the standard cohomology theories called *generalized cohomology theories*. These satisfy the Eilenberg-Steenrod axioms which characterize cohomology, but differ in that $H^\bullet(\mathrm{pt})$ can be more interesting. Oriented bordism is such an example.

Poincaré duality

If X is a compact oriented manifold without boundary of dimension n, and if $Y \subset X$ is a submanifold of dimension k, then it is possible to construct a closed differential form $\omega_Y \in \Omega^{n-k}(X)$ such that $d\omega_Y = 0$, and ω_Y vanishes outside an arbitrarily small neighborhood of Y. Moreover, one arranges that ω_Y integrates to 1 on each fiber of the normal bundle. In this case, the wedge product corresponds to oriented intersection. If Y' is another submanifold which is transverse to Y, then $\omega_Y \wedge \omega_{Y'}$ vanishes outside a small neighborhood of $Y \cap Y'$, and also integrates to 1 on each fiber of the normal bundle. In the case that Y and Y' intersect in a finite set of points, $[\omega_Y] \smile [\omega_{Y'}] \in H^n$, and $\langle [X], [\omega_Y] \smile [\omega_{Y'}] \rangle$ gives the signed count of these points. In this way, cohomology encodes intersection theory.

B.3 General coefficients for cohomology

There are no fractions involved in the definition of singular cohomology. Thus it makes sense to replace all instances of $\mathbb{R}$ by $\mathbb{Z}$ in the definition of singular cohomology. This version resolves the sheaf of locally constant $\mathbb{Z}$-valued functions, and is more delicate and powerful than its $\mathbb{R}$ counterpart. More generally, this construction makes sense over any abelian group A, which we denote with cochain groups $C^\bullet(X; A)$ and cohomology groups $H^\bullet(X; A)$. There is a cup product structure whenever A is a ring R. Most common are $R \in \{\mathbb{R}, \mathbb{Z}, \mathbb{Z}_2\}$.

Sadly, de Rham theory is capable only of computing $H^\bullet(X; \mathbb{R})$.

There is a "universal coefficient theorem" which computes cohomology with general coefficients, but first we need homology.

B.4 Singular homology

We observe that $C^p(X; A)$ is the dual space

$$C^p(X; A) = C_p(X; A)^* := \mathrm{Hom}(C_p(X; A), A),$$

where $C_p(X; A)$ is the vector space with a basis element corresponding to each $\sigma \in \Delta_p(X)$. In other words, finite formal linear combinations

$$C_p(X; A) := \sum_{\sigma \in \Delta_p(X)} c_\sigma\, \sigma, \text{ where finitely many } c_\sigma \in A \text{ are nonzero.}$$

Furthermore, there is a linear differential $\partial : C_p(X; A) \to C_{p-1}(X; A)$ determined by

$$\partial\sigma := \sum_{i=0}^{p} (-1)^i F_i \sigma,$$

satisfying $\partial^2 = 0$. This fits into a sequence

$$\cdots \xrightarrow{\partial} C_2(X) \xrightarrow{\partial} C_1(X) \xrightarrow{\partial} C_0(X),$$

where we leave A implicity for brevity. The dual of this sequence is precisely the sequence

$$C^0(X) \xrightarrow{d} C^1(X) \xrightarrow{d} C^2(X) \xrightarrow{d} \cdots$$

of singular cohomology. We define singular homology

$$H_p(X) := \frac{\ker\left(C_p(X) \to C_{p-1}(X)\right)}{\operatorname{image}\left(C_{p+1}(X) \to C_p(X)\right)}.$$

For example, $H_0(X; A) \cong A^{\#\text{components}(X)}$. Also, if X is connected, then $H_1(X; \mathbb{Z}) \cong \pi_1^{\text{ab}}(X, x_0)$, where π_1^{ab} denotes the abelianization of the fundamental group. In particular, for the Poincaré homology sphere P, $H_1(P; \mathbb{Z}) = 0$.

B.5 Universal coefficients and Poincaré duality

Cohomology is dual to homology in two distinct ways: universal coefficients and Poincaré duality.

The more straightforward is the universal coefficient theorem. One might hope that since cochains are dual to chains, maybe cohomology is dual to homology. This is almost true, but not quite. The situation is described by the split exact sequence

$$0 \to \operatorname{Ext}(H_{i-1}(X; \mathbb{Z}), A) \to H^i(X; A) \to \operatorname{Hom}(H_i(X; \mathbb{Z}), A) \to 0.$$

Whenever we have a *short exact sequence* of abelian groups $0 \to A \to B \to C \to 0$, it follows that A can be identified with its image $A \subset B$ and $C \cong B/A$. (These are straightforward consequences of the definition of exactness.) The term *split* means that there is a subgroup of B representing the quotient group C, so that $B \cong A \oplus C$. (However, this splitting it rarely unique.)

To be more concrete regarding the universal coefficient theorem, suppose that X is closed. Then $H_i(X; \mathbb{Z})$ is a finitely generated abelian group. According to the classification theorem for finitely generated abelian groups, there is a split exact sequence

$$0 \to T_i(X) \to H_i(X; \mathbb{Z}) \to H_i^{\text{free}}(X) \to 0,$$

where the *torsion subgroup* $T_i(X) \subset H_i(X; \mathbb{Z})$ is the finite subgroup consisting of elements of finite order. Furthermore, $T_i(X)$ is isomorphic to a direct sum of finite cyclic groups $T_i(X) \cong \mathbb{Z}_{i_1} \oplus \cdots \oplus \mathbb{Z}_{i_k}$.

The quotient group $H_i(X;\mathbb{Z})/T_i(X)$ is a free group denoted $H_i^{\text{free}}(X)$. Since the sequence is split, there exists isomorphisms

$$H_i(X;\mathbb{Z}) \cong H_i^{\text{free}}(X) \oplus T_i(X).$$

Upon choosing a basis for $H_i^{\text{free}}(X)$, we obtain an isomorphism

$$H_i(X;\mathbb{Z}) \cong \mathbb{Z}^{b_i(X)} \oplus T_i(X),$$

for some nonnegative integer $b_i(X)$.

Recall the universal coefficient theorem

$$0 \to \operatorname{Ext}(H_{i-1}(X;\mathbb{Z}), A) \to H^i(X;A) \to \operatorname{Hom}(H_i(X;\mathbb{Z}), A) \to 0,$$

which is split, so there is an isomorphism $H^i(X;A) \cong \operatorname{Ext}(\cdots) \oplus \operatorname{Hom}(\cdots)$. By classification of finitely generated abelian groups, when X is closed,

$$H_i(X;\mathbb{Z}) \cong \mathbb{Z}^{b_i(X)} \oplus T_i(X),$$

for some integer $b_i(X)$ and some torsion subgroup $T_i(X) \cong \mathbb{Z}_{i_1} \oplus \cdots \oplus \mathbb{Z}_{i_k}$.

To compute real cohomology

$$H^i(X;\mathbb{R}) \cong \operatorname{Ext}(\cdots) \oplus \operatorname{Hom}(\cdots),$$

we get $\operatorname{Ext}(H_{i-1}(X;\mathbb{Z}),\mathbb{R}) = 0$ and $\operatorname{Hom}(H_i(X;\mathbb{Z}),\mathbb{R}) \cong \mathbb{R}^{b_i(X)}$, thus

$$H^i(X;\mathbb{R}) \cong \mathbb{R}^{b_i(X)},$$

and we see that $b_i(X) = b^i(X)$, where $b^i(X)$ are the familiar Betti numbers from de Rham cohomology. For integer cohomology, we compute $\operatorname{Ext}(H_{i-1}(X;\mathbb{Z}),\mathbb{Z}) = T_{i-1}$ and $\operatorname{Hom}(H_i(X;\mathbb{Z}),\mathbb{Z}) \cong \mathbb{Z}^{b^i(X)}$, so

$$H^i(X;\mathbb{Z}) \cong \mathbb{Z}^{b^i(X)} \oplus T_{i-1}.$$

We now be more precise regarding the relation between real and integer cohomology for compact manifolds. We have natural isomorphisms

$$H^i(X;\mathbb{R}) = \operatorname{Hom}(H_i(X;\mathbb{Z}),\mathbb{R}) = \operatorname{Hom}(H_i(X;\mathbb{Z}),\mathbb{Z}) \otimes \mathbb{R} = H^i_{\text{free}}(X;\mathbb{Z}) \otimes \mathbb{R},$$

where we use the fact that torsion disappears under tensor or hom with $\mathbb{R}$. Thus, we may view $H^i_{\text{free}}(X;\mathbb{Z})$ as an integer lattice inside of the vector space $H^i(X;\mathbb{R})$.

Poincaré duality is a different identification of homology with cohomology, giving an isomorphism

$$H^i_c(X;\mathbb{Z}) \to H_{n-i}(X;\tilde{\mathbb{Z}}),$$

where $\tilde{\mathbb{Z}}$ denotes homology with "twisted coefficients." If X is closed, then $H^k_c(X;A) = H^k(X;A)$. If X is oriented, then $\tilde{\mathbb{Z}} = \mathbb{Z}$, and consequently,

$$H^i(X;\mathbb{Z}) \cong \mathbb{Z}^{b^{n-i}(X)} \oplus T_{n-i}(X).$$

Thus

$$T_i(X) := T(H_i(X;\mathbb{Z})) \overset{\text{UC}}{\cong} T(H^{i+1}(X;\mathbb{Z})) \overset{\text{PD}}{\cong} T(H_{n-i-1}(X)) =: T_{n-i-1}(X).$$

Note that since $H_0(X;\mathbb{Z}) = \mathbb{Z}^{\#\text{components}(X)}$ is free, $T_0 = 0$. Therefore, $T_{n-1} \cong T_0 = 0$, and $T_n \cong T_{-1} = 0$.

Now the homology and cohomology groups are very tightly constrained. For example, for a connected, compact oriented 4-manifold, by the universal coefficient theorem,

i	0	1	2	3	4
$H_i(X;\mathbb{Z})$	$\mathbb{Z}$	$\mathbb{Z}^{b^1(X)} \oplus T_1$	$\mathbb{Z}^{b^2(X)} \oplus T_2$	$\mathbb{Z}^{b^3(X)} \oplus T_3$	$\mathbb{Z}^{b^4(X)} \oplus T_4$
$H^i(X;\mathbb{Z})$	$\mathbb{Z}$	$\mathbb{Z}^{b^1(X)}$	$\mathbb{Z}^{b^2(X)} \oplus T_1$	$\mathbb{Z}^{b^3(X)} \oplus T_2$	$\mathbb{Z}^{b^4(X)} \oplus T_3$

and then by Poincaré duality,

i	0	1	2	3	4
$H_i(X;\mathbb{Z})$	$\mathbb{Z}$	$\mathbb{Z}^{b^1(X)} \oplus T_1$	$\mathbb{Z}^{b^2(X)} \oplus T_1$	$\mathbb{Z}^{b^1(X)}$	$\mathbb{Z}$
$H^i(X;\mathbb{Z})$	$\mathbb{Z}$	$\mathbb{Z}^{b^1(X)}$	$\mathbb{Z}^{b^2(X)} \oplus T_1$	$\mathbb{Z}^{b^1(X)} \oplus T_1$	$\mathbb{Z}$

For the Poincaré homology sphere P, since $H_1(X;\mathbb{Z}) \cong \pi_1^{\text{ab}}(P, x_0) = \tilde{I}^{\text{ab}} = 0$, one easily computes

i	0	1	2	3
$H_i(X;\mathbb{Z})$	$\mathbb{Z}$	0	0	$\mathbb{Z}$
$H^i(X;\mathbb{Z})$	$\mathbb{Z}$	0	0	$\mathbb{Z}$

Since there are no possibilities for nontrivial cup products, in both homology and cohomology, P looks exactly like S^3, even in the ring structure of cohomology.

Exercise. Compute the homology and cohomology groups of a connected closed oriented 3-manifold X in terms of $\pi_1(X)$.

In the compact case, an orientation is equivalent to a choice of generator of $H_i(X;\mathbb{Z}) \cong \mathbb{Z}$, known as the fundamental class $[X]$. This homology class should be thought of as a (oriented) triangulation of X. Concretely, a nowhere vanishing element of $\omega \in \Omega^n(X)$ determines a positive atlas, in which ω is positive in each chart. Thus $\int_X \omega > 0$. Clearly $\omega \in \ker d$, since $\Omega^{n+1}(X) = 0$. But $[\omega] \neq 0 \in H^n(X;\mathbb{R})$ since by Stokes' theorem,

$$[\omega] = 0 \implies \omega = d\eta \implies \int_X \omega = \int_X d\eta = \int_{\partial X = \emptyset} \eta = 0.$$

Thus ω must generate the one-dimensional vector space $H^n(X;\mathbb{R})$. After rescaling ω by the appropriate positive constant, ω determines a generator for $H^n(X;\mathbb{Z}) \subset H^n(X;\mathbb{R})$. By the universal coefficient theorem, $H_n(X;\mathbb{Z})$ is the dual lattice inside the dual vector space $H_n(X;\mathbb{R})$, and we have the corresponding dual basis element $[X] \in H_n(X;\mathbb{Z})$.

In the compact oriented case, the Poincaré duality isomorphism $H^i(X;\mathbb{Z}) \to H_{n-i}(X;\mathbb{Z})$ is *cap product* with the fundamental class $[X]$:

$$a \mapsto [X] \frown a.$$

Specifically, for any ring R, the cap product can be defined as partial evaluation $H_{i+j}(X;R) \times H^i(X;R) \to H_j(X;R)$. Ignoring torsion, the cap product is dual to the cup product. If we view $H^{\text{free}}_{n-i}(X;\mathbb{Z})$ as the dual group to $H^{n-i}_{\text{free}}(X;\mathbb{Z})$, then the map induced by the cap product and universal coefficients

$$H^i_{\text{free}}(X;\mathbb{Z}) \to H^{\text{free}}_{n-i}(X;\mathbb{Z}) \xrightarrow{\cong} H^{n-i}_{\text{free}}(X;\mathbb{Z})^*$$

corresponds to the "cup product and integrate" map

$$\begin{aligned} H^i(X;\mathbb{R}) &\to H^{n-i}(X;\mathbb{R})^* \\ a &\mapsto \left(b \mapsto \int_X (a \smile b)\right) \end{aligned}$$

which we used to define the intersection form via de Rham cohomology. This gives the intersection form the structure of a unimodular integer bilinear form, as was previously claimed.

B.6 Representing homology classes via submanifolds

Homology classes are represented by simplicial "cycles" i.e. chains without boundary. The "Steenrod problem" asks whether a class $a \in H_p(X;\mathbb{Z})$ can be represented by a manifold. Specifically, is there a closed oriented smooth manifold M and a continuous map $f : M \to X$ such that the image of a fundamental class $f_*([M]) = a$? In his work on cobordism, Thom showed that this is not always possible. However, there is always an integer multiple of a which is representable. This question can be strengthened to require that f be either an immersion (locally an embedding, but globally there can be self-intersections) or an embedding. In particular, when we discuss the minimal genus problem, we want to know that classes $a \in H_2(X;\mathbb{Z})$ are representable by embedded submanifolds. Assuming some homotopy theory, we can prove via Poincaré duality that this is always possible when X is a closed oriented 4-manifold.

To warm up, consider X closed and oriented, and $a \in H_{n-1}(X;\mathbb{Z})$. We can write any such a as the Poincaré dual $a = \mathrm{PD}(\alpha)$ to $\alpha \in H^1(X;\mathbb{Z})$. The homotopy theory fact we require is that cohomology groups are *representable*, i.e. $H^p(X;A)$ is in bijection with homotopy classes of maps from X into some space $K(A,p)$, called an Eilenberg-MacLane space.

$$H^p(X;A) \overset{\text{bij}}{\cong} [X, K(A,p)]\,.$$

In particular,

$$H^1(X;\mathbb{Z}) \overset{\text{bij}}{\cong} [X, K(\mathbb{Z},1)]\,,$$

and $K(\mathbb{Z},1) = S^1$. Choosing a representative map $h : X \to S^1$, the corresponding cohomology class is given by pullback of the generator $\xi = [d\theta/2\pi] \in H^1(S^1;\mathbb{Z})$. Thus each cohomology class $\alpha \in H^1(X;\mathbb{Z})$ is $\alpha = h^*(\xi)$ for some map $h : X \to S^1$. We can represent the homology class $\mathrm{PD}(\xi)$ by a point pt, so that $\mathrm{PD}(\xi) = [\mathrm{pt}] \in H^0(S^1)$.

Poincaré duality is functorial in the sense that

$$\mathrm{PD}(\alpha) = \mathrm{PD}(h^*(\xi)) = h^*(\mathrm{PD}(\xi)) = h^*([\mathrm{pt}]) = \left[h^{-1}\left(\mathrm{pt}\right)\right],$$

where h^{-1} denotes the inverse image, assuming that h is transverse to the given representative pt of $\mathrm{PD}(\xi)$.

In summary, to represent a homology class $a \in H_{n-1}(X; \mathbb{Z})$, write $a = \mathrm{PD}(\alpha)$, and then choose a map $h : X \to K(\mathbb{Z}, 1)$ such that $\alpha = h^*(\xi) \in H^1(X; \mathbb{Z})$, where ξ is the generator in $\mathbb{Z} = H^1(K(\mathbb{Z}, 1); \mathbb{Z})$. Since $K(\mathbb{Z}, 1)$ happens to be a manifold, we can look for an explicit codimension 1 submanifold $Y \subset K(\mathbb{Z}, 1)$ corresponding to $\mathrm{PD}(\xi) \in H_1(K(\mathbb{Z}, 1); \mathbb{Z})$. (Here, Y happens to be a point.) After perturbing h to make it transverse to Y, a is the fundamental class of the preimage $h^{-1}(Y)$.

With slight modification, the same argument carries through for $a \in H_{n-2}(X; \mathbb{Z})$. In this case, $K(\mathbb{Z}, 2) = \mathbb{CP}^\infty$, where $\mathbb{CP}^\infty$ is the union of the inclusions $\mathbb{CP}^1 \subset \mathbb{CP}^2 \subset \cdots \subset \bigcup_i \mathbb{CP}^i = \mathbb{CP}^\infty$. While this is not a manifold in the traditional sense, the "cellular approximation theorem" of homotopy theory allows us to homotope h to some $\mathbb{CP}^N$ with N finite. In this case, for ξ the generator of $H^2(\mathbb{CP}^N; \mathbb{Z}) \cong \mathbb{Z}$, the Poincaré dual $\mathrm{PD}(\xi) = \left[\mathbb{CP}^{N-1}\right]$. Thus when h is transverse to $\mathbb{CP}^{N-1}$, we obtain a suitable embedded submanifold $h^{-1}(\mathbb{CP}^{N-1})$ whose fundamental class represents a.

B.7 Self-intersection number

B.8 Cohomological definition of orientation

Bibliography

[BT82] R. Bott and L. W. Tu. *Differential forms in algebraic topology*, volume 82 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982. MR658304.

[DK90] S. K. Donaldson and P. B. Kronheimer. *The geometry of four-manifolds*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1990. Oxford Science Publications. MR1079726.

[FM94] R. Friedman and J. W. Morgan. *Smooth four-manifolds and complex surfaces*, volume 27 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1994. doi:10.1007/978-3-662-03028-8, MR1288304.

[Fra12] T. Frankel. *The geometry of physics*. Cambridge University Press, Cambridge, third edition, 2012. An introduction. MR2884939.

[FU91] D. S. Freed and K. K. Uhlenbeck. *Instantons and four-manifolds*, volume 1 of *Mathematical Sciences Research Institute Publications*. Springer-Verlag, New York, second edition, 1991. doi:10.1007/978-1-4613-9703-8, MR1081321.

[Iga02] K. Iga. What do topologists want from Seiberg-Witten theory? *Internat. J. Modern Phys. A*, 17(30):4463–4514, 2002. arXiv:hep-th/0207271, doi:10.1142/S0217751X0201217X, MR1941517.

[KM94] P. B. Kronheimer and T. S. Mrowka. The genus of embedded surfaces in the projective plane. *Math. Res. Lett.*, 1(6):797–808, 1994. doi:10.4310/MRL.1994.v1.n6.a14, MR1306022.

[KM07] P. Kronheimer and T. Mrowka. *Monopoles and three-manifolds*, volume 10 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2007. doi:10.1017/CBO9780511543111, MR2388043.

[Law85] H. B. Lawson, Jr. *The theory of gauge fields in four dimensions*, volume 58 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1985. MR799712.

[Law97] T. Lawson. The minimal genus problem. *Exposition. Math.*, 15(5):385–431, 1997. MR1486407.

[LM89] H. B. LAWSON, JR. AND M.-L. MICHELSOHN. *Spin geometry*, volume 38 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1989. MR1031992.

[Mar99] M. MARCOLLI. *Seiberg-Witten gauge theory*, volume 17 of *Texts and Readings in Mathematics*. Hindustan Book Agency, New Delhi, 1999. With an appendix by the author and Erion J. Clark. MR1700273.

[Moo01] J. D. MOORE. *Lectures on Seiberg-Witten invariants*, volume 1629 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001. MR1830497.

[Mor96] J. W. MORGAN. *The Seiberg-Witten equations and applications to the topology of smooth four-manifolds*, volume 44 of *Mathematical Notes*. Princeton University Press, Princeton, NJ, 1996. MR1367507.

[Mor98] J. W. MORGAN. An introduction to gauge theory. In *Gauge theory and the topology of four-manifolds (Park City, UT, 1994)*, volume 4 of *IAS/Park City Math. Ser.*, pages 51–143. Amer. Math. Soc., Providence, RI, 1998. MR1612968.

[Mül08] D. MÜLLNER. *Orientation reversal of manifolds*. PhD thesis, Bonn, 2008.

[Sal] D. SALAMON. *Spin geometry and Seiberg-Witten invariants*.

[Sco05] A. SCORPAN. *The wild world of 4-manifolds*. American Mathematical Society, Providence, RI, 2005. MR2136212.

[Tau98] C. H. TAUBES. The geometry of the Seiberg-Witten invariants. In *Surveys in differential geometry, Vol. III (Cambridge, MA, 1996)*, pages 299–339. Int. Press, Boston, MA, 1998. doi:10.4310/SDG.1996.v3.n1.a5, MR1677891.

[War83] F. W. WARNER. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition. MR722297.